The RAG1 Ubiquitin Ligase Domain Enhances Assembly and Selection of T Cell Receptor Genes to Restrain the Autoimmune Hazard of Generating T Cell Receptor Diversity

Thomas N. Burn^{*,†,1}; Charline Miot^{†,1}; Scott Gordon^{*,§}; Erica J. Culberson[‡]; Tamir Diamond ^{*,¶}; Portia Kreiger[‡]; Katharina E. Hayer^{||}; Anamika Bhattacharyya^{#,**}, Jessica M. Jones[#], Craig H. Bassing^{*,‡,2}; Edward M. Behrens^{*,‡,2}

^{*}Institute for Immunology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

[†]Division of Rheumatology, The Children's Hospital of Philadelphia, Philadelphia, PA 19104

[‡]Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA 19104

[§]Division of Neonatology, The Children's Hospital of Philadelphia, Philadelphia, PA 19104

[¶]Division of Gastroenterology, Hepatology, and Nutrition, The Children's Hospital of Philadelphia, Philadelphia, PA 19104

^{||}Department of Biomedical and Health Bioinformatics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104

[#]Department of Biochemistry and Molecular and Cellular Sciences, Georgetown University, Washington, DC, 20057

^{**}Present Address: Vyome Therapeutics Lt., 501 Patparganj Industrial Area, Dehli, India

^{††}Department of Pathology and Laboratory Medicine, The Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA 19104

¹Co-first Authors

²Co-corresponding Authors:

Craig H. Bassing, Ph.D.
Children's Hospital of Philadelphia
4054 Colket Translational Research Building
3501 Civic Center Blvd.
Philadelphia, PA 19104
Phone: 267-426-0311
Email: bassing@email.chop.edu

Edward M. Behrens, M.D.
Children's Hospital of Philadelphia
1102 Abramson Research Center
3615 Civic Center Boulevard
Philadelphia, PA 19104
Phone: 267-426-0142
Email: behrens@email.chop.edu

Running Title: RAG1 enhances TCR assembly and thymic selection

Abstract

RAG1/RAG2 (RAG) endonuclease-mediated assembly of diverse lymphocyte antigen receptor genes by V(D)J recombination is critical for the development and immune function of T and B cells. However, this process creates highly self-reactive cells that must be properly selected to suppress autoimmunity. The RAG1 protein contains a ubiquitin ligase domain that stabilizes RAG1 and stimulates RAG endonuclease activity *in vitro*. We report that mice with a mutation that inactivates the RAG1 ubiquitin ligase *in vitro* exhibit modestly decreased assembly and altered repertoires of T cell receptor (TCR) β and α genes in thymocytes and impaired thymocyte developmental transitions that require assembly of these genes and signaling by their proteins. The mice also exhibit a combined phenotype of: 1) decreased expression of TCR signaling proteins in thymocytes, 2) less efficient positive selection of conventional $\alpha\beta$ T cells and agonist selection of iNKT cells, 3) impaired superantigen-mediated negative selection of highly self-reactive thymocytes, and 4) CD4⁺ $\alpha\beta$ T cells with elevated autoimmune potential. Our findings demonstrate that the RAG1 ubiquitin ligase domain functions *in vivo* to stimulate TCR β and TCR α gene assembly and $\alpha\beta$ TCR selection, thereby establishing replete $\alpha\beta$ TCR diversity and $\alpha\beta$ T cell lineages while restraining the inherent autoimmune hazard of creating diverse antigen specificities.

Key Points

The RAG1 ubiquitin ligase domain stimulates TCR α and β gene assembly *in vivo*.

The RAG1 ubiquitin ligase domain enhances $\alpha\beta$ TCR selection of thymocytes.

The RAG1 ubiquitin ligase domain limits autoimmune potential of $\alpha\beta$ T cells.

Introduction

The ability of T and B lymphocyte populations to express antigen receptors able to recognize a potentially unlimited number of distinct pathogens is the fundamental basis of adaptive immunity. RAG1/RAG2 (RAG) endonuclease-mediated assembly of T cell receptor (TCR) and immunoglobulin (Ig) genes through V(D)J recombination establishes this vast pool of diverse receptors. Germline TCR and Ig loci consist of variable (V), joining (J), and sometimes diversity (D), gene segments residing upstream of constant (C) region exons. During T and B cell development, the lymphocyte-specific RAG complex cooperates with lineage- and developmental stage-specific activation of TCR and Ig loci to perform V(D)J recombination (1). RAG induces DNA double strand breaks (DSBs) between coding sequences and flanking recombination signal sequences of two participating gene segments, liberating hairpin-sealed coding ends and blunt signal ends (2). In a subsequent reaction phase guided by RAG, DNA repair factors ligate signal ends to generate a signal join and open, process, and ligate coding ends to create a V(D)J coding join, which comprises the second exon of an assembled V(D)J-C antigen receptor gene (2). The high number of V(D)J joining events, broad utilization of gene segments, and inherent imprecision in opening and processing coding ends all cooperate to yield a vast population and diversity of receptors. The imprecise formation of coding joins creates the third hypervariable complementarity-determining region (CDR3)-encoding nucleotides of Ig and TCR variable regions. Due to imprecise means of opening and processing coding ends, the outcomes of V(D)J recombination include out-of-frame genes unable to make protein and receptors that recognize self-antigens and have potential for autoimmunity. Accordingly, developing T and B lymphocytes engage quality control checkpoints to select for potentially beneficial antigen receptor proteins and against hazardous highly self-reactive receptors.

Both immature T and B cells employ developmental programs that link functional protein expression from in-frame V(D)J rearrangements to signal either survival and continued differentiation or apoptosis (1). The maturation of $\alpha\beta$ T cells in the thymus provides a paradigm for how V(D)J recombination and antigen receptor protein quality control checkpoints cooperate to create a large pool of lymphocytes with diverse receptors that provides immunity from pathogens without causing overt autoimmunity. After entering the thymus, early thymocyte progenitor cells differentiate into $CD44^+CD25^+$ stage 2 and then into $CD44^-CD25^+$ stage 3 $CD4^-CD8^-$ double negative (DN) thymocytes (3, 4). Both DN2 and DN3 cells express RAG and conduct V(D)J recombination of *Tcrb*, *Tcrd*, and/or *Tcrg* loci (3, 4). Although DN2-to-DN3 thymocyte maturation can occur independent of RAG and V(D)J recombination, an in-frame *Tcrb* rearrangement is necessary for development beyond the DN3 stage (3, 4). The resultant TCR β protein associates with invariant pT α protein to produce pre-TCR complexes that signal antigen-independent survival, proliferation, and differentiation of DN3 cells (β -selection) into $CD44^-CD25^-$ stage 4 DN thymocytes and then $CD4^+CD8^+$ double-positive (DP) thymocytes (3, 4). DP cells express RAG and conduct V(D)J recombination of *Tcra* loci, which typically proceeds through multiple successive rounds of V-to-J rearrangement on both alleles (5). After an in-frame VJ rearrangement, the resulting TCR α protein can pair with TCR β protein and yield surface $\alpha\beta$ TCRs, which must signal to prevent death by neglect (6-9). This signal activation requires appropriate physical interactions between $\alpha\beta$ TCRs and self-peptide/MHC (pMHC) complexes displayed on thymic epithelial cells (TECs) or dendritic cells (DCs), meaning that a functional $\alpha\beta$ TCR is inherently self-reactive (6-9). Interactions below a particular low threshold of affinity (or avidity) cannot activate signalling to block apoptosis (6-9). At the other extreme, contacts above a higher threshold trigger strong signals that causes apoptosis (negative selection) to delete highly self-reactive cells with substantial autoimmune potential (6-9). Contacts between these thresholds activate TCR signals of a strength within a range that promotes survival and differentiation of DP cells (positive selection) into $CD4^+$ or $CD8^+$ single-positive (SP)

thymocytes, which lack RAG expression and exit the thymus as mature $\alpha\beta$ T cells (6-9). Strong TCR signals in DP cells can also activate an alternative agonist selection process that promotes differentiation of unconventional $\alpha\beta$ T cell lineages: regulatory T (T reg) cells, intestinal intraepithelial lymphocytes (IELs), and invariant natural killer T (iNKT) cells (6, 7). Notably, the $\alpha\beta$ TCRs of iNKT cells have limited diversity and recognize lipids displayed by MHC-like CD1d proteins on another DP thymocyte (10). This limited diversity renders developing iNKT cells a useful thymocyte population to probe issues of TCR signal strength without confounding by contribution of TCR specificity.

The RAG1 protein N-terminus has a Zinc-binding RING finger ubiquitin ligase domain of undetermined relevance for antigen receptor gene assembly. *In vitro*, RAG1 autoubiquitylates on lysine 233 (11), which enhances RAG endonuclease activity *in vitro* as evidenced by impaired cleavage upon mutation of this residue (12). *In vitro*, RAG1 also ubiquitylates other proteins including histone H3 within chromatin (13, 14). RAG1 RING domain mutations that disrupt ubiquitin ligase activity impair RAG-mediated cleavage and joining of chromatin substrates in cell lines, which could be due to loss of histone H3 ubiquitylation and/or alteration in RAG1 protein structure (14-16). The C328Y mutation of a critical structural zinc-binding cysteine of human RAG1 impairs V(D)J recombination and causes severe T and B cell immunodeficiency associated with uncontrolled proliferation of activated oligoclonal $\alpha\beta$ T cells (17). The analogous C325Y mutation of mouse Rag1 protein (C325Y) *in vitro* destabilizes Rag1 tertiary structure, abrogates ubiquitin ligase activity, reduces V(D)J recombination (22). Mice with homozygous Rag1 C325Y mutation have a near complete block of $\alpha\beta$ T cell development at the DN3 thymocyte stage due to impaired *Tcrb* recombination (15, 16). RING domains have a conserved proline (proline 326 in mouse Rag1) that is critical for ubiquitin ligase activity by permitting functional interaction with ubiquitin-conjugating enzymes (18, 19). Relative to C325Y mutation, P326G mutation abrogates Rag1 ubiquitin ligase activity equivalently, but has substantially less severe effects on both destabilizing

tertiary structure of the Rag1 RING domain and reducing RAG endonuclease activity (16). Accordingly, P326 mutation of endogenous mouse Rag1 protein allows an approach to elucidate potential physiologic roles of the RAG1 ubiquitin ligase domain beyond promoting V(D)J recombination by stabilizing RAG1 protein structure.

To determine potential *in vivo* functions of RAG1 ubiquitin ligase activity, we created and analyzed mice carrying a homozygous Rag1 P326G mutation. We show that these mice exhibit decreased levels and altered repertoires of *Tcrb* or *Tcra* rearrangements in DN3 or DP thymocytes, respectively, correlating with lower fractions of these cells selected for further differentiation. When assaying $\alpha\beta$ TCRs expressing specific TCR β or TCR α chains, we observe diminished efficiencies for superantigen-mediated negative selection of conventional $\alpha\beta$ T cells and agonist selection of iNKT cells. We find that the $\alpha\beta$ TCR signaled upregulation of CD69 is diminished during initiation of positive selection, and the TCR signaling protein Zap70 is expressed at lower than normal levels in DP cells but not mature $\alpha\beta$ T cells. Finally, we show that mature CD4⁺ effector $\alpha\beta$ T cells of these mice exhibit normal immune responses to activation, yet confer greater potential for autoimmunity. Our data demonstrate that the RAG1 ubiquitin ligase domain functions *in vivo* to stimulate TCR β and TCR α gene assembly and $\alpha\beta$ TCR selection, thereby establishing replete $\alpha\beta$ TCR diversity and $\alpha\beta$ T cell lineages while restraining the inherent autoimmune hazard of creating diverse TCR specificities. We discuss how the RAG1 ubiquitin ligase domain might function *in vivo* to regulate the assembly of TCR β and TCR α genes and the selection of $\alpha\beta$ TCRs.

Materials and Methods

Mice

C57BL/6 background *PG* mice were made through gene-targeting as performed by InGenious Targeting Laboratory (2200 Smithtown Avenue, Ronkonkoma, NY). The 9.5 kb genomic DNA used to construct the targeting vector was sub-cloned from a 129Svev BAC clone. The CC>GG double point mutation was shuttled from pJMJ029[P326G](16) into the targeting vector using unique restriction sites, then confirmed by sequencing. Linearized targeting vector was electroporated into BA1 (C57BL/6 x 129/SvEv) embryonic stem (ES) cells. After selection with G418 antibiotic, clones were expanded for PCR analysis to identify recombinant ES clones. Positive clones were screened for short arm integration and then confirmation of the point mutation was performed by PCR and sequencing. Positive clones were analyzed by Southern blot to confirm long arm integration. Injected blastocysts were grown in foster mothers and chimeras were backcrossed against C57BL/6 mice to generate heterozygotes. These were crossed with CMV-*cre* mice (JAX #006054)(20) to remove the *neo* reporter cassette and generate the *Rag1*^{P326G} allele. After confirmation of *neo* cassette removal, the strain was backcrossed against C57BL/6, and removal of CMV-*cre* was confirmed by PCR. Sequencing of the entire *Rag1* locus in progeny confirmed the intended CC>GG mutation and an additional T>C point mutation that generates a V238A substitution outside of the RING finger domain. The strain was deposited at the Mutant Mouse Regional Resource Center at the University of North Carolina (RAG1-tm1, MMRRC #: 37105). Before initiating studies described herein, the *Rag1*^{P326G} allele was backcrossed for nine generations onto the

206 C57BL/6 background. To generate NOD background *PG* mice, CRISPR/Cas9-mediated gene editing
 207 was conducted in zygotes from *in vitro* fertilization of NOD eggs with NOD sperm. A guide RNA (5'-
 208 AAGCACGGATATCGGCAAGA-3') that hybridizes near *Rag1* sequences that encode proline 326 was
 209 generated as described (21). The following mix was electoporated into zygotes: 4μM of purified guide
 210 RNA, 4μM of Cas9-NLS nuclease (qb3 MacroLab Berkeley), 10μM of ssDNA 5'-
 211 C*A*G*AAAGGACTTCACTGGGCTCTCCAGGTCAGTAGGGAAGCACGGATATCGGCAGGAT
 212 CCACAATAGCTGCCCATGACTTTGAGACATCTGAGAATGCAGATCCTACAGAATAGATGCT
 213 T-3' (IDT) that contains nucleotides encoding glycine in place of proline at amino acid 326 and a silent
 214 point mutation of the next codon which introduce a cleavage site for BamHI restriction enzyme. Mice
 215 were screened for introduction of the mutation by PCR with primers 5'-
 216 CGTCGCAAGAGAACTCAGGCT-3' and 5'-TCTCGTCAGTGACAGGAGATGCT-3', followed by
 217 BamHI digestion of the PCR product. Sanger sequencing of PCR products confirmed incorporation of
 218 the desired mutation and no other mutations in the *Rag1* coding sequence. Founders with mutations were
 219 mated with NOD mice to establish NOD *Rag1*^{P326G/+} mice, which were bred with each other to generate
 220 NOD *PG* mice and *WT* littermate controls. NOD mice, C57BL/6 background wild-type and *Rag1*^{-/-} mice
 221 and BALB/c background *Rag1*^{-/-} mice were purchased from Jackson Laboratories. All experimental
 222 mice were littermates or age-matched with control wild-type animals. All experiments were conducted
 223 in accordance with national guidelines and approved by the Institutional Animal Care and Use
 224 Committee (IACUC) of the Children's Hospital of Philadelphia.

225

226 **Flow Cytometry**

227 Single cell suspensions of all organs were ACK lysed to remove erythrocytes before being stained with
 228 LIVE/DEAD™ Fixable Aqua Dead Cell Stain Kit (Invitrogen) and antibodies against respective surface
 229 antigens (BD Biosciences, eBioscience, and Biolegend). Cells were stained for respective intracellular
 230 proteins using Cytofix/Cytoperm Fixation/Permeabilization solution kit (BD Biosciences). For staining

of intracellular cytokines, cells (1×10^6) were cultured in the absence or presence of 50 ng/ml PMA (Sigma) and 1 μ g/ml Ionomycin (Cell Signaling Technology), with 2 μ g/ml brefeldin A (Sigma) and 2 μ M monensin (eBioscience) for 4 hours at 37°C. After staining for LIVE/DEAD and surface antigens, cells were stained for respective cytokines using the Cytofix/Cytoperm kit (BD Bioscience) or transcription factors using the FoxP3/Transcription Factor staining kit (eBiosciences). Data were acquired on a MACSQuant flow cytometer (Miltenyi Biotec) or LSRII Fortessa (BD Biosciences) and analyzed using Flowjo software version 10.5.3 (Tree Star). Analyses were conducted with the following antibodies: anti-CD4 (APC-eFluor 780, clone RM4-5, 1:100), anti-CD25 (BUV395, clone PC61, 1:100), anti-TCR β (BV711, clone H57-597, 1:100), anti-CD19 (PE, clone 1D3, 1:100), anti-CD11b (PE, clone M1/70, 1:100), anti-CD11c (PE, clone HL3, 1:100), anti-TER (PE, clone TER-119, 1:100), anti-TCR γ/δ (PE, clone GL3, 1:100), anti-NK-1.1 (PE, clone PK136, 1:100), and anti-B220 (PE, clone RA3-6B2, 1:100) from BD Biosciences; or anti-CD8 α (Pacific Blue, clone 53-6.7, 1:100), anti-CD44 (APC, PerCP/Cy5.5, clone IM7, 1:100), anti-CD24 (PE/Cy7, clone M1/69, 1:100), anti-CD28 (FITC, clone E18, 1:50), anti-CD45.1 (APC, clone A20, 1:100), anti-CD45.2 (FITC, clone 104, 1:100) from BioLegend.

Cell Sorting

To isolate DN3 cells, thymocytes were stained with PE-labelled CD4, CD8, CD11b, CD11c, NK1.1, Gr1, and Ter119. Non-labelled cells were enriched by MACS depletion using anti-PE microbeads and LS columns (Miltenyi). Enriched cells were then stained with CD4, CD8, CD44, CD25 and anti-lineage/stage markers (TCR β , B220, CD19, CD11b, CD11c, TCR δ , NK1.1, Ter119) antibodies. DN3 thymocytes were sorted on Lin⁻CD4⁻CD8⁻CD44⁻CD25⁺ phenotype. To isolate DP cells, thymocytes were stained with CD4, CD8, and anti-lineage (B220, CD19, CD11b, CD11c, TCR δ , NK1.1, Ter119) antibodies. DP thymocytes were sorted on Lin⁻CD4⁺CD8⁺ phenotype. The cells of interest were isolated using a FACS Aria Fusion sorter (BD Biosciences).

255

256 **Analyzing *Tcrb* and *Tcra* Rearrangements**

257 Genomic DNA was extracted from DN3 or DP cells using the DNeasy Blood and Tissue kit (Qiagen).
 258 To measure *Tcrb* rearrangements in DN3 cells, a Taqman PCR assay was used to quantify V β -to-DJ β 1.1
 259 and V β -to-DJ β 2.1 rearrangement levels with a primer specific for each V β paired and a probe, FAM or
 260 HEX, specific for J β 1.1 or J β 2.1, respectively. Taqman PCR was performed with conditions according
 261 to the manufacturer's instructions (IDT DNA) on the ViiA 7 system (Applied Biosystems). PCR of
 262 CD19 was used for normalization. Primers, probes, and reaction conditions are as described (22). To
 263 assay *Tcrb* repertoire, DNA from DN3 thymocytes were sent to Adaptive Biotechnologies, who used
 264 multiplex PCR to amplify and deep sequence V β -D β -J β rearrangements. Gene segment usage was
 265 analyzed by ImmunoSEQ Analyzer software (Adaptive Biotechnologies). To assess *Tcra*
 266 rearrangements in DP cells, representative V α -J α rearrangements were quantified using a QuantiFast
 267 SYBR Green PCR kit (Qiagen) on the ViiA 7 system (Applied Biosystems) as described (23, 24). PCR
 268 of β 2M was used for normalization. We quantified V α 14-J α 18 rearrangements as described (25).

269

270 **RT-qPCR for mRNA Expression**

271 Respective cell populations were lysed in RLT buffer (Qiagen) containing 2-mercaptoethanol or TRIzol
 272 (Life Technologies) immediately after cell sorting. Total RNA was isolated using RNeasy Mini kit
 273 (Qiagen), treated with DNase (RNase-Free DNase Set, Qiagen), and reverse transcribed to generate
 274 cDNA with High-Capacity RNA-to-cDNA™ Kit (Thermo-Fisher Scientific) according to manufacturer's
 275 directions. The cDNA was subjected to qPCR using Power SYBR Green kit (Applied Biosystems) and
 276 actin, HPRT, ZAP70, and SYK primers (Qiagen). Relative expression was calculated using the ddCt
 277 method, using actin or HPRT as a housekeeping gene, and relevant calibrator sample (explained in
 278 figure legends for respective samples).

279

280 **Competitive Mixed Bone Marrow Chimeric Mice**

281 6-week-old CD45.1 B6.SJL (WT) hosts were lethally irradiated (950 Rad on an X-RAD irradiator).
 282 Bone marrow from at least 3 pooled donors (CD45.1 WT only, CD45.2 RAG P326G mutant only, or
 283 mixed) was sterilely prepped and RBCs lysed. Acceptable mixing (CD45.1/CD45.2 ratio) was
 284 confirmed by flow cytometry. Then, $3\text{-}5 \times 10^6$ total cells were injected intravenously into hosts 6 hours
 285 after irradiation. Mice were monitored over 8 weeks for graft rejection and graft-versus-host disease
 286 prior to euthanasia for analysis. Engraftment was measured using the following antibodies from
 287 Biolegend: anti-CD45.1 (APC-Cy7, clone A20, 1:100), anti-CD45.2 (FITC, Clone 104, 1:100), anti-
 288 NK1.1 (PE, clone PK136, 1:200), anti-CD44 (PerCP-Cy5.5, clone IM7, 1:100), anti CD24 (PE-Cy7,
 289 clone M1/69, 1:100), and anti-TCR β (Pacific blue, clone H57-597, 1:100). CD1d tetramer from the NIH
 290 Tetramer Core (APC, 1:100) was a gift of Dr. Hamid Bassiri. Data were collected on a MACSQuant 10
 291 (Miltenyi Biotec) and were analyzed using FlowJo software, version 10.

292

293 ***In vitro* T Cell Differentiation**

294 Splenocytes from C57BL/6 background *Rag1*^{-/-} mice were irradiated with 2,500 Rads (X-rad irradiator)
 295 and 3×10^5 cells were seeded into each well of a 96 well plate. Naïve T cells (CD90.2⁺, CD4⁺,
 296 CD62L⁺, CD44⁺, CD25⁻) were sorted from C57BL/6 background WT or *Rag1*^{-/-} mice using a FACSaria
 297 Fusion cell sorter (BD Biosciences), and 3×10^4 sorted cells were added to each well containing
 298 irradiated feeder cells. All cells were cultured with 10 $\mu\text{g}/\text{ml}$ anti-CD3 (Biolegend) and 3 $\mu\text{g}/\text{ml}$ anti-
 299 CD28 (Biolegend) in RPMI with 10% FBS, 1% PSG, and 1x NEAA (Invitrogen), 1x Sodium Pyruvate
 300 (Invitrogen), and 0.001% 2-mercaptoethanol. rhIL-2 was used at 30 IU/mL, rmIL-12 and rmIL-4 at 10
 301 ng/mL, rmIL-6 at 20 ng/mL, mTGF β 1 at 1 ng/mL, anti-mIL-4, anti-mIFN γ and anti-mIL-12 at 10
 302 $\mu\text{g}/\text{ml}$ (from Biolegend except rhIL-2, rmIL-12, and rmIL-6 from Prepotech). The following
 303 cytokines/blocking antibodies were used to skew towards respective Th-subsets: Th0 condition: rhIL-2;

Th1 condition: rhIL-2, rmIL-12 and anti-mIL-4; Th2 condition: rhIL-2, rmIL-4, anti-mIFN γ and anti-mIL-12; Th17 condition: rhIL-2, mIL-6, mTGF β 1, anti-mIL-4 and anti-mIL-12; Treg condition: rhIL-2, mTGF β 1, anti-mIL-4 and anti-mIL-12.

LCMV Infections

Mice were infected intraperitoneally with 2×10^5 plaque-forming units of LCMV-Armstrong and euthanized at indicated timepoints. Virus and gp33-tetramer were kindly provided by E. John Wherry (University of Pennsylvania).

T Cell Transfer Model of Colitis

This model was performed as described (26). Briefly, naïve CD4 $^+$ T cells were sorted from respective donor mice (CD90.2 $^+$, CD4 $^+$, CD62L $^+$, CD44 $^-$, CD25 $^-$) on a FACS Aria Fusion cell sorter. 0.5×10^6 cells were transferred in 100 μ l of PBS into each of the donor mice via intraperitoneal injection. Non-transferred mice were injected with PBS. Mice were weighed weekly for the first 4 weeks, and then twice weekly for 4 weeks. Mice were euthanized 8 weeks post-transfer unless severe morbidity enforced premature euthanasia (in accordance with CHOP IACUC approved guidelines) or spontaneous death occurred.

Diabetes Onset in NOD Background Mice

NOD background *WT* and *PG* mice were aged and assessed for spontaneous development of diabetes. Mice were aged and monitored for disease to a maximum of 30 weeks of age. Mice older than 10 weeks of age were weekly tested for glycosuria (Uristix $^{\text{®}}$, Siemens). For positive mice, glycosuria was assayed by a second independent test one to three days later depending on the health of the animal. Mice exhibiting glycosuria from two independent readings and deteriorating health were subject to facial vein

328 bleeding for direct measurement of blood glucose levels using an electronic glucose meter. All such
 329 mice exhibited a blood glucose concentration of ≥ 250 mg/dL and were considered diabetic and
 330 immediately euthanized.

331

332 **Statistical analysis**

333 All data was analyzed in Graphpad Prism 8 using statistical tests indicated in the figure legends. Error
 334 bars indicate mean $\pm$ SEM unless otherwise stated. n.s.= $p>0.05$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$,
 335 **** $p<0.0001$.

Results

Homozygous Rag1 P326G mutation diminishes TCR gene assembly and subsequent thymocyte developmental stage transitions.

To determine potential *in vivo* functions of the RAG1 ubiquitin ligase, we created and analyzed C57BL/6 strain mice homozygous for the Rag1 P326G mutation (*Rag1*^{P326G/P326G} mice; see Materials and Methods) that abrogates RAG1 ubiquitin ligase activity but minimally destabilizes RAG1 protein (16). We initially analyzed $\alpha\beta$ T cell development and TCR gene rearrangements in *Rag1*^{P326G/P326G} (*PG*) mice and wild-type (*WT*) mice. As compared to *WT* mice, *PG* mice have a ~2-fold higher percentage and number of DN3 thymocytes and correspondingly lower percentage and number of DN4 thymocytes (Fig. 1, A and B). *PG* mice also exhibit ~10% reductions in the percentage and number of DP thymocytes and the ratio of CD4⁺ SP thymocytes to DP thymocytes (Fig. 1, C-E). Despite these minor impairments of the DN3-to-DN4 and DP-to-SP thymocyte developmental transitions, *PG* mice have normal numbers of total thymocytes and splenic $\alpha\beta$ T cells (Fig. 1 F). Using Taqman PCR to amplify V β rearrangements to D β 1J β 1.1, D β 1J β 2.1, or D β 2J β 2.1 complexes in non-selected DN3 cells, we observed ~50% lower levels of rearrangement for nearly all individual V β segments in *PG* mice (Fig. 2, A and B). Notably, the recombination of some V β segments involving J β 2.1 were undetectable in *PG* mice (Fig. 2, A and B), indicating altered J β repertoire from *Rag1*^{P326G} mutation. *Rag1* or *Rag2* mutations that substantially decrease RAG endonuclease activity, thymocyte development, and thymic cellularity also alter the usage of individual V β segments (27, 28). However, high-throughput sequencing of *Tcrb* rearrangements in DN3 thymocytes shows similar usage of each V β segment in *PG* and *WT* mice (Fig. 2 C), reflecting the modest reduction of *Tcrb* gene assembly in *PG* mice. This analysis also reveals modest alterations in D β and J β usage, but no difference in CDR3 β lengths, in *Tcrb* rearrangements of *PG* mice (Fig. 2, D and E; data not shown). Using Taqman PCR to amplify a subset of possible V α -to-J α rearrangements in non-selected DP thymocytes, we detected lower levels of

rearrangements for many of the V α /J α combinations, particularly those involving the two most distal J α gene segments (J α 17, J α 2) tested, in *PG* mice (Fig. 2, F). Together, these data indicate that homozygous Rag1^{P326G} mutation produces RAG endonuclease complexes that catalyze decreased levels of *Tcrb* and *Tcra* rearrangements, alter the primary repertoires of *Tcrb* and *Tcra* genes, and support less efficient thymocyte developmental transitions that depend on the functional assembly of a *Tcrb* or *Tcra* gene. The lower efficiency of overall *Tcrb* and *Tcra* gene assembly certainly contributes to the diminished DN3-to-DP and DP-to-SP thymocyte developmental progressions observed in *PG* mice.

Notably, the degrees to which TCR gene rearrangements and $\alpha\beta$ T cell development are reduced in *PG* mice are dramatically less severe than reported for homozygous Rag1 C235Y mutation (15). For example, Rag1^{P236G} mice generate 90% of the normal number of DP thymocytes, whereas homozygous Rag1^{C235Y} mice produce only 1% of the normal number of DP thymocytes. The markedly different severities of these phenotypes suggest that abrogation of RAG1 ubiquitin ligase activity without substantial disruption of RAG1 protein structure only slightly impairs V(D)J recombination and $\alpha\beta$ T cell development. Notably, the Rag1^{P326G} protein is expressed at modestly higher levels in thymocytes than wild-type Rag1 protein (Beilinson *et al.*, DOI 0.1084/jem.20210250), indicating that reduced Rag1 protein expression levels do not contribute to the slightly impaired development of $\alpha\beta$ T cells in *PG* mice. As we were not able to detect Rag1 ubiquitin ligase activity in *WT* mice, we could not confirm the expected abrogation of Rag1 ubiquitin ligase activity in *PG* mice that was observed *in vitro*, leaving open the possibility that the mutant Rag1^{P326G} protein exhibits some level of ubiquitin ligase activity *in vivo*.

To confirm that the impaired development of $\alpha\beta$ T cells in *PG* mice is due to cell intrinsic properties, we made and analyzed competitive bone marrow chimeric mice. These data show no significant differences in contributions of *WT* cells and *PG* cells at each of DN thymocyte stages (Supplemental Figure 1A). In

marked contrast, there were greater contributions of *WT* cells at the DP and both SP thymocyte stages and within the $CD4^+$ and $CD8^+$ splenic $\alpha\beta$ T cell populations (Supplemental Figure 1A, B). The data support the notion that the decreased progression of $Rag1^{P326G}$ thymocytes beyond the DN stage is due to cell intrinsic alterations, most likely from the decreased assembly of TCR β genes.

Reduced *Tcrb* gene assembly impairs DN thymocyte development in homozygous *Rag1* P326G mice.

The greater than normal number of DN3 thymocytes in *PG* mice is not typical of a mutation that impairs V(D)J recombination. It is notable that thymocytes lacking the Syk pre-TCR signaling protein display this phenotype (29, 30), and that RAG DSBs induced in pre-B cells trigger a genetic program that transcriptionally represses Syk protein expression to dampen pre-BCR signaling (31). As *PG* DN cells have reduced numbers of RAG DSBs from decreased *Tcrb* rearrangements, we determined whether these cells express lower than normal levels of Syk protein. We detected a slightly reduced level of Syk protein in *PG* mice (Fig. 3, A and B); however, unexpectedly, this correlated with an increase in the expression of Syk mRNA (Fig. 3 C). Thus, it appears that *Rag1* P326G mutation results in a dissociation between the levels of transcript and protein of the Syk signaling molecule through as yet undetermined mechanisms. Regardless, these data open the possibility that pre-TCR signaled developmental progression of DN3 cells is impaired in *PG* mice.

To distinguish potential contributions of reduced *Tcrb* gene assembly versus decreased pre-TCR-signaling on impaired DN thymocyte development in *PG* mice, we conducted a more refined analysis of DN stage cells based on CD28 protein expression. After successful *Tcrb* gene assembly in $CD28^-CD25^{high}$ DN3a thymocytes, pre-TCR signaling drives differentiation of DN3b and then DN3c thymocytes characterized by upregulation of CD28 and downregulation of CD25 (32). Mirroring relative levels of *Tcrb* rearrangements in DN3 thymocytes, we observed decreased progression of DN3a

cells to the DN3b and DN3c stages, as well as a smaller fraction of DN3a cells expression TCR β protein, in *PG* mice relative to *WT* mice (Fig. 3, D - I). We compared the percentage of DN3a cells expressing TCR β protein to the percentage of these cells that develop into DN3b cells. We found equivalent values between *PG* and *WT* mice (Fig. 3, J). We also compared the percentage of total DN3 cells expressing TCR β protein to the fraction of these cells that progress to the DN4 stage. Here again we found equivalent values between *PG* and *WT* mice (Supplemental Figure 2). Together, these data provide no evidence for diminished pre-TCR signaled development of β -selected DN3 thymocytes. Therefore, the most plausible explanation for impaired DN thymocyte development in *PG* mice is decreased levels of TCR β gene rearrangements.

Homozygous Rag1 P326G mutation lowers expression of the Zap70 signaling protein in thymocytes.

Given the evidence of the decreased levels of Syk protein but not mRNA in DN thymocytes of *PG* mice, we assessed whether the related Zap70 protein that transduces $\alpha\beta$ TCR signals might show a similar pattern in DP thymocytes. To this end, we quantified levels of Zap70 protein and mRNA in thymocytes at each of four stages of $\alpha\beta$ TCR-signaled positive selection as defined by expression of CD69 and CCR7 proteins (Supplemental Figure 3) (33). In *PG* mice, we observed lower than normal Zap70 protein levels, but normal levels of TCR β protein, at each stage of selection (Fig. 4, A-D). We detected normal levels of Zap70 mRNA in cells at each stage assayed in *PG* mice (Fig. 4 E), implying that lower Zap70 mRNA translation and/or Zap70 protein stability is the basis for lower Zap70 expression. RAG DSBs induced during lymphocyte development signal gene expression changes that persist in mature NK cells lacking RAG1 and RAG2 expression (34). We therefore considered whether the Zap70 alterations caused by mutant RAG1 during development would also persist in mature T cells after RAG1 is no longer expressed. We detected normal Zap70 protein levels in splenic CD4⁺ and CD8⁺ $\alpha\beta$ T cells

of *PG* mice (Fig. 4, F and G), indicating that $Rag1^{P326G}$ mutation decreases Zap70 expression specifically in immature $\alpha\beta$ T cells. These data open the possibility that Zap70-mediated $\alpha\beta$ TCR signaling might be weaker than normal in thymocytes and contribute to impaired DP-to-SP thymocyte transition in *PG* mice, beyond that resulting from reduced *Tcra* gene assembly and altered primary repertoires of assembled *Tcrb* and *Tcra* genes.

Homozygous *Rag1* P326G mutation decreases efficiencies of positive, negative, and agonist selection.

In developing B cells, RAG DSBs generated during *Igk* gene assembly signal transcriptional upregulation of a proteins involved in lymphocyte selection and activation to modulate pre-BCR signaling (35). Considering this, the decreased levels of RAG DSBs in *PG* thymocytes, and the important roles of Zap70 in mediating positive, negative, and agonist selection (36, 37), we investigated whether any of these $\alpha\beta$ TCR quality control processes might be impaired in *PG* mice. As a proxy to evaluate positive selection in *PG* mice, we determined the frequency of $\alpha\beta$ TCR expressing ($TCR\beta^+$) Stage 1 thymocytes that initiate positive selection as indicated by progression to Stage 2 in *PG* and *WT* mice (Fig. 5 A). By analyzing only $\alpha\beta$ TCR expressing cells, we circumvent potential effects of decreased *Tcra* gene rearrangements reducing the fraction of DP thymocytes that express $\alpha\beta$ TCRs and are subject to selection. This frequency is lower in *PG* mice (Fig. 5 B), consistent with modestly less efficient positive selection of $\alpha\beta$ TCR expressing DP thymocytes upon homozygous $Rag1^{P326G}$ mutation. Yet, we note that this is an imperfect means to quantify positive selection because it cannot account for possible differences in progression through subsequent stages or cellular fate decisions from specific TCR/pMHC interactions. Thus, we also quantified numbers of cells at each of the four stages of positive selection, finding fewer cells at each stage in *PG* mice relative to *WT* mice (Fig. 5C). Importantly, there is a greater decrease of cellularity in stage 2 (1.71x) as compared to stage 1 (1.25x), indicating fewer

cells are moving from stage 1 to stage 2 in *PG* mice. Finally, we quantified CD69 expression and the ratio of CD69 to TCR β expression at each stage as CD69 upregulation correlates directly with $\alpha\beta$ TCR signaling strength. Although we observed no difference of absolute levels of CD69 expression, *PG* mice have decreased CD69 expression relative to TCR β expression in stage 2 thymocytes that are initiating positively selection (Fig. 5D). Collectively, these analyses indicate that $\alpha\beta$ TCR signaled upregulation of CD69 and positive selection is less efficient in *PG* mice. However, they do not distinguish between potential effects from altered primary repertoires of *Tcrb* and *Tcra* genes versus diminished intrinsic $\alpha\beta$ TCR signaling capacity. They also do not address the possibility that the <10% reduction in numbers of total thymocytes in *PG* mice alters thymic architecture in a manner that increases the fraction of DP cells that engage pMHC presented by TECs or DCs.

To determine whether the Rag1^{P326G} mutation impairs negative selection, we utilized a naturally occurring system of superantigen-mediated negative selection (38). Laboratory mice express remnant mouse mammary tumor virus (MMTV) retroviral products. Some of these products can bind the MHCII I-E^d protein and certain V β peptides, including V β 5.1-5.2 (encoded by *Trbv12.1-12.2*) and V β 12 (encoded by *Trbv15*), to trigger exceptionally strong $\alpha\beta$ TCR signaling that promotes negative selection (38, 39). Our sequencing analysis of *Tcrb* gene rearrangements in DN3 thymocytes shows equivalent representation of *Trbv12.1*, *Trbv12.2*, and *Trbv15* in the primary V β repertoire (Fig. 2 C). Considering that V β repertoire remains constant during DN-to-DP thymocyte development (40), assaying superantigen-mediated deletion of V β 5.1-5.2 or V β 12 accounts for differences of J β repertoire of *Tcrb* genes assembled in *WT* and *PG* mice. Moreover, the reduced level and altered repertoire of *Tcra* genes assembled in DP cells of *PG* mice would not influence superantigen-mediated deletion of thymocytes expressing $\alpha\beta$ TCRs with specific V β peptides. BALB/c, but not C57BL/6, strain mice express I-E^d protein. As a result, thymocytes expressing V β 5.1-5.2 or V β 12 are efficiently deleted in BALB/c, but

not B6, genetic backgrounds. We bred BALB/c background *Rag1*^{-/-} mice with C57BL/6 background *Rag1* *PG* or *WT* mice to generate and analyze *Rag1*^{PG/-} and *Rag1*^{WT/-} mice expressing I-E^d protein. To ascertain whether these mice exhibit a difference in superantigen-mediated negative selection, we measured the frequencies of DP thymocytes that have not completed αβ TCR selection and post-selection SP thymocytes expressing I-E^d:MMTV-reactive (Vβ5⁺ or Vβ12⁺) or -unreactive (Vβ6⁺, Vβ8⁺, or Vβ14⁺ encoded by *Trbv19*, *Trbv13.1-13.3*, or *Trbv31*, respectively) αβ TCRs (Fig. 6, A-F). We calculated the percentages of DP and SP thymocytes expressing each Vβ peptide (Fig. 6, B-F). Our analysis indicates that greater fractions of I-E^d:MMTV-reactive thymocytes proceed through DP-to-SP development in *Rag1*^{PG/-} mice as compared to *Rag1*^{WT/-} mice (Fig. 5, B, C, and G), consistent with impaired negative selection of highly self-reactive thymocytes. In contrast, we found that the types of un-reactive thymocytes assayed develop equivalently in each genotype (Fig. 5, D -G). While the latter finding does not reflect less efficient positive selection, we note that this experiment monitors a small subset of thymocytes expressing specific Vβ peptides and thus is not as sensitive as determining the frequencies of overall DP thymocytes that initiate positive selection. Regardless, the data indicate that *PG* mice exhibit less efficient deletion of thymocytes that express αβ TCRs containing superantigen-reactive Vβ peptides. The plausible reasons for this impaired form of negative selection include that intrinsic αβ TCR signaling capacity is dampened and/or reduced numbers of DP thymocytes prevents some from contacting or maintaining interaction with superantigen.

Finally, to ascertain potential effects of *Rag1*^{P326G} mutation on agonist selection, we focused on assaying iNKT cell development. This lineage develops when TCRs containing an invariant Vα14-Jα18 TCRα chain and TCRβ chains with Vβ regions encoded by *Trbv1*, *Trbv13.1-13.3*, or *Trbv19* bind lipids presented by the MHC-like CD1d protein on different DP thymocytes (10, 41, 42). These interactions induce strong TCR signals that drive agonist selection and differentiation of iNKT cells through distinct

stages that can be identified by binding lipid/tetramers and expression of defined proteins (10, 41, 42). We first used a CD1d tetramer loaded with PBS57, an analogue of the α -galactosylceramide (α -GalCer) glycolipid, to identify α -GalCer-reactive iNKT cells in *WT* and *PG* mice. As compared to *WT* mice, *PG* mice had fewer absolute numbers of thymic α -GalCer-reactive iNKT cells (Fig. 7, A and B). While *PG* mice also tended to have fewer splenic iNKT cells (Fig. 7, A and B), this difference did not reach statistical significance ($p=0.06$) with the numbers of mice analyzed. The lower number of thymic iNKT cells in *PG* mice could arise solely from decreased levels of *Tcra* rearrangements, as is observed by other mutations that decrease these rearrangements over an order of magnitude (43, 44). The levels of V α 14-to-J α 18 rearrangements in total thymocytes of *PG* mice are ~40% relative to *WT* mice (Fig. 7C), which is consistent with reduced generation of V α 14-J α 18 TCR α chains contributing to the lower numbers of thymic α -GalCer-reactive iNKT cells in *PG* mice. However, one advantage of using tetramers is that they identify DP thymocytes that have assembled and express the appropriate *Tcrb* and *Tcra* genes for development into iNKT cells of defined antigen specificity. Thus, as a means to quantify agonist selection of α -GalCer-reactive iNKT cells independent of V α 14-J α 18 recombination levels, we quantified tetramer-staining cells at each stage of iNKT cell development. We observed no significant differences in numbers of Stage 0 (CD24^{hi}CD44⁻NK1.1⁻) and Stage 1 (CD24^{lo}CD44⁻NK1.1⁻) iNKT cells in *WT* and *PG* mice (Fig. 7, D and E). However, *PG* mice had lower numbers of Stage 2 (CD24^{lo}CD44⁺NK1.1⁻) and Stage 3 (CD24^{lo}CD44⁺NK1.1⁺) iNKT cells (Fig. 7, D and E) and diminished expansion of iNKT cells during progression from Stage 0 through Stage 4 (Fig. 7, F). Notably, the median fluorescence intensities of TCR β expression and PBS57/CD1d tetramer binding of thymocytes was similar or slightly higher in *PG* mice relative to *WT* mice (Fig. 7 A, data not shown), indicating that α -GalCer-reactive iNKT cells of *PG* mice have a normal range of intrinsic affinity for this lipid presented by CD1d. Therefore, the plausible reasons for the impaired development of α -GalCer-reactive iNKT cells in *PG* mice are diminished engagements of TCRs with lipid/CD1d

complexes and/or TCR signaling following lipid/CD1d recognition. We also analyzed iNKT development in the competitive setting of mixed bone marrow chimeras comparing mice that received donor bone marrow that was either 100% *PG* or at 1:1 mix of *PG* and *WT* marrow. The very few remnant *WT* cells in the 100% *PG* recipient mice very effectively outcompeted the *PG* cells starting at Stage 1 and increasing in later stages (Supplemental Figure 1, C). The same effect can be seen in the 1:1 mix recipient mice (Supplemental Figure 1, C). Thus, these data indicate that the impaired iNKT cell development in *PG* mice is cell intrinsic, making impaired TCR signaling the most plausible explanation. The molecular basis for impaired TCR signaling could be reduced expression of Zap70 and/or other signaling proteins, altered TCR β CDR3-encoding sequences that generate TCRs less effective at signaling upon recognition of lipids presented by CD1d, or both.

Mature $\alpha\beta$ T cells of homozygous Rag1 P326G mice exhibit normal responses to activation.

RAG endonuclease activity during lymphocyte ontogeny triggers permanent changes in gene expression that correlate with impaired responses of mature lymphocytes (34). Considering that RAG endonuclease function is diminished in thymocytes of *PG* mice, we sought to elucidate whether the mature $\alpha\beta$ T cells of these animals have normal or altered response to activation. We first conducted *in vitro* analyses. Upon stimulation with PMA and ionomycin, splenic CD44^{hi} $\alpha\beta$ T cells from *PG* mice produce normal levels of IFN γ , TNF α , IL-2, and IL-17A (Fig. 8 A). These cells also show normal proliferation upon anti-CD3/anti-CD28 stimulation (Fig. 8 B). Moreover, stimulated CD4⁺ and CD8⁺ $\alpha\beta$ T cells of *PG* mice normally express their canonical transcription factors (Fig. 8, C and D) and cytokines (Fig. 8, E and F) when cultured in respective skewing conditions. We next monitored $\alpha\beta$ T cell responses *in vivo* following infection of mice with an acute strain of lymphocytic choriomeningitis virus (LCMV-Armstrong). We analyzed mice either seven days later to monitor at the immune response peak or 35 days later to monitor memory cells. At each timepoint, we detected similar frequencies and numbers of

antigen-specific CD8⁺ αβ T cells that stain with gp33:H-2Db tetramer in *PG* and *WT* mice (Fig. 8 G), indicating that *PG* mice display normal numbers and dynamic responses of αβ T cells that recognize LCMV-Armstrong. We also observed equivalent dynamic expression of Tbet, Eomes, KLRG1, and CD127 proteins on αβ T cells from *PG* and *WT* mice (Fig. 8 H), indicating normal contraction of short-lived effector cells (SLECs), and sustained memory/memory precursor effector cells (MPECs) in *PG* mice. Finally, we detected similar numbers of antigen specific αβ T cells able to respond to the gp33, NP396, or gp61 viral peptides within a peptide re-stimulation assay (Fig. 8, I and J). Together, these data demonstrate that the mature αβ T cells of *PG* mice have normal responses to activation, at least for the stimuli assayed.

Mature αβ T cells of homozygous Rag1^{P326G} mice exhibit increased predisposition to autoimmunity.

Within a setting of less efficient negative selection as with homozygous Rag1^{P326G} mutation, an increased predisposition to αβ T cell mediated autoimmunity would be expected. However, we did not observe overt autoimmune symptoms within *PG* mice, possibly because mice of the C57BL/6 background are resistant to spontaneous autoimmune diseases (45, 46). Thus, to investigate whether the Rag1^{P326G} mutation can predispose to autoimmunity, we utilized an induced model of autoimmune colitis in C57BL/6 mice and a spontaneous model of autoimmune diabetes in NOD mice.

The colitis model involves transferring into *Rag1*^{-/-} mice naïve CD4⁺ effector αβ T cells, which become activated and trigger cytokine-driven colitis and associated systemic pathologies including weight loss due to the absence of Tregs (26, 47, 48). We observed accelerated wasting of *Rag1*^{-/-} mice receiving *PG* cells relative to *WT* cells or no transfer (Fig. 9 A). We analyzed mice at eight weeks after transfer. As compared to *WT* recipients, we detected greater numbers of total splenic CD4⁺ T cells and of CD4⁺ T

cells producing IFN γ , TNF α , and/or GM-CSF cytokines (Fig. 9, B and C). Colon length is used as a measure of disease severity in this model, and mice receiving transferred *PG* cells had shorter colons relative to mice receiving *WT* cells, consistent with worse disease caused by *PG* cells (Fig. 9 D). Histologic assessment and scoring in this colitis model has been formalized (48), and can be grossly broken down into scores related to inflammatory cell infiltration and scores related to injury of tissue architecture. There was no statistical difference in inflammation scores between mice receiving *WT* and *PG* cells. However, tissue damage scores were higher in mice receiving *PG* cells, suggesting that, on a per cell basis, *PG* cells induce greater tissue injury in this model (Fig. 9, E-F). Notably, specifically in *PG*-transferred animals, we observed autoimmune pathologies in addition to colitis, including dermatitis, chylous ascites, severe small intestine enlargement, and neurological stress (Fig. 9 H). Dermatitis was most common and involved massive inflammatory infiltration and tissue damage (Fig. 9 I). Consistent with more widespread autoimmune conditions in *PG*-transferred mice, we significant splenomegaly in these mice (Fig. 9 J). Finally, while not statistically significant, we observed an increased rate of premature death in the mice transferred with *PG* cells compared to *WT* cells (Fig. 9 K), which was especially surprising because death is an unexpected outcome in this model. Collectively, the data from this colitis model indicate that CD4⁺ effector T cells from *PG* mice have greater than normal intrinsic autoimmune potential.

For the spontaneous model of autoimmune diabetes, we utilized CRISPR/Cas9-mediated gene editing to establish NOD mice carrying the *Rag1*^{P326G} mutation. NOD mice develop spontaneous $\alpha\beta$ T cell driven diabetes at an incidence of 60–80% in females and 20–30% in males (49, 50). We bred together NOD:*Rag1*^{P326G/WT} mice to create cohorts of NOD *PG* and *WT* females, which we aged and screened for diabetes by monitoring glycosuria and blood glucose levels, euthanizing diabetic mice and confirming disease by blood glucose levels. We observed accelerated onset of diabetes in *PG* mice compared to *WT* mice, with median age of disease-free survival 120 and 175 days, respectively (Fig. 9 L). We started this

604 study with the assumption that NOD *RagI*^{P326G} mice will experience an early event due to autoimmunity
 605 risk as compared to NOD *RagI*^{WT} mice, and if events do not occur early then they are unlikely to occur
 606 at all. Therefore, we evaluated potential significance of the different median disease-free survival using
 607 the Gehan-Breslow-Wilcoxon test that models such un-proportional hazard, finding the difference
 608 significant (p=0.049). However, we also applied the long-rank Mantel-Cox test that is more stringent,
 609 but should be used when the assumption is of proportional hazards between groups, finding the
 610 difference near, but not reaching, significance (p=0.1151). Unfortunately, the disruption of research
 611 from the Covid19 pandemic prevented us from continuing to generate and monitor mice so that we
 612 could more rigorously evaluate the difference in diabetes-free survival, as well as pathologically and
 613 immunologically characterize disease in afflicted mice. Finally, we note that this experiment does not
 614 distinguish between potential abnormalities of T and B cells of the *PG* background.

Discussion

Our analysis of homozygous Rag1^{P326G} mice indicates roles for the integrity of the RAG1 RING domain ubiquitin ligase in stimulating TCR α and TCR β gene assembly and enhancing both positive selection of conventional $\alpha\beta$ T cells and agonist selection of iNKT cells. The assembly and expression of functional *Tcrb* genes is rate limiting for thymocyte development from the DN3a stage through the DP stage. Our data show that the lower levels of V β rearrangements predominantly, if not solely, underlie the impaired progression of thymocytes through these early developmental stages in *PG* mice. The assembly and expression of *Tcra* genes is rate limiting for DP-to-SP thymocyte development. On average, multiple rounds of successive V α -to-J α rearrangements occur in DP thymocytes before resulting $\alpha\beta$ TCRs signal positive selection, negative selection, or agonist selection. For agonist selection of iNKT cells, V α -to-J α rearrangements must proceed until generation of V α 14-J α 18 rearrangements. A small subset of DP cells with V α 14-J α 18 rearrangements can produce $\alpha\beta$ TCRs that recognize lipid/CD1d complexes and signal development of iNKT cells from Stage 0 through Stage 3. The lower levels of V α rearrangements in *PG* mice likely contributes to less robust positive selection of conventional $\alpha\beta$ T cells and agonist selection of iNKT cells by reducing the generation of $\alpha\beta$ TCR-expressing DP thymocytes. Although the modestly decreased level of V α 14-to-J α 18 rearrangements in *PG* thymocytes does not reduce the number of α -GalCer-reactive Stage 0 iNKT that arise that can be subject to agonist selection, our analysis of iNKT cells of a single specificity leaves open the possibility that decreased V α 14-to-J α 18 rearrangements impairs the genesis of a replete iNKT cell population in *PG* mice. The amino acids encoded by individual gene segments and CDR3 nucleotides influence $\alpha\beta$ TCR antigen recognition and ensuing TCR signaling. While we did not comprehensively study *Tcrb* and *Tcra* gene repertoires, our Taqman and Adaptive Immunoseq analyses of *Tcrb* gene rearrangements indicate minor differences of D β and J β repertoire between *WT* and *PG* mice. These in turn could generate subtle differences in TCR β CDR3 repertoire that might contribute to decrease efficiencies of

positive selection and agonist selection. Based on *in vitro* experiments (16), the molecular basis for reduced RAG endonuclease activity in *PG* mice could result from P326G mutation altering RAG complex tertiary structure, abrogating Rag1 ubiquitin ligase activity, and/or impairing Rag1 interaction with a protein that regulates RAG endonuclease activity. Moreover, while an intrinsically less efficient RAG endonuclease could account entirely for the reduced assembly of *Tcrb* and *Tcra* genes in *PG* mice, our analyses do not address potential roles of the RAG1 ubiquitin ligase in promoting chromatin accessibility and/or RAG binding of these loci. Distinguishing among all of these possibilities requires identifying all thymocyte proteins that bind Rag1 and/or are substrates of the Rag1 ubiquitin ligase, and then determining the effects of preventing their binding or ubiquitylation by Rag1.

Our findings of less efficient superantigen-mediated negative selection, impaired agonist selection of α -GalCer-reactive iNKT cells, and mature $\alpha\beta$ T cells of higher autoimmune potential in *PG* mice imply potential roles for the RAG1 ubiquitin ligase domain beyond stimulating TCR gene recombination and shaping TCR repertoire. The molecular basis for superantigen-mediated negative selection is that these antigens interact simultaneously with MHC proteins on TECs and specific V β peptides on DP cells to illicit very strong TCR signaling that triggers apoptosis. The reduced levels of V β gene rearrangements in *PG* mice do not alter primary V β repertoire, and to our knowledge there is no evidence that differences among D β -, J β -, or CDR3 β encoded nucleotides influence superantigen-directed negative selection. The initiation and progression of iNKT cell development is driven by strong TCR signaling upon recognition of lipid/CD1d complexes. In theory, ongoing *Tcra* gene recombination before Stage 0 or Stage 1 iNKT cells silence RAG expression could replace the invariant V α 14-J α 18 rearrangement and divert these cells into conventional $\alpha\beta$ T cells. If this scenario occurs, the lower levels of *Tcra* gene rearrangements in *PG* mice would be predicted to diminish the chance of this happening and thereby promote developmental progression of iNKT cells. However, the slightly altered D β and J β usage in

663 *Tcrb* gene rearrangements of *PG* mice could generate $\alpha\beta$ TCRs less effective at signaling upon
664 recognition of lipid/CD1d complexes, leading to impaired iNKT cell development. The higher
665 autoimmune potential of $\alpha\beta$ T cells in *PG* mice most likely results from impaired negative selection of
666 high self-reactive thymocytes. In mice with normal RAG endonuclease activity, highly self-reactive $\alpha\beta$
667 TCRs signal negative selection by apoptosis before further *Tcra* gene recombination can replace these
668 hazardous receptors with new ones (51). The reduced levels of *Tcra* gene rearrangements in *PG* mice
669 would be expected to increase the likelihood that negatively-selecting TCRs remain expressed and
670 trigger apoptosis. Considering all of the above, the most plausible reasons for less efficient negative
671 selection, impaired iNKT cell development, and higher autoimmune potential of CD4⁺ T cells in *PG*
672 mice are: 1) less robust TCR signaling upon recognition of self-antigens, and/or 2) decreased
673 frequencies or durations that $\alpha\beta$ TCRs engage self-antigens due to the 50% fewer DP thymocytes and
674 potentially altered thymic microenvironment. Our mixed bone marrow chimera analysis indicates that
675 such cell extrinsic mechanisms cannot explain impaired agonist selection of α -GalCer-reactive iNKT
676 cells, leaving decreased $\alpha\beta$ TCR signaling as a plausible mechanism. Genetic mutations in mice that
677 substantially impair V(D)J recombination also dramatically (>10-fold) reduce thymic cellularity and
678 consequently alter thymic architecture through defective maturation of TECs and/or distribution of DCs,
679 which collectively impairs negative selection and increases development of hazardous autoreactive $\alpha\beta$ T
680 cells (52). Although the slightly lower than normal levels of TCR gene rearrangements in *PG* mice
681 cause only a minor decrease in DP thymocyte cellularity, it will be important to determine if the
682 Rag1^{P326G} mutation has any effects on the thymic microenvironment that could underlie less efficient
683 negative selection.

684

685 The lower than normal expression of Zap70 protein specifically in thymocytes of *PG* mice offers
686 support and a potential basis for impaired $\alpha\beta$ TCR signaling contributing to less efficient thymocyte

687 selection. As reduced Zap70 activity impairs TCR signaling, positive selection, negative selection, and
688 agonist selection of iNKT cells (53-55), attenuated TCR signaling from lower Zap70 expression, even
689 the subtle decrease observed, is a plausible explanation for these phenotypes of *PG* mice. However,
690 definitive conclusions necessitate elucidating how Rag1^{P326G} mutation decreases Zap70 protein levels
691 and determining effects of disrupting these mechanisms without lowering RAG endonuclease activity.
692 RAG DSBs induced during Igκ gene recombination in pre-B cells activate diverse signaling pathways
693 that transcriptionally alter the expression of several antigen receptor signaling proteins to diminish pre-
694 BCR signaling (31, 56). Thus, it is conceivable that RAG DSBs induced during *Tcra* gene
695 rearrangements in DP thymocytes could trigger nucleus-to-cytoplasm signaling cascades that post-
696 transcriptionally increase Zap70 protein expression. In this scenario, the decreased number of RAG
697 DSBs generated during *Tcra* gene rearrangements in *PG* DP thymocytes might impair activation of such
698 signals. A not-mutually exclusive possibility is that Rag1^{P326G} mutation prevents the Rag1-directed
699 ubiquitylation of nuclear proteins that amplify the signals from RAG DSBs that up-regulate Zap70
700 protein expression. Another possibility is that integrity of the Rag1 ubiquitin ligase domain is necessary
701 for binding a factor that influences signals from RAG DSBs. We did not detect differences in Zap70
702 activation or other aspects of TCR signaling between *WT* and *PG* thymocytes after anti-CD3 stimulation
703 (data not shown). However, this approach for activating αβ TCR signaling is not as sensitive for
704 detecting differences as using tetramers or DCs to present peptides to transgenic αβ TCRs of fixed
705 specificity (9, 57, 58). Such experiments, which are beyond the scope of our current study, are critical to
706 determine whether Rag1^{P326G} mutation diminishes αβ TCR signaling specifically in DP thymocytes.
707 Considering that Zap70 might not be the only, or even a relevant, TCR signaling protein affected by this
708 Rag1 mutation, findings that confirm diminished TCR signaling would warrant identifying all changes
709 in thymocyte protein expression triggered by RAG DSBs and elucidating the effect of Rag1^{P326G}
710 mutation on these. This information would suggest approaches to disrupt signaling from RAG DSBs
711 without reducing RAG endonuclease activity to separate impacts of Rag1^{P326G} mutation on altering

expression of TCR signaling proteins versus reducing *Tcrb* and *Tcra* gene rearrangements, altering assembled *Tcrb* and *Tcra* gene repertoires, and decreasing thymic cellularity.

The requirement for TCR selection on self-peptides displayed by self-proteins presents major challenges for establishing conventional and non-conventional $\alpha\beta$ T cell populations that recognize and respond to a vast array of diverse foreign peptides, but not self-antigens. Thus, cell intrinsic mechanisms, at least some of which are Rag1-independent, have evolved to render thymocytes more sensitive to TCR stimulation than their mature $\alpha\beta$ T cell counterparts (57, 59-62). Several factors have been implicated in regulating this “thymic TCR” tuning (58, 63-67). Our observations of reduced thymic Zap70 expression and less efficient positive selection, negative selection, and agonist selection in *PG* mice are consistent with the possibility that the RAG1 protein augments thymic TCR signal tuning. Regardless, our study provides a phenotypic framework from which to elucidate precise molecular mechanisms by which the integrity of the RAG1 ubiquitin ligase domain stimulates TCR gene assembly and proper selection $\alpha\beta$ TCRs to establish replete populations of $\alpha\beta$ T lineage cells and central tolerance.

731 **Acknowledgements**

732 We thank Kyutae Lee for designing the gene-editing strategy and Adele Harman of the CHOP
733 Transgenic Core to generate the NOD *PG* mice. The authors have no competing financial interests to
734 declare.

References

1. Bassing, C. H., W. Swat, and F. W. Alt. 2002. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* 109 Suppl: S45-55.
2. Schatz, D. G., and P. C. Swanson. 2011. V(D)J recombination: mechanisms of initiation. *Annu Rev Genet* 45: 167-202.
3. Bhandoola, A., H. von Boehmer, H. T. Petrie, and J. C. Zuniga-Pflucker. 2007. Commitment and developmental potential of extrathymic and intrathymic T cell precursors: plenty to choose from. *Immunity* 26: 678-689.
4. von Boehmer, H., and F. Melchers. 2010. Checkpoints in lymphocyte development and autoimmune disease. *Nat Immunol* 11: 14-20.
5. Carico, Z., and M. S. Krangel. 2015. Chromatin Dynamics and the Development of the TCRalpha and TCRdelta Repertoires. *Adv Immunol* 128: 307-361.
6. Hogquist, K. A., and S. C. Jameson. 2014. The self-obsession of T cells: how TCR signaling thresholds affect fate 'decisions' and effector function. *Nat Immunol* 15: 815-823.
7. Gascoigne, N. R., V. Rybakina, O. Acuto, and J. Brzostek. 2016. TCR Signal Strength and T Cell Development. *Annu Rev Cell Dev Biol* 32: 327-348.
8. Klein, L., B. Kyewski, P. M. Allen, and K. A. Hogquist. 2014. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol* 14: 377-391.
9. Fu, G., V. Rybakina, J. Brzostek, W. Paster, O. Acuto, and N. R. Gascoigne. 2014. Fine-tuning T cell receptor signaling to control T cell development. *Trends Immunol* 35: 311-318.
10. Krovi, S. H., and L. Gapin. 2018. Invariant Natural Killer T Cell Subsets-More Than Just Developmental Intermediates. *Front Immunol* 9: 1393.
11. Jones, J. M., and M. Gellert. 2003. Autoubiquitylation of the V(D)J recombinase protein RAG1. *Proc Natl Acad Sci U S A* 100: 15446-15451.
12. Singh, S. K., and M. Gellert. 2015. Role of RAG1 autoubiquitination in V(D)J recombination. *Proc Natl Acad Sci U S A* 112: 8579-8583.
13. Jones, J. M., A. Bhattacharyya, C. Simkus, B. Vallieres, T. D. Veenstra, and M. Zhou. 2011. The RAG1 V(D)J recombinase/ubiquitin ligase promotes ubiquitylation of acetylated, phosphorylated histone 3.3. *Immunol Lett* 136: 156-162.
14. Grazini, U., F. Zanardi, E. Citterio, S. Casola, C. R. Goding, and F. McBlane. 2010. The RING domain of RAG1 ubiquitylates histone H3: a novel activity in chromatin-mediated regulation of V(D)J joining. *Mol Cell* 37: 282-293.
15. Deng, Z., H. Liu, and X. Liu. 2015. RAG1-mediated ubiquitylation of histone H3 is required for chromosomal V(D)J recombination. *Cell Res* 25: 181-192.
16. Simkus, C., P. Anand, A. Bhattacharyya, and J. M. Jones. 2007. Biochemical and folding defects in a RAG1 variant associated with Omenn syndrome. *J Immunol* 179: 8332-8340.
17. Villa, A., C. Sobacchi, L. D. Notarangelo, F. Bozzi, M. Abinun, T. G. Abrahamsen, P. D. Arkwright, M. Baniyash, E. G. Brooks, M. E. Conley, P. Cortes, M. Duse, A. Fasth, A. M. Filipovich, A. J. Infante, A. Jones, E. Mazzolari, S. M. Muller, S. Pasic, G. Rechavi, M. G. Sacco, S. Santagata, M. L. Schroeder, R. Seger, D. Strina, A. Ugazio, J. Valiaho, M. Vihinen, L. B. Vogler, H. Ochs, P. Vezzoni, W. Friedrich, and K. Schwarz. 2001. V(D)J recombination defects in lymphocytes due to RAG mutations: severe immunodeficiency with a spectrum of clinical presentations. *Blood* 97: 81-88.
18. Zheng, N., P. Wang, P. D. Jeffrey, and N. P. Pavletich. 2000. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* 102: 533-539.
19. Albert, T. K., H. Hanzawa, Y. I. Legtenberg, M. J. de Ruwe, F. A. van den Heuvel, M. A. Collart, R. Boelens, and H. T. Timmers. 2002. Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. *EMBO J* 21: 355-364.

- 784 20. Schwenk, F., U. Baron, and K. Rajewsky. 1995. A cre-transgenic mouse strain for the ubiquitous
785 deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic Acids Res* 23:
786 5080-5081.
- 787 21. Lee, K. D., and C. H. Bassing. 2020. Two Successive Inversional Vbeta Rearrangements on a
788 Single Tcrb Allele Can Contribute to the TCRbeta Repertoire. *J Immunol* 204: 78-86.
- 789 22. Gopalakrishnan, S., K. Majumder, A. Predeus, Y. Huang, O. I. Koues, J. Verma-Gaur, S.
790 Loguercio, A. I. Su, A. J. Feeney, M. N. Artyomov, and E. M. Oltz. 2013. Unifying model for
791 molecular determinants of the preselection Vbeta repertoire. *Proc Natl Acad Sci U S A* 110:
792 E3206-3215.
- 793 23. Shih, H. Y., J. Verma-Gaur, A. Torkamani, A. J. Feeney, N. Galjart, and M. S. Krangel. 2012.
794 Tcr gene recombination is supported by a Tcr enhancer- and CTCF-dependent chromatin hub.
795 *Proc Natl Acad Sci U S A* 109: E3493-3502.
- 796 24. Chen, L., Z. Carico, H. Y. Shih, and M. S. Krangel. 2015. A discrete chromatin loop in the
797 mouse Tcr-Tcrd locus shapes the TCRdelta and TCRalpha repertoires. *Nat Immunol* 16: 1085-
798 1093.
- 799 25. Gapin, L., J. L. Matsuda, C. D. Surh, and M. Kronenberg. 2001. NKT cells derive from double-
800 positive thymocytes that are positively selected by CD1d. *Nat Immunol* 2: 971-978.
- 801 26. Ostanin, D. V., J. Bao, I. Koboziev, L. Gray, S. A. Robinson-Jackson, M. Kosloski-Davidson, V.
802 H. Price, and M. B. Grisham. 2009. T cell transfer model of chronic colitis: concepts,
803 considerations, and tricks of the trade. *Am J Physiol Gastrointest Liver Physiol* 296: G135-146.
- 804 27. Liang, H. E., L. Y. Hsu, D. Cado, L. G. Cowell, G. Kelsoe, and M. S. Schlissel. 2002. The
805 "dispensable" portion of RAG2 is necessary for efficient V-to-DJ rearrangement during B and T
806 cell development. *Immunity* 17: 639-651.
- 807 28. Horowitz, J. E., and C. H. Bassing. 2014. Noncore RAG1 regions promote Vbeta rearrangements
808 and alphabeta T cell development by overcoming inherent inefficiency of Vbeta recombination
809 signal sequences. *J Immunol* 192: 1609-1619.
- 810 29. Colucci, F., D. Guy-Grand, A. Wilson, M. Turner, E. Schweighoffer, V. L. Tybulewicz, and J. P.
811 Di Santo. 2000. A new look at Syk in alpha beta and gamma delta T cell development using
812 chimeric mice with a low competitive hematopoietic environment. *J Immunol* 164: 5140-5145.
- 813 30. Palacios, E. H., and A. Weiss. 2007. Distinct roles for Syk and ZAP-70 during early thymocyte
814 development. *J Exp Med* 204: 1703-1715.
- 815 31. Bednarski, J. J., R. Pandey, E. Schulte, L. S. White, B. R. Chen, G. J. Sandoval, M. Kohyama, M.
816 Haldar, A. Nickless, A. Trott, G. Cheng, K. M. Murphy, C. H. Bassing, J. E. Payton, and B. P.
817 Sleckman. 2016. RAG-mediated DNA double-strand breaks activate a cell type-specific
818 checkpoint to inhibit pre-B cell receptor signals. *J Exp Med* 213: 209-223.
- 819 32. Williams, J. A., K. S. Hathcock, D. Klug, Y. Harada, B. Choudhury, J. P. Allison, R. Abe, and R.
820 J. Hodes. 2005. Regulated costimulation in the thymus is critical for T cell development:
821 dysregulated CD28 costimulation can bypass the pre-TCR checkpoint. *J Immunol* 175: 4199-
822 4207.
- 823 33. Kimura, M. Y., J. Thomas, X. Tai, T. I. Guintier, M. Shinzawa, R. Etzensperger, Z. Li, P. Love, T.
824 Nakayama, and A. Singer. 2016. Timing and duration of MHC I positive selection signals are
825 adjusted in the thymus to prevent lineage errors. *Nat Immunol* 17: 1415-1423.
- 826 34. Karo, J. M., D. G. Schatz, and J. C. Sun. 2014. The RAG recombinase dictates functional
827 heterogeneity and cellular fitness in natural killer cells. *Cell* 159: 94-107.
- 828 35. Bednarski, J. J., A. Nickless, D. Bhattacharya, R. H. Amin, M. S. Schlissel, and B. P. Sleckman.
829 2012. RAG-induced DNA double-strand breaks signal through Pim2 to promote pre-B cell
830 survival and limit proliferation. *J Exp Med* 209: 11-17.

36. Negishi, I., N. Motoyama, K. Nakayama, K. Nakayama, S. Senju, S. Hatakeyama, Q. Zhang, A. C. Chan, and D. Y. Loh. 1995. Essential role for ZAP-70 in both positive and negative selection of thymocytes. *Nature* 376: 435-438.
37. Zhao, M., M. N. D. Svensson, K. Venken, A. Chawla, S. Liang, I. Engel, P. Mydel, J. Day, D. Elewaut, N. Bottini, and M. Kronenberg. 2018. Altered thymic differentiation and modulation of arthritis by invariant NKT cells expressing mutant ZAP70. *Nat Commun* 9: 2627.
38. Jordan, M. S., J. E. Smith, J. C. Burns, J. E. Austin, K. E. Nichols, A. C. Aschenbrenner, and G. A. Koretzky. 2008. Complementation in trans of altered thymocyte development in mice expressing mutant forms of the adaptor molecule SLP76. *Immunity* 28: 359-369.
39. Surh, C. D., and J. Sprent. 1994. T-cell apoptosis detected in situ during positive and negative selection in the thymus. *Nature* 372: 100-103.
40. Wilson, A., C. Marechal, and H. R. MacDonald. 2001. Biased V beta usage in immature thymocytes is independent of DJ beta proximity and pT alpha pairing. *J Immunol* 166: 51-57.
41. Rossjohn, J., D. G. Pellicci, O. Patel, L. Gapin, and D. I. Godfrey. 2012. Recognition of CD1d-restricted antigens by natural killer T cells. *Nat Rev Immunol* 12: 845-857.
42. Dashtsoodol, N., S. Bortoluzzi, and M. Schmidt-Supprian. 2019. T Cell Receptor Expression Timing and Signal Strength in the Functional Differentiation of Invariant Natural Killer T Cells. *Front Immunol* 10: 841.
43. Callen, E., R. B. Faryabi, M. Luckey, B. Hao, J. A. Daniel, W. Yang, H. W. Sun, G. Dressler, W. Peng, H. Chi, K. Ge, M. S. Krangel, J. H. Park, and A. Nussenzweig. 2012. The DNA damage- and transcription-associated protein pax1 controls thymocyte development and emigration. *Immunity* 37: 971-985.
44. D'Cruz, L. M., J. Knell, J. K. Fujimoto, and A. W. Goldrath. 2010. An essential role for the transcription factor HEB in thymocyte survival, Tcr rearrangement and the development of natural killer T cells. *Nat Immunol* 11: 240-249.
45. Jiang, W., M. S. Anderson, R. Bronson, D. Mathis, and C. Benoist. 2005. Modifier loci condition autoimmunity provoked by Aire deficiency. *J Exp Med* 202: 805-815.
46. Jin, Y., A. Lee, J. H. Oh, H. W. Lee, and S. J. Ha. 2019. The R229Q mutation of Rag2 does not characterize severe immunodeficiency in mice. *Sci Rep* 9: 4415.
47. Mottet, C., H. H. Uhlig, and F. Powrie. 2003. Cutting edge: cure of colitis by CD4+CD25+ regulatory T cells. *J Immunol* 170: 3939-3943.
48. Eri, R., M. A. McGuckin, and R. Wadley. 2012. T cell transfer model of colitis: a great tool to assess the contribution of T cells in chronic intestinal inflammation. *Methods Mol Biol* 844: 261-275.
49. Anderson, M. S., and J. A. Bluestone. 2005. The NOD mouse: a model of immune dysregulation. *Annu Rev Immunol* 23: 447-485.
50. Delovitch, T. L., and B. Singh. 1997. The nonobese diabetic mouse as a model of autoimmune diabetes: immune dysregulation gets the NOD. *Immunity* 7: 727-738.
51. Kreslavsky, T., H. J. Kim, S. B. Koralov, D. Ghitza, T. Buch, H. Cantor, K. Rajewsky, and H. von Boehmer. 2013. Negative selection, not receptor editing, is a physiological response of autoreactive thymocytes. *J Exp Med* 210: 1911-1918.
52. Marrella, V., P. L. Poliani, L. D. Notarangelo, F. Grassi, and A. Villa. 2014. Rag defects and thymic stroma: lessons from animal models. *Front Immunol* 5: 259.
53. Sakaguchi, N., T. Takahashi, H. Hata, T. Nomura, T. Tagami, S. Yamazaki, T. Sakihama, T. Matsutani, I. Negishi, S. Nakatsuru, and S. Sakaguchi. 2003. Altered thymic T-cell selection due to a mutation of the ZAP-70 gene causes autoimmune arthritis in mice. *Nature* 426: 454-460.
54. Mallaun, M., G. Zenke, and E. Palmer. 2010. A discrete affinity-driven elevation of ZAP-70 kinase activity initiates negative selection. *J Recept Signal Transduct Res* 30: 430-443.

55. Hsu, L. Y., Y. X. Tan, Z. Xiao, M. Malissen, and A. Weiss. 2009. A hypomorphic allele of ZAP-70 reveals a distinct thymic threshold for autoimmune disease versus autoimmune reactivity. *J Exp Med* 206: 2527-2541.
56. Bredemeyer, A. L., B. A. Helmink, C. L. Innes, B. Calderon, L. M. McGinnis, G. K. Mahowald, E. J. Gapud, L. M. Walker, J. B. Collins, B. K. Weaver, L. Mandik-Nayak, R. D. Schreiber, P. M. Allen, M. J. May, R. S. Paules, C. H. Bassing, and B. P. Sleckman. 2008. DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes. *Nature* 456: 819-823.
57. Davey, G. M., S. L. Schober, B. T. Endrizzi, A. K. Dutcher, S. C. Jameson, and K. A. Hogquist. 1998. Preselection thymocytes are more sensitive to T cell receptor stimulation than mature T cells. *J Exp Med* 188: 1867-1874.
58. Fu, G., J. Casas, S. Rigaud, V. Rybakina, F. Lambalez, J. Brzostek, J. A. Hoerter, W. Paster, O. Acuto, H. Cheroutre, K. Sauer, and N. R. Gascoigne. 2013. Themis sets the signal threshold for positive and negative selection in T-cell development. *Nature* 504: 441-445.
59. Lucas, B., I. Stefanova, K. Yasutomo, N. Dautigny, and R. N. Germain. 1999. Divergent changes in the sensitivity of maturing T cells to structurally related ligands underlies formation of a useful T cell repertoire. *Immunity* 10: 367-376.
60. Yachi, P. P., C. Lotz, J. Ampudia, and N. R. Gascoigne. 2007. T cell activation enhancement by endogenous pMHC acts for both weak and strong agonists but varies with differentiation state. *J Exp Med* 204: 2747-2757.
61. Pircher, H., U. H. Rohrer, D. Moskophidis, R. M. Zinkernagel, and H. Hengartner. 1991. Lower receptor avidity required for thymic clonal deletion than for effector T-cell function. *Nature* 351: 482-485.
62. Sant'Angelo, D. B., and C. A. Janeway, Jr. 2002. Negative selection of thymocytes expressing the D10 TCR. *Proc Natl Acad Sci U S A* 99: 6931-6936.
63. Fu, G., S. Vallee, V. Rybakina, M. V. McGuire, J. Ampudia, C. Brockmeyer, M. Salek, P. R. Fallen, J. A. Hoerter, A. Munshi, Y. H. Huang, J. Hu, H. S. Fox, K. Sauer, O. Acuto, and N. R. Gascoigne. 2009. Themis controls thymocyte selection through regulation of T cell antigen receptor-mediated signaling. *Nat Immunol* 10: 848-856.
64. Lesourne, R., S. Uehara, J. Lee, K. D. Song, L. Li, J. Pinkhasov, Y. Zhang, N. P. Weng, K. F. Wildt, L. Wang, R. Bosselut, and P. E. Love. 2009. Themis, a T cell-specific protein important for late thymocyte development. *Nat Immunol* 10: 840-847.
65. Johnson, A. L., L. Aravind, N. Shulzhenko, A. Morgun, S. Y. Choi, T. L. Crockford, T. Lambe, H. Domschens, E. M. Kucharska, L. Zheng, C. G. Vinuesa, M. J. Lenardo, C. C. Goodnow, R. J. Cornall, and R. H. Schwartz. 2009. Themis is a member of a new metazoan gene family and is required for the completion of thymocyte positive selection. *Nat Immunol* 10: 831-839.
66. Wang, D., M. Zheng, L. Lei, J. Ji, Y. Yao, Y. Qiu, L. Ma, J. Lou, C. Ouyang, X. Zhang, Y. He, J. Chi, L. Wang, Y. Kuang, J. Wang, X. Cao, and L. Lu. 2012. Tspal is involved in late thymocyte development through the regulation of TCR-mediated signaling. *Nat Immunol* 13: 560-568.
67. Li, Q. J., J. Chau, P. J. Ebert, G. Sylvester, H. Min, G. Liu, R. Braich, M. Manoharan, J. Soutschek, P. Skare, L. O. Klein, M. M. Davis, and C. Z. Chen. 2007. miR-181a is an intrinsic modulator of T cell sensitivity and selection. *Cell* 129: 147-161.

928
929
930

931 **Footnotes**

932 This work was supported by the Fondation ARC pour la recherche sur le cancer (C.M.), La Ligue contre
 933 le cancer (C.M.), Fondation Pasteur Mutualite (C.M.), Nancy Taylor Foundation for Chronic Disease
 934 (E.M.B.) and NIH grants R03 AI097611 (J.M.J.), R01 HL112836 (E.M.B.), R01 AI121250-A1 (E.M.B.),
 935 R21 A1 135435 (C.H.B.), and RO1 AI 112621 (C.H.B.).

936
 937
 938

Figure Legends

Figure 1. *PG* mice exhibit impaired thymocyte developmental transitions that require TCR gene assembly. **(A)** Representative flow cytometry plots of DN stages of thymocyte development showing the frequency of cells at each stage. **(B)** Quantification of numbers of cells in each DN stage. **(C)** Representative flow cytometry plots of DN, DP, CD4⁺ SP, and CD8⁺ SP thymocytes. **(D)** Quantification of numbers of cells in each stage. **(E)** Quantification of the ratio of CD4⁺ SP or CD8⁺ SP thymocytes to DP thymocytes. **(F-G)** Quantification of the numbers of total thymocytes (F) and splenic αβ T cells (G). (A-G) All data were collected from 6-10 weeks old mice and (B, D, E, F, and G) are combined from four independent experiments including a total of 14-16 mice per genotype. Bars indicate mean +/- SEM. Stats: 2-way ANOVA with Sidak multiple comparison test. n.s. p>0.05; * p<0.05, **p<0.01, *** p<0.001, **** p<0.0001.

Figure 2. Thymocytes of *PG* mice support diminished levels of *Tcrb* and *Tcra* recombination. **(A-B)** Taqman qPCR quantification of rearrangements of each Vβ segment to DJβ1.1 (A) or DJβ2.1 (B) complexes conducted on genomic DNA from DN3 thymocytes. Signals from each assay were normalized to values from an assay for the invariant CD19 gene. Shown are average values +/- SEM from three independent DN3 thymocyte preparations, each from a different mouse. **(C-E)** Adaptive Immunoseq quantification of the percentage usage of each Vβ (C), Dβ (D), or Jβ (E) segment in *Tcrb* rearrangements of DN3 thymocytes. Shown are average values from two independent DN cell preparations. **(F)** qPCR quantification of specific Vα/Jα rearrangements performed on genomic DNA from DP thymocytes. Signals from each assay were normalized to values from an assay for the invariant β2m gene. Shown are the average values +/- SEM from three independent DP cell preparations. For all graphs in figure: #, not detected, n.s. p>0.05; * p<0.05, **p<0.01, *** p<0.001, **** p<0.0001

Figure 3. *PG* DN3 thymocytes exhibit reduced Syk protein expression but normal efficiency of pre-TCR signaled differentiation. **(A-B)** Representative flow cytometry histogram plots (A) and quantification (B) of Syk protein in DN3 thymocytes. (B) Shown are all data points and average values $\pm$ SEM from one of two independent experiments ($n = 3-4$ mice per genotype, unpaired t-test). **(C)** qRT-PCR quantification of Syk mRNA in DN3 cells. Relative mRNA levels were calculated using ddCt method with signals from each assay normalized to values from an assay for the *Hprt* gene. **(B-C)** Shown are all data points and average values $\pm$ SEM from two independent experiments ($n = 8-9$ mice per genotype, 2-way ANOVA with experiment an independent variable, stats indicate difference as a result of genotype). **(D)** Representative flow cytometry plots of DN3a, DN3b, and DN3c stages of thymocyte development showing the frequency of cells at each stage. **(E-F)** Quantification of percentages (E) or numbers (F) of total DN3 cells in each stage. Shown are all data points and average values $\pm$ SEM from three independent experiments each with three mice of each genotype ($n = 9$ mice per genotype, analyzed by multiple t tests using the Holm-Sidak method; n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.005$, **** $p < 0.00005$). **(G)** Quantification of percentages of DN3a cells that progress into DN3b cells. Shown are all data points and average values $\pm$ SEM from three independent experiments each with three mice of each genotype ($n = 9$ mice per genotype, analyzed by unpaired t-test with Welch's correction; ** $p < 0.005$). **(H)** Representative flow cytometry plots of intracellular TCR β protein expression in DN3a thymocytes. **(I)** Quantification of percentages of DN3a cells expressing intracellular TCR β protein. **(J)** Ratios comparing the percentages of DN3a cells expressing intracellular TCR β protein to the percentages of DN3a cells that transition into DN3b cells. **(H-J)** Shown are all data points and average values $\pm$ SEM from three independent experiments each with three mice of each genotype ($n = 9$ mice per genotype, analyzed by unpaired t-test with Welch's correction; n.s. $p > 0.05$, *** $p < 0.0005$).

989 **Figure 4.** Thymocytes of PD mice have reduced Zap70 protein levels. **(A-D)** Representative flow
 990 cytometry histogram plots (A, C) and quantification (B, D) of Zap70 (A, B) and TCR β (C, D) protein in
 991 Stages 1-4 of $\alpha\beta$ TCR selection. (C, D) Shown are all data points and average values $\pm$ SEM from one
 992 of three independent experiments (n = 4-5 mice per genotype, 2-way ANOVA with Sidak's multiple
 993 comparison post-test) **(E)** qRT-PCR quantification of Zap70 mRNA in DN3 cells or Stage 2 or 4 cells.
 994 Relative mRNA levels were calculated using ddCt method with signals from each assay normalized to
 995 values from an assay for the Hprt gene. Shown are all data points and average values $\pm$ SEM from one
 996 experiment (n = 4-5 mice per genotype, 2-way ANOVA with sidak multiple comparison post-test. **(F-G)**
 997 Representative flow cytometry histogram plots (F) and quantification (G) of Zap70 protein in naïve
 998 CD4 $^{+}$ or CD8 $^{+}$ T cells. Shown are all data points and average values $\pm$ SEM from two independent
 999 experiments (n = 3 mice per genotype, 2-way ANOVA with Sidak multiple comparison post-test). For
 1000 all graphs in figure: n.s. $p > 0.05$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

1001
 1002 **Figure 5.** PG mice exhibit impaired positive selection. **(A)** Gating strategy for measuring positive
 1003 selection of thymocytes showing Stage 1 and 2 gates. The TCR β^{+} gate was set using wild-type
 1004 thymocytes by gating on live, singlet, CD4 $^{+}$ CD8 $^{+}$ cells, preparing a histogram showing TCR β staining
 1005 on cells within this gate, and gating on the shoulder that represents cells expressing TCR β . **(B)**
 1006 Quantification of the ratio of the number of Stage 2 cells to the total number of Stage 1 and 2 cells.
 1007 Shown are all data points and global mean from three independent experiments (n = 10-11 mice per
 1008 genotype, 2-way ANOVA indicating effect of genotype). **(C)** Quantification of numbers of cells in each
 1009 stage of positive selection. Dots indicate individual mice. Data pooled from 2-independent experiments
 1010 (n = 7-8 mice per genotype, bars indicate mean $\pm$ SEM analyzed by Man-Whitney test). **(D)**
 1011 Quantification of CD69 expression alone and in comparison to TCR β expression on cells at each stage
 1012 of positive selection. Dots indicate individual mice. Data pooled from 2-independent experiments (n = 8

1013 mice per genotype, bars indicate mean $\pm$ SEM. Analysed by 2-way ANOVA and Sidak multiple
1014 comparison post-test) For all graphs in figure: n.s. $p>0.05$; * $p<0.05$, ** $p<0.01$, *** $p<0.001$, ****
1015 $p<0.0001$.

1016
1017 **Figure 6.** *PG* mice exhibit impaired superantigen-mediated negative selection. (A) Gating strategy form
1018 measuring negative selection of specific $V\beta^+$ thymocytes in F1 progeny of C57BL/6 background *WT* or
1019 *PG* mice crossed with BALB/c background *Rag1*^{-/-} mice. Pre-gated on live singlets. (B-F)
1020 Representative analyses and quantification of the indicated $V\beta^+$ cells in DP thymocytes that have not
1021 completed $\alpha\beta$ TCR selection or post-selection SP thymocytes. (G) Data from B-F displayed as LOG₂
1022 ratio of respective percent $V\beta^+$ cells in SP versus DP gates. Shown are all data points and average $\pm$
1023 SEM from two independent experiments (n = 9 mice per genotype 2-way ANOVA with Sidak multiple
1024 comparison post-test). For any significantly different combinations, fold change is indicated. For all
1025 graphs in figure: * $p<0.05$, **** $p<0.0001$.

1026
1027 **Figure 7.** *PG* mice have impaired agonist selection of iNKT cells. (A-B) Representative flow cytometry
1028 plots (A) and quantification (B) of PBS57/CD1d tetramer binding iNKT cells in thymuses or spleens.
1029 Gated on live, singlet lymphocytes. (B) Data from three independent experiments with 1-4 mice per
1030 group. Mann-Whitney test: *, $p=0.01$ (thymus); ns, not significant (spleen). (C) Taqman PCR
1031 quantification of $V\alpha 14$ - $J\alpha 18$ rearrangements in total thymocytes. The data were generated from three
1032 technical replicates of three independent biological samples from mice of each genotype. Data are
1033 represented as Mean $\pm$ SEM analyzed by an unpaired parametric t test with Welchs correction.
1034 $p=0.0094$. (E-F) Representative flow cytometry plots and gating strategy (E) or quantification (F) of
1035 Stage 0 ($CD24^{hi}$ $CD44$ - $NK1.1$ -), Stage 1 ($CD44$ - $NK1.1$ -), Stage 2 ($CD44$ + $NK1.1$ -), and Stage 3
1036 ($CD44$ + $NK1.1$ +) iNKT cells. (G) Data from three independent experiments with 1-4 mice per group.
1037 Two way ANOVA repeated measures test for genotype x stage interaction: **, $p<0.0011$.

1038

1039 **Figure 8.** Mature $\alpha\beta$ T cells of PG mice exhibit normal immune response. **(A)** Representative plot of
 1040 cytokine expression on CD4⁺CD44^{hi} $\alpha\beta$ T cells following *ex vivo* stimulation of splenocytes with PMA
 1041 and ionomycin. **(B-F)** Representative analyses of naïve splenic CD4⁺ or CD8⁺ T cells after *ex vivo*
 1042 simulation with anti-CD3/anti-CD28 antibodies and additional antibodies to skew towards indicated
 1043 lineages. Shown is CFSE incorporation (B) or expression of relevant transcription factors (C, D) or
 1044 cytokines (E, F). **(G, H)** Representative flow data showing (G) frequencies of splenic $\alpha\beta$ T cells that
 1045 stain with the gp33:H-2Db tetramer or (H) express indicated cytokines after LCMV infection of mice.
 1046 Data are representative of two independent experiments with at least four mice of each genotype. **(I, J)**
 1047 Representative plots (J) and (I) quantification of antigen-specific $\alpha\beta$ T cells able to respond to indicated
 1048 viral peptides within a peptide re-stimulation assay. (A-F) Data are representative of two independent
 1049 experiments. (G-J) Data are combined from four independent experiments (two D7 and two D35
 1050 timepoints) with 8-11 mice of each genotype per time-point.

1051

1052

1053 **Figure 9.** Mature $\alpha\beta$ T cells of PG mice possess higher than normal autoimmune hazard. **(A)** Graph of
 1054 body weights of *Rag1*^{-/-} hosts over time after transfer of WT or PG CD4⁺ T cells or no cells (NT). Data
 1055 combined from three independent experiments with a total of 13-17 mice of each genotype. Lines
 1056 indicate fitted segmented-linear regression lines +/- 95% confidence interval. Day 15 was fixed as the
 1057 inflection point. Difference between gradients of second slope was determined by linear regression
 1058 analysis. **(B)** Numbers of CD4⁺ T cells in spleen eight weeks post-transfer. **(C)** Total number of IFN γ ,
 1059 TNF α , and/or GM-CSF producing cells per spleen. Shown are all data points from three independent
 1060 experiments with 12-14 mice per group. Bar indicates global mean. Pre-gating on live singlets, CD90.2⁺,
 1061 CD4⁺, and CD44⁺. **(D)** Colon length of mice sacrificed at eight weeks post-infection. **(E)** Representative
 1062 hemotoxylin and eosin stained colon sections from respective mice at eight weeks post-transfer. 10x

1063 magnification. **(F-G)** Total colitis score as assessed by blinded pathologist. Scored as per previously
1064 described guidelines for the five phenotypes (48): **(F)** Colon inflammation score (contribution of scores
1065 for crypt abscesses and inflammatory cell infiltrate). **(G)** Colon tissue damage score (contribution of
1066 scores for crypt architecture, tissue damage, goblet cell loss). **(H)** Non-colitis pathologies described for
1067 mice transferred with *WT* versus *PG* cells. **(I)** Hemotoxylin and eosin stained skin section from *PG*-
1068 transferred mouse with severe dermatitis at eight weeks post-transfer. 10x magnification. **(J)** Spleen
1069 weight at eight weeks post-transfer normalized to total mouse bodyweight prior to transfer. **(K)** Survival
1070 curves of mice transferred with *WT* or *PG* cells. Data combined from three independent experiments.
1071 Statistical analysis = Log-rank Mantel-Cox test. Overall: 0/13 NT mice died, 1/15 *WT*-transferred mice
1072 died, and 4/17 *PG*-transferred mice died. (B, C, F, G, and J) Shown are all data points from three
1073 independent experiments, 12-14 mice per group. Experiment indicated by shape of data point. Bar
1074 indicates global mean. Analyzed by 2-way ANOVA with Tukey HSD post-test. n.s. $p>0.05$; * $p<0.05$,
1075 ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$ **(L)** Curves showing diabetes free survival of female NOD *WT*
1076 (n=16) and *PG* (n=14) mice. Mice were generated by mating together NOD *Rag1*^{WT/PG} mice and
1077 monitored for their first 30 weeks of life. The percent diabetes survival curves include mice that have
1078 passed 20 weeks of age or gotten diabetes before 20 weeks of age. Analyzed by Gehan-Breslow-
1079 Wilcoxon test ($p=0.049$) and log-rank Mantel Cox test ($p=0.1151$).

1080
1081
1082

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

Figure 9