Inherent reactivity of unselected TCR repertoires to peptide-MHC molecules

S. Harsha Krovi[a], John W. Kappler[a,b], Philippa Marrack[a,b,1], and Laurent Gapin[a,1]

[a]Department of Immunology and Microbiology, University of Colorado School of Medicine, Aurora, CO 80045; and [b]Department of Biomedical Research, National Jewish Health, Denver, CO 80206

Contributed by Philippa Marrack, September 2, 2019 (sent for review June 3, 2019; reviewed by Remy Bosselut and K. Christopher Garcia)

The repertoire of αβ T cell antigen receptors (TCRs) on mature T cells is selected in the thymus where it is rendered both self-tolerant and restricted to the recognition of major histocompatibility complex molecules presenting peptide antigens (pMHC). It remains unclear whether germline TCR sequences exhibit an inherent bias to interact with pMHC prior to selection. Here, we isolated TCR libraries from unselected thymocytes and upon reexpression of these random TCR repertoires in recipient T cell hybridomas, interrogated their reactivities to antigen-presenting cell lines. While these random TCR combinations could potentially have reacted with any surface molecule on the cell lines, the hybridomas were stimulated most frequently by pMHC ligands. The nature and CDR3 loop composition of the TCRβ chain played a dominant role in determining pMHC-reactivity. Replacing the germline regions of mouse TCRβ chains with those of other jawed vertebrates preserved reactivity to mouse pMHC. Finally, introducing the CD4 coreceptor into the hybridomas increased the proportion of cells that could respond to pMHC ligands. Thus, αβ TCRs display an intrinsic and evolutionary conserved bias for pMHC molecules in the absence of any selective pressure, which is further strengthened in the presence of coreceptors.

T cell antigen receptor | major histocompatibility complex | T cell selection

Antigen specificity is a hallmark of the adaptive immune system. This specificity is driven by the diverse repertoire of antigen receptors expressed by lymphocytes. During lymphocyte development, antigen receptor loci undergo a series of complex and coordinated genetic rearrangements that give rise to receptors not encoded within the germline (1). Overall, the diversity of these novel receptors has been estimated to exceed 10^{15} unique sequences, which ensures that the immune system can respond to the similarly large world of foreign antigens (2).

Although the antigen receptor loci of B and T lymphocytes are located on different chromosomes, the genes belonging to each locus are organized similarly (1). Additionally, developing B and T lymphocytes employ the same recombination machinery in order to generate their respective antigen receptors, which are likewise structurally similar (3). Despite this similarity, the manner in which these 2 types of antigen receptors recognize their antigens is vastly different. B cell antigen receptors (BCRs) recognize 3D conformational epitopes, while the vast majority of αβ T cell antigen receptors (TCRs) expressed by mature T cells are specific for linear peptide fragments presented by the major histocompatibility complex (MHC) proteins. This phenomenon displayed by αβ TCRs has been dubbed MHC restriction (4). Although MHC restriction is foundational to αβ T cell biology, the cause of this phenomenon remains unclear (5).

Two nonmutually exclusive hypotheses have been proposed to explain MHC restriction. The selection hypothesis posits that TCRs, like BCRs, have the ability to react with any type of antigen, but that thymic selection allows only T lymphocytes with TCRs that can specifically recognize peptides bound to MHC molecules to mature (6). During thymocyte development, a completely rearranged αβ TCR is first expressed on cells at the $CD4^+$ $CD8^+$ (DP) stage. At this stage the TCRs are interrogated for their ability to interact with pMHC (MHC molecules presenting peptide antigens) complexes (7, 8). Expression of the CD4 and CD8 coreceptors serves 2 purposes. First, CD4 and CD8 can interact extracellularly with conserved regions of MHC-II and MHC-I molecules, respectively, thereby assisting TCRs in finding their cognate ligands (9, 10). Second, both coreceptors intracellularly interact with and sequester the kinase Lck, which is responsible for initiating the TCR signaling cascade (11). Thus, according to the selection hypothesis, although the preselection αβ TCR repertoire contains non-MHC specificities, the coreceptors do not allow cells with such specificities to mature because the coreceptors cannot engage these ligands and thus bring Lck into close enough proximity for it to contribute to the reaction. Without any coreceptors, Lck would be free to initiate signaling for TCRs of all specificities (12). Support for this hypothesis is provided by data from $Quad^{KO}$ mice, which lack both coreceptors and MHC molecules and possess postselected T cells with TCRs that are not MHC-restricted but instead recognize native epitopes similar to how BCRs recognize their targets (13, 14).

The germline hypothesis instead proposes that the preselection αβ TCR repertoire has an evolutionary bias for interacting with pMHC complexes (4). TCRs mainly interact with their targets using 3 loops, the complementarity determining regions (CDRs) (3). The CDR1 and CDR2 loops are encoded within the germline portion of the variable (V) genes, while the CDR3 loops are the

Significance

Central to adaptive immunity is the interaction between the αβ T cell antigen receptor (TCR) and peptides presented by the major histocompatibility complex (MHC) molecules. It is fundamental to understand any potential generalities regarding a TCR-pMHC recognition event. Two theories have been proposed to explain how TCR recognition of pMHC is specified: The germline-encoded theory, which proposes an evolutionary "hardwiring" of the TCR for recognition of MHC molecules through germline-encoded motifs, and the selection theory of TCR recognition, which suggests that TCR selection during development imposes the constraint of MHC recognition. Here, we show that TCRs have intrinsic specificity for MHC molecules, inherent to TCRαβ sequences, and enhanced but not determined by coreceptor expression.

Author contributions: S.H.K., P.M., and L.G. designed research; S.H.K. performed research; S.H.K. and J.W.K. contributed new reagents/analytic tools; S.H.K. and L.G. analyzed data; and S.H.K., J.W.K., P.M., and L.G. wrote the paper.

Reviewers: R.B., NIH; and K.C.G., Stanford University.

The authors declare no competing interest.

See Commentary on page 21969.

[1]To whom correspondence may be addressed. Email: marrackp@njhealth.org or Laurent.gapin@CUAnschutz.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1909504116/-/DCSupplemental.

First published September 30, 2019.

hypervariable portions of the antigen receptor that are created by DNA recombination. In the myriad TCR/pMHC complex structures that have now been solved, TCRs dock onto MHC molecules in a fairly conserved manner, with the germline-encoded CDR1/2 loops interacting with the α-helices of the MHC grooves, while the variable CDR3 loops interact primarily with the bound peptide, the most variable component of the ligand (3). Thus, TCR and MHC molecules might have coevolved over time for interacting with each other in a fairly conserved manner, even prior to any formal selective event.

Studies from numerous laboratories have provided support for this hypothesis over the years. First, coexpression of TCRα and TCRβ chains originating from different MHC-restricted TCRs might have been expected to abrogate MHC reactivity altogether. However, such novel chimeric TCRs continued to interact with MHC, suggesting that MHC restriction might have been a feature independent of peptide specificity (15, 16). In addition, 2 independent studies reported that the preselection repertoire contains a larger-than-expected proportion of MHC-reactive TCRs (17, 18). Furthermore, many (~7%) postselected TCRs display alloreactivity to nonself MHC molecules, indicating a common mode of MHC recognition independent of the actual selecting MHC molecule (19, 20). More recently, we identified residues within the CDR2β loop that were both evolutionarily conserved and critical for efficient selection by MHC molecules (21, 22). Conversely, mutating conserved residues on the α-helices of MHC molecules altered the peripheral TCR repertoire, as only the TCRs capable of compensating for the absence of these interactions were selected (23). Other studies have also corroborated the predisposition of TCR molecules to interact with pMHC complexes (24–27). Finally, genetic quantitative trait loci analyses in humans demonstrated that polymorphisms unique to a given MHC haplotype are linked to TCR V-gene usage in mature T cells across numerous individuals (28). Importantly, these polymorphisms code for residues located near where TCRs interact with the MHC molecules. Thus, while different MHC haplotypes may uniquely bias TCR V-gene usage, this bias is likely maintained across individuals through conserved modes of interaction between certain TCR V-genes and MHC haplotypes.

The selection and germline hypotheses are neither mutually exclusive nor absolute. The preselection TCR repertoire might indeed contain some TCRs that can recognize native antigens as well as TCRs that are predisposed for MHC recognition. The fundamental issue that remains is whether the frequency of MHC-biased TCRs in the preselection repertoire is high enough to warrant the conclusion that evolution has played a role in their appearance. To answer this question, we isolated TCR repertoires from thymocytes of TCRα$^{-/-}$ mice (29). Because thymocytes in such mice do not express complete TCRαβ heterodimers, neither the TCRα nor the TCRβ chains could have been subjected to MHC selection, even though TCR rearrangements occur normally in these mice, allowing for the isolation and cloning of TCRα libraries. Such TCRα libraries were expressed and paired with various TCRβ chains into a TCR$^-$ hybridoma engineered to report TCR stimulation. Such a strategy ensured that the TCR repertoires that were recreated in vitro had never been selected for any particular reactivity and that the responses of the TCR-expressing hybridomas could be monitored at the single-cell level. In this way, we tested the reactivity of >1 × 10^5 different unselected TCRs. Different combinatorial pairing of various TCRβ and TCRα chains demonstrated different levels of reactivity toward diverse antigen-presenting cells (APCs). Nevertheless, when responses were detected, they were largely directed at pMHC complexes with little evidence for non-MHC reactivity. Furthermore, the nature of the TCRβ chain played a key role in determining whether the αβ TCR pair was MHC-reactive. Expression of a high-affinity variant of the human CD4 coreceptor increased the proportion of TCRs that responded to pMHC ligands (30, 31). Altogether, our results demonstrate that the intrinsic specificity of TCRs for MHC molecules is inherent to TCRαβ sequences that are enhanced but not determined by coreceptor expression.

Results

Diverse TCRα Sequences Are Present within DP Cells from TCRα$^{-/-}$ Mice. Since the theoretical diversity of the TCR repertoire is estimated to be ~10^{15} unique sequences, it was not feasible to test the reactivities of all of those receptors. Instead, in order to restrict the total number of αβ TCR pairs tested, we elected to pair libraries of TCRα chains with fixed TCRβ chains. To ensure that we would be interrogating the reactivity of TCRs that have never been selected on pMHC complexes or any other surface ligands, we isolated and cloned TCRα rearrangements from DP thymocytes of TCRα$^{-/-}$ mice. The TCRα rearrangements are not impacted in these mice but the translated proteins are not functional and degraded (29). Consequently, these TCR rearrangements truly represent unselected chains and provide a unique opportunity to study the extent of TCR reactivities in the absence of any selection events.

We extracted RNA and amplified TCRα rearrangements to generate 5 different TCRα libraries from TCRα$^{-/-}$ DP cells, with each library corresponding to a different TCRα variable (TRAV) gene member (TRAV5, TRAV6, TRAV7, TRAV9, and TRAV14). Each of these libraries was then independently cloned into retroviral vectors, placing the rearrangements in the context of a functional TCRα constant region to ensure that if the CDR3 rearrangements had been productive in vivo they could make proteins upon expression in hybridomas in vitro. Deep sequencing of the libraries showed ~5 × 10^4 total unique TCRα sequences (Fig. 1), with a large diversity of TCRα joining (TRAJ) gene usage and CDR3 length distributions (*SI Appendix*, Fig. S1). As expected, approximately two-thirds of these sequences were out of frame and could not produce functional proteins. Upon removing these nonproductive rearrangements, we were left with a total of ~1.6 × 10^4 unique sequences whose reactivities could be examined (Fig. 1).

Postselected TCRβ Chains Paired to Unselected TCRα Libraries Display Strong MHC Reactivity. Because we aimed to evaluate the reactivity of thousands of different TCRs cloned from unselected thymocytes, we had to establish a robust and sensitive in vitro system allowing for the detection of TCR-mediated responses at the single-cell level. We used a TCR$^-$ T cell hybridoma line expressing an NFAT-GFP reporter construct (32, 33) into which we could retrovirally introduce TCR chains. Proof-of-principle experiments demonstrated that this system induced GFP upon stimulation in antigen-specific (Fig. 2), sensitive (*SI Appendix*, Fig. S2), and cell-intrinsic (*SI Appendix*, Fig. S3) manners. We next proceeded to test if pairing the unselected TCRα libraries with fixed postselection TCRβ chains in the hybridomas would produce any reactivities when cultured with APCs. To do so, we used TCRβ chains from 3 different TCRs: 75-55, YAe5-62.8, and DO-11.10. The first 2 displayed strong alloreactivity/peptide promiscuity while the third had a defined pMHC specificity (34, 35). However, their known reactivities were observed only when these TCRβ chains were paired with their original TCRα chains. Thus, by pairing the TCRβ with the random TCRα chains in the libraries, we generated a multitude of potentially novel specificities.

We paired each of the 3 TCRβ chains with the 5 different TCRα libraries and sorted those cells that expressed surface TCRs, thereby eliminating cells that had been transduced with genes expressing nonfunctional TCRα chains, and normalized for equivalent levels of TCR expression between libraries (*SI Appendix*, Fig. S4). The hybridomas expressing the different TCRα libraries were then independently cocultured with the H-2^b–expressing APCs used previously. Under these conditions, a

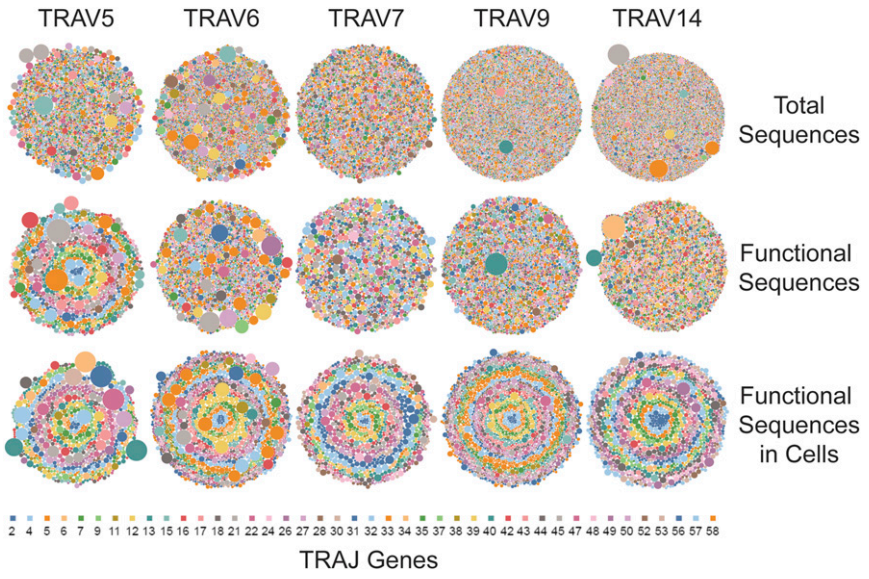


Fig. 1. Cloned unselected TRAV libraries are diverse. Plots displaying the diversity for each of the TRAV libraries generated. In every plot, each circle represents 1 specific clone (as determined by TRAV/TRAJ usage and a unique CDR3 sequence). The size of the circle represents the frequency with which the clone is represented in the library. Finally, the color of the circle corresponds to the TRAJ gene utilized by the clone. (*Top*) Plots for all of the sequences in each of the 5 TRAV libraries. (*Middle*) Plots for only the functional sequences (sequences that are in-frame and lack stop codons) in each of the 5 TRAV libraries. (*Bottom*) Plots for the functional sequences recovered from transduced hybridomas for each of the 5 TRAV libraries.

large proportion of the cells expressing the 75-55β or the YAe5-62.8β chains paired to the TCRα libraries up-regulated GFP (Fig. 3). Interestingly, addition of MHC blocking antibodies (but not isotype control mAbs) significantly inhibited the observed response of each of the libraries (over 90% in some cases), hinting that these TCRβ chains conferred MHC reactivity even when not paired to their original TCRα chains. Hybridomas expressing TCRα libraries paired with the DO-11.10β chain did not respond to the APCs so frequently. Only the cells expressing the TRAV5 library paired with the DO-11.10β chain responded to H-2^b APCs, with almost all of the response inhibited by MHC-blocking antibodies. Interestingly, the original DO-11.10α chain also uses a TRAV5 rearrangement and this pairing may yield greater pMHC-specific reactivity.

Although transductions of the hybridomas with the TCRα libraries were performed at low multiplicity of infection, the possibility exists that multiple TCRα chains might end up expressed in any given cell, potentially affecting the response of the hybridoma to a given antigen. To find out how this potentially confounding problem might affect our experiments, we took advantage of the fact that the DO-11.10β chain engenders reactivity to the OVA$_{323–339}$ peptide presented by the MHC-II molecule I-A^b when paired to the DO-11.10α chain, while the same DO-11.10β chain reacts with the lipid α-galactosylceramide (αGC) presented by the MHC-I–like molecule CD1d when paired to the iNKTα chain (22). Hybridomas expressing the DO-11.10β chain were infected with different ratios of retroviruses coding for either the DO-11.10α or the iNKTα chains (*SI Appendix*, Fig. S5*A*). In this way, we aimed to generate populations of cells with identical surface TCR levels but with different ratios of the iNKT TCR or DO-11.10 TCR per single cell. Staining the hybridomas generated under these conditions with CD1d tetramers (*SI Appendix*, Fig. S5*B*) revealed that hybridomas expressing 100% of the DO11.10 TCR did not bind the tetramer, as expected. However, when a fraction of the TCRα chains expressed in the hybridomas are derived from the iNKT TCRα chain, PBS57-CD1d tetramer binding is observed and increases as the proportion of this chain rises within the cell (*SI Appendix*, Fig. S5*B*). When each of these hybridomas were cultured with I-A^{b+} CHB-2.4.4 cells and increasing concentration of the OVA$_{323–339}$ peptide, we observed that reduction of the functional avidity of the cells (by reducing the proportion of specific TCRs) strongly influenced the ability of the cells to respond to their cognate antigen (*SI Appendix*, Fig. S5*C*). In essence, this mimics the loss of avidity that would be achieved if multiple TCRα chains were incorporated into the same cell when infecting with the TCRα libraries and indicates that any response measured in our system actually underestimates the proportion of cells that can potentially respond to a given stimulatory condition.

Altogether, our data show that the random pairing of TCRα libraries with different TCRβ chains derived from postselected TCRs in hybridomas produced reactivities to APCs. These TCRs

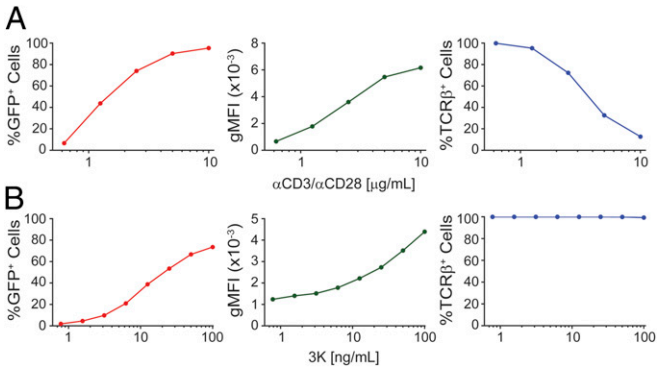


Fig. 2. Hybridomas with the NFAT-GFP reporter respond in a dose-dependent manner upon stimulation. Hybridomas expressing the B3K506 TCR (specific for the 3K peptide presented by the I-A^b MHC-II molecule) were stimulated at different concentrations of αCD3/αCD28 (*A*) or by the H-2^b haplotype bearing CHB-2.4.4 cells with various 3K peptide concentrations (*B*) for 18 to 24 h. Subsequently, they were stained with an antibody targeting TCRβ and the proportion of GFP$^+$ cells was assessed. Summary of the results from the antibody stimulation or from the antigen presentation assay with plots depicting the proportion of GFP$^+$ cells, the geometric mean fluorescent intensity (gMFI) of the GFP$^+$ cells, and the proportion of TCRβ$^+$ cells as a function of stimulation dose are shown.

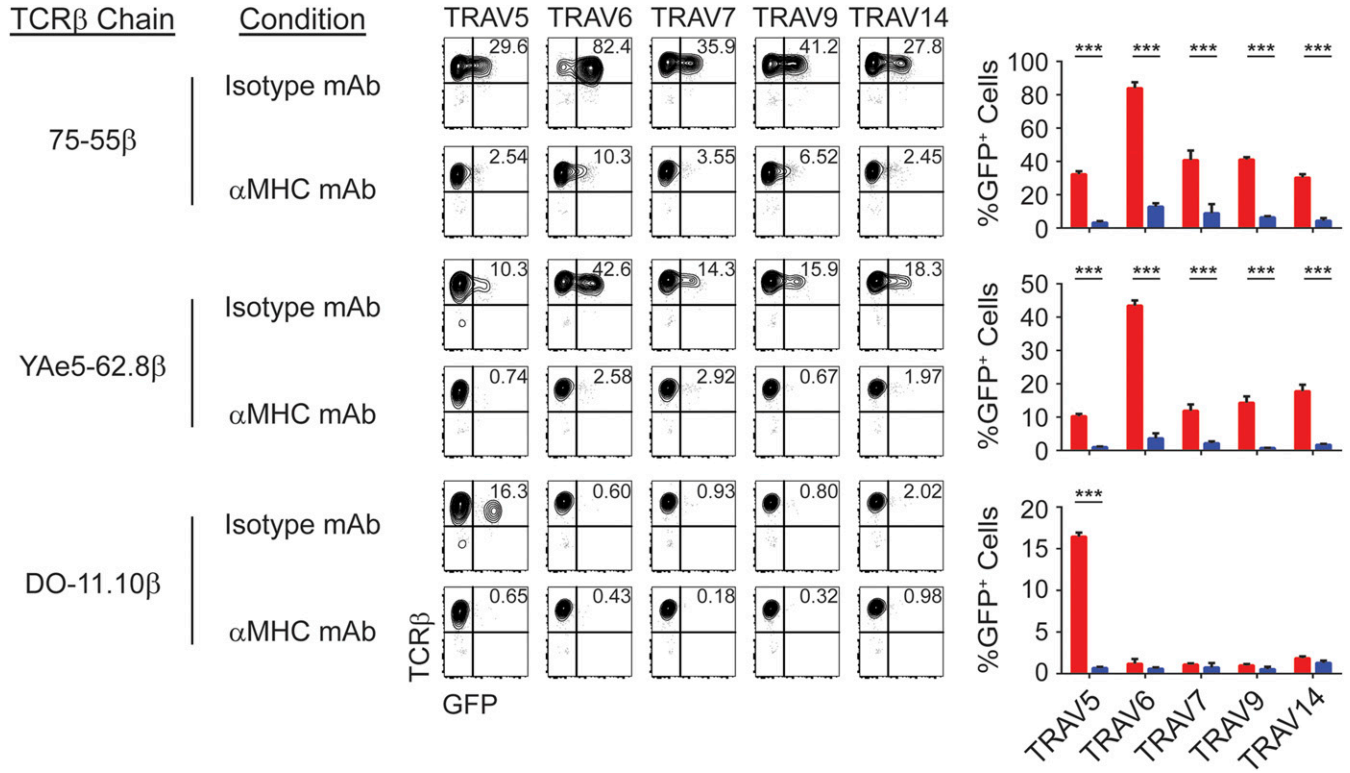
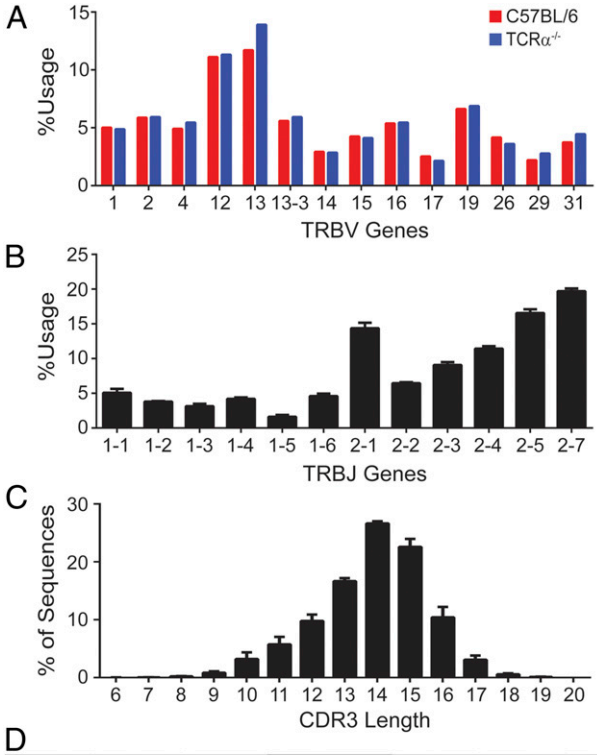


Fig. 3. Unselected TRAV libraries paired to postselected TCRβ chains are MHC-reactive. (*Left*) TRAV libraries were paired to the 75-55β, the YAe5-62.8β, and the DO-11.10β chains and stimulated by H-2^b–bearing CHB-2.4.4 cells in the presence of isotype or MHC-blocking antibodies. After 18 to 24 h of stimulation, the cells were stained with an antibody targeting TCRβ and the proportion of GFP$^+$ cells was assessed in each stimulation condition. Results are representative of 2 independent experiments. (*Right*) The results are summarized for each stimulation condition with red representing the isotype antibody condition and blue representing the MHC-blocking antibody condition (***$P < 0.001$ by multiple t tests; error bars are mean ± SD).

could have potentially reacted to any surface molecules expressed by these APCs. However, the reactivity of the hybridomas was completely abrogated by the addition of MHC blocking antibodies, suggesting that although new combinations of TCRα and TCRβ chains were generated, the reactivities to pMHC complexes were conserved. These reactivities were easily observed when the TCRβ chains were obtained from polyspecific TCRs but were less frequent with TCRβ chains derived from TCRs with more restricted specificities. Because the 3 TCRβ chains tested in these experiments all use the TRBV13-2 protein and, consequently, share their CDR1 and CDR2 loops but differ in the composition of their CDR3 loops, the nature of the CDR3β loop appears to play a dominant and crucial role in determining pMHC reactivity of the complete αβ TCR. The TCRα chain, on the other hand, appears to work within the reactivity framework established by the TCRβ chain by fine-tuning the overall interaction based on the composition of its CDRs.

Highly Represented TCRβ Sequences Are Primarily Germline Rearrangements with Short CDR3 Length. Since the 75-55β, Ae5-62.8β, and DO-11.10β chains are all derived from TCRs that have been previously positively selected on pMHC complexes in vivo, it was possible that they were predisposed to pMHC recognition. Therefore, we set out to identify unselected TCRβ chains that were commonly found in the DP repertoire and shared between individual mice. We hypothesized that these TCRβ chains, upon pairing with TCRα chains in vivo, would audition for selection more often than other sequences. Intracellular Vβ staining of preselection DP thymocytes from C57BL/6 or TCRα$^{-/-}$ mice revealed that TRBV12 and TRBV13 chains were the most prevalent proteins expressed (Fig. 4*A*). Since the TRBV12 gene products are known to interact with mouse mammary tumor virus-derived superantigens (36), we concentrated on TRBV13 rearrangements. We deep-sequenced TRBV13 gene rearrangements from DP cells of 3 separate TCRα$^{-/-}$ mice to identify the sequences that were the most prevalent and shared across all 3 samples. We observed substantial diversity in the sequences, as might be expected in the absence of any selection (*SI Appendix*, Fig. S6). The overall TRBJ gene usage and CDR3 lengths were fairly consistent for each of the samples (Fig. 4 *B* and *C*). TRBJ2-7 was the most common gene rearranged to TRBV13-2, and the CDR3 lengths appeared to be near-normally distributed with a peak centered at a length of 14 amino acids. The most represented sequences within the samples (named 13-2_1β, 13-2_3β, and 13-2_4β) all involved rearrangements with TRBJ2-7, and each one of them had a length shorter than the observed median CDR3 length (Fig. 4*D*). These sequences are also observed at high frequencies in unselected C57BL/6 DP thymocytes (37). Furthermore, when the nucleotide rearrangements were analyzed, 2 of the 3 sequences primarily used a germline-exclusive rearrangement with no nontemplated nucleotides. Thus, frequently occurring rearrangements were public, used TRBJ2-7, had short CDR3 lengths, and did not include nontemplated nucleotide additions at the junctions. We proceeded to test the reactivities of these TCRβ chains by pairing them to the TCRα libraries in our engineered hybridomas.

Random TCRβ Chains Also Possess Intrinsic MHC Reactivity That Is Revealed in the Presence of Coreceptors. We independently expressed the 3 TCRβ chains identified above, paired them to each of the 5 different TCRα libraries in the hybridomas, and checked the cells for similar TCR expression levels. Coculturing these hybridomas with CHB-2.4.4 APCs bearing the H-2^b (Fig. 5*A*) or the BALB/c



D

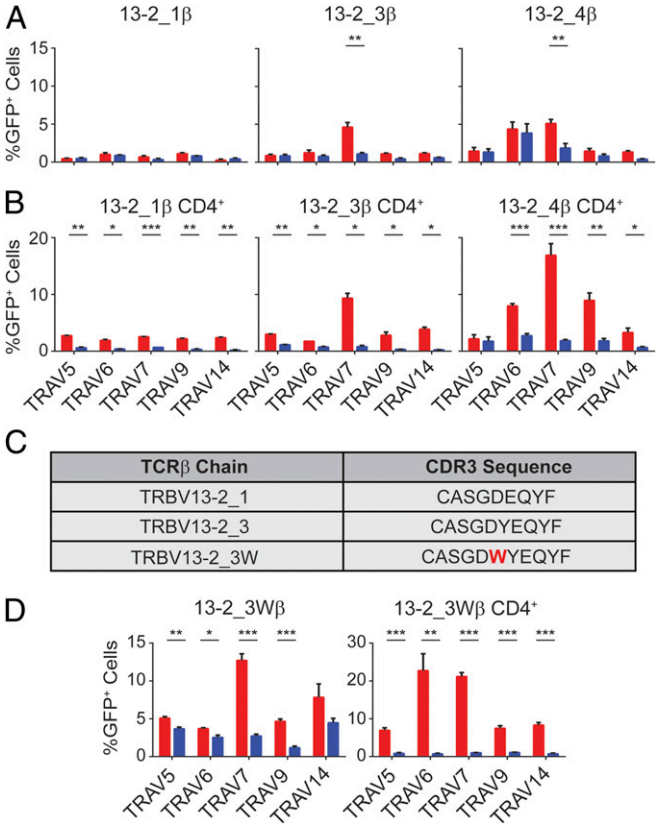
Sequence ID	V Gene	J Gene	CDR3	Mouse 1 %	Mouse 2 %	Mouse 3 %
13-2_1	TRBV13-2	TRBJ2-7	CASGDEQYF	0.642	1.509	0.760
13-2_3	TRBV13-2	TRBJ2-7	CASGDYEQYF	0.621	1.509	0.716
13-2_4	TRBV13-2	TRBJ2-7	CASGDAGEQYF	0.627	1.280	0.560

Fig. 4. TCRβ repertoire of TCRα$^{-/-}$ DP thymocytes. (*A*) Intracellular Vβ (TRBV) staining of immature, nonsignaling DP thymocytes from C57BL/6 and TCRα$^{-/-}$ mice. (*B*) TRBJ gene usage of the sequenced TRBV13-2 rearrangements from DP thymocytes of 3 different TCRα$^{-/-}$ mice (error bars are mean ± SD). (*C*) CDR3 length distribution of the rearrangements between TRBV13-2 and TRBJ2-7 summarized for all 3 mice (error bars are mean ± SD). (*D*) Table depicting the 3 most frequently occurring and shared sequences. Percentages for each sequence represent the proportion of the given sequence within all TRBV13-2/TRBJ2-7 rearrangements for any given mouse.

MutuDC APCs bearing H-2^{d} (*SI Appendix*, Fig. S7*A*) revealed minimal reactivity. Only a few of the libraries had observable reactivity (e.g., the TRAV7 library paired to the 13-2_3β and 13-2_4β chains). Therefore, unlike those displayed by the 75-55β and YAe5-62.8β chains, the unselected random TCRβ chains did not display extensive reactivities toward any potential ligand provided by the APCs, including pMHC.

We next tested the idea that low TCR affinities for their pMHC ligands on the APCs might cause the poor observed responses in these transduced hybridomas. To increase the sensitivity to low-affinity ligands, we cotransduced the cells with the high-affinity mutant human CD4 coreceptor (30, 31). When the experiments shown in Fig. 5*A* and *SI Appendix*, Fig. S7*A* were repeated with the mutant CD4-expressing libraries, they all showed increased response to the APCs, which was inhibitable by MHC-blocking mAbs (Fig. 5*B* and *SI Appendix*, Fig. S7*B*). These results demonstrate that prior to coreceptor expression, the TCR libraries do not contain random specificities to non-MHC ligands. In contrast, in the presence of this CD4 coreceptor, the same TCR libraries react to pMHC complexes expressed by the APCs, suggesting that these random TCRs are indeed pMHC-specific but that observable responses by the hybridomas require the expression of the CD4 coreceptor, which presumably strengthens the interaction between TCR and MHC-II molecules.

Self-reactive TCRs tend to possess a hydrophobic doublet motif in their CDR3β sequences (38). Unfortunately, none of our random TCRβ chains possessed such a motif. However, we did notice that by introducing a tryptophan residue into the 13-2_3β chain sequence, we could generate a TCRβ chain with the described motif (Fig. 5*C*), allowing us to test whether the presence of more hydrophobic residues within the CDR3 loop might affect the reactivity of the hybridomas. Indeed, hybridomas expressing each of the TCRα libraries paired to this engineered TCRβ chain reacted strongly to APCs in a MHC-dependent manner (Fig. 5*D* and *SI Appendix*, Fig. S7*C*). The pMHC-specific



C

TCRβ Chain	CDR3 Sequence
TRBV13-2_1	CASGDEQYF
TRBV13-2_3	CASGDYEQYF
TRBV13-2_3W	CASGDWYEQYF

Fig. 5. The CD4 coreceptors strengthen intrinsic TCR specificity for MHC. (*A*) Bar graphs depicting the proportion of GFP^{+} cells after 18 to 24 h of stimulation by H-2^{b}–bearing CHB-2.4.4 cells for TRAV libraries expressed in hybridomas bearing the labeled TCRβ chain. Red represents the isotype antibody condition and blue represents the MHC-blocking antibody condition. Each experiment was repeated 2 independent times (**$P < 0.01$ by multiple *t* tests; error bars are mean ± SD). (*B*) Bar graphs depicting the proportion of GFP^{+} cells after 18 to 24 h of stimulation by H-2^{b} bearing CHB-2.4.4 cells for TRAV libraries expressed in hybridomas bearing the labeled TCRβ chain and expressing the CD4 coreceptor. Red represents the isotype antibody condition and blue represents the MHC-blocking antibody condition. Each experiment was repeated 2 independent times (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$ by multiple *t* tests; error bars are mean ± SD). (*C*) Table displaying the location of the introduced tryptophan residue (W, in red) in the engineered TCRβ chain and how this CDR3 sequence compares with those of 2 of the frequently observed TCRβ chains. (*D*) Bar graphs depicting the proportion of GFP^{+} cells after 18 to 24 h of stimulation by H-2^{b}–bearing CHB-2.4.4 cells for TRAV libraries expressed in hybridomas bearing the labeled TCRβ chain and either expressing the CD4 coreceptor or not. Red represents the isotype antibody condition and blue represents the MHC-blocking antibody condition. Each experiment was repeated 2 independent times (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$ by multiple *t* tests; error bars are mean ± SD).

response was even greater when the mutant human CD4 was added to the hybridomas, recapitulating what was detected with the other random TCRβ chains. Moreover, there was an incremental increase in pMHC reactivity as the number of hydrophobic residues increased within the CDR3 (compare the 13-2_1β, 13-2_3β, and 13-2_3Wβ library responses), suggesting that the presence of hydrophobic residues at the tip of the CDR3β loop does indeed favor pMHC reactivity.

Evolutionarily Conserved Reactivity between TCRs and MHC Molecules. Previous work by our group demonstrated that chimeric TCRβ chains comprising germline-encoded regions derived from distantly related jawed vertebrates continued to interact with mouse MHC molecules (22). These interactions were strongly dependent on specific CDR2 residues that are evolutionarily conserved, despite only ∼30% overall cross-species sequence similarity. We asked whether similar reactivities could be observed when such TCRβ chains were paired to unselected TCRα libraries. To do so, we generated TCRβ chains where the mouse TRBV13-2 germline-encoded regions were replaced with those of frog (from the TRBV2 gene), shark (from the Vβ3-Hf73 gene), or trout (from the TRBV8S2 gene). We coupled these chains with the 13-2_3Wβ and the 13-2_4β CDR3 sequences since they provided observable responses by the hybridomas when paired to the TCRα libraries (Fig. 5). Additionally, because CD4 boosted the responses dramatically, we assessed how the hybridoma expressing any of these TCR combinations responded in the presence of this coreceptor.

We independently expressed each of these chains (frog, shark, or trout) with either the 13-2_3Wβ or the 13-2_4β CDR3 for a total of 6 different TCRβ chains. Each of these was then paired to the different TCRα libraries and sorted for comparable TCR levels, demonstrating that each combination can pair efficiently with mouse TCRα chains. When cocultured with H-2^b–bearing APCs, each of the library-expressing hybridomas responded in a characteristic manner. Cells expressing the combination of TRAV6 and TRAV7 libraries paired with the 13-2_3Wβ chains from each species responded stronger than cells containing the same TCRβ chains but paired to the other TCRα libraries (Fig. 6). However, the responses generated by these chimeric TCRβ chains were substantially improved compared to their mouse counterparts, with over 50% of the cells stimulated by the APCs in some cases when the frog TCRβ chain was present. Indeed, all of the expressed libraries, irrespective of the species of origin of the TCRβ chain, conferred greater stimulation to the cells than what was observed in the cells bearing the TCRα libraries paired to the mouse TCRβ chain. Furthermore, these responses were all pMHC-specific as they were completely inhibited in the presence of MHC-blocking antibodies. Similar results were obtained when TCRβ chains bearing the 13-2_4β CDR3 were tested (*SI Appendix*, Fig. S8). Thus, even though there is little sequence conservation between these genes originating from species that diverged ∼400 million years ago, the residues that are conserved can provide MHC recognition. In fact, reactivity with mouse MHC was even stronger than what was observed when a mouse TCRβ was used, perhaps hinting at fine-tuning of sequence selection within a species to ensure recognition of MHC while simultaneously preventing autoreactivity.

Random TCRβ Libraries Paired to Unselected TCRα Libraries Are Primarily MHC-Reactive. In our previous experiments (described

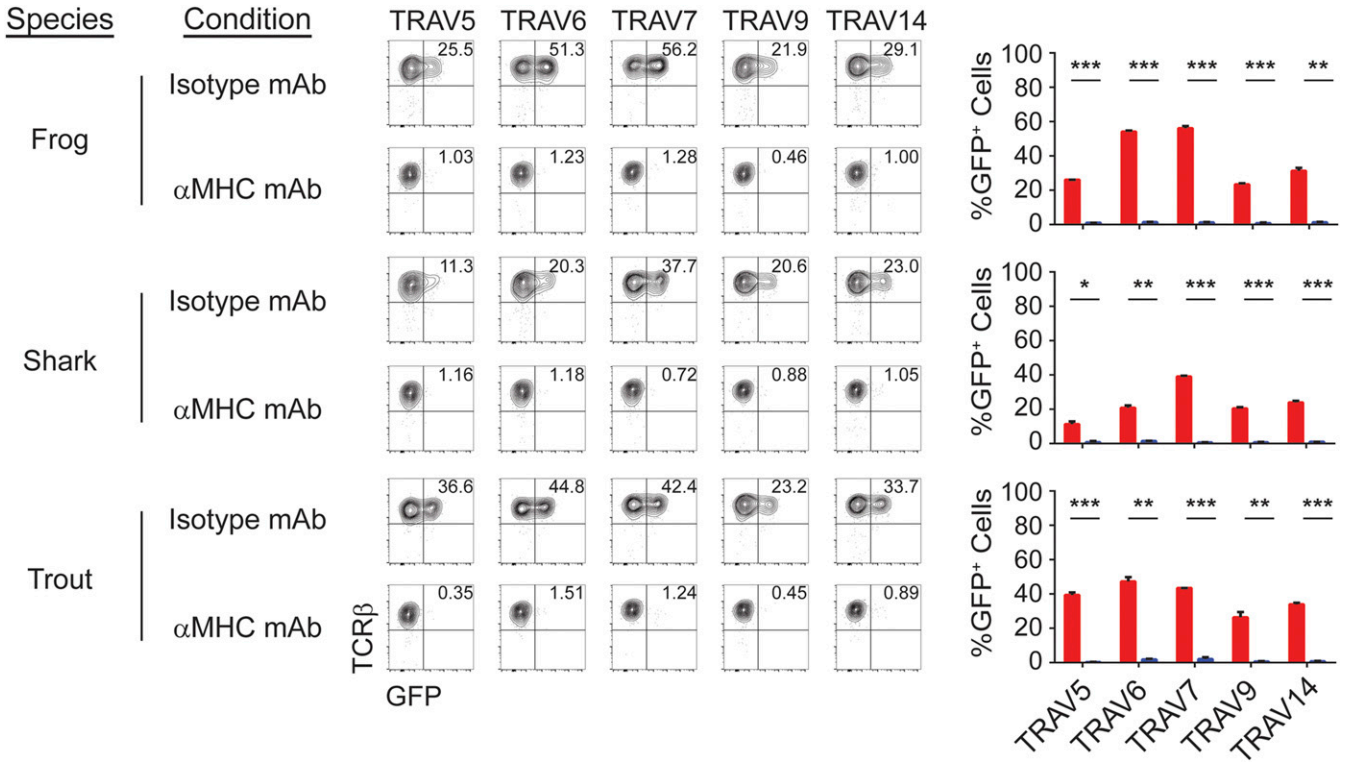


Fig. 6. Cross-species conservation of reactivity between TCR and MHC molecules. (*Left*) The entire germline region of mouse TRBV13-2 was replaced by either frog (from the TRBV2 gene), shark (from the Vβ3-Hf73 gene), or trout (from the TRBV8S2 gene). The 13-2_3W CDR3 was coupled to each of these TCRβ chains to create chimeric TCRβ chains. These chains were then independently expressed in the hybridomas expressing the CD4 coreceptor, paired to each of the libraries and then stimulated by H-2^b–bearing CHB-2.4.4 cells in the presence of isotype or MHC-blocking antibodies. After 18 to 24 h of stimulation, the cells were stained with an antibody targeting TCRβ and the proportion of GFP$^+$ cells was assessed in each stimulation condition. Results are representative of 2 independent experiments. (*Right*) The results are summarized for each stimulation condition with red representing the isotype antibody condition and blue representing the MHC-blocking antibody condition (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$ by multiple t tests; error bars are mean ± SD).

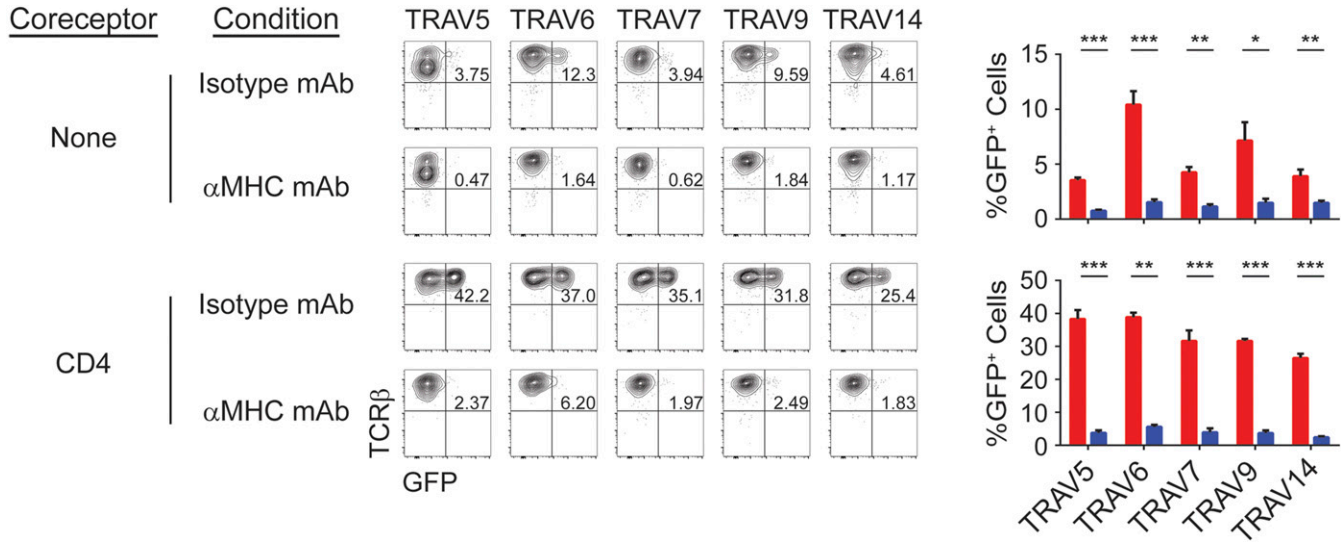


Fig. 7. TCRβ libraries paired to unselected TCRα libraries react with MHC. (*Left*) TRBV13-2 libraries were paired to each of the libraries and expressed in hybridomas either expressing the CD4 coreceptor or not. They were then stimulated by H-2^b–bearing CHB-2.4.4 cells in the presence of isotype or MHC-blocking antibodies. After 18 to 24 h of stimulation, the cells were stained with an antibody targeting TCRβ and the proportion of GFP$^+$ cells was assessed in each stimulation condition. Results are representative of 2 independent experiments. (*Right*) The results are summarized for each stimulation condition with red representing the isotype antibody condition and blue representing the MHC-blocking antibody condition (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$ by multiple *t* tests; error bars are mean ± SD).

above), we tested the reactivities of hybridomas expressing TCRα libraries paired with fixed TCRβ chains and showed that each of the TCRβ chains tested imparts its own unique bias based on its specific CDR3 sequence. This is in good agreement with previous work showing that the effects mediated by one specific TCRβ chain using a certain TRBV gene segment cannot be generalized to all rearrangements using that TRBV gene (39). Accordingly, to further extend our conclusions, instead of testing TCRβ chains one at a time, we cloned TRBV13-2–containing TCRβ chains from TCRα$^{-/-}$ DP cells and paired them with the various TCRα libraries previously described. Unexpectedly, when the TRBV13-2 library expressing hybridomas were cocultured with H-2^b–bearing APCs, we observed varying degrees of responses based on the TRAV library involved (Fig. 7). Perhaps even more surprising, these responses were specific to pMHC as they were completely abrogated in the presence of MHC-blocking antibodies. Thus, even when random, bulk TCRβ chains are paired with random TCRα chains, pMHC-reactivities can be readily observed. Expression of the CD4 coreceptor further boosted the responses just as we had previously observed in the context of fixed TCRβ chains. Indeed, proportionally more cells responded in these stimulatory conditions for each of the TCRα libraries and the responses were again pMHC-specific. Similar results, albeit with reduced responses, were achieved with APCs bearing the H-2^d MHC haplotype (*SI Appendix*, Fig. S9). Thus, although not all cells responded in any stimulatory condition, the responding cells were largely responding to pMHC ligands. Overall, our data provide evidence for a model wherein TCRs possess an intrinsic reactivity for MHC molecules prior to any selection event in the thymus, and this reactivity is further enhanced by the expression of the high-affinity CD4 coreceptor.

Discussion

The randomly generated preselection TCR repertoire can be expected to have 3 general populations vis-a-vis pMHC specificity. The first would be those TCRs that have no detectable affinity for pMHC ligands in any setting. This population could conceivably be specific for other non-MHC ligands, but because the cells expressing these TCRs do not receive the proper signals from both antigen and coreceptors, they are subjected instead to apoptotic "death by neglect" in the thymus (40). The second population would have high affinity for self-pMHC or be generally MHC-reactive in a peptide or MHC-allele promiscuous way. While thymocytes with such TCRs could potentially be initially positively selected, they are very likely to be eventually negatively selected by one or more of the host pMHC complexes (34). The final population would display an affinity for some self-pMHC complexes sufficient for positive selection, but not for subsequent negative selection, and cells expressing TCRs belonging to this population could be exported from the thymus to seed the peripheral T cell pool.

The selection vs. germline hypotheses differ by their predictions regarding the relative sizes of these 3 populations. The selection hypothesis predicts that, since the preselection repertoire is generated by the same random mechanisms that generate antibodies, the frequency of TCRs specific for a given pMHC complex should be no greater than that for epitopes of other proteins. Therefore, these frequencies are very low prior to selection; that is, most of the preselection repertoire would not be pMHC-specific and would lie in the "death-by-neglect" population described above. The germline hypothesis predicts that germline-encoded features of TCR CDR1 and CDR2 loops give them a predilection for pMHC recognition, which is fine-tuned during the somatic generation of the nongermline encoded CDR3 loops. Therefore, this hypothesis predicts that pMHC recognizing TCRs should be frequent in the initial preselection repertoire.

Distinguishing between these 2 hypotheses, therefore, requires techniques to assess the frequency of TCRs in preselection repertoire that have affinities required to participate in the 2 subsequent pMHC driven steps: Positive and negative selection. To ensure we were testing TCR sequences that had never been selected on pMHC or any other ligand, we isolated and cloned preselection repertoires using DP cells from TCRα$^{-/-}$ mice as a source material. We combined our preselection TCRα libraries with a variety of TCRβ chains, including some from TCRs highly promiscuous for pMHC, others from preselection thymocytes,

and others from other species. Since preselection thymocytes display sensitivities, especially in the case of positive selection, higher than those required to activate peripheral T cells or T cell hybridomas in vitro (41, 42), we had to develop ways to increase the sensitivity of our in vitro assays to properly assess the reactivities of the hybridomas expressing TCRs either partially or completely constructed from sequences observed in preselection thymocytes. We accomplished this in several ways. We used a sensitive fluorescent reporter assay that allowed us to assess the response of individual T cells in a population. Additionally, we incorporated a mutant form of human CD4 with high affinity for MHC-II that had previously been shown to increase the antigen sensitivity of T cell hybridomas without sacrificing antigen dependence.

This combination of approaches showed that there was a higher frequency of pMHC-specific TCRs among the preselection thymocytes than would be predicted by the selection hypothesis. These TCRs included those predicted to have high-enough affinities to drive negative selection in the thymus and also, especially revealed with the mutant CD4, those likely to allow T cell-positive selection and export from the thymus. In contrast, we did not find much evidence for nonpMHC reactivity from the preselection repertoire. This would have been detected upon GFP expression over background when any of the hybridomas that we generated were cultured either by themselves or with the APCs used, since both express high levels of numerous other surface proteins. In this context, we should note that the hybridomas express low levels of MHC-I and no detectable MHC-II. Once stimulated, some T cell hybridomas die (43); therefore, one might argue that hybridomas bearing TCRs that could react with other surface proteins expressed on the hybridomas themselves might have died before they could have been assayed in our experiments. We find this to be unlikely. First, although some of the hybridomas died when stimulated using plate-coated antibodies, there was little observable death in the antigen-presentation assays. Additionally, once isolated, the GFP^+ hybridomas showed no signs of apoptosis in culture. These sorted GFP^+ hybridomas eventually lost GFP expression (the half life of GFP is ∼26 h) (44), and could then be retested. Upon restimulation, we observed that they were enriched for response to their original targets, suggesting that the sorted cells preserve their reactivities without dying. Finally, to prevent untoward death of stimulated cells, we overexpressed Bcl-2 in the hybridomas. Introduction of Bcl-2 did not allow a higher frequency of GFP expression by the hybridomas when the cells were cultured by themselves or with APCs. However, the overexpression of Bcl-2, perhaps through its nonapoptotic function in regulating calcium signaling (45, 46), negatively affected the sensitivity of the NFAT-GFP reporter in response to stimulation by ∼8-fold (*SI Appendix*, Fig. S10), which precluded its use further.

Other studies have previously interrogated the reactivity of the preselection TCR repertoire (17, 18). In these experiments, a substantial fraction of DP thymocytes (15%) derived from MHC-deficient mice was shown to up-regulate CD69 24 h after exposure to MHC^+ thymic stroma cells (18). Similarly, the reactivities of DP thymocytes from β2m and I-Aβ doubly deficient mice that had been matured nonselectively with anti-TCRβ and anti-CD4 mAbs and fused to generate hybridomas revealed that a high proportion of pMHC-specificity exists within the preselection repertoire (17). However, these preselection repertoires might not have been completely MHC-naive. D^b molecules have been shown to traffic to the cell surfaces even in the absence of β2m (47), and other nonclassic MHC-I molecules—such as CD1d and MR1, which usually pair with β2m—can be expressed on cell surfaces without β2m (48, 49). Conventional T cells can be selected on these nonclassic MHC molecules in the absence of classic MHC (50, 51). Furthermore, in C57BL/6 mice, I-A is the only MHC-II molecule expressed on cells due to a deletion in I-Eα (52). However, I-Aα could still pair with the functional I-Eβ chain, potentially creating a hybrid MHC-II molecule (53, 54). In contrast, in the experiments described here, the TCR components we used had never been surface-expressed, ensuring that the observed reactivity of the preselection repertoire was not biased due to previous exposure to pMHC complexes or any other molecules. In addition, because the previous studies (17, 18) directly used DP cells (or hybridomas generated with them) in their stimulations, the coreceptors continued to be expressed, which may have influenced the proportion of cells they determined to be pMHC-reactive. We avoided this obstacle by instead reexpressing the TCRs from the DP cells in our hybridomas, which do not endogenously express any functional coreceptors.

Hydrophobic residues within the CDR3β have been proposed to promote pMHC-reactivity of TCRs (38). Our results support these findings and confirm previous ideas that, although TCR V regions may have been evolutionary selected to bind MHC, this bias may be affected by amino acids in the somatically generated portions of the CDR3 (55). Because bulky aromatic amino acids promote MHC reactivity, thymocytes expressing TCRs that include such residues may be deleted by negative selection in the thymus to prevent autoreactive cells from entering the periphery, regardless of the MHC allele expressed in the animal. Indeed, thymocytes in the YAe5-62.8 TCR transgenic mouse undergo negative selection and do not mature past the DP stage (34). Large residues have the potential to create more contacts with their targets and, consequently, stronger interactions and TCRs with such features are preferentially excluded from the mature TCR repertoire (56). This finding raises the question as to the frequency with which such residues promoting self-reactivity are observed in the preselection TCR repertoire. Interestingly, these residues are primarily introduced by nontemplated nucleotide addition during recombination. Our sequencing data show that large residues are not commonly found in those positions in the preselection repertoire (*SI Appendix*, Fig. S6). Instead, smaller and more polar residues dominate. Additionally, the most frequent and shared sequences are those that are short and lacking nontemplated nucleotides, indicating that large residues are excluded from those positions even in the sequences most often interrogated for selection. Thus, the germline sequences capable of contributing to the CDR3 for the different TCRβ genes may have evolved to temper strong autoreactivity, which is overcome only on the infrequent occasions when random nucleotide additions introduce codons coding for large, hydrophobic residues.

Although variations in the features that underpin TCR/pMHC recognition have been observed (57), a large proportion of the solved TCR/pMHC complex structures that populate the Protein Data Bank database display a diagonal orientation with the CDRβ loops over the MHC-II α-chain (or the MHC-I α1 domain) and the CDRα loops over the MHC-II β-chain (or the MHC-I α2 domain) (3, 57). In general, the CDR3 loops interact with the solvent-exposed residues of the peptide. Depending on the length and conformation of the CDR3 loops, they can participate in binding not only the peptide but the MHC helices as well. The CDR2 loops mainly interact with the MHC helices and the CDR1 loops usually interact with both peptide and MHC helices. However, the exact angle, forward/backward pitch, or left/right tilt can lead to certain CDR loops dominating the actual contacts with peptide and MHC. The structures that employ such orientations include a number of different human or mouse TRBV regions, including the TRBV gene used in most of the experiments presented in this report, mouse TRBV13-2, which is present in many such structures. It is possible that some TCR V regions have evolved to bind MHC in orientations other than those observed in most solved structures. This might be the case for human TRBV6-2 or mouse TRBV17, which are used by TCRs described to bind MHC in an orientation that is opposite to that usually seen (58, 59). Such TCRs have extremely weak

affinities for their pMHC complexes and upon engaging their ligands, poorly activate the T cells bearing them. This may be illustrative of other strategies used by TCRs to interact with pMHC molecules but nevertheless does not negate the idea that TCR V regions have evolved to react with MHC. The data obtained using the species-chimeric TCRs indicate that evolutionarily conserved residues were likely positively selected over evolutionary time to help mediate pMHC recognition. Further structural and high-throughput experiments testing TCR reactivity will be required to determine the relative frequency of "common" vs. "unconventional" TCR-pMHC recognition in the preselection repertoire. Here, we demonstrate that the preselection repertoire of mouse TCRs is inherently biased toward MHC recognition.

Methods

Mice. C57BL/6 and TCR$\alpha^{-/-}$ mice (29) were purchased from Jackson Laboratories. Mice were raised in a specific pathogen-free environment at the University of Colorado Anschutz Medical campus. All animal procedures were approved by the University of Colorado Institutional Animal Care and Use Committees (protocol 00065) and were carried out in accordance with the approved guidelines.

Lymphocyte Isolation. Thymi from mice between 6 and 10 wk of age were harvested and single-cell suspensions were prepared via manual disruption with a syringe plunger.

TCR Constructs and Retroviral Vectors. The NFAT-GFP retroviral reporter construct has been previously described (32) and was kindly provided by Kenneth Murphy (Washington University in St. Louis, St. Louis, MO). Murine CD3 WTδ-F2A-γ-T2A-ε-P2A-ζ pMIA II was a gift from Dario Vignali (University of Pittsburgh, Pittsburgh, PA) (Addgene plasmid # 52093; http://addgene.org/52093/; RRID:Addgene_52093) (60). The iNKTα, 75-55β, YAe5-62.8β, and DO-11.10β cloned TCR constructs have been described previously (21, 33). The TRBV13-2 TCRβ sequences were synthesized by Genscript while the TRBV13-2 library was generated by PCR using primers annealing to the TRBV13-2 leader and a primer annealing to the TCRβ constant (TRBC) gene. These sequences all contained engineered restriction sites for downstream cloning. Subsequently, the sequences were cloned into a murine stem cell virus (MSCV)-based retroviral plasmid with an internal ribosomal entry site and human NGFR as a reporter gene. The TCRα libraries were initially generated by PCR using primers annealing to the leader sequences of the various TRAV genes and a common primer annealing to the TCRα constant (TRAC) gene. The primers also contained engineered restriction sites, which were used to clone the libraries into a MSCV-based retroviral plasmid without any reporter gene. Retroviral plasmids containing the frog, shark, and trout Vβ regions have been described previously (22). Here, the CDR3 regions were introduced into those chimeric constructs by PCR using overlapping primers to mutate the CDR3. The resulting PCR products were then cloned into a MSCV-based retroviral plasmid with human NFGR as a reporter gene. A high-affinity mutated CD4 molecule (30, 31) was cloned into a MSCV-based retroviral plasmid with an internal ribosomal entry site and eBFP as a reporter gene.

Cell Lines. TCR retroviral constructs were expressed in the 58$\alpha^{-}\beta^{-}$ T cell hybridoma (32) and the cells were grown in complete DMEM. CHB-2.4.4 cells served as the APCs bearing the H-2^{b} MHC haplotype and were grown in complete S-MEM (21), while the BALB/c derived MutuDC cell line (a kind gift from Hans Acha-Orbea, Universite de Lausanne, Lausanne, Switzerland) served as the APCs bearing the H-2^{d} MHC haplotype haplotype and were grown in Iscove's Modified Dulbecco's Medium (61).

Retroviral Packaging and Transductions. Phoenix cells were used to produce retroviral particles used for infecting the hybridomas. Retroviral plasmids were cotransfected with pCL-Eco using Lipofectamine 2000 (Thermo Fisher Scientific) into Phoenix cells, as previously described (33). Supernatants containing the virus were harvested 24 h after transfection and filtered using a 0.22-μm filter. Hybridomas were then spinfected with the virus using standard techniques.

Sorting. The following mAbs were used to stain cells in this study: anti-CD4 (GK1.5), anti-CD8α (53-6.7), anti-CD69 (H1.2F3), anti-CD24 (M1/69), anti-TCRβ (H57-597), and anti-hNGFR (ME20.4) (all from BioLegend). Cells were sorted using the BD FACSAria Fusion III (BD Biosciences) directly into FBS. Cells were then either processed for RNA isolation or placed directly into culture.

Hybridoma Stimulations. For plate-coated antibody stimulations, anti-CD3 (145-2C11) and anti-CD28 (37.51) from BioLegend were added to individual wells in a flat-bottom 96-well microtiter plate at the indicated concentrations. The plate was then incubated at 37 °C for 3 h and then washed twice with 1× PBS. Subsequently, 3.5×10^4 hybridomas were added to each well in 250 μL of media. For antigen presentation assays, 3.5×10^4 hybridomas were cocultured with 3.5×10^5 CHB-2.4.4 cells or 1×10^6 MutuDC cells. MutuDC cells were stimulated with Poly I:C (GE Healthcare) at 8 μg/mL for 24 h before addition of hybridomas. In the conditions where reactivity to specific antigens was assessed, the antigens at the described concentrations were added to the culture concurrently with the hybridomas and the APCs.

After 18 to 24 h, the cells were washed and stained for 30 min with anti-TCRβ antibody at 4 °C. Subsequently, stimulated hybridomas in the various conditions were identified by their induction of GFP on a BD LSRFortessa device (BD Biosciences). The data were analyzed using the FlowJo software (Treestar). For measuring secretion of IL-2, the supernatants from these stimulations were collected and IL-2 concentration was measured by IL-2 ELISA or using the HT-2 cells via standard protocols. Blocking of MHC reactivities was achieved using a mixture of M5/114.15.2 and 28-8-6 from BioLegend (for H-2^{b} APCs) or M5/114.15.2, SF1-1.1.1 from Thermo Fisher, and 34-5-8S from Novus Biologicals (for H-2^{d} APCs), each at 15 μg/mL, while control stimulations contained the isotype antibodies (RTK4530 and MOPC-173 from BioLegend for both H-2^{b} and H-2^{d} APCs), also each at 15 μg/mL.

High-Throughput Sequencing. For TRBV13-2 sequencing, RNA from DP cells of 3 independent TCR$\alpha^{-/-}$ mice was isolated and TRBV13-2 rearrangements from each sample were independently amplified by PCR using primers targeting TRBV13-2 and TRBC. The primers also contained IonTorrent adapters and the resulting products were sequenced on the IonTorrent Personal Genome Machine. The sequences were analyzed and the CDR3 regions were identified using in-house software. For TRAV library sequencing, the libraries were amplified by PCR using primers that introduce the Illumina adapters and sample barcodes. Subsequently, the products were sequenced on the NovaSeq platform and demultiplexed and analyzed for CDR3 diversity using in-house software. To sequence the TRAV libraries present in the hybridomas, RNA was isolated from cells expressing the 13-2_3Wβ chain and amplified by PCR using the aforementioned Illumina primers.

ACKNOWLEDGMENTS. We thank members of our laboratories for thoughtful discussions and critical comments on the manuscript; Jennifer Matsuda and James Scott Browne for critical comments and support; the National Jewish Health flow cytometry facility and the University of Colorado flow cytometry shared resource facility; and the Center for Genes, Environment, and Health at National Jewish Health and the Genomics and Microarray core at the University of Colorado, Denver for sequencing. This work was supported by National Institutes of Health Grants AI121761 and AI130198 (to L.G.), AI018785 and AI135374 (to P.M.), and Cancer Center Support Grant P30CA046934.

1. D. G. Schatz, Y. Ji, Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* **11**, 251–263 (2011).
2. M. M. Davis, P. J. Bjorkman, T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395–402 (1988).
3. M. G. Rudolph, R. L. Stanfield, I. A. Wilson, How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* **24**, 419–466 (2006).
4. P. Marrack, J. P. Scott-Browne, S. Dai, L. Gapin, J. W. Kappler, Evolutionarily conserved amino acids that control TCR-MHC interaction. *Annu. Rev. Immunol.* **26**, 171–203 (2008).
5. N. L. La Gruta, S. Gras, S. R. Daley, P. G. Thomas, J. Rossjohn, Understanding the drivers of MHC restriction of T cell receptors. *Nat. Rev. Immunol.* **18**, 467–478 (2018).
6. F. Van Laethem, A. N. Tikhonova, A. Singer, MHC restriction is imposed on a diverse T cell receptor repertoire by CD4 and CD8 co-receptors during thymic selection. *Trends Immunol.* **33**, 437–441 (2012).
7. A. Singer, S. Adoro, J. H. Park, Lineage fate and intense debate: Myths, models and mechanisms of CD4- versus CD8-lineage choice. *Nat. Rev. Immunol.* **8**, 788–801 (2008).
8. G. L. Stritesky, S. C. Jameson, K. A. Hogquist, Selection of self-reactive T cells in the thymus. *Annu. Rev. Immunol.* **30**, 95–114 (2012).
9. C. Doyle, J. L. Strominger, Interaction between CD4 and class II MHC molecules mediates cell adhesion. *Nature* **330**, 256–259 (1987).
10. A. M. Norment, R. D. Salter, P. Parham, V. H. Engelhard, D. R. Littman, Cell-cell adhesion mediated by CD8 and MHC class I molecules. *Nature* **336**, 79–81 (1988).

11. A. Veillette, M. A. Bookman, E. M. Horak, J. B. Bolen, The CD4 and CD8 T cell surface antigens are associated with the internal membrane tyrosine-protein kinase p56lck. *Cell* **55**, 301–308 (1988).
12. F. Van Laethem *et al*., Lck availability during thymic selection determines the recognition specificity of the T cell repertoire. *Cell* **154**, 1326–1341 (2013).
13. F. Van Laethem *et al*., Deletion of CD4 and CD8 coreceptors permits generation of alphabetaT cells that recognize antigens independently of the MHC. *Immunity* **27**, 735–750 (2007).
14. A. N. Tikhonova *et al*., αβ T cell receptors that do not undergo major histocompatibility complex-specific thymic selection possess antibody-like recognition specificities. *Immunity* **36**, 79–91 (2012).
15. M. Blackman *et al*., The T cell repertoire may be biased in favor of MHC recognition. *Cell* **47**, 349–357 (1986).
16. H. L. Parrish, N. R. Deshpande, J. Vasic, M. S. Kuhns, Functional evidence for TCR-intrinsic specificity for MHCII. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3000–3005 (2016).
17. J. Zerrahn, W. Held, D. H. Raulet, The MHC reactivity of the T cell repertoire prior to positive and negative selection. *Cell* **88**, 627–636 (1997).
18. M. Merkenschlager *et al*., How many thymocytes audition for selection? *J. Exp. Med.* **186**, 1149–1158 (1997).
19. N. J. Felix, P. M. Allen, Specificity of T-cell alloreactivity. *Nat. Rev. Immunol.* **7**, 942–953 (2007).
20. E. J. Suchin *et al*., Quantifying the frequency of alloreactive T cells in vivo: New answers to an old question. *J. Immunol.* **166**, 973–981 (2001).
21. J. P. Scott-Browne, J. White, J. W. Kappler, L. Gapin, P. Marrack, Germline-encoded amino acids in the αβ T-cell receptor control thymic selection. *Nature* **458**, 1043–1046 (2009).
22. J. P. Scott-Browne *et al*., Evolutionarily conserved features contribute to αβ T cell receptor specificity. *Immunity* **35**, 526–535 (2011).
23. D. Silberman *et al*., Class II major histocompatibility complex mutant mice to study the germ-line bias of T-cell antigen receptors. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5608–E5617 (2016).
24. L. A. Colf *et al*., How a single T cell receptor recognizes both self and foreign MHC. *Cell* **129**, 135–146 (2007).
25. D. Feng, C. J. Bond, L. K. Ely, J. Maynard, K. C. Garcia, Structural evidence for a germline-encoded T cell receptor-major histocompatibility complex interaction 'codon'. *Nat. Immunol.* **8**, 975–983 (2007).
26. J. J. Adams *et al*., Structural interplay between germline interactions and adaptive recognition determines the bandwidth of TCR-peptide-MHC cross-reactivity. *Nat. Immunol.* **17**, 87–94 (2016).
27. B. D. McDonald, J. J. Bunker, S. A. Erickson, M. Oh-Hora, A. Bendelac, Crossreactive αβ T cell receptors are the predominant targets of thymocyte negative selection. *Immunity* **43**, 859–869 (2015).
28. E. Sharon *et al*., Genetic variation in MHC proteins is associated with T cell receptor expression biases. *Nat. Genet.* **48**, 995–1002 (2016).
29. P. Mombaerts *et al*., Mutations in T-cell antigen receptor genes alpha and beta block thymocyte development at different stages. *Nature* **360**, 225–231 (1992).
30. T. Williams *et al*., Development of T cell lines sensitive to antigen stimulation. *J. Immunol. Methods* **462**, 65–73 (2018).
31. X. X. Wang *et al*., Affinity maturation of human CD4 by yeast surface display and crystal structure of a CD4-HLA-DR1 complex. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15960–15965 (2011).
32. W. Ise *et al*., CTLA-4 suppresses the pathogenicity of self antigen-specific T cells by cell-intrinsic and cell-extrinsic mechanisms. *Nat. Immunol.* **11**, 129–135 (2010).
33. S. Sundararaj *et al*., Differing roles of CD1d2 and CD1d1 proteins in type I natural killer T cell development and function. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E1204–E1213 (2018).
34. E. S. Huseby *et al*., How the T cell repertoire becomes peptide and MHC specific. *Cell* **122**, 247–260 (2005).
35. K. Haskins *et al*., The major histocompatibility complex-restricted antigen receptor on T cells. I. Isolation with a monoclonal antibody. *J. Exp. Med.* **157**, 1149–1169 (1983).
36. K. J. Gollob, E. Palmer, Divergent viral superantigens delete V beta 5+ T lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5138–5141 (1992).
37. J. Lu *et al*., Molecular constraints on CDR3 for thymic selection of MHC-restricted TCRs from a random pre-selection repertoire. *Nat. Commun.* **10**, 1019 (2019).
38. B. D. Stadinski *et al*., Hydrophobic CDR3 residues promote the development of self-reactive T cells. *Nat. Immunol.* **17**, 946–955 (2016).
39. M. Correia-Neves, C. Waltzinger, J. M. Wurtz, C. Benoist, D. Mathis, Amino acids specifying MHC class preference in TCR V alpha 2 regions. *J. Immunol.* **163**, 5471–5477 (1999).
40. L. Klein, B. Kyewski, P. M. Allen, K. A. Hogquist, Positive and negative selection of the T cell repertoire: What thymocytes see (and don't see). *Nat. Rev. Immunol.* **14**, 377–391 (2014).
41. G. M. Davey *et al*., Preselection thymocytes are more sensitive to T cell receptor stimulation than mature T cells. *J. Exp. Med.* **188**, 1867–1874 (1998).
42. B. Lucas, I. Stefanová, K. Yasutomo, N. Dautigny, R. N. Germain, 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 (1999).
43. T. Brunner *et al*., Cell-autonomous Fas (CD95)/Fas-ligand interaction mediates activation-induced apoptosis in T-cell hybridomas. *Nature* **373**, 441–444 (1995).
44. P. Corish, C. Tyler-Smith, Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng.* **12**, 1035–1040 (1999).
45. G. P. Linette, Y. Li, K. Roth, S. J. Korsmeyer, Cross talk between cell death and cell cycle progression: BCL-2 regulates NFAT-mediated activation. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9545–9552 (1996).
46. G. Monaco *et al*., Selective regulation of IP3-receptor-mediated Ca2+ signaling and apoptosis by the BH4 domain of Bcl-2 versus Bcl-Xl. *Cell Death Differ.* **19**, 295–309 (2012).
47. T. A. Potter, C. Boyer, A. M. Verhulst, P. Golstein, T. V. Rajan, Expression of H-2Db on the cell surface in the absence of detectable beta 2 microglobulin. *J. Exp. Med.* **160**, 317–322 (1984).
48. S. P. Balk *et al*., Beta 2-microglobulin-independent MHC class Ib molecule expressed by human intestinal epithelium. *Science* **265**, 259–262 (1994).
49. J. Lion *et al*., MR1B, a natural spliced isoform of the MHC-related 1 protein, is expressed as homodimers at the cell surface and activates MAIT cells. *Eur. J. Immunol.* **43**, 1363–1373 (2013).
50. B. D. Stadinski *et al*., A role for differential variable gene pairing in creating T cell receptors specific for unique major histocompatibility ligands. *Immunity* **35**, 694–704 (2011).
51. R. E. Berg *et al*., Positive selection of an H2-M3 restricted T cell receptor. *Immunity* **11**, 33–43 (1999).
52. D. J. Mathis, C. Benoist, V. E. Williams, 2nd, M. Kanter, H. O. McDevitt, Several mechanisms can account for defective E alpha gene expression in different mouse haplotypes. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 273–277 (1983).
53. J. R. Dorfman, I. Stefanová, K. Yasutomo, R. N. Germain, CD4+ T cell survival is not directly linked to self-MHC-induced TCR signaling. *Nat. Immunol.* **1**, 329–335 (2000).
54. B. Martin, C. Bourgeois, N. Dautigny, B. Lucas, On the role of MHC class II molecules in the survival and lymphopenia-induced proliferation of peripheral CD4+ T cells. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6021–6026 (2003).
55. P. Marrack *et al*., The somatically generated portion of T cell receptor CDR3α contributes to the MHC allele specificity of the T cell receptor. *eLife* **6**, e30918 (2017).
56. A. Kosmrlj, A. K. Jha, E. S. Huseby, M. Kardar, A. K. Chakraborty, How the thymus designs antigen-specific and self-tolerant T cell receptor sequences. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16671–16676 (2008).
57. J. Rossjohn *et al*., T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200 (2015).
58. D. X. Beringer *et al*., T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. *Nat. Immunol.* **16**, 1153–1161 (2015).
59. S. Gras *et al*., Reversed T cell receptor docking on a major histocompatibility class I complex limits involvement in the immune response. *Immunity* **45**, 749–760 (2016).
60. J. Holst *et al*., Scalable signaling mediated by T cell antigen receptor-CD3 ITAMs ensures effective negative selection and prevents autoimmunity. *Nat. Immunol.* **9**, 658–666 (2008).
61. S. A. Fuertes Marraco *et al*., Novel murine dendritic cell lines: A powerful auxiliary tool for dendritic cell research. *Front. Immunol.* **3**, 331 (2012).