Supplementary Materials for

Pre–T cell receptors topologically sample self-ligands during thymocyte β-selection

Xiaolong Li*, Réka Mizsei*, Kemin Tan, Robert J. Mallis, Jonathan S. Duke-Cohan, Aoi Akitsu, Paul W. Tetteh, Abhinav Dubey, Wonmuk Hwang, Gerhard Wagner, Matthew J. Lang, Haribabu Arthanari, Jia-huai Wang†, Ellis L. Reinherz†

*These authors contributed equally to this work.
†Corresponding author. Email: jwang@crystal.harvard.edu (J.W.); ellis_reinherz@dfci.harvard.edu (E.L.R.)

Published 17 December 2020 on *Science* First Release
DOI: 10.1126/science.abe0918

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.abe0918/DC1)

Materials and Methods

Molecular cloning and inclusion bodies expression of proteins

K^b, β-2 microglobulin (β2m), K^b-t variants and N15β variants, were produced as detailed in (*17, 18*) with modifications for protecting the free cysteine residue in case of K^b-t and N15β variants. N15β used for crystallography and NMR denotes the N15β-c1 as detailed in (*18*). K^b-t2 is Q121C variant of K^b-t, the mutation was generated to change conserved Gln121 to Cys to facilitate single incorporation of Cys for chemical linkage. All protein constructs were cloned into pET-11d expression vector (New England BioLabs). Single-cysteine mutations were generated by site-directed mutagenesis using standard protocols (Invitrogen). Recombinant plasmids were transformed to *Escherichia coli* BL21(DE3), and unlabeled proteins for crystallography or isotopically labeled proteins for NMR were expressed into inclusion bodies (ib) using LB (unlabeled, 25 mg/ml Luria broth, Sigma, 100 ug/ml Carbenicillin), M9 ($^2H^{13}C^{15}N$-labeled; 50 mM Na_2HPO_4, 20 mM KH_2PO_4, 10 mM NaCl, 18 mM $^{15}NH_4Cl$, 2 mg/ml $^2H^{13}C$-glucose, 2 mM $MgSO_4$, 20 μM $CaCl_2$, 100 mg/L carbenicillin, 10 μM $FeCl_2$, and 2 ml Vitamin Cocktail in 99.9% D_2O). Isotopes were from Cambridge Isotope Laboratories, Inc.

Refolding

The octapeptides VSV8 (RGYVYQGL) and its L4 variant (RGYLYQGL) and Q4H7 (SIIQFEHL) were chemically synthesized (United Biosystems). The identity of the molecular ions was confirmed by mass spectrometry and the purity of >95% was verified by HPLC. The inclusion body preparations of all protein variants (K^b, β2m, K^b-t, and

N15β) were washed thoroughly in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl (TBS), and TBS + 1% Triton X-100 (TBS-T) and dissolved in 5.4 M Guanidine-HCl, 0.1 M Tris-HCl (pH 8.0).

K^b, β2m and Q4H7 were mixed in1:3:0.5 mass ratio and diluted into 20 mM Tris-HCl (pH 8.0), 8 M Urea buffer, and serially dialyzed against 2 M, 1 M, 0.5 M, and 0 M urea in 20 mM Tris-HCl (pH 8.0) for 2 hours with a final additional dialysis against 0 M urea, 20 mM Tris-HCl (pH 8.0) overnight. VSV8-K^b-t, VSV8-K^b-t2-G56C, VSV8-K^b-t2-R62C and L4-K^b-t2-G56C were refolded identically to Q4H7-K^b with the exception that β2m was not included; 1 mM reduced glutathione (GSH) and 0.1 mM oxidized glutathione (GSSG) were added to 8 M Urea, and Q4H7 was substituted with the appropriate peptide. N15β chain ib preparations were washed thoroughly in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl (TBS) and (TBS-T), dissolved in 6 M Guanidine-HCl and refolded by dilution in 5.4 M Guanidine-HCl, 0.1 M Tris-HCl (pH 8.0), 1 M arginine, 1 mM reduced glutathione (GSH), and 0.1 mM oxidized glutathione (GSSG). This was followed by dialysis in TBS for >16 hours. All proteins were purified by successive rounds of size-exclusion chromatography.

Covalent linkage of heterodimers

Each protein (30 μM) containing a single non-disulfide bonded cysteine was prepared for linkage by reduction using 25 mM (β chains) or 5 mM (K^b-t2 variants) dithiothreitol (DTT) in phosphate buffered saline (PBS, pH 7.4) for 30 min at 25°C. After reduction, the protein pairs were mixed at a 1:1 molar ratio and immediately separated via analytical size exclusion chromatography (Superdex S200A, GE Healthcare Life Sciences) to

remove the DTT. The corresponding fractions were collected immediately and concentrated to 15 μM (of each subunit). A 45-μM 1,11-bis(maleimido)triethylene glycol (BMPEG3, Thermo Fisher Scientific) linker at three times excess was then added. The mixture was incubated for 20 min at 25°C and was purified by two rounds of size-exclusion chromatography.

Materials for crystallization screening

Four protein samples were crystallized. Three preTCR–pMHC constructs were generated via linking N15β-S62C to VSV8-K^b-t2-G56C, VSV8-K^b-t2-R62C, or L4-K^b-t2-G56C using BMPEG3 crosslinker and dubbed N15β–K^b-t2-G56, N15β–K^b-t2-R62, and N15β–K^b-t2-L4, respectively. The fourth pMHC sample contains Q4H7, full extracellular domain of K^b and β2m and dubbed Q4H7-K^b. The commercial crystallization screens from Hampton Research, Qiagen and Microlytic were used for initial trials, including Index, Screen, JCSG Core Suites and Top96. The screening and optimization for crystallization conditions were set up with a Formulatrix NT8 robot using the sitting drop vapor diffusion technique in INTELLI-PLATE® 96 Well (Art Robbins Instruments) in the crystallography facility core of Dana-Farber Cancer Institute.

Crystallization condition

N15β–K^b-t2-G56, N15β–K^b-t2-R62, and N15β–K^b-t2-L4 complex samples were concentrated to ~8 mg/ml in 0.5 X PBS buffer. Full length Q4H7-K^b was concentrated to 12 mg/ml in 1X PBS buffer. Initial crystal screening was carried out using sitting drops with 0.1 μl of protein and 0.1 μl of crystallization solution set up by NT8 robot

(FORMULATRIX); then the mixture was equilibrated against 50 µl of the crystallization solution in each reservoir well. Crystals of N15β–K^b-t2-G56 appeared under multiple conditions, including (i) 12% (w/v) PEG 20K, 0.1 M MES (pH 6.5); (ii) 10% (w/v) PEG 6K, 5% (v/v) (±)-2-methyl-2,4-pentanediol, 0.1 M HEPES (pH 7.5); and (iii) 10% (w/v) PEG 8K, 8% (v/v) ethylene glycol 0.1 M HEPES (pH 7.5). Crystals of N15β–K^b-t2-R62 appeared under the conditions of 0.1 M CHES:NaOH (pH 9.5), 20% (w/v) PEG 8K as well as 0.2 M Calcium chloride dihydrate, 0.1 M Bis-Tris (pH 5.5), 45% (v/v) (±)-2-Methyl-2,4-pentanediol. Crystals of N15β–K^b-t2-L4 appeared under the condition of 0.1 M CHES:NaOH (pH 9.5), 20% (w/v) PEG 8K. Different PEGs and pH were optimized for the three complexes, and improved needle-like crystal clusters appeared in 15 to 20% PEG 8K, Tris-HCl (pH 8.5). Then, crystal macro-seed stocks were created in 500 ul 18% PEG 8K, Tris-HCl (pH 8.5) by using Seed Bead Kit from Hampton Research. 0.1µl protein samples were mixed with equal amounts of the seed stocks in INTELLI-PLATE® 96-well plates and equilibrated against the seed stocks. Crystals of Q4H7-K^b grew under the condition of 0.2 M Potassium Sodium Tartrate 20% (w/v) PEG 3350.

X-ray diffraction data collection

Single crystals were harvested by using Mounted CryoLoop (Hampton Research) and transferred to cryoprotectant solution that contains 10 to 20% (v/v) glycerol in addition to crystallization buffer for a few minutes and then cryocooled directly in liquid nitrogen. All X-ray diffraction data were remotely collected using the 19-ID beamline of the Structural Biology Center at the Advanced Photon Source at Argonne National Laboratory.

Structural determination, refinement and structural analysis

Three complex structures, N15β–K^b-t2-G56, N15β–K^b-t2-R62, and N15β–K^b-t2-L4 as well as the Q4H7-K^b structure were determined. Diffraction data were integrated, scaled and merged by using the HKL-3000 program suite (*32*). The CCP4i program suite (*33*) was used for data transformation etc., when necessary. The structure of N15β–K^b-t2-G56 was first determined using molecular replacement method with the program Phaser (*34*) in the PHENIX program suite (*35*). The search template for the pMHC component of K^b-t2 was truncated from the full-length K^b structure (PDB code:1KPU), whereas the template used to search N15β was from the TCRαβ complex structure (PDB code:1NFD) containing the same full-length TCR β-chain. The molecular replacement solutions were inspected, and the final model was selected and improved with the program COOT (*36*) based on electron density maps. The structures of N15β–K^b-t2-R62, and N15β–K^b-t2-L4 were subsequently determined based on the resolved N15β–K^b-t2-G56 structure. The refinements of all three complex structures were performed by using Phenix.refine (*37*) with the refined model fitting the electron density well. The chemical linker introduced no bias. In all three crystals, there were only weak and discontinuous electron densities for the chemical linker, suggesting that the BMPEG3 linker is mobile and, hence unlikely to affect the binding between N15β and VSV8-K^b-t2 (fig. S2) and concordant with the unlinked proteins by NMR analysis (fig. S4)

The structure of Q4H7-K^b also was determined using the molecular replacement method (*34*,*35*), and the search model was the modified 1KPU structure. The structure refinement and inspection followed the same approach described above.

Structural presentation and analysis were performed with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). The buried surface and interface residue-contact of preTCR–VSV8-K^b-t2 were calculated by using CCP4i (*33*). The interface shape complementarity was measure following reference (*38*).

For the three essentially isomorphous complex structures, N15β–K^b-t2-G56, N15β–K^b-t2-R62, and N15β–K^b-t2-L4, each asymmetric unit contains three N15β–VSV8-K^b-t2 complexes: ABC, DEF, and GHI, in which the chains A, D and G designate the truncated MHC molecule K^b-t2, the chains B, E and H are the VSV8 peptide, and the chains C, F and I are N15β. Collectively there are nine crystallographically independent complexes in the three crystals. Among all possible molecular contacts with buried surface area (BSA) more than 900 $Å^2$ in these crystals, the N15β–VSV8-K^b-t2 interface seen in Fig. 1A is the only one that is conserved, suggesting its physiological relevance. This fact can be appreciated in table S2. In contrast, fig. S3 gives other possible interacting pairs having interface BSA values above 900 $Å^2$**.** As discussed in fig. S3, all these interacting pairs result from crystallographic packing artifact.

For additional analyses we combined the N15 Vβ domain with the VSV8-K^b-t2 α1α2 module as an interacting structural unit. Using the complex ABC from N15β–K^b-t2-G56 as a reference, superimposing the other eight complexes to the ABC complex results in backbone RMSD values of ~1 Å (table S3). Thus, the binding observed between N15β

and VSV8-K^b-t2 is independent of constructs, molecular packing in crystals and crystallization conditions.

NMR data acquisition and processing

NMR spectra were acquired using Topspin (Bruker) or VNMRJ (Varian) and processed using NMRPipe (*39*) together with hmsIST (*40*). Peak fitting and quantitation as well as general spectral analysis used CARA (*41*).

Relaxation Dispersion (Rex)

The TROSY version of ^{15}N Rex (*42, 43*) using 400 μM $^{2}H^{13}C^{15}N$ N15β alone or with 400 μM unlabeled VSV8-K^b-t was executed on 600 and 700 MHz Varian spectrometers at 25°C. The pulse program used interleaved randomized sequence of spin-echo repeats of 0, 2, 4, 10, 16, 20, 26, 34, and 40 cycles of 25 ms each with 0, 4, 16, and 34 cycles repeated for error analysis. Data were analyzed using the Relax program for relaxation dispersion analysis (*44, 45*) with data fit to the two-site fast-exchange model LM63 (*46*).

Chemical Exchange Saturation Transfer (CEST)

CEST experiments (*47*) were acquired on a 750 MHz Bruker spectrometer with a 400 μM $^{2}H^{13}C^{15}N$ N15β sample at 25°C. A series of 64 2D planes were acquired with different offsets in the ^{15}N CEST which utilized saturating field strength of 25 Hz and duration of saturation as 300 ms. The recycle delay was set to 1.5 s. Normalized intensity data were plotted and analyzed with an in-house script using R (R Core Team (2019) R: A Language and Environment for Statistical Computing. Vienna, Austria).

Mice

Six-week-old C57Bl/6 (B6) mice and B6*Rag2*$^{-/-}$ (RAGN12) were purchased from Taconic Farms Inc. and housed at the DFCI Animal Facility, accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). All maintenance, breeding, and experimental procedures were approved under Dana-Farber Cancer Institute Institutional Animal Care and Use Committee (IACUC) protocols 03-138 and 04-113. Euthanasia was by CO_2 inhalation followed by cervical dislocation. Following removal from the uterus, E14.5 fetuses were euthanized by decapitation with surgical scissors. Where appropriate, no gender preference was expressed for experimental animal use.

Analysis of B6 thymocyte development in vitro

Isolation of wild-type hematopoietic stem cells (HSC) followed the procedure described previously (*18*). Briefly, fetal liver cells from E14.5 B6 embryos were depleted of B cells using anti-CD24 and complement lysis followed by staining with anti-CD4-Pacific Blue, anti-CD8-PE, anti-ScaI-FITC and anti-CD117(c-Kit)-APC. The $CD4^-CD8^-$ (lineage-negative, lin^-) $ScaI^+$ $c\text{-}Kit^+$ cells (HSC) were isolated by a Becton-Dickinson FACS Aria II cell sorter and seeded onto 70 to 90% confluent layers of wild-type OP9-DL4 cells or MHC-negative OP9-DL4 cells (*16*) at 2000 HSC/well in six-well plates in α-MEM without nucleosides + 15% FCS (OP9 media), HEPES (10 mM), and gentamycin supplemented with Flt3 (5 ng/ml) and IL-7 (1 ng/ml). After growth for 9 days and 13 days, cells were isolated from the culture, counted and then stained (fig. S9) with anti-

CD45-APC, anti-CD4-Pacific Blue, anti-CD8-PE, anti-CD25-PE/Cy7 and anti-CD44-APC/Cy7 for simultaneous collection of DN3 and DN4 thymocytes on a FACS Aria II cell sorter. Contaminating OP9 cells expressed GFP permitting their exclusion while selection for CD45 expression ensured only hematopoietic cells were used for subsequent subset delineation. Cells were further selected on the basis of being $CD4^-CD8^-$ (double-negative, DN) from which 10,000 DN3 ($CD25^+CD44^-$) cells, and 10,000 DN4 ($CD25^-CD44^-$) cells, respectively, were collected. For each population, the collected cells were immediately deposited into TCL lysis buffer (Qiagen) supplemented with 2-mercaptoethanol (1%) on ice, snap-frozen by immersion in dry-ice–methanol and stored at –80°C until processed for RNA extraction and TCRβ chain repertoire analysis.

N15β mutant functional analysis

HSCs from the fetal liver of B6*Rag2*$^{-/-}$ mice (E14.5) were cultured on OP9-DL4 cells for 6 days, and then transduced with cDNAs encoding N15β or mutant N15β followed by overnight incubation as described (*18*). Two thousand transduced GFP^+DN3 cells were sorted and plated onto single chain VSV8-K^b expressing scOP9 stromal cells (scOP9) (*16*), which were plated the day before at 5×10^4 cells in OP9 medium and cultured with Flt-3 (5 ng/ml), IL-7 (1 ng/ml), and gentamicin (5 μg/ml). After 6 days and 10 days, cells were isolated from the culture, counted and then stained with anti-CD45-APC, anti-CD4-Brilliant violet 711, anti-CD8-PerCP-Cy5.5, anti-CD25-PE/Cy7, anti-CD44-APC/Cy7, and anti-CD28-PE (table S8). Dead cells were stained with Zombie Aqua. All cells were gated on lymphocyte FSC/SSC, live cells, $CD45^+$, and GFP^+ cells to analyze the transfected cells. $CD4^-CD8^-$ (DN), and $CD4^+CD8^+$ (DP) cells were subsequently

analyzed. DN cells were further gated on $CD44^-$, and DN3a ($CD25^+CD28^-$), DN3b ($CD25^+CD28^+$), and DN4 cells ($CD25^-CD28^+$) were analyzed. Cells were analyzed on LSRFortessa (BD Biosciences). Data were analyzed with FlowJo software (Tree Star). To assess preTCR complex surface expression levels on thymocytes, N15 WT and mutant TCRβ retroviruses were transfected into SCID.adh cells using Lipofectamine (Invitrogen) as previously described (*18*). After overnight incubation, cells were collected, stained with anti-PE-H57-597 mAb, and analyzed on LSRFortessa (BD Biosciences).

TCR β-chain repertoire analysis of DN3 and DN4 thymocytes.

Total RNA was extracted from each sample of 10^4 cells using the PicoPure column purification system (Applied Biosystems). Subsequently, the procedure followed precisely that described in (*48*). Briefly, using a 3′ *Trbc* (TCRβ constant region) universal primer, 1st strand cDNA was synthesized from the starting RNA and a universal "Switch" primer ligated to the 5′ ends. Nested/extended PCR amplification through the universal ends yielded unbiased amplification of transcripts containing the complete V(D)J region and a 5′ segment of the *Trbc*. In the second PCR, pentanucleotide bar codes were introduced to tag each library with unique barcodes at both 5′ and 3′ ends. Following quality control using the Agilent 2100 Bioanalyzer and Illumina adapter addition, samples were sequenced (150 PE) on the MiSeq platform. Library sequences were deconvoluted from the fastx sequence output files using the barcode splitter module of the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). The deconvoluted library sequences were aligned to TCR Vβ regions in the

GRCm38.p6/mm10 mouse genome followed by clone assembly and CDR3 extraction using the MiXCR suite running under Java (*49*). Output provided V, D, J, and Cβ usage, CDR3 nucleotide and amino acid sequence, sequence quality, and relative representation by read count (attached table S7, representative of five experiments). The VDJtools analytical package was used to track and compare clonotypes through samples (*50*). CDR3 hydropathy was calculated using the Kyte-Doolittle scale (*51, 52*).

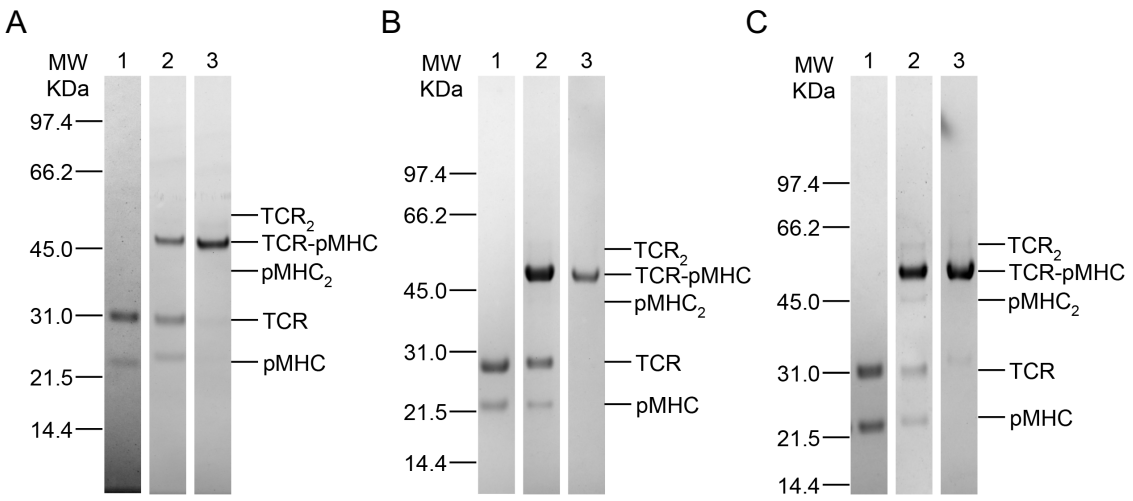

Fig. S1. N15β–K^b-t2 Linkage reactions. **A)** N15β–K^b-t2-G56, **B)** N15β–K^b-t2-R62 and **C)** N15β–K^b-t2-L4 using the bifunctional 1,11-bis(maleimido)triethylene glycol (BMPEG3) linker analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Each panel shows the mixture of individual protein components (pMHC, TCR) (lane 1), the composition of linkage-reaction-mixture (lane 2) and the preTCR–pMHC heterodimers used for crystallography (lane 3) purified from (lane 2) by two successive rounds of size exclusion chromatography (Superdex S200A, GE Healthcare Life Sciences, Chicago, IL, USA). The SDS-PAGE experiments were carried out under reducing conditions using 4 to12% Bis–Tris gradient gels (Invitrogen, Carlsbad, CA, USA, NuPAGE, Cat. No.: NP0323BOX). The molecular weights were referenced to Low Range SDS-PAGE Molecular Weight Standards (Bio-Rad, Hercules, CA, USA, Cat. No. 161-0304, containing: lysozyme (14.4 KDa), trypsin inhibitor (21.5 KDa), carbonic anhydrase (31.0 KDa), ovalbumin (45.0 KDa), BSA (66.2 KDa), phosphorylase b (97.4 KDa), visualized by Coomassie blue R-250 dye (Imperial™ Protein Stain, Fisher Scientific, Hampton, NH, USA), imaged, and quantified using Image Lab, (Bio-Rad).

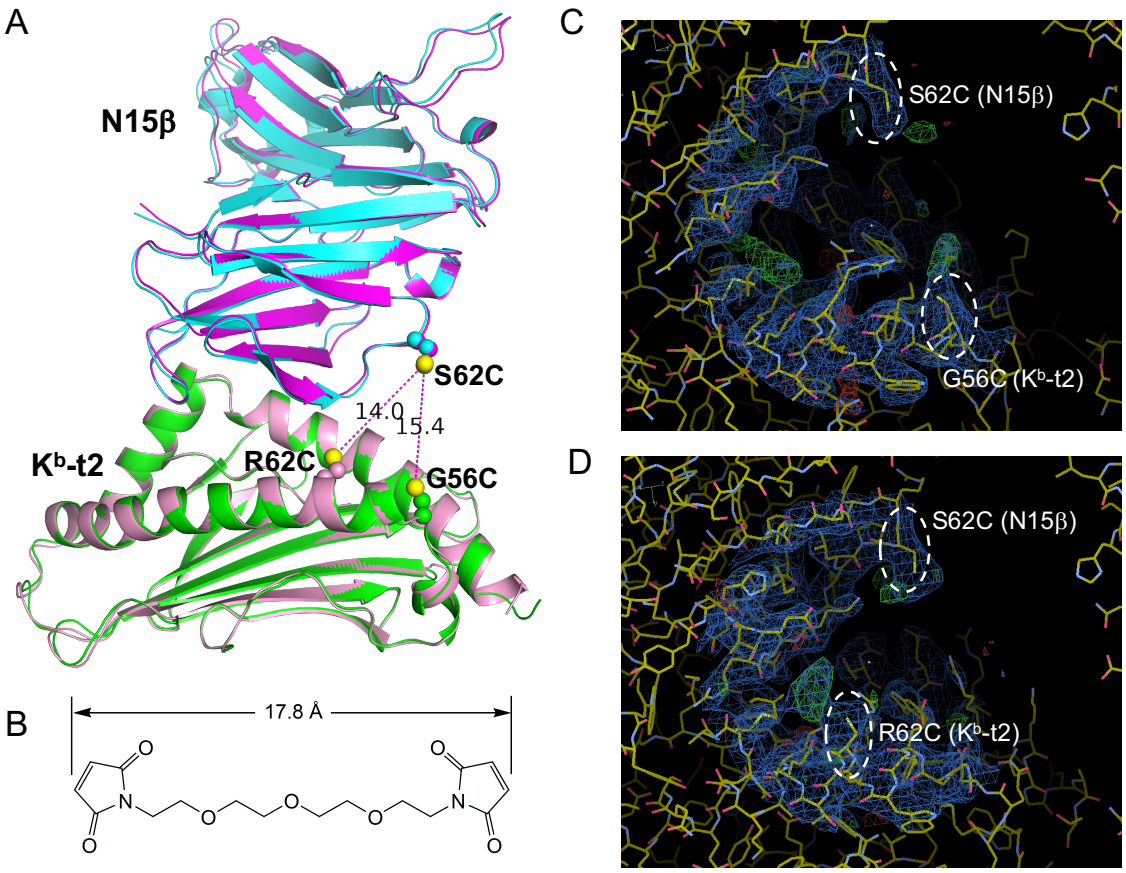

Fig. S2. Linkage sites in the N15β–VSV8-K^b-t2 complexes. A) Two N15β–VSV8-K^b-t2 complex structures were superimposed to show their different engineered linkages. For both linked complexes, N15β has the same mutation site to introduce cysteine but pairs with different mutation sites for cysteine along the α1-helix of K^b-t2, giving rise to slightly different distances between the two sulfur atoms (yellow spheres) of the pairwise cysteine residues. Specifically, the cyan and green ribbons represented the structure of N15β–K^b-t2-G56 which has an N15β S62C mutation and a K^b-t2 α1-helix G56C mutation (distance, 15.4 Å). The magenta and pink ribbons represented the structure of N15β–K^b-t2-R62 which has the same N15β S62C mutation but a K^b-t2 α1-helix R62C

mutation (distance, 14 Å). **B)** The chemical structure of the applied linker, BMPEG3. The length of the extended BMPEG3 linker is 17.8 Å. **C)** The electron density map of one representative of N15β–K^b-t2-G56 complexes. The two mutated cysteine residues are highlighted with broken circles. The 2Fo–Fc map is in blue (σ, 1.0), and the Fo–Fc difference map is in green, and red(σ, 3.0). **D)** The electron density map of one representative of N15β–K^b-t2-R62 complexes. The two mutated cysteine residues are highlighted with broken circles. The 2Fo–Fc map is in blue (σ, 1.0), and the Fo–Fc difference map is in green, and red(σ, 3.0). The absent density for linker in both constructs implies segmental mobility.

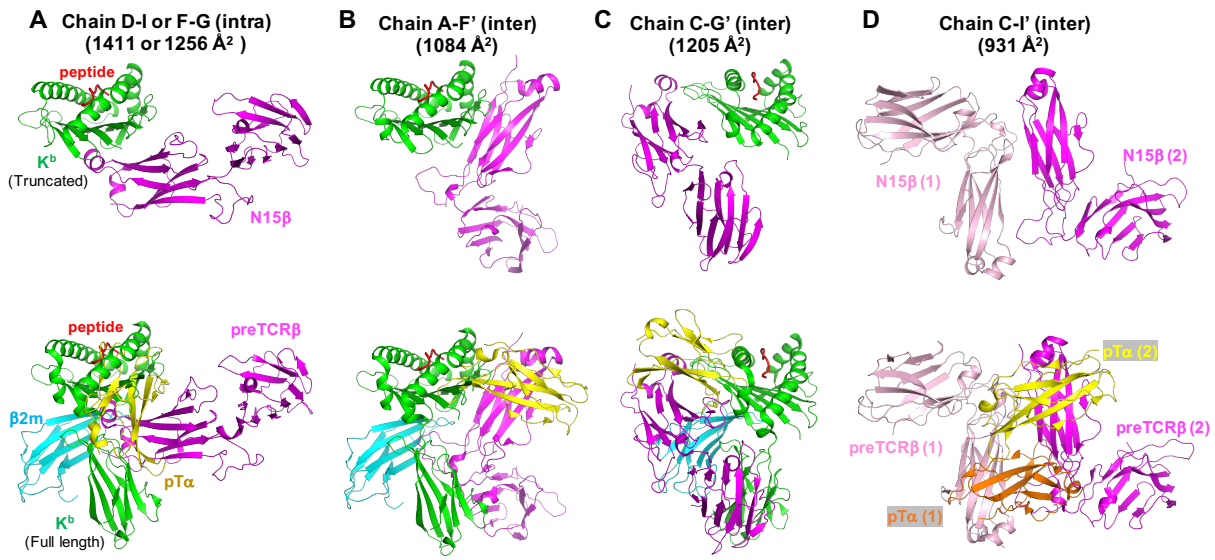

Fig. S3. Crystal-packing analysis. Besides the observed interacting interface between N15β and VSV8-K^b-t2 shown in Fig.1A, there are other molecular interfaces resulted from crystal packing. Shown here are those interactions with buried surface area (BSA) greater than 900 Å^2 (table S2). Top row displays these interacting pairs observed within an asymmetric unit (intra) or between symmetry-related ones (inter). None of these interfaces involves VSV8 peptide. In addition, these crystallographic interactions would be blocked by other subunits of the biological peptide-MHC such as the β2m and the α3-domain in the full-length K^b (panels **A-C**) and/or the pTα component of the preTCR (panel **A-D**). Hence, none of the interacting pairs shown in this figure would be physiological. They are all crystallographic packing artifacts. The interface shown in Fig. 1A is the only one that involves peptide interaction, and also the only one present in all three independent pairs in the asymmetric unit with a large BSA value and without such collision. Therefore, only the interacting pair observed in Fig. 1A represents physiological binding between a preTCR to its pMHC ligand.

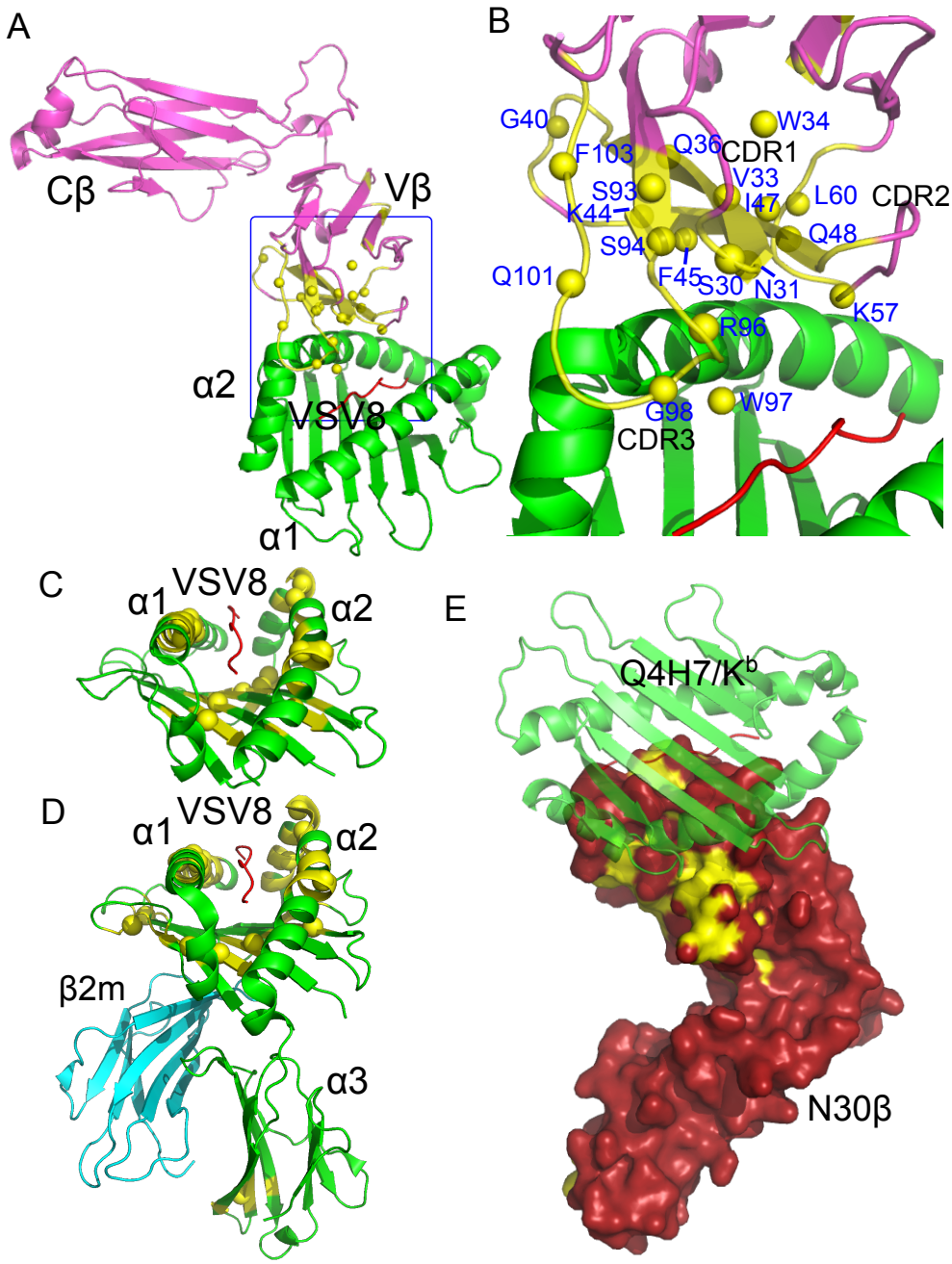

Fig. S4. Concordance of NMR resonance perturbation on complexation for N15β and N30β with pMHC and the crystallographic model of preTCRβ interaction with ligands. A-B: Overlay of NMR signal changes of N15β + VSV8-K^b-t. **A)** X-ray crystal structure is consistent with Vβ chemical shift perturbations as previously published (*17,*

18). Perturbed residues (top 10%) are yellow spheres with structural elements within 5 Å also colored yellow. **B)** Closeup view of A with CDR1, CDR2 and CDR3 as well as perturbed residues labeled. Note that residues in CDR1 that are changed are proximal to CDR3 and the binding interface. Similarly, Q48 is the only *bona fide* residue adjacent to the CDR2 loop in the interface. Residues 49 to 55 were not assigned in NMR data for N15β. CDR2 was previously inferred to modulate binding based on functional data with a single mutation in thymic stromal culture (*18*) perhaps due to its close packing with CDR3 including W97 (*14, 18*). **C-D)** Intensity loss of VSV8-K^b-t **C)** and VSV8-K^b **D)** residues are consistent with a peptide groove-centric binding interface (*17*). Although it is difficult to definitively assign sidedness to changes in the floor of the peptide groove in the truncated K^b-t, given the changes to the peptide-flanking helices, it is likely that these alterations arise from transmitted changes through the peptide. In the full-length construct, with one exception the largest changes are found in proximity to the peptide-binding groove and not the distal α3 domain, corroborating the binding mode of the X-ray structure. **E)** NMR chemical shift and peak intensity changes of N30β binding Q4H7-K^b-t (*19*) denoted as Q4H7/K^b mapped onto the current X-ray structure suggest that the X-ray-reported binding mode is also operative for an unrelated preTCR model protein binding to a different peptide bound to K^b-t.

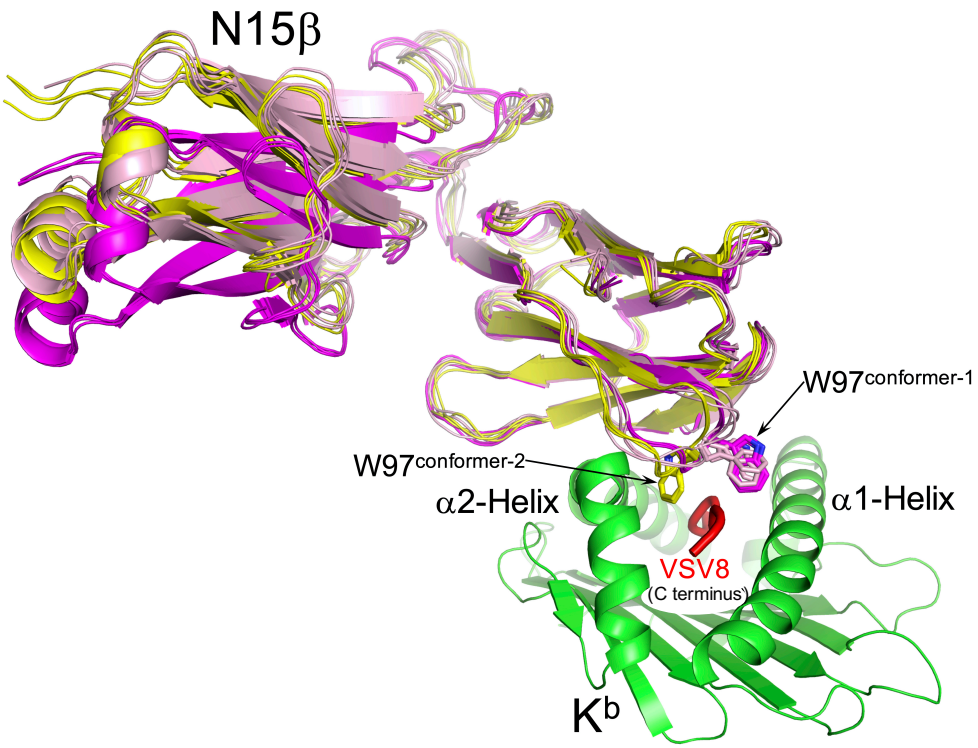

Fig. S5. Two conformer groups of N15β–VSV8-K^b-t2. Nine N15β–VSV8-K^b-t2 complexes from three crystal structures were superimposed on their K^b-t2 components (green ribbon) only. There are two distinguishable conformer groups in the binding of N15β to K^b-t2, shown in magenta and yellow ribbons, respectively. They differ in the W97 side chain conformation at the top of the CDR3 loop and have a slightly different bending angle between V and C domains of N15β. The VSV8 peptide is shown in red worm.

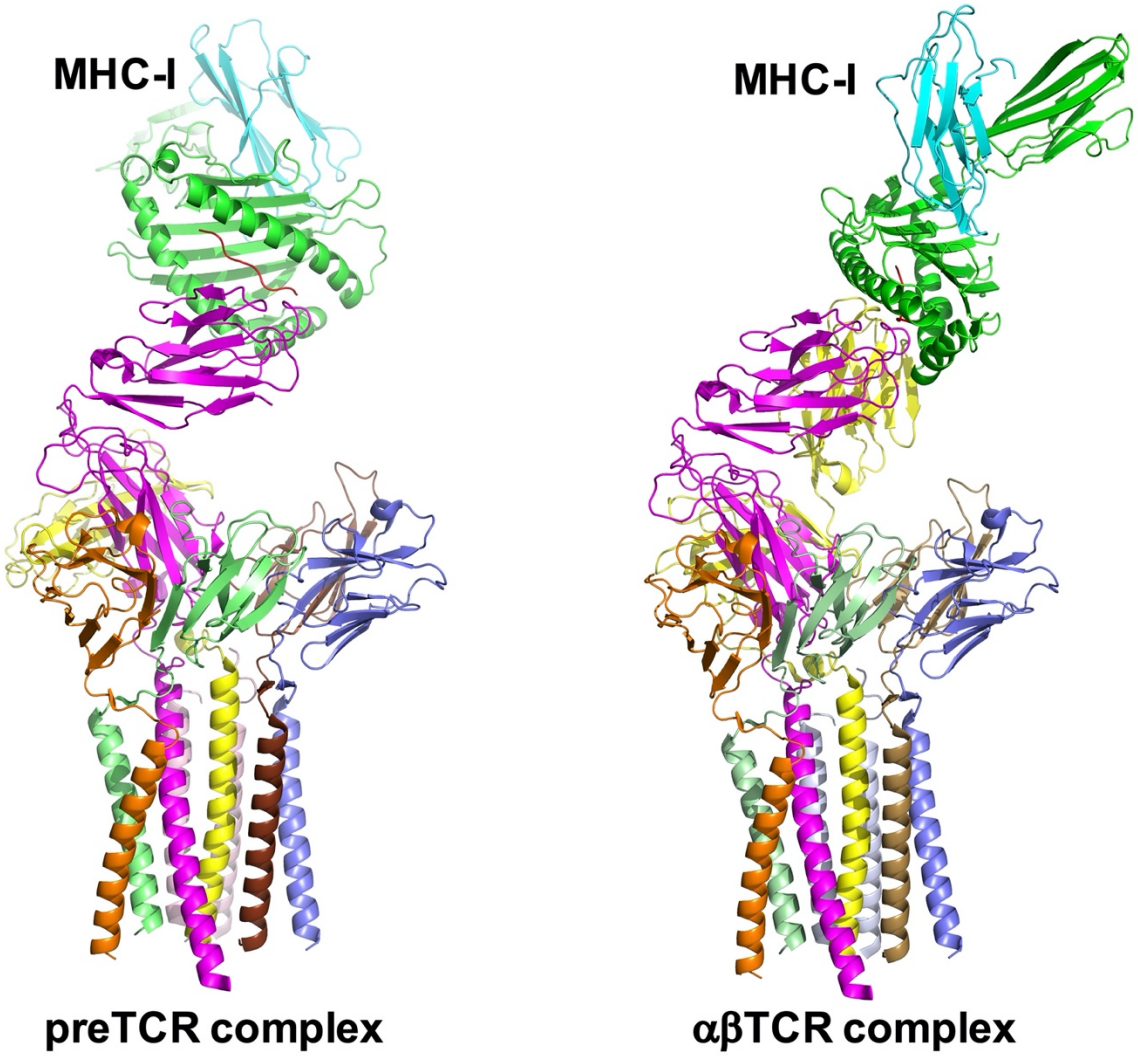

Fig. S6. Models of full complexes of preTCR vs αβTCR and their interaction with pMHC. Models were constructed by overlaying the N15β–VSV8-K^b-t2 structure reported here and the human preTCR structure (PDB: 3OF6) (left panel) as well as TCRαβ–pMHC structure (PDB: 2CKB) (right panel) onto the full αβTCR complex structure including the CD3 subunits (PDB: 6JXR). The magenta ribbon represents TCR β chain, whereas the yellow ribbon represents TCR α (or pTα) chains. CD3 γ, δ, ε, and ζ

subunits are in brown, green, blue, and light blue ribbons, respectively. The helical bundles represent the transmembrane segments that span the plasma membrane of the T-lineage cell. This figure demonstrates that the horizontally bound preTCR–pMHC is as compatible as the vertically bound TCRαβ–pMHC when in the full receptor complex associated with CD3 subunits. Notably, in the two models, the relative orientations of the receptors are nearly the same while the relative orientations of the interacting pMHC molecules are distinct. By contrast, in Fig. 2C, the schematic diagram reveals differential docking orientation of receptors when the pMHC orientation is fixed. Given a single pMHCI transmembrane segment, the MHC ectodomain is anticipated to be mobile and hence adaptable so as to assume different orientations for its interaction with preTCR or αβTCR complexes.

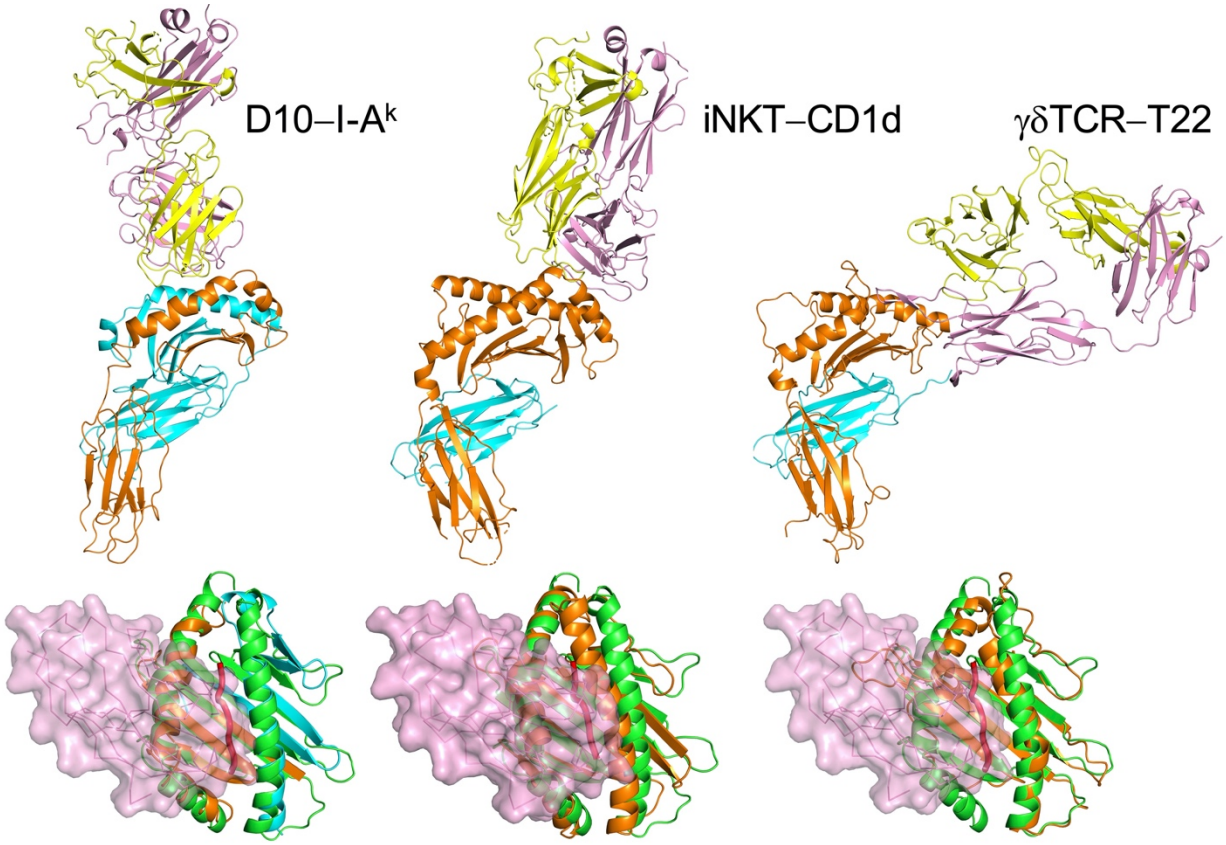

Fig. S7. PreTCR binding is restricted to a subset of MHC ligands. Three representative TCR-MHC complexes, D10–I-A^k, iNKT–CD1d and γδTCR–T22 (derived from PDB: 1FYT, 3TZV and 1YPZ, respectively) are shown in ribbon diagram from left to right. The upper row is a side view of the three complexes. The lower row is a top view of the binding groove region of various MHC molecules (orange ribbon) overlaid with K^b (green ribbon). Also overlaid is the N15β V-domain in the pink surface representation. The other color codes are: cyan for β2 microglobulin and I-A^k β-chain, orange for I-A^k α-chain and the heavy chain of CD1d and T22, yellow for TCR α- and γ-chains, and magenta for TCR β- and δ- chains. Red worm represents VSV8 peptide. In the lower row, left panel, a clear fit between MHC-I (K^b) and MHC-II (I-A^k) molecules is

observed, implying that the preTCR could bind I-A^k. In the middle panel, by contrast, the narrower groove of CD1d likely prevents the preTCR from binding. On the right panel, note that T22 lacks the N-terminal half of α2-helix necessary for preTCR binding.

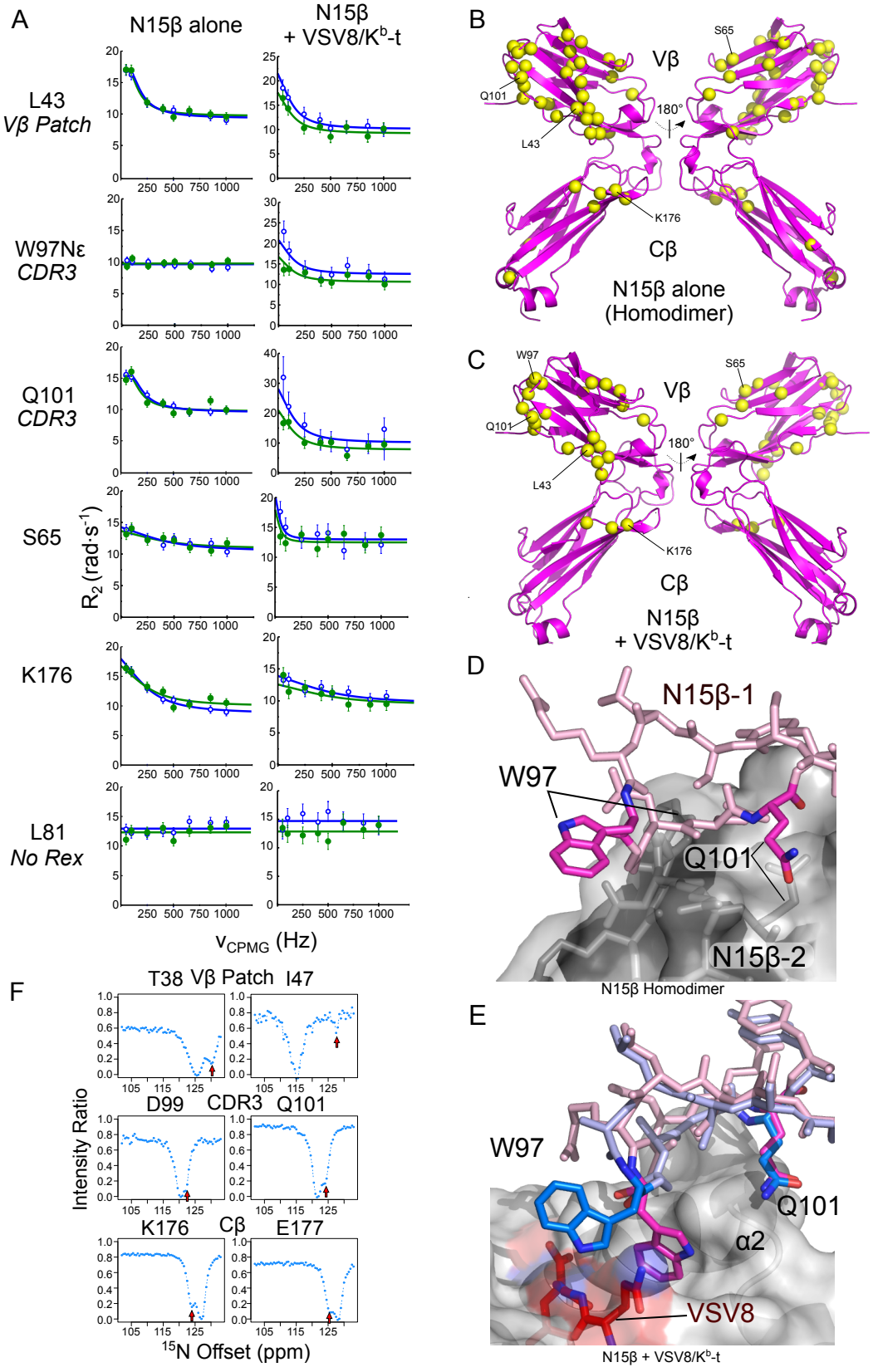
A
N15β alone
N15β
+ VSV8/Kb-t
L43
Vβ Patch
W97Nε
CDR3
Q101
CDR3
S65
K176
L81
No Rex
R2 (rad·s-1)
vCPMG (Hz)
B
Vβ
S65
Q101
180°
L43
K176
Cβ
N15β alone
(Homodimer)
C
W97
Vβ
S65
Q101
180°
L43
K176
Cβ
N15β
+ VSV8/Kb-t
D
N15β-1
W97
Q101
N15β-2
N15β Homodimer
E
W97
Q101
α2
VSV8
N15β + VSV8/Kb-t
F
T38 Vβ Patch I47
D99 CDR3 Q101
K176 Cβ E177
Intensity Ratio
15N Offset (ppm)

Fig. S8. Microsecond–millisecond dynamics confirm binding mode of N15β–VSV8-K^b-t in solution as detected by NMR. A) Representative ^{15}N relaxation dispersion profiles of residues in ligand binding regions (L43, W97, and Q101) show changes consistent with increased exchange parameters on VSV8-K^b-t addition depicted as VSV8/K^b-t herein. Blue and green points represent measured transverse relaxation (R_2) at a given Carr–Purcell–Meiboom–Gill (νCPMG) pulse frequency acquired at 700 and 600 MHz, respectively. Lines are interpolated curves. Samples were 400 μM $^2H^{13}C^{15}N$ N15β with or without 400 μM unlabeled VSV8-K^b-t. Notably, W97Nε does not exhibit μs-ms Rex effects unless ligand pMHC is added. Distal region residues show either decreased Rex effect with binding (S65, K176) or no Rex when either bound or unbound (L81). **B)** Ribbon model of N15β with yellow spheres shown for residues exhibiting ^{15}N R_{ex} effect as a consequence of homodimerization. **C)** Model of N15β with residues exhibiting exchange when in the presence of VSV8-K^b-t. The magnitude of the Rex effect for Q101 is greater when pMHC is added, consistent with a stronger interaction for N15β–VSV8-K^b (K_D=400 μM) versus N15β-N15β homodimerization (K_D>1 mM) and in line with previous results that pMHC binding rather than homodimerization occurred for the heterodimeric pTα–β preTCR (*16-18*). Residues in Cβ proximal to K176 participate in homodimerization and show a correspondingly weaker Rex effect with pMHC addition, consistent with an absent interaction with pMHC in the crystal structure. Analysis of dispersion data to determine the exchange rate (k_{ex}) of residues in the respective interfaces (see methods) finds a slower exchange rate for N15β–pMHC interaction (872 ± 130/s) versus N15β-N15β homodimerization (1180 ± 79/s). **D)** Homodimeric interface of N15β (*10*) highlighting CDR3 residues W97 and Q101. Subunit 1 (N15β) is shown in

stick representation and subunit 2 is shown as surface representation. W97 is not in that interface, whereas a Q101−Q101 contact is formed. **E)** N15β CDR3 (pink and light blue stick representation) residue W97 and Q101 (magenta and dark blue) interact with the surface of VSV8-K^b-t (gray surface with VSV8 in red) in two conformations observed in the current study. For clarity only N15β CDR3 is shown. **F)** Chemical exchange saturation transfer (CEST) measurements, an orthogonal method for detection of ms timescale dynamics of N15β alone, show chemical exchange at resonances distant from their ground state values for residues in CDR3, Vβ patch and the C domain confirming the observations with Rex. CEST plots show chemical exchange peaks in residues located within the V and C domains of N15β including Vβ patch residues T38 and I47, CDR3 D99 and Q101 and C domain K176 and E177. Spectra were collected at 800 mHz. Exchange peaks are highlighted with red arrows.

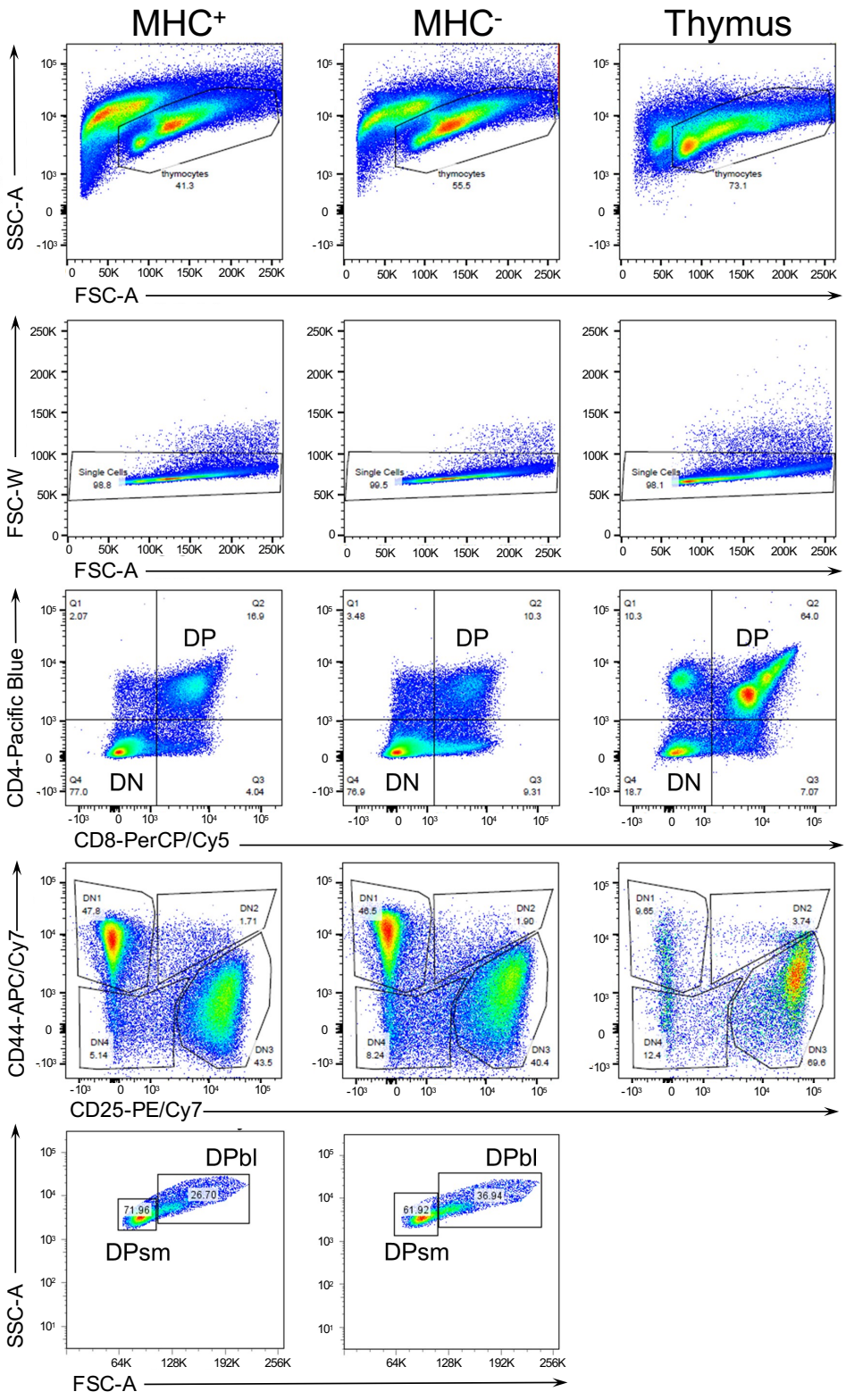

Fig. S9. Gating strategy for B6 thymocyte–thymic epithelial cultures.

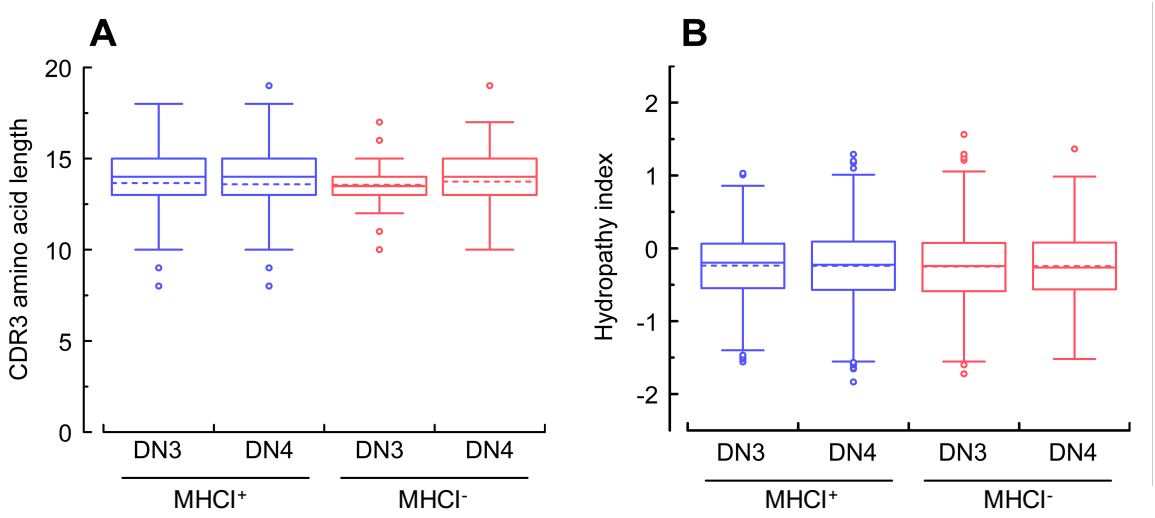

Fig. S10. Comparison of β-chain CDR3 features in DN3 and DN4 thymocytes developing on MHC^{+} and MHC^{-} epithelial stromal cultures. **A)**. CDR3 length. **B)**. CDR3 hydropathy index. Data for CDR3 hydropathy index calculated using the Kyte-Doolittle method (from attached table S7 representative of five experiments) (*51*) and CDR lengths are plotted as fenced box plots for thymocytes growing on either MHCI-competent stromal cells (468 DN3, and 2528 DN4, clonotypes, respectively) or on MHCI-deficient stromal cells (953 DN3, and 402 DN4, clonotypes, respectively). For each condition, the box bounds the 25th and 75th percentile with internal 50th percentile (solid line) and mean (dotted line). Whiskers with bars define upper and lower fenced values for each condition with outliers depicted by open circles.

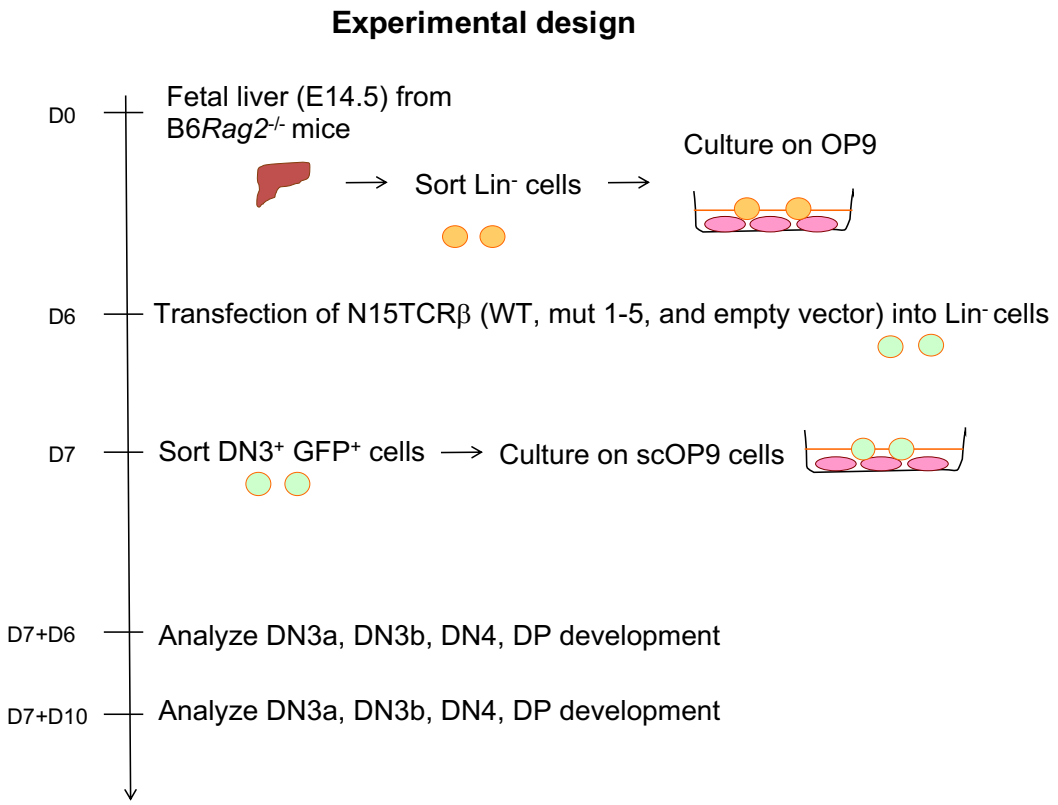

Fig. S11. Experimental protocol and analysis scheme of B6 *Rag2*$^{-/-}$ thymocytes transduced with wt and variant N15β chains representing crystallographically defined interface contact residue mutations. Development was assessed after in vitro culture at indicated times on OP9-DL4 epithelia expressing a single pMHC, VSV8-K^{b} (scOP9)

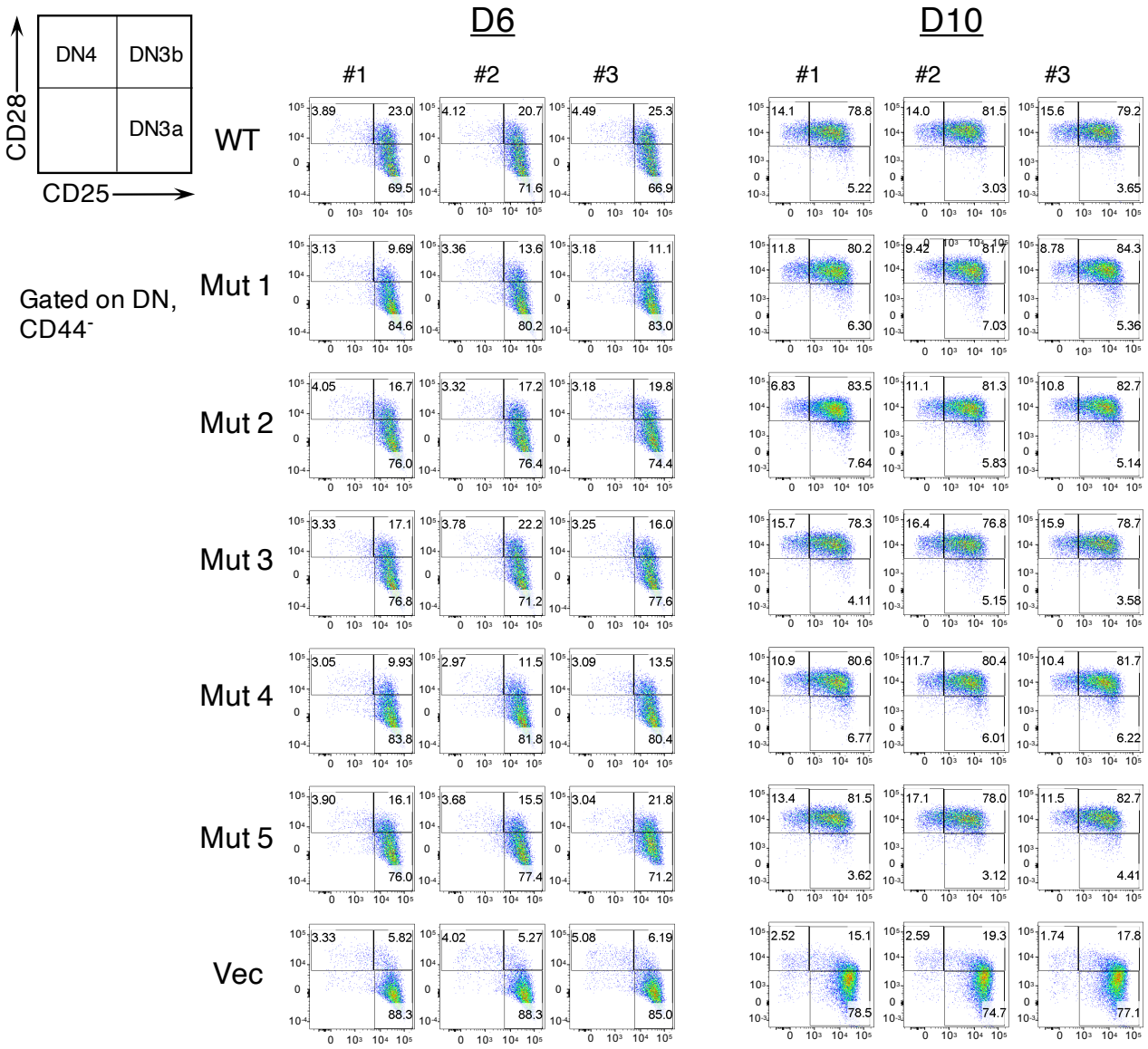

Fig. S12. Gating strategy for DN thymocytes for fig. S11 development. Refer to Fig. 4 for location of mutations in the N15β protein sequence.

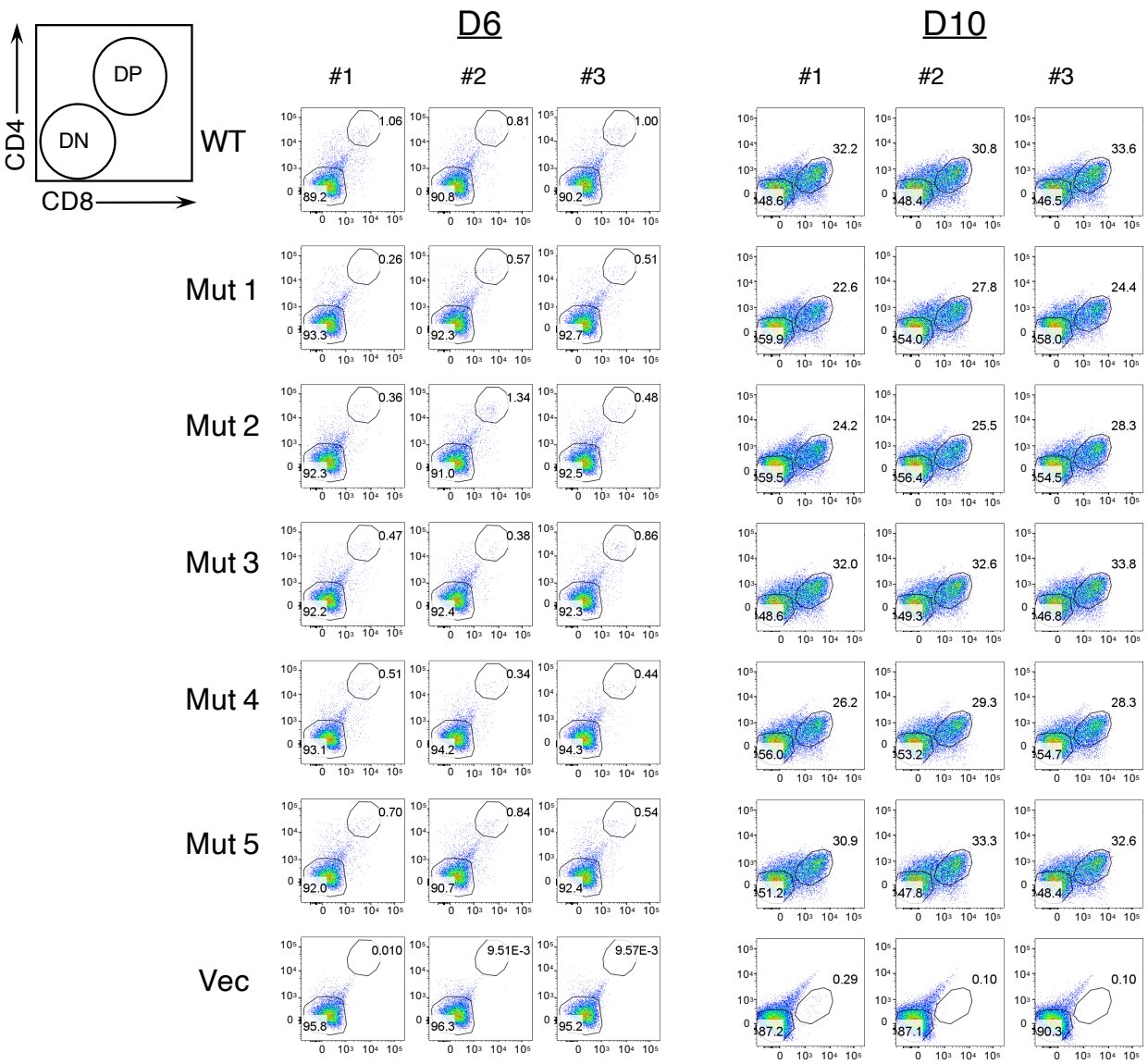

Fig. S13. Gating strategy for DP thymocytes for fig. S11 development.

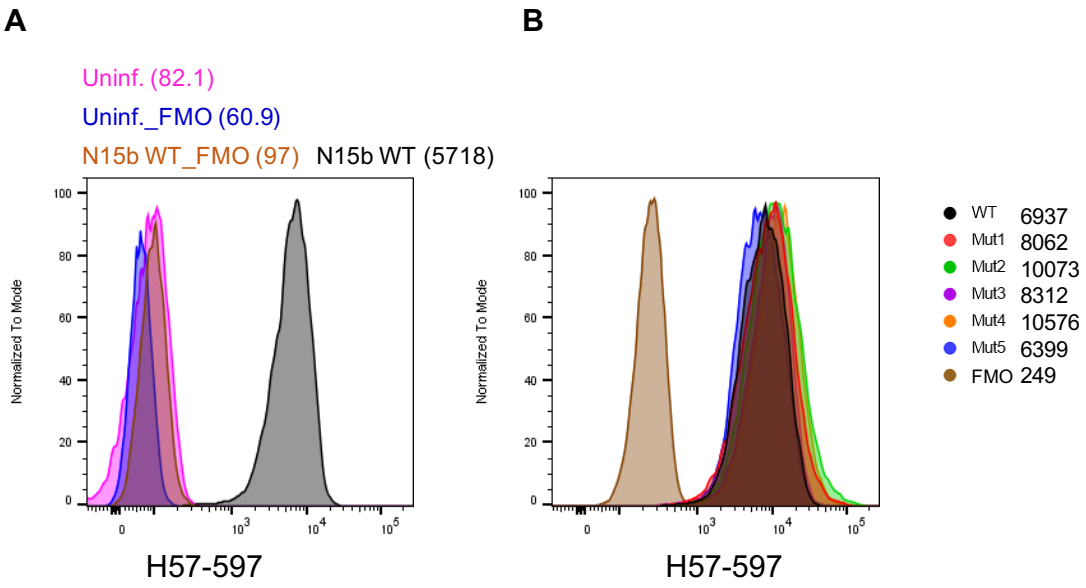

Fig. S14. N15β WT and mutants express comparable levels of TCRβ after transfection. A) TCRβ expression on SCID.adh cells transfected N15β WT (black) or untransfected (magenta) is shown. Fluorescence minus one (FMO) is used for negative control. The parenthesis represents geometric mean fluorescence intensity (gMFI). **B)** TCRβ expression on SCID.adh cells transfected N15β WT or mutants 1-5 is shown.

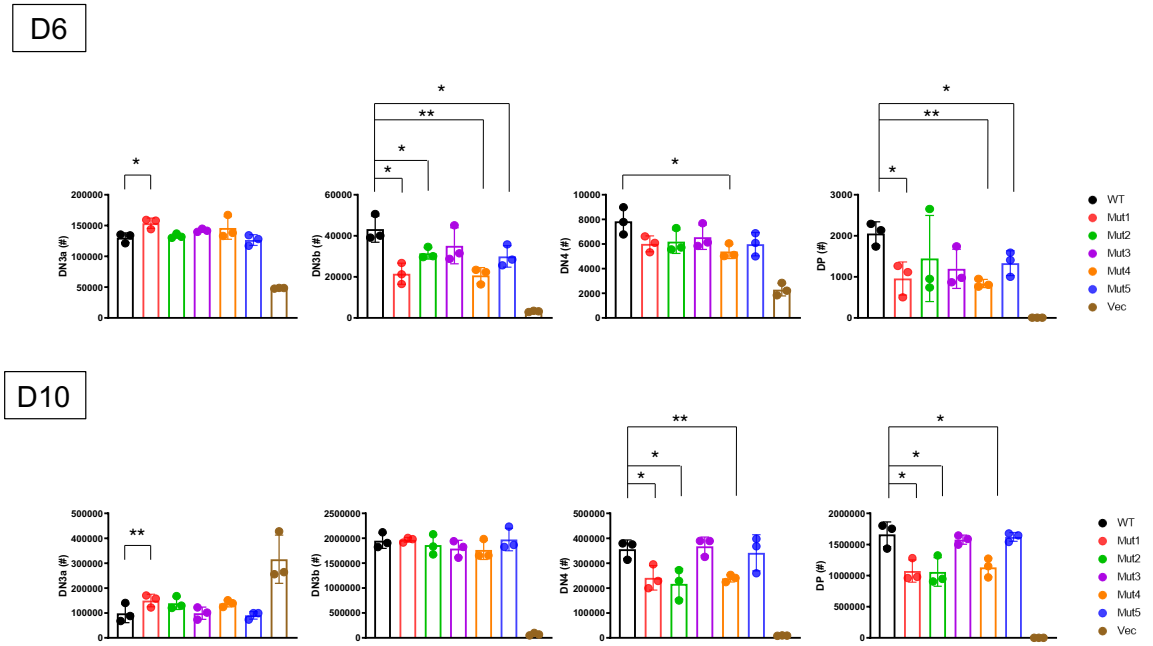

Fig. S15. Display of absolute cell numbers corresponding to cell percentages displayed in Fig. 4E. *: $P<0.05$ and **: $P<0.01$ by 2-tailed Student's *t*-test.

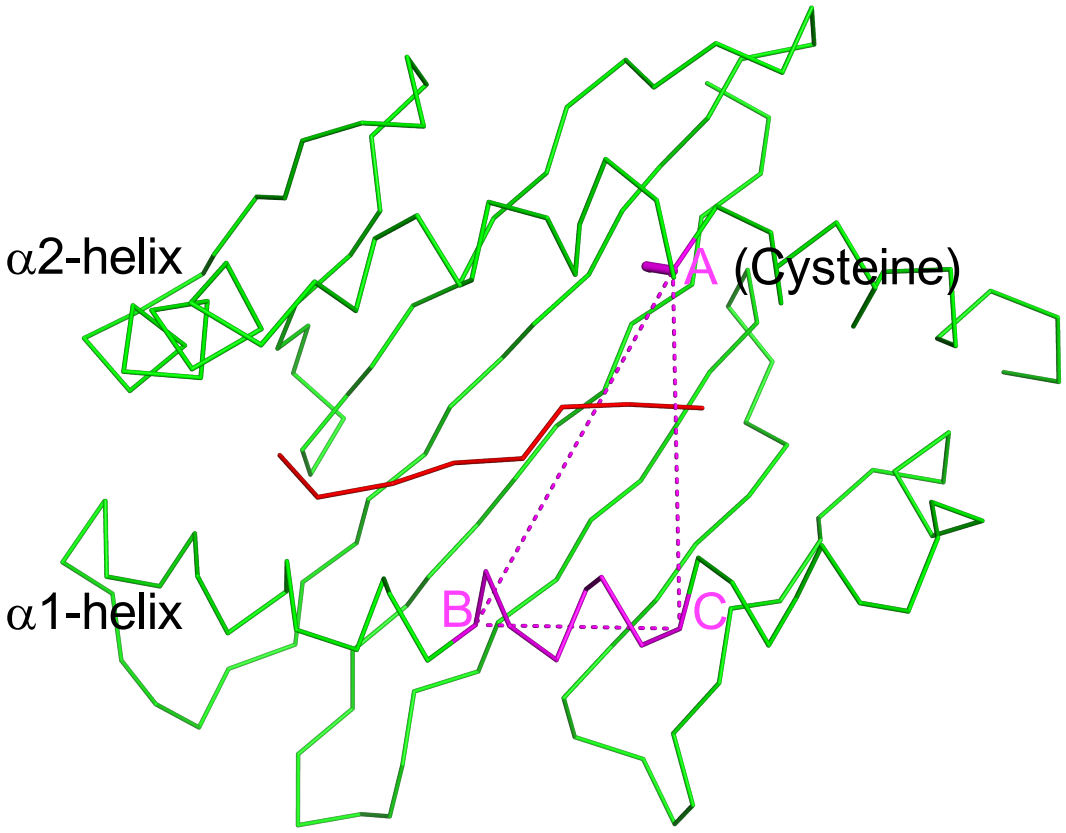

Fig. S16. Definition of the interhelical width of MHC and MHC-fold proteins. An MHC-I binding groove is drawn in green skeleton with two helices, α1 and α2 labeled. The red skeleton between the two helices represents a peptide bound to the MHC. Position A indicates the Cα position of a cysteine (Cys164 in K^b) conserved in MHC-I, MHC-II, non-classic MHC and MHC like proteins within α2 helix. Position B (Glu71 in K^b) and C (Thr64 in K^b) are the Cα locations of the two residues at the bottom of α1 helix. The helical region between B and C is structurally well aligned in all MHC-fold proteins. The dashed lines AB, AC, and BC link corresponding Cα atoms of these residues. The vertical distance from the position A to the line BC is designated as the interhelical width.

Table S1. Data collection, refinement and structural determination

	N15β–K^b-t2-G56 (6WL2)	**N15β–K^b-t2-L4 (6WL3)**	**N15β–K^b-t2-R62 (6WL4)**	**Q4H7-K^b (7JI2)**
Data collection				
Space group	P 4_3 2_1 2	P 4_3 2_1 2	P 4_3 2_1 2	P 1 2_1 1
Cell dimensions				
a,b,c (Å)	108.9 108.9 316.2	109.4 109.4 316.5	110.6 110.6 316.9	66.4 90.0 89.7
α,β,γ (°)	90 90 90	90 90 90	90 90 90	90 112 90
Resolution range (Å)	50.00 - 3.30	50.00 - 3.45	50.00 - 3.60	50.00 - 1.95
	(3.36-3.30)	(3.52-3.45)	(3.71-3.60)	(1.98-1.95)
No. unique reflections	29728	26388	23833	71792
Rpim (%)	5.9 (58.3)	7.0 (38.3)	9.9 (50.8)	4.0 (25.7)
R-merge (%)	8.4 (112)	11.5 (100)	14.4 (119)	9.0 (55.1)
I/σ(I)	12.3 (1.1)	9.3 (1.5)	7.9 (1.5)	19.5 (2.4)
Completeness (%)	99.9 (100)	98.6 (99.6)	99.1 (99.5)	98.1 (98.4)
Redundancy	3.9 (3.8)	10.1 (9.6)	8.7 (8.6)	6.9 (7.0)
Refinement				
Resolution range (Å)	43.49 - 3.30	48.94 - 3.45	49.44 - 3.60	45.01 - 1.95
No. unique reflections	29536 (2868)	25896 (2496)	23483 (2311)	70095 (6978)
Rwork/ Rfree (%)	22.8/27.0	22.0/25.6	22.6/26.9	16.3/19.7
No. non-hydrogen atoms	10279	10292	10281	6855
B-factor				
Protein	112	97	107	31
R.m.s. deviations				
Bond lengths (Å)	0.003	0.003	0.002	0.009
Bond angles (°)	0.58	0.57	0.55	0.78
Ramachandran favored (%)	94.87	95.44	94.56	96.83
Ramachandran allowed (%)	4.97	4.40	5.12	2.64
Ramachandran outliers (%)	0.16	0.16	0.32	0.53

Values in parentheses are for the highest resolution shell.

Table S2. Buried surface area (BSA) calculation for all N15β contacted interfaces in the crystal packing.

	Intra-packing		Inter-packing	
Chain (Molecule)	**Contacted chain**	**BSA (Å²)**	**Contacted chain**	**BSA (Å²)**
A (K^b-t2)	**C**	1389	**F′**	1084
	I	504	**C′**	765
Chain C (N15β)	**D**	832	**G′**	1205
	F	171	**I′**	931
Chain D (K^b-t2)	**F**	1370	**Null**	N/A
	I	1412	**Null**	N/A
Chain F (N15β)	**G**	1257	**C′**	311
Chain G (K^b-t2)	**I**	1406	**Null**	N/A
Chain I (N15β)	**Null**	N/A	**G′**	105

Note: physiological interfaces observed in the crystal packing were highlighted in green. Other non-biological interfaces with BSA larger than 900 Å^2 resulted from crystal packing artifact were highlighted in orange. fig. S3 addressed the functional irrelevance of those non-biological interfaces. Here, A, C, D, F, G, and I represent chains within one asymmetric unit. A′, C′, F′, G′, and I′ represent corresponding chains from symmetry-related neighboring molecules.

Table S3. Least-squares superpositions of N15β-V–K^b-t2

Alignment RMSD	Complex ABC (Conformer I)	Complex DEF (Conformer II)	Complex GHI (Conformer II)
N15β–K^b-t2-G56		0.95 Å	0.78 Å
N15β–K^b-t2-R62	0.50 Å	1.07 Å	0.97 Å
N15β–K^b-t2-L4	0.49 Å	0.88 Å	0.69 Å

Note: Each of the crystal structures listed in this table consists of three crystallographically independent N15β–K^b-t2 complexes. Each complex contains three chains. For example, complex ABC contains chains A, B and C, representing K^b-t2, peptide and N15β, respectively. In order to evaluate the variation of the N15β docking onto K^b-t2, the N15β V-domain and K^b-t2 components (N15β-V–K^b-t2) from the Complex ABC of N15β–K^b-t2-G56 structure were used as a reference. The backbone Cα atoms of the N15β-V–K^b-t2 parts were superimposed to the reference using least-squares method within the program Coot (*36*). The exclusion of the N15β C-domain in the structural alignment is to exclude the potential variation between N15β V and C domains. To perform the superposition using the program (*36*), the chain ID and the residue numbering of N15β in each complex was manipulated so that N15β-V and K^b-t2 appear to be two domains of one molecule. This calculation demonstrates that the interacting pairings between N15β Vβ module and K^b are the same among all nine complexes in three crystals, suggestive of a physiological binding.

Table S4. Buried surface area (BSA) and shape complementarity (Sc) Analysis

Protein (PDB)	Complex	BSA (Å^2)	Sc
N15β–K^b-t2-G56 (6WL2)	ABC	1389	0.53
	DEF	1370	0.65
	GHI	1406	0.59
N15β–K^b-t2-R62 (6WL4)	ABC	1351	0.49
	DEF	1495	0.65
	GHI	1302	0.52
N15β–K^b-t2-L4 (6WL3)	ABC	1352	0.50
	DEF	1399	0.64
	GHI	1343	0.50
Average	Mean (RMSD)	1379 (54.1)	0.56 (0.07)
2C–DEV8-K^b (2CKB)	TCRβ–DEV8-K^b	916	0.44
2C TCR (2CKB)	Vα/Vβ	1525	0.55

To validate the topological fitting between N15β and K^b-t2, both BSA and Sc for each N15β–K^b-t2 complex in each crystal structure were calculated by using the PISA and Sc in CCP4 *(28)*. In the table, calculated data from three crystal structures: N15β–K^b-t2-G56, N15β–K^b-t2-R62, and N15β–K^b-t2-L4 were summarized together, and each crystal structure contains 3 N15β–K^b-t2 complex: ABC, DEF, and GHI as described. For the calculation, each peptide was combined with its corresponding heavy chain of K^b as one ensemble. The BSA and Sc of the interface between TCRαβ and pMHC as well as the interface between Vα and Vβ of a TCRαβ–pMHC complex structure, 2C-K^b (PDB, 2CKB), were also calculated for comparison.

Table S5. Detail contact information of N15β–VSV8-K^b-t2

VSV8 or K^b-t2	N15β	Contacts number in each asymmetric unit			Total
		N15β–K^b-t2-G56 (6WL2)	N15β–K^b-t2-R62 (6WL4)	N15β–K^b-t2-L4 (6WL3)	
K131	E42	1			1
K146	W97	2	2	1	5
	G98	3	3	3	9
	D99			1	1
W147	W97	1	1	1	3
Q149	G98		1		1
	D99	1	1		2
	E100	3	3	3	9
	Q101	3	3	3	9
A150	Y35	2	1	1	4
	G98	3	2	3	8
	D99	2	1	2	5
	Q101	3	3	3	9
G151	Y35		1	2	3
	Q101	2	2	2	6
E152	W97	1	1	1	3
E154	Y35	1		2	3
	L43	3	2		5
	F45	3	3	3	9
R155	F45	3	3	3	9
	Q48	3	1	2	6
	D56	3	3	2	8
R157	E42	3	3	3	9
A158	F45	3	2	2	7
	F59	3	2	3	8
E161	E42		1		1
	F59	2	2	1	5
G162	F59		2	2	4
T163	K57	1	1	1	3
V/L4 (VSV8)	R55	1	2	3	6
	D56		1		1
Q6 (VSV8)	Q48	3	3	3	9
	R55	2	2	2	6
	W97	3	3	3	9
	G98			1	1
G7 (VSV8)	W97	2	2	2	6

Note: 1) contacts within 4 Å existing in 7-9 complexes are highlighted in yellow; 2) CDR1 loop encompasses residues P25–S30 and CDR2 loop contains Y50–V53. They do not contact K^b-t2.

Table S6. Inter-helical widths of MHC and MHC fold proteins.

Classification	Protein (PDB code)	Width (Å)
MHC-I	HLA-A2 (1DUZ)	17.2
	HLA-B5701 (5VUE)	17.4
	HLA-C0801 (4NT6)	17.0
	H-2K^{b} (1LEG)	17.2
	H-2K^{d} (1VGK)	16.7
	H-2D^{d} (5T7G)	17.2
	Mean (RMSD)	**17.1 (0.25)**
MHC-II	HLA-DQ (6DIG)	17.8
	HLA-DR (6HBY)	18.4
	HLA-DP (3LQZ)	18.4
	I-A^{d} (1IAO)	18.0
	I-A^{b} (6MNN)	18.3
	I-A^{u} (1K2D)	18.2
	I-A^{k} (1D9K)	17.9
	Mean (RMSD)	**18.2 (0.25)**
Non-classical MHCs, MHC related and MHC like Proteins	HLA-E (3BZE)	16.6
	HLA-G (1YDP)	17.1
	Qa1 (3VJ6)	17.3
	MR1 (4GUP)	18.1
	TL (1NEZ)	14.4
	T22 (1YPZ)	13.3
	M10 (1ZS8)	17.0
	ZAG (1T7V)	17.0
	HFE (1A6Z)	14.3
	FcRn (6C97)	13.5
	CD1d (1ZT4)	13.5
	ULBP3 (1KCG)	11.2
	MICA (1HYR)	14.4
	MICB (1JE6)	13.4
	UL18 (3D2U)	17.0

Note: Three human MHC-I and three mouse MHC-I structures were selected from Protein Data Bank to calculate the interhelical width of class I MHC (defined in **fig. S16**).

The average width for the six MHC is 17.1 Å (RMSD 0.25). Similarly, three human MHC-II and three mouse MHC-II structures were selected to calculate the interhelical width of class II MHC. The average width of the six MHCs is 18.2 Å (RMSD 0.25). Finally, interhelical widths of 15 individual non-classical MHC, MHC-related, and MHC-like proteins were calculated for comparison with classic MHCs. Two distinct subgroups based on the interhelical width seemingly appear among the 15 sampled molecules.

Table S7 (uploaded as abe0918_Suppl.Excel_seq1_v1.xlsx)

This MS Excel file contains TCRβ chain repertoire results for a representative experiment included in Fig. 4A-C. The file is organized into 5 tabbed sheets labeled:

1. WT-DN3
2. WT-DN4
3. H-2neg-DN3
4. H-2neg-DN4
5. Statistics

where "WT" indicates TCRβ chain repertoire of cells developing on stroma expressing MHCI, and "H-2neg" indicates that the cells were developing on MHCI-negative stroma. "DN3" indicates that the cells were phenotypically sorted by FACS as being DN3 ($CD44^-CD25^+$) and "DN4" indicates that the cells were sorted as being DN4 ($CD44^-CD25^-$).

For each of the first 4 sheets, the data are derived from targeted TCRβ sequencing of RNA isolated from 10,000 cells as described in Ref. 49 and analyzed and processed as described in Refs. 50 and 51.

Within each sheet, the data are organized by clonotype read counts (column B). Non-productive reads (e.g. during recombination of V, D and J region, internal stop codons, out-of-frame sequence reads, or incomplete recombination events occurred) were excluded.

Productive sequence (column C) was then analyzed by the MiXCR suite (50) to identify V, D, and J regions (columns D, E, F). Note that due to targeted sequence within the β chain constant region, the software could not discriminate between Trbc1 and Trbc2 and thus all were dominantly characterized as Trbc2 (column G). For columns D-G, the

numbers in parentheses are a relative measure of the confidence of identification where numbers >= 100 may be considered as “reliable” and numbers <100 as “likely”. The translated CDR3 region is depicted in column H. Columns I and J list CDR3 length, and hydropathy by Kyte-Doolittle analysis (52), respectively.

Sheet 5, “Statistics” provides parametric statistical analysis of the CDR3 and hydropathy values listed in the preceding sheets and were used to generate fig. S10.

Table S8. Monoclonal Antibodies Used in the Current Studies.

Antibody	Clone	Label	Manifacture	Cat #	RRID	Final concentration
anti-mouse CD4	RM4-5	Pacific Blue	BioLegend	100531	AB_493374	5ug/ml
anti-mouse CD8b.2	53-5.8	PE	BioLegend	140408	AB_10644002	2ug/ml
anti-mouse Ly-6A/E (Sca-1)	D7	FITC	eBioscience	11-5981-81	AB_465332	5ug/ml
anti-mouse CD117	2B8	APC	BD Biosciences	553356	AB_398536	2ug/ml
anti-mouse CD45	30-F11	APC	BioLegend	103112	AB_312977	2ug/ml
anti-mouse CD25	PC61.5	PE-Cyanine7	eBioscience	25-0251-82	AB_468733	1ug/ml
anti-mouse CD44	IM7	APC-Cy7	BD Biosciences	560568	AB_1727481	1ug/ml
anti-mouse CD4	RM4-5	Brilliant Violet 711	BioLegend	100549	AB_11219396	2ug/ml
anti-mouse CD8a	53-6.7	PerCP/Cyanine5.5	BioLegend	100734	AB_2075238	2ug/ml
anti-mouse CD28	E18	PE	BioLegend	122010	AB_604078	2ug/ml
anti-mouse TCR β chain	H57-597	PE	BioLegend	109207	AB_313430	2ug/ml

References and Notes

1. K. Shortman, M. Egerton, G. J. Spangrude, R. Scollay, The generation and fate of thymocytes. *Semin. Immunol.* **2**, 3–12 (1990). Medline

2. U. Koch, E. Fiorini, R. Benedito, V. Besseyrias, K. Schuster-Gossler, M. Pierres, N. R. Manley, A. Duarte, H. R. Macdonald, F. Radtke, Delta-like 4 is the essential, nonredundant ligand for Notch1 during thymic T cell lineage commitment. *J. Exp. Med.* **205**, 2515–2523 (2008). doi:10.1084/jem.20080829 Medline

3. H. R. Rodewald, M. Ogawa, C. Haller, C. Waskow, J. P. DiSanto, Pro-thymocyte expansion by c-kit and the common cytokine receptor gamma chain is essential for repertoire formation. *Immunity* **6**, 265–272 (1997). doi:10.1016/S1074-7613(00)80329-5 Medline

4. T. Kreslavsky, M. Gleimer, M. Miyazaki, Y. Choi, E. Gagnon, C. Murre, P. Sicinski, H. von Boehmer, β-Selection-induced proliferation is required for αβ T cell differentiation. *Immunity* **37**, 840–853 (2012). doi:10.1016/j.immuni.2012.08.020 Medline

5. H. von Boehmer, The thymus in immunity and in malignancy. *Cancer Immunol. Res.* **2**, 592–597 (2014). doi:10.1158/2326-6066.CIR-14-0070 Medline

6. C. Saint-Ruf, K. Ungewiss, M. Groettrup, L. Bruno, H. J. Fehling, H. von Boehmer, Analysis and expression of a cloned pre-T cell receptor gene. *Science* **266**, 1208–1212 (1994). doi:10.1126/science.7973703 Medline

7. B. H. Koller, P. Marrack, J. W. Kappler, O. Smithies, Normal development of mice deficient in beta 2M, MHC class I proteins, and CD8+ T cells. *Science* **248**, 1227–1230 (1990). doi:10.1126/science.2112266 Medline

8. M. J. Grusby, H. Auchincloss Jr., R. Lee, R. S. Johnson, J. P. Spencer, M. Zijlstra, R. Jaenisch, V. E. Papaioannou, L. H. Glimcher, Mice lacking major histocompatibility complex class I and class II molecules. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3913–3917 (1993). doi:10.1073/pnas.90.9.3913 Medline

9. H. J. Fehling, A. Krotkova, C. Saint-Ruf, H. von Boehmer, Crucial role of the pre-T-cell receptor alpha gene in development of alpha beta but not gamma delta T cells. *Nature* **375**, 795–798 (1995). doi:10.1038/375795a0 Medline

10. B. A. Irving, F. W. Alt, N. Killeen, Thymocyte development in the absence of pre-T cell receptor extracellular immunoglobulin domains. *Science* **280**, 905–908 (1998). doi:10.1126/science.280.5365.905 Medline

11. S. S. Pang, R. Berry, Z. Chen, L. Kjer-Nielsen, M. A. Perugini, G. F. King, C. Wang, S. H. Chew, N. L. La Gruta, N. K. Williams, T. Beddoe, T. Tiganis, N. P. Cowieson, D. I. Godfrey, A. W. Purcell, M. C. Wilce, J. McCluskey, J. Rossjohn, The structural basis for autonomous dimerization of the pre-T-cell antigen receptor. *Nature* **467**, 844–848 (2010). doi:10.1038/nature09448 Medline

12. S. Yamasaki, E. Ishikawa, M. Sakuma, K. Ogata, K. Sakata-Sogawa, M. Hiroshima, D. L. Wiest, M. Tokunaga, T. Saito, Mechanistic basis of pre-T cell receptor-mediated autonomous signaling critical for thymocyte development. *Nat. Immunol.* **7**, 67–75 (2006). doi:10.1038/ni1290 Medline

13. P. Smelty, C. Marchal, R. Renard, L. Sinzelle, N. Pollet, D. Dunon, T. Jaffredo, J.-Y. Sire, J. S. Fellah, Identification of the pre-T-cell receptor alpha chain in nonmammalian vertebrates challenges the structure-function of the molecule. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19991–19996 (2010). doi:10.1073/pnas.1010166107 Medline

14. B. Zhou, Q. Chen, R. J. Mallis, H. Zhang, J. H. Liu, E. L. Reinherz, J. H. Wang, A conserved hydrophobic patch on Vβ domains revealed by TCRβ chain crystal structures: Implications for pre-TCR dimerization. *Front. Immunol.* **2**, 5 (2011). doi:10.3389/fimmu.2011.00005 Medline

15. A. M. Michie, A. C. Chan, M. Ciofani, M. Carleton, J. M. Lefebvre, Y. He, D. M. Allman, D. L. Wiest, J. C. Zúñiga-Pflücker, D. J. Izon, Constitutive Notch signalling promotes CD4 CD8 thymocyte differentiation in the absence of the pre-TCR complex, by mimicking pre-TCR signals. *Int. Immunol.* **19**, 1421–1430 (2007). doi:10.1093/intimm/dxm113 Medline

16. D. K. Das, R. J. Mallis, J. S. Duke-Cohan, R. E. Hussey, P. W. Tetteh, M. Hilton, G. Wagner, M. J. Lang, E. L. Reinherz, Pre-T cell receptors (Pre-TCRs) leverage Vβ complementarity determining regions (CDRs) and hydrophobic patch in mechanosensing thymic self-ligands. *J. Biol. Chem.* **291**, 25292–25305 (2016). doi:10.1074/jbc.M116.752865 Medline

17. R. J. Mallis, H. Arthanari, M. J. Lang, E. L. Reinherz, G. Wagner, NMR-directed design of pre-TCRβ and pMHC molecules implies a distinct geometry for pre-TCR relative to αβTCR recognition of pMHC. *J. Biol. Chem.* **293**, 754–766 (2018). doi:10.1074/jbc.M117.813493 Medline

18. R. J. Mallis, K. Bai, H. Arthanari, R. E. Hussey, M. Handley, Z. Li, L. Chingozha, J. S. Duke-Cohan, H. Lu, J. H. Wang, C. Zhu, G. Wagner, E. L. Reinherz, Pre-TCR ligand binding impacts thymocyte development before αβTCR expression. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 8373–8378 (2015). doi:10.1073/pnas.1504971112 Medline

19. R. J. Mallis, K. N. Brazin, J. S. Duke-Cohan, W. Hwang, J. H. Wang, G. Wagner, H. Arthanari, M. J. Lang, E. L. Reinherz, NMR: An essential structural tool for integrative studies of T cell development, pMHC ligand recognition and TCR mechanobiology. *J. Biomol. NMR* **73**, 319–332 (2019). doi:10.1007/s10858-019-00234-8 Medline

20. X. Yang, R. A. Mariuzza, Pre-T-cell receptor binds MHC: Implications for thymocyte signaling and selection. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 8166–8167 (2015). doi:10.1073/pnas.1510127112 Medline

21. C. López-Rodríguez, J. Aramburu, R. Berga-Bolaños, Transcription factors and target genes of pre-TCR signaling. *Cell. Mol. Life Sci.* **72**, 2305–2321 (2015). doi:10.1007/s00018-015-1864-8 Medline

22. 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). doi:10.1038/nri3667 Medline

23. M. G. Rudolph, R. L. Stanfield, I. A. Wilson, How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* **24**, 419–466 (2006). doi:10.1146/annurev.immunol.23.021704.115658 Medline

24. J. H. Wang, E. L. Reinherz, The structural basis of αβ T-lineage immune recognition: TCR docking topologies, mechanotransduction, and co-receptor function. *Immunol. Rev.* **250**, 102–119 (2012). doi:10.1111/j.1600-065X.2012.01161.x Medline

25. L. Lo Conte, C. Chothia, J. Janin, The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198 (1999). doi:10.1006/jmbi.1998.2439 Medline

26. C. Zhang, A. Anderson, C. DeLisi, Structural principles that govern the peptide-binding motifs of class I MHC molecules. *J. Mol. Biol.* **281**, 929–947 (1998). doi:10.1006/jmbi.1998.1982 Medline

27. H. Hosokawa, E. V. Rothenberg, Cytokines, transcription factors, and the initiation of T-cell development. *Cold Spring Harb. Perspect. Biol.* **10**, a028621 (2018). doi:10.1101/cshperspect.a028621 Medline

28. F. Vasseur, A. Le Campion, C. Pénit, Scheduled kinetics of cell proliferation and phenotypic changes during immature thymocyte generation. *Eur. J. Immunol.* **31**, 3038–3047 (2001). doi:10.1002/1521-4141(2001010)31:10<3038:AID-IMMU3038>3.0.CO;2-3 Medline

29. D. K. Das, Y. Feng, R. J. Mallis, X. Li, D. B. Keskin, R. E. Hussey, S. K. Brady, J. H. Wang, G. Wagner, E. L. Reinherz, M. J. Lang, Force-dependent transition in the T-cell receptor β-subunit allosterically regulates peptide discrimination and pMHC bond lifetime. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 1517–1522 (2015). doi:10.1073/pnas.1424829112 Medline

30. Y. Feng, E. L. Reinherz, M. J. Lang, αβ T cell receptor mechanosensing forces out serial engagement. *Trends Immunol.* **39**, 596–609 (2018). doi:10.1016/j.it.2018.05.005 Medline

31. M. K. Teng, A. Smolyar, A. G. Tse, J. H. Liu, J. Liu, R. E. Hussey, S. G. Nathenson, H. C. Chang, E. L. Reinherz, J. H. Wang, Identification of a common docking topology with substantial variation among different TCR-peptide-MHC complexes. *Curr. Biol.* **8**, 409–414 (1998). doi:10.1016/S0960-9822(98)70160-5 Medline

32. W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, HKL-3000: The integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859–866 (2006). doi:10.1107/S0907444906019949 Medline

33. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011). doi:10.1107/S0907444910045749 Medline

34. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007). doi:10.1107/S0021889807021206 Medline

35. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010). doi:10.1107/S0907444909052925 Medline

36. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). doi:10.1107/S0907444910007493 Medline

37. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012). doi:10.1107/S0907444912001308 Medline

38. M. C. Lawrence, P. M. Colman, Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950 (1993). doi:10.1006/jmbi.1993.1648 Medline

39. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995). doi:10.1007/BF00197809 Medline

40. S. G. Hyberts, A. G. Milbradt, A. B. Wagner, H. Arthanari, G. Wagner, Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J. Biomol. NMR* **52**, 315–327 (2012). doi:10.1007/s10858-012-9611-z Medline

41. R. Keller, *The Computer Aided Resonance Assignment* (Cantina, 2004).

42. J. P. Loria, M. Rance, A. G. Palmer 3rd, A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J. Biomol. NMR* **15**, 151–155 (1999). doi:10.1023/A:1008355631073 Medline

43. J. P. Loria, M. Rance, A. G. Palmer, A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332 (1999). doi:10.1021/ja983961a

44. M. Bieri, E. J. d'Auvergne, P. R. Gooley, relaxGUI: A new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and μs motion of proteins. *J. Biomol. NMR* **50**, 147–155 (2011). doi:10.1007/s10858-011-9509-1 Medline

45. S. Morin, T. E. Linnet, M. Lescanne, P. Schanda, G. S. Thompson, M. Tollinger, K. Teilum, S. Gagné, D. Marion, C. Griesinger, M. Blackledge, E. J. d'Auvergne, relax: The analysis of biomolecular kinetics and thermodynamics using NMR relaxation dispersion data. *Bioinformatics* **30**, 2219–2220 (2014). doi:10.1093/bioinformatics/btu166 Medline

46. Z. Luz, S. Meiboom, Nuclear magnetic resonance study of the protolysis of trimethylammonium ion in aqueous solution-order of the reaction with respect to solvent. *J. Chem. Phys.* **39**, 366–370 (1963). doi:10.1063/1.1734254

47. D. Long, G. Bouvignies, L. E. Kay, Measuring hydrogen exchange rates in invisible protein excited states. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8820–8825 (2014). doi:10.1073/pnas.1405011111 Medline

48. I. Z. Mamedov, O. V. Britanova, I. V. Zvyagin, M. A. Turchaninova, D. A. Bolotin, E. V. Putintseva, Y. B. Lebedev, D. M. Chudakov, Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling. *Front. Immunol.* **4**, 456 (2013). doi:10.3389/fimmu.2013.00456 Medline

49. D. A. Bolotin, S. Poslavsky, I. Mitrophanov, M. Shugay, I. Z. Mamedov, E. V. Putintseva, D. M. Chudakov, MiXCR: Software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015). doi:10.1038/nmeth.3364 Medline

50. M. Shugay, D. V. Bagaev, M. A. Turchaninova, D. A. Bolotin, O. V. Britanova, E. V. Putintseva, M. V. Pogorelyy, V. I. Nazarov, I. V. Zvyagin, V. I. Kirgizova, K. I. Kirgizov, E. V. Skorobogatova, D. M. Chudakov, VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLOS Comput. Biol.* **11**, e1004503 (2015). doi:10.1371/journal.pcbi.1004503 Medline

51. J. Kyte, R. F. Doolittle, A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982). doi:10.1016/0022-2836(82)90515-0 Medline

52. S. Fuchs, "GRAVY calculator" (2020); http://www.gravy-calculator.de/.