Supplementary methods.

Mice. Rag1 deficient mice\(^{33}\), OT-I as well as AND TCR transgenic mice\(^{34,35}\) have been described. To generate a conditional Rad21 allele, *LoxP* sites were inserted into introns 4 and 6 to allow Cre-mediated excision of exons 5 and 6. Correctly targeted clones were detected by Southern blotting of *BsmI* and *PacI* digested genomic DNA to confirm the presence of 5’ and 3’ *LoxP* sites, respectively. A *puro\(\Delta\)tk* cassette flanked by *FRT* sites was excised to generate the conditional allele (supplementary fig. 1a). For genomic deletion analysis DNA was extracted\(^{39}\) and 50ng were used per real time PCR reaction. Primers were designed to detect the intact and not the deleted *Rad21* allele. Data were normalised to the geometric mean of two unrelated genomic sites on chromosome 2 and 3, respectively.

**Rad21lox**
forward CATGGTTGCGAGATGAGCAAC
reverse CTA CTT TCC CGC TAG CAA CTG

**tnah chr2**
forward CCCGGGCTAATTCTCTATGTC
reverse GCTGTAAGTCGTCGCTCGTG

**tnah chr3**
forward TGTGCCAGCATCTTTTGCC
reverse GCGTGCTGGAATTAAAAGCC

Mice were on a mixed C57BL/6 and 129 background and matched for age and *Tcra* haplotypes in all experiments involving *Tcra* expression or rearrangement. *Tcra* haplotypes were determined by genomic PCR.

**Primer 1:** GAGGAAAAATGGCCCGGTAG, **primer R:** TTCAGTAGTCCCTCTCCACG. PCR amplification 94°C for 5 min, 5 cycles of 94°C for 30 s, 68°C for 30 s, 72°C for 1 min, 5 cycles of 94°C for 30 s, 65°C for 30 s, 72°C for 1 min, 5 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 1 min, 20 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 1 min, 10-min extension at 72°C, HindIII digestion yields a 300 bp band for 129 and a 150 bp band for C57BL/6.

Rag1-deficient thymocytes were induced to progress to the CD4\(^+\)CD8\(^+\) DP stage by injecting 4-6 week old Rag1-deficient mice once with 100µg anti-CD3 (145-2C11) intraperitoneally. Thymocytes were harvested 10 days later. For BrdU labeling experiments, mice were given a single intraperitoneal injection of 2mg BrdU (pulse-
chase) or 0.6 mg BrdU every 12 hours (continuous labeling). Mouse work was done under project licence PPL70/6845 issued by the Home Office, UK.

Jα segment usage was assessed by RT-PCR using Vα8 and Cα primers, following which Southern blots of PCR products were hybridized with radiolabeled Jα-specific and Cα-specific probes.

**Ligation-mediated PCR.** This was done as described using Jα-specific primers and probes (5′ to 3′)

- Jα61: primer (internal) AGTCCCTTCTCAGGCAAGATGG, probe TGAGGAACACGGAGATCTC
- Jα58: primer ATGGCTTTGGACCATGGATG, probe TCTGAAACTCGCACAGTGGA
- Jα49: primer AGGGAAATGACAACCAGGC, probe AGGTGTAGATTCTAGCTCTG
- Jα27: primer ATGGCAGATAGAATGGAGCGG, probe TACTCCTCAGCTG
- Jα6: primer ATCAGACCAGACTGTCTGCCC, probe GACCAATGGCAAAGGAGGT

**Confocal analysis.** Mitotic figures were stained for alpha tubulin (DM1A, Sigma) and DAPI as described.

**Flow cytometry.** For the simultaneous detection of BrdU and cell surface markers we used a modification of described protocols (ref. 21, 22). Cells (10⁶) were stained for surface markers (CD4, CD8, TCR), fixed in 4% formaldehyde in PBS, 0.01% Tween 20 for 5 min at 37°C, permeabilised in 0.1% Triton X100, 2% FCS for 1 hour on ice, treated with 50 Kunitz units/ml DNase I (bovine pancreas, Sigma) in pre-warmed phenol-free IMDM for 10 min at 37°C water bath, stained with 5µl of anti-BrdU-FITC (Becton Dickinson) for 10 min at room temperature and washed twice before analysis.

**Real time PCR** was carried out in a 20µl reaction volume using iQ SYBR Green Supermix (Bio-Rad) and 0.5µM primers. Amplification was 95°C for 3 min, then 40 cycles of 94°C for 15s, 60°C for 30s and 72°C for 30s followed by plate read.

**RT-PCR and transcript copy number analysis.** Total RNA was extracted using RNAbee (Tel-Test) and 1µg was treated with 1µl Turbo DNase (Ambion). Following DNase inactivation the RNA was reverse-transcribed using 200U Superscript III...
(Invitrogen) and 0.5µg oligo(dT) or 3µg random primers in a 20µl reaction volume. 1/80 was used per real time PCR reaction.

To estimate the number of transcripts, standard curves were generated for each pair of primers. The PCR amplicons generated by individual primer pairs were separated by agarose gel electrophoresis, gel-purified and their concentration was determined using Quant-iT Picogreen (Invitrogen). The molar weight of the amplicons was calculated based on their sequence using OligoCalc (ref. 44), which allowed the enumeration of copies in each sample. Two-fold serial dilutions from 2,500,000 to 1,220 copies (spliced transcripts) and 2,000,000 to 976 copies (intronic transcripts) were analysed by real time PCR. The standard curves were compiled by plotting the real time PCR C(t) values against seeded copy numbers and subsequently fitting a logarithmic function. Only primers that generated a log fit $R^2>0.99$ were used. These standard curves were subsequently used to extrapolate the number of transcripts in the biological samples based on the C(t) values obtained after real time RT-PCR. For the analysis of spliced transcripts the forward primer was specific to the respective Jα and the reverse primer to Cα.

eJα56 forward CTG GAG GCA ATA ATA AGC TGA CT
eJα52 forward AAC ACT GGA AAG CTC ACG TTT GG
eJα44 forward GGC AGT GGT GGA AAA CTC ACT
eJα38 forward AGC TGA TTT GGG GCT TGG GGA
eJα31 forward GGA CGC AG C TGG TGG TGA AGC
eJα23 forward GCT TAT CTT TGG ACA GGG AAC CA
eJα18 forward AAC TTA CCT TCG GGA CAG GG
eJα9 forward ACA AAC TTA CCT TCG GGA CAG G
TCRa-C reverse TCC ATA GCT TTC ATG TCC AGC

Germline transcripts were analysed using primer pairs specific to introns adjacent to the Jα segments included in the spliced transcripts analysis. No amplification was detected without reverse transcriptase.

iJα56 forward CCCTTGAACCCTGATATGC
iJα56 reverse CAGCCATTGTTTGGATTGGA
iJa52 forward CTGTGGGCTGTTGCTTTCTG
iJa52 reverse AGGGAAATGGGTGGGTACACG
iJa44 forward GAAGAGGCCTCATGGAGGAA
iJa44 reverse TCTCACAGACCGAAGGGACA
iJa38 forward AGGACTTGGGCTGCTCTCCA
iJa38 reverse GGCTCCTCAGGACAGACCAG
iJa31 forward GCAGCCTGCCAGCTATCTTT
iJa31 reverse CAAAAGCAGCAACCCAACAA
iJa23 forward AGAGGAGGCCGAAAGTCTCC
iJa23 reverse TGCAAAAATCCAGGCCTATG
iJa3 forward TCTGGACCCTTGGCAATCAT
iJa3 reverse CCCTGCCCTGGTCTACTGTG
iJ52-Cδ forward GGGTTTTGAGTTGGCTTTGG
iJ52-Cδ reverse GGCATGTATTTTTGCGGTTGA

ChIP-sequencing. Single read high throughput sequencing libraries were prepared from 10ng of ChIP DNA and sequenced according to the manufacturer's protocols (Illumina Genome Analyzer II). Reads of 38 bases were aligned to the Mus musculus mm9 genome assembly using ELAND (Illumina) and wig files generated with FindPeaks (ref. 45) were visualised with the UCSC Genome Browser (http://www.genome.ucsc.edu/).

RNA sequencing. 10µg total RNA were depleted of ribosomal RNA using RiboMinus (Invitrogen). Depletion was verified using an Agilent 2100 Bioanalyser and the retention of mRNA by real time RT-PCR. The samples were treated with DNasel, and the RNA was fragmented for 2 minutes at 94°C in RNA fragmentation buffer (final concentration: 40mM Tris acetate, pH 8.2, 100mM potassium acetate and 30mM magnesium acetate). The reaction was stopped by placing the samples on ice and the fragmented RNA was column purified (RNAeasy Minelute Kit, Qiagen), eluted in 10 µl, and used as template for cDNA synthesis using 5 µg random hexamers in a total volume of 20 µl. Second-strand synthesis was performed by adding 91.8 µl water, 30 µl 5x Second strand buffer (Invitrogen), 3 µl 10 mM dNTPs, 4 µl E. coli polymerase I (10 U/µl), 1 µl E. coli DNA ligase (10 U/µl) and 0.2 µl RNase H (10 U/µl), followed by 2h at 16°C. 1 µl T4 DNA Polymerase (10 U/µl) was added followed by an additional 10 min at 16°C. The ds cDNA was purified using the Minelute Reaction Cleanup Kit (Qiagen), eluting twice in 10 µl.
Following quantification, 20ng ds cDNA was used for the preparation of a single read high throughput sequencing library according to Illumina protocols.

**Rearrangement frequency analysis by semi-quantitative PCR**

Three-fold serial dilutions of 100ng genomic DNA were used for PCR amplification with forward primers specific to Vα5 or Vα7 and reverse primers specific to Jα56, Jα38 or Jα27, respectively\(^46\) (this reference refers to Vα5 as TRAV3-4 and Vα7 as TRADV15-1). Primers for the Calreticulin locus were used as loading control. PCR products were separated by agarose gel electrophoresis and stained with ethidium bromide.

Vα7 forward GCA GGA TCT AAT GTG GCC CAG AAA GTG ATT CA  
Vα5 forward AGG TGA TCA CAG AGG CAT CCT

Jα56 reverse ACG TAC CTG GTA TAA CAC TCA GAA C  
Jα38 reverse CAA AGA CGA CTT TGT TGG  
Jα27 reverse TTA AGA GCC CAA GCA GAT GCA TAA G

Calreticulin forward TCATGAGTTCCCCACATCTTTG  
Calreticulin reverse CTGCCCTATCCTGAGTCTGACA

**Chromosome conformation capture (3C)**

Cells were fixed in 10% FCS, 1% formaldehyde for 10 minutes at room temperature and fixation was stopped with glycine (0.125M). 10\(^7\) cells per sample were lysed in 10mM Tris, pH8, 10mM NaCl, 5mM MgCl\(_2\), 0.2% NP-40 for 30 min on ice. The nuclei were pelleted and re-suspended in 0.5ml 1.2x digestion buffer (NEB2, New England Biolabs) and permeabilised with SDS (0.5% final concentration) for 1 hr at 37\(^0\)C, shaking at 800rpm and 3.3% Triton X-100 were added for an additional 1 hour at 37\(^0\)C. 2000U HindIII (New England Biolabs) were added before incubation over night at (37\(^0\)C, 800rpm) and inactivated with SDS (1.5%, 65\(^0\)C, 30 min). The reaction was diluted in 6.2 ml 1.1x T4 ligase buffer (New England Biolabs) and incubated at 37\(^0\)C for 1h after addition of 1% Triton X-100. 800U T4 DNA ligase (New England Biolabs) was added for 4 hrs at 16\(^0\)C, crosslinking was reversed by 300µg proteinase K (65\(^0\)C, 16h). 300µg RNase A was added for 1h at 37\(^0\)C. DNA was isolated by phenol/chloroform extraction and ethanol precipitation, quantified using Quant-iT PicoGreen (Invitrogen) and 200ng
DNA were used per TaqMan PCR reaction (QuantiFast, Qiagen). Data were normalised to the cross-linking frequency between the anchor and the neighbouring HindIII fragment. HindIII digestion efficiency was calculated as described\(^2\) (Suppl. Fig. 8a). The efficiency and linearity of 3C primers was tested on templates obtained by HindIII digestion and religation of genomic PCR products spanning the HindIII restriction sites. The products of each primer pair were gel-purified and quantified to calculate copy numbers. Standard curves were constructed by 10-fold serial dilution (2\(\times\)10\(^6\) to 20 copies). 3C primers were selected for > 90% efficiency and efficiency and a logarithmic coefficient of determination (R\(^2\)) >0.99 (Suppl. Fig. 8b,c).

| Distance from anchor (kb) | 3C primer | Reverse primer |
|--------------------------|-----------|----------------|
| -75                      | CTGCCCAAAAGAGAGAGCTTTGG | GCATGTGAATGAGAGAGATGG |
| -73 (TEA)                | GCCGATTGTTGCGCAAGAATAA | CACAGGCTGCTCCAGAGATG |
| -50                      | TCTCGAAGGTCACACAAAGCA | GGACACGCTCTTTGGCTAC |
| -44                      | AGTCTGTGGCGTGAACCTTTC | TGGGTGTTTTATGGGTTTGAAG |
| -26                      | TGGCCAGGGGACCATATTA | CTGCTTCCGCTGATGCTG |
| -19                      | CACAGGAGAATAGGCGCTCAGA | CTCAGGAGGCCCTCCAGATAA |
| -13                      | TGTCCCTGTCCACACAGTGCT | TAGGAGCAGAGGAGGCTGA |
| 0 (Anchor)               | AGCATGGGAGACACGTAAATG | ATTCGGTCCTCCGCTGGA |

Normalisation

|               | 3C primer                  | Reverse primer                  |
|---------------|----------------------------|---------------------------------|
| 8             | AATGCCGGGCATCTCAGACTT      | GAGACTCCCACGGCTTACC             |
| 9             | GGACTTGTCTATACTGCCCAAAGA  | ATGTCTGGAACGGCACC              |
| 13            | ATTGCTGTGGAAGGGCCACCA      | GCACCTGTCTATACTGCACCAA         |
| 19            | GGCGGAAAACACACATCAAGA      | GGGTCATCATCATTACCACTTCCCAGT    |
| 23            | TTTAAAAAGGCCAGACTGCAAACA  | CTCCCACATCCCTTTACCAA           |

TaqMan Probe [FAM]-CCTGCCTGCCTGAG GACTGC-[BHQ1]

Supplementary references

33. Spanopoulou E, Roman CA, Corcoran LM, Schlissel MS, Silver DP, Nemazee D, Nussenzweig MC, Shinton SA, Hardy RR, Baltimore D. Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. Genes Dev. 8, 1030-1042 (1994)

34. Hogquist KA, Jameson SC, Heath WR, Howard JL, Bevan MJ, Carbone FR. T cell receptor antagonist peptides induce positive selection. Cell. 76, 17-27 (1994)
35. Kaye J, Hsu ML, Sauron ME, Jameson SC, Gascoigne NR, Hedrick SM. Selective development of CD4+ T cells in transgenic mice expressing a class II MHC-restricted antigen receptor. Nature 341, 746-749 (1989)

36. Srinivas S, Watanabe T, Lin CS, William CM, Tanabe Y, Jessell TM, Costantini F. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev Biol. 1, 4 (2001)

37. Gregson HC, Schmiesing JA, Kim JS, Kobayashi T, Zhou S, Yokomori K. A potential role for human cohesin in mitotic spindle aster assembly. J Biol Chem 276, 47575-47582 (2001)

38. Kong X, Ball AR Jr, Sonoda E, Feng J, Takeda S, Fukagawa T, et al. Cohesin associates with spindle poles in a mitosis-specific manner and functions in spindle assembly in vertebrate cells. Mol Biol Cell 20, 1289-1301 (2009)

39. Laird PW, Zijderveld A, Linders K, Rudnicki MA, Jaenisch R, Berns A. Simplified mammalian DNA isolation procedure. Nucleic Acids Res 19, 4293 (1991).

40. Hawwari A, Bock C, Krangel MS. Regulation of T cell receptor a gene assembly by a complex hierarchy of germline Ja promoters. Nat. Immunol. 6, 481-489 (2005)

41. Abarrategui I, Krangel MS. Regulation of T cell receptor-a gene recombination by transcription. Nat. Immunol. 7, 1109-1115 (2006)

42. Jackson A, Kondilis HD, Khor B, Sleckman BP, Krangel MS. Regulation of T cell receptor b allelic exclusion at a level beyond accessibility. Nat Immunol 6, 189-197 (2005)

43. Toyoda Y, Yanagida M. Coordinated requirements of human topo II and cohesin for metaphase centromere alignment under Mad2-dependent spindle checkpoint surveillance. Mol Biol Cell.17, 2287-2302 (2006)
44. Kibbe, W.A. OligoCalc: an online oligonucleotide properties calculator. Nucl. Acids Res. 35, W43-46 (2007).

45. Fejes, AP, Robertson, G, Bilenky, M, Varhol, R, Bainbridge, M, Jones, SJM. FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology, Bioinformatics 24, 1729–1730 (2008).

46. Lee YN, Alt FW, Reyes J, Gleason M, Zarrin AA, Jung D. Differential utilization of T cell receptor TCRα/TCRδ locus variable region gene segments is mediated by accessibility. Proc. Natl. Acad. Sci. USA 106, 17487-17492 (2009).