Kinase-dependent structural role of DNA-PKcs during immunoglobulin class switch recombination

Jennifer L. Crowe[a,b,1], Zhengping Shao[a,1], Xiaobin S. Wang[a,b], Pei-Chi Wei (魏珮琪)[c,d], Wenxia Jiang[a], Brian J. Lee[a], Verna M. Estes[a], Frederick W. Alt[c,d], and Shan Zha[a,e,f,g,2]

[a]Institute for Cancer Genetics, Columbia University, New York, NY 10032; [b]Graduate Program of Pathobiology and Molecular Medicine, College of Physicians and Surgeons, Columbia University, New York, NY 10032; [c]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115; [d]Program in Cellular and Molecular Medicine, The Children's Hospital at Boston, Harvard Medical School, Boston, MA 02115; [e]Department of Pathology and Cell Biology, Columbia University, New York, NY 10032; [f]Herbert Irvine Comprehensive Cancer Center, Columbia University, New York, NY 10032; and [g]Division of Pediatric Oncology, Department of Pediatrics, Columbia University Medical Center, New York, NY 10032

Edited by Michel C. Nussenzweig, Howard Hughes Medical Institute and The Rockefeller University, New York, NY, and approved July 10, 2018 (received for review May 17, 2018)

The catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) is a classical nonhomologous end-joining (cNHEJ) factor. Loss of DNA-PKcs diminished mature B cell class switch recombination (CSR) to other isotypes, but not IgG1. Here, we show that expression of the kinase-dead DNA-PKcs (*DNA-PKcs*$^{KD/KD}$) severely compromises CSR to IgG1. High-throughput sequencing analyses of CSR junctions reveal frequent accumulation of nonproductive interchromosomal translocations, inversions, and extensive end resection in *DNA-PKcs*$^{KD/KD}$, but not *DNA-PKcs*$^{-/-}$, B cells. Meanwhile, the residual joints from *DNA-PKcs*$^{KD/KD}$ cells and the efficient Sμ-Sγ1 junctions from *DNA-PKcs*$^{-/-}$ B cells both display similar preferences for small (2–6 nt) microhomologies (MH). In *DNA-PKcs*$^{-/-}$ cells, Sμ-Sγ1 joints are more resistant to inversions and extensive resection than Sμ-Sε and Sμ-Sμ joints, providing a mechanism for the isotype-specific CSR defects. Together, our findings identify a kinase-dependent role of DNA-PKcs in suppressing MH-mediated end joining and a structural role of DNA-PKcs protein in the orientation of CSR.

DNA-PKcs | immunoglobulin class switch recombination | classical nonhomologous end joining | alternative end joining | somatic mutation

Upon contact with antigens, naïve B lymphocytes undergo class switch recombination (CSR) to achieve different effector functions (isotypes). CSR is initiated by activation-induced cytidine deaminase (AID), which introduces mismatches that are eventually converted to double-strand breaks (DSBs) within the switch (S) region preceding each set of constant region (C_H) exons. The joining between a DSB at Sμ and a downstream S region completes the CSR. While CSR primarily utilizes the classical nonhomologous end-joining (cNHEJ) pathway for repair, in the absence of cNHEJ factors (e.g., Xrcc4 or Lig4), up to 50% of CSR can be mediated by the alternative end-joining (A-EJ) pathways (1, 2) that preferentially use microhomology (MH) at the junctions. The relative extent of MH usage differs in Ku- vs. Xrcc4-deficient B cells, suggesting that more than one type of A-EJ pathways might exist (2).

The catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs) is a vertebrate-specific cNHEJ factor. Upon DSBs, KU70-KU80 heterodimer (KU) binds to DNA and recruits DNA-PKcs, which further recruits and activates Artemis endonuclease to open hairpin ends. DNA-PKcs and Artemis are not essential for direct ligation of blunt DNA ends (3–5). Accordingly, *DNA-PKcs*$^{-/-}$ mice are of normal size (3–5), in contrast to embryonic lethality of Lig4 or Xrcc4-deficient mice. However, expression of the kinase-dead DNA-PKcs protein completely blocks cNHEJ, resulting in embryonic lethality similar to *Lig4*$^{-/-}$ mice (6). Deletion of Ku, especially the C-terminal of Ku80 that recruits DNA-PKcs to the DNA ends, rescues the embryonic development and end ligation in *DNA-PKcs*$^{KD/KD}$ mice (6), suggesting that the DNA-PKcs protein physically blocks cNHEJ in the absence of its own kinase activity. Consistent with the dispensable role of DNA-PKcs in direct end ligation, *DNA-PKcs*$^{-/-}$ mature B cells with preassembled immunoglobulin heavy (IgH) and light chains (IgL) (HL) undergo efficient CSR to IgG1 (at nearly 80% of WT levels). However, switching to other isotypes are much more severely compromised (7–10). The reason for this isotype-dependent CSR defect in *DNA-PKcs*$^{-/-}$ B cells remains elusive. Analyses of rare Sμ-Sα junctions from severely immunocompromised DNA-PKcs–deficient patients ($n = 2$) or DNA-PKcs null mice (without rescue by HL) suggest an increase of large (>7 bp), but not small (1–6 bp), MH at the junctions (11). In contrast to cNHEJ, MH-mediated A-EJ often requires DNA end resection to expose the flanking MHs (12) and KU suppresses A-EJ by blocking EXO1 mediated end resection (13, 14). So, we asked whether the presence of DNA-PKcs-KD would block end resection and therefore A-EJ in switching B cells. In this context, DNA-PKcs–specific kinase inhibitors (NU7441 or NU7026) promote A-EJ without blocking cNHEJ in WT cells (15–18). Notably DNA-PKcs inhibitors have a off rate and can inhibit other related kinases at 5- to 15-μM ranges (19).

To ascertain how different DNA-PKcs mutations (null vs. KD) affect CSR in an isotype-dependent manner, we employed the

Significance

To combat pathogens, B lymphocytes switch the types of antibody they express through a DNA break and repair event termed class switch recombination (CSR). To do so, DNA breaks introduced in specific genomic regions within the antibody gene locus are joined together by the nonhomologous end-joining (NHEJ) DNA repair pathway. Here, we show that mutation and deletion in one NHEJ factor, the catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs), leave different molecular signatures at the CSR junctions, uncovering activation-dependent structural function of DNA-PKcs in DNA repair and antibody maturation.

Author contributions: J.L.C. and S.Z. designed research; J.L.C., B.J.L., V.M.E., and S.Z. performed research; Z.S., X.S.W., P.-C.W., W.J., F.W.A., and S.Z. contributed new reagents/analytic tools; W.J. generated the mouse model; J.L.C., Z.S., X.S.W., and S.Z. analyzed data; and J.L.C., X.S.W., and S.Z. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo/ (accession no. GSE117628).

[1]J.L.C. and Z.S. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: sz2296@columbia.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1808490115/-/DCSupplemental.

Published online August 2, 2018.

high-throughput genome translocation sequencing (HTGTS) (20) method to analyze CSR junctions in *DNA-PKcs*$^{KD/KD}$ and *DNA-PKcs*$^{-/-}$ B cells with preassembled IgH and IgL chains (HL). In contrast to *DNA-PKcs*$^{-/-}$ cells, *DNA-PKcs*$^{KD/KD}$ B cells display severe switching defects in IgG1, like the cNHEJ-deficient *Xrcc4*$^{-/-}$ B cells. However, CSR junctions from *DNA-PKcs*$^{KD/KD}$ and *DNA-PKcs*$^{-/-}$ B cells have similar increases of small MH (2–7 nt) as the price of blunt joints, suggesting that DNA-PKcs suppresses MH-mediated A-EJ via its kinase activity. Despite similar MH usage, Sμ-Sγ1 joints from *DNA-PKcs*$^{-/-}$ B are much more resilient to inversions and deletions than both Sμ-Sμ and Sμ-Sε junctions, suggesting differential preference to the productive orientations might contribute to the isotype-dependent switching defects in DNA-PKcs–deficient cells. Finally, our analyses also identified long MH-mediated interchromosomal translocations in *DNA-PKcs*$^{KD/KD}$ B cells and a reduced number of G mutations in 5′Sμ in repair-deficient B cells.

Results

B Cells Expressing Kinase-Dead DNA-PKcs Display Severe CSR Defects. To circumvent the requirement for DNA-PKcs in V(D)J recombination and early B cell development, we generated *DNA-PKcs*$^{KD/KD}$ mice carrying the germ-line knock-in IgH and Igκ(kappa) chains (referred to as *HL*$^{k/k}$) (21, 22). Similar approaches have been used to study the role of other cNHEJ factors, including Ku70, Xrcc4, Ligase4, and Artemis, in CSR (1, 2, 23, 24). Tp53 deficiency was included to rescue the embryonic development of *DNA-PKcs*$^{KD/KD}$ mice (6). Consistent with previous reports (25), Tp53 deficiency, heterozygous or homozygous, does not affect CSR efficiency (*SI Appendix*, Fig. S1*A*) or junction properties (see below). So, we combined the result from *Tp53*$^{+/-}$ and *Tp53*$^{-/-}$ mice together. For reasons that are yet to be understood, the majority of *DNA-PKcs*$^{KD/KD}$*Tp53*$^{-/-}$ mice died shortly after 21 d of age. Therefore, the CSR analyses were performed on splenic B cells derived from young (~21 d old) *DNA-PKcs*$^{KD/KD}$*Tp53*$^{-/-}$ HL or young adult (up to 6 wk) *DNA-PKcs*$^{KD/KD}$*Tp53*$^{+/-}$ HL mice with controls. The splenic B cells were activated by anti-CD40 and IL4 to initiate CSR to IgG1 and IgE. As shown in Fig. 1 *A* and *B*, ~50% HL and Tp53-deficient/HL B cells became IgG1^{+} by day 4. *DNA-PKcs*$^{-/-}$ B cells undergo γ1 switch at ~80% of the WT levels ($P = 0.0387$), while *DNA-PKcs*$^{KD/KD}$*Tp53*$^{-/-}$HL and *Xrcc4*$^{-/-}$*Tp53*$^{-/-}$HL B cells can only undergo γ1 switch at 25–30% of the WT levels ($P < 0.0001$) (Fig. 1 *A* and *B*) (1, 7–9, 26). Productive CSR requires cell proliferation, so we traced cell division via the Cell Trace Violet (CTV) dye. Despite moderate reduction of overall proliferation, the relative frequency of IgG1^{+} cells in *DNA-PKcs*$^{KD/KD}$ B cells after three cell divisions is also ~25% of the WT levels, suggesting *DNA-PKcs*$^{KD/KD}$ B cells have proliferation-independent defects in CSR (Fig. 1*C*). As in *DNA-PKcs*$^{+/+}$ cells, Tp53 status does not affect CSR efficiency in *DNA-PKcs*$^{KD/KD}$ B cells (*SI Appendix*, Fig. S1*B*). Together, these data suggest that the presence of kinase-dead DNA-PKcs protein further blocks CSR, beyond the loss of DNA-PKcs.

Increased Interchromosomal Translocations in B Cells from *DNA-PKcs*$^{KD/KD}$ but Not *DNA-Pkcs*$^{-/-}$ Mice. To determine whether DNA-PKcs-KD blocks MH-mediated A-EJ, we applied the HTGTS approach (20, 27, 28) to analyze thousands of junctions from a single 5′Sμ bait to other partners (preys) (Fig. 2*A*). For each genotype, we analyzed several independently stimulated and prepared libraries (*SI Appendix*, Fig. S2*A*). As previously reported (20, 28, 29), HTGTS analyses recover internal deletions (prey sites within Sμ), productive CSR (prey sites within Sγ1 or Sε), and other nonproductive joints, including inversions and intrachromosomal and interchromosomal translocations (outside Chr12). In addition to junctional property (MH, blunt and insertion), HTGTS also provides location and orientation (deletion vs. inversion) information on the prey DNA breaks (28) and allows the comparison between different switch regions (29). The vast majority of prey-break sites from wild-type (90.9 ± 3.1%), Tp53-deficient (93.4 ± 0.3%), and DNA-PKcs null (84 ± 6.9%) B cells were within the IgH locus (Fig. 2*B*), consistent with low frequency cytogenetic aberration (~10%) previously documented (10). In contrast, nearly 20% of prey sites identified from *Xrcc4*$^{-/-}$ B cells (18.9 ± 4.2%) and nearly 40% of prey breaks from *DNA-PKcs*$^{KD/KD}$ cells (39.7 ± 2.8%) reside outside of IgH, and even outside of Chr12, suggesting frequent interchromosomal translocations (Fig. 2*B*).

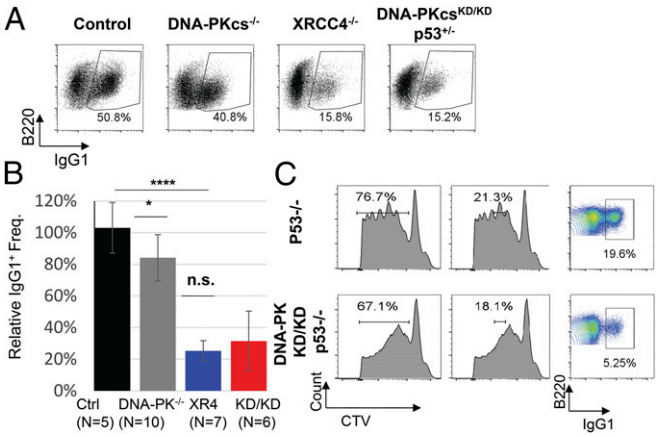

Fig. 1. Class switch recombination analyses of DNA-PKcs–deficient and mutated B cells. (*A*) Representative CSR switch percentage (defined as B220^{+}IgG1^{+}) on day 4 of stimulation. All mice carry the knocked-in productive IgH and IgL. (*B*) Quantification of relative B220^{+}IgG1^{+} B cells (against the littermate control in each experiments) on day 4 of stimulation. The bars represent the average and the SD of 5 or more mice per genotype. (*C*) Representative Cell Trace Violet proliferation stain in *Tp53*$^{-/-}$ (Control) and *PKcs*$^{KD/KD}$*Tp53*$^{-/-}$ splenic B cells stimulated with anti-CD40 and IL-4 for CSR. *$P < 0.05$, ****$P < 0.0001$, n.s., not significant. *P* value was determined by two-tailed Student's *t* test with equal variance.

Sε Switching Is Preferentially Affected over Sγ1 Switching in DNA-Pkcs or NHEJ-Deficient B Cells. Next, we examined the distribution of prey-break sites among Sμ, Sγ1, and Sε (Fig. 2*A*). In WT and Tp53-deficient B cells, ~53–55% of the IgH prey fall into Sγ1, 25% in Sε, with another 25% reflecting internal deletion in Sμ. In *DNA-PKcs*$^{-/-}$ cells, while there are still ~58% of the IgH prey in Sγ1, only 6.3% is now in Sε, a sharp reduction from the 25% in WT cells. Despite severe Sγ1 switching defect in *Xrcc4*$^{-/-}$ and *DNA-PKcs*$^{KD/KD}$ cells (Fig. 1*C*), nearly 50% of preys in IgH remains in Sγ1, while the frequency of Sε prey breaks dropped to 3–4% (Fig. 2*C*). Even considering all junctions (including translocations), the frequency of Sγ1 junctions does not drop much, while Sε junctions almost diminished in *Xrcc4*$^{-/-}$ and *DNA-PKcs*$^{KD/KD}$ B cells (*SI Appendix*, Fig. S2*B*), suggesting cNHEJ deficiency preferentially compromises Sε over Sγ1 junctions, which explains the isotype-dependent CSR defects in *DNA-PKcs*$^{-/-}$ B cells.

Relative Increase of Inversion Events in All IgH Junctions from *DNA-PKcs*$^{KD/KD}$ Cells. Next, we consider the orientation of the joints (28). We define the prey sequence that reads from the junction as centromere to the telomere orientation as (+) and from telomere to centromere as (−) (Fig. 2*A*). Since the bait primer resides at the 5′ of Sμ and reads from the telomere to the centromere toward the junction, the internal deletion and productive rearrangement would give rise to prey of (−)

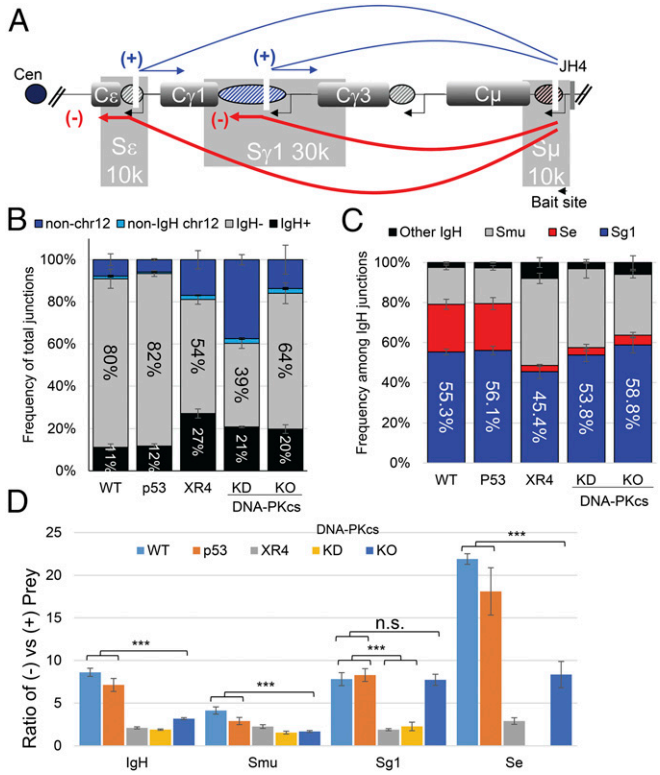

Fig. 2. Distribution of the prey sites recovered from a common 5′Sμ bait site. (*A*) Diagram of representative preys identified in HTGTS using a common bait at the 5′Sμ region. The centromere (cen) and JH4 exon (closer to telomere) are marked. The blue arrows indicate possible configuration of (+) strand/inversional prey sites, and the red arrows indicate the possible configuration of (−) strand/deletional prey sites. The relative size and position of the Sμ, Sγ1, and Sε switch region are marked by gray boxes. The data in this figure represent the average and SE of several independently derived libraries of each genotype (see *SI Appendix*, Fig. S2*A* for sample summary). (*B*) Frequency of all HTGTS-identified prey-break sites including IgH (+) in black, IgH (−) in gray, non-IgH within chromosome 12 (in light blue), and outside of chromosome 12 (in blue, interchromosomal translocations). (*C*) The relative frequency of prey break sites in Sμ, Sγ1, and Sε and other IgH regions among all IgH prey sites. The relative frequency of prey in each switch region among all junctions (including non-IgH) can be found in *SI Appendix*, Fig. S2*B*. (*D*) The (−) vs. (+) strand prey break sites ratio in IgH and among each switch regions (μ, γ1, and ε). ***$P < 0.001$ by two-tailed Student's t test with unequal variance and n.s., not significant ($P > 0.05$).

orientations, and inversion would give rise to prey of (+) orientation (Fig. 2*A*). As previously reported (28), the IgH preys from WT B cells are strongly biased toward deletions (−) [Fig. 2*B*, with the (−):(+) ratio = 8.61 ± 1.16]. Preys in the non-IgH region and outside chromosome 12 are roughly evenly distributed between (+) and (−) orientations (*SI Appendix*, Fig. S2*C*), suggesting that the orientation bias is mediated by an intrachromosomal mechanism (30). Since our assay could not distinguish the paternal and maternal chromosome 12, the IgH prey breaks also include a small number of interallelic switching events, presumably evenly distributed between (+) and (−) orientation like other interchromosomal translocations (31). Tp53 deficiency does not alter the orientation bias within IgH prey-break sites (−):(+) = 7.13 ± 1.52 ($P > 0.15$ vs. WT) (Fig. 2*D*). However, the ratio between (−) and (+) prey decreased markedly among IgH junctions recovered from *Xrcc4*$^{-/-}$ (−:+ = 2.10 ± 0.25, $P = 1.77 \times 10^{-5}$ vs. WT), *DNA-PKcs*$^{KD/KD}$ (−:+ = 1.89 ± 0.11), and surprisingly also *DNA-PKcs*$^{-/-}$ B cells (−:+ = 3.17 ± 0.24, $P = 4.07 \times 10^{-5}$ vs. WT) (Fig. 2*D* and *SI Appendix*, Fig. S2*B*). In WT B cells, the orientation bias (−):(+) is most prominent in Sε (21.90 ± 1.53), followed by Sγ1 (7.82 ± 1.86), and lowest in Sμ (4.13 ± 1.04) (Fig. 2*D*). This orientation bias decreased in all switch regions in *Xrcc4*$^{-/-}$ (Sε = 2.90 ± 0.78, Sγ1 = 1.88 ± 0.24, and Sμ = 2.25 ± 0.43), and *DNA-PKcs*$^{KD/KD}$ (Sε = 1.90 ± 0.58, Sγ1 = 1.54 ± 0.24, and Sμ = 2.27 ± 0.71) B cells (Fig. 2*D*). Between the different switch regions, the orientation bias decreases more (by fold) among Sε prey breaks than those from Sγ1 prey breaks. In joints recovered from *DNA-PKcs*$^{-/-}$ cells, the (−):(+) ratio decreases in both Sμ (1.67 ± 0.22) and Sε1 (8.34 ± 3.10), but not in Sγ1 (7.73 ± 1.30, P value vs. WT > 0.5), again highlighting the resilience of Sγ1 switching in *DNA-PKcs*$^{-/-}$ cells (Fig. 2*D*).

Increased Distal Switching Events in *DNA-Pkcs*$^{-/-}$ and *DNA-PKcs*$^{KD/KD}$ B Cells. Next, we examined the spatial distribution of the prey-break sites within Sμ, Sγ1, and Sε regions. Of the Sμ junctions derived from *DNA-PKcs*$^{+/+}$ and *Tp53*$^{-/-}$ cells, most prey-break sites were located near the bait sites favoring small internal deletions (Fig. 3). Sμ junctions derived from *Xrcc4*$^{-/-}$, *DNA-PKcs*$^{-/-}$, and *DNA-PKcs*$^{KD/KD}$ B cells spread farther centromeric (distal from the 5′Sμ), indicative of larger internal deletions even into the Cμ exons (Fig. 3). Similar spreading, particularly of the (−) strand junctions, was also noted in the Sγ1 and Sε region junctions derived from *Xrcc4*$^{-/-}$, *DNA-PKcs*$^{-/-}$, and *DNA-PKcs*$^{KD/KD}$ cells (Fig. 4 and *SI Appendix*, Fig. S3). Notably, the spreading of Sγ1 prey is comparable between those from *Xrcc4*$^{-/-}$ and those from *DNA-PKcs*$^{-/-}$ cells, while the spreading of Sμ prey to downstream regions is significantly less in *DNA-PKcs*$^{-/-}$ cells. Finally, the majority of (−) and (+) prey breaks within Sγ1 nicely correlate with the frequency of AID hot spots (plot as the number of "RGYW" sequence motif per 50 bp) in Sγ1, indicating that they are AID-initiated events.

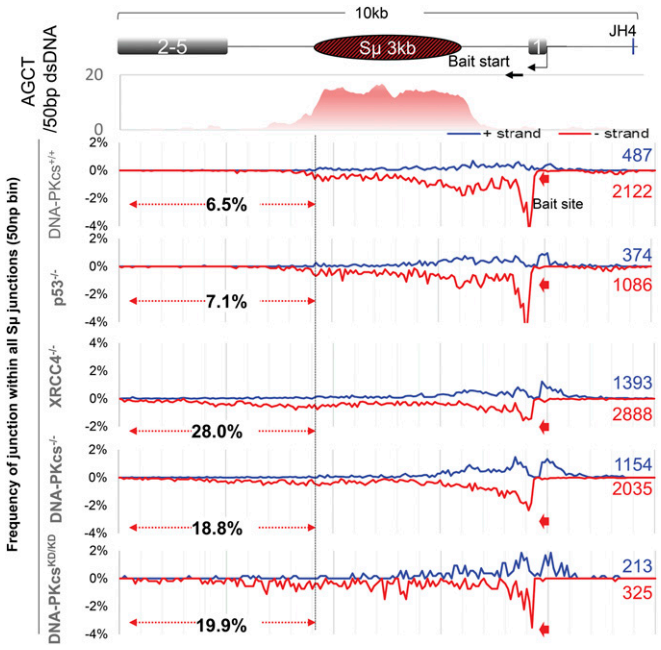

Fig. 3. The spatial distribution of prey-break sites in Sμ. The frequency (%) of Sμ prey breaks that falls into each 50-bp bin were plotted. The pool of all data from each genotype was plotted. All (−, red, below 0) and (+, blue, above 0) strand prey breaks added up to 100%. Schematic of the Sμ region is at *Top*. For each genotype, the number of (+) strand (blue) and (−) strand (red) junctions are marked at the right. The number of AGCT motifs per 50-bp double-strand DNA is plotted. The bait site at the 5′ Sμ is marked by an arrow. The dashed lines indicate the percentage of (−) strand Sμ prey that fall outside the core Sμ region.

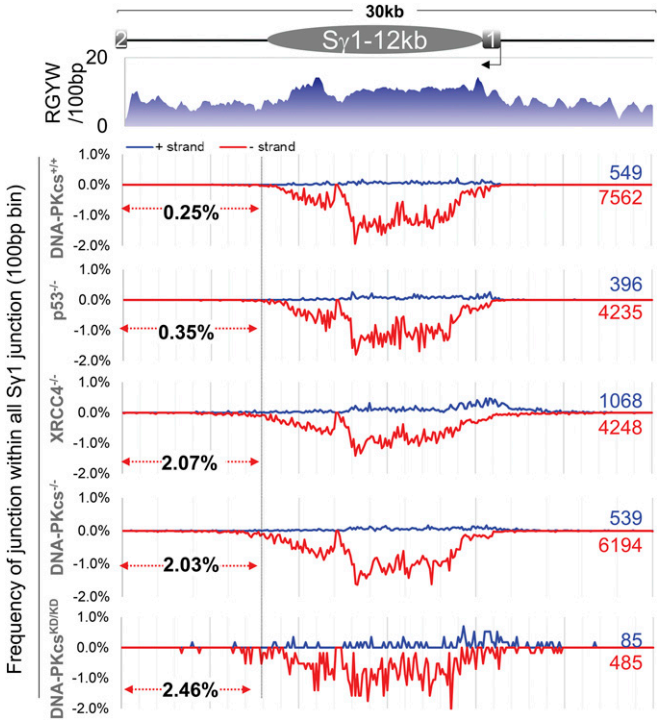

Fig. 4. The spatial distribution of prey break sites in Sγ1. The plot represents the frequency (%) of Sγ1 prey-break sites that falls into each 100-bp bin. The pool of all data from each genotype was plotted. All (–, red, below 0) and (+, blue, above 0) strand-prey breaks add up to 100%. For each genotype, the number of total (+) strand (blue) and (–) strand (red) junctions are marked at the right. The number of RGYW motifs per 100-bp double-strand DNA are shown at *Top*. The dashed lines indicate the percentage of (–) strand Sγ1 prey that fall outside the core Sγ1 region.

Residual CSR and Internal Deletions in *DNA-PKcs*$^{KD/KD}$ or *DNA-Pkcs*$^{-/-}$ B Cell Junctions Are Mediated by Small MH. To determine the contribution of A-EJ to CSR, we divided junctions recovered from HTGTS analyses into three mutually exclusive categories—MH, blunt, and insertion (INS) and then further divided MH or INS junctions based on the size of MH or INS (Fig. 5*A*). For sequences with insertions, the usage of MH cannot be determined by HTGTS (29). The joints involving the IgH prey breaks are consistently biased toward small MH than the joints involving the non-IgH prey sites from the same genotype, likely reflecting the CG-rich and repetitive nature of the switch region. For this reason, the IgH and non-IgH prey breaks were analyzed independently (*SI Appendix*, Fig. S4*A*). IgH junctions derived from both WT and p53-deficient B cells included ∼25% blunt, 22% 1-nt MH, 32% 2- to 7-nt MH. Approximately 20% of IgH junctions from WT or p53-deficient cells had insertions (1–30 nt) (Fig. 5 *A* and *B*). In comparison, the IgH junctions derived from *Xrcc4*$^{-/-}$ and *DNA-PKcs*$^{KD/KD}$ cells contained only 5% blunt, 11% 1-nt MH, and a remarkable 66–70% with 2- to 7-nt MH (Fig. 5 *A* and *B*). Perhaps most surprisingly, despite the nearly normal switching to IgG1, the total IgH as well as the productive Sγ1(−) junctions from *DNA-PKcs*$^{-/-}$ cells were also dominated by small MHs (2–7 nt), similar to those from XRCC4-deficient cells (Fig. 5 *A* and *B* and *SI Appendix*, Fig. S4*B*), suggesting that small MH-mediated A-EJ is not just a backup pathway, but can indeed catalyze robust CSR to IgG1, at least in *DNA-PKcs*$^{-/-}$ cells.

In contrast to IgH junctions, non-IgH junctions (also from the common 5′Sμ bait site) recovered from WT and Tp53-deficient cells contained ∼20–25% blunt junctions along with a wider spread of MH size (from 1 to >15 bp) (*SI Appendix*, Fig. S4*C*). The non-IgH junctions with small MH (2–5 nt) increased significantly in the absence of Xrcc4 or DNA-PKcs. In *DNA-PKcs*$^{KD/KD}$ cells, the non-IgH junctions displayed a unique shift to large MH (>10 nt) (*SI Appendix*, Fig. S4*C*), potentially reflecting extensive resection to bypass both Ku and DNA-PKcs KD at the DNA ends. Such large MHs (>10 nt) are rare in IgH junctions.

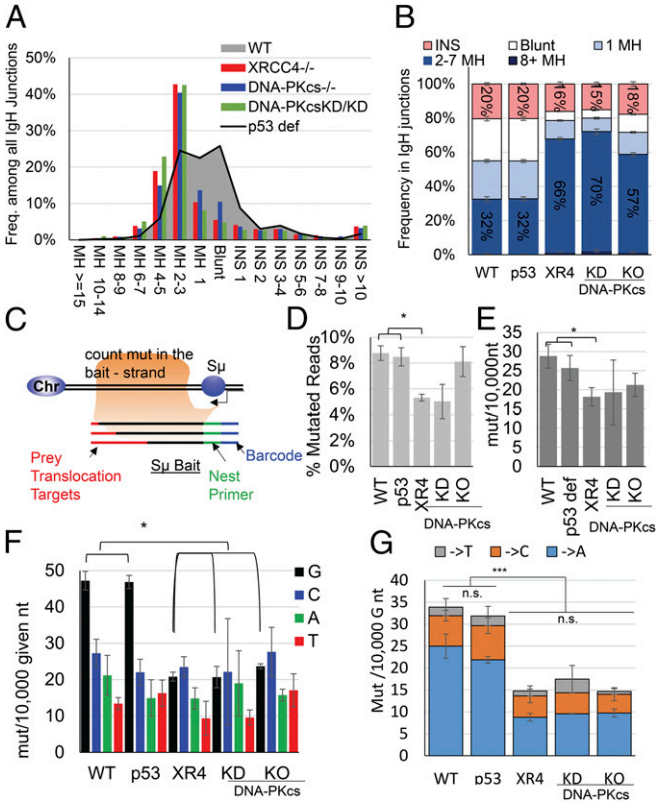

Fig. 5. The usage of MH, blunt, and insertion in switch junctions. (*A*) The distribution of IgH junctions by junctional type (INS, insertion). This graph represents the pool of IgH junctions from each genotype (see *SI Appendix*, Fig. S2*A* for sample information). The junctions were collected from three *DNA-PKcs*$^{+/+}$, four p53-defecient, three *Xrcc4*$^{-/-}$*p53*$^{-/-}$, four *DNA-Pkcs*$^{-/-}$*p53*$^{+/-}$, and two *DNA-PKcs*$^{KD/KD}$*p53*$^{-/-}$ mice (all with HL knock-in). The bars in *B* and *D*–*G* represent the average and the SEs of several independently derived libraries by genotype. (*B*) The distribution of the IgH junctions by junctional features (both + and – orientation). INS includes a 1- to 30-nt insertion. The bar represents the average and SEs of multiple libraries derived from independently stimulated B cells of each genotype. The bars in *B* and *D*–*G* represent the average and the SE of several independently derived libraries by genotype. (*C*) Diagram of the bait site and nucleotide counting scheme for mutational analysis. The barcode, the common nesting primer binding site, and the prey sites (including any possible insertions) were excluded from the mutation analyses. The top strand relative to Sμ germ-line transcript (telomere to centromere direction, –) was used as the germ-line sequence for base analyses. Only bait reads from IgH preys were included, and over 5,000 reads and more than 270K nt were analyzed per genotype (except *DNA-PKcs*$^{KD/KD}$). See *SI Appendix*, Fig. S4*D* for sample information. Two *DNA-PKcs*$^{-/-}$ libraries with less than 200 IgH reads per library were excluded from the mutation analyses due to extreme absolute number of mutations (due to low reads and nucleotide numbers). (*D*) Percentage of IgH reads with any mutation in the bait region by genotype. (*E*) The rate of mutation per 10,000 nt in the bait region of all IgH junctions by genotype. (*F*) The rate of mutations per 10,000 of each nucleotide type (A, T, C, or G). The mutation per 10,000 G is significantly higher in WT and Tp53-deficient B cells than in *Xrcc4*$^{-/-}$, *DNA-PKcs*$^{-/-}$, or *DNA-PKcs*$^{KD/KD}$ B cells. (*G*) Rate of mutations per 10,000 G nucleotides broken down by the type of mutation (G to A, to C and to T). **P* < 0.05, ****P* < 0.001, n.s., not significant. *P* value was determined by two-tailed Student's *t* test with equal variance.

Reduced G Mutations in the 5′Sμ Bait Region from IgH Junctions Recovered from Repair-Deficient Cells. Finally, we analyzed mutations at the 5′Sμ region among the junctions (Fig. 5*C* and *SI Appendix*, Fig. S4*D*). To compare with prior studies using Sanger sequencing, we present the data based on the top strand relative to the Sμ germ-line transcription (Fig. 5*C*). The estimated mutation rate for the MiSeq platform is 0.1% (10 per 10,000). The overall mutation rate (mut/10,000 bases) and frequency of reads with mutations were both higher in IgH junctions derived from *DNA-PKcs*$^{+/+}$ and T*p53*$^{-/-}$ cells (~27 mut/10,000 nt) than those from *Xrcc4*$^{-/-}$ and *DNA-PKcs*$^{KD/KD}$ cells that were also Tp53 deficient (~16 mut/10,000 nt) (Fig. 5 *D* and *E*). Junctions from *DNA-PKcs*$^{-/-}$ cells also had a slightly lower mutation rate (Fig. 5*E*). Despite a similar distribution of nucleotide types (*SI Appendix*, Fig. S4*E*), G was most frequently mutated in the 5′Sμ region from *DNA-PKcs*$^{+/+}$ and T*p53*$^{-/-}$ B cells (47 mut/10,000 G), and reductions in G mutations account for the lower mutation rate in *Xrcc4*$^{-/-}$ and *DNA-PKcs*$^{KD/KD}$ cells (20 mut/10,0000 G) (Fig. 5 *F* and *G*). The pattern of G mutations (relative frequency to A, T, C) did not change, suggesting the reduction is likely due to AID targeting, rather than differences in secondary processing by uracil-DNA glycosylase (UNG), mismatch repair (MMR), or other pathways (Fig. 5*G* and *SI Appendix*, Fig. S4*F*). Altogether, we found a reduction of G nucleotide mutations in the 5′Sμ region in *Xrcc4*$^{-/-}$, *DNA-PKcs*$^{KD/KD}$, and *DNA-PKcs*$^{-/-}$ cells.

Discussion

Here, we present the high-throughput analyses of CSR and translocation junctions from two types of DNA-PKcs–deficient B cells: loss of DNA-PKcs (null) and the expression of the kinase-dead DNA-PKcs (*DNA-PKcs*$^{KD/KD}$). Our results identified a role of DNA-PKcs kinase activity in suppressing MH-mediated A-EJ, and a kinase-dependent structural function of DNA-PKcs in end ligation that ensures the orientation of CSR. Moreover, despite similar MH usages, Sμ-Sγ1 joints are more resistant to inversion and deletions, providing an explanation for the robust IgG1 switching in *DNA-PKcs*$^{-/-}$ cells.

Specifically, we found that *DNA-PKcs*$^{KD/KD}$ B cells display much more severe CSR defects than *DNA-PKcs*$^{-/-}$ cells. The level of residual CSR, the frequency, and size of MHs among CSR junctions are indistinguishable between *DNA-PKcs*$^{KD/KD}$ and *Xrcc4*$^{-/-}$ B cells, suggesting the DNA-PKcs KD protein, while blocking cNHEJ, does not block MH-mediated A-EJ. A-EJ requires end resection to generate single-strand DNA overhangs to initiate MH-mediated annealing and eventually end ligation. Single-stranded DNA overhangs also preclude KU binding and protect the DNA ends from cNHEJ. In contrast to cNHEJ, which has to work at the DNA ends, end resection can be initiated at the end by exonuclease (e.g., EXO1) or internally by endonucleases, such as the MRE11–RAD50–NBS1 (MRN) complex with the help of CtIP (12). KU at the DNA ends prevents EXO1 cleavage, but not MRN/CtIP-mediated endonuclease cleavage. In fact, in vitro studies suggest that end binding by the KU dimers stimulates the endonuclease activity of the MRN complex (32, 33). During meiosis, MRN/CtIP removes the DNA ends covalently linked Spo11 protein, further supporting the ability of MRN/CtIP to bypass protein block at the DNA ends (12). Given that the DNA-PKcs-KD protein is recruited to the DNA ends by KU, we envision that MRN/CtIP can release the stalled KU and DNA-PKcsKD proteins to initiate end resection and commit the ends to A-EJ. Alternatively, ongoing AID activity and the processing of AID-generated U-G mismatches can also generate additional breaks with single-strand overhangs without bona fide end resection, and such ssDNA overhangs would prevent KU and DNA-PKcs KD binding and favor A-EJ. Correspondingly, extensive usage of the downstream (distal) switch regions is noted in Sμ and Sγ1 prey break sites recovered from *DNA-PKcs*$^{KD/KD}$ cells and *DNA-PKcs*$^{-/-}$ cells (Fig. 3).

In both *DNA-PKcs*$^{-/-}$ (7–10) and AID-mutated cells (34), γ1 switching is less affected than other isotypes tested. HTGTS analyses indicated that the size and extent of MH usage are similar among Sμ-Sγ1, Sμ-Sμ, and Sμ-Sε junctions (*SI Appendix*, Fig. S4*B*) and are not affected by orientation (−/+). Orientation preference during CSR is determined by cis-IgH organization features and enforced by ATM and 53BP1-mediated DNA damage responses (28). In addition, increased translocations, including interhomologous switches, can also attenuate the orientation preference (29). In *DNA-PKcs*$^{-/-}$ B cells, the orientation bias is compromised in distal Sμ-Sε and proximal Sμ-Sμ joints, but not the Sμ-Sγ1 joints, suggesting proximity to 5′Sμ alone cannot explain the robustness of Sγ1. In addition to orientation, the relative frequency of Sμ-Sγ1 joints remain at ~50% of all IgH joins in all cNHEJ-deficient cells, while Sμ-Sε joints diminished ~10-fold. The length of the Sγ1 region (12.5 kb in 129SV strain), the dense RGYW/AGCT sequence motifs (Fig. 3), and the sequence-independent 3D interactions (synapses) between Sμ-Sγ1 (35–38) could all contribute to this sequence independent, robust switching to Sγ1.

HTGTS also allowed us to analyze SHM patterns at the 5′Sμ bait site. The estimated error rate of the Illumina Miseq platform is 0.1%, which is lower than the estimated and reported SHM rate (39). Consistent with previous reports (40, 41), we found that G mutations are nearly two times more frequent than C mutations on the top strand of the 5′Sμ region (the nontemplate strand relative to Sμ germ-line transcription). This enrichment of G mutations is lost in cNHEJ-deficient cells, while the relative pattern of G mutations (to A, C, T) is not affected, suggesting that cNHEJ affects AID targeting to 5′Sμ, but not the processing of the U:G mismatch. The relative abundance of sense vs. antisense transcripts (42), and the preference of APOBEC family enzymes to deaminate the lagging strand during DNA replication (39, 43), could all contribute to this strand bias.

Materials and Methods

Mice. *DNA-PKcs*$^{-/-}$, *DNA-PKcs*$^{KD/KD}$, *Xrcc4*$^{-/-}$, *Tp53*$^{+/-}$, and *HL*$^{k/k}$ mice and cells have been described before (4, 21, 44–48). Mice were housed in a pathogen-free facility and approved by the Institutional Animal Care and Use Committee of Columbia University.

Class Switch Recombination. CSR was performed on 6- to 10-wk-old mice as described (49). Briefly, splenic cells were sorted with CD43 magnetic beads (MACS; Miltenyi), cultured in RPMI medium (Gibco; catalog no. 11875-093) at a density of 1×10^6 cells/mL in medium containing anti-CD40 (1 ng/mL; BD Pharmingen) and IL-4 (20 ng/mL; R&D). Cells were analyzed by flow cytometry. Given the low splenocyte count in 3-wk-old mice, CD43 purification was waived for Tp53$^{-/-}$*DNA-PKcs*$^{KD/KD}$ splenic B cells.

HTGTS. HTGTS was performed as described (20, 28, 29). Genomic DNA was collected from activated B cell after 4 d, was sonicated (Diagenode Bioruptor) and amplified with Sμ-specific biotin primer (5′/5BiosG/CAGACCTGGGAATGTATGGT3′) and nested (5′CACACAAAGACTCTGGACCTC3′) primers. AflII is used to remove germ-line sequence. Since all our experimental mice carry the preassembled IgH on the 129 background (48), we replaced the IgH switch region (from JH4 to the last Cα exon, chr12 114, 494, 415–114, 666, 816) of the C57/BL6-based mm9 with the corresponding region in the AJ851868.3 (GenBank accession no. AJ851868.3) 129 IgH sequence (1415966–1592715) to generate the mm9sr (switch region replacement) genome. Sequences were analyzed as detailed before (20, 28). The best-path searching algorithm (related to YAHA; ref. 50) was used to identify optimal sequence alignments from Bowtie2-reported top alignments (alignment score > 50). The reads are then filtered to exclude mispriming events, germ-line (unmodified) sequence, sequential joints, and duplicated reads. A "duplicate read" is defined by bait and prey alignment coordinates within 2 nt of another read's bait and prey alignments. To plot all of the S-region junctions, including those in the repeats but unequivocally mapped to an individual switch region, we combined the ones filtered by a mappability filter but unequivocally mapped to S regions with "good" reads passing both the mappability (both deduplicated) filters (please see ref. 28 for simulation and details on plotting). MHs are defined as regions of 100% homology between the bait and prey-break site. Insertions

are defined as regions containing nucleotides that map to neither the bait and prey-break site. Blunt junctions are considered to have no MHs or insertions.

Mutational Analysis Calculations. Mutation rate was calculated by custom Excel integrated VBA script. Mutations were determined by comparing the actual bait sequence of each IgH junction with the germ-line bait sequence based on the IgH sequence from 129/Sv strain (GenBank accession no. AJ851868.3). Only true mismatches (no insertion or deletion within 7 nt) were counted as a mutation.

ACKNOWLEDGMENTS. We thank Dr. Jiazhi Hu and the members of the F.W.A. laboratory for their technical assistance and advice for the HTGTS analyses. This work was supported by NIH Grants 5R01CA184187, 5R01CA158073 (to S.Z.), and R01AI077595 (to F.W.A.). J.L.C. is supported by NIH/National Cancer Institute Grant F31CA183504 and Clinical and Translational Science Awards/NIH Grant TL1 TR000082. P.-C.W. is supported by the Charles A. King Trust Postdoctoral Research Fellowship Program. S.Z. is a Leukemia Lymphoma Society Scholar. F.W.A. is an Investigator of the Howard Hughes Medical Institute.

1. Yan CT, et al. (2007) IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* 449:478–482.
2. Boboila C, et al. (2010) Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4. *J Exp Med* 207:417–427.
3. Taccioli GE, et al. (1998) Targeted disruption of the catalytic subunit of the DNA-PK gene in mice confers severe combined immunodeficiency and radiosensitivity. *Immunity* 9:355–366.
4. Gao Y, et al. (1998) A targeted DNA-PKcs-null mutation reveals DNA-PK-independent functions for KU in V(D)J recombination. *Immunity* 9:367–376.
5. Blunt T, et al. (1995) Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine scid mutation. *Cell* 80:813–823.
6. Jiang W, et al. (2015) Differential phosphorylation of DNA-PKcs regulates the interplay between end-processing and end-ligation during nonhomologous end-joining. *Mol Cell* 58:172–185.
7. Bosma GC, et al. (2002) DNA-dependent protein kinase activity is not required for immunoglobulin class switching. *J Exp Med* 196:1483–1495.
8. Manis JP, Dudley D, Kaylor L, Alt FW (2002) IgH class switch recombination to IgG1 in DNA-PKcs-deficient B cells. *Immunity* 16:607–617.
9. Cook AJ, et al. (2003) Reduced switching in SCID B cells is associated with altered somatic mutation of recombined S regions. *J Immunol* 171:6556–6564.
10. Franco S, et al. (2008) DNA-PKcs and Artemis function in the end-joining phase of immunoglobulin heavy chain class switch recombination. *J Exp Med* 205:557–564.
11. Björkman A, et al. (2015) DNA-PKcs is involved in Ig class switch recombination in human B cells. *J Immunol* 195:5608–5615.
12. Symington LS, Gautier J (2011) Double-strand break end resection and repair pathway choice. *Annu Rev Genet* 45:247–271.
13. Mimitou EP, Symington LS (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* 455:770–774.
14. Mimitou EP, Symington LS (2010) Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2. *EMBO J* 29:3358–3369.
15. Callén E, et al. (2009) Essential role for DNA-PKcs in DNA double-strand break repair and apoptosis in ATM-deficient lymphocytes. *Mol Cell* 34:285–297.
16. Zha S, et al. (2011) Ataxia telangiectasia-mutated protein and DNA-dependent protein kinase have complementary V(D)J recombination functions. *Proc Natl Acad Sci USA* 108:2028–2033.
17. Gapud EJ, et al. (2011) Ataxia telangiectasia mutated (Atm) and DNA-PKcs kinases have overlapping activities during chromosomal signal joint formation. *Proc Natl Acad Sci USA* 108:2022–2027.
18. Bhargava R, Carson CR, Lee G, Stark JM (2017) Contribution of canonical non-homologous end joining to chromosomal rearrangements is enhanced by ATM kinase deficiency. *Proc Natl Acad Sci USA* 114:728–733.
19. Zhao Y, et al. (2006) Preclinical evaluation of a potent novel DNA-dependent protein kinase inhibitor NU7441. *Cancer Res* 66:5354–5362.
20. Hu J, et al. (2016) Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat Protoc* 11:853–871.
21. Sonoda E, et al. (1997) B cell development under the condition of allelic inclusion. *Immunity* 6:225–233.
22. Pelanda R, Schaal S, Torres RM, Rajewsky K (1996) A prematurely expressed Ig(kappa) transgene, but not V(kappa)J(kappa) gene segment targeted into the Ig(kappa) locus, can rescue B cell development in lambda5-deficient mice. *Immunity* 5:229–239.
23. Pan-Hammarström Q, et al. (2005) Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. *J Exp Med* 201:189–194.
24. Rucci F, et al. (2010) Homozygous DNA ligase IV R278H mutation in mice leads to leaky SCID and represents a model for human LIG4 syndrome. *Proc Natl Acad Sci USA* 107:3024–3029.
25. Franco S, et al. (2006) H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. *Mol Cell* 21:201–214.
26. Kiefer K, et al. (2007) The catalytic subunit of DNA-protein kinase (DNA-PKcs) is not required for Ig class-switch recombination. *Proc Natl Acad Sci USA* 104:2843–2848.
27. Chiarle R, et al. (2011) Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* 147:107–119.
28. Dong J, et al. (2015) Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. *Nature* 525:134–139.
29. Panchakshari RA, et al. (2018) DNA double-strand break response factors influence end-joining features of IgH class switch and general translocation junctions. *Proc Natl Acad Sci USA* 115:762–767.
30. Zhang Y, et al. (2012) Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* 148:908–921.
31. Reynaud S, et al. (2005) Interallelic class switch recombination contributes significantly to class switching in mouse B cells. *J Immunol* 174:6176–6183.
32. Wang W, Daley JM, Kwon Y, Krasner DS, Sung P (2017) Plasticity of the Mre11-Rad50-Xrs2-Sae2 nuclease ensemble in the processing of DNA-bound obstacles. *Genes Dev* 31:2331–2336.
33. Reginato G, Cannavo E, Cejka P (2017) Physiological protein blocks direct the Mre11-Rad50-Xrs2 and Sae2 nuclease complex to initiate DNA end resection. *Genes Dev* 31:2325–2330.
34. Cheng HL, et al. (2009) Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. *Proc Natl Acad Sci USA* 106:2717–2722.
35. Wuerffel R, et al. (2007) S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* 27:711–722.
36. Jhunjhunwala S, van Zelm MC, Peak MM, Murre C (2009) Chromatin architecture and the generation of antigen receptor diversity. *Cell* 138:435–448.
37. Zarrin AA, Tian M, Wang J, Borjeson T, Alt FW (2005) Influence of switch region length on immunoglobulin class switch recombination. *Proc Natl Acad Sci USA* 102:2466–2470.
38. Zarrin AA, et al. (2004) An evolutionarily conserved target motif for immunoglobulin class-switch recombination. *Nat Immunol* 5:1275–1281.
39. Haradhvala NJ, et al. (2016) Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell* 164:538–549.
40. Neuberger MS, Rada C (2007) Somatic hypermutation: Activation-induced deaminase for C/G followed by polymerase eta for A/T. *J Exp Med* 204:7–10.
41. Xue K, Rada C, Neuberger MS (2006) The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in msh2-/- ung-/- mice. *J Exp Med* 203:2085–2094.
42. Lim J, et al. (2017) Nuclear proximity of Mtr4 to RNA exosome restricts DNA mutational asymmetry. *Cell* 169:523–537.e15.
43. Hoopes JI, et al. (2016) APOBEC3A and APOBEC3B preferentially deaminate the lagging strand template during DNA replication. *Cell Rep* 14:1273–1282.
44. Gao Y, et al. (1998) A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 95:891–902.
45. Gu Y, et al. (1997) Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity* 7:653–665.
46. Jacks T, et al. (1994) Tumor spectrum analysis in p53-mutant mice. *Curr Biol* 4:1–7.
47. Nussenzweig A, et al. (1996) Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* 382:551–555.
48. Pelanda R, et al. (1997) Receptor editing in a transgenic mouse model: Site, efficiency, and role in B cell tolerance and antibody diversification. *Immunity* 7:765–775.
49. Liu X, Shao Z, Jiang W, Lee BJ, Zha S (2017) PAXX promotes KU accumulation at DNA breaks and is essential for end-joining in XLF-deficient mice. *Nat Commun* 8:13816.
50. Faust GG, Hall IM (2012) YAHA: Fast and flexible long-read alignment with optimal breakpoint detection. *Bioinformatics* 28:2417–2424.