HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2020 April 30.

Published in final edited form as:

Nature. 2019 November ; 575(7782): 385–389. doi:10.1038/s41586-019-1723-0.

Fundamental Roles of Chromatin Loop Extrusion in Antibody Class Switching

Xuefei Zhang[1], **Yu Zhang**[1,2], **Zhaoqing Ba**[1], **Nia Kyritsis**[1], **Rafael Casellas**[3], **Frederick W. Alt**[1,#]

[1]Howard Hughes Medical Institute, Program in Cellular and Molecular Medicine, Boston Children's Hospital, and Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

[2]Center for Immunobiology, Department of Biomedical Sciences, Western Michigan University Homer Stryker M.D. School of Medicine, Kalamazoo, MI 49007, USA

[3]Lymphocyte Nuclear Biology, NIAMS, Center of Cancer Research, NCI, NIH, Bethesda, MD 20892, USA

Abstract

Antibody class switch recombination (CSR) in B lymphocytes replaces immunoglobulin heavy chain locus (*Igh*) Cμconstant region exons (C_Hs) with one of 6 C_Hs lying 100-200kb downstream[1]. Each C_H is flanked upstream by an I-promoter and long repetitive switch (S) region[1]. Cytokines/activators induce Activation-Induced Cytidine Deaminase (AID)[2] and I-promoter transcription, with 3'IgH regulatory region (3'IgHRR) enhancers controlling the latter via I-promoter competition for long-range 3'IgHRR interactions[3-8]. Transcription through donor Sμ and an activated downstream acceptor S region targets AID-generated deamination lesions at, potentially, any of 100s of individual S region deamination motifs[9-11]. General DNA repair pathways convert these lesions to DSBs and join an Sμ upstream DSB-end to an acceptor S region downstream DSB-end for deletional CSR[12]. AID-initiated DSBs at targets spread across activated S regions routinely participate in such deletional CSR joining[11]. Here, we report that chromatin loop

[#]Address Correspondence to Alt@enders.tch.harvard.edu.

AUTHOR CONTRIBUTIONS

X.Z., and F.W.A. designed the study; X.Z. performed all experiments; Y.Z., Z.B., and R.C. provided reagents and advice; N.K. designed some of the bioinformatics pipelines; X.Z., and F.W.A analyzed and interpreted data; X.Z., and F.W.A. designed figures and wrote the manuscript; Y.Z., Z.B., and R.C. helped polish the manuscript.

The authors declare no competing financial interests.

Data availability

CSR-HTGTS-Seq, 3C-HTGTS, GRO-Seq and ChIP-Seq sequencing data analyzed here has been deposited in the GEO database under the accession number GSE130270. Specifically, the GEO accession number for Fig. 1d, 1e, 2d, 2e, 2f, 3g, 4e, 4f, 4g, and Extended Date Fig. 2b, 2c, 2d, 2e, 5a, 5b, 6a, 6b, 6c, 7b, 7c, 8d, 10a, 10b, 10c is GSE130263. The GEO accession number for Fig. 1c, 2c, 4h and Extended Date Fig. 2a, 4a, 4b, 10d is GSE130266. The GEO accession number for Fig. 2b, 3c, 3d, 4b, 4c, 4d and Extended Date Fig. 3f, 8b, 9a, 9b is GSE130265. The GEO accession number for Fig. 1f, 1g, 3e and Extended Date Fig. 2f, 2g, 5c, 5d, 7a is GSE130264.

Code availability

The CSR-HTGTS-Seq and 3C-HTGTS pipelines used in this study had been published by our lab (http://robinmeyers.github.io/transloc_pipeline/). The Bowtie2 v2.2.8 used in this study has been published (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml). MACS(2.1.0) used in this study has been published (http://github.com/taoliu/MACS).

extrusion underlies the "unprecedented" mechanism[11] by which IgH organization *in cis* promotes deletional CSR. In naive B cells, loop extrusion dynamically juxtaposes 3'IgHRR enhancers with the 200kb upstream Sμ to generate a CSR center ("CSRC"). In CSR-activated primary B cells, I-promoter transcription activates cohesin loading, leading to generation of dynamic sub-domains that directionally align a downstream S region with Sμ for deletional CSR. During constitutive Sα CSR in CH12F3 B lymphoma cells, inversional CSR can be activated by insertion of a CTCF-binding element ("CBE")-based impediment in the extrusion path. CBE insertion also inactivates upstream S region CSR, while converting adjacent downstream sequences into an ectopic S region by, respectively, inhibiting or promoting their dynamic alignment with Sμ in the CSRC. Our findings suggest that, in a CSRC, dynamically impeded cohesin-mediated loop extrusion juxtaposes proper ends of AID-initiated donor and acceptor S region DSBs for deletional CSR. Such a mechanism might also contribute to pathogenic DSB joining genome-wide.

Treating resting B cells with αCD40/IL4 induces Iγ1 plus Iε promoter transcription and CSR to Sγ1 and Sε (Fig. 1a). To test a transcription-influenced chromatin loop extrusion CSR mechanism (Extended Data Fig. 1; Supplementary Video 1), we first used GRO-Seq to assess transcription through the C_H-containing *Igh* sub-domain in resting and αCD40/IL4-stimulated splenic B cells. All GRO-Seq, as well as 3C-HTGTS, and ChIP-Seq studies, were done in an AID-deficient background to obviate confounding effects of CSR-related genomic rearrangements. In resting B cells, robust sense/anti-sense transcription occurred at the iEμ/Iμ locale with sense transcription continuing through Sμ-Cδ, and also within the 3'IgHRR, most notably at the HS1,2, HS3b, and HS4 enhancers; however, there was little transcription across the 150kb intervening C_H-containing sequence (Fig. 1b, c upper; Extended Data Fig. 2a). In αCD40/IL4-stimulated B cells, substantial transcription was induced across Iγ1-Sγ1 and, to a lesser extent, Iε-Sε locales; but, unexpectedly, transcription across the iEμ-Cδ and 3'IgHRR locales was reduced (Fig. 1c, bottom; Extended Data Fig. 2a).

In resting B cells, high resolution 3C-HTGTS[13] with an iEμ/Iμ bait revealed broad interactions with transcribed downstream sequences, including Sμ, 3'IgHRR HS3a, HS1,2, HS4 and 3'CBEs; but, corresponding to transcription, little interaction with intervening non-transcribed C_H-containing sequences (Fig. 1d, upper; Extended Data Fig. 2b). Likewise, 3'IgHRR HS4 had broad interactions with 3'IgHRR and iEμ-Sμ locales, but largely lacked interactions with intervening C_H-containing sequences (Fig. 1e, upper; Extended Data Fig. 2c). Interaction patterns of HS4 with other 3'IgHRR enhancers suggest that internal extrusions frequently synapse these enhancers and proximal 3'CBEs, which could facilitate combined interactions with upstream sequences in the CSRC (Fig. 1e; Extended Data Fig. 1a, 2c). In αCD40/IL4-stimulated B cells, iEμ/Iμ and HS4 had similar interactions across transcribed Iμ-Cδ and 3'IgHRR sequences as in non-stimulated B cells, but gained broad interactions with transcribed Iγ1-Cγ1 and Iε-Cε locales that peaked over S regions (Fig. 1d, e, bottom; Extended Data Fig. 2b, c). Such interactions with regions tens of kb in length likely reflect combined interactions in all single cells in which impeded sequence extrusion across synapsed regions progressed varying distances in different cells. We refer to such interaction patterns as "dynamic"[14], consistent with interpretation of related Hi-C patterns ("stripes")[15]. In stimulated B cells, the "first" of 10 3'CBEs had similar upstream

interactions as flanking HS4; while the 4.3kb downstream "seventh" 3'CBE had modest interactions with upstream C_H-containing sequences (Extended Data Fig. 2d, e). Overall, these data implicate transcribed Iμ-Sμ, Iγ1-Sγ1, Iε-Sε, and 3'IgHRR sequences as being dynamically aligned in the CSRC by impeding loop extrusion between them, perhaps with 3'CBEs contributions (Extended Data Fig. 1a-i).

In resting B cells, ChIP-Seq revealed accumulation of NIPBL, a cohesin loader, across transcribed Iμ-Sμ and 3'IgHRR HS1,2, HS3b, and HS4 enhancers, along with accumulation at HS7 in the 3'CBEs (Fig. 1f; Extended Data Fig. 2f). Correspondingly, Rad21, a cohesin sub-unit, markedly accumulated at iEμ/Iμ (and Sμ-Cμ), the 3'IgHRR enhancers, a cryptic Cα CBE[16] and at 3'CBEs (Fig. 1g; Extended Data Fig. 2g). In αCD40/IL4-stimulated B cells, NIPBL accumulated at the same locales as in resting B cells, plus new accumulations at Sγ1 and Sε locales (Fig. 1f; Extended Data Fig. 2f). Notably, Rad21 also accumulated at Sγ1 and Sε locales; but, corresponding to transcription patterns, had decreased accumulation at iEμ-Sμ and 3'IgHRR enhancers (Fig. 1g; Extended Data Fig. 2g). These findings are consistent with cohesin-loading at transcriptionally-activated S region locales contributing to ongoing 3'IgH domain extrusion that synapse S regions for CSR (Extended Data Fig.1h-i). The intriguing decreases in transcription and cohesin accumulation at iEμ and 3'IgHRR might reflect competition for enhancer activities by synapsed, activated I promoters (Supplementary Discussion). Overall, our GRO-Seq, 3C-HTGTS and ChIP-Seq studies suggest that linear competition of I promoters for transcriptional activation via the 3'IgHRR[3-5] occurs via loop-extrusion and further imply that transcriptional activation of I promoters generates impediments to induced internal extrusions within the "basal" 3'IgH sub-loop that promote directional alignment, *in cis*, of Sμ and transcribed acceptor S regions within the CSRC (Extended Data Fig. 1f-i).

Upon αCD40/IL4/TGFβ activation, CH12F3 B lymphoma cells undergo CSR between Sμ and Sα (Extended Data Fig. 3a)[17]. The mechanism of exclusive CH12F3 Sα CSR has been elusive. To employ CH12F3 cells for further mechanistic studies, we generated sub-clones lacking the non-coding allele C_H domain to focus assays *in cis* on the productive allele ("CH12F3NCΔ"; Extended Data Fig. 3b, c)[11]. CSR-HTGTS-Seq[11] confirmed exclusive, predominantly deletional Sα CSR in CH12F3NCΔ cells (Fig. 2a, b upper; Extended Data Fig. 3f). GRO-Seq analyses of non-activated and activated AID-deficient CH12F3NCΔ lines (Extended Data Fig. 3d, e) revealed, in both, transcription of Iμ-Cδ and 3'IgHRR regions (with major sense peaks over HS3a, HS1,2 and HS4), constitutive Iα transcription through Sα-Cα, and little transcription of the 150kb intervening region (Fig. 2c, upper, middle; Extended Data Fig. 4a-c). Relative to activated primary B cells (Fig 1. c), 3'IgHRR transcription was more robust, had an additional enhancer peak (HS3a), and extended 24kb downstream (Fig. 2c, upper and middle; Extended Data Fig. 4a, b), suggesting constitutive CH12F3 Iα transcription is driven by ectopic 3'IgHRR activation. With or without activation, AID-deficient CH12F3NCΔ cells, had broad, dynamic iEμ/Iμ and HS4 interactions with transcribed Iμ-Cμ, 3'IgHRR, and Iα-Cα locales (peaking over Sα), but limited interactions with non-transcribed sequences between Iμ-Cμ and 3'IgHRR or transcribed sequences downstream of 3'CBEs (Fig. 2d, e, upper and middle; Extended Data Fig. 5a, b). Highest iEμ/Iμ interactions occurred with transcribed Sα, HS3a, HS1,2, HS4, and proximal 3'CBEs; while highest HS4 interactions were with iEμ/Iμ. ChIP-Seq revealed

enriched NIPBL and cohesin binding in CH12F3NCΔ at the same locales as in resting B cells, with key differences being NIPBL accumulation at transcribed HS3a, both NIPBL and cohesin accumulation at Iα, and cohesin accumulation across the Iα-HS4 locale (Extended Data Fig. 5c, d). These CH12F3NCΔ findings are consistent with constitutive Iα transcription leading to extrusion-based synapsis of Sα and Sμ to form a constitutive Sα-containing CSRC without external stimulation. Activation then induces AID, leading to Sα CSR by a mechanism related to that of primary B cell Sγ1 CSR (Extended Data Fig. 5e).

To assess potential roles of Iα transcription in Sα CSR beyond AID-targeting, we deleted Iα, including the promoter, from CH12F3NCΔ cells to generate IαΔ cells. As anticipated Sα CSR was abrogated in IαΔ cells, but surprisingly, moderate CSR to Sγ3 and low-level CSR to Sγ2b and Sγ2a was activated (Fig. 2b, bottom; Extended Data Fig. 3f, g). To address mechanism, we performed GRO-Seq in both activated and non-activated IαΔ cells which confirmed abrogation of Iα-Cα transcription, but revealed low-level anti-sense transcription across the upstream C_H-containing domain and moderately activated sense transcription across Iγ3-Cγ3 and, at lower levels, γ2b and γ2a locales (Fig. 2c, bottom; Extended Data Fig. 4a-d). In both, Iα-deletion eliminated iEμ/Iμ, HS4 and, therefore, CSRC interactions with Iα-Cα, leading to significantly increased interactions of these locales with the long upstream region between Cδ and Sα (Fig. 2d-f; Extended Data Fig. 6a-c). Together, these findings indicate that elimination of Iα transcription-mediated domination of CSRC interactions allows I promoters within the upstream C_H-containing region extruded past the CSRC to be transcriptionally activated, at least modestly, by dynamic CSRC enhancer interactions, targeting AID and significant CSR to their corresponding S regions (Extended Data Fig. 5e). Activation of both anti-sense and sense transcription across upstream C_H sequences may also contribute to AID-targeting by generating convergent S region transcription[18].

Insertion of CBEs between Sα and Sμ in CH12F3NCΔ cells would be predicted to impact loop extrusion-mediated CSR[13]. Indeed, insertion of three sequential CBEs (“i3CBEs”) downstream of Cγ2α in a convergent orientation to the 3’CBEs reduced Sα CSR to 65% of controls, while, strikingly, increasing inversional Sα to Sμ joining to approximately 20% (Fig. 3a-d; Extended Data Fig. 8a, b). ChIP-Seq and 3C-HTGTS revealed that the i3CBEs insertion-site accumulated Rad21 and gained interactions with 3’CBEs and other extrusion impediment locales, including iEμ and Iα-Cα (Fig. 3e, f; Extended Data Fig. 7a, b). While iEμ/Iμ interacted with i3CBEs, it also interacted with Iα-Cα, 3’IgHRR, and 3’CBEs downstream of i3CBEs site (Fig. 3g; Extended Data Fig. 7c), consistent with i3CBEs impeding, but not blocking, loop extrusion, as occurs for certain other inserted or endogenous CBEs[13,19-21,]. Thus, the i3CBEs-inserted CH12F3NCΔ cell population likely comprises cells with CSRCs containing activated Sα and Sμ directly synapsed for deletional CSR and cells with activated Sα and Su in close proximity, but not directly synapsed due to the i3CBEs impediment (Extended Data Fig. 7d). In the latter, increased inversional joining may result from Sα DSB ends gaining access to both upstream and downstream Sμ DSB ends via diffusion-related mechanisms[11,22-24]. Deletion of 3’CBEs in i3CBEs-inserted CH12F3NCΔ cells did not affect CSR patterns (Extended Data Fig. 8a, b), potentially because requisite 3’CBEs interactions are replaced by interactions with downstream transcribed non-CBE sequences (Fig. 2c) or convergent CBEs (Extended Data Fig. 8c, d).

We examined impact of the i3CBEs on loop extrusion-mediated CSR in IαΔ-CH12F3NCΔ in which Iα-promoter domination of CSRC interactions is abrogated (Fig. 4a). Remarkably, the i3CBEs activated "CSR" of Sμ DSBs to low density AID deamination targets in non-S region sequences immediately downstream of the i3CBE insertion site with 20% of junctions being inversional (Fig 4b-d; Extended Data Fig. 9a-c), implicating a CSR mechanism similar to that of Sα in i3CBEs-inserted CH12F3NCΔ cells (Extended Data Fig. 7d). In i3CBEs-inserted IαΔ cells, general interactions of i3CBEs with sequences in the 3'IgH domain were similar to those of i3CBE-inserted CH12F3NCΔ cells, except for expected lack of interactions with Iα-Cα and notable gain of interactions with HS4 (Fig. 4e; Extended Data Fig. 10a). In addition, iEμ/Iμ and HS4 both interacted broadly across the ectopic S region, while maintaining interactions with each other and adjacent sequences, consistent with frequent combined location of these sequences in a dynamic CSRC (Fig. 4f, g. Extended Data Fig. 10b, c). The i3CBEs also consistently activated downstream "sense" transcription from the insertion locale and upstream antisense transcription from the Sε locale, with the latter extending through the C_H-containing region (Fig. 4h; Extended Data Fig. 10d). This ectopic transcription may be driven by the observed 3'IgHRR (HS4) interaction with the i3CBEs locale in the absence of Iα promoter competition, which may contribute to extrusion-based alignment of the ectopic S region with Sμ in the CSRC. Combined sense/anti-sense transcription may promote AID access to the synapsed ectopic S region via convergent transcription[18]. Notably, i3CBEs abrogated sense transcription of and CSR to upstream C_Hs in IαΔ cells (Fig. 4b, h; Extended Data Fig. 9a, b, d; 10d). Correspondingly, iEμ and, particularly, HS4 had substantially dampened interactions with this entire upstream C_H-containing region (Fig. 4f, g; Extended Data Fig. 10b, c), possibly due to robust ectopic transcription initiation from the i3CBEs site competing for enhancer interactions in the CSRC.

Our findings support a cohesin-mediated chromatin loop extrusion model that addresses unanswered questions regarding the enigmatic CSR mechanism (Extended Data Fig. 1; Supplementary Video 1; Supplementary Discussion). In primary B cells, iEμ and 3'IgHRR enhancers, which are cohesin loading sites, dynamically impede loop extrusion which, thereby, leads to their juxtaposition along with Iμ-Sμ to form a CSRC. CSR-activation then primes the I-promoter of a targeted acceptor S region, which becomes highly transcribed when associated with CSRC enhancers via ongoing loop extrusion. Hi-level transcription promotes cohesin loading and additional extrusions for synapsis with Sμ. CSR-activation also induces AID expression, which may target constitutively transcribed Sμ before synapsis causing frequent internal deletions[25]. Downstream S regions become robust targets mainly upon transcriptional activation in the CSRC where their DSBs are aligned for deletional joining to Sμ DSBs, explaining the recognized paucity of internal deletions in downstream S regions[26] (Supplementary Discussion). This general CSRC model provides an explanation for how Sμ and acceptor S region DSBs are properly synapsed in time and space for deletional-orientation joining. Thus, one or both ends of a given synapsed S region DSB could be extruded into an associated cohesin ring, with extrusion stalling when the end reaches the ring. Then, ends of a DSB in the second synapsed S region could be similarly extruded into cohesin ring(s), aligning donor and acceptor DSB ends for deletional joining. This model is consistent with cohesin accumulation at DSBs[27,28]; which previously was

considered to reflect DSB recruitment of cohesin rather than *vice versa*. The model also is consistent with proposed cohesin complex protein roles in CSR end-joining based on knock-down effects or Cornelia de Lange Syndrome mutations[29,30] (Supplementary Discussion). Finally, related DSB joining mechanisms in other extrusion-impeded genomic regions might contribute to pathogenic deletions or expansions.

METHODS

Experimental procedures.

No statistical methods were used to predetermine sample size. Experiments of cell lines and mice were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Reagents.

APC IgM antibody[31] (17-5790-82; eBioscience™), PE IgA antibody[32] (1100-09S; SouthernBiotech), PE IgG3 antibody[33] (1100-09S; SouthernBiotech), and PE IgG2b antibody[34] (1090-09S; SouthernBiotech) were used for flow cytometric analysis. Rad21 antibody (ab992, Abcam) and NIPBL antibody (A301-779A, Bethyl) were used for ChIP-Seq[15]. AID antibody[9], which we had made by Covance company and Actin antibody[35] (3700s, Cell Signaling Technology) were used for western blotting. The pX330 vector (Addgene plasmid ID 42230) was used for constructing CRISPR/Cas9 plasmids. The NlaIII restriction enzyme (NEB, R0125) was used for 3C-HTGTS[13]. All the oligonucleotides were synthesized by Integrated DNA Technologies and listed in Supplementary Information Table 1.

Mice.

We isolated splenic mature B cells from 8-12 weeks old $AID^{-/-}$ C57BL/6 mice[2] for 3C-HTGTS, GRO-Seq and ChIP-Seq assays. For each experiment, we employed at least 3 mice including both males and females and experiments were repeated three times. Mouse work was in compliance with ethical regulation established by the Boston Children's Hospital Institutional Animal Care and Use Committee (IACUC) and Institutional Biosafety Committee (IBC). Mouse work was performed under protocols approved by the Boston Children's Hospital and Rockefeller University Institutional Animal Care and Use Committees.

Cell culture.

Mature splenic B cells isolated from $AID^{-/-}$ C57BL/6 mice[2] via a CD43-negative selection kit (MACS) were cultured in R15 medium (RPMI1640, 10% FBS, L-glutamate, 1 × penicillin and streptomycin), and stimulated with αCD40 (1 μg/ml, eBioscience) plus IL4 (20 ng/ml, PeproTech) for 48 hrs before doing GRO-Seq, 3C-HTGTS and ChIP-seq assays. CH12F3 cells[17] were cultured in medium R15 and stimulated with αCD40 (1 μg/ml, eBioscience), IL4 (20 ng/ml, PeproTech) and TGFβ (0.5 ng/ml, R&D systems). CSR-HTGTS-Seq was performed in AID-proficient CH12F3 cells after stimulation for 72 hrs. GRO-Seq, 3C-HTGTS, and ChIP-Seq were performed in AID-deficient CH12F3 cells after stimulation for 24 hrs. Cells were not test for mycoplasma contamination.

Generation of mutant CH12F3 cell lines.

A Cas9/gRNA approach was employed to generate the various mutant strains used in this study as described previously[36]. A 4D-nucleofector X (Lonza, solution SF, protocol CA-137) was used for nucleofection procedures used to generate all mutant cell lines. Generation of CH12F3NCΔ strains has been described[11]. IαΔ strains were generated by deleting the Iα region from the CH12F3NCΔ strains and confirmed by PCR genotyping and sequencing. The i3CBEs strains were generated by inserting triple CBEs just downstream of Cγ2a in the CH12F3NCΔ lines through an approach involving Cas9/gRNA targeting combined with short single-stranded DNA oligonucleotide (ssODN) templates[37] and confirmed by PCR genotyping and sequencing. The i3CBEs-3'CBEΔ was generated by using a Cas9/gRNA approach to delete the entire 3'CBEs region (9kb) including all 10 3'CBEs from i3CBEs strains and confirmed by genotyping and sequencing. The IαΔ-i3CBEs line was generated by inserting triple CBEs just downstream of Cγ2a from IαΔ strains via Cas9/gRNA targeting combined with short single-stranded DNA oligonucleotide (ssODN) templates[37] and confirmed by PCR genotyping and sequencing. All AID-deficient strains were generated by deleting the *Aicda* gene from the above strains via a Cas9/gRNA approach and confirmed by PCR genotyping, western blotting and flow cytometric analyses after stimulation with αCD40, IL4 and TGFβ for 72 hrs. At least two independent clones were obtained for each derivative mutant genotype. All gRNA oligonucleotides for CRISPR/Cas9 used for the targeting outlined above were cloned into the pX330 vector (Addgene plasmid ID 42230). All the oligonucleotides employed for these experiments are listed in Supplementary Information Table 1.

Flow cytometric analysis.

Flow cytometric analysis was used for measuring IgH class switching in CH12F3 cells stimulated with αCD40/IL4/ TGFβ for 72 hrs. The cells were passaged several times to ensure that they grew well at the time of stimulation. After stimulation, cells were collected and washed once with PBS. Then, the cells were stained for the surface makers with the indicated antibodies (APC-IgM/ PE-IgA; APC-IgM/PE-IgG3; APC-IgM/PE-IgG2b; The APC-IgM antibody was diluted 100 times from stock concentration, while the other antibodies were diluted 200 times) at room temperature for 5 mins. The stained cells were washed once with PBS and resuspended in PBS for flow cytometric analysis with a BDFACSCalibur (BD bioscience). CellQuest Pro alias software was used for collecting the data and FlowJo software (10.0.6) was used for analyzing the data. Live cells were gated from the FSC/SSC gate for further analysis. Then, we gated the $IgM^{-}IgA^{+}$, $IgM^{-}IgG3^{+}$ and $IgM^{-}IgG2b^{+}$ cells respectively to show the percentage of the cell populations. Primary data of the bar graphs for all FACS experiments is Supplementary Information Table 2.

CSR-HTGTS-Seq and data analysis.

CSR-HTGTS-Seq libraries with 5'Sμ bait were prepared from different CH12F3 mutants stimulated with αCD40/IL4/TGFβ for 72 hours as described previously[11]. Briefly, 50 μg gDNA from αCD40/IL4/TGFβ-stimulated CH12F3 cells was sonicated (25s ON and 60s OFF, 2 cycles with low energy input) on Diagenode Bioruptor sonicator. The sonicated DNA segments were amplified by LAM-PCR with biotinylated 5'Sμ primer. The amplified biotin-

labeled LAM-PCR products were enriched with streptavidin C1 beads (Thermo Fisher Scientific, #65001) for 4 hrs at room temperature, followed by adaptor ligation with the following PCR program: 25 ℃ 1 hour, 22 ℃ 3 hours, 16 ℃ overnight. The adaptor-ligated products were subjected to nested-PCR with barcode primers and followed by tag-PCR with P5-I5 and P7-I7 primers. 500-1000bp tag-PCR products were selected by separation on 1% TBE gel. CSR-HTGTS-Seq libraries were sequenced by paired-end 300bp on a Mi-Seq™ (Illumina).

Libraries were processed via our published pipeline[38] and mapped against the mm9 genome or modified-mm9 genomes including three CBEs inserted convergent to 3'CBEs (i3CBEs). Data were analyzed and plotted after removing the duplicates as described[11]. Each experiment was repeated at least three times from at least two independent clones. The junction numbers within the C_H-containing portion of the *Igh* (i.e. from iEμ through the 3'CBEs) and within different S regions, as well as the percentage analysis of different S region junctions with respect to total junctions within the C_H-containing portion of the *Igh* are listed in the Supplementary Information Table 3. Primers used for CSR-HTGTS-Seq are listed in Supplementary Information Table 1.

3C-HTGTS.

3C-HTGTS analyses were performed as previously described[13] on $AID^{-/-}$ mature splenic B cells stimulated with αCD40/IL4 for 48 hrs and with $AID^{-/-}$ CH12F3 cells stimulated with αCD40/IL4/TGFβ for 24 hrs. Briefly, 10 million cells were crosslinked with 2% formaldehyde for 10 minutes at room temperature and quenched with glycine at a final concentration of 125 mM. Then, the crosslinked cells were lysed in the 3C lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1% Triton X-100, protease inhibitors) and nuclei were digested with NlaIII enzyme (NEB, R0125) at 37 ℃ overnight. They were then brought to 16℃, 100U of T4 ligase added (Promega, M1801) and they incubated again overnight at 16℃. The ligated products were de-crosslinked with Proteinase K (Roche, #03115852001) at 56 ℃ overnight and the 3C templates were purified by phenol/chloroform. The 3C-HTGTS libraries were then sequenced by paired end 300 bp or paired end 150 bp sequencing on either a Mi-Seq™ or Next-Seq™550 (Illumina) and data processed as previously described[30]. Three 3C-HTGTS baits (iEμ/Iμ, 3'IgHRR(HS4) and the CBE insertion locale) were used for this study and each experiment was repeated three times from at least two independent clones. Before plotting the data for comparison, 3C-HTGTS libraries were normalized by random selection to the numbers of junctions at bait locales of the smallest library for each set of libraries (Supplementary Information Table 4). For statistical analyses, we counted the number of junctions within the indicated bait-interacting locales for both control and experimental samples. For bar graph presentations in Extended Data Figures 2, 5, and 9, the junction number recovered from control (e.g. non-stimulated primary cells or indicated non-stimulated or genetic mutant CH12F3 cells) samples was normalized to represent 100% and relative experimental values listed as a percentage of the control values. Primers used for 3C-HTGTS are listed in Supplementary Information Table 1.

GRO-Seq analysis.

GRO-Seq libraries were prepared as described previously[39,40] from AID$^{-/-}$ mature splenic B cells stimulated with αCD40/IL4 for 48 hrs or AID$^{-/-}$ CH12F3 cells stimulated with αCD40/IL4/TGFβ for 24 hrs. Briefly, 10 million cells were collected and permeabilized with the buffer (10 mM Tris-HCl pH 7.4, 300 mM sucrose, 10 mM KCl, 5 mM $MgCl_2$, 1 mM EGTA, 0.05% Tween-20, 0.1% NP40 substitute, 0.5 mM DTT, protease inhibitors and Rnase inhibitor). The permeabilized cells were resuspended in 100 μl of storage buffer (10 mM Tris-HCl pH 8.0, 25% (V/V) glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA and 5 mM DTT) for nuclear run-on with 2X run-on mix (5 mM Tris-HCl PH 8.0, 2.5 mM $MgCl_2$, 0.5 mM DTT, 150 mM KCl, 0.5 mM ATP, 0.5 mM CTP, 0.5 mM GTP, 0.5 mM BrUTP, RNase inhibitor, 1% sarkosyl) at 37o for 5 min. RNA was extracted by Trizol and followed by hybrolyzation with NaOH at a final concentration of 0.2 N on ice for 18 min. After quenching with ice-cold Tris-HCl PH6.8 at a final concentration of 0.55 M and exchanging buffer via Bio-Rad P30 columns, the RNA was incubated with Br-dU antibody-conjugated beads (Santa Cruz biotechnology, sc-32323-ac) for 1 hr. The enriched run-on samples were incubated with RppH (NEB, M0356S) and hydroxyl repair with T4 PNK (NEB, M0201S), followed by ligating the 5' and 3' RNA adaptor. RT-PCR was performed from the adaptor-ligated RNA to obtain cDNA. Half of the cDNA was subjected to making GRO-Seq libraries by two rounds of PCR with barcode primers. 200-500 bp products from the first round of PCR were subjected to the second round of PCR with the number of PCR cycles determined by test PCR amplification. The second round of PCR products were size-selected by SPRIselect beads (Beckman Coulter, B23318). GRO-Seq libraries were sequenced via paired end 75 bp sequencing on a Next-Seq™550 and normalized to a coverage of 10 million 100nt reads for display. Relative transcriptional activity of specific regions was calculated as Reads Per Million Reads (RPM). Each experiment was repeated three times from at least two independent clones.

ChIP-Seq.

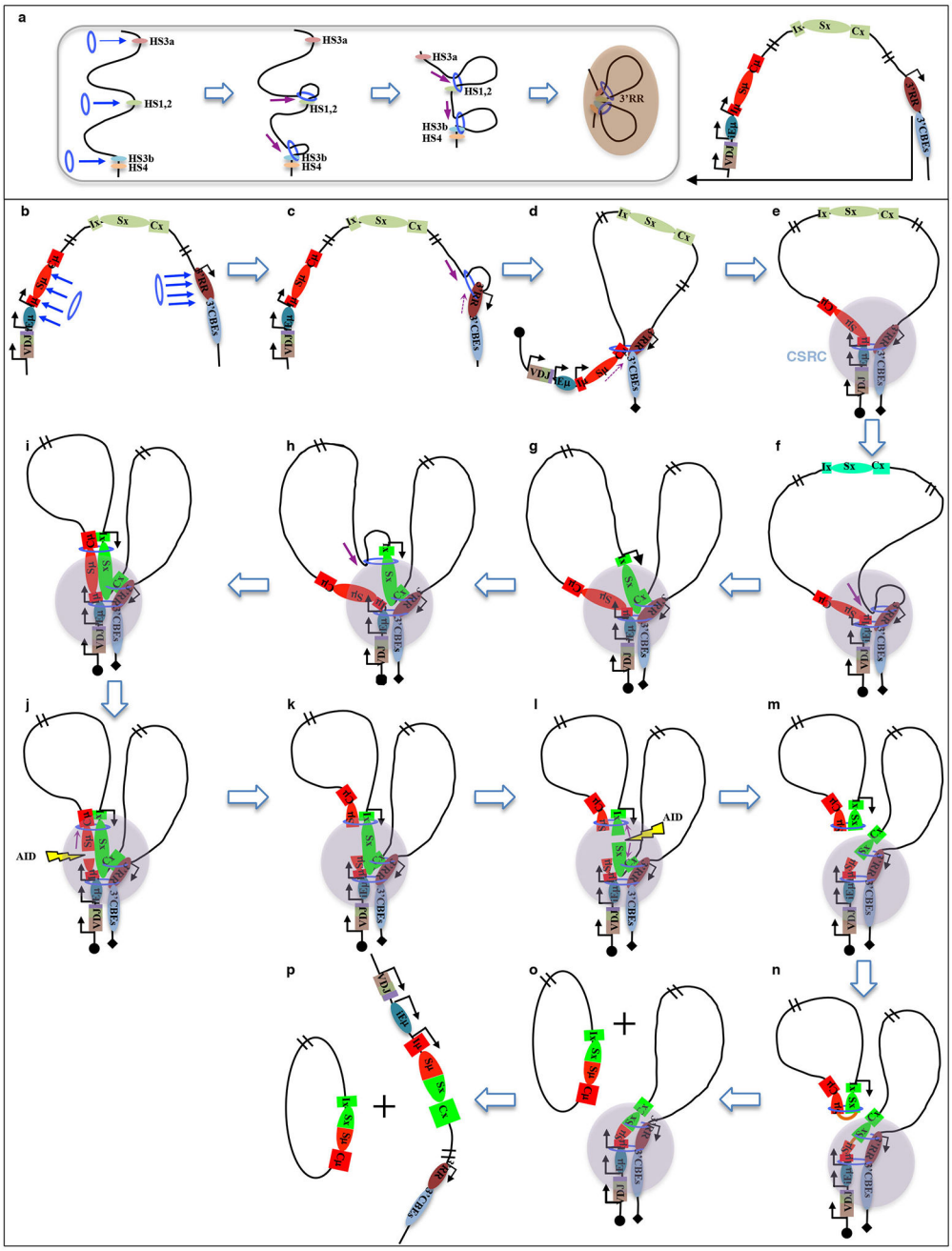
ChIP-Seq was performed as described previously[41] with AID$^{-/-}$ splenic mature B cells stimulated with αCD40/IL4 for 48 hrs or with AID$^{-/-}$ CH12F3 cells stimulated with αCD40/IL4/TGFβ for 24 hrs. Briefly, 20 million cells were crosslinked with 1% formaldehyde for 10 min at room temperature and quenched with glycine at a final concentration of 125 mM. Then, the crosslinked cells were lysed in cell lysis buffer (5mM PIPES pH 8, 85mM KCl, 0.5% NP-40) nuclei collected by centrifugation, followed by lysis of the nuclei in nuclear lysis buffer (50mM Tris-HCl pH 8.1, 10mM EDTA, 1% SDS). The lysed nuclei were sonicated to achieve chromatin segments with an average size of 200-300 bp. The 200-300 bp chromatin segments were pre-cleared with 20 μl Dynabeads protein A (Invitrogen, 10002D) in the total volume of 1.5 ml binding buffer (0.01% SDS, 1% Triton X-100, 16.7 mM Tris-HCl PH8.0, 167 mM NaCl) at 4 ℃ for 2 hrs. 50 μls of pre-cleared samples were collected as input and the rest of the precleared samples was incubated with 5 μg Rad21 (ab992, Abcam) or NIPBL (A301-779A, Bethyl) at 4 ℃ overnight, followed by adding 40 μl Dynabeads protein A for another 2-4 hrs. The enriched segments were washed, eluted and de-crosslinked to obtain the ChIP-Seq templates with the de-crosslinked input as negative control. ChIP-Seq libraries were prepared with Illumina Truseq ChIP Sample Preparation Kit (Illumina, 15032488) and sequenced by paired-end 75 bp sequencing on a

Next-Seq™550. Libraries were normalized to 1 million reads for display. Relative Rad21 and NIPBL accumulation of specific regions was calculated as Reads Per Million Reads (RPM). Shown are three repeats for all of the ChIP-Seq experiments except the two repeats of the NIPBL ChIP-Seq for the CH12F3NCΔ-AID$^{-/-}$ cells with or without αCD40/IL4/TGFβ stimulation.

Quantification and statistical analysis

Statistical analyses for flow cytometric analysis, CSR-HTGTS-Seq, 3C-HTGTS, GRO-Seq and ChIP-Seq were performed via two-tailed, unpaired t test. $P < 0.05$ is considered significant. At least three repeats were done for each statistical analysis. *P* values are shown in the bar graphs in main and Extended Data Figures.

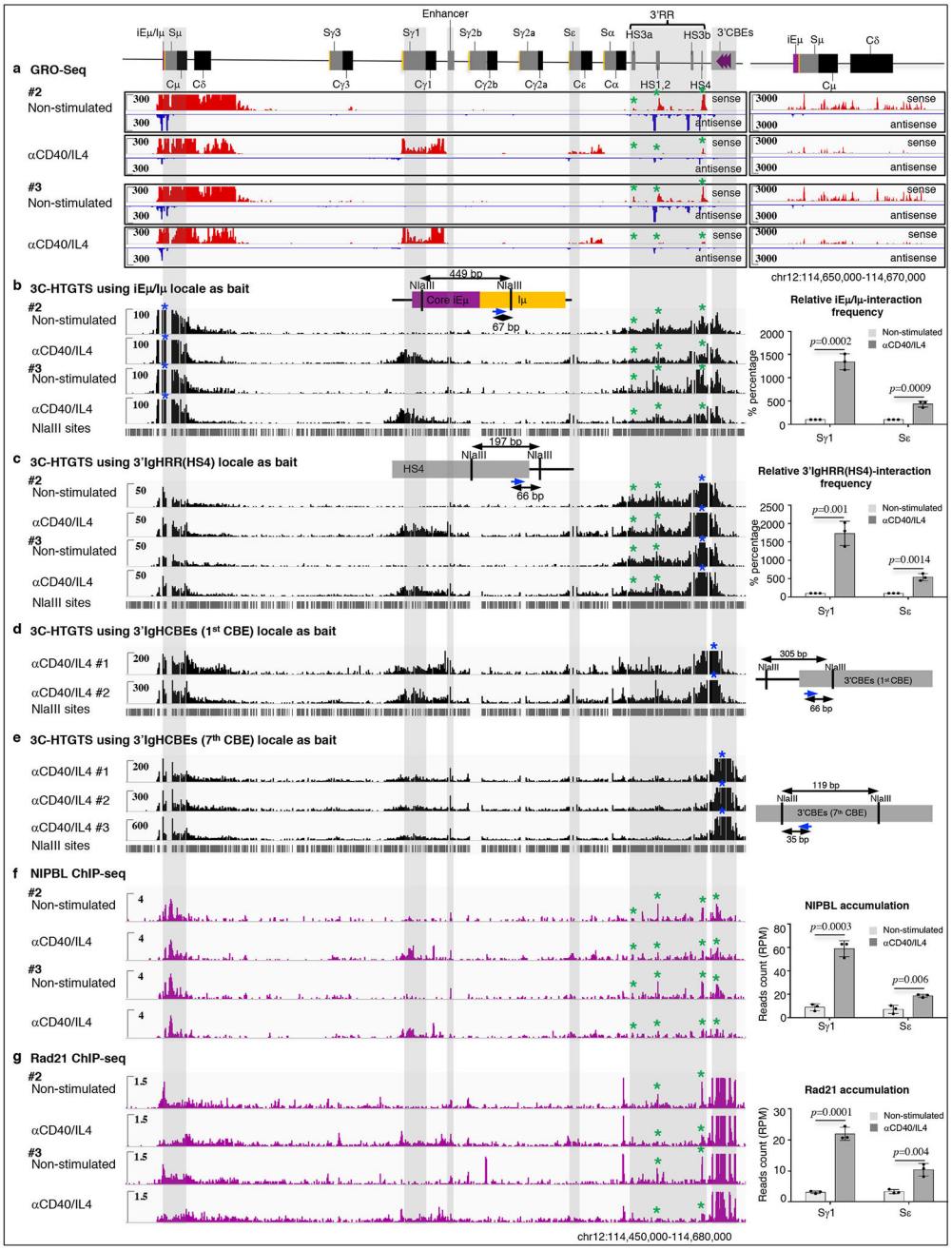
Extended Data



Extended Data Figure 1. Working model for cohesin-mediated chromatin loop extrusion-driven deletional CSR joining.

a, Cohesin (blue rings) loaded at the indicated HS sites within the 3'IgHRR dynamically extrude 3'IgHRR chromatin which aligns the HS sites as transient loop anchors (brown oval). **b-e,** In resting B cells, cohesin is loaded (blue arrows) at either Iμ-Sμ (red rectangles) or the 3'IgHRR. While similar models could be drawn for both, we illustrate one in which loading occurs at 3'IgHRR (brown oval) and downstream extrusion is impeded by 3'IgHRR/3'CBEs chromatin to generate a dynamic impediment for extrusion of upstream chromatin that brings iEμ/Iμ/Sμ into proximity with the 3'IgHRR to generate a CSRC (Grey circle). In

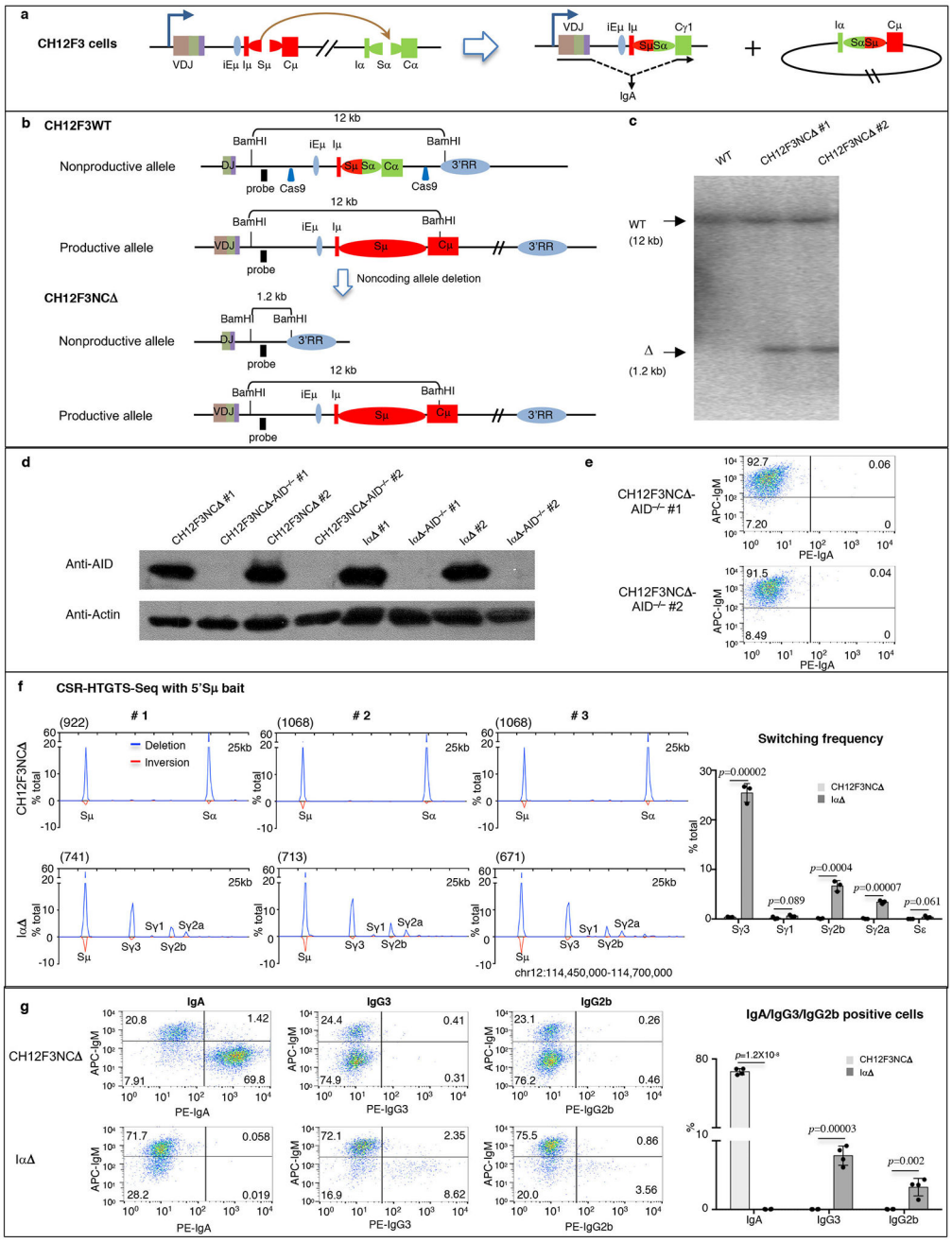
this process, upstream extrusion is strongly impeded at the V(D)J/iEμ locale. **f, g,** B cell activation primes a targeted S region promoter (light green becoming darker green, which is activated for high level transcription (bright green) after extrusion into proximity with the 3'IgHRR. **h, i,** Downstream extrusion of cohesin loaded at the activated S region is impeded by activated S region chromatin allowing extrusion of upstream chromatin to dynamically align targeted S region with Sμ. **j-p,** Activation-induced AID is transcriptionally targeted to Sμ and the activated S region leading to DSBs (lightning bolts) in one or the other and, ultimately, in both. Cohesin-mediated loop extrusion pulls S region DSB ends into cohesin rings stalling extrusion and aligning them for deletional end-joining. DSBs in the Sμ and activated S regions need not occur at the same spatial location or time in this model. See also Text and Supplementary Video 1.



Extended Data Figure 2. Cytokine/activator-induced S region transcription promotes dynamic loop formation and S-S synapsis during CSR.

a, (left) Additional repeats of GRO-Seq profiles shown in Fig. 1c from non-stimulated and αCD40/IL4 stimulated AID$^{-/-}$ mature splenic B cells. (right), Zoom-in view of the GRO-Seq profiles on the left to better reveal the transcription level around the iEμ-Cδ locale from non-stimulated and αCD40/IL4-stimulated AID$^{-/-}$ mature splenic B cells. **b, c,** Additional repeats of 3C-HTGTS profiles shown in Fig. 1d, e from non-stimulated and αCD40/IL4-stimulated AID$^{-/-}$ mature splenic B cells using iEμ/Iμ (**b**) or 3'IgHRR(HS4) (**c**) locale as baits (blue asterisks). Bar graphs on the right of 3C-HTGTS profiling show the relative iEμ/

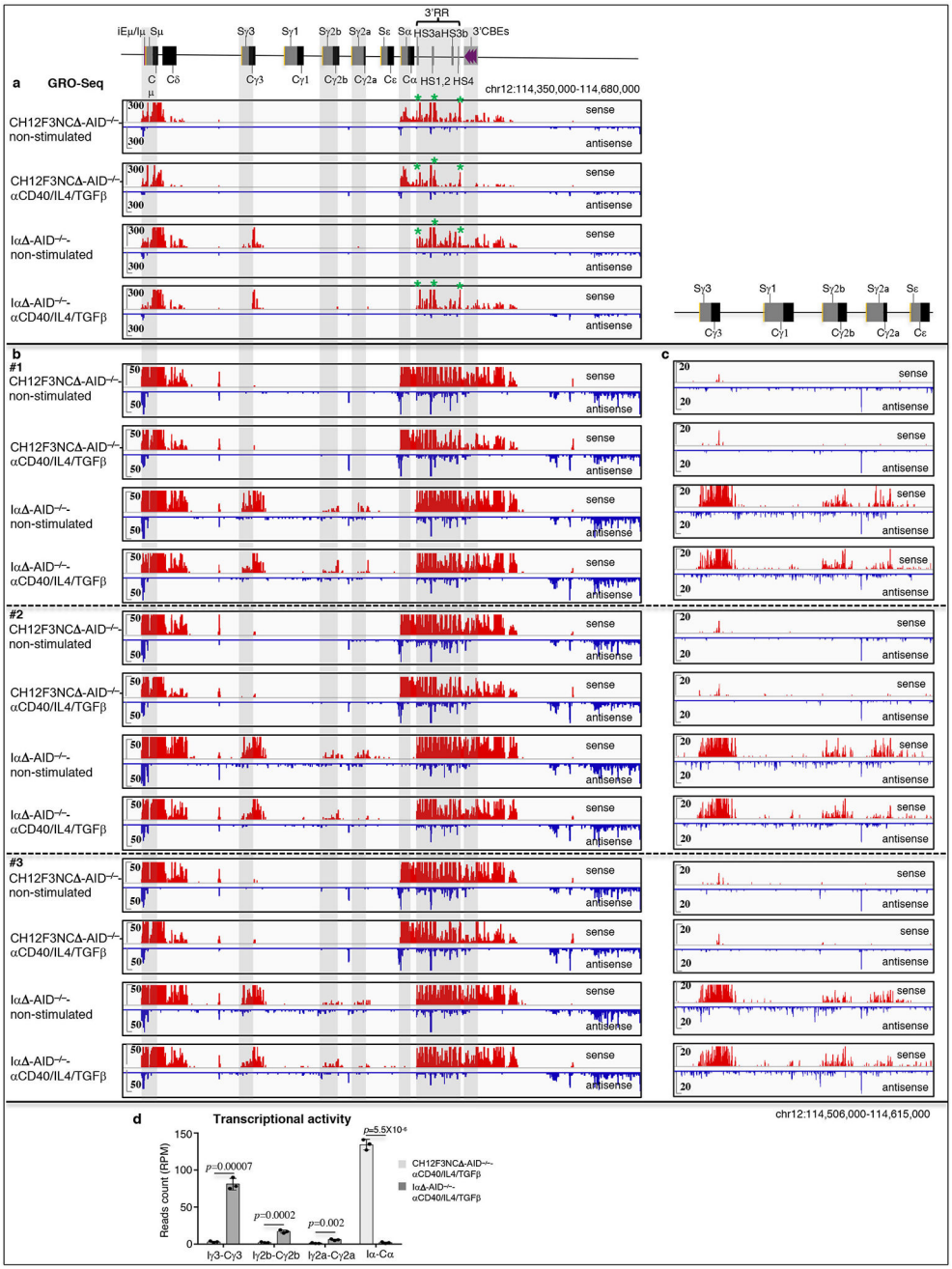
Iμ or 3'IgHRR(HS4) interaction frequency with Sγ1 and Sε in αCD40/IL4-stimulated mature splenic B cells. Diagrams on the top of the 3C-HTGTS profiling show the digestion/bait strategies used for the 3C-HTGTS experiments. **d, e,** 3C-HTGTS profiles of interactions within the 3'IgH locus domain in αCD40/IL4-stimulated $AID^{-/-}$ mature splenic B cells using the first (**d**) and seventh (**e**) 3'CBE locale as baits (blue asterisk). Diagrams on the right of the 3C-HTGTS profiling show the digestion/bait strategies used for the 3C-HTGTS experiments. **f, g,** Additional repeats of NIPBL (**f**) and Rad21 (**g**) ChIP-Seq shown in Fig. 1f, g from non-stimulated and αCD40/IL4 stimulated $AID^{-/-}$ mature splenic B cells. Bar graphs on the right of the ChIP-Seq profiling show NIPBL and Rad21 accumulation of indicated regions. All bar graph data represent mean ± s.d. in panels **b-c** and **f-g** from three biologically independent repeats. *P* values were calculated via an unpaired two-tailed *t*-test. All other bars and symbols are as indicated in Figure 1 legend.



Extended Data Figure 3. Iα deletion promotes CSR to upstream S regions.

a, Illustration of dominant, deletional CSR between Sμ and Sα in CH12F3 cells. **b,** Illustration of Cas9/gRNA targeting (lightning bolts) used to generate the CH12F3NCΔ line. **c,** Southern blot confirmation (using BamHI digestion and a J_H4 probe) of the CH12F3NCΔ lines (done twice independently with similar results). **d,** Western blot confirmation of AID expression or lack of expression, respectively, in AID sufficient and deficient (via targeted deletion) CH12F3NCΔ and IαΔ lines following stimulation with αCD40/IL4/TGFβ for 72 hrs (done twice independently with similar results). **e,** FACS analysis for surface IgA expression in CH12F3NCΔ-AID$^{-/-}$ cells stimulated with αCD40/IL4/TGFβ for 72 hrs (done

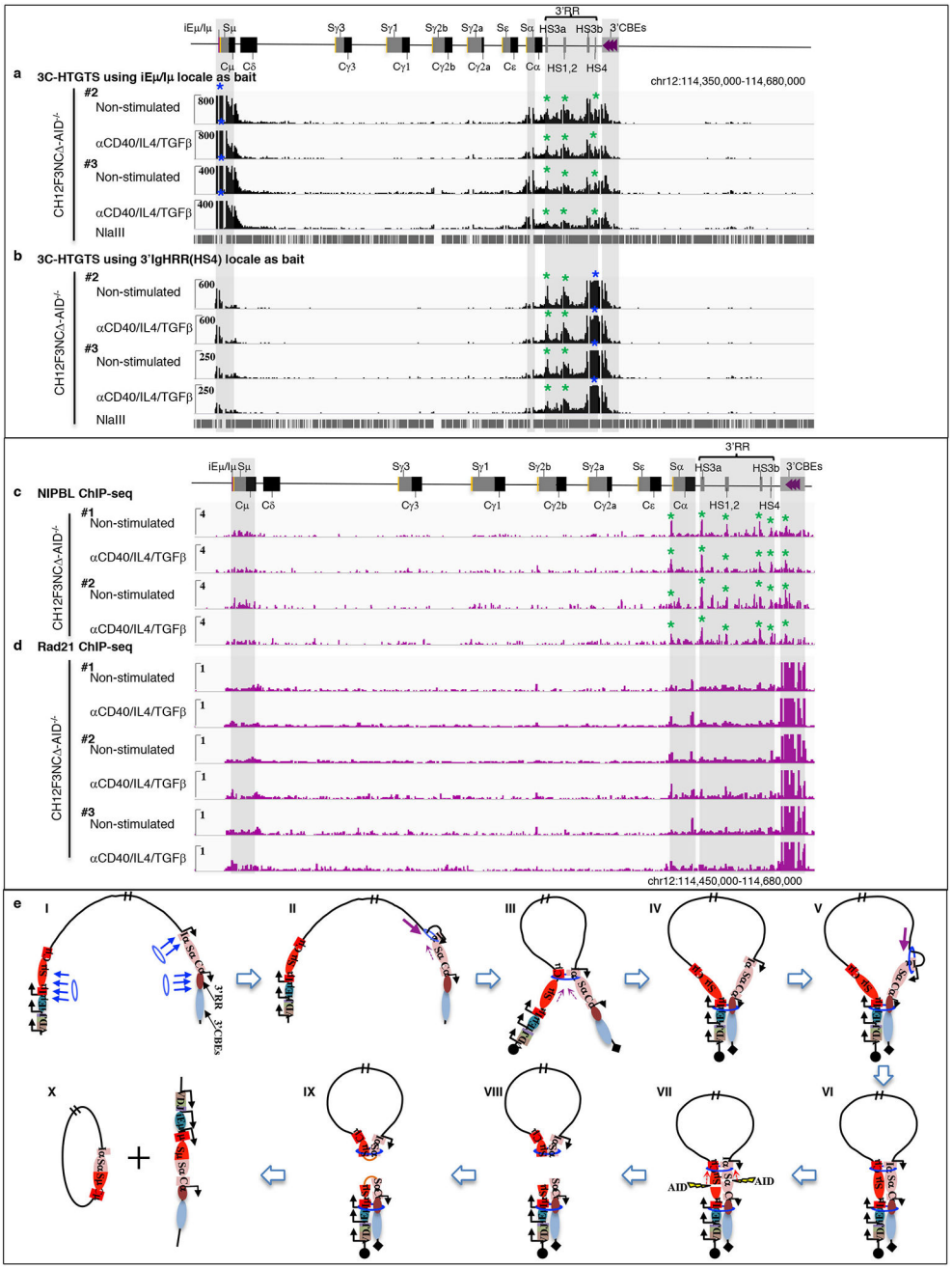
three times independently with similar results). **f,** Three repeats of CSR-HTGTS-Seq data shown in Fig. 2b for αCD40/IL4/TGFβ-stimulated CH12F3NCΔ and IαΔ cells (3 biologically independent repeats). Junctions are plotted at 2.5 kb bin size. The blue lines indicate deletional joining and the red lines indicate inversional joining. Bar graph shows percentages of junctions located in different regions from CH12F3NCΔ and IαΔ cells. Data represents mean ± s.d. from three biologically independent repeats. *P* values were calculated via an unpaired two-tailed *t*-test. **g,** FACS analysis of IgA, IgG3 and IgG2b surface expression in CH12F3NCΔ and IαΔ cells stimulated with αCD40/IL4/TGFβ for 72 hrs (4 biologically independent repeats). Bar graph shows percentages of IgA, IgG3 and IgG2b expression on CH12F3NCΔ and IαΔ cells. Data represents mean ± s.d. from four biologically independent repeats. *P* values were calculated via an unpaired two-tailed *t*-test.



Extended Data Figure 4. Iα deletion promotes transcription to upstream S regions.

a, GRO-Seq profile of repeat #1 (shown immediately below it) with an enlarged scale to allow better comparison of relative transcription levels of different portions of the IgH constant region in CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells with or without αCD40/IL4/TGFβ stimulation (3 biologically independent repeats with similar results). Green asterisks indicate the HS3a, HS1,2 and HS4 sites within 3'IgHRR. **b**, Three repeats of the GRO-Seq profiles with a smaller scale to better reveal low, but significant transcription of Cγ2b and Cγ2a (upon Iα-deletion) in CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells with or without αCD40/IL4/TGFβ stimulation. **c,** Higher zoom-in view of the three repeats of GRO-Seq

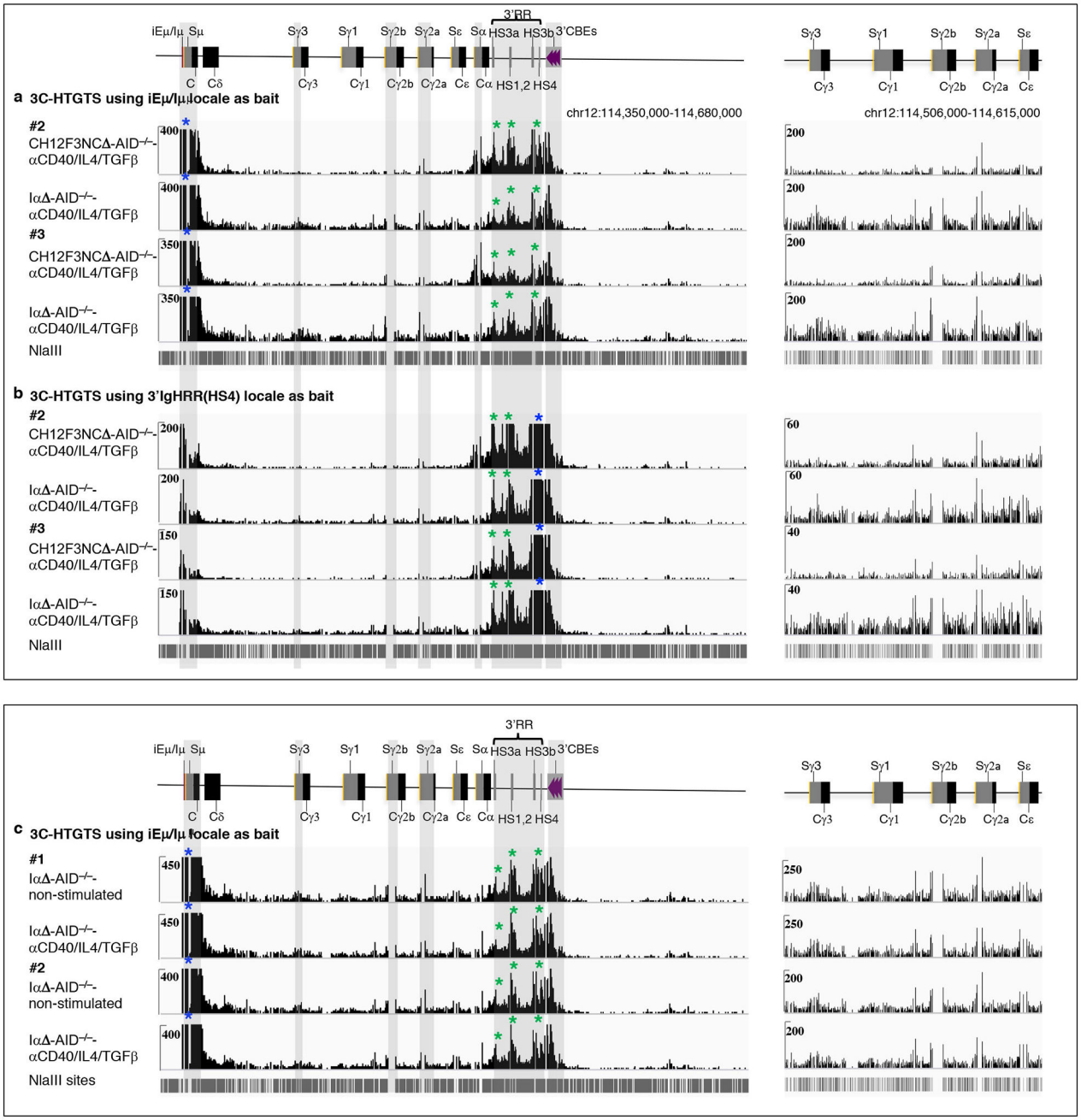
profiles to better reveal induced anti-sense transcription in the region between Sγ3 to Sε in IαΔ-AID$^{-/-}$ versus CH12F3NCΔ-AID$^{-/-}$ cells with or without αCD40/IL4/TGFβ stimulation. **d**, Bar graph shows GRO-Seq transcriptional activity (calculated as RPM) of the different indicated S regions in αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ cells and IαΔ-AID$^{-/-}$ cells (3 biologically independent repeats). Bar graph panel represents mean ± s.d. from three biologically independent repeats. *P* values were calculated via unpaired two-tailed *t*-test. Grey Bars highlight the iEμ-Cμ, Iγ3-Cγ3, Iγ2b-Cγ2b, Iγ2a-Cγ2a, Iα-Cα, 3'IgHRR and 3'CBEs.



Extended Data Figure 5. Constitutively transcribed Sα leads to constitutively synapsis of Sα with Sμ in CH12F3 cells.

a, b, Additional repeats for the 3C-HTGTS profiles shown in Fig. 2d, e, from non-stimulated and αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ cells using iEμ/Iμ (**a**) or HS4 (**b**) locale as baits (blue asterisks). Green asterisks indicate the HS3a, HS1,2 and HS4 sites within 3'IgHRR. Grey Bars highlight the iEμ-Cμ, Sα, 3'IgHRR and 3'CBEs. **c, d,** NIPBL (**c**) and Rad21 (**d**) ChIP-Seq profiles of non-stimulated and αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ cells. Green asterisks indicate the Iα, HS3a, HS1,2, HS3b, HS4 and HS7 sites that were implicated by this experiment as targets for cohesin loading. Grey Bars

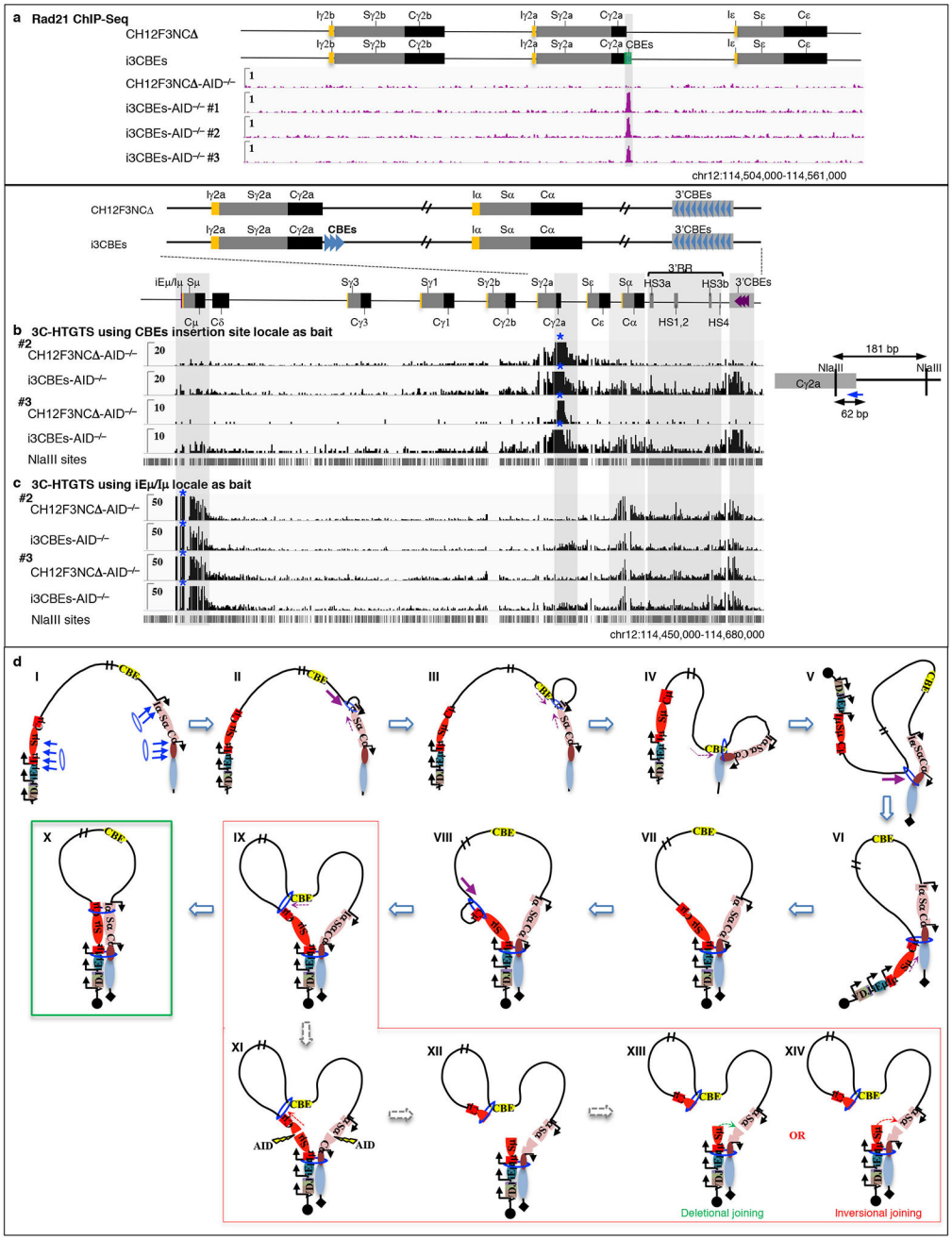
highlight the broader regions around Sμ, Sα, the 3'IgHRR and the 3'CBEs. **e,** Loop extrusion-mediated Sμ-Sα synapsis in CH12F3 cells. **I,** Cohesin is loaded at various *Igh* locations including transcriptionally activated Iα-Sα. **II-IV,** For cohesin loaded at Iα locale downstream extrusion is impeded by transcribed Sα allowing upstream extrusion to proceed until being dynamically impeded by transcribed iEμ-Sμ-Cμ locale resulting in Sμ and Sα being brought into proximity without complete alignment. During upstream extrusion, the activated Iα promoter blocks extrusion-mediated activation of upstream I promoters by the 3'IgHRR via promoter competition. **V, VI,** Continued loading of cohesin at the Iα locale is impeded for downstream extrusion allowing continued upstream extrusion until reaching the transcribed Sμ locale causing dynamic Sα/Sμ synapsis. **VII-X,** Activation-induced AID is transcriptionally targeted to Sμ and the activated Sα leading to DSBs (lightning bolts) in one or the other and, ultimately, in both. Cohesin-mediated loop extrusion pulls S region DSB ends into cohesin rings stalling extrusion and aligning them for deletional end-joining. This model could be explained by other variations including cohesin loading at Sμ or the 3'IgHRR or a process like that in Extended Data Fig. 1.



Extended Data Figure 6. Iα deletion increases iEμ/Iμ and HS4 interactions across the upstream C_H domain.

a, b, (left) Additional repeats for the 3C-HTGTS profiles shown in Fig. 2d, e, from αCD40/IL4/TGFβ-stimulated CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells using iEμ/Iμ (**a**) or HS4 (**b**) locale as baits (blue asterisks). Green asterisks indicate 3'IgH RR HS sites in all panels. Grey Bars highlight the iEμ-Cμ, Sγ3, Sγ2b, Sγ2a, Sα, 3'IgHRR and 3'CBEs. (right), Zoom-in view of the 3C-HTGTS profiles on the left to better reveal the interaction patterns in the region from Iγ3 to Cε in αCD40/IL4/TGFβ-stimulated CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells. **c,** (left) 3C-HTGTS profiles of interactions across the indicated domain of

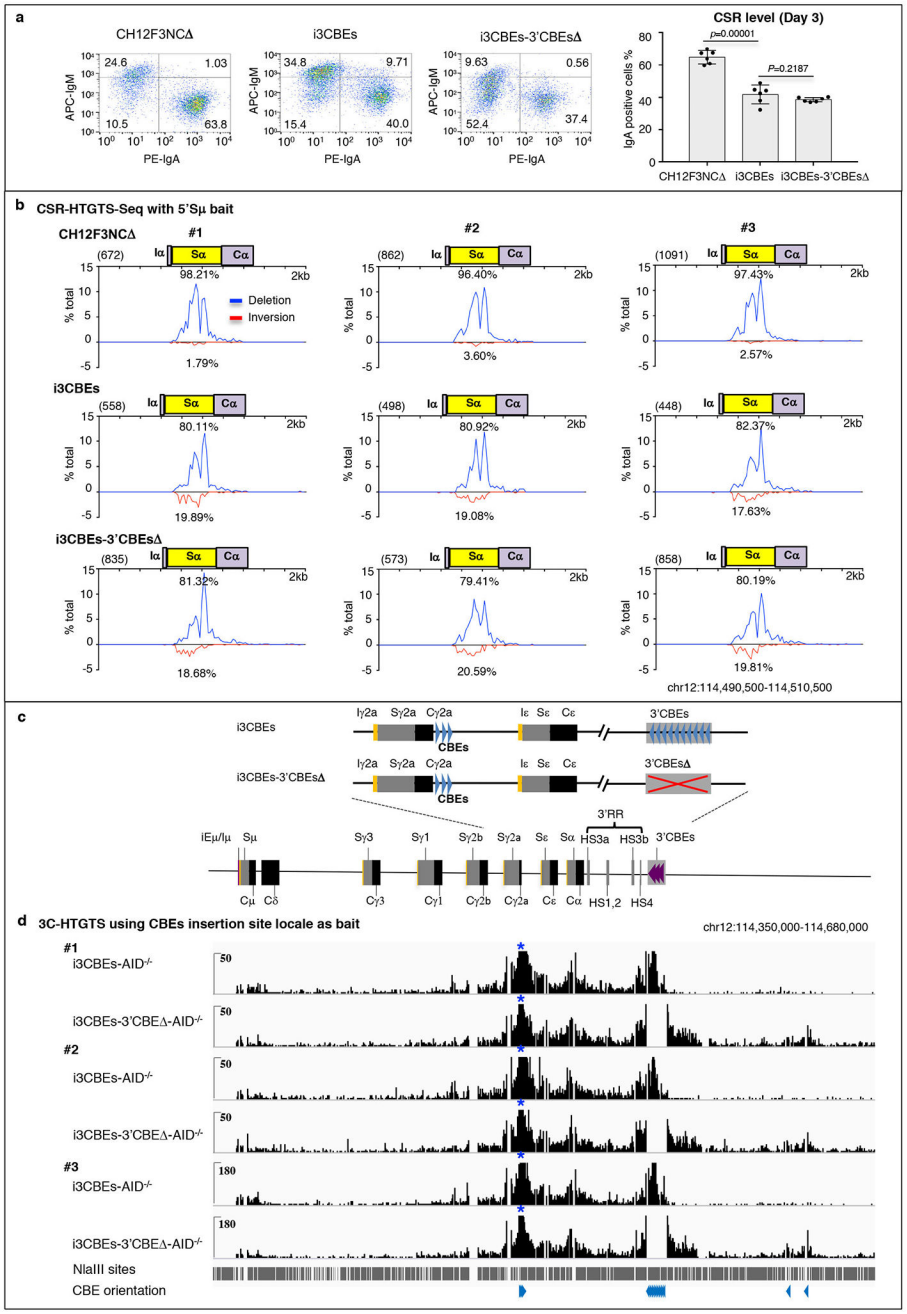
non-stimulated and αCD40/IL4/TGFβ-stimulated IαΔ-AID$^{-/-}$ cells using the iEμ/Iμ locale as bait (blue asterisks). Grey Bars highlight the iEμ-Cμ, Sγ3, Sγ2b, Sγ2a, Sα, 3'IgHRR and 3'CBEs. (right), Zoom-in view of the 3C-HTGTS profiles on the left to better reveal the interaction patterns in the region from Sγ3 to Sε in non-stimulated and αCD40/IL4/TGFβ stimulated IαΔ-AID$^{-/-}$ cells.



Extended Data Figure 7. CBEs inserted upstream of Iα lead to increased inversional Sα CSR.

a, Three repeats of Rad21 ChIP-Seq profiles of CD40/IL4/TGFβ stimulated i3CBEs-AID $^{-/-}$. **b, c,** Additional repeats of the 3C-HTGTS profiles shown in Fig. 3f, g, from CD40/IL4/TGFβ-stimulated CH12F3NCΔ-AID$^{-/-}$ and i3CBEs-AID$^{-/-}$ cells using CBE insertion (**c**) or iEμ/Iμ (**d**) locale as baits (blue asterisk). Diagram on the right of panel **c** 3C-HTGTS profiling shows the digestion/bait strategies used; **e,** Model to address increased inversional Sα CSR in CH12F3 cells with CBEs inserted upstream of Iα. **I,** Cohesin is loaded at various *Igh* locations including transcriptionally activated Iα-Sα. **II-VII,** For cohesin loaded at Iα locale, extrusion past the CBE impediment allows a significant subset of cells to reach step

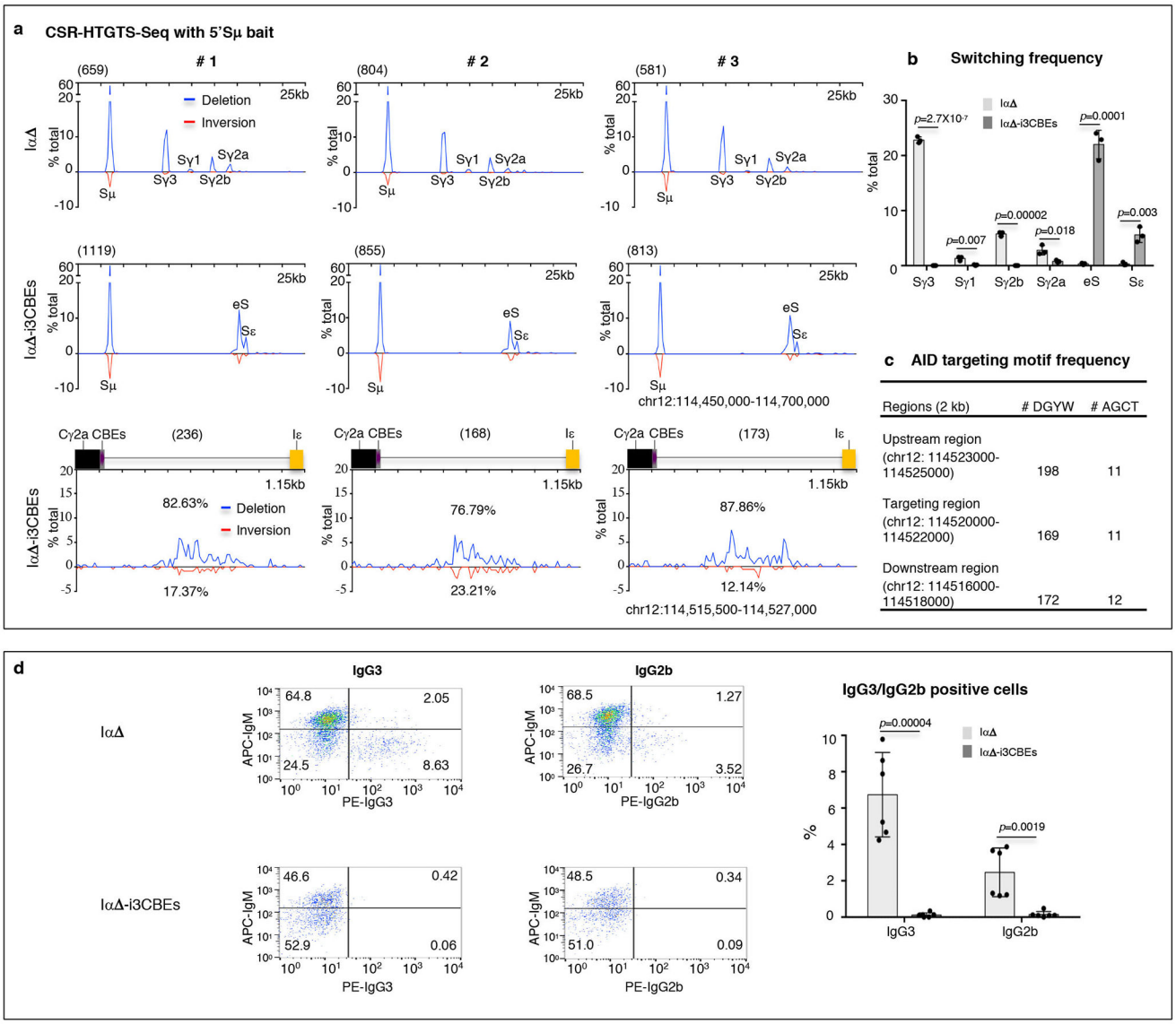
VII to generate CSRC. **VIII-X,** In these cells, a significant portion of continued upstream extrusion passes by the CBE impediment to yield cells in the population with configurations shown steps **IX** and **X**. **XI-XIV,** The cells with the configuration shown in **IX** will have both deletional (**XIII)** and inversional (**XIV**) joining mediated by a diffusion-related process in the absence of complete Sμ-Sα synapsis (See text for more details). Those with the configuration shown in **X** will join via deletion as described in Extended Data Fig. 5e. This working model could be explained by other variations as indicated for other model figures.



Extended Data Figure 8. 3'CBEs deletion in i3CBEs cells has little effect on the Sα CSR and Sα inversional joining.

a, Representative FACS analyses for IgH class switching from IgM to IgA for CH12F3NCΔ, i3CBEs and i3CBEs-3'CBEsΔ cells stimulated with αCD40/IL4/TGFβ for 72hrs. Bar graph on right shows FACS data from six biological independent repeats plotted as mean ± s.d. *P* values were calculated via an unpaired two-tailed *t*-test. **b,** CSR-HTGTS-Seq of three repeats that use 5'Sμ bait for analyses of αCD40/IL4/TGFβ-stimulated CH12F3NCΔ, i3CBEs and i3CBEs-3'CBEsΔ cells. Junctions are plotted at 200 bp bin size. Blue lines indicate deletional joining and red lines indicate inversional joining. **c,** Schematic of *Igh* C_H

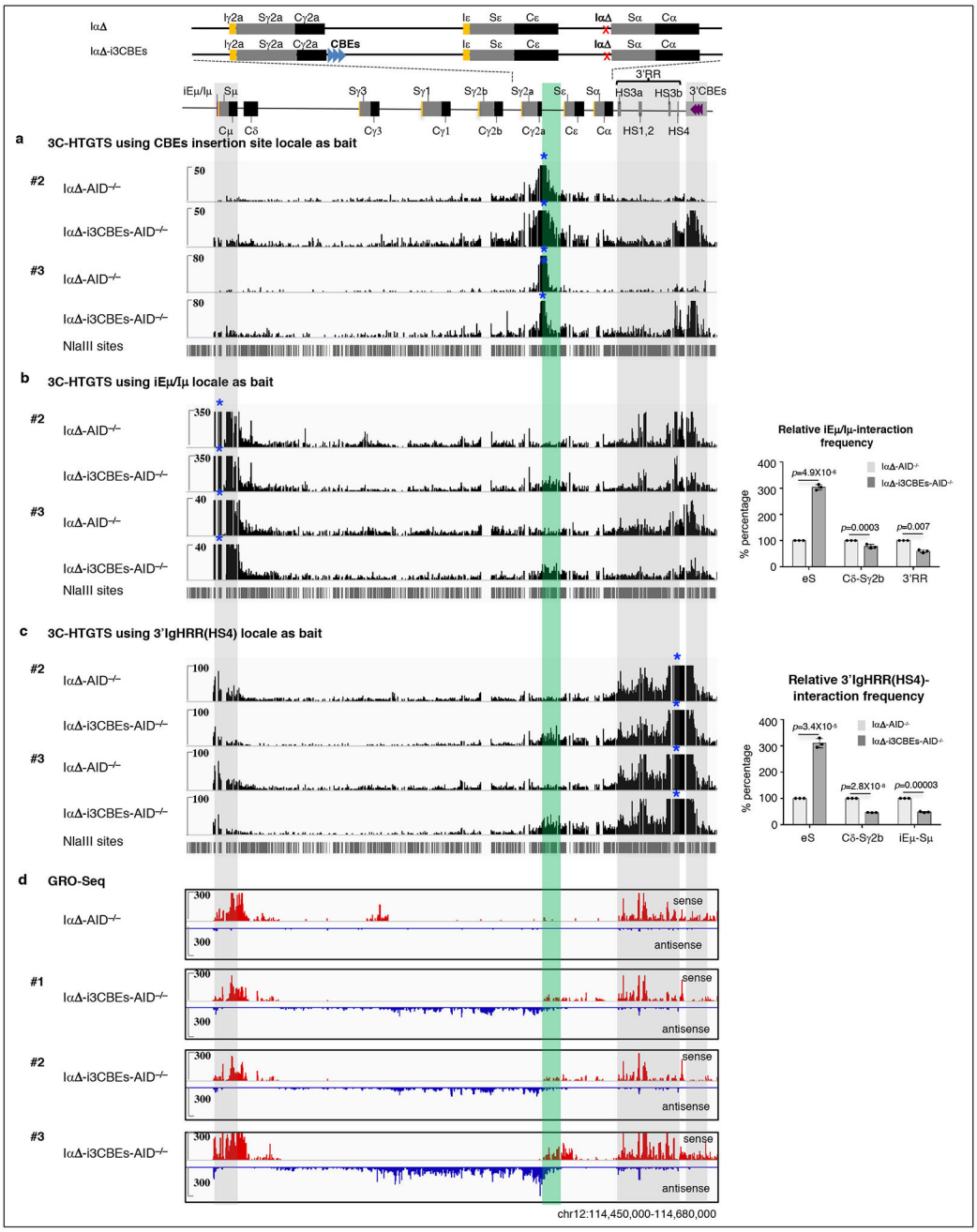
locus from iEμ to 33kb downstream of 3'CBEs. Zoom-in view on top illustrates 3'CBEs deletion in i3CBEs lines to generate i3CBEs-3'CBEΔ lines. **d,** Three repeats of 3C-HTGTS profiles of αCD40/IL4/TGFβ-stimulated i3CBEs-AID$^{-/-}$ and i3CBEs-3'CBEΔ-AID$^{-/-}$ cells using the CBEs insertion locale as bait (blue asterisks).



Extended Data Figure 9. CBE-insertion in Iα-deleted CH12F3NCΔ cells impedes IgH class-switching/CSR and creates ectopic S region.

a, (upper and middle panels) Three individual repeats of CSR-HTGTS-Seq experiments shown in Fig. 4b that employ a 5'Sμ bait to assay αCD40 region /IL4/TGFβ-stimulated IαΔ and IαΔ-i3CBEs cells. Junctions are plotted at 2.5 kb bin size. (bottom panels) Zoom-in views of three repeats of data in Fig. 4c showing junctions located in the AID-targeted ectopic S region between Cγ2a and Iε in assays of IαΔ-i3CBEc cells. Junctions are plotted at 115 bp bin size. Blue lines indicate deletional joining and red lines indicate inversional joining. **b,** Bar graph shows percentages of junctions located in indicated AID targeted regions from IαΔ and IαΔ-i3CBEs cells. Data represents mean ± s.d. from three biologically independent repeats. *P* values were calculated via an unpaired two-tailed *t*-test based on the three repeats. **c,** AID targeting motif frequency analysis of 2 kb ectopic S region targeting peak and in comparably sized region just upstream and downstream of the targeting peak. **d,** FACS analysis of IgG3 and IgG2b surface expression in IαΔ and IαΔ-i3CBEs cells

stimulated with αCD40/IL4/TGFβ for 72 hrs (6 biologically independent repeats). Bar graph shows percentages of IgG3 and IgG2b production from IαΔ and IαΔ-i3CBEs cells. Data represents mean ± s.d. from six biologically independent repeats. *P* values were calculated via an unpaired two-tailed *t*-test.



Extended Data Figure 10. Repeats of Figure 4e-h showing CBE-insertion in Iα-deleted CH12F3Δ cells impedes upstream transcription and looping.

a, b, c, Additional repeats of Fig. 4e-h 3C-HTGTS profiles from αCD40/IL4/TGFβ-stimulated IαΔ-AID$^{-/-}$ and IαΔ-i3CBEs-AID$^{-/-}$cells using the CBEs insertion (**a**) (3 biologically independent repeats), the iEμ/Iμ (**b**) (3 biologically independent repeats), or the 3'IgHRR(HS4) (**c**) (3 biologically independent repeats) locale as baits (blue asterisks). Bar graphs on the right of the 3C-HTGTS profiles show the relative iEμ/Iμ or 3'IgHRR(HS4) interaction frequency with eS and the region between Cδ to Sγ2b. Data represents mean ± s.d. in panels **b-c** from three biologically independent repeats. *P* values were calculated via

paired two-tailed *t*-test. **d,** Three repeats of Fig. 4h GRO-Seq profiles with larger scale from αCD40/IL4/TGFβ stimulated IαΔ-AID$^{-/-}$ and IαΔ-i3CBEs-AID$^{-/-}$cells. All bars and other notations as described in the legend to Fig. 4.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank Tasuku Honjo for the CH12F3 cell line and AID$^{-/-}$ C57BL/6 mice. This work was supported by NIH R01AI077595. F.W.A. is an investigator of the Howard Hughes Medical Institute. Y.Z. is a special fellow of the Leukemia and Lymphoma Society. Z.B. was a Cancer Research Institute Irvington fellow.

REFERENCES

1. Matthews AJ, Zheng S, DiMenna LJ & Chaudhuri J Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. Adv Immunol 122, 1–57, doi:10.1016/B978-0-12-800267-4.00001-8 (2014). [PubMed: 24507154]
2. Muramatsu M et al. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell 102, 553–563 (2000). [PubMed: 11007474]
3. Cogne M et al. A class switch control region at the 3' end of the immunoglobulin heavy chain locus. Cell 77, 737–747 (1994). [PubMed: 8205622]
4. Manis JP et al. Class switching in B cells lacking 3' immunoglobulin heavy chain enhancers. J Exp Med 188, 1421–1431 (1998). [PubMed: 9782119]
5. Seidl KJ et al. Position-dependent inhibition of class-switch recombination by PGK-neor cassettes inserted into the immunoglobulin heavy chain constant region locus. Proc Natl Acad Sci U S A 96, 3000–3005 (1999). [PubMed: 10077626]
6. Wuerffel R et al. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. Immunity 27, 711–722, doi:10.1016/j.immuni.2007.09.007 (2007). [PubMed: 17980632]
7. Feldman S et al. 53BP1 Contributes to Igh Locus Chromatin Topology during Class Switch Recombination. J Immunol 198, 2434–2444, doi:10.4049/jimmunol.1601947 (2017). [PubMed: 28159901]
8. Rocha PP et al. A Damage-Independent Role for 53BP1 that Impacts Break Order and Igh Architecture during Class Switch Recombination. Cell Rep 16, 48–55, doi:10.1016/j.celrep.2016.05.073 (2016). [PubMed: 27320916]
9. Chaudhuri J et al. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature 422, 726–730, doi:10.1038/nature01574 (2003). [PubMed: 12692563]
10. Rogozin IB & Diaz M Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. J Immunol 172, 3382–3384 (2004). [PubMed: 15004135]
11. Dong J et al. Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. Nature 525, 134–139, doi:10.1038/nature14970 (2015). [PubMed: 26308889]
12. Methot SP & Di Noia JM Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. Adv Immunol 133, 37–87, doi:10.1016/bs.ai.2016.11.002 (2017). [PubMed: 28215280]
13. Jain S, Ba Z, Zhang Y, Dai HQ & Alt FW CTCF-Binding Elements Mediate Accessibility of RAG Substrates During Chromatin Scanning. Cell 174, 102–116 e114, doi:10.1016/j.cell.2018.04.035 (2018). [PubMed: 29804837]

14. Zhang Y et al. The fundamental role of chromatin loop extrusion in physiological V(D)J recombination. Nature, doi:10.1038/s41586-019-1547-y (2019).
15. Vian L et al. The Energetics and Physiological Impact of Cohesin Extrusion. Cell 173, 1165–1178 e1120, doi:10.1016/j.cell.2018.03.072 (2018). [PubMed: 29706548]
16. Braikia FZ et al. Inducible CTCF insulator delays the IgH 3' regulatory region-mediated activation of germline promoters and alters class switching. Proc Natl Acad Sci U S A 114, 6092–6097, doi: 10.1073/pnas.1701631114 (2017). [PubMed: 28533409]
17. Nakamura M et al. High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA + cells. Int Immunol 8, 193–201 (1996). [PubMed: 8671604]
18. Meng FL et al. Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. Cell 159, 1538–1548, doi:10.1016/j.cell.2014.11.014 (2014). [PubMed: 25483776]
19. Haarhuis JHI et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell 169, 693–707 e614, doi:10.1016/j.cell.2017.04.013 (2017). [PubMed: 28475897]
20. Fudenberg G et al. Formation of Chromosomal Domains by Loop Extrusion. Cell Rep 15, 2038–2049, doi:10.1016/j.celrep.2016.04.085 (2016). [PubMed: 27210764]
21. Sanborn AL et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A 112, E6456–6465, doi:10.1073/pnas.1518552112 (2015). [PubMed: 26499245]
22. Lucas JS, Zhang Y, Dudko OK & Murre C 3D trajectories adopted by coding and regulatory DNA elements: first-passage times for genomic interactions. Cell 158, 339–352, doi:10.1016/j.cell.2014.05.036 (2014). [PubMed: 24998931]
23. Gostissa M et al. IgH class switching exploits a general property of two DNA breaks to be joined in cis over long chromosomal distances. Proc Natl Acad Sci U S A 111, 2644–2649, doi:10.1073/pnas.1324176111 (2014). [PubMed: 24550291]
24. Hwang JK, Alt FW & Yeap LS Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. Microbiol Spectr 3, MDNA3-0037-2014, doi:10.1128/microbiolspec.MDNA3-0037-2014 (2015).
25. Dudley DD et al. Internal IgH class switch region deletions are position-independent and enhanced by AID expression. Proc Natl Acad Sci U S A 99, 9984–9989, doi:10.1073/pnas.152333499 (2002). [PubMed: 12114543]
26. Zhang T et al. Downstream class switching leads to IgE antibody production by B lymphocytes lacking IgM switch regions. Proc Natl Acad Sci U S A 107, 3040–3045, doi:10.1073/pnas.0915072107 (2010). [PubMed: 20133637]
27. Kim JS, Krasieva TB, LaMorte V, Taylor AM & Yokomori K Specific recruitment of human cohesin to laser-induced DNA damage. J Biol Chem 277, 45149–45153, doi:10.1074/jbc.M209123200 (2002). [PubMed: 12228239]
28. Strom L, Lindroos HB, Shirahige K & Sjogren C Postreplicative recruitment of cohesin to double-strand breaks is required for DNA repair. Mol Cell 16, 1003–1015, doi:10.1016/j.molcel.2004.11.026 (2004). [PubMed: 15610742]
29. Thomas-Claudepierre AS et al. The cohesin complex regulates immunoglobulin class switch recombination. J Exp Med 210, 2495–2502, doi:10.1084/jem.20130166 (2013). [PubMed: 24145512]
30. Enervald E et al. A regulatory role for the cohesin loader NIPBL in nonhomologous end joining during immunoglobulin class switch recombination. J Exp Med 210, 2503–2513, doi:10.1084/jem.20130168 (2013). [PubMed: 24145515]
31. Deng R et al. Extrafollicular CD4(+) T-B interactions are sufficient for inducing autoimmune-like chronic graft-versus-host disease. Nat Commun 8, 978, doi:10.1038/s41467-017-00880-2 (2017). [PubMed: 29042531]
32. Usui T et al. Overexpression of B cell-specific activator protein (BSAP/Pax-5) in a late B cell is sufficient to suppress differentiation to an Ig high producer cell with plasma cell phenotype. J Immunol 158, 3197–3204 (1997). [PubMed: 9120274]
33. Cook AJ et al. Reduced switching in SCID B cells is associated with altered somatic mutation of recombined S regions. J Immunol 171, 6556–6564 (2003). [PubMed: 14662857]

34. Gliddon DR & Howard CJ CD26 is expressed on a restricted subpopulation of dendritic cells in vivo. Eur J Immunol 32, 1472–1481, doi:10.1002/1521-4141(200205)32:5<1472::AID-IMMU1472>3.0.CO;2-Q (2002). [PubMed: 11981836]

35. Lahrouchi N et al. Homozygous frameshift mutations in FAT1 cause a syndrome characterized by colobomatous-microphthalmia, ptosis, nephropathy and syndactyly. Nat Commun 10, 1180, doi: 10.1038/s41467-019-08547-w (2019). [PubMed: 30862798]

36. Cong L et al. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823, doi:10.1126/science.1231143 (2013). [PubMed: 23287718]

37. Yang L et al. Optimization of scarless human stem cell genome editing. Nucleic Acids Res 41, 9049–9061, doi:10.1093/nar/gkt555 (2013). [PubMed: 23907390]

38. Hu J et al. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. Nat Protoc 11, 853–871, doi: 10.1038/nprot.2016.043 (2016). [PubMed: 27031497]

39. Core LJ, Waterfall JJ & Lis JT Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322, 1845–1848, doi:10.1126/science.1162228 (2008). [PubMed: 19056941]

40. Mahat DB et al. Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nat Protoc 11, 1455–1476, doi:10.1038/nprot.2016.086 (2016). [PubMed: 27442863]

41. Marinov GK ChIP-seq for the Identification of Functional Elements in the Human Genome. Methods Mol Biol 1543, 3–18, doi:10.1007/978-1-4939-6716-2_1 (2017). [PubMed: 28349419]

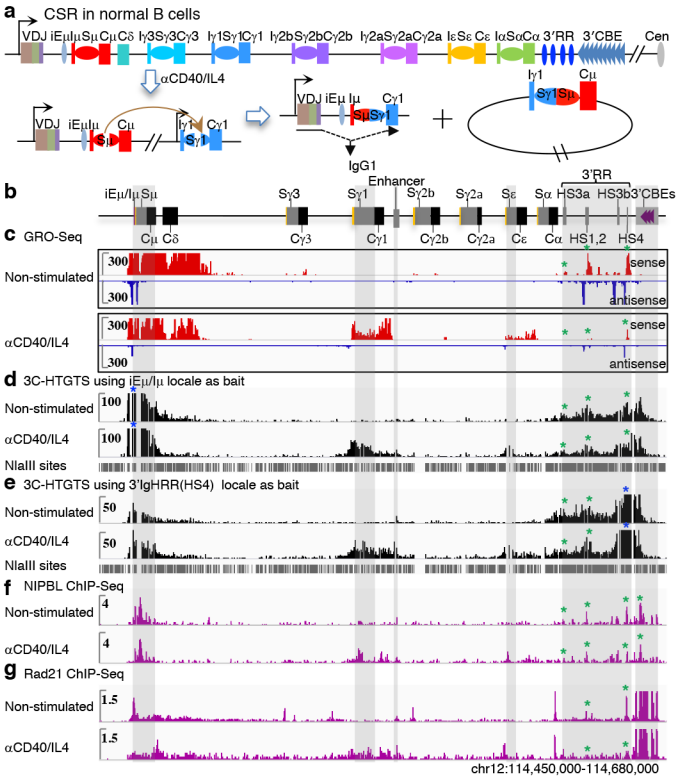


Figure 1. Cytokine -induced target S region transcription promotes synapsis with Sμ during CSR.

a, Illustration of *Igh* C_H locus (top) and activation of CSR in normal B cells stimulated with αCD40/IL4, which induces AID and activates transcription of Iγ1(shown) and Iε (not shown)[11]. As indicated, the vast majority of CSR events are deletional, with an upstream end of an Sμ DSB joining to the downstream end of an acceptor S region DSB[11]. **b,** Schematic of 3'*Igh* locus domain from iEμ to 3'CBEs. **c,** GRO-Seq profiles of AID$^{-/-}$ mature splenic B cells without stimulation or with αCD40/IL4 stimulation (3 biologically independent repeats with similar results). Sense transcription is shown above in red and antisense transcription is shown below in blue lines. **d, e,** High resolution 3C-HTGTS[13] profiles of interactions within the *3'Igh* locus domain in AID$^{-/-}$ mature splenic B cells without stimulation or with αCD40/IL4 stimulation as indicated using the iEμ/Iμ (**d**) (3 biologically independent repeats with similar results) or 3'IgHRR HS4 (**e**) (3 biologically independent repeats with similar results) locale as baits (blue asterisk). As portions of Sμ and certain other S regions cannot be mapped due to lack of NlaIII sites, their interactions are inferred from mappable sequences. **f, g,** NIPBL (**f**) (3 biologically independent repeats with similar results) and Rad21 (**g**) (3 biologically independent repeats with similar results) ChIP-Seq profiles of AID$^{-/-}$ mature splenic B cells without stimulation or with αCD40/IL4 stimulation as indicated. Grey Bars highlight the iEμ-Cμ, Sγ1, Sε, 3'IgHRR and 3'CBEs. Green asterisks indicate the HS3a, HS1,2 and HS4 sites within 3'IgHRR. Repeat experiments for all panels are in Extended Data Fig. 2.

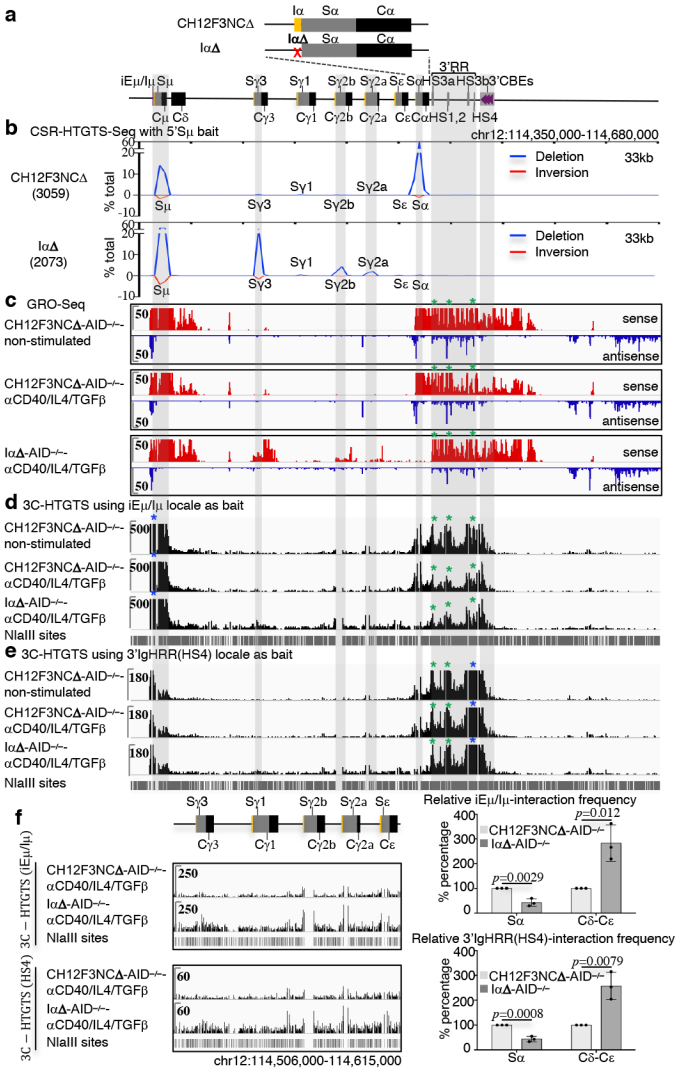


Figure 2. Constitutive CH12F3 Sα transcription causes dominant Sα CSR and impedes long-range interactions and CSR to upstream S regions.

a, Schematic of *Igh* C_H locus from iEμ to 33kb downstream of the 3'CBEs. Zoom-in view on top illustrates deletion of Iα in CH12F3NCΔ lines to generate IαΔ lines. **b,** Results of CSR-HTGT-Seq analysis[11] which measures joining of the 5' end of AID-initiated DSBs in upstream 5' region of Sμ to either upstream (inversional) or downstream (deletional) ends of AID-initiated DSB ends in downstream acceptor S regions in CH12F3NCΔ cells and IαΔ cells stimulated with αCD40/IL4/TGFβ for 72 hrs. Junctions are plotted at 3.3 kb bin size. Blue line indicates deletional joining and red line indicates inversional joining. **c,** GRO-Seq profiles across the indicated *Igh* domain of non-stimulated CH12F3NCΔ-AID$^{-/-}$ cells, αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells, other details as in Fig. 1c legend. **d, e,** 3C-HTGTS profiles of interactions across the indicated *Igh* domain of non-stimulated CH12F3NCΔ-AID$^{-/-}$ cells, αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells using the iEμ/Iμ (**d**) or 3'IgHRR HS4 (**e**) locale as baits (blue asterisks). All panels in **b-e** were repeated three times independently and showed similar results. Grey Bars highlight the iEμ-Cμ, Sγ3, Sγ2b, Sγ2a, Sα,3'IgHRR and 3'CBEs. **f.**

Zoom-in view 3C-HTGTS profiles in bottom two tracks of panels d and e (iEμ/Iμ and HS4 baits) to better reveal the interaction patterns in the region between Cδ to Cε in αCD40/IL4/TGFβ stimulated-CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells. Bar graphs on right show the relative iEμ/Iμ or 3'IgHRR(HS4) interaction frequency with Sα and the intervening sequence between Cδ and Cε in αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ and IαΔ-AID$^{-/-}$ cells. Data represents mean ± s.d. from three biologically independent samples. *P* values were calculated via an unpaired two-tailed *t*-test. Repeat experiments for all panels are in Extended Data Fig. 3, 4, 5, 6.

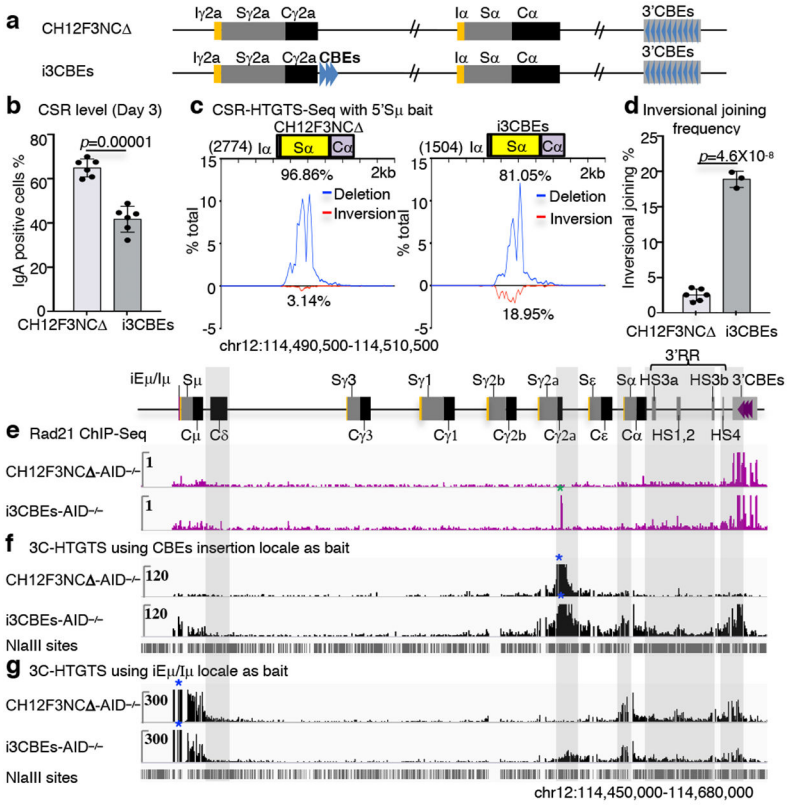


Figure 3. Inserting CBEs upstream of Iα activates inversional CH12F3 CSR.

a, Illustration of i3CBEs insertion in CH12F3NCΔ lines to generate i3CBEs lines. CBEs and their orientation are indicated by blue arrowheads. **b,** IgH class-switching to IgA in αCD40/IL4/TGFβ stimulated (72 hrs) in CH12F3NCΔ and i3CBEs cells (6 biologically independent repeats). Data represents mean ± s.d. from six biologically independent repeats. *P* values were calculated via an unpaired two-tailed *t*-test. **c,** CSR-HTGT-Seq analysis of the CH12F3NCΔ (6 biologically independent repeats with similar results) and i3CBEs (3 biologically independent repeats with similar results) cells stimulated with αCD40/IL4/TGFβ for 72 hrs. Junctions are plotted at 200 bp bin size. **d,** Bar graph showing the percentages of inversional end joining between 5'Sμ DSB ends and Sα DSBs in αCD40/IL4/TGFβ stimulated (72 hrs) CH12F3NCΔ (*n*=6 libraries) and i3CBEs cells (*n*=3 libraries). Data represents mean ± s.d. from biological independent samples. *P* values were calculated via an unpaired two-tailed *t*-test. **e,** Rad21 ChIP-Seq profiles of αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ and i3CBEs-AID$^{-/-}$ cells (3 biologically independent repeats with similar results). Green asterisk indicates cohesin accumulation at the CBEs insertion site. **f, g,** 3C-HTGTS profiles of αCD40/IL4/TGFβ stimulated CH12F3NCΔ-AID$^{-/-}$ and i3CBEs-AID$^{-/-}$ cells using either the CBEs insertion (**f**) (3 biologically independent repeats with similar results) or iEμ/Iμ (**g**) (3 biologically independent repeats with similar results) locale as bait (blue asterisk). Grey Bars highlight the broader regions around the Sμ, i3CBEs, Sα, 3'IgHRR and 3'CBEs. Repeat experiments for all panels are in Extended Data Fig. 7, 8.

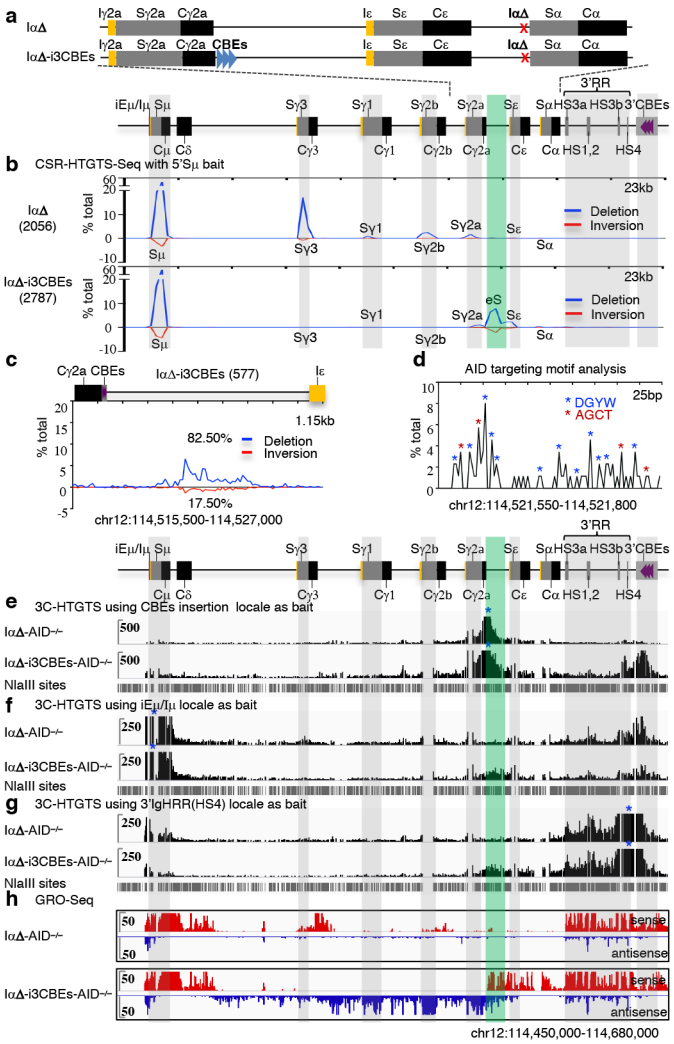


Figure 4. CBE-insertion in Iα-deleted CH12F3Δ cells impedes upstream transcription, looping and CSR and creates ectopic S region.

a, Schematic of *Igh* C_H locus from iEμ to 3'CBEs. Zoom-in view shows i3CBEs insertion site in IαΔ lines to generate IαΔ-i3CBEs lines. **b,** CSR-HTGTS-Seq analysis of break joining between 5'Sμ and downstream acceptor S or non-S regions in IαΔ and IαΔ-i3CBEs cells stimulated with αCD40/IL4/TGFβ for 72 hrs (3 biological independent repeats with similar results). Junctions are plotted at 2.3 kb bin size. Blue line indicates deletional joining and red line indicates inversional joining. In the lower panel the location of the AID-targeted ectopic S region (labeled as "eS") between Cγ2a and Iα is highlighted by a transparent green bar through all panels. Grey Bars highlight the iEμ-Cμ, Sγ3, Sγ1, Sγ2b, Sγ2a, Sε, 3'IgHRR and 3'CBEs. **c,** Zoom-in view of CSR-HTGTS-Seq junctions located in the AID-targeted ectopic S region between Cγ2a and Iε from IαΔ-i3CBEs cells (3 biological independent repeats with similar results). Junctions are plotted at 115 bp bin size. **d,** AID targeting motif analysis for the junctions located in a 250bp region within AID-targeted ectopic S region from IαΔ-i3CBEs cells (3 biological independent repeats with similar results). Blue asterisks in this panel indicate DGYW motifs and red asterisks indicate AGCT motifs. **e, f, g,** 3C-HTGTS profiles of αCD40/IL4/TGFβ stimulated IαΔ-AID$^{-/-}$ and IαΔ-

i3CBEs-AID$^{-/-}$ cells using, respectively, the CBEs insertion (**e**) (3 biological independent repeats with similar results), the iEμ/Iμ (**f**) (3 biological independent repeats with similar results), or the 3'IgHRR(HS4) **(g)** (3 biological independent repeats with similar results) locale as baits (bait sites denoted by blue asterisk). **h,** GRO-Seq analysis of αCD40/IL4/TGFβ stimulated IαΔ-AID$^{-/-}$ and IαΔ-i3CBEs-AID$^{-/-}$ cells (3 biological independent repeats with similar results). Repeat experiments for all panels are in Extended Data Fig. 9, 10.