Immunoglobulin switch μ sequence causes RNA polymerase II accumulation and reduces dA hypermutation

Deepa Rajagopal,[1,3] Robert W. Maul,[1] Amalendu Ghosh,[2] Tirtha Chakraborty,[2] Ahmed Amine Khamlichi,[4] Ranjan Sen,[2] and Patricia J. Gearhart[1]

[1]Laboratory of Molecular Gerontology and [2]Laboratory of Cellular and Molecular Biology, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224
[3]Graduate Program in Molecular Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD 21201
[4]Centre national de la recherche scientifique UMR 5089-IPBS, Université de Toulouse, Toulouse Cedex F-31077, France

Repetitive DNA sequences in the immunoglobulin switch μ region form RNA–containing secondary structures and undergo hypermutation by activation–induced deaminase (AID). To examine how DNA structure affects transcription and hypermutation, we mapped the position of RNA polymerase II molecules and mutations across a 5–kb region spanning the intronic enhancer to the constant μ gene. For RNA polymerase II, the distribution was determined by nuclear run-on and chromatin immunoprecipitation assays in B cells from uracil–DNA glycosylase (UNG)–deficient mice stimulated ex vivo. RNA polymerases were found at a high density in DNA flanking both sides of a 1–kb repetitive sequence that forms the core of the switch region. The pileup of polymerases was similar in unstimulated and stimulated cells from $Ung^{-/-}$ and $Aid^{-/-}Ung^{-/-}$ mice but was absent in cells from mice with a deletion of the switch region. For mutations, DNA was sequenced from $Ung^{-/-}$ B cells stimulated in vivo. Surprisingly, mutations of A nucleotides, which are incorporated by DNA polymerase η, decreased 10–fold before the repetitive sequence, suggesting that the polymerase was less active in this region. We propose that altered DNA structure in the switch region pauses RNA polymerase II and limits access of DNA polymerase η during hypermutation.

CORRESPONDENCE
Patricia J. Gearhart:
gearhartp@grc.nia.nih.gov

Class switch recombination is initiated by activation–induced deaminase (AID) in switch (S) regions preceding each C_H gene (1). AID acts by deaminating cytosine to uracil in DNA, and the rogue uracil causes mutations and DNA strand breaks by error-prone processing. Mutations of C:G bp could be produced by replication past uracil or an abasic site caused by removal of uracil by uracil-DNA glycosylase (UNG). Mutations can also arise from the combined actions of MSH2-MSH6 mismatch repair proteins and the low fidelity DNA polymerase (pol) η to produce mutations of A:T bp (2). Strand breaks produced by cleavage of abasic sites on both strands by an abasic endonuclease (3) would produce substrates for nonhomologous end joining to different S regions. In the S region preceding the Cμ gene,

mutations start downstream of an intronic (Iμ) promoter located in the enhancer (Eμ) region (4), accumulate for 2 kb, and then decrease before the Cμ gene (5). A similar pattern of mutations occurs downstream of intronic promoters in the S regions of other C_H genes as well (5). This distribution supports the prevailing hypothesis that AID either travels with the transcription complex or is brought to the single-strand regions formed during transcription (6, 7).

The Sμ DNA sequence is unique in two aspects: it contains an abundance of hot spot motifs, and it has a stable secondary structure that exposes single-strand DNA for AID to act on. The substructure of this region contains a 1–kb highly repetitive sequence, which is comprised of tandem WGCW (W is A or T) motifs and

clusters of three to four Gs on the nontranscribed strand. WGCW is a hot spot for AID activity, and the G clusters have been proposed to form secondary structures during transcription, such as R–loops and G–loops containing G–rich RNA and C–rich transcribed DNA (8, 9). Although the nontranscribed strand is mostly single stranded in these structures, the spectrum of mutations in $Ung^{-/-}\ Msh2^{-/-}$ mice indicates that both strands are targeted for mutation by AID (5). The mechanism by which AID deaminates the transcribed strand containing an RNA–DNA hybrid is unknown, although three theories have been proposed: (a) the DNA upstream of an elongating RNA pol II may be supercoiled and unwound, which would allow AID access to both strands (10); (b) the DNA in R–loops may be collapsed by endogenous RNase H digestion, which would expose single-strand regions on the transcribed strand (11); and (c) single-strand DNA on the transcribed strand could be generated during antisense transcription. Antisense transcripts have been detected at low levels in S regions, but it is not known if they participate in hypermutation (12).

The DNA region upstream of the repetitive sequence has fewer WGCW motifs and G clusters, but it can also form secondary structure and undergo hypermutation (5, 13, 14). Huang et al. (11) have recently mapped the 5′ and 3′ boundaries of R–loops in the Sμ region and found that they begin 600 bp upstream and end 600 bp downstream of the 1-kb repetitive sequence. The existence of R–loops and associated single-strand DNA in the upstream region may also limit the participation of MSH2-MSH6 and DNA pol η during hypermutation, which would reduce the frequency of mutations of A:T bp.

RESULTS AND DISCUSSION
Sμ DNA structure affects the location of RNA pol II molecules

In view of the requirement of transcription for hypermutation and the possibility that AID may associate with the transcription complex, we systematically mapped the position of RNA pol II molecules across the Sμ locus using nuclear run-on and chromatin immunoprecipitation (ChIP) assays. The former measures RNA polymerase activity in a region for a given time, whereas the latter measures the amount of RNA pol II molecules bound to DNA at the instant of cell lysis. B cells from $Ung^{-/-}$ mice were studied because the entire Sμ region can be analyzed without deletions caused by switching (15). The 5-kb Sμ region was divided into a series of contiguous 500-bp probes for run-on analysis (Fig. 1 A, S1–9). However, the 1-kb region containing highly repetitive G:C-rich DNA could not be amplified by our PCR techniques, although we tried a variety of DNA polymerases and conditions (unpublished data). We prepared nuclei from naive B cells or cells stimulated ex vivo with LPS and IL-4 for 2 d and initiated de novo transcription by the addition of ribonucleotides and radioactive UTP for 30 min. The labeled nascent RNA molecules were hybridized to membranes containing DNA probes of the Sμ region, and membrane-bound radioactivity was quantified by phosphorimager analysis (Fig. 1 B). Data were normalized to β-actin transcription and corrected for dTTP content because radiolabeled UTP was incorporated. In nuclei from days 0 and 2, the strongest signals were obtained from probes S5 and S7, which flank the repetitive region and are within the R–loop region. Hybridizations were also performed with single-strand DNA probes made in M13 vectors. There was hybridization to transcribed strand probes but no detectable hybridization to nontranscribed strand probes (unpublished data), which indicates that the majority of RNA transcripts are complementary to the transcribed strand, as expected.

To determine if the accumulation of polymerases around the repetitive region was caused by mutations initiated by AID, nuclear run-ons were performed in B cells from $Aid^{-/-}Ung^{-/-}$ mice 2 d after stimulation. As seen in Fig. 1 C, there was still an accumulation of polymerases in S5 and S7 in the absence of AID. We also saw a similar pattern of hybridization in nuclei from hybridoma cells that do not express AID (unpublished data). There were twofold fewer transcripts in S5 and S7 in day-2 cells from $Ung^{-/-}$ mice, which express the AID protein, compared with day-0 and day-2 cells from $Ung^{-/-}$ and $Aid^{-/-}Ung^{-/-}$ cells, respectively. This implies that AID affects the level of transcription, but the molecular basis for this is unknown. To see if the pileup of polymerases was directly caused by the Sμ DNA sequence, we performed run-on assays in cells from mice which had a 3.7-kb deletion in Sμ ($Sμ\ del$) (16) and have no detectable R-loop formation (11). As shown in Fig. 1 D, after 2 d of stimulation, there was no accumulation of RNA pol II when the Sμ sequence was virtually deleted.

To confirm the nuclear run-on data, the density of RNA pol II molecules was directly assayed by ChIP with anti-RNA pol II antibodies, which were used to enrich polymerase-bound chromatin. After purification, the associated genomic DNA was amplified by PCR with 10 primer sets located across the Sμ region (Fig. 2 A, Sa–j). The polymerase distribution across the region was the same in day-0 and day-2 activated cells. The highest levels of RNA pol II were detected in the amplicons that were closest to the repetitive region, Sg, Sh, and Si, which showed a twofold increase compared with adjacent amplicons. In contrast, there was no accumulation of RNA pol II in mice devoid of the Sμ region (Fig. 2 B). Therefore, both the nuclear run-on and ChIP assays indicate that RNA pol II accumulates close to the repetitive region. This pattern is directly related to the DNA sequence encoding Sμ and is not related to the activation status of the cells or to the presence of AID-dependent mutation. It should be noted that the quantity of RNA pol II flanking the repetitive region differs between the two assays in that the ChIP assay shows a twofold increase, whereas the run-on assay detects a four- to eightfold increase. This could happen if the ChIP assay depicts the loading of RNA pol II at a given instant, whereas the run-on assay shows the activity of the polymerases during a 30-min incubation time, which may produce a more intense signal, as demonstrated in Fig. S1. In both cases, the increase in polymerase density around the repetitive region supports the notion that DNA structure retards the movement of the polymerases.

To evaluate where the RNA pol II molecules initiated transcription, we measured the quantity of transcripts starting
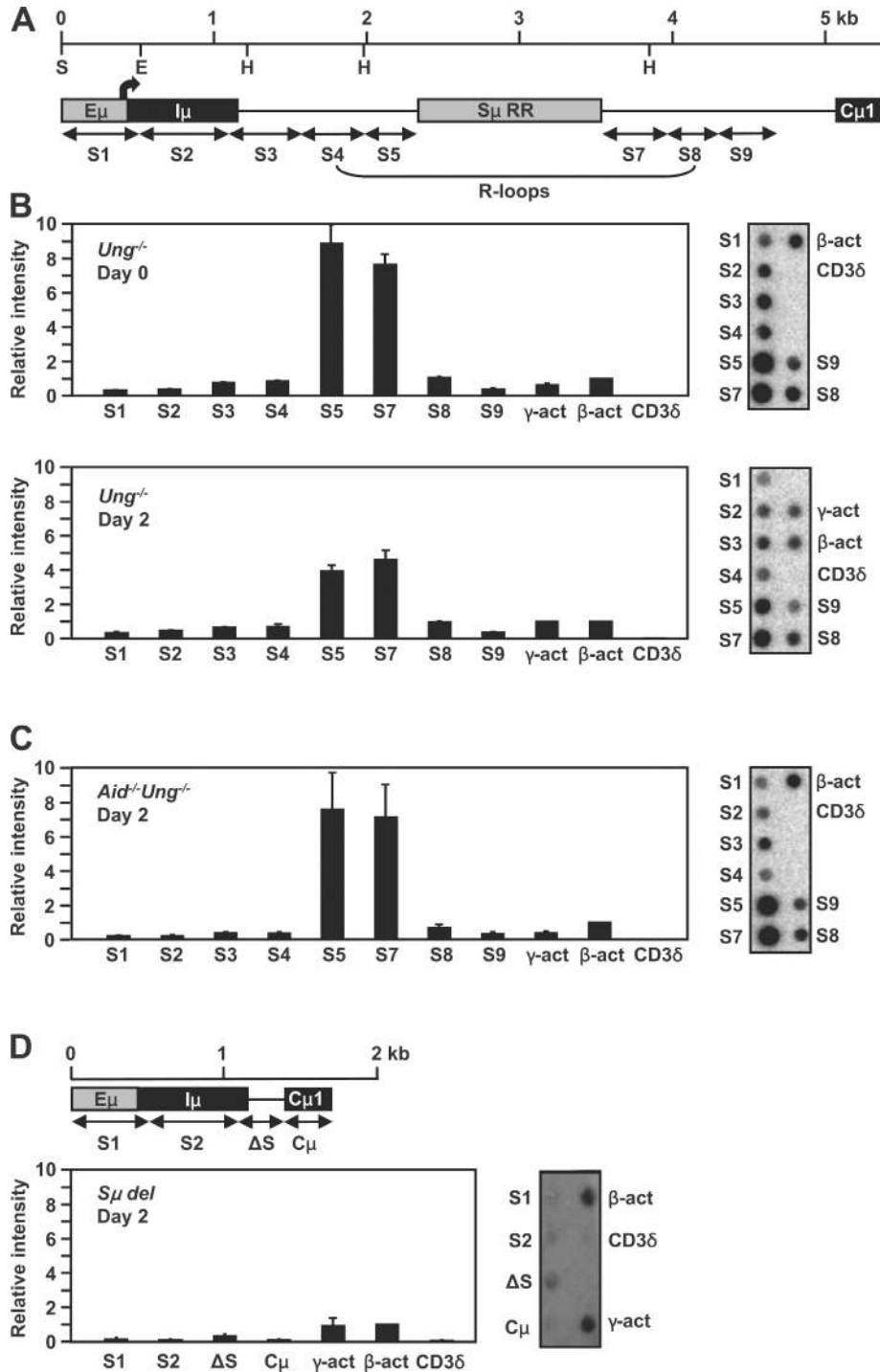
Figure 1. **Location of RNA polymerases by nuclear run–on.** (A) A map of the fragments that were analyzed is shown; the Sμ repetitive region (RR) was not analyzed. Position 1 corresponds to nt 135,091 in GenBank/EMBL/DDBJ under accession no. AC073553. The restriction sites ScaI (S), EcoRI (E), and HindIII (H) are depicted. The bent arrow shows the position of the Iμ promoter. The region forming the proposed R-loops is indicated. For graphs in B–D, representative blots for each graph are shown on the right. The intensity of hybridization of labeled nascent RNA to each of the fragments was graphed based on the value of β-actin, which was set as 1. γ-Actin and CD3δ were positive and negative controls, respectively. (B) Ung⁻/⁻ splenic B cells before and after stimulation with LPS and IL-4. The error bars represent the SD of the values from triplicate blots from one experiment on days 0 and 2 with five mice. Three independent experiments with five mice each were performed and showed similar patterns and intensities. (C) Aid⁻/⁻Ung⁻/⁻ cells after stimulation. Data are from triplicate blots from one experiment with three mice, with error bars showing SD. Three independent experiments with three mice each were performed with similar results. (D) Sμ del cells after stimulation. The ΔS probe corresponded to the 5′ and 3′ ends flanking the 3.7-kb deletion (16). Data are from triplicate blots from one experiment with three mice. Error bars indicate SD.

at the VDJ versus Iμ promoters. Expression of Iμ transcripts has been shown to be constitutive before and after mitogen stimulation (17), but a direct comparison to VDJ transcripts has not been performed. RNA was isolated from $Ung^{-/-}$ B cells and analyzed as described in the supplemental materials and methods. There was a threefold increase in RNA levels
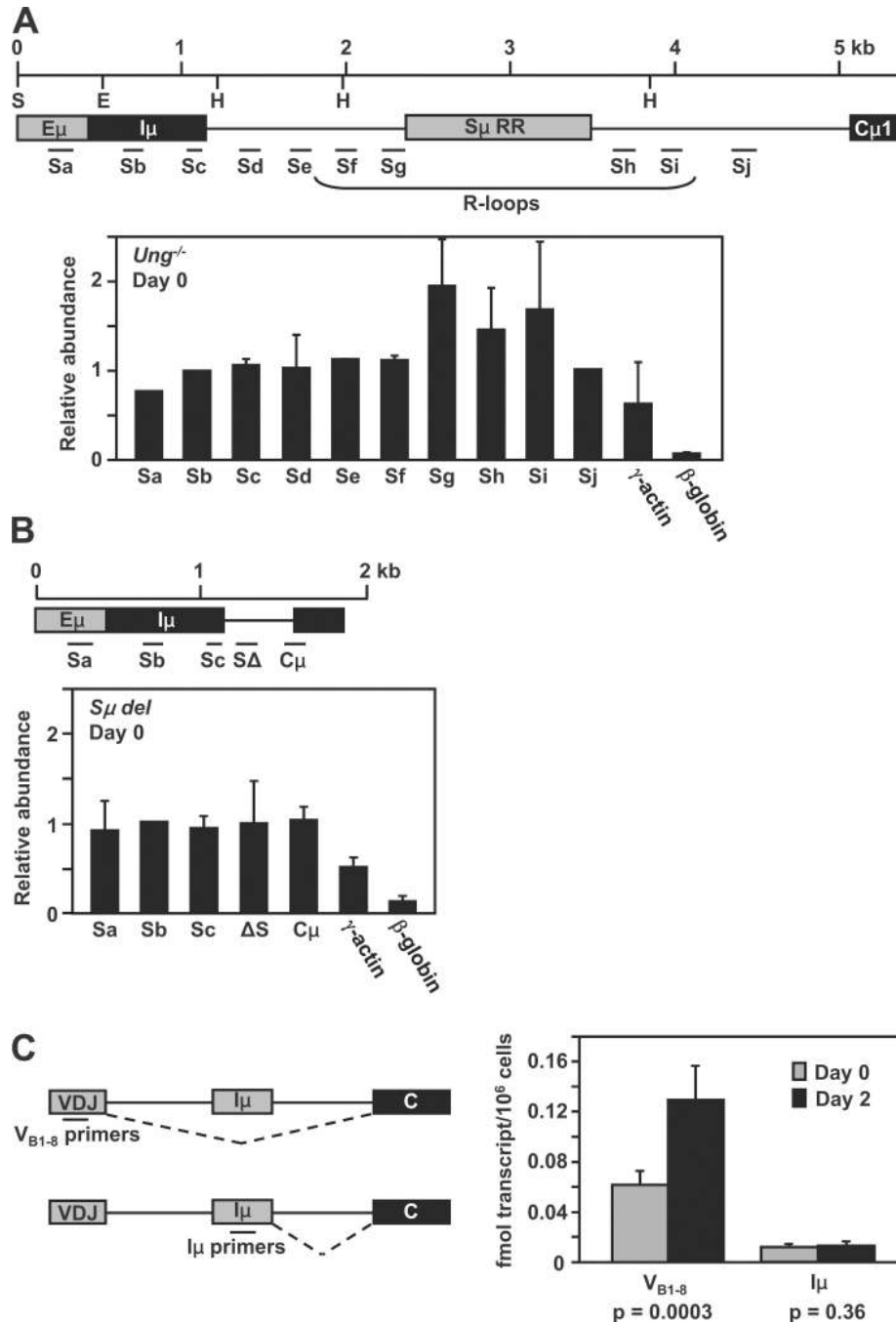

Figure 2. RNA analyses by ChIP and absolute qPCR. (A) A map of the position of PCR primers for the ChIP analysis is diagrammed. Data from $Ung^{-/-}$ splenic B cells at day 0 is shown below the map. The abundance of amplified sequences was graphed relative to the input DNA. The data are shown compared with a positive signal from γ-actin and negative signal from β-globin. Error bars represent the SD of values from two independent experiments with three mice each. (B) ChIP analysis in $S\mu$ del splenic B cells on day 0. The position of primers is shown with the data below. Error bars represent the SD of values from one experiment with three mice. (C) Absolute qPCR analysis of V_{B1-8} and Iμ RNA spliced transcripts from B1-8 cells before and after stimulation. The position of primers and splicing patterns are shown to the left. The data on the right is expressed as fmols of transcripts per million cells. Error bars represent the SD of values from two independent experiments with three mice each. Statistical significance comparing days 0 and 2 was determined by the student's t test.

per cell 2 d after activation (Fig. S2 A), which is likely the result of increased transcription of many genes in cells preparing for G1 progression. For instance, transcription of the β-actin gene increased after stimulation as measured by PCR (unpublished data) and Northern assays (Fig. S2 B). Northern blots also showed that Cμ spliced transcripts originating from VDJ promoters were in much greater abundance compared with Iμ spliced transcripts before and after stimulation (Fig. S2 B). We then quantified the amount of transcripts per cell by absolute quantitative PCR (qPCR; Fig. 2 C) using B1-8 mice, which have a homozygous knockin of a rearranged V$_H$J558 gene (18). Total cellular RNA from cells before and after ex vivo stimulation was measured with V$_{B1-8}$ and Iμ primers. Although these cells can switch, the primers should detect transcripts from any rearranged C$_H$ gene. The results show that VDJ transcripts increased twofold after ex vivo stimulation, whereas Iμ transcript levels remained constant. The data also show that spliced VDJ transcripts were 5–13-fold higher than Iμ transcripts before and after stimulation, respectively. Thus, transcripts initiating from the VDJ promoter may contribute substantially to forming secondary structure in the Sμ region.

Mutations increase before the Sμ repetitive region and decrease afterward

To evaluate the relationship between DNA structure and somatic hypermutation, we sequenced DNA in the Sμ region. This region has been previously sequenced in *Ung*$^{-/-}$*Msh2*$^{-/-}$ mice, and it showed that the DNA sequence upstream of the repetitive region was targeted for deamination by AID, as measured by G:C to A:T transitions (5). However, somatic hypermutation also consists of mutations of A:T bp, which are made during error-prone processing of uracils by MSH2-MSH6 and DNA pol η. To assess where the A:T mutations occurred, we analyzed the pattern in *Ung*$^{-/-}$ mice, which have a normal frequency of A:T mutations (15), using DNA from Peyer's patch and immunized splenic B cells. A 4.3-kb region spanning Eμ to downstream of the repetitive region was amplified in ∼500-bp segments that corresponded to the location of probes used for the run-on analysis. As shown in Fig. 3 A and Fig. S3, mutations were sparse around the Eμ-Iμ region and increased sequentially up to the repetitive region, with the highest frequency in S5 (Fig. 3 B). Although we could not sequence the repetitive region because of difficulties in PCR amplification, a high frequency of mutations is likely to be found there as well. After the repetitive region, the frequency fell precipitously, with the exception of a hot spot at T in an AGGCTGGGA motif in S7. Because no mutations were seen in S8, we did not analyze sequences further downstream.

Sμ DNA structure may limit activity of DNA pol η

As expected in *Ung*$^{-/-}$ mice, the vast majority of G:C mutations were transitions (Fig. S4). After correction for nucleotide composition, it was noted that the frequency of mutations of C increased relative to mutations of G in S5 as recorded from the nontranscribed strand (Fig. 3 C), which has also been reported in previous studies of this region (19–21). This is consistent

with R–loop structures exposing the nontranscribed strand for cytosine deamination. Furthermore, there was an unexpected decrease in the frequency of mutations of A bases in S4 and S5. As shown in Fig. 3 D, the frequency of mutations of A plummeted 10-fold to just 3% before the repetitive region, even though the germline dA density was high, and the WGCW density increased only twofold. We conclude that DNA pol η, which incorporates mutations of A during somatic hypermutation, may not be active in S4 and S5. The observation that the decrease in mutations of A was greater than the decrease in mutations of T is also consistent with a role for DNA pol η to introduce mutations on the nontranscribed strand (22, 23). The decline in A mutations begins at the proposed start of R–loops (11), implicating DNA structure as a reason for reduced recruitment of pol η. One way this could happen is by the inability of MSH2-MSH6, which interacts with pol η (2), to recognize single-strand DNA within R–loops and thereby not enlist pol η. Alternatively, pol η could be at the site but unable to copy in a repair patch if the template-transcribed strand is stably complexed to RNA. In either case, pol η would have less activity on R–loop structures. A potential benefit to preventing repair and synthesis by pol η would be to retain uracils for removal by UNG and subsequent nicking of the abasic site to induce switch recombination.

Setting up the Sμ locus for AID

Our interpretation is that the DNA sequence of the Sμ region influences the location of RNA pol II and hypermutation. Recently, Conticello et al. (24) identified a protein that interacts with AID and affects hypermutation, class switching, and gene conversion, possibly via the transcriptional spliceosome complex. One speculative model incorporating DNA structure, transcription, and mutation is presented in Fig. S5. RNA pol II molecules move down the DNA and then pause when they encounter the R–loop region. At the other end of the sequence, they are no longer stalled and proceed to the Cμ gene. Why would the polymerases pause? R–loops contain long RNA transcripts forming a stable hybrid with the transcribed DNA strand, and RNA polymerases may slow down when trying to unwind the tight RNA–DNA hybrid. Such blocks in transcription have been reported in vitro when the nontranscribed strand is G rich and the transcribed strand is C rich and can form R–loops (25, 26). If AID associates with the transcription complex, it would be paused as well and have the opportunity to bind to DNA. Along these lines, a recent study has described an in vitro system where AID-generated mutations were increased when T7 RNA polymerase was stalled (27). The decrease in mutation frequency after the repetitive region could occur if prolonged stalling caused AID to dissociate from the transcription complex in an unknown manner, resulting in fewer mutations. Alternatively, the local concentration of AID molecules may become depleted after being bound to DNA in the repetitive sequence. This model could be tested by detecting AID molecules bound to DNA before and after the repetitive region.

We also show that secondary structure affects the activity of DNA pol η because mutations of dA decrease at the start of R–loop formation. Therefore, both the pileup of RNA
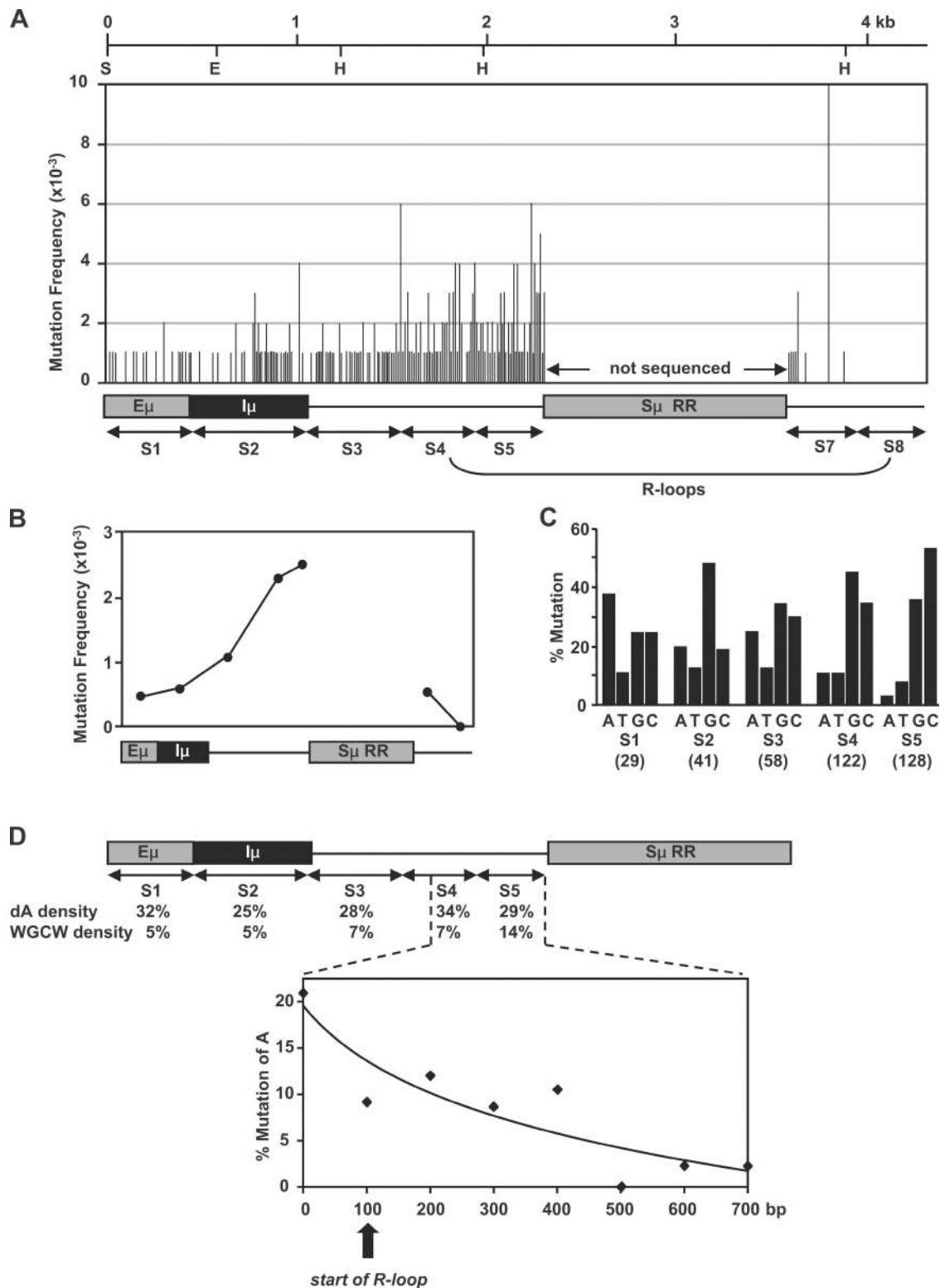
Figure 3. Location of hypermutation in the Sμ region from *Ung⁻/⁻* mice. The region was sequenced in ∼500-bp segments (S1–8). DNA from Peyer's patch and splenic B cells was analyzed; the Sμ repetitive region (RR) was not sequenced. (A) Mutation frequency (mutations/bp) at each nucleotide position. (B) Total mutation frequency for each segment. (C) The indicated number of mutations in parentheses was corrected for nucleotide composition of each fragment, and the percentage of mutation at each base is shown for S1–5. The detailed mutation analysis is recorded in Fig. S4. (D) Mutations of A bases decreased in S4 and S5. The density of dA and WGCW motifs is calculated for each segment. Data for S4 and S5 are then plotted below in 100-bp increments. The vertical arrow shows the proposed start of R-loop formation (11).

polymerases and the reduction of dA mutations provide independent and functional measures of DNA secondary structure in Sμ. In contrast, frog Sμ regions, which do not have G-clusters or R-loops, do not switch as efficiently as mouse S regions (28). Thus, the mouse Sμ structure appears to be the perfect substrate to amass mutations and associated strand breaks for switching. The pattern of RNA pol II accumulation preexists before B cell stimulation at day 0, which is most likely caused by R-loop formation from RNA polymerases transcribing membrane IgM. Both the stalling of transcription complexes and presence of single-strand DNA on the nontranscribed strand would give AID molecules optimal access to initiate the mutagenic cascade. This secondary structure may also explain why mutations readily arise in the Sμ region but not in nearby variable regions in B cells stimulated ex vivo.

MATERIALS AND METHODS

Mice. $Ung^{+/-}$ mice on a mixed C57BL/6J × 129SV background were obtained from H. Krokan (Norwegian University of Science and Technology, Trondheim, Norway). $Aid^{+/-}$ mice on a C57BL/6 background were provided by T. Honjo (Kyoto University, Kyoto, Japan), courtesy of M. Scharff (Albert Einstein College of Medicine, Bronx, NY). B1-8 mice on a C57BL/6 background were obtained from R. Casellas (National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD). The mice were bred in our animal facilities and used at 3–6 mo of age. Spleens from $S\mu$ del mice were provided by A. Khamlichi (16). All animal procedures were reviewed and approved by the National Institute on Aging Animal Care and Use Committee.

Stimulation of splenic B cells. Spleen cells were treated with ACK lysing buffer (Quality Biological, Inc.) to lyse red blood cells, and B cells were purified by negative selection using anti-CD43 and anti-CD11b antibodies coupled to magnetic beads (Miltenyi Biotec). Cells were cultured at 10^6 cells/ml and stimulated with 5 μg/ml of *Escherichia coli* lipopolysaccharide serotype 0111:B4 (Sigma-Aldrich) and 5 ng/ml of recombinant mouse IL-4 (BD). The cells were harvested on days 0 and 2 after stimulation.

Nuclear run-on. Isolation of nuclei and preparation of nuclear RNA were performed as previously described (29) with some modifications. In brief, ~5 × 10^8 cells were pelleted, washed with PBS, resuspended in ice-cold lysis buffer (20 mM Tris-HCl, pH 7.4, 20 mM NaCl, 5 mM MgCl$_2$, and 0.25% vol/vol NP-40), and incubated on ice for 10 min. Nuclei were spun down and resuspended in storage buffer (50 mM Tris-HCl, pH 8.0, 5 mM MgCl$_2$, 0.1 mM EDTA-NaOH, pH 8.0, and 45% vol/vol glycerol). The nuclei in 200-μl aliquots were then mixed with 200 μl of reaction buffer (300 mM KCl, 10 mM MgCl$_2$, and 1 mM each of ATP, CTP, and GTP) plus 500 μCi α-[^{32}P]UTP (3,000 Ci/mmol, 10 mCi/ml; PerkinElmer) and incubated for 30 min at 30°C. Samples were then incubated with 100 U of RNase-free DNase I (Applied Biosystems) for 30 min at 37°C and 20 μg/ml of proteinase K for 45 min at 37°C. Finally, the labeled nascent RNA was purified by Sephadex G-25 column filtration (GE Healthcare) and hybridized to DNA probes as described in the following section.

Preparation of DNA probes and hybridization. The Sμ region was amplified from genomic DNA with the primers listed in Table S1 and Taq DNA polymerase (Takara Bio Inc.). 1 μg PCR products was then printed on Hybond N$^+$ positively charged nylon membranes (GE Healthcare) as described in the manufacturer's protocol (Bio-Dot Microfiltration apparatus; Bio-Rad Laboratories). The membranes were prehybridized overnight with 3 ml MicroHyb hybridization solution (Invitrogen), 10 μg of Cot DNA (Invitrogen), and 8 μg of poly A DNA (Invitrogen). The labeled RNA was then added in 1 ml of hybridization solution for 24 h in a rotisserie-style incubator at 62°C. The membranes were rinsed in 2× SSC and 0.1% SDS at

62°C, followed by washes in SSC and 0.1% SDS twice at 62°C. Membranes were exposed for 2 d and scanned using a Phosphorimager (GE Healthcare). ImageQuant software (GE Healthcare) was used to convert the hybridization signals into raw intensity values which were analyzed.

ChIP. ChIP assays were performed as previously described (30), using anti-RNA pol II antibodies (Millipore). The DNA from immunoprecipitation was quantified using Pico green (Invitrogen) and amplified by qPCR using the primers listed in Table S2. Each primer set was examined for >90% amplification efficiency and for the lack of secondary products.

Absolute qPCR. Total cellular RNA was collected from unstimulated (day 0) and stimulated (day 2) B cells from B1-8 mice. Complementary DNA (cDNA) was produced using 0.5 μg and 1.0 μg RNA and the iScript cDNA synthesis kit (Bio-Rad Laboratories). The Iμ standard was amplified using cDNA template and primers Iμ-5′ 51019 (5′-GCTTGAGTAGTTC-TAGTTTCCCCAAACTTAAG-3′) and Iμ-3′ 50615 (5′-GAGTTGGTG-GTTGGTCGTACAAGTTG-3′) and cloned into a pGEM-T-easy (Promega) cloning vector. Similarly, the V$_{B1-8}$ standard was amplified using primers V$_{B1-8}$-2F (5′-CTGAGCACACAGGACCTCACC-3′) and V$_{B1-8}$-2R(5′-GGACT-CACCTGAGGAGACTGTG-3′). For the standard curve, plasmid standards were linearized with ScaI-HF (New England Biolabs, Inc.) and purified by agarose gel electrophoresis and gel extraction (QIAGEN). qPCR was performed using IQ SYBR green supermix (Bio-Rad Laboratories) with primers Iμ-5′ 50827 (5′-CCAATACCCGAAGCATTTACAGTGAC-3′) and Iμ-3′ 50726 (5′-GTGAAGCCGTTTTGACCAGAATGTC-3′) and with primers V$_{B1-8}$-3F (5′-GACGAGGCCTTGAGTGGATTG-3′) and V$_{B1-8}$-3R (5′-CATGTAGGCTGTGCTGGAGG-3′). Each primer set was examined for >90% amplification efficiency and for the lack of secondary products. Reactions were performed using a 7900HT real-time instrument (Applied Biosystems) and data analyzed using SDS 2.3 software (Applied Biosystems).

Hypermutation. Activated B cells were obtained from the Peyer's patches of three mice, and they were obtained from spleens of four mice immunized with sheep red blood cells for 1 mo and sacrificed 4 d after a boost. Cells were stained with phycoerythrin-labeled antibody to B220 (eBioscience) and fluorescein-labeled peanut agglutinin (PNA; E-Y Laboratories). The cells were separated by flow cytometry, and DNA was prepared from B220$^+$PNA$^+$ cells. Libraries were made by amplifying DNA using the primers listed in Table S1. 100 ng DNA was amplified using Pfx DNA polymerase and PCR enhancer (Invitrogen) in a 50-μl vol using the external primers for 35 cycles. Nested PCR was performed with 5 μl of the first reaction and the internal primers for another 35 cycles. The products were digested, cloned into pBluescript (Agilent Technologies), and sequenced. Only clones with unique patterns of mutation were recorded.

Online supplemental material. The supplemental materials and methods describe RNA isolation and Northern assays. Fig. S1 presents a model to compare polymerase density by ChIP and run-on assays. Fig. S2 shows the promoter analysis of RNA by Northern blots. Fig. S3 shows the location of mutations in the Sμ region. Fig. S4 displays the spectrum of mutations for each base in the S1–7 regions. Fig. S5 shows a potential model for AID and DNA pol η activity in the Sμ region. Table S1 lists the primers used for sequencing and run-on assays. Table S2 lists the primers used for ChIP analysis. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20082514/DC1.

REFERENCES

1. Stavnezer, J., J.E. Guikema, and C.E. Schrader. 2008. Mechanism and regulation of class switch recombination. *Annu. Rev. Immunol.* 26:261–292.

2. Wilson, T.M., A. Vaisman, S.A. Martomo, P. Sullivan, L. Lan, F. Hanaoka, A. Yasui, R. Woodgate, and P.J. Gearhart. 2005. MSH2–MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations in antibody genes. *J. Exp. Med.* 201:637–645.

3. Guikema, J.E., E.K. Linehan, D. Tsuchimoto, Y. Nakabeppu, P.R. Strauss, J. Stavnezer, and C.E. Schrader. 2007. APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. *J. Exp. Med.* 204:3017–3026.

4. Lennon, G.G., and R.P. Perry. 1985. C_μ-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5′-nontranslatable exon. *Nature.* 318:475–478.

5. Xue, K., C. Rada, and M.S. Neuberger. 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.

6. Peters, A., and U. Storb. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity.* 4:57–65.

7. Nambu, Y., M. Sugai, H. Gonda, C.G. Lee, T. Katakai, Y. Agata, Y. Yokota, and A. Shimizu. 2003. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science.* 302:2137–2140.

8. Yu, K., F. Chedin, C.L. Hsieh, T.E. Wilson, and M.R. Lieber. 2003. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* 4:442–451.

9. Duquette, M.L., P. Handa, J.A. Vincent, A.F. Taylor, and N. Maizels. 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* 18:1618–1629.

10. Shen, H.M., and U. Storb. 2004. Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled. *Proc. Natl. Acad. Sci. USA.* 101:12997–13002.

11. Huang, F.T., K. Yu, B.B. Balter, E. Selsing, Z. Oruc, A.A. Khamlichi, C.L. Hsieh, and M.R. Lieber. 2007. Sequence dependence of chromosomal R-loops at the immunoglobulin heavy-chain Sμ class switch region. *Mol. Cell. Biol.* 27:5921–5932.

12. Perlot, T., G. Li, and F.W. Alt. 2008. Antisense transcripts from immunoglobulin heavy-chain locus V(D)J and switch regions. *Proc. Natl. Acad. Sci. USA.* 105:3843–3848.

13. Nagaoka, H., M. Muramatsu, N. Yamamura, K. Kinoshita, and T. Honjo. 2002. Activation-induced deaminase (AID)–directed hypermutation in the immunoglobulin Sμ region: implication of AID involvement in a common step of class switch recombination and somatic hypermutation. *J. Exp. Med.* 195:529–534.

14. Reina-San-Martin, B., S. Difilippantonio, L. Hanitsch, R.F. Masilamani, A. Nussenzweig, and M.C. Nussenzweig. 2003. H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. *J. Exp. Med.* 197:1767–1778.

15. Rada, C., G.T. Williams, H. Nilsen, D.E. Barnes, T. Lindahl, and M.S. Neuberger. 2002. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr. Biol.* 12:1748–1755.

16. Khamlichi, A.A., F. Glaudet, Z. Oruc, V. Denis, M. Le Bert, and M. Cogne. 2004. Immunoglobulin class-switch recombination in mice devoid of any S mu tandem repeat. *Blood.* 103:3828–3836.

17. Li, S.C., P.B. Rothman, J. Zhang, C. Chan, D. Hirsh, and F.W. Alt. 1994. Expression of I mu-C gamma hybrid germline transcripts subsequent to immunoglobulin heavy chain class switching. *Int. Immunol.* 6:491–497.

18. Shih, T.A., M. Roederer, and M.C. Nussenzweig. 2002. Role of antigen receptor affinity in T cell-independent antibody responses in vivo. *Nat. Immunol.* 3:399–406.

19. Zeng, X., G.A. Negrete, C. Kasmer, W.W. Yang, and P.J. Gearhart. 2004. Absence of DNA polymerase η reveals targeting of C mutations on the nontranscribed strand in immunoglobulin switch regions. *J. Exp. Med.* 199:917–924.

20. Martomo, S.A., W.W. Yang, R.P. Wersto, T. Ohkumo, Y. Kondo, M. Yokoi, C. Masutani, F. Hanaoka, and P.J. Gearhart. 2005. Different mutation signatures in DNA polymerase eta- and MSH6-deficient mice suggest separate roles in antibody diversification. *Proc. Natl. Acad. Sci. USA.* 102:8656–8661.

21. Martomo, S.A., W.W. Yang, and P.J. Gearhart. 2004. A role for Msh6 but not Msh3 in somatic hypermutation and class switch recombination. *J. Exp. Med.* 200:61–68.

22. Mayorov, V.I., I.B. Rogozin, L.R. Adkison, and P.J. Gearhart. 2005. DNA polymerase eta contributes to strand bias of mutations of A versus T in immunoglobulin genes. *J. Immunol.* 174:7781–7786.

23. Unniraman, S., and D.G. Schatz. 2007. Strand-biased spreading of mutations during somatic hypermutation. *Science.* 317:1227–1230.

24. Conticello, S.G., K. Ganesh, K. Xue, M. Lu, C. Rada, and M.S. Neuberger. 2008. Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. *Mol. Cell.* 31:474–484.

25. Tous, C., and A. Aguilera. 2007. Impairment of transcription elongation by R-loops in vitro. *Biochem. Biophys. Res. Commun.* 360:428–432.

26. Tornaletti, S., S. Park-Snyder, and P.C. Hanawalt. 2008. G4-forming sequences in the non-transcribed DNA strand pose blocks to T7 RNA polymerase and mammalian RNA polymerase II. *J. Biol. Chem.* 283:12756–12762.

27. Canugovi, C., M. Samaranayake, and A.S. Bhagwat. 2009. Transcriptional pausing and stalling causes multiple clustered mutations by human activation-induced deaminase. *FASEB J.* 23:34–44.

28. Zarrin, A.A., F.W. Alt, J. Chaudhuri, N. Stokes, D. Kaushal, L. Du Pasquier, and M. Tian. 2004. An evolutionarily conserved target motif for immunoglobulin class-switch recombination. *Nat. Immunol.* 5:1275–1281.

29. Cheadle, C., J. Fan, Y.S. Cho-Chung, T. Werner, J. Ray, L. Do, M. Gorospe, and K.G. Becker. 2005. Control of gene expression during T cell activation: alternate regulation of mRNA transcription and mRNA stability. *BMC Genomics.* 6:75.

30. Chakraborty, T., D. Chowdhury, A. Keyes, A. Jani, R. Subrahmanyam, I. Ivanova, and R. Sen. 2007. Repeat organization and epigenetic regulation of the D_H–Cμ domain of the immunoglobulin heavy-chain gene locus. *Mol. Cell.* 27:842–850.