Target DNA Sequence Directly Regulates the Frequency of Activation-Induced Deaminase-Dependent Mutations

Zhangguo Chen, Sawanee S. Viboolsittiseri, Brian P. O’Connor and Jing H. Wang

*J Immunol* 2012; 189:3970-3982; Prepublished online 7 September 2012;
doi: 10.4049/jimmunol.1200416
http://www.jimmunol.org/content/189/8/3970
Target DNA Sequence Directly Regulates the Frequency of Activation-Induced Deaminase-Dependent Mutations

Zhangguo Chen,* Sawanee S. Viboolsittiseri,* Brian P. O’Connor,† and Jing H. Wang*†

Activation-induced deaminase (AID) catalyzes class switch recombination (CSR) and somatic hypermutation (SHM) in B lymphocytes to enhance Ab diversity. CSR involves breaking and rejoining highly repetitive switch (S) regions in the IgH (Igh) locus. S regions appear to be preferential targets of AID. To determine whether S region sequence per se, independent of Igh cis regulatory elements, can influence AID targeting efficiency and mutation frequency, we established a knock-in mouse model by inserting a core Sy1 region into the first intron of the proto-oncogene Bcl6, which is a non-Ig target of SHM. We found that the mutation frequency of the inserted Sy1 region was dramatically higher than that of the adjacent Bcl6 endogenous sequence. Mechanistically, S region-enhanced SHM was associated with increased recruitment of AID and RNA polymerase II, together with Spt5, albeit to a lesser extent. Our studies demonstrate that target DNA sequences influence mutation frequency via regulating AID recruitment. We propose that the nucleotide sequence preference may serve as an additional layer of AID regulation by restricting its mutagenic activity to specific sequences despite the observation that AID has the potential to access the genome widely. The Journal of Immunology, 2012, 189: 3970–3982.

I

mmunoglobulin genes undergo somatic hypermutation (SHM) and class switch recombination (CSR) during the course of B lymphocyte activation in well-orchestrated structures termed germinal centers (GCs) (1). IgH and L chain V region exons acquire point mutations during the SHM process at a rate of \( \sim 1 \times 10^{-3} \) per base pair per generation (2, 3). Transcription of Ig genes is required but not sufficient for SHM (4, 5). SHM targets regions between \( \sim 200 \) and 1500 bp downstream of the transcription start site to the intronic region containing the J sequences, generally sparing Ig C region exons (4, 5). These mutations increase the DNA sequence diversity of V region exons, which allows the selection of B cell clones with increased affinity for Ag (2, 3). SHM can also target non-Ig gene loci, albeit at much lower frequency (4, 6). Among them, Bcl6 is a frequently mutated non-Ig target of SHM (4, 6–8). Bcl6 encodes a zinc finger transcriptional repressor expressed in GC B cells and is required for GC formation (9, 10). Rearrangements of the Bcl6 gene are found in 30–40% of diffuse large B cell lymphomas, clustering in a 4-kb major breakpoint region in the first noncoding exon and 5′ region of the first intron (11, 12). SHM can only be induced by in vivo immunization or infection and occurs efficiently in GC B cells (1).

IgH CSR modulates the class and effector functions of Abs. Newly generated naïve B cells express IgM on the surface. CSR allows the assembled V(D)J exon to be expressed with one of the sets of downstream C\(_H\) exons (referred to as C\(_H\) genes) and enables production of different IgH classes (e.g., IgG, IgE, and IgA) encoded by different C\(_H\) genes (e.g., C\(_{\gamma}\), C\(_{\varepsilon}\), and C\(_{\alpha}\)) (13). CSR occurs between repetitive switch (S) regions located 5′ of each C\(_{\varepsilon}\) exon except C\(_{\alpha}\). Mammalian S regions are 2–10 kb in length, usually rich G rich on the nontemplate strand, and primarily composed of tandem repetitive sequences within which certain motifs predominate (13, 14). S\(_{q}\) is exceptionally repetitive and enriched in GAGCT motifs, with the AGCT palindrome representing a canonical RGYW (where R = purine, Y = pyrimidine, and W = A or T nucleotide) SHM motif. Sy1 is the largest S region and, like other S\(_{q}\) sequences, carries multiple 49-bp repeats that are rich in RGYW motifs (13, 14). Transcribed mammalian S regions form R-loop structures (15, 16) and appear to be the main CSR targets (17–19). CSR can be induced by both in vivo immunization and in vitro activation with different combinations of cytokines that enable the accessibility of a given S region for recombination (13, 14).

Both SHM and CSR require activation-induced deaminase (AID) (20), which deaminates deoxycytidines to deoxyuridines and creates deoxyuridine:deoxyguanosine lesions (21). The regulation of AID targeting efficiency and specificity is not completely understood (13, 22). Specific AID targeting could reflect factors such as transcription rate, specific features of DNA sequences, effects of cis elements, or chromatin modifications. Unlike V(D)J recombinase RAG, which specifically targets recombination signal sequences (23), no specific sequences have been identified for preferential AID targeting. It has been proposed that special DNA motifs such as RGYW/WRCY hotspots play an important role in attracting SHM (24). These studies were largely based on molecular analysis of endogenous V region sequences and suggest a correlation between hotspot positions and mutations (2). S regions appear to be preferential AID targets during CSR (18, 25), likely because of several unique aspects associated with their highly GC-rich repetitive sequences (26). An intriguing question is whether possessing such specific sequence features alone is sufficient to increase the mutability of the sequence, thereby influencing muta-

*Integrated Department of Immunology, University of Colorado School of Medicine and National Jewish Health, Denver, CO 80206; and †Center for Genes, Environment and Health, National Jewish Health, Denver, CO 80206

Received for publication February 1, 2012. Accepted for publication August 13, 2012.

This work was supported by University of Colorado School of Medicine startup funds (J.H.W.), the Leukemia Research Foundation (J.H.W.), and a Boettcher Foundation Webb-Waring Biomedical Research award (to J.H.W.).

Address correspondence and reprint requests to Dr. Jing H. Wang, University of Colorado/National Jewish Health, 1400 Jackson Street, K-8, Denver, CO 80206. E-mail address: jing.wang@ucdenver.edu

The online version of this article contains supplemental material.

Abbreviations used in this article: AID, activation-induced deaminase; ChIP, chromatin immunoprecipitation; CSR, class switch recombination; ES, embryonic stem; GC, germinal center; IP, immunoprecipitation; KLH, keyhole limpet hemocyanin; PNA, peanut agglutinin; qPCR, quantitative PCR; RNAPII, RNA polymerase II; S switch; SHM, somatic hypermutation; SRBC, sheep RBC; wt, wild type.

Copyright © 2012 by The American Association of Immunologists, Inc. 0022-1767/12/$16.00

www.jimmunol.org/cgi/doi/10.4049/jimmunol.1200416
tion frequency. However, a model to test this possibility is lacking, and it has not been clearly addressed whether target DNA sequence, as the substrate of AID’s mutagenic activity, influences mutation frequency.

Before the discovery of AID, it was proposed that the “mutator factor” travels with transcribing RNA polymerase and is deposited on the target DNA sequence when the polymerase pauses (4, 27). Recent studies show that RNA polymerase II (RNAPII) pauses at S regions in cytokine-activated B cells during CSR; furthermore, S region sequences are important contributors to the increased occupancy of RNAPII (28, 29). Apart from pausing at S regions, RNAPII associates with pause-inducing factors such as DRB sensitivity-inducing factor shortly after transcription initiation, and the resulting RNAPII complex often pauses in the promoter proximal region and does not progress into elongation (30–33). In humans and mice, DRB sensitivity-inducing factor is composed of Spt4 and Spt5, the latter of which has been shown to function as an AID cofactor during CSR via interacting with AID and facilitating its association with RNAPII (34). However, it remains to be determined whether and how Spt5 contributes to RNAPII pausing at S regions. Of note, most of these studies were performed in the context of CSR. It thus remains largely unknown how target DNA sequence regulates AID activity in B cells undergoing SHM. So far, no experimental evidence has demonstrated that target DNA sequence can regulate mutation frequency via recruitment of AID, RNAPII, or Spt5 in the context of SHM in vivo.

To determine whether S region sequences per se, independent of Igh cis regulatory elements, can regulate AID targeting efficiency, we inserted a core S region (35) into the first intron of the mouse Bcl6 gene via a gene-targeting approach. This strategy excluded the influence of Igh cis regulatory elements, thereby allowing a direct test of the role of DNA sequences in regulating AID activity. The core S region (4 kb) was chosen because it supports CSR efficiently in vivo in its endogenous Igh locus (35). We generated Bcl6S<sup>S<i>ω</i></sup> knock-in mice and assayed the ability of the wild type (wt) and Bcl6S<sup>S<i>ω</i></sup> alleles to serve as AID targets in primary GC B cells undergoing SHM. We found that S region sequence per se, independent of Igh cis regulatory elements, enhanced AID targeting efficiency; moreover, target DNA sequence could determine the mutation frequency in the context of SHM.

**Materials and Methods**

**Embryonic stem cells targeting and generation of Bcl6<sup>S<ω</sup></sup> knock-in mice**

A 4-kb Xbal-Not I S<sup>ω</sup> homology arm from a plasmid containing Bcl6 genomic DNA (a gift from Dr. R. Dalla-Favera, Columbia University, New York, NY) was cloned into an XhoI site in the pLNtk targeting vector. A 3.5-kb NotI-SalI fragment containing Bcl6 genomic DNA was used as the 3<sup>ω</sup> homology arm. The configuration of the targeted allele was further confirmed by additional Southern blots with probes that hybridized to the upstream of the 3<sup>ω</sup> homology arm. The targeting construct was transfected into TC1 embryonic stem (ES) cells. Correctly targeted clones were detected by Southern blot with two probes that hybridized upstream of the 5<sup>ω</sup> homology arm or downstream of the 3<sup>ω</sup> homology arm. The 5<sup>ω</sup> probe was a 1.5-kb XbaI-Xbal fragment upstream of the first exon of Bcl6, and the 3<sup>ω</sup> probe was a 1.4-kb BamHI-HindIII fragment from the first intron of Bcl6. The configuration of the targeted allele was further confirmed by additional Southern blots with probes that hybridized to the S region (35). For the deletion of neo<sup>c</sup> cassette through two flanking loxp sites, targeted ES clones were infected with recombiant adenovirus that expressed Cre recombinase. The targeted ES cells were injected into RAG<sup>−/−</sup> mice and analyzed by Southern blot using the 5<sup>ω</sup>-A TCACCACACTTCCACCTCTCCTGGCTAACCTAT-3<sup>ω</sup> and 5<sup>ω</sup>-GACTA-GAGTGTGGGGGCTGGTGTCATCATC-3<sup>ω</sup> primers (35); primer (35) S region, 5′-ATCACCACACTTCCACCTCTGCTGATACCTAT-3′; primer (Bcl6), 5′-GAATCCAAAAAGTACCCCTCTCCCTCCTC-3′; primer C (S region), 5′-ATCACCACACTTCCACCTCTGCTGATACCTAT-3′. For additional Southern blot analysis, DNA was extracted from sorted GC B cells using the 5′-A TCACCACACTTCCACCTCTGCTGATACCTAT-3′ and 5′-GAATCCAAAAAGTACCCCTCTCCCTCCTC-3′ primers. Sequencing was performed on the S region using the S<sup>ω</sup> and wt intronic primers.

**Cell sorting and flow cytometry**

For SHM studies, spleen and mesenteric lymph nodes were dissected from immunized mice and single-cell suspensions were made. Cells were stained with PNA-FTTC (Vector) and B220-PE, and CD3-FITC and B220-PE, and were sorted for B220<sup>+</sup>/PNA<sup>+</sup> cells. For analysis of GC B cells, mice were immunized with SRBC Ag on days 0 (10<sup>9</sup>) and 5 (10<sup>8</sup>). Ten days after immunization, four to eight mice were sacrificed, and spleen single-cell suspensions were prepared. After depletion of RBCs with RBC lysis buffer (Sigma, St. Louis, MO), B cells were purified using a mouse B cell enrichment kit (Stem Cell Technologies, Vancouver, BC, Canada) according to manufacturer’s instructions with the purity usually reaching >90%. Purified B cells from four to eight mice were pooled into 1.5 ml FACS buffer (2% FBS in PBS). Cells were stained

**Immunofluorescent staining and immunohistochemistry**

For immunofluorescence analyses of GCs, spleens were harvested and embedded in OCT compound (Miles, Elkhart, IN), and frozen at −70˚C. Frozen sections (8–10 µm thick) mounted on slides were fixed in cold acetone. After washing in PBS, the sections were stained for B cells by incubation with PE-conjugated anti-B220 (BD. San Diego, CA); Slides were mounted with Moviol medium and analyzed with a Leica fluorescent microscope; pictures were obtained with a Sony digital photo camera DKC-5000 and processed with Adobe Photoshop software. For GC analysis in Peyer’s patch, small intestine samples were fixed in 10% buffered formalin and processed for paraffin embedding using standard methods. Bcl6 expression was detected with a rabbit antibody against Bcl6 (Santa Cruz, CA) followed by addition of the alkaline phosphatase substrate p-nitrophenyl phosphate (Sigma) at 1 mg/ml. The mean OD at 405 nm from duplicate wells is presented as the mean ± SD.

**Mutation analysis**

For Bcl6 locus mutation analysis, DNA was extracted from sorted GC B cells (B220<sup>+</sup> PNA<sup>+</sup>) from the immunized mice or from kidney as a control for germline sequence. iProof High-Fidelity DNA Polymerase (Bio-Rad, Hercules, CA) was used to amplify the first intron region of Bcl6 on SRBC immunized mice. Primer sets (5′-CAGTGTTTTGGGAAGGCTTCGCTTGC-3′, forward; 5′-GGGGTGGGGTCATCATCCT-3′, reverse) were designed to amplify the first intron region as a NotI-SalI fragment (35). PCR products were cloned into the pGEM easy vector (Promega), and clones were sequenced. Sequences were aligned with DNA-Star/SeqMan software and were compared with the corresponding genomic sequences of Bcl6 (accession number: MGI 107187, http://www.informatics.jax.org/marker/MGI:107187) or VpJ558/FR3 distal sequence (accession number: A851868, http://www.ncbi.nlm.nih.gov/nuccore/126349412). Primers for wt and S<sup>ω</sup> allele of Bcl6 locus were: forward primer A (5′-CAGTGTTTTGGGAAGGCTTCGCTTGC-3′); reverse primer B (Bcl6), 5′-GAATCCAAAAAGTACCCCTCTCCCTCCTC-3′; primer C (S region), 5′-ATCACCACACTTCCACCTCTGCTGATACCTAT-3′. For primers VpJ558/FR3 and JH4 intronic region were: forward primer, 5′-CAGTGTTTTGGGAAGGCTTCGCTTGC-3′; reverse primer, 5′-GACTAGTTCCTTCTACAGTTTGCCTGCTGCTG-3′.

**Mutation frequency**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Control wt</th>
<th>Bcl6&lt;sup&gt;S&lt;i&gt;ω&lt;/i&gt;&lt;/sup&gt;</th>
<th>p&lt;sub&gt;&lt;i&gt;value&lt;/i&gt;&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Bcl6</td>
<td>0.07 ± 0.02</td>
<td>0.12 ± 0.01</td>
<td>0.46</td>
</tr>
<tr>
<td>Bcl6</td>
<td>0.10 ± 0.03</td>
<td>0.05 ± 0.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*P<sub>value</sub>* values were calculated using the Wilcoxon rank-sum test. The mean ± SD of the data is shown in Table 2.
with FITC-conjugated anti-PNA (Vector Laboratories, Burlingame, CA) and PE-conjugated anti-B220 (BD Biosciences, Sparks, MD) for 20 min at 4°C. After one wash, cells were resuspended in 1 or 2 ml FACS buffer and sorted by Beckman Coulter MoFloXDP Sorter (Miami, FL) for B220+ PNA<sup>αβ</sup> (GC B cells) and B220+PNA<sup>αβ</sup> (naïve B cells) populations. The PNA<sup>αβ</sup>B220<sup>+</sup> (>90% purity) population was collected for subsequent chromatin immunoprecipitation (ChIP) assay.

Chromatin preparation. Three to 6 × 10<sup>6</sup> freshly isolated GC B cells were resuspended in 1 ml lymphocyte culture media (RPMI 1640 with non-essential amino acids, sodium pyruvate, 10% FBS, 10 mM HEPES, and 3 μM 2-ME). Chromatin was cross-linked by adding 30 μl of 37% formaldehyde (Fisher Scientific, Fairlawn, NJ), and the tubes were rocked at room temperature for 7 min. Cross-linking was stopped immediately by adding 120 μl of 10× glycine, and the tubes were rocked at room temperature for 7 min. The ChIP-IT Express Kit (Active Motif, Carlsbad, CA) was used for the following steps according to manufacturer’s instructions. In brief, cells were pelleted by centrifugation for 10 min at 2500 rpm at 4°C; 2) cells were washed twice with ice-cold PBS; 3) cells were lysed with 600 μl ice-cold lysis buffer (supplemented with 6 μl Protease Inhibitor Cocktail [Roche Applied Science, Indianapolis, IN] + 6 μl PMSF); 4) tubes were briefly vortexed and rotated at 4°C for 30 min; 5) lysates were spun at 5000 rpm for 10 min at 4°C to pellet nuclei; 6) nuclei pellet was resuspended in 300 μl shearing buffer supplemented with 3 μl Protease Inhibitor Cocktail + 3 μl PMSF; 7) chromatin was sheared with a Covaris S2 sonicator (Covaris, Woburn, MA) using the following settings: 20% Duty cycle; 5 Intensity; 200 cycles per burst; 30 s per cycle; total cycle for 24; bath temperature at 4°C; 8) sheared chromatin samples were spun at 15,000 rpm at 4°C for 10 min; and 9) supernatant was collected and aliquoted as chromatin.

Immunoprecipitation. Anti-AID Ab was described previously (36). Anti-RNAPII Ab was purchased from Abcam (Cambridge, MA) and anti-SPT5 Abs. IP from Santa Cruz Biotechnology (Santa Cruz, CA). Control Igs were selected and aliquoted as chromatins.

Quantitative PCR. Input and IP samples were purified with a QIAGEN PCR purification kit. The eluted DNA was used as a template for quantitative PCR (qPCR). Template and primers were added into the Absolute QPCR SYBR Green ROX Mix (Thermo Scientific) in a total volume of 20 μl per reaction. The reactions were run on 7900 Fast Real-Time PCR system (AB Applied Biosystems, Carlsbad, CA). PCR conditions were as follows: 1) 50°C 2 min, 95°C 10 min; 2) 95°C 10 s, 62°C 30 s, 72°C 45 s, 40 cycles; 3) dissociation curve 95°C 1 s, 60°C 15 s, 95°C 5 s. Dissociation curves were evaluated after each reaction to ensure the specificity of the amplification. The primer sets used for qPCR are listed in Supplemental Fig. 2A.

Normalization of ChIP-qPCR data. Two common methods are used to normalize ChIP-qPCR data, namely, the percentage input and the fold enrichment. Our ChIP-qPCR data were normalized with the former method because it includes normalization for the amount of total chromatin present within each ChIP sample. The percentage input was calculated based on the equation 100 × [2<sup>-Ct(input) − Ct(IP) − Ct(Control IP)] and is presented in Figs. 6 and 7. In addition, another level of normalization was used based on the assumption that the IP of a specific Ab contains both specific signal and background, whereas the IP of a negative control-Im represents the background only. The background/input was subtracted from signal/input, and the remaining value corresponded to net pull-down of a specific DNA/chromatin region by a given specific Ab normalized to both total chromatin and nonspecific (negative control) IP. The normalized ChIP-qPCR data calculated with this secondary method are presented in Supplemental Fig. 3B-D.

Hybridoma generation

Hybridomas were generated from GC B cells sorted as described earlier. For each fusion, GC B cells and NS1 myeloma cells were fused with 50% PEG 1500 (Roche) and were subsequently cultured in 15% FBS, 2 mM glutamine, 100 U/ml penicillin-streptomycin, 100 μM 2-ME and hypoxanthine-aminopterin-thymidine (Sigma). Clones that secreted IgG or IgM were identified by ELISA. DNA was isolated from hybridomas and used for Southern blot analysis. For rearrangements of the Bcl6<sup>Syl</sup> allele, DNA was digested with HindIII and hybridized with a probe for the first intron of Bcl6.

Results

Generation of Bcl6<sup>Syl</sup> knock-in mice

A 4-kb core Syl region, which has been shown to support efficient CSR in vivo (35), was inserted into the first intron of Bcl6 by homologous recombination in ES cells to generate the Bcl6<sup>Syl</sup> allele (Fig. 1A). The Bcl6 locus was chosen because it is an AID target in GC B cells (6–8, 37) and a frequent translocation target in human B cell lymphomas (11, 12). The first intron region of mouse Bcl6 was chosen for targeting because the SHM and translocation breakpoints cluster in the first intron of human Bcl6 (6, 38). Thus, this experimental design allows us to test whether an S region can regulate the efficiency of AID targeting in a locus where a low level of SHM normally occurs. 5′ and 3′ homologous arms were cloned into the pLNTK targeting vector, and the 4kb Syl region was inserted downstream of the pgk-Neo<sup>+</sup> cassette (Fig. 1A). The targeting construct was transfected into 129Sv ES cells.
cells. The floxed pgk-neo cassette was deleted by Cre recombination, and the deletion was confirmed by Southern blot analysis using the Sy1 probe, which hybridizes to the endogenous Sy1 region (23 kb), the targeted (8 kb) and Cre-deleted Bcl6SY1 alleles (10 kb; Fig. 1B). The Cre-deleted ES cells were injected for germline transmission to generate Bcl6SY1/SY1 heterozygous (Het) knock-in mice, which were bred to yield homozygous (Hom) knock-in mice (Bcl6SY1/SY1). The insertion of the Sy1 region was confirmed by Southern blot analysis using 5’ and 3’ probes hybridizing to the endogenous Bcl6 locus (Fig. 1C). Bcl6SY1/SY1 homozygous knock-in mice were viable, fertile, and had normal lymphocyte development and CSR in vitro (data not shown).

Normal immune response in Bcl6SY1/SY1 homozygous mice

BCL6 is expressed in GC B cells and is required for GC formation (9, 10). To test whether Sy1 insertion affects the function of the Bc16 gene, we immunized the Bcl6SY1/SY1 mice and wt littermates with T cell-dependent Ags SRBC and KLH, and assessed GC and Ab responses. The GC reactions in immunized Bcl6SY1/SY1 and wt mice were comparable, whereas the unimmunized wt mice did not form GCs (Fig. 2A). Flow cytometry data also showed that the GC population defined as B220+PNAhigh was indistinguishable between immunized wt and Bcl6SY1/SY1 mice (data not shown). More importantly, similar levels of Ab response were observed in wt and Bcl6SY1/SY1 mice after immunization (Fig. 2B). Finally, we directly examined the expression of BCL6 protein in the Peyer’s patch of the small intestine by immunohistochemistry and found that the level of BCL6 expression was comparable between wt and Bcl6SY1/SY1 mice (Fig. 2C). Taken together, these data demonstrate that insertion of the core Sy1 region into the first intron of the Bc16 gene had no obvious effects on the expression and function of Bc16.

Mouse Bc16 is a target of AID-dependent mutation process

To compare the mutation frequency of the Bc16 first intron in the wt endogenous locus with that in the Bcl6SY1 knock-in locus, we first assayed for the mutation frequency of this region in the wt Bc16 locus. Genomic DNA was purified from the sorted B220+PNAhigh GC or B220+PNAlow naive B cells of the immunized Bcl6SY1/+ heterozygous mice. We used a PCR approach with specific primers designed for the wt versus Bcl6SY1 allele (Fig. 3A), and the amplified PCR products were subcloned and sequenced for the mutational analysis. The mutation frequency of the wt allele in the B220+PNAhigh population was significantly higher than that of the B220+PNAlow population or of genomic DNA from kidneys (2.389 × 10−4 versus <3.534 × 10−5 or 3.658 × 10−5; Table I), and most of the Bc16 locus mutations are substitutions except for one deletion and one insertion (Fig. 3B). Thus, our results further confirm that mouse Bc16 is a target of SHM (7, 8). Consistent with previous studies (7, 8), we found that no mutations in the Bc16 locus were detected in the GC B cells of AID−/− mice (Table I). Therefore, we conclude that mouse Bc16 is susceptible to an AID-dependent mutational process.

SHM occurs in the V regions of Ig genes, with the highest mutation rate immediately downstream of the V promoter, as well as a substantial rate in the downstream intron immediately after
To compare the level of mutations in the wt Bcl6 locus with that in the Igh locus in our immunization setting, we used the same GC B cell genomic DNA samples isolated from Bcl6\textsuperscript{Sg1/+} heterozygous mice described earlier and assayed for the mutation frequency in the JH4 intronic region using a forward primer common to FR3 of most VHJ558 family members and a reverse primer located in the JH4 intronic region (39). An extremely high load of mutations was found in the JH4 intronic region of B220\textsuperscript{+}PNA\textsuperscript{high} GC B cells compared with that of the B220\textsuperscript{+}PNA\textsuperscript{low} naive B cells (128 mutations in 17,194 bp, 7.44 \times 10\textsuperscript{-3} versus 0 mutations in 15,726 bp, 6.36 \times 10\textsuperscript{-5}; Table I). The mutation frequency of the JH4 intronic region was \( \sim 30\)-fold higher than that of the Bcl6 first intron region (7.44 \times 10\textsuperscript{-3} versus 2.39 \times 10\textsuperscript{-4}; Table I). These data demonstrate that the
Sorted B220+PNAhigh GC or B220+PNAlow naive B cells from

frequency of the two hotspots in the inserted S region (303–422 bp), and 68 in

identified in naive B cells (B220+PNAlow), and three of them were

than A:T pairs (66 versus 20; Fig. 4A). Only four mutations were found in the inserted S region.

Increased mutation frequency in the knock-in Bcl6S1 allele

To test whether the inserted S1 region influences the mutation frequency of the Bcl6 locus, we purified genomic DNA from the sorted B220+PNAhigh GC or B220+PNAlow naive B cells from immunized Bcl6S1/S1 heterozygous or Bcl6S1/S1 homozygous mice. The Bcl6S1 allele was examined for mutations or deletions using primers A and C (Fig. 3A). A diagram showing the endogenous Bcl6 first intron sequence, the targeting vector, and the inserted S1 sequence with the position of the primers used for amplification is presented, with the wt version of the locus included for comparison (Fig. 3A, details in Supplemental Fig. 1A, 1B). We found that the mutation frequency of the Bcl6S1 allele was ~5-fold higher than that of the wt allele in GC B cells (B220+PNAhigh, 1.201 x 10^-3 in S1 allele versus 2.389 x 10^-4 in wt allele; Table I). Of the 93 mutations identified in the Bcl6S1 allele, 12 were found in the endogenous Bcl6 first intron sequence (1–302 bp), 13 in the targeting vector sequence (303–422 bp), and 68 in the inserted S1 region (423–602 bp; Fig. 3C; see later). We noted two hotspots in the inserted S1 region that were frequently targeted at positions 486 and 535, with both as C residues (Fig. 3C). These mutations were identified in unique subclones from independent donor mice and did not represent the repeated counts, because the clones carrying these mutations also had independent mutations at other locations in the sequences. Eighty-six of 93 mutations were substitutions targeting to both A:T and C:G pairs; however, there were many more mutations occurring at G:C pairs than A:T pairs (66 versus 20; Fig. 4A). Only four mutations were identified in naive B cells (B220+PNAlow), and three of them were found in the inserted S1 region, in particular, the C residue at position 486 (Table I). No mutation was detected in the Bcl6S1 allele of the genomic DNA isolated from kidney (Table I). These data demonstrate that the Bcl6S1 allele is accessible to AID targeting, and the S region integrated in a non-Ig location can be an efficient target of SHM in vivo.

Next, we compared the mutation frequency of the same endogenous Bcl6 first intron sequences in wt versus Bcl6S1 allele, and found that the mutation frequency was not significantly different between these two alleles (2.99 x 10^-4 in the wt allele versus 3.09 x 10^-4 in the Bcl6S1 allele; Fig. 4B, 1–302 bp in the analyzed sequences). However, within the knock-in Bcl6S1 allele, the mutation frequency of the inserted S1 region was ~10-fold higher than that of the adjacent endogenous Bcl6 first intron sequence (29.4 x 10^-4 versus 3.09 x 10^-4; Fig. 4B). These results demonstrate that the inserted S1 region is a much more efficient AID target than the Bcl6 first intron sequence. Thus, we conclude that S region sequence per se is sufficient to enhance AID targeting efficiency to a non-Ig locus, and it does not require Ig cis regulatory elements such as the VH promoter and IEM enhancer.

The endogenous Sµ region undergoes internal deletions frequently in normal B cells and B cell lines (40–44). Next, we investigated whether the S1 region inserted into Bcl6 locus can undergo internal deletions or other types of genomic rearrangements such as translocations. Using a PCR approach, we identified 4 independent deletional events out of 133 sequences from PNAhigh GC B cells, with the lengths of deletion being 47, 60, 78, and 50 bp, and no such deletion was detected in kidney DNA, thus arguing against potential PCR artifacts (Fig. 4C and data not shown). These results suggest the occurrence of DNA double-strand breaks in the inserted S1 region. However, we did not detect any large deletions or genomic rearrangements in the knock-in Bcl6S1 allele by Southern blot analysis. A total of 270 hybridomas were generated from GC B cells of the immunized Bcl6S1/S1 and Bcl6S1/S1 mice (Table II), and hybridoma genomic DNA was analyzed by Southern blotting. The size of the targeted Bcl6S1 allele was not altered significantly in any of the samples analyzed (data not shown). Thus, our data suggest that the inserted S1 region undergoes infrequent microdeletions, but not large deletions or genomic rearrangements.

In addition, we designed primers D and E to amplify the endogenous S1 region in the Igh locus (Fig. 5A). The total length of analyzed sequence is 1229 bp, and the mutation frequency of this region is 2.107 x 10^-4 (Table III). The distal 166-bp sequence (1064–1229 bp) is exactly the same as the inserted S1 region we sequenced for the Bcl6S1 allele, and the mutation frequency of this region is 5.917 x 10^-4 in the endogenous Igh locus. This level of mutation is lower than what we observed for the same region in the Bcl6S1 allele (29.4 x 10^-4; Fig. 4B). The mutation spectra of the S1 sequence are quite similar in both loci, with many more mutations occurring at G:C pairs than A:T pairs (Figs. 4B, 5B). Given that these two sequences are identical but located at two different loci, our results suggest that locus-specific cis regulatory elements might contribute to SHM targeting to a non-Ig locus.

<table>
<thead>
<tr>
<th>Row Number</th>
<th>Mice</th>
<th>Locus</th>
<th>Cell Types</th>
<th>Total Mutations</th>
<th>Total Length of Sequence (bp)</th>
<th>Mutation Frequency</th>
<th>Independent Mutated Clones</th>
<th>Donor Mice</th>
<th>p Value*</th>
<th>Fold Change*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bcl6S1/+</td>
<td>Wt allele</td>
<td>B220+PNAhigh</td>
<td>9</td>
<td>37,666</td>
<td>2.389 x 10^-4</td>
<td>8/81</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Bcl6S1/+</td>
<td>Wt allele</td>
<td>B220+PNAlow</td>
<td>0</td>
<td>28,300</td>
<td>&lt;3.534 x 10^-5</td>
<td>0/50</td>
<td>2</td>
<td>p &lt; 0.05 (r1, 2)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Bcl6S1/+</td>
<td>Wt allele</td>
<td>Kidney</td>
<td>1</td>
<td>27,336</td>
<td>3.658 x 10^-5</td>
<td>1/66</td>
<td>4</td>
<td>p &lt; 0.05 (r1, 3)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Bcl6S1/+</td>
<td>S1 allele</td>
<td>B220+PNAhigh</td>
<td>93</td>
<td>77,445</td>
<td>1.201 x 10^-3</td>
<td>43/133</td>
<td>5</td>
<td>p &lt; 0.0001 (r1, 4)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Bcl6S1/+</td>
<td>S1 allele</td>
<td>B220+PNAlow</td>
<td>4</td>
<td>44,432</td>
<td>9.003 x 10^-5</td>
<td>4/78</td>
<td>3</td>
<td>p &lt; 0.0001 (r4, 5)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Bcl6S1/+</td>
<td>S1 allele</td>
<td>Kidney</td>
<td>0</td>
<td>53,353</td>
<td>&lt;1.874 x 10^-5</td>
<td>0/53</td>
<td>3</td>
<td>p &lt; 0.0001 (r4, 6)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Bcl6S1/AID−/−</td>
<td>S1 allele</td>
<td>B220+PNAhigh</td>
<td>1</td>
<td>29,498</td>
<td>3.390 x 10^-5</td>
<td>1/49</td>
<td>4</td>
<td>p &lt; 0.0001 (r4, 7)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Bcl6S1/AID−/−</td>
<td>S1 allele</td>
<td>B220+PNAhigh</td>
<td>0</td>
<td>31,304</td>
<td>&lt;3.194 x 10^-5</td>
<td>0/52</td>
<td>4</td>
<td>p &lt; 0.0001 (r4, 8)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Bcl6S1/+</td>
<td>J4 region</td>
<td>B220+PNAhigh</td>
<td>128</td>
<td>17,194</td>
<td>7.444 x 10^-3</td>
<td>15/17</td>
<td>3</td>
<td>p &lt; 0.0001 (r1, 9)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Bcl6S1/+</td>
<td>J4 region</td>
<td>B220+PNAhigh</td>
<td>0</td>
<td>15,726</td>
<td>&lt;6.359 x 10^-5</td>
<td>0/15</td>
<td>3</td>
<td>p &lt; 0.0001 (r9, 10)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>AID−/−</td>
<td>Wt Bcl-6</td>
<td>B220+PNAhigh</td>
<td>0</td>
<td>29,498</td>
<td>&lt;3.341 x 10^-5</td>
<td>0/64</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>AID−/−</td>
<td>Wt Bcl-6</td>
<td>Kidney</td>
<td>0</td>
<td>23,012</td>
<td>&lt;3.436 x 10^-5</td>
<td>0/51</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistics were calculated by Fisher’s exact test.

*aThe p values are shown for mutation frequencies between two rows (r) as denoted in the parentheses.

*bFold change is calculated between mutation frequencies of two rows (r) as denoted in the parentheses.

Bcl6 first intron region is a much less efficient target for SHM compared with the JH intronic region downstream of the assembled V region exon.

The Journal of Immunology
activity of AID, we bred the knock-in Bcl6^{Sy1} allele into AID-deficient background to generate double-mutant Bcl6^{Sy1/AID−/−} mice. GC and naive B cells were sorted from the immunized Bcl6^{Sy1/AID−/−} mice, and genomic DNA was isolated from both populations and used for mutational analysis as described earlier. We found no mutations in the knock-in Bcl6^{Sy1} allele of the naive B cells from the double-mutant Bcl6^{Sy1/AID−/−} mice and an extremely low level of mutation in the GC B cells (Table I). Therefore, our results unequivocally demonstrate that the S region mediated SHM in Bc16 locus is also dependent on the function of AID.

**Increased RNAPII and AID recruitment in the inserted Sy1 region**

To investigate the mechanism of S region-enhanced SHM, we hypothesized that the inserted Sy1 region may induce an increased accumulation of the RNAPII molecule, which, in turn, recruits more AID to this unique and specific sequence to cause mutations. To test our hypothesis, we isolated GC B cells from immunized mice and used for mutational analysis as described earlier. We developed a novel experimental approach that allowed us to determine the density of RNAPII on both alleles in the same sample, thereby eliminating the effects of differential IP efficiency observed between multiple samples. Anti-RNAPII was used to enrich the bound chromatin, and qPCR was performed using multiple sets of primers specific for the wt or Bcl6^{Sy1} alleles. wt allele-specific primers are in the targeting vector, and 5′ end of the inserted Sy1 region is marked along the sequence derived from 303–602 bp of the Bc16 gene.

### Mutations in the Bc16^{Sy1} allele are dependent on AID activity

To definitively demonstrate that the mutations identified in the inserted Sy1 region of the knock-in Bc16^{Sy1} allele depend on the activity of AID, we conducted experiments using Bc16^{Sy1/AID−/−} mice. GC and naive B cells were sorted from the immunized Bc16^{Sy1/AID−/−} mice, and genomic DNA was isolated from both populations and used for mutational analysis as described earlier. We found no mutations in the knock-in Bc16^{Sy1} allele of the naive B cells from the double-mutant Bc16^{Sy1/AID−/−} mice and an extremely low level of mutation in the GC B cells (Table I). Therefore, our results unequivocally demonstrate that the S region mediated SHM in Bc16 locus is also dependent on the function of AID.

### FIGURE 5. Mutation spectrum of the endogenous Sy1 region in the Igh locus.

(A) Schematic of PCR strategy for the endogenous Sy1 region in the Igh locus. Sy1, Cy1 exons, and the positions of the primers D and E are indicated. (B) Pattern of the nucleotide substitutions in the analyzed endogenous Sy1 region. Patterns were deduced by counting all the nucleotide substitutions identified in the unique sequences (n = 4 donor mice).

### Table II. No large rearrangements detected in the inserted Sy1 allele

<table>
<thead>
<tr>
<th>Fusion</th>
<th>Mice</th>
<th>Hybrids/loci</th>
<th>Rearrangement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bc16^{Sy1/−}</td>
<td>64</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Bc16^{Sy1/−}</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Bc16^{Sy1/−}</td>
<td>91</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Bc16^{Sy1/Sy1}</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Bc16^{Sy1/Sy1}</td>
<td>57</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Bc16^{Sy1/Sy1}</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>Bc16^{Sy1}</td>
<td>270</td>
<td>0</td>
</tr>
</tbody>
</table>
GC B cells. Notably, the DNA of the inserted Sy1 region was dramatically enriched in the anti-RNAPII IP compared with the control-Ig IP sample (Fig. 6B). After additional normalization, we found that the level of RNAPII occupancy in the inserted Sy1 region was increased 5.6-fold compared with the wt Bc16 first intron region (Supplemental Fig. 3B). Consistent results were obtained with multiple sets of allele-specific primers (data not shown). Thus, we conclude that the inserted Sy1 sequence per se, independent of IgH cis regulatory elements, increases RNAPII recruitment to a non-Ig locus in primary GC B cells during SHM.

Next, to directly assess the recruitment of AID to the inserted Sy1 region, we performed an AID ChIP-qPCR assay. Consistent with the RNAPII enrichment, the DNA of the inserted Sy1 region was dramatically enriched in the anti-AID IP compared with the control-Ig IP sample (Fig. 6C). Further normalized ChIP-qPCR data showed that the level of AID accumulation in the inserted Sy1 region was increased 7.8-fold compared with the wt Bc16 first intron region (Supplemental Fig. 3C). We confirmed our qPCR data via semiquantitative PCR and gel electrophoresis (Supplemental Fig. 3A). To further rule out the possibility that AID enrichment is an artifact mediated by the anti-AID Ab, we performed ChIP-qPCR and semiquantitative PCR analysis using GC B cells from immunized Bc16Sy1/AID−/− double-mutant mice. Our results showed that RNAPII but not AID was still enriched in the inserted Sy1 region in Bc16Sy1/AID−/− mice, demonstrating that its accumulation is independent of AID (Fig. 6B, 6C, Supplemental Fig. 3A–C). Thus, we conclude that the inserted Sy1 region can directly enhance the recruitment of both RNAPII and AID, which could serve as the underlying mechanism of S region-enhanced SHM.

To address whether Spt5 is involved in RNAPII pausing at the inserted Sy1 region, we performed ChIP-qPCR analysis to investigate its occupancy in this region. Indeed, we detected an increase of Spt5 occupancy in the inserted Sy1 region. Our semiquantitative RT-PCR analysis showed that the level of RNAPII was closely associated with the inserted Sy1 region (Fig. 6D). These data are also consistent with previous findings of critical role of Spt5 sequence in regulating RNAPII density in the endogenous IgH locus (28, 29).

### Discussion

During SHM, apart from the V region exons of IgH and IgL loci, AID targets a group of non-Ig genes but generally spares the rest of the genome, although the mutation frequency of these genes is much lower than that of V regions (2, 4, 5, 6, 8, 37, 45–47). Multiple mechanisms may operate to regulate the specificity and efficiency of AID targeting (48, 49). Whether and how the target DNA sequences influence mutation frequency has not been clearly
addressed. Studies using artificial substrates driven by Ig-specific promoter and enhancer suggest that nucleotide sequence might influence mutation frequency (50–52); however, several tested non-Ig genes mutated at a similar frequency to the V region exon, implicating an unimportant role of the sequences (53). Nevertheless, these experiments were inconclusive due to technical limitations of the transgenic approach (54). Thus, the role of nucleotide sequence in SHM remains unresolved. In addition, these studies were performed before the discovery of AID, and it remains to be determined whether target DNA sequences regulate the efficiency of AID targeting in the context of SHM.

To address these fundamental questions, we established a model system via the gene-targeting approach to introduce a core S<sub>γ</sub>1 region into the first intron of <i>Bcl6</i>. Consistent with previous analysis of a similar region (7), we found that the mouse <i>Bcl6</i> first intron region mutated at a frequency of 2.389 ± 2.4 (Table I). In

FIGURE 6. Enhanced recruitment of RNAPII, AID, and Spt5 to the inserted Sγ1 region of the Bcl6<sup>Sγ1</sup> allele. (A) Schematics of wt and Bcl6<sup>Sγ1</sup> knock-in loci of the Bcl6 gene. The positions of allele-specific primers for ChIP-qPCR analysis are indicated (see details in Supplemental Fig. 2A). ChIP analysis for RNAPII (B), AID (C), and Spt5 (D) occupancy in the wt Bcl6 first intron region versus the inserted Sγ1 region. GC B cells were isolated from Bcl6<sup>Sγ1/+</sup> heterozygous mice [left panels in (B) and (C)] or Bcl6<sup>Sγ1/AID<sup>−/−</sup> double-mutant mice [right panels in (B)and (C); n = 4–8 per experiment]. Input and IP samples were analyzed by qPCR for the wt Bcl6 first intron, inserted Sγ1, GAPDH, and CD8 promoter regions. Three independent experiments were performed, and representative data from one experiment are shown. ChIP-qPCR data are presented as percentage input (see details in Materials and Methods). SD was calculated from qPCRs performed in triplicate. Statistical analyses were calculated by a Student t test with two-tailed distribution and equal variance.
contrast, the mutation frequency of the inserted Sy1 region was ∼10-fold higher than that of the adjacent endogenous Bcl6 sequence (Fig. 4B). Thus, we conclude that S region sequence per se, independent of Igh cis regulatory elements, enhances AID targeting efficiency; more importantly, nucleotide sequence of a gene locus itself has a direct impact on mutability. We chose an S region as the first test sequence because, among the sequences analyzed so far, S regions appear to be preferential AID targets (18, 19, 25). AID’s preference for S regions is likely attributable to the unique aspects of their highly GC-rich repetitive sequences. Our data demonstrate that possessing such specific sequence features is sufficient to enhance the mutability of a target DNA sequence. Given that certain motifs predominate in S regions (26), future studies to modify these motifs via mutagenesis will further elucidate the mechanism of S region-mediated mutability.

We analyzed the Sy1 sequence in the endogenous Igh locus and found that it mutated at a lower frequency compared with the same sequence in the Bcl6 locus (Fig. 4B, Table III). We reason that this differential effect is likely caused by the locus-specific cis regulatory elements, although we cannot exclude the influence of other factors, such as the distance to promoter or the absence of the analyzed Sy1 sequence in switched B cells because of its deletion during CSR. The effects of cis-acting elements of Ig loci in SHM have been investigated extensively (2, 55–58). However, none of the known transcriptional control elements in mouse B cells appears to be preferential targets of AID such as Bcl6.

Our ChIP-qPCR data indicate that the enhanced recruitment of RNAPII and AID might provide the mechanistic basis for the observed mutational phenotypes of the inserted S region (Fig. 6). Moreover, we show that the higher level of RNAPII accumulation is closely correlated with the inserted Sy1 region and detected at both the 5′ and 3′ ends of this region, supporting a higher density of RNAPII across the entire knock-in S region. These data appear to be highly consistent with the previous findings with the endogenous S region in the Igh locus (29). Why does RNAPII accumulate at the S region? RNAPII pausing is often observed at highly regulated genes (62), which have higher amounts of RNAPII at the promoter versus gene body. The meaning of promoter-proximal RNAPII pausing is not completely understood; the current consensus view is that it regulates gene activation via transcription (63). Recent studies suggest an interplay between promoter DNA sequences, nucleosome occupancy, and RNAPII pausing (64). However, S regions are often located several kilo-
bases downstream of promoter regions, and the length of individual S regions varies from 2 to 10 kb (e.g., Sy1). Thus, the meaning of RNAPII accumulation across S regions may be distinct from its promoter-proximal pausing. Previous studies showed that, in the absence of transcription, nucleosomes prevented AID-mediated cytidine deamination (65), whereas with transcription, AID readily gained access to DNA in nucleosomes on both strands (65). Therefore, we propose that RNAPII pausing at S regions probably facilitates the repositioning of repressive nucleosomes to create a permissive chromatin architecture that allows AID to access target DNA sequences. Our AID ChIP data seem to be consistent with this notion, because the higher level of AID accumulation peaks at the inserted Sy1 region, with an even more focused pattern compared with RNAPII (Fig. 7). Several possible mechanisms have been proposed that might induce RNAPII accumulation at S regions, including R-loop structures, G-quartet structures, and antisense RNA expression in the S region (29). It remains to be determined which mechanism(s) leads to increased RNAPII accumulation, which may require additional studies focused on disrupting unusual aspects of the S region sequence via mutagenesis. Our hypothesis is also consistent with the involvement of histone chaperone facilitates chromatin transcription complex (66) and distinct chromatin modifications in CSR (29, 66–68). Overall, these studies reveal a complex picture of chromatin modification patterns in the Igh locus during CSR, with active or repressive histone marks associated with S or C regions (29, 66–68). Thus, it is likely that a specific combination of histone modifications may be responsible for the overall mutational phenotypes we observed.

Our ChIP-qPCR data also show increased recruitment of Spt5 in the inserted Sy1 region in primary GC B cells, suggesting that Spt5 may contribute to the recruitment of RNAPII and AID during SHM in vivo. Spt5 has been shown to interact with AID (34), and its genome localization pattern is highly correlated to that of RNAPII in the in vitro cytokine-activated B cells (34). Apart from associating with RNAPII paused at the promotor proximal region, Spt5 links RNAPII to other factors such as splicing factors, capping enzyme, and transcription coupled repair factors (69–73). Thus, it remains to be determined how Spt5 contributes to RNAPII pausing at S regions. We also notice that the level of Spt5 enrichment is not as high as that of AID and RNAPII (Fig. 6). Therefore, we propose that robust mutation probably requires additional AID cofactors, including factors recruiting RNAPII or mediating RNAPII/AID interaction that are yet to be identified in primary GC B cells.

Comprehensive ChIP-Seq analysis shows that AID is recruited to 5910 target genes in cytokine-activated AID+/+ B cells (74). However, the mutation frequency of non-Ig genes in cytokine-activated wt B cells is extremely low; in fact, the V region exon of the Igh locus does not even undergo SHM in this setting (75, and data not shown). Although AID recruitment pattern remains to be investigated in GC B cells, analysis of hundreds of non-Ig genes shows that their mutation frequency is much lower than that of Ig V region exons (8). Therefore, although AID has the potential to access the genome widely, its mutagenic activity appears to be highly restricted to a few loci. The choice of downstream DNA repair mechanism (error-free versus error-prone) contributes to the tight regulation of AID’s mutagenic activity (8). In addition to this mechanism, we propose that the nucleotide sequence preference may serve as an additional layer of AID regulation by restricting its mutagenic activity to specific sequences. We show that the inserted Sy1 sequence manifests a higher mutability and enhances AID recruitment but has rather minimal effects on the mutability of adjacent endogenous Bcl6 sequence (Fig. 4); thus, these data demonstrate that AID’s mutagenic activity depends on its target sequence. This regulatory mechanism might explain the phenomenon that although AID is recruited to 5910 target genes (74), it would hardly induce mutations in most of them because of the absence of preferential target sequences of AID. Consistently, in the context of CSR, the occupancy of Spt5 and RNAPII in cytokine-activated B cells extends throughout the Cμ region within the Igh locus (34); similarly, the occupancy of another AID cofactor, RPA, also extends throughout the Cμ and Cγ regions (74), whereas neither of these regions is targets of AID. Thus, we propose that these AID cofactors might function as “accessibility factors” to create a permissive chromatin architecture for AID targeting; subsequently, the outcome of AID’s mutagenic activity is further regulated by local control mechanisms including target DNA sequences, DNA repair pathways, posttranslational regulation of AID, or other AID-targeting factors.

Acknowledgments

We thank Drs. F.W. Alt (Children’s Hospital Boston, Howard Hughes Medical Institute) for generous support of the project, A.A. Zarrin for help with generating the Bcl6α knock-in mice, and J. Manis for helpful discussions. We are grateful to Y. Fujiwara, T. Borjeson, and A. Williams for mouse work and to the flow cytometry core facility at National Jewish Health.

Disclosures

The authors have no financial conflicts of interest.

References


