A Role for Msh6 But Not Msh3 in Somatic Hypermutation and Class Switch Recombination

Stella A. Martomo, William W. Yang, and Patricia J. Gearhart

Laboratory of Molecular Gerontology, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224

Abstract

Somatic hypermutation is initiated by activation–induced cytidine deaminase (AID), and occurs in several kilobases of DNA around rearranged immunoglobulin variable (V) genes and switch (S) sites before constant genes. AID deaminates cytosine to uracil, which can produce mutations of C:G nucleotide pairs, and the mismatch repair protein Msh2 participates in generating substitutions of downstream A:T pairs. Msh2 is always found as a heterodimer with either Msh3 or Msh6, so it is important to know which one is involved. Therefore, we sequenced V and S regions from Msh3- and Msh6-deficient mice and compared mutations to those from wild-type mice. Msh6-deficient mice had fewer substitutions of A and T bases in both regions and reduced heavy chain class switching, whereas Msh3-deficient mice had normal antibody responses. This establishes a role for the Msh2-Msh6 heterodimer in hypermutation and switch recombination. When the positions of mutation were mapped, several focused peaks were found in $Msh6^{-/-}$ clones, whereas mutations were dispersed in $Msh3^{-/-}$ and wild-type clones. The peaks occurred at either G or C in WGCW motifs (W = A or T), indicating that C was mutated on both DNA strands. This suggests that AID has limited entry points into V and S regions in vivo, and subsequent mutation requires Msh2-Msh6 and DNA polymerase.

Key words: DNA repair enzymes • cytosine deamination • class switching • mutation hotspots • B lymphocytes

Introduction

Activation-induced cytidine deaminase (AID) causes hypermutation in both immunoglobulin variable (V) and switch (S) regions and recombination in S regions (1, 2). AID deaminates cytosine to uracil on a variety of DNA substrates in vitro (3–7), and overexpression of AID causes mutations of C:G base pairs in vivo (8–13). However, in B cells, substitutions of all four bases occur at similar levels in V and S regions, indicating that other proteins are required to generate mutations at A:T base pairs. The low fidelity DNA polymerase η participates because humans deficient for the enzyme have low frequencies of A:T mutations and high levels of C:G mutations in V and S regions (14–16). The mismatch repair protein Msh2 is also engaged because Msh2-deficient mice have few A:T mutations in V and S regions (17–20). Msh2 is always associated as a heterodimer in vivo with either Msh3 or Msh6, and the two dimers have distinct and overlapping functions in mismatch repair (21). Therefore, it is important to know if Msh3 and/or Msh6 are involved in hypermutation. A paper by Wiesendanger et al. (22) implicates Msh6, but not Msh3, in V gene mutations; however, mutations in the S region and heavy chain class switching to IgG have not been examined in these mice.

How AID is directed to V and S regions is a question of intense interest. Two lines of evidence support targeting by cis DNA sequences in the locus. First, transcription complexes may bring AID to the V and S regions because mutations are found downstream of the promoters preceding V genes and S regions, and require components of the intron and 3′ enhancers (for review see reference 23). Recent studies show that histone acetylation of chromatin in the V (12) and S (24) regions precedes hypermutation and switching. Thus, DNA would be transcribed and become accessible to AID activity. Second, the DNA motifs RGYW/WRCY (R = purine, Y = pyrimidine, W = A or T) are hotspots because mutations are frequently found there in immunoglobulin genes (25). In support of hotspots, AID was shown recently to preferentially deaminate C in WRC motifs on

single-stranded substrates in vitro (26, 27). However, in B cells, mutations are frequently found at other positions in V and S regions, suggesting that other proteins either repair the resulting mismatches and/or generate mutations past the original C lesion.

To address the question of which DNA sites are recognized by AID in vivo, we looked at hypermutation in the absence of mismatch repair proteins, which may reveal sites of initial targeting. Because switching has not been examined in mice deficient for Msh3 or Msh6, first we measured IgG production by splenic B cells stimulated in culture. Mutations in rearranged J_H4 introns and $S\mu$ regions from Peyer's patch B cells were identified by sequencing. Sequences containing peaks of mutation were analyzed to determine why these sites are targeted.

Materials and Methods

Mice. $Msh3^{-/-}$ and $Msh6^{-/-}$ mice on a C57BL/6 background were provided by W. Edelmann (Albert Einstein College of Medicine, Bronx, NY), and C57BL/6 mice were purchased from Jackson ImmunoResearch Laboratories. Mice were used at 3–7 mo of age. For immunization, some mice were injected intraperitoneally with 100 μg KLH (Calbiochem) in adjuvant (RIBI Immunochem Research), boosted after 3 mo with 50 μg, and bled 4 d later.

Isotype Analysis. For in vitro switching, spleen cells were treated with ACK lysing buffer (Quality Biological) to lyse red blood cells. The cells were mixed with MicroBeads conjugated to rat monoclonal anti–mouse/human CD11b and anti–mouse CD43 (Ly-48, Leukosialin; Miltenyi Biotec) to remove macrophages, T cells, and activated B cells. The mixture was incubated at 4°C for 20 min and added to a magnetic cell separating column (MACS CS; Miltenyi Biotec). The unbound population, containing resting B cells, was cultured at 10^6 cells/ml in RPMI 1640 supplemented with 10% heat-inactivated FBS, antibiotics, L-glutamine, and 5×10^{-5} M β-mercaptoethanol. To induce switching, cells were stimulated with 50 μg/ml *Escherichia coli* LPS (Sigma-Aldrich) or LPS and 50 ng/ml recombinant mouse IL-4 (R&D Systems). The cells were harvested after 3 d, washed twice with RPMI 1640, and stained with phycoerythrin-conjugated rat monoclonal anti–mouse CD45R (B220; BD Biosciences) and either FITC-conjugated rat monoclonal anti–mouse IgG3 or anti–mouse IgG1 (BD Biosciences) in RPMI 1640 supplemented with 1% heat-inactivated serum. Flow cytometry analysis was gated on live cells as determined from forward and side scatter analysis.

For serum isotypes, an ELISA was performed to measure anti-KLH antibodies. Wells were coated with 100 ng KLH in carbonate buffer containing 0.016 M Na_2CO_3, 0.034 M $NaHCO_3$, pH 9.8. Wells were blocked with 1% ELISA grade BSA and 0.1% Tween 20 in 1× PBS (blocking buffer). A 1:100 dilution of serum from immunized mice was used to bind to the antigen, and after washing with 0.1% Tween 20 in PBS (washing buffer), a 1:10 dilution of rabbit anti–mouse IgM, 1:5 of IgG1, or 1:2 of IgG3 antibody (Bio-Rad Laboratories) in blocking buffer was added to the wells. After incubation, the wells were washed twice with washing buffer, and donkey anti–rabbit Ig horseradish peroxidase–linked whole antibody (Amersham Biosciences) diluted 1:5,000 in blocking buffer was added. The wells were washed, and bound antibody was detected with *o*-phenylenediamine dihydrochoride (Sigma-Aldrich) followed by termination with 3M

H_2SO_4. All incubations were done at 37°C for 30 min. The absorbance was read at 490 nm. Values were normalized to total IgM, IgG1, and IgG3 in the serum as determined by ELISA using purified immunoglobulins (BD Biosciences).

Libraries of V and S Regions for Sequencing. Cells from the Peyer's patches of two to three mice of each genotype were stained with phycoerythrin-labeled antibody to B220 and fluorescein-labeled peanut agglutinin (PNA; E-Y Laboratories). The cells were separated by flow cytometry, and DNA was prepared from $B220^+PNA^+$ cells. For V regions, the intron region downstream of rearranged V, diversity (D), and joining (J) gene segments on the heavy chain locus was sequenced. DNA was amplified using nested 5′ primers for the third framework region of V_HJ558 gene segments and 3′ primers for 344 nucleotides downstream of the J_H4 gene segment as described previously (28). For $S\mu$ regions, a 561-base region located upstream of the core μ S region was sequenced (29). The following sets of nested switch
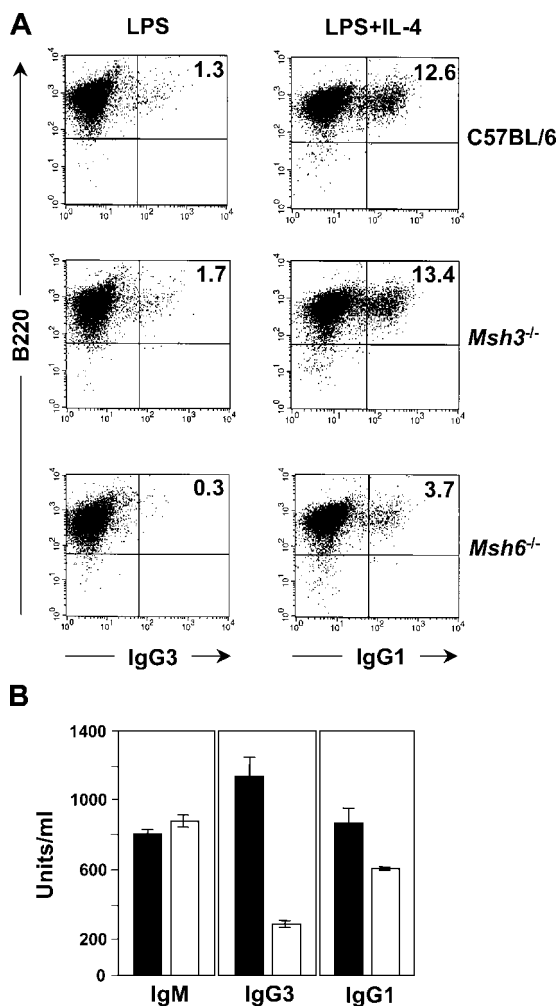

Figure 1. Diminished heavy chain class switching in $Msh6^{-/-}$ but not $Msh3^{-/-}$ B cells. (A) Flow cytometry analysis. Spleen cells were stimulated with LPS or LPS plus IL-4 to induce switching, and surface immunoglobulin was measured 3 d later. The percentage of $B220^+$ cells that switched to IgG3 or IgG1 is shown in each box. (B) Serum antibody analysis. An ELISA was used to quantify anti-KLH antibodies from immunized mice. Black bars, C57BL/6; white bars, $Msh6^{-/-}$. Units are defined as the antibody titer that was normalized to the total amount of each isotype.

primers were used: first set, forward (nucleotides 4,560–4,579 of Genbank/EMBL/DDBJ under accession no. J00440), 5′-AGA-TAAAATGGATACCTCAG-3′; reverse (nucleotides 5,183–5,202), 5′-TAGTTTAGCTTAGCGGCCCA-3′; and second set, forward (nucleotides 4,580–4,599) with XbaI addition in italics, 5′-AC*TCTAGA*TGGTTTTTAATGGTGGGTTT-3′; reverse (nucleotides 5,152–5,181) with EcoRI addition in italics, 5′-AC*GAATTC*CTCATTCCAGTTCATTACAG-3′. 20 ng of genomic DNA was amplified with Platinum Pfx polymerase and PCR enhancer (Invitrogen) in a 50-μl volume using the first set of primers for 30 cycles of 95°C for 30 s, 55°C for 30 s, and 68°C for 1 min, followed by a final incubation at 68°C for 10 min. Nested PCR was performed with 5 μl of the first reaction and the second set of primers with an annealing temperature of 50°C for 30 s for another 30 cycles. Products were digested, cloned into pBluescript (Stratagene), and sequenced.

Online Supplemental Material. Figs. S1–S3 present the sequences of J_H4 introns from C57BL/6, $Msh3^{-/-}$, and $Msh6^{-/-}$ clones, and Figs. S4–S6 show the sequences of Sμ regions from the three sets of clones. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20040691/DC1.

Results

$Msh6^{-/-}$ B Cells Had Diminished IgG Switching. To see if the Msh3 or Msh6 protein affected heavy chain class switching in vitro, the ability of B cells to switch isotypes in cell culture was measured by flow cytometry. Spleen cells were stimulated with LPS to induce IgG3 switching and with LPS plus IL-4 to induce IgG1 switching. After 3 d, B cells from C57BL/6, $Msh3^{-/-}$, and $Msh6^{-/-}$ mice had equal percentages of 77% live cells as assessed by scatter analysis. Cells were stained to identify IgG isotypes, and flow cytometry analysis was gated on the live cells. As shown in Fig. 1 A, compared with C57BL/6 cells, Msh3-deficient cells had normal levels of switching to IgG3 and IgG1, but Msh6-deficient cells had reduced levels of switching to both isotypes (P < 10^{-8}, Fisher's exact test). To confirm the defect in switching, serum antibody was measured by ELISA from immunized C57BL/6 and Msh6-deficient mice. The level of IgM anti-KLH antibody was equal from both groups, but IgG3 and IgG1 antibodies were reduced in Msh6-deficient serum (Fig. 1 B, P < 10^{-6}). These results demonstrate a role for Msh6 in recombination at S regions.

$Msh6$-deficient Mice Had Fewer A and T Mutations in J_H4 Introns and Sμ Regions. For V regions, we sequenced a 344-bp intron region downstream of J_H4 gene segments rearranged to V_HJ558 gene segments in B220+PNA+ B cells from Peyer's patches. The frequencies of mutation were similar from both unimmunized and immunized mice, so the data were combined. The following number of clones with unique VDJ junctions and mutations were obtained: C57BL/6 mice, 60% of 42 clones were mutated at a frequency of 6.3 × 10^{-3} mutations/bp with 91 substitutions and one deletion (1 bp); $Msh3^{-/-}$ mice, 16% of 134 clones were mutated at 1.6 × 10^{-3} mutations/bp with 73 substitutions and one deletion (1 bp); and $Msh6^{-/-}$ mice, 14% of 154 clones were mutated at 0.8 × 10^{-3} mutations/bp with
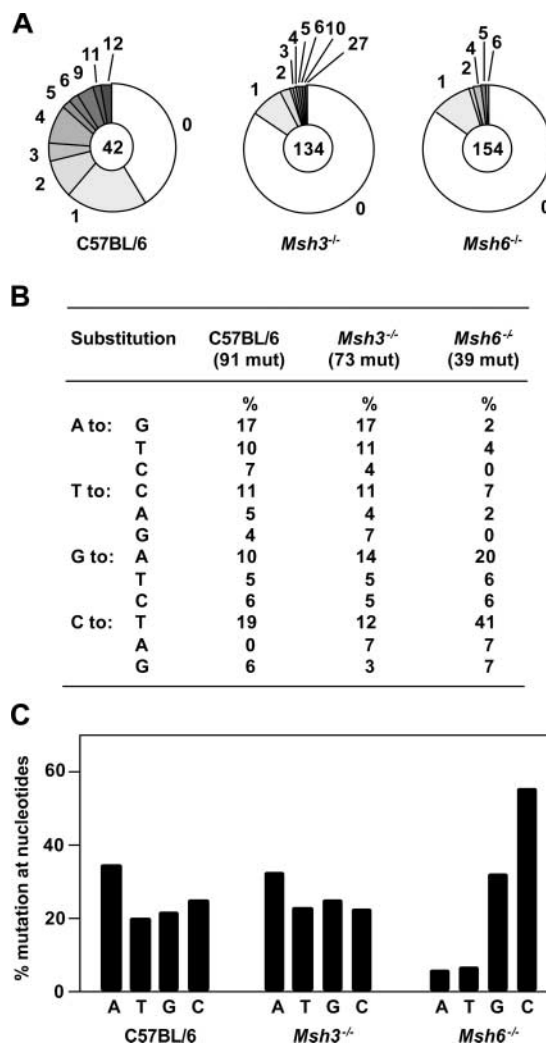

Figure 2. Fewer A:T substitutions in J_H4 introns from $Msh6^{-/-}$ clones. (A) The total number of clones analyzed is shown in the center of each circle. Segments represent the proportion of clones that contain the indicated number of mutations. (B) The sequence contains 26% A, 32% T, 28% G, and 14% C; values were corrected to represent a sequence with equal amounts of the four nucleotides. Mutations are shown from the nontranscribed strand. (C) Total mutations for each nucleotide are grouped.

39 substitutions and one deletion (15 bp; Fig. 2 A). The overall frequencies of mutation may be different because they reflect arbitrary exposure of Peyer's patch B cells to environmental antigens in the gut. However, among the mutated sequences, the $Msh6^{-/-}$ clones had a significantly lower frequency of mutation (5 × 10^{-3} mutations/bp, P < 0.001) compared with C57BL/6 and $Msh3^{-/-}$ clones (1 × 10^{-2} mutations/bp for both). In $Msh6^{-/-}$ clones, there were strikingly fewer mutations of A and T nucleotides (P < 10^{-3}), and correspondingly more mutations of G and C bases than in C57BL/6 and $Msh3^{-/-}$ clones (Fig. 2, B and C). Most of the G:C mutations were transitions. These data support a role for the Msh2-Msh6 heterodimer in causing A:T substitutions in the V region (22).

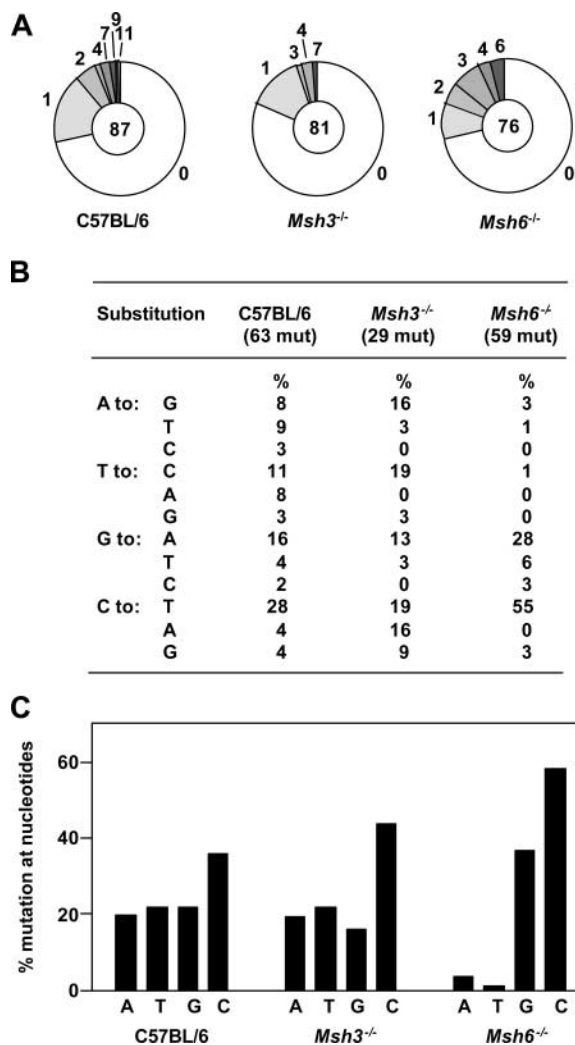For S regions, we focused on mutations in Sμ because Msh6-deficient B cells had impaired switching to IgG. A

Figure 3. Reduced A:T mutations in Sμ from *Msh6*[−/−] clones. (A) The proportion of clones containing mutations is shown. (B) Values are corrected for nucleotide composition of the sequence, which contained 33% A, 23% T, 28% G, and 16% C. (C) Total mutations of each nucleotide are grouped.

561-bp region located upstream of the core S region has been reported to accumulate mutations in B cells stimulated in culture (29, 30). To obtain a distribution profile that more closely resembles mutations occurring in vivo, we sequenced this region in PNA⁺ B cells from Peyer's patches. The following number of clones and mutations were identified: C57BL/6 mice, 31% of 87 clones were mutated at a frequency of 1.3×10^{-3} mutations/bp with 63 substitutions and four deletions (5, 15, >100, and >100 bp); *Msh3*[−/−] mice, 19% of 81 clones were mutated at 0.6×10^{-3} mutations/bp with 29 substitutions and one deletion (>100 bp); and *Msh6*[−/−] mice, 29% of 76 clones were mutated at 1.4×10^{-3} mutations/bp with 59 substitutions and two deletions (1 and 27 bp; Fig. 3 A). Among the mutated sequences, there was no significant difference between the frequencies of mutation in all three genotypes ($3–5 \times 10^{-3}$ mutations/bp). The large deletions may be signatures of internal recombination in the Sμ region (29).

Only clones with unique mutations were analyzed for the types of substitutions. As shown in Fig. 3 (B and C), all three types of mice had more mutations of C than G on the nontranscribed strand. The *Msh6*[−/−] clones had a dramatic decrease in the frequency of A and T mutations ($P < 10^{-3}$) compared with G and C mutations, with G:C transitions being the major category of mutations. Differences between the clones in individual categories, such as C to A, were not statistically significant. Thus, Msh6 participates in hypermutation of A:T nucleotides in the Sμ region.

Hot Spot Focusing of Mutations in V and S Regions from Msh6-deficient Mice. Mutations were plotted along V and S sequences to determine if certain positions were targeted. In J$_H$4 introns, three peaks were observed in Msh6-deficient clones, but not in Msh3-deficient and C57BL/6 clones (Fig. 4). Approximately 50% of the *Msh6*[−/−] mutations were in these peaks, compared with only 8% of the mutations from each of the *Msh3*[−/−] and C57BL/6 groups. The positions are nucleotides 57–58, 62–63, and 253–254 (clones shown in Figs. S1–S3, available at http://www.jem.org/cgi/content/full/jem.20040691/DC1) and consist of adjacent G and C bases in WGCW motifs as recorded from the nontranscribed strand. In Sμ, five major peaks were seen in *Msh6*[−/−] clones that contained 50% of the total mutations, whereas these positions had only 16% mutations in the *Msh3*[−/−] and C57BL/6 clones (Fig. 5). The positions are nucleotides 216–217, 282–283, 392–393, 452–453, and 462–463 (clones shown in Figs. S4–S6, available at http://www.jem.org/cgi/content/full/jem.20040691/DC1), which are G-C bases in WGCW motifs.
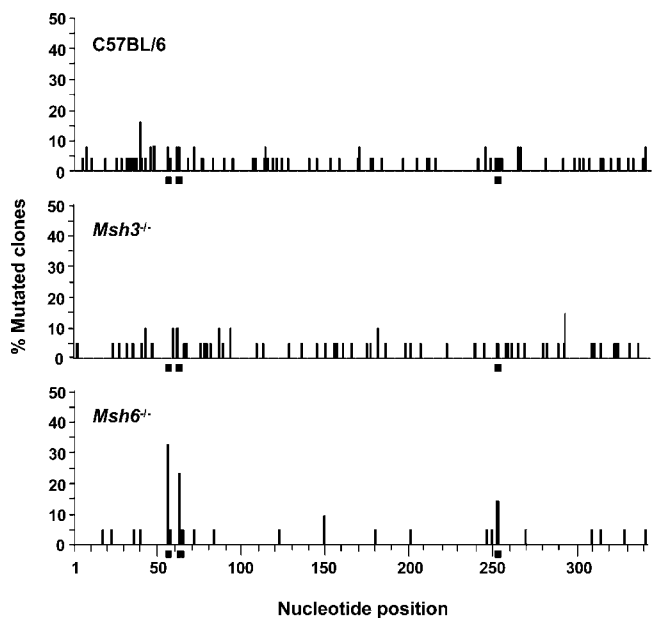

Figure 4. Distribution of mutations in J$_H$4 intron reveals peaks in *Msh6*[−/−] clones. Nucleotides are numbered from the first base after the J$_H$4 coding sequence, with 1 corresponding to 2,340 of Genbank/EMBL/DDBJ under accession no. J00440. Data were calculated as the number of mutations at a nucleotide position divided by the number of mutated clones. Bars below the abscissa depict WGCW motifs.
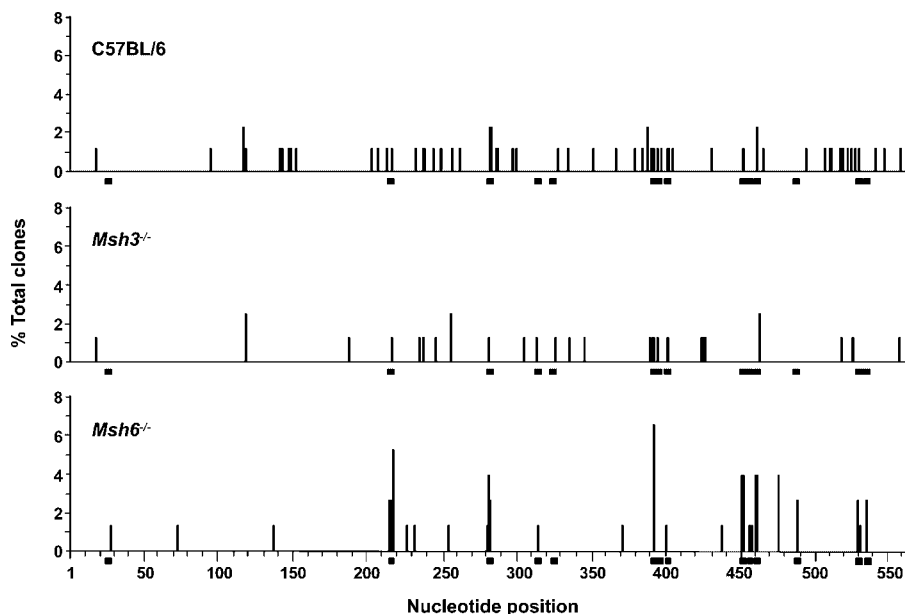
Figure 5. Location of mutations in Sμ shows targeting in $Msh6^{-/-}$ clones. Nucleotides are numbered with 1 corresponding to 4,600 of Genbank/EMBL/DDBJ under accession no. J00440. Data represent the number of mutations at a position divided by the total number of clones. Bars below the abscissa represent WGCW motifs.

The number of mutations of G and C in WGCW sequences from $Msh6^{-/-}$ clones is listed in Table I. Mutations in individual clones occurred at either G or C, but not in tandem. In the J_H4 intron region containing 344 bp, only three WGCW motifs were found, and they all had mutations at either the G or C positions. Other motifs such as RGCW or WGCY had far fewer mutations of G or C. In the Sμ region containing 561 bp, 14 WGCW sequences were found; 12 of them had substitutions in either the G or C positions. Fewer mutations were located in RGCW or WGCY motifs. Therefore, the WGCW subset of the canonical RGYW/WRCY motif was highly targeted, whereas other variants were avoided.

Discussion

Msh6, But Not Msh3, Affects Heavy Chain Class Switching. As reported for Msh2-deficient mice (31–33), Msh6-deficient mice had diminished switching to IgG1 and IgG3, compared with Msh3-deficient and wild-type cells. This indicates that the Msh2-Msh6 heterodimer, but not Msh2-Msh3, is involved in S recombination. Mismatch repair proteins could be involved in switching by two potential mechanisms. First, they could be involved in the recombination process itself by recruiting nucleases to process cleavage intermediates (34). Msh2-, Pms2-, and Mlh1-deficient mice all had reduced switching and different lengths of microhomology at the S junctions compared with wild-type mice (31, 32, 35, 36). In particular, $Msh2^{-/-}$ clones had short lengths of overlapping nucleotides, and Min et al. (37) postulated that the protein is needed for handling long staggered ends before recombination. Although S junctions in $Msh6^{-/-}$ mice were not analyzed in this work due to the low frequency of μ-γ switching, it is possible that the diminished switching is due to inefficient recombination.

Second, Msh6 may facilitate nicking at many positions to create more DNA breaks. Breaks would occur when cytosine on single-stranded DNA is deaminated to uracil, uracil is removed by uracil glycosylase, and the abasic site is cut by an abasic endonuclease (3–7, 13, 38, 39). The restricted focusing of mutations observed in Msh6-deficient clones suggests that Msh6 is needed to expose more cytosines to deamination. This may require an exonuclease to digest DNA from the nick, which would expose C on the opposite strand for AID to generate a double strand break. In support of this model, Bardwell et al. (40) have shown that mice deficient for exonuclease 1 have reduced heavy chain class switching. Because Msh2 binds to exonuclease 1 (41, 42), the Msh2-Msh6 dimer could recruit it to the uracil lesion. Thus, reduced switching in the absence of Msh2-Msh6 could be due to both limited processing of recombination intermediates and creation of fewer nicks.

Msh6 Is Necessary for A-T Mutations in V and S Regions. We analyzed mutations in the intron region downstream of rearranged V-D-J_H4 genes and the Sμ region upstream of the core repeats for types of substitutions. For the V region, the mutational frequency in the mutated sequences was significantly lower in $Msh6^{-/-}$ B cells compared with C57BL/6 and $Msh3^{-/-}$ cells; a lower frequency in this region has also been reported in earlier studies on Msh6 and Msh2 (18, 19, 22). However, for the S region, the frequency of mutation for the mutated sequences was the same for all three genotypes. We propose that in the absence of Msh6, mutations are targeted to WGCW motifs, and because there are fewer of these in the V region versus the S region, there will be fewer mutations in the V region.

Mutations in both V and S regions from the $Msh6^{-/-}$ clones showed a dramatic decrease in substitutions of A and T bases, and a rise in mutations of G and C bases, compared with $Msh3^{-/-}$ and C57BL/6 clones. There were more substitutions of C on the nontranscribed strand from

Table I. *Mutations Are Focused in WGCW Motifs in J$_H$4 Intron and Sμ Regions in Msh6$^{-/-}$ and Msh2$^{-/-}$ Clones*

			No. mutations at G/C	
Nucleotide position[a]	Sequence	Target sequence	Msh6$^{-/-}$	Msh2$^{-/-}$[b]
J$_H$4 intron				
56	TTCTGA**GC**ATTGCA	**WGCW**	7/1	7/3
61	AGCATT**GC**AGACTA	**WGCW**	0/5	2/5
252	GAAGAT**GC**TAAAAC	**WGCW**	3/3	1/4
269	CCTATG**GC**TGGAGG	RGCW	1/0	ND
98	GAGCCG**GC**TGAGAG	RGCW	0/0	1/0
235	TTGGAG**GC**TCATTT	RGCW	0/0	0/0
287	GTTGGG**GC**TGTAGT	RGCW	0/0	ND
94	GAGGGA**GC**CGGCTG	WGCY	0/0	0/0
139	CTCAGA**GC**CTTTAG	WGCY	0/0	1/0
Sμ region				
391	CATAAA**GC**TTGCTG	**WGCW**	5/4	
215	GACAGT**GC**TTAGAT	**WGCW**	2/4	
451	TAAAAA**GC**ACAGCT	**WGCW**	3/3	
461	AGCTGA**GC**TGAGAT	**WGCW**	3/3	
281	GCCACA**GC**TGTACA	**WGCW**	3/2	
487	CTGAGT**GC**TTCTAA	**WGCW**	2/2	
529	AGGTAA**GC**AAAGCT	**WGCW**	2/1	
534	AGCAAA**GC**TGGGCT	**WGCW**	2/0	
456	AGCACA**GC**TGAGCT	**WGCW**	1/1	
25	TTGGAA**GC**TAATTT	**WGCW**	0/1	
400	TGCTGA**GC**AAAATT	**WGCW**	1/0	
312	AGACCT**GC**AGTTGA	**WGCW**	0/1	
325	AGGCCA**GC**AGGTCG	**WGCW**	0/0	
395	AAGCTT**GC**TGAGCA	**WGCW**	0/0	
475	GGGTGG**GC**TTCTCT	RGCW	0/3	
56	ACCCAG**GC**TAAGAA	RGCW	0/0	
65	AAGAAG**GC**AATCCT	RGCW	0/0	
248	ACAGGG**GC**TGGGGT	RGCW	0/0	
333	AGGTCG**GC**TGGACT	RGCW	0/0	
441	AGCGAG**GC**TCTAAA	RGCW	0/0	
166	GAAGTT**GC**CAGTAA	WGCY	0/0	
545	GCTTGA**GC**CAAAAT	WGCY	0/0	

[a]All of the WGCW, RGCW, and WGCY sequences in both regions are listed. The position of the first base in the underlined sequence is taken from Figs. S1 and S4, and shown in Figs. 4 and 5.
[b]Data from reference 18.

all three genotypes in the S region, which suggests that secondary structures in this region, such as stable R-loops, may enhance deamination of single-stranded C (16). These data demonstrate that the Msh2-Msh6 heterodimer is involved in generating mutations of A:T bp in V and S regions. Because exonuclease 1- and DNA polymerase η-deficient B cells also have fewer A:T mutations, all of these proteins may interact at the uracil lesion. For example, Msh2-Msh6 could bind to a U:G mismatch, and re-

cruit the other proteins to extend mutation past the original C target to include A:T pairs.

AID May Be Initially Targeted to WGCW Sites. Focused peaks of mutation were observed in the J$_H$4 intron in *Msh6$^{-/-}$* clones, which exactly coincided with the peaks noted for *Msh2$^{-/-}$* clones (Table I; references 18 and 19). Peaks of mutation were also identified in the Sμ region in *Msh6$^{-/-}$* clones, but not in *Msh3$^{-/-}$* or C57BL/6 clones (Fig. 5). The peaks occurred exclusively at WGCW motifs,

which are a subset of RGYW/WRCY hotspots: in the J_H4 intron, there are three motifs with three corresponding peaks of mutation; and in the Sμ region, there are 14 motifs with major peaks at five of them. Substitutions were found at either the G or C bases in WGCW, but were rarely found in other variants of RGYW such as GGCT, GGCA, AGCC, and TGCC. Recombination junctions are also focused at the frequent WGCW sequences (AGCT) in S regions from Msh2-deficient mice, but not from wild-type mice (31). AID has been shown to preferentially deaminate C in WRC on single-stranded DNA in vitro (26, 27), indicating that the protein has high affinity for this sequence. Thus, WGCW represents overlapping WGC motifs on both strands, which suggests that C is targeted by AID on either strand. W may be necessary on both ends because A:T has a lower melting temperature than G:C, which would be R or Y. Thus, unwinding of WGCW in DNA during transcription could expose C on both strands as a transcription bubble is being formed. Crystal structures of transcription of DNA by RNA polymerase II confirm that both strands are single stranded at the edges of the transcription bubble, particularly at the trailing edge where RNA is extruded (43).

Subsequent dispersal of mutations from the entry points in wild-type cells could be caused by a secondary round of AID attacks, where Msh2-Msh6 may facilitate the processive deamination of cytosines in other sequence contexts. For example, exonuclease 1 could generate long stretches of single-strand DNA that resemble the DNA substrates used in vitro, which show that C is deaminated in other WRC motifs (26, 27). Dispersed mutations could also result from error-prone synthesis by DNA polymerase η, which can synthesize in a gapped substrate. It remains to be seen how Msh2-Msh6 and pol η are recruited to uracil lesions, and how they function in the presence of other players, such as uracil glycosylase, abasic endonuclease, and exonuclease 1, in this error-prone DNA repair scenario.

Note added in proof. A defect in class switch recombination in $Msh6^{-/-}$ mice has also been reported (Li, Z., S.J. Scherer, D. Ronai, M.D. Iglesias-Ussel, J.U. Peled, P.D. Bardwell, M. Zhuang, K. Lee, A. Martin, W. Edelmann, and M.D. Scharff. 2004. *J. Exp. Med.* 47–59.).

References

1. Muramatsu, M., K. Kinoshita, S. Faragasan, S. Yamada, Y. Shinkai, and T. Honjo. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell.* 102:553–563.
2. Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Piebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Lagelouse, A. Gennery, et al. 2000. Activation-induced cytidine deaminase (AID) causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell.* 102:565–575.
3. Bransteitter, R., P. Pham, M.D. Scharff, and M.F. Goodman. 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. USA.* 100:4102–4107.
4. Chaudhuri, J., M. Tian, C. Khuong, K. Chua, E. Pinaud, and F.W. Alt. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature.* 422:726–730.
5. Ramiro, A.R., P. Stavropoulos, M. Jankovic, and M.C. Nussenzweig. 2003. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat. Immunol.* 4:452–456.
6. Dickerson, S.K., E. Market, E. Besmer, and F.N. Papavasiliou. 2003. AID mediates hypermutation by deaminating single stranded DNA. *J. Exp. Med.* 197:1291–1296.
7. Sohail, A., J. Klapacz, M. Samaranayake, A. Ullah, and A.S. Bhagwat. 2003. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res.* 31:2990–2994.
8. Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature.* 418:99–103.
9. Yoshikawa, K., I. Okazaki, T. Eto, K. Kinoshita, M. Muramatsu, H. Nagaoka, and T. Honjo. 2002. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science.* 296:2033–2036.
10. Martin, A., and M.D. Scharff. 2002. Somatic hypermutation of the AID transgene in B and non-B cells. *Proc. Natl. Acad. Sci. USA.* 99:12304–12308.
11. Okazaki, I., H. Hiai, N. Kakazu, S. Yamada, M. Muramatsu, K. Kinoshita, and T. Honjo. 2003. Constitutive expression of AID leads to tumorigenesis. *J. Exp. Med.* 197:1173–1181.
12. Woo, C.J., A. Martin, and M.D. Scharff. 2003. Induction of somatic hypermutation is associated with modifications in immunoglobulin variable region chromatin. *Immunity.* 19:479–489.
13. Poltoratsky, V.P., S.H. Wilson, T.A. Kunkel, and Y.I. Pavlov. 2004. Recombinogenic phenotype of human activation-induced cytosine deaminase. *J. Immunol.* 172:4308–4313.
14. Zeng, X., D.B. Winter, C. Kasmer, K.H. Kraemer, A.R. Lehmann, and P.J. Gearhart. 2001. DNA polymerase η is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat. Immunol.* 2:537–541.
15. Faili, A., S. Aoufouchi, S. Weller, F. Vuillier, A. Stary, A. Sarasin, C.-A. Reynaud, and J.-C. Weill. 2004. DNA polymerase η is involved in hypermutation occurring during immunoglobulin class switch recombination. *J. Exp. Med.* 199:265–270.
16. 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.
17. Phung, Q.H., D.B. Winter, A. Cranston, R.E. Tarone, V.A. Bohr, R. Fishel, and P.J. Gearhart. 1998. Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from mice deficient in the MSH2 mismatch repair protein. *J. Exp. Med.* 187:1745–1751.

18. Frey, S., B. Bertocci, F. Delbos, L. Quint, J.C. Weill, and C.A. Reynaud. 1998. Mismatch repair deficiency interferes with the accumulation of mutations in chronically stimulated B cells and not with the hypermutation process. *Immunity.* 9:127–134.

19. Rada, C., M.R. Ehrenstein, M.S. Neuberger, and C. Milstein. 1998. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity.* 9:135–141.

20. Schrader, C.E., S.P. Bradley, J. Vardo, S.N. Mochegova, E. Flanagan, and J. Stavnezer. 2003. Mutations occur in the Ig Sμ region but rarely in Sγ regions prior to class switch recombination. *EMBO J.* 22:5893–5903.

21. Acharya, S., T. Wilson, S. Gradia, M.F. Kane, S. Guerrette, G.T. Marsischky, R. Kolodner, and R. Fishel. 1996. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc. Natl. Acad. Sci. USA.* 93:13629–13634.

22. Wiesendanger, M., B. Kneitz, W. Edelmann, and M.D. Scharff. 2000. Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2-MSH6 heterodimer in modulating the base substitution pattern. *J. Exp. Med.* 191:579–584.

23. Li, Z., C.J. Woo, M.D. Iglesias-Ussel, D. Ronai, and M.D. Scharff. 2004. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev.* 18:1–11.

24. 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.

25. Rogozin, I.B., and N.A. Kolchanov. 1992. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim. Biophys. Acta.* 1171:11–18.

26. Pham, P., R. Bransteitter, J. Petruska, and M.F. Goodman. 2003. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature.* 424:103–107.

27. Yu, K., F.T. Huang, and M.R. Lieber. 2004. DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. *J. Biol. Chem.* 279: 6496–6500.

28. McDonald, J.P., E.G. Frank, B.S. Plosky, I.B. Rogozin, C. Masutani, F. Hanaoka, R. Woodgate, and P.J. Gearhart. 2003. 129-derived strains of mice are deficient in DNA polymerase ι and have normal immunoglobulin hypermutation. *J. Exp. Med.* 198:635–643.

29. 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.

30. 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.

31. Ehrenstein, M.R., and M.S. Neuberger. 1999. Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class switch recombination: parallels with somatic hypermutation. *EMBO J.* 18:3484–3490.

32. Schrader, C.E., W. Edelmann, R. Kucherlapati, and J. Stavnezer. 1999. Reduced isotype switching in splenic B cells from mice deficient in mismatch repair enzymes. *J. Exp. Med.* 190: 323–330.

33. Martin, A., Z. Li, D.P. Lin, P.D. Bardwell, M.D. Iglesias-Ussel, W. Edelmann, and M.D. Scharff. 2003. Msh2 ATPase activity is essential for somatic hypermutaiton in A-T base-pairs and for efficient class switch recombination. *J. Exp. Med.* 198:1171–1178.

34. Sugawara, N., F. Pâques, M. Colaiácovo, and J.E. Haber. 1997. Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc. Natl. Acad. Sci. USA.* 94:9214–9219.

35. Ehrenstein, M.R., C. Rada, A.M. Jones, C. Milstein, and M.S. Neuberger. 2001. Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. *Proc. Natl. Acad. Sci. USA.* 98:14553–14558.

36. Schrader, C.E., J. Vardo, and J. Stavnezer. 2002. Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. *J. Exp. Med.* 195:367–373.

37. Min, I.M., C.E. Schrader, J. Vardo, R.M. Luby, N. D'Avirro, J. Stavnezer, and E. Selsing. 2003. The Sμ tandem repeat region is critical for Ig isotype switching in the absence of Msh2. *Immunity.* 19:515–524.

38. 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.

39. Di Noia, J.M., and M.S. Neuberger. 2004. Immunoglobulin gene conversion in chicken DT40 cells largely proceeds through an abasic site intermediate generated by excision of the uracil produced by AID-mediated deoxycytidine deamination. *Eur. J. Immunol.* 34:504–508.

40. Bardwell, P.D., C.J. Woo, K. Wei, Z. Li, A. Martin, S.Z. Sack, T. Parris, W. Edelmann, and M.D. Scharff. 2004. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. *Nat. Immunol.* 5: 224–229.

41. Schmutte, C., R.C. Marinescu, M.M. Sadoff, S. Guerrete, J. Overhauser, and R. Fishel. 1998. Human exonuclease I interacts with the mismatch repair protein hMSH2. *Cancer Res.* 58:4537–4542.

42. Rasmussen, L.J., M. Rasmussen, B. Lee, A.K. Rasmussen, D.M. Wilson III, F.C. Nielsen, and H.C. Bisgaard. 2000. Identification of factors interacting with hMSH2 in the fetal liver utilizing the yeast two-hybrid system. In vivo interaction through the C-terminal domains of hEXOI and hMSH2 and comparative expression analysis. *Mutat. Res.* 460:41–52.

43. Gnatt, A.L., P. Cramer, J. Fu, D.A. Bushnell, and R.D. Kornberg. 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science.* 292:1876–1882.