

# Absence of DNA Polymerase $\eta$ Reveals Targeting of C Mutations on the Nontranscribed Strand in Immunoglobulin Switch Regions

Xianmin Zeng, George A. Negrete, Cynthia Kasmer, William W. Yang, and Patricia J. Gearhart

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

## Abstract

Activation-induced cytosine deaminase preferentially deaminates C in DNA on the nontranscribed strand in vitro, which theoretically should produce a large increase in mutations of C during hypermutation of immunoglobulin genes. However, a bias for C mutations has not been observed among the mutations in variable genes. Therefore, we examined mutations in the  $\mu$  and  $\gamma$  switch regions, which can form stable secondary structures, to look for C mutations. To further simplify the pattern, mutations were studied in the absence of DNA polymerase (pol)  $\eta$ , which may produce substitutions of nucleotides downstream of C. DNA from lymphocytes of patients with xeroderma pigmentosum variant (XP-V) disease, whose polymerase  $\eta$  is defective, had the same frequency of switching to all four  $\gamma$  isotypes and hypermutation in  $\mu$ - $\gamma$  switch sites (0.5% mutations per basepair) as control subjects. There were fewer mutations of A and T bases in the XP-V clones, similar to variable gene mutations from these patients, which confirms that polymerase  $\eta$  produces substitutions opposite A and T. Most importantly, the absence of polymerase  $\eta$  revealed an increase in C mutations on the nontranscribed strand. This data shows for the first time that C is preferentially mutated in vivo and pol  $\eta$  generates hypermutation in the  $\mu$  and  $\gamma$  switch regions.

Key words: somatic hypermutation • activation-induced cytosine deaminase • xeroderma pigmentosum variant • gap-filling repair • translesion replication

## Introduction

Immunoglobulin diversity is achieved in mammals at three molecular levels: joining of variable (V), diversity, and joining gene segments; hypermutation of rearranged V genes; and switching of heavy chain constant (C) genes. The latter two processes depend on the activation-induced cytosine deaminase (AID) protein. Mice and humans without AID have neither V gene mutations nor C gene switching (1, 2). This implies that the two processes, which involve different mechanisms that generate point mutations and double strand breaks, share common enzymes. Biochemical and genetic experiments indicate that AID functions to deaminate cytosine in DNA to uracil (3–10). Using gene-deficient mice and humans, several other proteins have been shown to alter the pattern of mutation and switching. First, uracil DNA glycosylase is required to remove uracil lesions in

DNA. Mice and humans deficient for the enzyme have altered V gene mutations and deficient heavy chain switching (5, 11). Second, the mismatch repair proteins Msh2, Msh6, Pms2, and Mlh1, participate in an unknown way to change the pattern of V gene mutations and C gene switching in gene-deficient mice (12–18). Third, DNA polymerase (pol)  $\eta$  plays a role in generating mutations in V genes at A and T nucleotides. Humans with xeroderma pigmentosum variant (XP-V) disease whose pol  $\eta$  is defective have fewer A:T substitutions (19).

Mutations have also been detected in introns containing the recombination sites of switched C genes (20). Mutations are even found in the  $\mu$  switch region before switching in murine B cells stimulated in culture, and they are dependent on AID expression (21, 22). Although mutations in V genes have the potential to change the coding sequence to produce high affinity antibodies, mutations in the switch regions likely reflect the footprints of AID deamination

The online version of this article contains supplemental material.

Address correspondence to Patricia J. Gearhart, Laboratory of Molecular Gerontology, National Institute on Aging, National Institutes of Health, 5600 Nathan Shock Drive, Baltimore, MD 21224. Phone: (410) 558-8561; Fax: (410) 558-8157; email: gearhartp@grc.nia.nih.gov

*Abbreviations used in this paper:* AID, activation-induced cytosine deaminase; pol, DNA polymerase; XP-V, xeroderma pigmentosum variant.

that precede DNA strand breaks. Actual recombination requires additional factors because defects in the carboxy-terminal region of the AID protein do not affect hypermutation but eliminate switching (23, 24).

In vitro, AID preferentially deaminates C on single stranded DNA or on the nontranscribed strand during transcription (6–10, 25). This leads to the conundrum that there should be a large increase in mutations of C compared with those of G, A, and T, as recorded from the nontranscribed strand. However, in vivo, mutations of all four nucleotides in V genes are approximately equal in frequency (26), suggesting that other proteins, e.g., mismatch repair proteins and pols, generate mutations downstream of deaminated cytosines. In this study, we looked for a C bias in mutations from switch regions, which can form stable secondary structures that expose single strands. To further simplify the pattern, we removed pol  $\eta$  and analyzed mutations in the  $\mu$ - $\gamma$  switch joins from three XP-V patients. We show that XP-V patients have normal switching, but as in V genes, there were fewer mutations of A and T nucleotides. Furthermore, the absence of pol  $\eta$  revealed preferential targeting of C bases for mutation.

## Materials and Methods

**DNA Preparation from Peripheral Blood Lymphocytes.** Peripheral blood was obtained from patients XP11BR, XP31BE, and XP7BR as previously described (19), and control donors as reported (27). DNA was extracted with phenol-chloroform and suspended in Tris-EDTA buffer.

**Libraries of  $\mu$ - $\gamma$  Switch Regions.** Hybrid switch sites containing  $\mu$  and  $\gamma$  switch (S) regions were amplified by PCR using primers flanking the repetitive core regions of S $\mu$  and S $\gamma$ . The S $\gamma$  primers were specific for a conserved region downstream of all S $\gamma$  sequences, and thus would detect  $\gamma$ 3,  $\gamma$ 1,  $\gamma$ 2, and  $\gamma$ 4 switches (28). The following sets of nested primers were used: first set, S $\mu$  forward (nucleotides 91–110 from GenBank/EMBL/DDBJ under accession no. 54713), 5'CAAGCAGGTCTGGTGGGCTG; S $\gamma$  reverse (nucleotides 2911–2933 from GenBank/EMBL/DDBJ under accession no. U39737), 5'CTTGCCAAGTCTCAGTGGGATG; second set, S $\mu$  forward (nucleotides 118–141 from GenBank/EMBL/DDBJ under accession no. X54713) with EcoRI addition in italics, 5'GCCGGAATTCCTGGCCATGACAACCTCCATCCAGC; S $\gamma$  reverse (nucleotides 2861–2882 of U39737) with BamHI addition in italics, 5'GCGGGATCCGGCTGCACTGCACTTTCACCAG. 15 ng genomic DNA was amplified with Pfu polymerase in 50  $\mu$ l volume using the first set of primers for 30 cycles of 95°C for 45 s, 55°C for 1 min, and 72°C for 2 min, followed by a final incubation at 72°C for 10 min. 5  $\mu$ l of this reaction was then reamplified in 50  $\mu$ l with the second set of primers for an additional 30 cycles. The PCR products were cloned into EcoRI and BamHI-digested pBlue-script and plasmids containing unique inserts were sequenced.

**Determination of PCR Error.** To assess PCR error, unrearranged S $\gamma$ 1 regions from all six subjects were sequenced from the same DNA using the same procedures as described above. The following sets of primers were used to generate a 1.37-kb product: first set, S $\gamma$  forward (nucleotides 1430–1449 from GenBank/EMBL/DDBJ under accession no. U39737), 5'AAGCAGAAA-GATCAGGGGTC; S $\gamma$  reverse as above; second set, S $\gamma$  forward (nucleotides 1505–1523 from GenBank/EMBL/DDBJ under accession no. U39737) with EcoRI addition in italics, 5'CGGA4-

TTCCTCAGCCTCAGGGAGCCAGG; S $\gamma$  reverse with BamHI as above. 20 ng genomic DNA was amplified with Platinum Pfx polymerase and PCRx enhancer (Invitrogen) in 50  $\mu$ l volume using the first set of primers for 30 cycles of 95°C for 30 s, 55°C for 1 min, and 68°C for 2 min, followed by a final incubation at 68°C for 10 min. Nested PCR was performed with 5  $\mu$ l of the first reaction and the second set of primers with an annealing and extension step of 68°C for 2 min for another 30 cycles. Products were digested, cloned, and sequenced.

**Online Supplemental Material.** Figs. S1–S6 present the data from three XP-V and three control groups of clones. They contain a summary of clone length, isotypes, and microhomology, sequences of the  $\mu$  and  $\gamma$  switch regions, and examples of microhomology at the  $\mu$ - $\gamma$  junctions. Figs. S1–S6 are available at <http://www.jem.org/cgi/content/full/jem.20032022/DC1>.

## Results

### XP-V Patients Have Normal Class Switch Recombination.

To analyze C $\mu$  to C $\gamma$  switching and identify mutations in the switch regions from DNA pol  $\eta$ -deficient humans, peripheral blood was obtained from three XP-V patients. The DNA repair defects in these patients have been described, and the mutations in their *POLH* genes were expected to inactivate pol  $\eta$  (19). Recombined switch sites associated with S $\mu$  and S $\gamma$  regions were sequenced from cloned PCR products derived from the DNA of the XP-V patients and three control individuals. To avoid nonspecific priming in the repetitive sequences, DNA was amplified with nested primers. Inserts ranged in size from 100 to 600 bp. The  $\gamma$  isotypes and mutations in the switching sites were identified by comparing the recombined S $\mu$ -S $\gamma$  regions to the germline S $\mu$  sequence and to all four S $\gamma$  sequences (28). As shown in Table I, usage of the  $\gamma$  isotypes for XP-V and control clones was similar, with over half of the clones containing  $\mu$  to  $\gamma$ 1 switches. Breakpoints in the  $\mu$  and  $\gamma$  regions occurred randomly, and there was no difference between the two groups in the lengths of microhomology at the joining sites. Approximately half of the clones had insertions or no microhomology at the  $\mu$ - $\gamma$  joins, and half had homology of one to five nucleotides. Summaries of all

**Table I.** Immunoglobulin G Switch Region Isotypes in XP-V and Control Clones

Subject	Number of clones					
	Total	$\mu$ to $\gamma$ 3	$\mu$ to $\gamma$ 1	$\mu$ to $\gamma$ 2	$\mu$ to $\gamma$ 4	$\mu$ to ? <sup>a</sup>
XP11BR	23	4	15	1	2	1
XP31BE	23	4	14	4	0	1
XP7BR	25	3	14	6	1	1
Control 1	31	4	14	3	2	8
Control 2	20	3	14	0	2	1
Control 3	16	0	13	0	3	0

<sup>a</sup>Nucleotide sequence after the  $\mu$  break was not identified.

**Table II.** Types of Mutation in  $\mu$ - $\gamma$  Switch Regions

Subject	Frequency (% mutation/bp)	Nucleotides sequenced	Number <sup>a</sup> of		
			Substitutions	Deletions	Insertions
XP11BR	0.59	5549	26	6	1
XP31BE	0.35	5681	20	0	0
XP7BR	0.51	5840	28	2	0
Control 1	0.75	6028	38	6	1
Control 2	0.49	5128	19	6	0
Control 3	0.56	4259	20	3	1
PCR error <sup>b</sup>	0.08	7248	5	1	0

<sup>a</sup>Does not include mutations or insertions at the  $\mu$ - $\gamma$  junctions.

<sup>b</sup>Mutations in recombined switch regions differed significantly from PCR error,  $P < 10^{-4}$ .

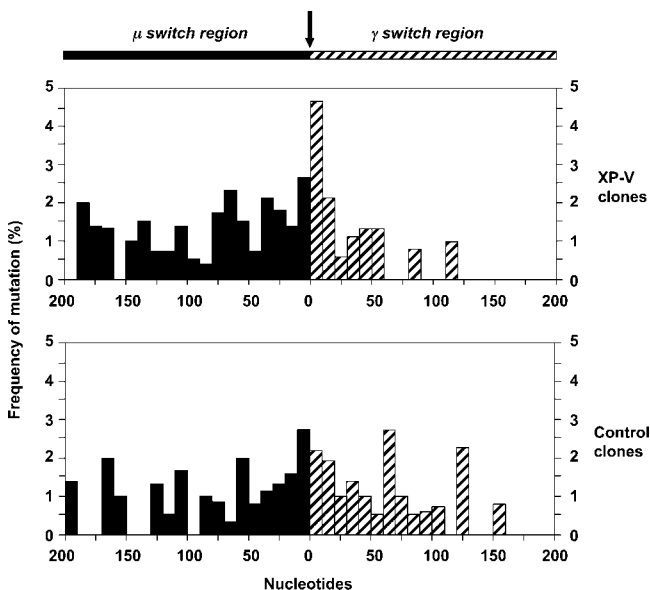
the clones, sequences, and examples of microhomology are included in Figs. S1–S6, available at <http://www.jem.org/cgi/content/full/jem.20032022/DC1>.

*Mutations Are Located throughout  $\mu$  and  $\gamma$  Switch Regions.* The sequenced  $S\mu$  region was nonpolymorphic among the six subjects, so that mutations were readily identified. The  $S\gamma$  regions were polymorphic, and germline  $S\gamma$  sequences were identified for each human. To avoid errors in assigning mutations for the other isotypes, only unique mutations were counted. Over two thirds of the clones from both study groups had mutations. As seen in Table II, the overall

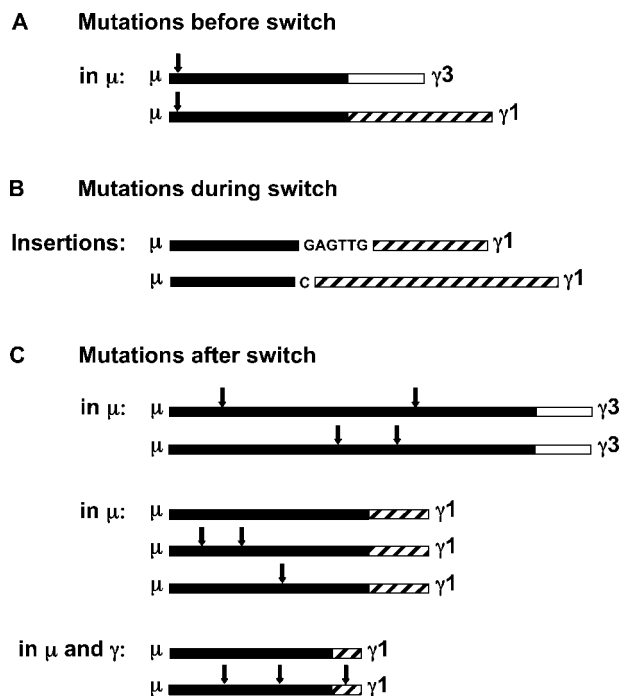
frequency of mutation was similar between XP-V and control clones, with an average of 0.5% mutations per bp. This was significantly higher than the mutation frequency in unrearranged  $S\gamma$  regions ( $P < 10^{-4}$ , Fisher's exact test), which is reported to be similar to the frequency of PCR error (29). As in V genes, the majority of mutations were base substitutions, although the frequency of deletions was much higher in the switch region ( $\sim 13\%$  of mutations) compared with that in introns flanking rearranged V genes (1%). Approximately 2% of the mutations within the switch sequences were insertions and 16% of the clones had insertions at the junction site. The distance and frequency of mutations from the recombination sites of  $S\mu$  and  $S\gamma$  are shown in Fig. 1. Mutations were scattered throughout the switch regions and occurred as far as 200 bp from the joining site. Preferential clustering of mutations around the break points was not observed.

*Mutations May Occur Before, During, and After Switching.* Related clones were identified in all the libraries and likely derive from clonal progeny that have undergone sequential mutations. Approximately one third of the clones were related as defined by the same length of  $\mu$  and/or  $\gamma$  sequences and the same site of joining. Based on the pattern of shared and unique mutations in related clones, it is possible to estimate when the mutations occurred. As shown in Fig. 2 A, mutations could happen before switching in the unrecombined  $\mu$  region, as demonstrated by two clones with identical substitutions that were joined to different  $S\gamma$  sequences. Mutations in the  $S\mu$  region before switching have been well documented in the literature (21, 22, 29), indicating that they precede recombination. In Fig. 2 B, mutations could occur during switching as shown by insertions of nontemplated nucleotides at the site of joining. In Fig. 2 C, mutations probably occurred after switching in three sets of clones that had identical  $\mu$  and  $\gamma$  joins and shared mutations, as well as unique mutations.

*Pol  $\eta$  Is an A-T Mutator in the  $\mu$ - $\gamma$  Switch Regions.* The types of base substitutions were examined to see if there was a difference between XP-V and control clones. All



**Figure 1.** Mutations are distributed throughout the  $\mu$ - $\gamma$  switch regions at a high frequency. The top schematic drawing shows the recombined  $\mu$  (solid) and  $\gamma$  (striped) switch regions with an arrow indicating the breakpoint. On the abscissa, mutations are plotted on either side of the breakpoint for XP-V and control clones. On the ordinate, the frequency of mutation per 10 nucleotide increments was calculated as the number of mutations in the increment, divided by 10, divided by the number of times the increment was sequenced, and multiplied by 100.



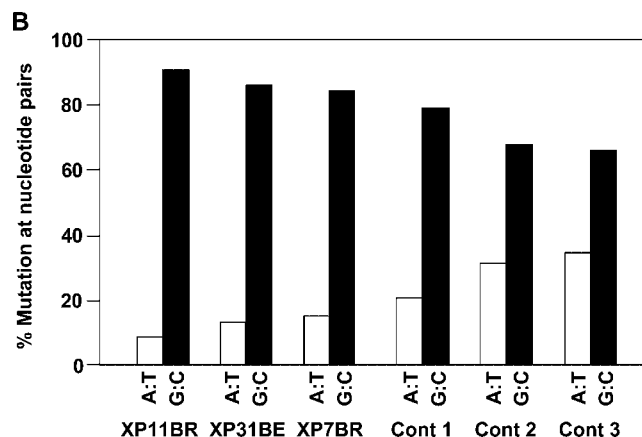
**Figure 2.** Mutations can occur sequentially in clonal progeny before and after switching. Clonal daughters were identified by identical break-points and unique mutations are depicted by arrows. (A) Before switching in the  $\mu$  region, two clones (control 1, 2–21 and 93) with identical mutations in  $\mu$  were joined to  $\gamma3$  and  $\gamma1$ . (B) During switching, two unique clones (control 1: 2–24 and 4–B4) had nucleotide insertions at the site of joining. (C) After switching, three examples are shown. In the  $\mu$  region, two clones (control 2: 11A9 and 11G9) with identical  $\gamma3$  junctions had different mutations in  $\mu$ . Also in the  $\mu$  region, three clones (XP7BR: 12D12, 65, and 67) with identical  $\gamma1$  regions had different mutations in  $\mu$ . In the  $\mu$  and  $\gamma$  regions, two clones (XP11BR: 5A1 and 5A2) with the same junction had different mutations in  $\mu$  and  $\gamma$ . Sequences of these clones are shown in Figs. S1, S3, and S4.

mutations were recorded from the nontranscribed strand. There was a decrease in mutations of A and T, and a corresponding increase in mutations of G and C in the XP-V clones (Fig. 3 A). All three XP-V patients had clones with decreased A:T mutations (Fig. 3 B), indicating that pol  $\eta$  is involved in generating mutations in the switch regions. Comparing mutations of A or T versus mutations of G or C, there is a highly significant difference among the XP-V and control clones ( $P = 0.0035$ , using a Monte Carlo analysis; reference 30). In addition to transitions of G and C, there was a strong bias for G to C and C to G transversions in both groups of clones. For example, of the mutations of G and C, 58% were G:C to A:T transitions, 8% were G:C to T:A transversions, and 34% were G:C to C:G transversions. There was no difference in the types of mutations located within 15 bp of the switch junction and those located farther away.

*C on the Nontranscribed Strand Is Targeted for Mutation in XP-V Clones.* Because the  $S\mu$  sequence was nonpolymorphic, we analyzed mutations there in more detail. Some 130 nucleotides of the  $S\mu$  region located upstream of the repetitive core sequences are shown in Fig. 4. The lo-

**A**

Substitution	XP-V (74 mutations)	Control (77 mutations)
	%	%
A to: G	0	10
T	1	5
C	3	6
T to: C	3	3
A	2	3
G	4	0
G to: A	18	15
T	3	4
C	8	9
C to: T	30	29
A	3	3
G	25	13

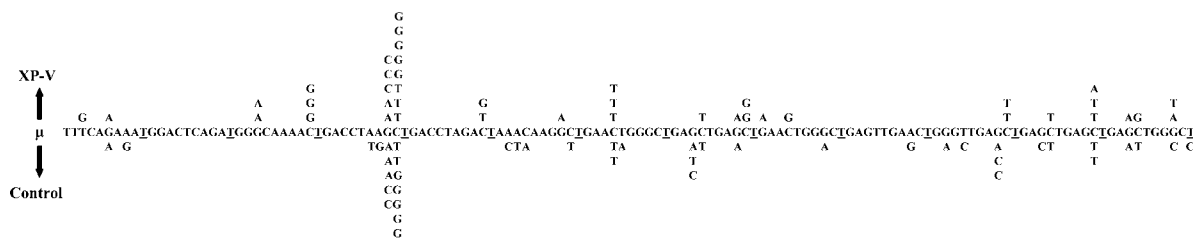


**Figure 3.** Pol  $\eta$  mutates A and T nucleotides in  $\mu$ - $\gamma$  switch regions. (A) Substitutions in clones from three XP-V and three control subjects show differences in spectra. Data have been corrected for nucleotide composition of the  $S\mu$ ,  $S\gamma1$ ,  $S\gamma2$ ,  $S\gamma3$ , and  $S\gamma4$  regions that were sequenced. (B) Substitutions from each individual are grouped as A:T and G:C pairs.

cation of mutations reveals a hotspot at C in position 39 (position 180 in GenBank/EMBL/DDBJ under accession no. X54713) in both XP-V and control clones. The position is in a WRC motif (31), although it is unclear why this particular position is favored over the other WRC sequences in Fig. 4. When all the substitutions in the 350-nucleotide  $S\mu$  region were tabulated (Fig. 5 A), it became apparent that C was preferentially mutated compared with G in the XP-V clones (Fig. 5 B). Some 70% of the 52 mutations are at C nucleotides in XP-V clones compared with 47% of 50 mutations in the control clones ( $P = 0.05$ ). Targeting of C bases is also observed when the  $S\gamma$  substitutions are included (Fig. 3 A). 58 and 45% are at C in XP-V and control clones, respectively, compared with 29 and 28% at G, respectively. Half of the C mutations in XP-V clones were C to T transitions and half were C to G transversions.

## Discussion

*Switch Regions Are Hotspots for Hypermutation.* The importance of the switch region as a hypermutation target has only recently been recognized. Mutations have been reported in  $S\mu$  regions from mice and humans (21, 22, 32),



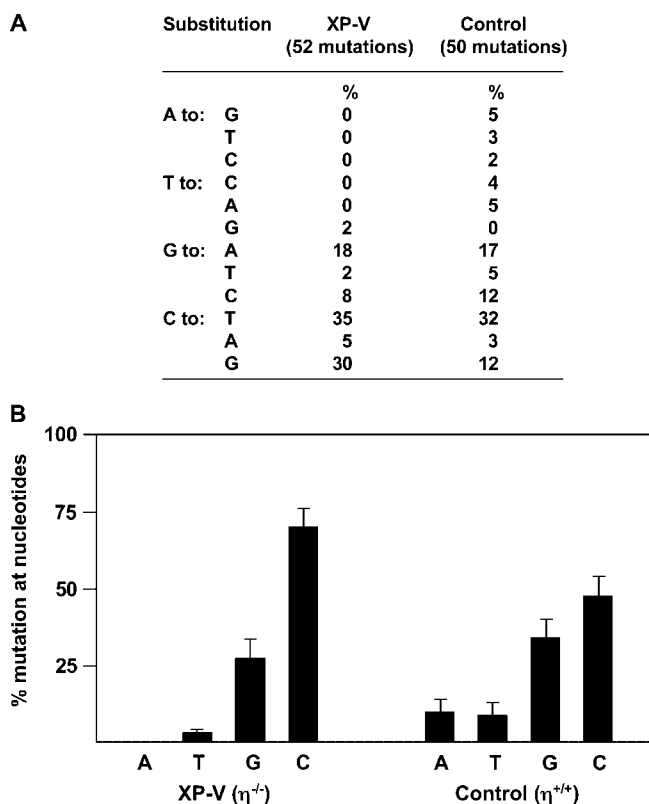
**Figure 4.** Location of substitutions in the  $\mu$  switch region. The 130-base region shown is from HSIGHMUS X54713 (nucleotides 142–271). Mutations from XP-V and control clones are shown above and below the germline sequence, respectively. Every 10th base is underlined.

and as we describe here, extensive mutations are found in the  $S\gamma$  regions as well. The overall frequency of mutation in these human clones is 0.5% mutations per bp, which is identical to the frequency of mutation seen in intron regions adjacent to rearranged  $V_H$  genes from humans (unpublished data). Therefore, the question of how AID is targeted must take into account the sequences of two very different regions of DNA. Targeting to the V region might be initiated by association with the transcription complex near the V gene promoter (33). Similarly, targeting to the distal switch sites may occur by association with the transcription complexes that form upstream of unrearranged switch regions (34). The location of mutations in the 130-nucleotide  $S\mu$  sequence from both XP-V and control

clones (Fig. 4) indicates a hotspot at 39C, which is also observed in another report (32) and may represent a major entry point for AID. This hotspot, located at the 5' beginning of the  $S\mu$  pentamer core motifs, is not in a palindromic stem-loop structure, but might be favored in the formation of other secondary structures.

Successive mutations in clonal progeny have been documented in V gene clones and indicate that mutations are generated over several rounds of division of B cells in germinal centers. Here we show that mutations also occur sequentially in the switch regions. As demonstrated in Fig. 2, related clones have the same site of joining in  $S\mu$  and  $S\gamma$ , and contain both shared and unique mutations. It appears that mutations can occur before, during, and after joining in  $S\mu$  and  $S\gamma$ . Mutations before switching have been reported by others (21, 22, 29) and suggest that switch regions undergo multiple deaminations, and repair of the lesions produces mutations and occasional deletions caused by strand breaks. Internal deletions in switch regions have been previously noted (35–37). Among the mutations in this study, 13% were internal deletions in  $S\mu$  and  $S\gamma$ . Presumably breaks of this type could also trigger recombination between different switch regions. Mutations during switching (20, 38) are clearly identified at the junctions between  $S\mu$  and  $S\gamma$ , where insertions are commonly found. Even after recombination, the hybrid  $S\mu$ – $S\gamma$  sequences continue to sustain more mutations, more breaks, and likely sequential switching (28, 39) to downstream C genes. Successive mutations would explain why the mutational pattern is spread out over 200 bp from the junction site (Fig. 1; reference 32) in these human clones. In contrast to short-term cultures of murine lymphocytes where the mutational pattern is more clustered around the junction site (40), human peripheral B cells are long-lived and could accumulate mutations long after switching.

*Absence of DNA Pol  $\eta$  Reveals Increased C Mutations.* DNA pol  $\eta$  is a low fidelity polymerase that preferentially generates mutations opposite A and T nucleotides in vitro (41, 42). Humans with XP-V disease are deficient in pol  $\eta$ , and their V genes have fewer mutations of A and T, which is consistent with pol  $\eta$  being an A-T mutator in hypermutation (19). To determine if pol  $\eta$  is also involved in generating mutations in switch regions, we examined the spectra of mutations in clones from three XP-V patients and three control subjects. Both groups had the same frequency of mutation and similar  $\gamma$  isotypes, but the XP-V clones had

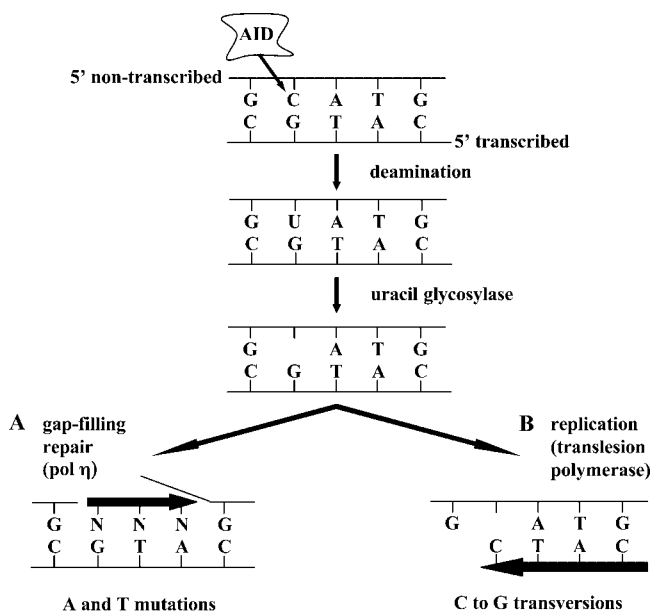


**Figure 5.** C bases in the  $\mu$  switch region are preferentially mutated in pol  $\eta$ -deficient clones. (A) Substitutions were corrected for nucleotide composition of each clone. (B) Total mutations of each nucleotide are shown for XP-V and control clones, with standard errors for individual variation.

fewer mutations of A and T bases, suggesting that pol  $\eta$  synthesizes mutations in both the V and switch regions. The mutation frequency is not decreased in the XP-V clones, perhaps because AID repeatedly deaminates C bases and produces more C:G substitutions in the absence of pol  $\eta$ .

Biochemical evidence indicates that AID preferentially deaminates C on the nontranscribed strand, which would be single stranded during transcription (7, 8, 10). However, the mutational spectra in V gene introns show an equal frequency of C and G mutations on the nontranscribed strand (5, 43). This suggests that the separation of DNA in V genes during transcription might be too transient to favor deamination on one strand versus the other, or that pols insert mutations downstream of the initial C lesion. In contrast, DNA in switch regions is rich in G and C nucleotides and has been suggested to form stable secondary structures including R loops (44, 45), G-quartets (46), and stem loops (47). These structures might be induced by transcription (48) and expose C on the nontranscribed strand for deamination for a length of time. In support of this model, we observed an increase in mutations of C on the nontranscribed strand in human S $\mu$  and S $\gamma$  regions. A slight increase was seen in the mutations from control clones, and this effect was dramatically augmented in XP-V clones where some 70% of mutations were at C bases in S $\mu$  (Fig. 5).

The three major categories of hypermutation in the switch regions from control subjects are substitutions of A and T, transversions of C:G to G:C, and transitions of C:G



**Figure 6.** Model of mutations introduced during repair and replication of deaminated cytosine in switch regions. AID deaminates C on the nontranscribed strand to uracil (U). U is removed by uracil glycosylase to produce an abasic site. (A) During gap-filling repair, pol  $\eta$  synthesizes a short distance (thick arrow). Mutations opposite T and A are introduced on the nontranscribed strand. (B) During replication, a translesion polymerase such as Rev1 inserts C opposite the abasic site as it synthesizes the transcribed strand (thick arrow). This would produce transversions of C to G as recorded from the nontranscribed strand.

to T:A. We propose a model to explain how these mutations are generated (Fig. 6). Substitutions of A and T could occur during gap-filling repair of the abasic site on the nontranscribed strand by pol  $\eta$ , which synthesizes DNA downstream of the initial C lesion to generate mutations opposite all four nucleotides. Pol  $\eta$  could synthesize a gap produced by exonuclease 1 (49) and/or perform strand displacement (43). In the absence of pol  $\eta$ , mutations would occur primarily at sites of deaminated cytosines, which is why C is targeted in the XP-V clones. In the second category, transversions of C:G to G:C could occur during replication past the abasic site by a translesion polymerase. Half of the mutations of C in XP-V clones were C to G transversions, and a high frequency of transversions was also found in the control clones (Fig. 5; reference 32). C to G transversions could be specifically generated by Rev1, which is a deoxycytidyl transferase that inserts C opposite an abasic site (50, 51). The chicken DT40 B cell line makes a preponderance of C:G to G:C transversions in immunoglobulin V genes (52), and deletion of Rev1 eliminates these mutations (53), indicating that this polymerase is involved in chicken hypermutation. Our data support the notion that Rev1 participates in human switch mutations as well. Rev1 associates with several polymerases, including pol  $\eta$ , pol  $\iota$ , and pol  $\zeta$  (50, 54), and might be recruited to the mutation site in a multiprotein complex. The third category of mutations is C:G to T:A transitions. Transitions could arise during error-prone repair by preferential insertion of T opposite G by pol  $\eta$ , or during replication past uracil by any polymerase.

In addition to pol  $\eta$ , the mismatch repair proteins Msh2 and Msh6 participate in generating mutations of A and T bases in V genes. It remains to be determined how the low fidelity pols and mismatch repair proteins are recruited to deaminated cytosines in the immunoglobulin locus, and how pol  $\eta$  and Msh2/Msh6 produce A:T mutations.

We are grateful to Robert Tarone and Igor Rogozin for statistical analyses. We also thank Ed Max for PCR strategies, Roger Woodgate and John Tainer for comments, and Vilhelm Bohr for support.

Submitted: 24 November 2003

Accepted: 5 February 2004

Note added in proof. Two recent papers have also observed a preference for C mutations on the non-transcribed strand in switch regions: Cook et al. (55) and Faili et al. (56).

## 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. 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.
4. Di Noia, J., and M.S. Neuberger. 2002. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature*. 419:43–48.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. Sohail, A., J. Klapacz, M. Samaranyake, 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.
11. Imai, K., G. Slupphaug, W.I. Lee, P. Revy, S. Nonoyama, N. Catalan, L. Yel, M. Forveille, B. Kavli, H.E. Krokan, et al. 2003. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nat. Immunol.* 4:1023–1028.
12. Winter, D.B., Q.H. Phung, A. Umar, S.M. Baker, R.E. Tarone, K. Tanaka, R.M. Liskay, T.A. Kunkel, V.A. Bohr, and P.J. Gearhart. 1998. Altered spectra of hypermutation in antibodies from mice deficient for the DNA mismatch repair protein PMS2. *Proc. Natl. Acad. Sci. USA*. 95:6953–6958.
13. 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.
14. 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.
15. Rada, C., M.R. Ehrenstein, M.S. Neuberger, and C. Milstein. 1998. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stage of mutational targeting. *Immunity*. 9:135–141.
16. Kim, N., G. Bozek, J.C. Lo, and U. Storb. 1999. Different mismatch repair deficiencies all have the same effects on somatic hypermutation: intact primary mechanism accompanied by secondary modifications. *J. Exp. Med.* 190:21–30.
17. 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.
18. 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.
19. Zeng, X., D.B. Winter, C. Kasmer, K.H. Kraemer, A.R. Lehmann, and P.J. Gearhart. 2001. DNA polymerase  $\eta$  is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat. Immunol.* 2:537–541.
20. Dunnick, W., M. Wilson, and J. Stavnezer. 1989. Mutations, duplication, and deletion of recombined switch regions suggest a role for DNA replication in the immunoglobulin heavy-chain switch. *Mol. Cell. Biol.* 9:1850–1856.
21. Petersen, S., R. Casellas, B. Reina-San-Martin, H.T. Chen, M.J. Difilippantonio, P.C. Wilson, L. Hanitsch, A. Celeste, M. Muramatsu, D.R. Pilch, et al. 2001. AID is required to initiate Nbs1/ $\gamma$ -H2AX focus formation and mutations at sites of class switching. *Nature*. 414:660–665.
22. Nagaoka, H., M. Muramatsu, N. Yamamura, K. Kinoshita, and T. Honjo. 2002. Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin S $\mu$  region: implication of AID involvement in a common step of class switch recombination and somatic hypermutation. *J. Exp. Med.* 195:529–534.
23. Ta, V.T., H. Nagaoka, N. Catalan, A. Durandy, A. Fischer, K. Imai, S. Nonoyama, J. Tashiro, M. Ikegawa, S. Ito, et al. 2003. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat. Immunol.* 4:843–848.
24. Barreto, V., B. Reina-San-Martin, A.R. Ramiro, K.M. McBride, and M.C. Nussenzweig. 2003. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol. Cell.* 12:501–508.
25. 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.
26. Smith, D.S., G. Creadon, P.K. Jena, J.P. Portanova, B.L. Kotzin, and L.J. Wysocki. 1996. Di- and trinucleotide target preferences of somatic mutagenesis in normal and auto-reactive B cells. *J. Immunol.* 156:2642–2652.
27. Rosner, K., D.B. Winter, C. Kasmer, G.L. Skovgaard, R.E. Tarone, V.A. Bohr, and P.J. Gearhart. 2001. Impact of age on hypermutation of immunoglobulin variable genes in humans. *J. Clin. Immunol.* 21:102–115.
28. Mills, F.C., M.P. Mitchell, N. Harindranath, and E.E. Max. 1995. Human Ig S $\gamma$  regions and their participation in sequential switching to IgE. *J. Immunol.* 155:3021–3036.
29. Schrader, C.E., S.P. Bradley, J. Vardo, S.N. Mochegova, E. Flanagan, and J. Stavnezer. 2003. Mutations occur in the Ig S $\mu$  region but rarely in S $\gamma$  regions prior to class switch recombination. *EMBO J.* 22:5893–5903.
30. Cariello, N.F., W.W. Piegorsch, W.T. Adams, and T.R. Skopek. 1994. Computer program for the analysis of mutational spectra: application to p53 mutations. *Carcinogenesis*. 15:2281–2285.
31. 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.
32. Pan-Hammarström, Q., S. Dai, Y. Zhao, I.F. van Dijk-Härd, R.A. Gatti, A.L. Børresen-Dale, and L. Hammarström. 2003. ATM is not required in somatic hypermutation of V<sub>H</sub>, but is involved in the introduction of mutations in the switch  $\mu$  region. *J. Immunol.* 170:3707–3716.
33. Storb, U., and J. Stavnezer. 2002. Immunoglobulin genes: generating diversity with AID and UNG. *Curr. Biol.* 12:

R725–R727.

34. Stavnezer-Nordgren, J., and S. Sirlin. 1986. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. *EMBO J.* 5:95–102.
35. Dunnick, W., G.Z. Hertz, L. Scappino, and C. Gritzmacher. 1993. DNA sequences at immunoglobulin switch region recombination sites. *Nucleic Acids Res.* 21:365–372.
36. Zhang, K., H.K. Cheah, and A. Saxon. 1995. Secondary deletion recombination of rearranged switch region in Ig isotype-switched B cells. *J. Immunol.* 154:2237–2247.
37. Dudley, D.D., J.P. Manis, A.A. Zarrin, L. Kaylor, M. Tian, and F.W. Alt. 2002. Internal IgH class switch region deletions are position-independent and enhanced by AID expression. *Proc. Natl. Acad. Sci. USA.* 99:9984–9989.
38. Dunnick, W., and J. Stavnezer. 1990. Copy choice mechanism of immunoglobulin heavy chain switch recombination. *Mol. Cell. Biol.* 10:397–400.
39. Gearhart, P.J., J.L. Hurwitz, and J.J. Cebra. 1980. Successive switching of antibody isotypes expressed within the lines of a B-cell clone. *Proc. Natl. Acad. Sci. USA.* 77:5424–5428.
40. Chen, X., K. Kinoshita, and T. Honjo. 2001. Variable deletion and duplication at recombination junction ends: implication for staggered double-strand cleavage in class-switch recombination. *Proc. Natl. Acad. Sci. USA.* 98:13860–13865.
41. Rogozin, I.B., Y.I. Pavlov, K. Bebenek, T. Matsuda, and T.A. Kunkel. 2001. Somatic mutation hotspots correlate with DNA polymerase  $\eta$  error spectrum. *Nat. Immunol.* 2:530–536.
42. Pavlov, Y.I., I.B. Rogozin, A.P. Galkin, A.Y. Aksenova, F. Hanaoka, C. Rada, and T.A. Kunkel. 2002. Correlation of somatic hypermutation specificity and A–T base pair substitution errors by DNA polymerase  $\eta$  during copying of a mouse immunoglobulin  $\kappa$  light chain transgene. *Proc. Natl. Acad. Sci. USA.* 99:9954–9959.
43. 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  $\iota$  and have normal immunoglobulin hypermutation. *J. Exp. Med.* 198:635–643.
44. Mizuta, R., K. Iwai, M. Shigeno, M. Mizuta, T. Uemura, T. Ushiki, and D. Kitamura. 2003. Molecular visualization of immunoglobulin switch region RNA/DNA complex by atomic force microscope. *J. Biol. Chem.* 278:4431–4434.
45. 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.
46. Hun, H., A. Yabuki, and N. Maizels. 2001. A human nuclease specific for G4 DNA. *Proc. Natl. Acad. Sci. USA.* 98:12444–12449.
47. Tashiro, J., K. Kinoshita, and T. Honjo. 2000. Palindromic but not G-rich sequences are targets of class switch recombination. *Int. Immunol.* 13:495–505.
48. Shinkura, R., M. Tian, M. Smith, K. Chua, Y. Fujiwara, and F.W. Alt. 2003. The influence of transcriptional orientation on endogenous switch region function. *Nat. Immunol.* 4:435–441.
49. 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.
50. Nelson, J.R., C.W. Lawrence, and D.C. Hinkle. 1996. Deoxycytidyl transferase activity of yeast REV1 protein. *Nature.* 382:729–731.
51. Lin, W., H. Xin, Y. Zhang, X. Wu, F. Yuan, and Z. Wang. 1999. The human *REV1* gene codes for a DNA template-dependent dCMP transferase. *Nucleic Acids Res.* 27:4468–4475.
52. Sale, J.E., D.M. Calandrini, M. Takata, S. Takeda, and M.S. Neuberger. 2001. Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic hypermutation. *Nature.* 412:921–926.
53. Simpson, L.J., and J.E. Sale. 2003. Rev1 is essential for DNA damage tolerance and non-templated immunoglobulin gene mutation in a vertebrate cell line. *EMBO J.* 22:1654–1664.
54. Guo, C., P.L. Fischhaber, M.J. Luk-Paszyc, Y. Masuda, J. Zhou, K. Kamiya, C. Kisker, and E.C. Friedberg. 2003. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *EMBO J.* 22:6621–6630.
55. Cook, A.J.L., L. Oganessian, P. Harumal, A. Basten, R. Brink, and C.J. Jolly. 2003. Reduced switching in *SCID* B cells is associated with altered somatic mutation of recombined S regions. *J. Immunol.* 171:6556–6564.
56. Faili, A., S. Aoufouchi, S. Weller, F. Vuillier, A. Sary, A. Sarasin, C.-A. Reynaud, and J.-C. Weill. 2004. DNA polymerase  $\eta$  is involved in hypermutation occurring during immunoglobulin class switch recombination. *J. Exp. Med.* 199:265–270.