DNA polymerase eta is involved in hypermutation occurring during immunoglobulin class switch recombination.

Ahmad Faili, Said Aoufouchi, Sandra Weller, Françoise Vuillier, Anne Stary, Alain Sarasin, Claude-Agnès Reynaud, Jean-Claude Weill

To cite this version:


HAL Id: pasteur-00177779

https://hal-pasteur.archives-ouvertes.fr/pasteur-00177779

Submitted on 27 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
DNA Polymerase $\eta$ Is Involved in Hypermutation Occurring during Immunoglobulin Class Switch Recombination

Ahmad Faili,1 Said Aoufouchi,1 Sandra Weller,1 Françoise Vuillier,2 Anne Stary,3 Alain Sarasin,3 Claude-Agnès Reynaud,1 and Jean-Claude Weill1

1INSERM U373, Faculté de Médecine Necker-Enfants Malades, Université Paris V, 75730 Paris, France
2Unité d’Immuno-Hématologie et d’Immunopathologie, Institut Pasteur, 75724 Paris, France
3Étude des Relations Instabilité Génétique et Cancer, CNRS UPR 2169, Institut Gustave Roussy, 94805 Villejuif, France

Abstract

Base substitutions, deletions, and duplications are observed at the immunoglobulin locus in DNA sequences involved in class switch recombination (CSR). These mutations are dependent upon activation-induced cytidine deaminase (AID) and present all the characteristics of the ones observed during V gene somatic hypermutation, implying that they could be generated by the same mutational complex. It has been proposed, based on the V gene mutation pattern of patients with the cancer-prone xeroderma pigmentosum variant (XP-V) syndrome who are deficient in DNA polymerase $\eta$ (pol $\eta$), that this enzyme could be responsible for a large part of the mutations occurring on A/T bases. Here we show, by analyzing switched memory B cells from two XP-V patients, that pol $\eta$ is also an A/T mutator during CSR, in both the switch region of tandem repeats as well as upstream of it, thus suggesting that the same error-prone translesional polymerases are involved, together with AID, in both processes.

Key words: somatic mutation • pol $\eta$ • translesional DNA polymerases • xeroderma pigmentosum variant syndrome

Introduction

Since the discovery of activation-induced cytidine deaminase (AID) and of its central role in the initiation of the three postrearrangement B cell diversification processes, somatic hypermutation (SHM), class switch recombination (CSR), and gene conversion, a major question is to unravel what are the common and the specific partners involved in these different pathways (1, 2). SHM implies the introduction mainly of base substitutions, but also of deletions and duplications into V genes to generate an antigen receptor with better affinity for the immunizing antigen. Translesional DNA polymerases, which are able to bypass specific DNA lesions in an error-free mode at the replication fork but are highly error-prone when copying undamaged DNA, have been suspected to be involved in an error-prone short patch DNA synthesis taking place during SHM (3). It has since been reported, based on a study of SHM in patients with xeroderma pigmentosum variant (XP-V), a cancer-prone disease corresponding to a deficiency in DNA polymerase $\eta$ (pol $\eta$; 4), that this enzyme might be responsible for a large part of the mutations occurring at A/T basepairs (5).

In CSR, the heavy chain $\mu$ constant region is switched to downstream constant region genes that are responsible for various effector functions once the antibody has bound its cognate antigen. The heavy chain $\mu$ locus region that is targeted for CSR, encompasses a large DNA region starting downstream of the I$\mu$ exon up to a region of repetitive pentamer motifs constituting the core $\mu$ switch region (S$\mu$; 6). Base substitutions, deletions, and duplications have been observed around switch junctions (7). More recently, mutations that display many characteristics of SHM, that is, base substitutions on RGYW hotspot motifs (with R for purine, Y for pyrimidines, and W for A or T; 8), bias for transitions and AID dependence have been described downstream of

C.-A. Reynaud and J.-C. Weill are senior authors on this paper.

Address correspondence to Jean-Claude Weill, INSERM U373, Faculté de Médecine Necker-Enfants Malades, Université Paris V, 156 Rue de Vaugirard, 75730 Paris, France. Phone: 33-1-40-61-53-80; Fax: 33-1-40-61-55-90; email: weil@necker.fr; or Claude-Agnès Reynaud, INSERM U373, Faculté de Médecine Necker-Enfants Malades, Université Paris V, 156 Rue de Vaugirard, 75730 Paris, France. Phone: 33-1-40-61-56-84; Fax: 33-1-40-61-55-90; email: reynaud@necker.fr

Abbreviations used in this paper: AID, activation-induced cytidine deaminase; CSR, class switch recombination; pol $\eta$, polymerase $\eta$; TLS, translesion DNA synthesis; S$\mu$, $\mu$ switch region; SHM, somatic hypermutation; S$\mu$, $\mu$ switch region; XP-V, xeroderma pigmentosum variant.

The Journal of Experimental Medicine Volume 199, Number 2, January 19, 2004 265–270
http://www.jem.org/cgi/doi/10.1084/jem.20031831
the \text{I} \mu \text{ exon}, outside the switch core region (9, 10). Here we show, by studying hypermutation at the heavy chain locus in memory B cells of two XP-V patients, that pol η is involved in hypermutation taking place during SHM and CSR, as well, implying that the same mutational complex is probably recruited during these two DNA transactions.

**Materials and Methods**

*Characteristics of XP-V Patients and Preparation of Cell Samples.* Blood samples were obtained from four healthy donors (34, 37, 41, and 50 yr old) and two French XP-V patients (55 and 56 yr old) after informed consent. The two XP-V patients studied were diagnosed clinically and biologically. Patient 1 showed numerous skin tumors since the age of 26. Patient 2 had her first malignant melanoma at 30 and numerous melanomas, basal cell carcinomas, and squamous cell carcinomas since then. Their cultured fibroblasts exhibit normal UV-C–induced unscheduled DNA synthesis and are UV-C sensitive in the presence of caffeine, a hallmark of XP-V cells (11). Heterozygous mutations of the POL H gene were found for both of them: a stop codon at amino acid 303 and a 4-bp deletion at amino acid 408 for patient 1, and a Gly to Arg amino acid substitution at position 295 and a 2-bp deletion (1,727–1,728) at amino acid 576 for patient 2. CD19–CD27+ IgD– memory B cells were purified by cell sorting as previously described (12) and DNA was extracted using proteinase K digestion.

PCR and Sequence Analysis. The \text{J} \mu 4 intronic sequence was amplified using an FR3 consensus primer CACGGCGTGT-HATTACTGTGC and a primer upstream from \text{J} \mu 5, AGGACCCAGGCAAGAAC (at 94°C for 45 s, 55°C for 1 min, and 72°C for 2 min for 40 cycles with Pfui polymerase; Stratagene). The sequence upstream from the \text{S} \mu 1 core was amplified using the following primers: 5′-\text{S} \mu 1, GAATGGATTCCATGCCAAAGC and 3′-\text{S} \mu 1, AGCTGGAATGGTTGTGCAAGTC (at 94°C for 45 s, 56°C for 1 min, and 72°C for 3 min for 35 cycles with Pfui polymerase), and the \text{S} \mu 1–core switch region (\text{S} \mu 1) switch junctions with \text{S} \mu 1-ext, GGGACCTGCTATTTCATTAC and \text{S} \mu 1-ext, CCCATCAGAACCCTAAAGAC (at 94°C for 45 s, 60°C for 1 min, and 72°C for 3 min for 40 cycles using Pfui Turbo polymerase). In this last case, multiple PCRs were performed on aliquots of 2,000–5,000 cells, and PCR products below 1 kb were selected for cloning. PCR products were cloned using the TOPO TA cloning kit (Invitrogen) and sequences were determined on an ABI prism 3100 Analyzer (Applied). All amplifications were performed using Pfui polymerase with a number of cycles to keep background mutations at a minimal level. Mutations induced during amplification by polymerases devoid of proofreading show an A/T over G/C bias that may confuse analysis, notably in the case of switch junction mutations. Error rate of Pfui polymerase is in our assay around $5 \times 10^{-5}$ for 40 cycles of amplification.

**Results and Discussion**

DNA pol η has been proposed as an A/T mutator of Ig V genes (5). This was based on the analysis of the mutation pattern of functional V\text{H}6 genes from three XP-V patients, with characterized mutations of pol η for two of them. However, the question of the exact contribution of pol η to hypermutation has been challenged by Dorner and Lipsky (13). By studying one of these three patients by a single cell PCR approach, designed at amplifying both functional and nonfunctional \text{V} \text{H} genes, they failed to observe such a deficiency at A/T targeting and concluded that it might be a bias introduced by the study of selected V gene sequences (14).

To redress this question, CD27+ IgD–memory B cells from two XP-V patients and from four normal individuals were purified by cell sorting and mutations were analyzed in three noncoding regions of the heavy chain locus: the \text{J} \text{H}4 intron, the \text{S} \mu 1 core upstream region, and \text{S} \mu 1–\text{S} \mu 1 switch junctions (Fig. 1).

The \text{J} \text{H}4 intron downstream from \text{R} \text{H} genes was amplified using a 5′ FR3 primer consensus for all human \text{V} \text{H} genes, and 283 bp starting at the border of \text{J} \text{H}4 intron were sequenced. Only sequences corresponding to different V-D-J junctions were considered. The \text{J} \mu 4 segment was used in 50% of rearranged \text{V} \text{H} genes; such an approach allows a large sampling of selected sequences (Fig. 2 a; reference 15). The mutation frequency was 2.0/100 bp for the controls, and 1.5 for the XP-V patients (Table I). In both groups, mutations bore the hallmark of somatic mutations: prevalence of base substitutions (with 6–9% of total mutations representing deletions or duplications), bias for transitions (47–51% of all mutations), and targeting of RGYW motifs (Table II). This targeting was calculated taking into account mutations occurring at the G position of the hotspot motif only (considered for both strands of DNA), as compared with mutations at other G/C positions. Moreover, normal controls showed a marked preference for mutations at A over T positions. In accordance with Zeng et al. (5), we observed a drastic reduction of mutations on A/T bases in XP-V patients as compared with the controls (10.7 vs. 54.0%; Fig. 2 a), thus confirming the implication of pol η. The lower number of events on A/T bases also correlates with a slight reduction in the total amount of mutations, obviously only noticeable on such unselected targets.

To find out whether the mutations observed during CSR and SHM were generated by the same mutational

---

**Figure 1.** Schematic representation of the regions selected for mutation analysis at the Ig heavy chain locus. The configuration selected represents a \text{V} \text{H} \text{H} gene rearranged to \text{J} \text{H}4 and switched to the \text{S} \mu 1–\text{S} \mu 1 switch region. Horizontal arrows above the locus represent transcription initiation sites. Dotted lines between arrows mark the DNA regions amplified and the bold line represents the DNA regions sequenced. \text{E} \mu , heavy chain enhancer; \text{L} , leader; \text{V} , variable; \text{D} , diversity; and \text{J} , joining coding elements; \text{I} \mu , intronic leader exon; \text{S} \mu 1 and \text{S} \mu 1, core switch regions.
complex, we analyzed two regions of the heavy chain locus that undergo hypermutation during CSR in memory B cells of both groups: the region upstream of the Sμ core region and Sμ–Sδ junctions. AID-dependent mutagenesis of the Sμ core upstream region has been shown to be induced after in vitro induction of switch in mouse splenic B cells (9, 10). Mutations in this region have also been detected in human chronic lymphocytic leukemia (16).

A region of 900 bp upstream from the core repetitive sequences of the Sμ region was amplified from controls and XP-V patients, and the sequence of the 3’ 295 bp was determined (16). A mutation frequency of 0.31/100 and 0.21/100 bp was found for the controls and the XP-V individuals, respectively, i.e., a value six- to sevenfold lower than for the JH4 intron (Table I). It should be noted that both alleles are scored in this assay and that the precise fraction of nonfunctional alleles that undergo isotype switch in memory B cells in vivo is not known. The pattern of mutation was very similar to the one observed at the VDJ locus considering the ratio of base substitutions over deletions and duplications, the bias for transitions and for A over T mutations (although somewhat reduced), and the targeting to G/C bases of the RGYW (WRCY) motifs (Table II and Fig. 2 b). This similarity extends to the selective deficiency of mutations at A and T positions observed for the two XP-V patients in this genomic se-

![Figure 2. Mutation pattern of JH4 intronic sequences and Sμ core upstream regions in normal controls and XP-V patients.](image)

**Table I. Somatic Mutations in JH4 Intronic Sequences and Sμ Core Upstream Sequences from Normal Individuals and XP-V Patients**

<table>
<thead>
<tr>
<th></th>
<th>JH4 intronic sequence (283 bp)</th>
<th>Sμ core upstream sequence (295 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>XP-V</td>
</tr>
<tr>
<td>Number of sequences</td>
<td>73</td>
<td>41</td>
</tr>
<tr>
<td>Total length sequenced (bp)</td>
<td>20,659</td>
<td>11,603</td>
</tr>
<tr>
<td>Unmutated sequences (percent)</td>
<td>2.7</td>
<td>2.4</td>
</tr>
<tr>
<td>Total number of mutations</td>
<td>414</td>
<td>168</td>
</tr>
<tr>
<td>(Number of deletions and duplications)</td>
<td>(23)</td>
<td>(15)</td>
</tr>
<tr>
<td>Mutations range per mutated sequence</td>
<td>1–25</td>
<td>1–19</td>
</tr>
<tr>
<td>Mutation frequency (per 100 bp)</td>
<td>2.0</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Figure 2. Mutation pattern of JH4 intronic sequences and Sμ core upstream regions in normal controls and XP-V patients. (a) Distribution and nucleotide substitution preference of mutations in JH4 intronic sequences. The distribution of mutations is represented along the 283 bp of intronic JH4 sequence, with the proportion of mutated clones in ordinates. Mutations from normal controls are represented above the nucleotide position, and those from XP-V patients are below. Nucleotide substitution preferences are corrected for base composition. A, 18.0%; G, 31.1%; C, 32.2%; T, 18.7%. (b) Distribution and nucleotide substitution preference of mutations in Sμ core upstream sequences. Mutations obtained for normal controls are represented above the 295 nucleotides of the Sμ core upstream sequence, and those from XP-V patients are below. The G position of the RGYW motif (and the C position of the complementary WRCY motif on the other strand) is highlighted in gray. Nucleotide substitution preferences are corrected for base composition. A, 28.5%; G, 28.8%; C: 20.0%; T, 22.7%.
sequence as well (10.2% in XP-V patients compared with 47.4% in controls).

\( \text{S}_{\mu}-\text{S}_{\alpha} \) switch junctions from either \( \text{S}_{\alpha}1 \) or \( \text{S}_{\alpha}2 \) loci were amplified and PCR products below 1 kb were isolated and sequenced. An average of 170 bp on the \( \text{S}_{\mu} \) side and 211 bp on the \( \text{S}_{\alpha}1 \) or \( \text{S}_{\alpha}2 \) side were determined for control sequences, and 200 bp on the \( \text{S}_{\mu} \) side and 317 bp on the \( \text{S}_{\alpha} \) side were determined for the XP-V individuals (Table III). A similar mutation frequency was observed for both control and XP-V patients (0.43/100 and 0.50/100 bp, respectively; Table III). Mutations were mainly base substitutions (deletions over 25 bp were not included and considered as an internal switch event), with a bias for transitions (51.8 and 59.4 for controls and XP-V patients, respectively; Fig. 3). The switch core region, which is made of pentamer repeats, differs from other regions of the heavy chain locus that are targeted for mutations under two aspects: first an extreme incidence of G/C bases within RGYW hotspots, as the switch repeat motif at both \( \text{S}_{\mu}-\text{S}_{\alpha} \) loci GAGCTGGGC includes three G/C bases out of seven in such a context. Accordingly, 79–89% of all G/C mutations are targeted to such hotspot positions. The second specific feature is a high G/C content with a marked G/C imbalance (40% G and 20% C in the reference sequence of the nontranscribed strand). Therefore, mutations have been tabulated in different ways: as raw proportions, thus reflecting their real occurrence in the sequence studied, and after correction for base composition, reflecting then their mechanistic occurrence (Fig. 3). The G/C targeting was higher than in the other regions of the heavy chain locus for the controls, even after correction for the base composition. Nevertheless, there was still a clear diminution of mutations on A/T bp in the XP-V patients with a threefold reduction compared with control values (6.6 vs. 19.6%), a figure similar to the one found for either the \( J_{H4} \) intron or the \( \text{S}_{\mu} \) core upstream region. The moderate intervention of \( \text{pol} \eta \) can probably also explain why the frequency of mutations is approximately similar for the two groups of individuals in this region. We also observed that cytosines were twice as often targeted than guanines.

### Table II. Somatic Mutations at G/C Positions of RGYW (or WRCY) Motifs

<table>
<thead>
<tr>
<th></th>
<th>( J_{H4} ) intronic sequences (283 bp)</th>
<th>( \text{S}_{\mu} ) core upstream sequences (295 bp)</th>
<th>( \text{S}_{\mu} ) core junctional sequences (250 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G/C within RGYW motifs</td>
<td>14%</td>
<td>19%</td>
<td>35%</td>
</tr>
<tr>
<td>Per total G/C bases</td>
<td>(22/157)</td>
<td>(27/144)</td>
<td>(50/142)</td>
</tr>
<tr>
<td>G/C mutated within RGYW motifs</td>
<td>36%</td>
<td>49%</td>
<td>79%</td>
</tr>
<tr>
<td>Per total G/C bases in sequences from normal controls</td>
<td>(89/246)</td>
<td>(45/86)</td>
<td>(57/72)</td>
</tr>
<tr>
<td>G/C mutated within RGYW motifs</td>
<td>31%</td>
<td>57%</td>
<td>89%</td>
</tr>
<tr>
<td>Per total mutated G/C bases in sequences from XP-V patients</td>
<td>(45/147)</td>
<td>(28/49)</td>
<td>(40/45)</td>
</tr>
</tbody>
</table>

\(^a\)Only the underlined G or C base within the motif is considered.

\(^b\)Sequences listed in Fig. 3 are tabulated.

Figure 3. Mutation pattern of \( \text{S}_{\mu}-\text{S}_{\alpha} \) junctions in normal controls and XP-V patients. (a) Mutation distribution along the first 250 bp of the \( \text{S}_{\mu} \) region in \( \text{S}_{\mu}-\text{S}_{\alpha} \) junctions. Mutations collected from normal controls are listed above the reference sequence, and mutations from XP-V patients are below. The G position of the RGYW motif (and the C position of the complementary WRCY motif on the other strand) is highlighted in gray. (b) Nucleotide substitution preference of mutations in \( \text{S}_{\mu}-\text{S}_{\alpha} \) junctions. Mutations collected on both \( \text{S}_{\mu} \) and \( \text{S}_{\alpha} \) regions are tabulated, first as raw proportions and then after correction for base composition, taking into account the base composition of the average region sequenced in each case. \( \text{S}_{\alpha} \) controls: 170 bp, 25.9% A, 30.0% G, 22.3% C, and 21.8% T; \( \text{S}_{\alpha} \) controls: 211 bp, 12.6% A, 47.6% G, 18.6% C, and 21.2% T; \( \text{S}_{\mu}-\text{XP-V} \) : 200 bp, 25.5% A, 31% G, 22% C, and 21.5% T; \( \text{S}_{\alpha}-\text{XP-V} \) : 350 bp, 12.9% A, 46.1% G, 19.6% C, and 21.4% T.
after correction for base composition, but that the absolute number of Cs and Gs mutated was the same. The Ig sterile transcript has been shown to form an R loop with its template strand (17), which would favor the attack of the single stranded nontranscribed strand of DNA by AID according to its in vitro substrate preference (18–22). Assuming that most of the G mutations scored are due to an AID-mediated attack of cytosines on the template strand, it is tempting to propose that the lower accessibility to AID of this strand that forms a DNA–RNA hybrid is compensated by its C richness, whereas the C poorness of the accessible strand is compensated by its preferential targeting. Both factors, AID substrate preference and evolution of target sequences, would ensure approximately the same amount of AID-mediated deamination on each strand, thus allowing for efficient generation of double stranded breaks, necessary for the switch recombination process (23).

In conclusion, we would like to propose that the mutations observed during CSR most probably involve the SHM mutasome. The pattern of mutation is clearly more G/C targeted in the S\(\mu\)/H9262 core region. This G/C targeting might be explained by the G/C richness within pentameric hotspots in this region and the unique accessibility provided by the R loop structure, so that less nuclease digestion, and/or less error-prone short patch DNA synthesis would be required to generate adjacent cleavage sites, i.e., double stranded breaks that can be engaged into a successful switch recombination event. A similar conclusion was drawn from a study on the role of MSH2 during CSR (24) showing that this molecule is more specifically required for switching occurring in the S\(\mu\)/H9262 upstream region. In this region, the lower density of hotspots, as compared with the S\(\mu\) core, may require more processing of the DNA ends to render them suitable for recombination, a processing that would rely on the strand displacement activity of MSH2.

This work also brings the confirmation that the translesion DNA synthesis (TLS) polymerase pol \(\eta\) is an A/T mutator. Nevertheless, there are still mutations on A/T bases in both SHM and CSR, TLS polymerases, such as pol \(\eta\) that interacts with pol \(\eta\) at replication foci, could be involved in both processes (25). Accordingly, in a human Burkitt’s B cell line inducible for SHM and whose pattern of mutations is strongly G/C biased, induction of SHM was shown to be dependent on pol \(\eta\) (26, 27). This result remains controversial because it was recently reported that 129/SvJ mice that lack a functional pol \(\eta\) did not show any significant change in the frequency and pattern of Ig mutations (28). To what extent the human in vitro and mice in vivo systems differ and to what extent TLS polymerases may compensate for each other remains to be explored.

The rationale for using error-prone polymerases during CSR is not clearly perceptible. Here we propose that these enzymes participate in an AID-triggered, short patch DNA synthesis ensuring the efficient fill-in of staggered double strand breaks. However, we cannot exclude that the mutations observed may represent the bystander recruitment of AID partners involved in the mutagenesis of Ig V genes, and that the involvement of TLS polymerases as cofactors of the nonhomologous end joining process might be dispensable. An in vitro system that would enable to assess the impact of these polymerases on the strict efficiency of switching may help resolve this issue.

We thank Dr. M.F. Avril for providing blood samples from XP-V patients, C. Debacker for cell cultures, A. Van der Linden, and the laboratory of Alan Lehmann for the determination of pol \(\eta\) mutations in XP-V patients 1 and 2, respectively. We thank Jérome Mégret for performing cell sorting, Claire Soudais and Claire Fieschi for their contribution, and Guillaume Dighiero for stimulating discussions.

This work was supported by the Ligue contre le Cancer (Equipe Labelisée) and the Fondation Princesse Grace de Monaco.

Submitted: 21 October 2003
Accepted: 24 November 2003

References


