An immunoglobulin mutator that targets G-C base pairs
(transfected target sequence/termination codon/reversion/lacZ)

JUERGEN BACHL AND MATTHIAS WABL

Department of Microbiology and Immunology, University of California, San Francisco, CA 94143-0670

Communicated by Philippa Marrack, National Jewish Center for Immunology and Respiratory Medicine, Denver, CO, September 26, 1995

ABSTRACT Hypermutation can be defined as an enhancement of the spontaneous mutation rate which the organism uses in certain types of differentiated cells where a high mutation rate is advantageous. At the immunoglobulin loci this process increases the mutation rate >10^6-fold over the normal, spontaneous rate. Its proximate cause is called the immunoglobulin mutator system. The most important function of this system is to improve antibody affinity in an ongoing response; it is turned on and off during the differentiation of B lymphocytes. We have established an in vitro system to study hypermutation by transfecting a rearranged μ gene into a cell line in which an immunoglobulin mutator has been demonstrated. A construct containing the μ gene and the 3′ χ enhancer has all the cis-acting elements necessary for hypermutation of the endogenous gene segments encoding the variable region. The activity of the mutator does not seem to depend strongly on the position of the transfected gene in the genome. The mutator is not active in transformed cells of a later differentiation stage. It is also not active on a transfected lacZ gene. These results are consistent with the specificity of the mutator system being maintained and make it possible to delineate cis and trans mutator elements in vitro. Surprisingly, the mutator preferentially targets G-C base pairs. Two hypotheses are discussed: (i) the immunoglobulin mutator system in mammals consists of several mutators, of which the mutator described here is only one; or (ii) the primary specificity of the system is biased toward mutation of G-C base pairs, but this specificity is obscured by antigenic selection.

Although mutations are, in general, detrimental to the individual, they are the substrate upon which evolution works. The intrinsic fidelity of the enzymes involved in DNA replication is not very high, and to reduce the number of mutations, elaborate error-checking and repair systems have evolved in mammals. But mutations in immunoglobulin genes add to antibody diversity and help the immune system to cope with the large diversity of threatening microorganisms, which is itself mutational in origin.

Over 35 years ago it was suggested that somatic mutations were the major source of antibody diversity (1). Later, Brenner and Milstein (2) proposed that an enzyme complex would actively place mutations and thus cause hypermutation at the immunoglobulin locus. The Brenner–Milstein model bears an uncanny resemblance to the way in which untemplated nucleotides, so-called N regions, are added at the DJ and VD junctions during immunoglobulin heavy-chain gene rearrangement (3), a process that was discovered almost 20 years after the model was proposed. The early work of Weigert et al. (4) showed that somatic point mutations indeed make a major contribution to the diversity of the λ light chain in the mouse. Although the tone had already been set by those studies on myelomas, it came as somewhat of a surprise that the genes encoding the immunoglobulin chains expressed in the second-ary immune response carried so many mutations of somatic origin (5–17). Because of the high frequency, the process that produced these somatic point mutations earned the epithet “hypermutation.” If hypermutation is to contribute to affinity maturation, it must be active during B-cell proliferation after antigenic stimulation. Indeed, cells undergoing hypermutation are found in the germinal centers, where the B-cell response in large part takes place (18, 19).

Recently, some important advances have been made in understanding the mechanism of hypermutation. Because mutator action is restricted to a region flanked by the V gene leader intron and intron enhancer (20–24), there ought to be one or more recognition sequences, which are presumably in or around the rearranged VDJ or VJ sequences. However, the recognition sequence does not seem to be contained in the 5′ portion containing the promoter, and both the κ intron and κ 3′ enhancer regions were found to be essential for full hypermutation at the κ light-chain locus (25, 26). Thus, there may well be more than one recognition sequence. Somatic mutation has also been linked to the direction of DNA replication (27).

Even though the final goal is description of hypermutation in the whole animal, the study of the mutator components in transgenic animals is cumbersome. Thus, mutations at the immunoglobulin loci were also studied early on in lymphocyte lines (28–34). We have defined hypermutation in vitro in the murine pre-B cell line 18-81 (refs. 35, 36). The proximate cause for hypermutation, the putative immunoglobulin mutator system (mutator), increases the mutation rate at the gene segments encoding the endogenous V region by a factor of at least 10^3. The mutator does not work efficaciously at the Cκ gene segment (37) nor at the B2m locus (38). It does not seem to be active at the plasma cell stage (35), which represents the final stage in the differentiation of a B lymphocyte.

We now report a system to study hypermutation rapidly and effectively by means of transfection in vitro.

MATERIALS AND METHODS

Plasmid Constructs. Plasmids phyp#1 and phyp#2 were generated by exchanging the original variable region of the plasmid pmu (39) with the rearranged V2 variable region of the cell line 18-81. The neomycin-resistance (neo') gene was cloned as a Sal I–Xho I cassette 5′ to the variable region. A Sac I–Xba I fragment containing the 3′ χ enhancer (40, 41) was cloned into the Xho I site 3′ of the μ membrane exons. For stable transfection the construct was linearized at the Sal I site adjacent to the neo' gene. To generate plasmid ptk-LacZ-Stop the thymidine kinase (tk) gene promoter from plasmid pMCIneo (42) was cloned as a Xho I–Pst I fragment together with a Sal I–Xho I fragment of the lacZ gene into the BS/KS vector (Stratagene). A termination codon had been introduced before into the lacZ gene by oligonucleotide site-directed mutagenesis at a position 161 nt 3′ to the Kpn I site. The neo' gene was cloned as a Xho I–Sal I fragment into the Spe I site

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside.
of BS/KS-tk-LacZ-Stop. For stable transfection the ptk-LacZ-Stop plasmid was linearized with Xho I.

Detection and Quantitation of Mutants. Plasmids phyp#1 and phyp#2 were transfected into clone 18-81 ΔM (43), and transfectants were subcloned under limiting dilution conditions. The clones were grown to 10^6 cells and μ-producing revertants were counted per 10^5 cells by immunofluorescence. For the clones transfected with lacZ containing the termination codon the cells were expanded in a 24-well plate to 1–2 × 10^6 cells per well, and all cells in a well were analyzed for β-galactosidase activity.

For immunofluorescence, 10^5 cells were spun onto a slide, ethanol fixed, and rehydrated in phosphate-buffered saline (PBS) with 1% bovine serum albumin; cells were stained with fluorescein isothiocyanate-conjugated goat anti-mouse IgM antibody (Fisher) and washed three times for 10 min in PBS.

β-Galactosidase activity was monitored with either the chromogenic substrate 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal) or the fluorogenic substrate fluorescein di-β-D-galactopyranoside. For staining with X-Gal, the cells were fixed for 5 min at room temperature in the culture wells or on slides with PBS containing 2% formaldehyde and 0.2% glutaraldehyde. After washing in PBS, the cells were stained with PBS containing 5 mM potassium ferrocyanide, 5 mM potassium ferrocyanide, 2 mM magnesium chloride, and X-Gal at a final concentration of 1 mg/ml. With this procedure, the cytoplasm of positive cells is stained bright blue within 1–24 hr at 37°C. To verify the nature of a β-galactosidase-expressing cell, one transfection clone was selected for enrichment and purification of its revertant cells by fluorescence-activated cell sorting (44) and subsequent subcloning. The revertant tested was missing the Bcl I restriction site that we had introduced and of which the opal codon (TGA) was part.

Cloning and Sequencing of Mutants. From six independent transfection clones (of 18-81ΔM, an 18-81 subclone with no endogenous Cμ) some revertants and some of their μ-negative sister cells were cloned and the sequence of their D region was determined. Clones containing construct phyp#1 or phyp#2 with 25–500 revertants per 10^5 cells were first sb-selected for enrichment of revertants and subsequently subcloned by seeding 30 cells per 96-well plate. Pure revertant clones were expanded and total RNA was isolated. cDNA was synthesized by using an oligo(dT) primer ([dT]17). Then DNA was amplified by using a specific primer pair that covered sequences specific to V81X (GAAGAGGCCTGGATGGTGGCAG) and Cμ4 (CACACTGATGTCTGCAGGAGAAGC), respectively. The PCR product was agarose gel-purified and directly sequenced by using the V81X primer. PCR procedures and sequencing were done on coded samples—i.e., without knowledge of to which clone they belonged, to one of the 7 revertants or to one of the 6 sister cells containing the TAG codon. In all cases the codon in position 101 corresponded to the status of μ-chain expression; i.e., the clones with TAC or TAT turned out to be μ-chain expressors, whereas the clones with TAG did not express μ.

RESULTS AND DISCUSSION

We constructed two plasmids for transfection into cell lines and thus probing their mutability. Into plasmid phyp#1 and phyp#2 we cloned the silent VDJ allele of the mutator-active cell line 18-81 and ligated it to the Cμ gene segment of the plasmid pμ (39); we also introduced neo' as a selectable marker. One copy of the 3' κ enhancer was cloned into phyp#1, while phyp#2 contains two tandem copies oriented in the opposite direction. As a result of an amber termination codon (TAG) in the D segment, translation is terminated prematurely, and no μ sequences are expressed. As a control, we also constructed a selectable plasmid containing a lacZ gene with an opal termination codon (TGA) introduced by site-directed mutagenesis. The structures of these plasmids are diagrammed in Fig. 1.

The phyp plasmids (and the control ptk-LacZ-Stop) were transfected into 18-81ΔM, a subclone of the 18-81 cell line that had lost both endogenous Cμ alleles (43). Any μ chain synthesized must, therefore, have been encoded by the transfectected μ gene. Indeed, upon reversion of the termination codon, full-length μ chain is expressed and can be detected by immunofluorescence (Fig. 2A). Reversion or in-frame deletion of the termination codon in the control transfectants results in expression of the enzyme β-galactosidase and conversion of the colorless X-Gal substrate into a blue dye by the transfectant (Fig. 2B).

Hypermutation of the Transfected μ Genes. When most transfectants were expanded from one or a few cells up to 10^6 cells, the revertants constituted more than 1 in 10^6 cells; in 20–30% of the cases, however, there were no revertants at all. When transfected with a functional μ gene construct, about 30% of cells resistant to the selection marker did not express μ. This presumably explains the fact that mutants are common in most clones, but absent in some.

Six independent transfectants with phyp plasmids into clone 18-81ΔM were subcloned. These clones contained at most one copy of the phyp plasmid, as assayed on Southern blots (data not shown). From these we determined the cumulative distribution of reversion frequencies in cultures of two independent clones—phyp#2 transfectant 6-6.2 and phyp#1 transfectant 11-9.3 (Fig. 3). The median reversion frequencies were 1.0 × 10^{-4} and 3.1 × 10^{-4}, respectively. It is not clear whether the difference is due to (i) fluctuation effects, (ii) the chromosomal location of the integrated plasmid, (iii) phyp#1 and phyp#2 structure, or (iv) instability of the transfected clones. If the last mentioned possibility were true, we should ignore the cultures with zero mutants, in which case the medians are even closer—2.3 × 10^{-4} and 3.4 × 10^{-4}. In any event, the difference is so small that neither chromosomal location nor orientation of the 3' κ enhancer can have a large effect on the frequency of mutants.

Fig. 3 also gives the cumulative distribution of reversion frequencies for lacZ in the control transfectants. The median revertant frequency is below 1 × 10^{-6}, at least 2 orders of magnitude lower than that for μ.
magnitude less than that of the \( \mu \) transfectants. Even if we ignore the cultures with zero transfectants (and there is no reason to do so), the median is \( 3 \times 10^{-6} \), which is 1.5 orders of magnitude less than that of the \( \mu \) transfectants. Thus, the mutator we are assaying does not work on just any gene. The data are consistent with it being specific for the VDJ segments of immunoglobulin genes, as was the mutator acting on the endogenous heavy-chain gene in 18-81 (35–38). That is, in both cases we are dealing with an immunoglobulin mutator.

There are several hints that hypermutation in the animal depends on transcription (25, 27, 46). \( \text{lacZ} \), with a \( \text{tk} \) promoter, may have fewer transcripts than VDJ, with the immunoglobulin promoter and enhancers. If this were the case the \( \text{lacZ} \) gene would be disadvantaged as a target for hypermutation. However, replacement of the immunoglobulin promoter by the \( \text{tk} \) promoter in the VDJ constructs did not decrease, but increased gene expression; it did not lower hypermutability (unpublished work).

All revertants are due to point mutations. To determine whether the \( \mu \)-producing variants are true revertants rather than small deletions that excise the chain-termination codon, we cloned several independent revertants by sib selection and then sequenced them. Fig. 4 shows the genealogy of the revertants and their sister cells as well as the codon at position 101 (45), which in the transfected construct was TAG. All seven \( \mu \)-producing subclones were true revertants, whereas their nonproducing sister clones retained the TAG termination codons. As the revertant frequency was at least as high as in an endogenous variable region in 18-81 (35), we conclude that the mutator was able to act with full efficacy on the transfected construct, which, except for the \( \kappa \) enhancer and plasmid sequences, contained only the 5' and 3' sequences necessary for \( \mu \)-chain expression.

Mutator inactive in a myeloma. It has been previously reported that immunoglobulin genes are not hypermutable in hybridomas (32–35). Wabi et al. (35) also reported that when 18-81 was fused to the mouse plasmacytoma Ag8.653, the resulting hybridoma showed no hypermutation. To find out whether our transfected construct would be a target for hypermutation in such cells, we transfected phyp\#2 into Ag8.653, which contains no functional immunoglobulin genes. Although \( \mu \) transcripts could be detected (data not shown), we found no revertants in any of 15 transfectants; the frequency of revertants could not have been as high as \( 10^{-7} \). When similar plasmids with no internal termination codon were transfected

**FIG. 2.** Detection and isolation of mutants. (A) Subclone 18-81/\( \Delta \)M of the mutator-positive cell line 18-81 was transfected with plasmid phyp as depicted in Fig. 1, and stained with fluorescence-labeled antibody to \( \mu \) chain. (Upper) A \( \mu \)-producing clone isolated by sib selection and subcloning under limiting dilution conditions. (Lower) A \( \mu \)-producing clone isolated by fluorescence-activated cell sorting and subcloned under limiting dilution conditions.

**FIG. 3.** Cumulative distribution of reversion frequencies in cultures of the two independent clones 6-6.2 (transfected with phyp\#2) and 11-9.3 (phyp\#1), and clone \( \text{lacZ} \) (transfected with ptk-LacZ-Stop). The \( y \) axis shows cumulative fraction of cultures, in percent on a linear scale. The \( x \) axis shows frequency of mutants (\( \mu \) producers and \( \beta \)-galactosidase producers, respectively) in a given culture, on a logarithmic scale.
into Ag8.653, the μ chain was detected by immunofluorescence (data not shown). Apparently the immunoglobulin mutator is turned off in these cells. These findings confirm that the VDJ sequences we have been studying do not have an intrinsically high mutability—they mutate because of the action of a mutator in the cell line.

**Mutator Preferentially Targets G-C Base Pairs.** An intriguing observation in the sequencing studies was that seven out of seven revertants contained TAC or TAT in place of the TAG termination codon. That is, only the G-C base pair was affected. In the course of the present work, we also sequenced three more independent mutations of the endogenous μ-chain gene in the 18-81 cell line; in all three, TAC had been replaced by TAC. Previous work showed that all of nine 18-81 mutations sequenced involved a G-C base pair (36); these included four mutations at sites other than the chain termination codon. Thus altogether 19 mutations have been sequenced, and in all cases a G-C base pair was involved. Each mutation arose in a separate culture well that had been seeded with a single nonrevertant cell; therefore we can be sure that they represent independent mutation events. Thus we conclude that the immunoglobulin mutator in 18-81 affects G-C pairs. One might argue that Tyr, which is encoded by TAT and TAC, can be detected in our assay, while Leu, Ser, or Trp, which would result from a change at the middle base, and Gin, Lys, or Glu, which would result from a change at the first base, all might not be. The counterargument is that the revertants are detected by reactivity with an anti-Cu antibody, an interaction that ought not to be influenced by the sequence of VDJ; moreover, there is no light chain that might select for VDJ pairing capability. Thus we consider it unlikely that mutations at only the G-C base pair would be detected.

**G-C Base-Pair Preference in Frogs.** In mouse hybridoma sequences, there is little overall bias for mutations of G-C base pairs (5–17). In view of this fact, the simplest interpretation of the G-C bias in 18-81 would be that several mutators are effective at the immunoglobulin loci, of which only one, the G-C mutator, is active in cell line 18-81. However, it has been reported that in the clawed frog *Xenopus*, >90% of all somatic point mutations at the immunoglobulin heavy-chain locus affect G-C base pairs (47). While the above interpretation is also possible here, the authors of the *Xenopus* study suggest a second interpretation. They postulate that the primary action of the mutator is biased toward G-C and against T-A in mice as well as in frogs, but that strong antigen selection for mutants with better affinity obscures the original spectrum of mutants in mice. Affinity maturation, which is driven by antigen selection, is poor in frogs, perhaps because they lack germinal centers, and, as a consequence, the high percentage of mutations at G-C base pairs is preserved. The same preference has been observed in another lower vertebrate, the shark (48). Because there is clearly no antigen selection in the 18-81 cell line, our finding would be consistent with the notions of the *Xenopus* study.

**G-C Base-Pair Preference in Mice and Humans.** In revisiting the issue in mouse and humans we turned to mutational hotspots. The primary data for Table 1 were taken from several papers on mutation at the heavy- and light-chain loci of mice and humans (46, 50–53). These data were tabulated by Betz et al. (49), who classified mutational hotspots into “selected” and “intrinsic”—i.e., into those that seemed to have been selected by antigen and those that seemed to have not been selected by antigen. Our Table 1 is in fact a summary of the results given in their table 2. While there is no obvious preference for G-C base pairs among the selected hotspots, 12 of the 14 unselected hotspots, with 111 of the 123 mutations, are at G-C base pairs. This striking finding is in line with the second interpretation given in the previous paragraph.

Thus, all of the data derived from mutations in VDJ coding sequences seem to be consistent with the notion that the immunoglobulin mutator system acts preferentially on G-C base pairs and that this preference is obscured by antigen selection. What is not in agreement with this notion are data derived from mutations in flanking regions (20–23, 54). We have no explanation for this discrepancy, except the vague idea that something might be wrong with the commonsense assumption that mutations in flanking regions cannot be subject to selection.

Wilson et al. (47) also presented an interesting argument, too complex to be restated here, that the fraction of mutations at G-C base pairs was highest in those experimental systems in which selection was the lowest. The systems considered were all mammalian and none of them involved genes of the immune system. And there are many claims in the literature (e.g., ref. 55; reviewed in ref. 56) that an oxidized form of guanine in DNA is the principal source of spontaneous mutations. Perhaps the amount of oxidized guanine in DNA can be experimentally manipulated. Of course, the fact that G-C base pairs are preferred does not necessarily mean that all G-C pairs are targeted with equal likelihood and no other sequence requirements are necessary. Thus, the sequence motifs RGYW and TAA (57) and CAGCT and AAGTT have been noted. At the λ locus of the mouse, hotspots seem to be contained in palindromes; no consensus sequence is immediately obvious, but all 11 hot base pairs reported are at a G-C base pair (24). The TAG codon in our studies is not contained in the sequence CAGCT, but in RGYW.

While this manuscript was in preparation, a paper from Scharff’s laboratory reported an approach similar to the one

---

**Table 1.** Selected and intrinsic hotspots in human and mouse immunoglobulin V gene segments (adapted from ref. 49)

<table>
<thead>
<tr>
<th>From</th>
<th>Selected</th>
<th>Intrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of hotspots</td>
<td>No. of mutations</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>T</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>G</td>
<td>3</td>
<td>32</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>60</td>
</tr>
<tr>
<td>All</td>
<td>9</td>
<td>164</td>
</tr>
</tbody>
</table>
reported here, without finding a bias for G-C base pairs (58). There are, however, significant differences in characteristics between their construct and our construct: Their construct contained a complete μ gene, but no 3′ κ enhancer. When the 3′ κ enhancer was left out of our construct, reversion frequencies dropped by 2 orders of magnitude (unpublished work). Furthermore, our construct did not hypermutate in a myeloma line, whereas their results were obtained in a myeloma line. And finally, even though the results cannot be compared directly in quantitative terms, the effect on our hypermutable construct was at least 2 orders of magnitude greater than that on our control construct, whereas the difference between V and C in the study from Scharff's lab was about 5-fold.

Conclusions. From our studies we draw the following conclusions: (i) A construct containing the μ gene and the κ enhancer has all elements necessary to recapitulate hypermutation as it occurs in the endogenous μ gene. (ii) Mutator action does not seem to depend strongly on the position of the transected gene in the genome. (iii) The mutator in 18-81 preferentially acts on VDJ gene segments; i.e., it is an immunoglobulin mutator. (iv) The mutator is turned off at later differentiation stages. (v) The mutator in 18-81 preferentially acts on G+C base pairs, which may need to be embedded in special sequence motifs.

We thank C. Steinberg for the suggestion to present the reversion frequency data as in Fig. 4 and for editing the manuscript. We thank K. Meyer and M. Neuberger for the 3′ κ enhancer. This work was supported by National Institutes of Health Grant 1R01 GM37699 and by funds from the Markey Trust to M.W. and by a Boehringer Ingelheim Fonds grant to J.B.