The strand bias paradox of somatic hypermutation at immunoglobulin loci

Andrew Franklin¹ and Robert V. Blanden²

1 The Immune Disease Institute, The Children’s Hospital, and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA
2 Division of Immunology and Genetics, The John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia

Somatic hypermutation has two phases: phase 1 affects cytosine-guanine (C/G) pairs and is triggered by the deamination of cytosine residues in DNA to uracil; phase 2 affects mostly adenine-thymine (A/T) pairs and is induced by the detection of uracil lesions in DNA. It is not known how, at V(D)J genes in mice, hypermutations accumulate at A/T pairs with strand bias without perturbing the strand unbiased accumulation of hypermutations at C/G pairs. Additionally, it is not known why, in contrast, at switch regions in mice, both C/G-targeted and A/T-targeted hypermutations accumulate in a strand unbiased manner. To explain the strand bias paradox, we propose that phase 1 and phase 2 hypermutations are generated at different stages of the cell cycle.

Somatic hypermutation (SHM) sequentially introduces single nucleotide substitutions, and occasionally insertions and deletions, at a high rate into rearranged V(D)J genes and switch (S) regions at the immunoglobulin loci (Figure 1; Box 1) [1–6]. The molecular mechanisms of SHM at V(D)J genes and S regions have yet to be entirely clarified. As anticipated earlier based on genetic evidence [7], SHM has two phases: phase 1 affects cytosine-guanine (C/G) pairs and phase 2 affects mostly adenine-thymine (A/T) pairs. Phase 1 of SHM is triggered by the deamination of cytosine residues in DNA to uracil by activation-induced deaminase (AID). Phase 2 of SHM is attributed to low-fidelity DNA-dependent DNA synthesis during mutagenic patch repair that is induced by detection of AID-generated uracil-guanine (U/G) mispairs.

How these phase 1 and phase 2 mechanisms integrate to produce the spectrum of hypermutations that accumulate at C/G pairs and A/T pairs in vivo has hitherto lacked an explicit description. Specifically, although the two-phase scheme of SHM has not been tightly defined since it was initially proposed, a linear interpretation of this model is still the most widely accepted explanation for the observed spectrum of hypermutations that accumulate at C/G pairs and A/T pairs during SHM at V(D)J genes and S regions.

Mechanics of hypermutagenesis

AID is essential for the induction of SHM at V(D)J genes in mice [9] and humans [10]. SHM at the S region located 5’ of the μ constant region exons (Sμ) has also been shown to be AID-dependent [2,3]. The ‘DNA deamination model’ [11] envisages that AID triggers SHM by deaminating cytosine residues in DNA to uracil. In vitro biochemical studies have demonstrated that AID preferentially deaminates C residues in single-stranded DNA substrates [12–14]. SHM correlates with transcription [15] so AID might access the TS and the NTS as single-stranded DNA substrates during transcription of the double-stranded DNA target.

Uracil residues that occur in DNA because of cytosine deamination are subject to removal from the deoxyribose-phosphate chain by a uracil-DNA glycosylase (UDG) during the initial stages of excision repair [16].

The DNA deamination model explains how most of the hypermutations at C/G pairs are produced in vivo: dATP incorporation opposite AID-generated uracil lesions generates transitions at C/G pairs; dNTP incorporation opposite non-instructional apyrimidinic (AP) sites that form because of UDG excision of AID-generated uracil lesions generates transversions and transitions at C/G pairs. This model is strongly supported by genetic data that mice and humans deficient in UDG activity encoded by the uracil N-glycosylase (Ung) gene display a substantial increase in transition substitutions at C/G pairs and a concomitant decrease in transversion substitutions at C/G pairs [17,18].

It has been shown since that UNG2, the nuclear isoform of...
SHM occurs at the assembled V(D)J genes of antigen-activated germinal centre B lymphocytes called centroblasts. At the SHM also occurs in the vicinity of S regions and S

Box 1. Physiological significance of somatic hypermutation.

The variable (V) region of an immunoglobulin (Ig) molecule directly binds antigen and is epitope specific. The somatic preimmune repertoire consists of an extremely large number of diverse Ig molecules. However, the Ig molecules in the preimmune repertoire that are selected by an immunizing antigen typically exhibit low binding affinity. The average binding affinity of Ig molecules recognizing a target antigen increases during the course of an immune response. This ‘affinity maturation’ of the humoral immune response results after (i) antigen-driven somatic hypermutation (SHM) diversifies the V(D)J genes that encode IgV regions (Figure 1) and (ii) the immunizing antigen positively selects mutants that express cognate Ig with increased binding affinity. Therefore, affinity maturation increases the efficiency of the humoral immune response to a target antigen and hence improves immunity against infection.

The heavy chain constant (CH) region of an Ig molecule determines its biological effector functions. Mice and humans produce five CH classes (IgM, IgD, IgG, IgE and IgA encoded by CHμ, CHγ, CHε, and CHδ genes), each of which is specialized in determining how an Ig molecule eliminates antigen. The VDJ exons that encode VH regions are initially expressed in association with Cμ (or Cα). Antigen-driven class switch recombination (CSR) enables Cμ (and Cδ) to be replaced by Cγ, Cε or Cλ. Therefore, class switching alters the biological effector functions of the species of Ig molecule that a B lymphocyte produces without altering antigenic specificity. Class switch recombination (CSR) adjoins the VDJ exon to one of the downstream sets of CH exons (thereby replacing the Cμ and Cδ exons) in activated, cytokine-stimulated B lymphocytes. Switch (S) regions are located 5’ of each set of CH exons (except Cδ). Each set of CH exons is a transcription unit with a cytokine-inducible promoter located 5’ of its associated S region; each promoter is inducible by specific cytokines. The initiation of germline transcription at a set of Cμ exons correlates with the CSR potential of the associated S region. CSR occurs through the introduction of double-strand breaks into the donor Sμ and 3’ acceptor S region (e.g. S3/4) followed by intrachromosomal joining of the broken S regions and deletion of the intervening sequence (not shown).

SHM may occur at the donor Sμ and 3’ acceptor S regions both before and during switching at the V(D)J-Cμ-3’ junctions after switching [6]. Black ovals with arrows indicate transcriptional promoters. EIμ, intronic enhancer (heavy chain locus); I, intronic leader exon; L, leader exon.

UNG [19], is the only UDG that normally excises AID-generated uracil residues during SHM at V(D)J genes and S regions in vivo [4]. UNG2 preferentially removes uracil residues from single-stranded DNA substrates so, if UNG2 efficiently accesses the TS and the NTS as single-stranded DNA substrates, it is possible that the activities of AID and UNG2 are orchestrated during transcription of the double-stranded DNA target [20].

The initiation of A/T-targeted hypermutagenesis at V(D)J genes and S regions needs either the excision of AID-generated uracil lesions by UNG2 or the recognition of AID-generated U/G mismaps by MSH2–MSH6, the MutS homologs that form a heterodimeric protein complex that identifies postreplicative base–base mispairs and initiates mismatch repair [4,21]. Hypermutations at A/T pairs are thought to be introduced during a type of mutagenic DNA-dependent patch repair of the initiating AID-generated U/G lesion. In this context, the mammalian AP endonuclease (APE1), which catalyses the incision of phosphodiester linkages 5’ of AP sites [16], might facilitate the initiation of mutagenic patch repair in the UNG2-mediated pathway. Furthermore, exonuclease 1 (EXO1), the only exonuclease that is thus far known to be needed for mismatch repair [22], might facilitate mutagenic patch repair in the MSH2–MSH6-mediated pathway; the incidence of A/T-targeted hypermutations at V(D)J genes is substantially diminished in EXO1-deficient mice [23].

Previous evidence indicated that DNA polymerase-η (pol-η) is substantially involved in the mechanism of A/T-targeted hypermutagenesis at V(D)J genes and S regions

![Figure 1](https://example.com/figure1.png)
in humans [1,6,24] and mice [25,26]. However, recent genetic evidence strongly suggests that pol-δ is the only DNA polymerase that generates hypermutations at A/T pairs at V(D)J genes in the normal physiological context [27]. Additionally, the error spectrum that pol-δ generates at A/T pairs, but not at C/G pairs, during DNA-dependent DNA synthesis in vitro correlates with the mutation spectrum at A/T pairs during SHM at V(D)J genes in vivo [28,29]. Therefore, the proposition that pol-δ introduces hypermutations at A/T pairs during DNA-dependent patch repair of the initiating AID-generated U/G lesion is favored within the SHM research community.

Hypermutagenesis at V(D)J genes

SHM at V(D)J genes occurs in antigen-activated germinal centre B lymphocytes called centroblasts. Centroblasts are large, surface immunoglobulin-negative B lymphocytes that occupy the germinal centre dark zone and divide rapidly with a cell cycle time of 6–7 h [30–32]. Centroblasts undergo a finite number of divisions, tentatively approximated to be three to four cell cycles [33], to generate progeny that come out of cell cycle and differentiate into centrocytes that re-express surface immunoglobulin and migrate through the germinal centre light zone where they undergo antigen-based selection, competing to bind the immunizing antigen [31,34]. Centrocytes undergo a high physiological cell death rate in vivo [35,36], via apoptosis [34], within a few hours of entering the light zone. Apoptosis occurs most frequently in the light zone but also occurs at high levels in the dark zone [36]; the molecular machinery that contributes to spontaneous apoptosis is turned on within centroblasts [37]. Evidence suggests that SHM generates mismatches during the G₁ phase of the cell cycle [38].

Hypermutations accumulate at C/G pairs and A/T pairs with approximately equal frequency at V(D)J genes in mice and humans. However, nucleotide substitutions accumulate at C and G residues with approximately equal frequency on the TS and the NTS, whereas, with respect to the NTS, nucleotide substitutions accumulate at A residues two times more frequently than at T residues. That is, at V(D)J genes, the targeting of C/G pairs is strand unbiased and the targeting of A/T pairs is strand biased [39]. The correlation of the in vivo pol-δ error spectrum with the in vitro SHM spectrum at A/T pairs specifically suggested that pol-δ performs mutagenic patch repair synthesis of the NTS twice as frequently as it does the TS during SHM in vivo to establish the strand bias that affects A/T pairs [28,29]. As of yet, it has been difficult to envisage how pol-δ generates hypermutations at A/T pairs in a substantially strand biased manner at V(D)J genes if AID generates the initiating lesion on the TS and NTS with approximately equal frequency [4] and C/G-targeted hypermutations accumulate in a substantially strand unbiased manner.

The observed SHM spectra at C/G pairs and A/T pairs at V(D)J genes suggests to us that the AID-generated U/G lesions that lead to phase 1 hypermutagenesis are inaccessible to the machinery that leads to phase 2 hypermutagenesis. We currently propose that, at V(D)J genes, the lesions that lead to so-called phase 1 hypermutagenesis accumulate mostly after the mechanism of phase 2 hypermutagenesis has operated during the G₁ phase of the centroblast cell cycle (Figure 2).

This might be achieved if the MSH2–MSH6–mediated pathway, the major pathway by which phase 2 of SHM is initiated at V(D)J genes [4,7,18], operates mostly during the G₁ phase of the centroblast cell cycle. Although the mismatch repair machinery is constitutively expressed (at least in normal cells) [40,41], the protein levels of components involved in coordinating events during mismatch repair are cell cycle regulated to some extent, increasing specifically in late G₁ and S phase [40]. Therefore, it is possible that at least one factor of the mismatch repair apparatus is limiting during most of G₁ phase in normal cells. We suspect that none of the components of the mismatch repair machinery are limiting in centroblasts during G₁ phase, possibly because the centroblast cell cycle time is very short. In this scenario, the processing of AID-generated U/G mispairs in the MSH2–MSH6–mediated pathway might reflect residual mismatch repair activity from the previous episode of replication. Genetic experiments in which components of the mismatch repair machinery become limiting before G₁ entry in centroblasts might provide supporting evidence for the proposed two-phase scheme of SHM at V(D)J genes.

The excision of AID-generated uracil lesions by UNG2 is the minor pathway by which phase 2 of SHM is initiated at V(D)J genes [4,18], possibly because UNG2 expression is strictly regulated during the cell cycle [42]. UNG expression starts at the G₁-S phase transition, reaches its zenith at the beginning of S phase and gradually declines to undetectable levels by the end of S phase [42,43]. This scheme could explain how pol-δ introduces hypermutations at A/T pairs during DNA-dependent patch repair of initiating AID-generated U/G lesions independently of the process by which hypermutations accumulate at C/G pairs as a consequence of identical lesions. Why should the MSH2–MSH6–mediated pathway be substantially inoperative during S phase? We suspect that the answer is that most of the AID-generated uracil lesions that persist through to S phase are either excised by UNG2 or replicated over before MSH2–MSH6 is able to detect them in the context of U/G mispairs.

Analysis suggests that AID preferentially generates uracil lesions on the NTS to some extent in vivo [44]. AID might gain unequal access to the TS and NTS because of the asymmetrical structure of a transcription bubble. However, this does not alter the fact that the TS and NTS accumulate an almost equal number of AID-generated lesions at V(D)J genes in mice [4]. Therefore, it would seem likely that an incision that facilitates mutagenic patch repair in the dominant MSH2–MSH6–mediated pathway is introduced by an as yet unidentified mismatch repair endonuclease into either the TS or the NTS with equal probability preceding EXO1-mediated strand degradation. Likewise, an APE1-generated incision might facilitate the initiation of mutagenic patch repair in the back-up UNG2-mediated pathway at either the TS or the NTS with equal probability.

If this is so, one simple explanation of strand bias at A/T pairs during transcription-dependent SHM at V(D)J genes

Opinion

Trends in Immunology Vol.29 No.4
Figure 2. Proposed two-phase scheme of transcription-dependent somatic hypermutation (SHM) at V(D)J genes in germinal centre centroblasts. The DNA deamination model envisages that activation-induced deaminase (AID) triggers SHM and leads to C/G-targeted hypermutagenesis (phase 1) by deaminating cytosine residues in DNA to uracil. Incorporation of dATP opposite AID-generated uracil lesions generates transitions at C/G pairs. Incorporation of dNTP opposite non-instructional apyrimidinic (AP) sites that form because of excision of AID-generated uracil lesions by the nuclear isoform of uracil N-glycosylase (UNG2) generates transversions and transitions at C/G pairs. The initiation of A/T-targeted hypermutagenesis (phase 2) requires either the excision of AID-generated uracil lesions by UNG2 (minor pathway) or the recognition of AID-generated U/G mispairs by MSH2–MSH6, the MutS homologs that form a heterodimeric protein complex that identifies postreplicative base–base mispairs and initiates mismatch repair (major pathway). Apyrimidinic endonuclease 1 (APE1) might cleave the phosphodiester bond 5'-0 of an apyrimidinic site that is caused by UNG2 removal of an AID-generated uracil lesion from DNA. Exonuclease 1 (EXO1) facilitates A/T-targeted hypermutagenesis in the MSH2–MSH6-mediated pathway. The polymerase error model envisages that polymerase-δ (Pol-δ) misincorporates nucleotides preferentially opposite adenine and thymine template residues during DNA-dependent patch repair of initiating AID-generated uracil lesions. After mutagenic patch repair, subsequent replication opposite a mispair generated by Pol-δ would maintain the misincorporated nucleotide: one daughter cell would have a wild-type A/T pair and the other would have a misincorporated nucleotide paired with its complementary nucleotide. Strand degradation and resynthesis would correct initiating, AID-generated uracil lesions. After mutagenic patch repair, subsequent replication opposite a mispair generated by Pol-δ would maintain the misincorporated nucleotide: one daughter cell would have a wild-type A/T pair and the other would have a misincorporated nucleotide paired with its complementary nucleotide. Strand degradation and resynthesis would correct initiating, AID-generated uracil lesions. Evidence strongly suggests that AID deaminates C residues to uracil on both DNA strands with approximately equal frequency. The targeting of C/G pairs is strand unbiased (i.e., correlates with the strand targeting behaviour of AID). By contrast, the targeting of A/T pairs is strand biased. We propose that the lesions that lead to so called phase 1 hypermutagenesis accumulate mostly after the mechanism of phase 2 hypermutagenesis has operated during the G1 phase of the centroblast cell cycle. We suspect that most of the uracil lesions that AID generates in centroblasts during late G1 are either excised by UNG2 or replicated over before MSH2–MSH6 is able to detect them in the context of U/G mispairs. This scheme might explain how phase 2 hypermutations could accumulate in a strand biased manner without perturbing the strand unbiased accumulation of phase 1 hypermutations.

Opinion Trends in Immunology Vol.29 No.4 170
in vivo might be that centroblasts are more likely to undergo apoptosis because of incisions that are introduced into the TS than into the NTS. That is, an incision introduced into the TS during transcription-dependent SHM at a V(D)J gene might increase the possibility of the centroblast undergoing apoptosis and hence might be less likely to lead to the initiation of mutagenic patch repair than an incision introduced into the NTS. Consequently, those centroblasts that survive the early stages of the G₁ phase of the cell cycle will have potentially acquired hypermutations at A/T pairs in a strand-biased manner. The U/G mispairs that AID subsequently generates during the later stages of the G₁ phase of the centroblast cell cycle would lead to the accumulation of hypermutations at C/G pairs in a strand-unbiased manner. Supporting evidence for this explanation of strand bias at A/T pairs during SHM at V(D)J genes might be obtained by the use of genetic experiments in which the molecular machinery that contributes to spontaneous apoptosis is turned off in centroblasts.

It should be noted that it was recently proposed that strand bias at A/T pairs during SHM at V(D)J genes in vivo is attributable to preferential targeting of the dominant MSH2–MSH6–mediated pathway of phase 2 of SHM to AID-generated uracil lesions situated on the NTS [45]. However, it is difficult to reconcile this explanation with evidence which suggests that the targeting of A/T pairs is strand biased at V(D)J genes in MSH2-deficient mice [7].

Hypermutagenesis at S regions

AID is also essential for the induction of class switch recombination (CSR) (Figure 1; Box 1) in mice [9] and humans [10]. It is likely that, as anticipated by the DNA deamination model [11], the double strand breaks (DSB) that are needed for CSR to occur are introduced during the processing of AID-generated U/G lesions at the donor and acceptor S regions to be joined [5,46]. CSR is severely impaired in UNG-deficient mice [18] and humans [17] and is completely ablated in UNG/MSH2 double-deficient mice [4]. Hence, as per the initiation of A/T-targeted hypermutagenesis at V(D)J genes and S regions, the initiation of CSR requires either the excision of AID-generated uracil lesions by UNG2 or the recognition of AID-generated U/G mispairs by MSH2–MSH6 [4]. In this context, APE1 might facilitate the initiation of DSB formation at S regions in the UNG2-triggered pathway, the major pathway during CSR in mice [4,18] and humans [17]. That is, APE1 might introduce adjacent incisions on opposing strands that result in DSB formation. An as yet unidentified mismatch repair endonuclease might similarly facilitate the initiation of DSB formation at S regions in the MSH2–MSH6-triggered pathway, the minor pathway during CSR [4].

One can predict that, as per CSR initiation, the UNG2- and MSH2–MSH6–dependent pathways are the major and minor pathways by which A/T-targeted SHM is initiated at S regions, respectively. As at V(D)J genes, AID deaminates C residues on the TS and NTS with approximately equal frequency at the Sµ core 5' flank of mice [4,5]. However, unlike at V(D)J genes, A/T-targeted hypermutations also accumulate in a substantially strand-unbiased manner at that region [4,25,26]. This suggests that, at the Sµ core 5' flank of mice, the probability with which mutagenic DNA-dependent patch repair is recruited to a particular DNA strand to generate hypermutations at A/T pairs is directly related to the strand targeting behaviour of AID. AID deaminates C residues on the TS and NTS with approximately equal frequency at the donor Sµ region and at Sy1 and Sy2a acceptor regions after appropriate cytokine stimulation [5]. Hence, during the processing of AID-generated uracil lesions in the UNG2-dominant pathway at S regions, APE1 would introduce incisions in a strand-symmetric manner.

We expect that, unlike at V(D)J genes, hypermutations accumulate at S regions in mice according to a linear interpretation of the two-phase model of SHM. Phase 1 hypermutations would accumulate at S regions during S phase as a consequence of the uracil lesions that AID generates throughout the G₁ phase of the cell cycle. Phase 2 hypermutation (and DSB formation) would be initiated at S regions during the G₁-S transition phase of the cell cycle when UNG expression commences. AID-generated uracil lesions would be either excised by UNG2 or replaced over before MSH2–MSH6 is able to detect them in the context of U/G mispairs. We anticipate that components of the postreplicative mismatch repair machinery are limiting to some extent during most of the G₁ phase of the cell cycle in B lymphocytes that undergo SHM at S regions (and CSR), possibly because these clones have a longer cell cycle time than centroblasts.

CSR clearly occurs in B lymphocytes that are to some extent tolerant of the incisions that are introduced into the TS and the NTS during germline transcription at S regions (in order that requisite DSB are able to form), so it is unlikely that CSR occurs in centroblast-like clones that are primed to undergo spontaneous apoptosis. Accordingly, because incisions introduced into the TS at S regions by APE1 would not lead to a large proportion of switching B lymphocytes expiring through apoptosis, the probability of mutagenic patch repair initiating at either the TS or the NTS would directly correlate with the strand-targeting behaviour of AID. We envisage that this explanation applies to SHM at S regions and Sµ–Sx junctions in mice (where x denotes any 3' acceptor S region). This scheme agrees with evidence that CSR within germinal centres probably occurs in selected, high-affinity centrocytes that have re-entered the cell cycle [47].

References

1 Faili, A. et al. (2004) DNA Polymerase {eta} is involved in hypermutation occurring during immunoglobulin class switch recombination. J. Exp. Med. 199, 265–270
3 Petersen, S. et al. (2001) AID is required to initiate Nbl1/gamma-H2AX focus formation and mutations at sites of class switching. Nature 414, 660–665

171
14 Dickerson, S.K. et al. (2003) AID mediates hypermutation by deaminating single stranded DNA. J. Exp. Med. 197, 1291–1296
20 Kavli, B. et al. (2005) B cells from hyper-IgM patients carrying UNG mutations lack ability to remove uracil from ssDNA and have elevated genomic uracil. J. Exp. Med. 201, 2011–2021
27 Delbos, F. et al. (2007) DNA polymerase eta is the sole contributor of A T modifications during immunoglobulin gene hypermutation in the mouse. J. Exp. Med. 204, 17–23
30 Hanna, M.G., Jr (1964) An autoradiographic study of the germinal center in spleen white pulp during early intervals of the immune response. Lab. Invest. 13, 90–104
43 Fischer, J.A. et al. (2004) Proteolytic degradation of the nuclear isoform of uracil-DNA glycosylase occurs during the S phase of the cell cycle. DNA Repair (Amst.) 3, 505–513
44 MayoroV, V.I. et al. (2005) Expression of human AID in yeast induces mutations in context similar to the context of somatic hypermutation at G-C pairs in immunoglobulin genes. BMC Immunol. 6, 10