Towards an understanding of somatic hypermutation

Heinz Jacobs and Linda Bross

How germinal center (GC) B cells diversify their rearranged immunoglobulin genes by somatic hypermutation is unknown. However, the GC-specific activation-induced cytidine deaminase has been identified as a key factor controlling two central GC-specific events: somatic hypermutation and class-switch recombination of immunoglobulin genes. This factor may function as a catalytic subunit of an RNA-editing complex or, more directly, on DNA as a deoxy-cytidine deaminase in the hypermutation domain and class-switch region. Deamination of deoxy-cytidines on both strands may result in staggered DNA double-strand breaks (DSBs) that, in the hypermutation domain, become processed by member(s) of newly identified error-prone DNA polymerases. Direct evidence for DSBs in hot-spots of hypermutating immunoglobulin genes has been provided, implicating DSBs as reaction intermediates of an error-prone DSB-repair pathway acting specifically in GC B cells. These recent findings are key to the identification of the hypermutation mechanism.

Addresses
Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland
Correspondence: Heinz Jacobs; e-mail: Jacobs@BII.CH

Current Opinion in Immunology 2001, 13:208–218

0952-7915/01/$ – see front matter

Abbreviations

ACF	Apobec-1 complementation factor
AID	activation-induced cytidine deaminase
CDR	complementarity-determining region
CSR	class-switch recombination
DSB	DNA double-strand break
Ei	intronic enhancer
GC	germinal center
HYM domain	hypermutation domain
IgH	immunoglobulin heavy chain
IgL	immunoglobulin light chain
kb	kilobasepairs
MAR	matrix-attachment region
MMR	mismatch repair
TdT	terminal-deoxynucleotidyl transferase

Introduction

As the adaptive immune system is founded on clonal selection of antigen-specific lymphocytes, the generation of antigen-receptor diversity is fundamental for the proper functionality of this system. The V(D)J recombination process that generates the primary antigen-receptor repertoire in B cell precursors in the fetal liver or adult bone marrow is understood in great detail [1]. In contrast, the molecular mechanism that allows antigen-activated B cells to further diversify their immunoglobulin genes by hypermutation has remained obscure since the first reports on affinity maturation [2] and the first experimental evidence for somatic hypermutation were provided [3]. Concomitantly, the number of models proposing putative mechanisms for these processes has been growing since these initial observations (for a review, see [4]).

Although the principal immunobiology of somatic hypermutation is understood, the question of how, on average, every B cell generation can acquire approximately one point mutation in their rearranged V(D)J segments of immunoglobulin heavy chain (IgH) and immunoglobulin light chain (IgL) chain genes is central to the understanding of hypermutation and of oncogenic translocation in follicular B cell lymphomas [5]; this question is the central issue of this review.

Somatic hypermutation: why, where, by what mechanisms and how does it occur?

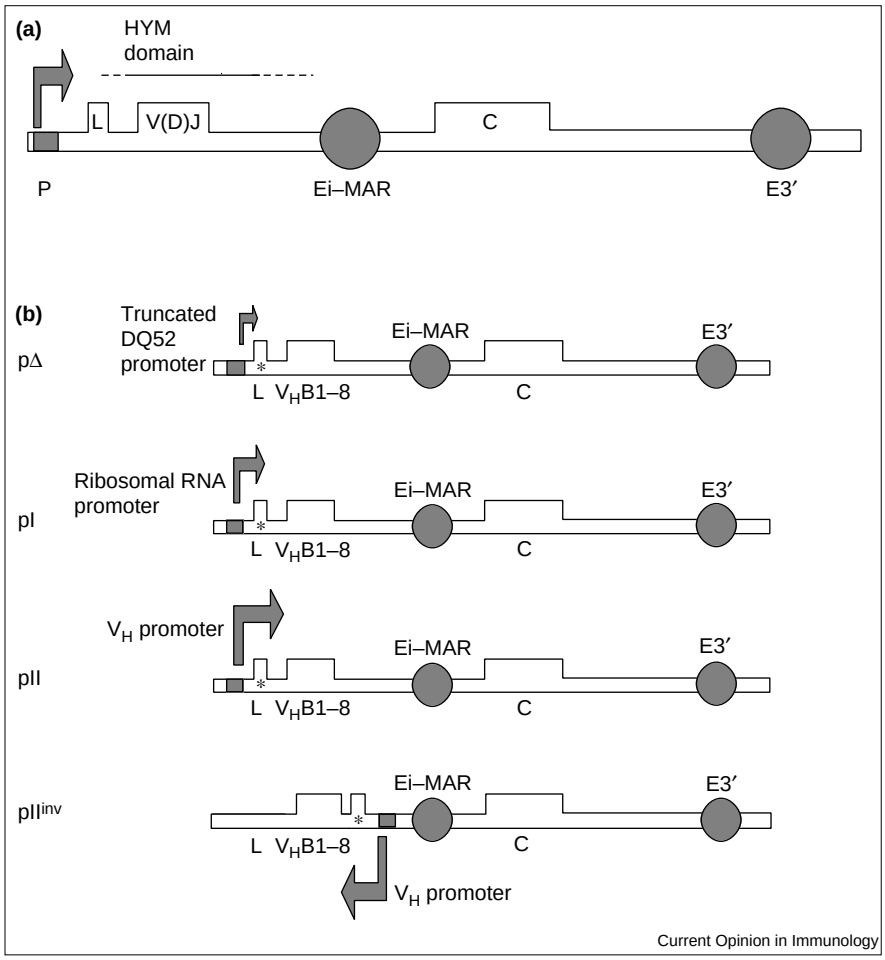
Following antigen challenge, activated B cells migrate into B cell follicles of secondary lymphatic organs, where they undergo rapid expansion and establish oligoclonal germinal centers (GCs). The GC is a specialized microenvironment, which arises during T-cell-dependent immune responses. Here GC B cells are able to modify their rearranged immunoglobulin genes through somatic hypermutation and to modify their constant region through class-switch recombination . The activation of a 'hypermutator' leads to the introduction of mutations in and around the V-region of IgH and IgL genes. A small proportion of B cells expressing an immunoglobulin with high affinity for antigen become selected and mature as 'survivors' of the GC reaction into memory B cells and plasma cells. Memory B cells and plasma cells shape the secondary immunoglobulin repertoire, which can be revealed serologically to have undergone 'affinity maturation' of antigen-specific immunoglobulins.

As well as increasing the number of lymphocytes, the expansion of GC B cells counteracts the substantial cell loss that occurs — as it does for precursor B and T cells [6]. In the case of GC B cells, the cells lost are precious antigen-specific B cells [7,8]. Thus, the GC can be visualized as a stock market where antigen-specific B cells take the risk and invest in their 'immunoglobulin options' and eventually improve them by hypermutation and clonal selection, along with increased effector function as a result of class-switch recombination (CSR).

Mutations: point mutations >> deletions > duplications

Mutations in hypermutated immunoglobulin genes are confined to the hypermutation domain (HYM domain), a region spaning about two kilobasepairs (kb) downstream of the immunoglobulin promoter region. Just downstream of the promoter, the frequency of mutations increases rapidly from 5′ to 3′ along the leader exon, peaks over the rearranged V(D)J exon and decreases in the J-C intron of immunoglobulin genes [9] (Figure 1a). The mutation rate has been estimated to approach 10^{-3} base pairs per generation, which is six orders of magnitude higher than that of

Figure 1



Elements controlling somatic hypermutation of immunoglobulin genes. The promoter (P), Ei–MAR and the 3′ enhancer (E3′) control somatic hypermutation. The promoter can be swapped with other promoters (see below). Ei is critical for somatic hypermutation. E3′ enhances not only transcription but also somatic hypermutation. **(a)** Mutations are restricted to the HYM domain, a region spaning the leader exon (L), the V(D)J region and the V(D)J intron. In principle, any sequence introduced in the HYM domain will undergo hypermutation. **(b)** Schematic representation of the promoter/V_HB1–8 transcription unit and the IgH enhancers in pΔ, pI, pII and pII^{inv} 'knockin' mice. B cells of pΔ mice express the targeted V_HB1–8 element under the control of a truncated DQ52 promoter [26]. B cells of pI knockin mice express the targeted V_HB1–8 element under the control of a ribosomal RNA-polymerase-I-dependent promoter. B cells of pII and pII^{inv} mice express the targeted V_HB1–8 element under the control of its regular IgH promoter. Whereas in pII the IgH-promoter/V_HB1–8 transcription unit is in its normal configuration, in pII^{inv} this unit is inverted, bringing the promoter adjacent to the intronic IgH enhancers [13••,25]. The asterisk in the leader intron represents an introduced stop codon, which renders all knockins into mutation reporters. A correlation between transcription and somatic hypermutation is found in these mice.

spontaneous mutations [10–12]. Although the majority of mutations are single point mutations, it is known that deletions occur at a frequency of approximately 4–7%, and duplications at a frequency of about 1%, of all mutations in memory B cells of 'passenger IgH knockin mice' and human tonsillar GC B cells [13••,14].

Intrinsic hot-spots of the hypermutator

The analysis of unselected mutations has provided information on the intrinsic specificity of the hypermutator and has indicated that somatic hypermutation is not a random process but is frequently targeted to hot-spots and exhibits characteristic nucleotide-substitution preferences. Transitions are more frequent than transversions and A nucleotides in the coding strand are replaced about two-fold more frequently than T nucleotides; the latter observation is referred to as DNA-strand bias. This strand bias is likely to relate to a short-patch DNA synthesis by an error-prone DNA polymerase with an asymmetry in the misincorporation frequency of template A versus T (see below). However, besides the DNA-strand bias, the analysis of hot-spot mutations indicates that both strands are hypermutation targets [15,16].

These hot-spots were identified on the basis of large databases of somatically mutated immunoglobulin genes. This led to the identification of the RGYW motif (where R = A or G; Y = C or T; and W = A or T) as a preferred target for the mutation machinery. In particular, AGC- and AGT-triplets (both of which are serine codons) have been identified as intrinsic hot-spots of the hypermutation machinery [17]. However, not all tetramers with an RGYW consensus or all AGC/T codons are equally targeted by the hypermutator. Whether these hot-spot motifs become frequent substrates of the hypermutator appears to depend on the surrounding sequence and might involve secondary structures.

In line with these observations, the complementarity-determining regions (CDRs) of antibody V-gene segments are biased in their codon usage. The serine codons AGC and AGT are used more frequently than the other four serine codons TCN (where N = A, G, C or T). This strongly implies the importance of evolving and maintaining hypermutable V-gene segments. Mutations in these intrinsic hot-spots can confer increased antigen-binding capacity and the hot-spots have often been identified as antigen-selected hot-spots [18]. Knowledge about the base-exchange pattern and intrinsic

hot-spots is extremely helpful in assessing reaction intermediates, such as DNA double-strand breaks (DSBs) [13••] or error-prone DNA polymerases [19•,20•,21,22] for their potential involvement in hypermutation.

Somatic hypermutation and transcription

The restricted occurrence of mutations to the hypermutation domain of rearranged immunoglobulin genes suggests the existence of specific *cis*-acting element(s) capable of targeting the hypermutator and controlling the mutability of immunoglobulin genes. In addition, the oncogenic risk imposed by an unspecific hypermutator calls for the existence of such elements. So far, the search for these elements has been related exclusively to transcriptional regulation. The promoter (reviewed in [9,23,24]; see also [25,26]) and the immunoglobulin enhancers [27], but not the rearranged V(D)J region itself [28–30], have been identified as critical components in promoting somatic hypermutation. Although the V-gene promoter is exchangeable by other RNA-polymerase-II-dependent promoters or even a ribosomal RNA-polymerase-I-dependent promoter, the strength of the promoter is critical for this process [25,26]. Placing an additional immunoglobulin promoter immediately 5′ of the constant portion of an IgL κ transgene leads to the occurrence of somatic hypermutation within 2 kb downstream of each promoter but not between these two areas. Thus, hypermutation is restricted to the first 2 kb of transcription elongation [4]. Whether the hypermutator has no access to the site of transcription initiation or, alternatively, whether mutations in the transcription initiation region are repaired more effectively as compared with downstream regions [31] is not clear.

These studies suggest a dependence of somatic hypermutation on transcription. To study this relationship in more detail, three mutant 'IgH knockin mice' were derived by gene targeting. In these mice, a targeted, nonfunctional V_HB1–8 passenger transgene was either placed under the transcriptional control of a weak, truncated, RNA-polymerase-II-dependent DQ52 promoter (pΔ), its own RNA-polymerase-II-dependent IgH promoter (pII) or an RNA-polymerase-I-dependent promoter (pI). A premature termination codon in the leader intron of the targeted V_HB1–8 renders all inserted transgenes nonproductive (Figure 1b). The transcribed regions of the targeted V_HB1–8 are identical in pI, pII and pΔ mice, excluding the possibility of differential secondary structures in the template strand, which is a prerequisite for a comparison between transcription and somatic hypermutation. The relative mutation-frequency of the V_HB1–8 passenger transgene in memory B cells of pΔ, pI and pII mice correlated with the relative levels of transgene-specific pre-mRNA expressed in GC B cells isolated from the mutant mice.

These data indicate that the mutation load of rearranged immunoglobulin genes correlates with — and can be tuned by — transcription [25,26]. Whether transcription is just enabling hypermutation to begin (by opening the chromatin and allowing the formation of stable secondary structures within the nontemplated strand; see below) or whether the hypermutator physically interacts with components of the RNA-polymerase complex [32] is a question that remains to be addressed.

Other hypermutation targets

Although hypermutation was initially believed to be an immunoglobulin-specific process, recent data indicate that the gene encoding the transcription factor Bcl6 mutates in B cells of the GC [33,34]. Although the pattern and distribution of mutations reflect those found in the HYM domain of immunoglobulin genes, the frequency of mutations is about 10–100-fold lower. These findings provide a direct link between translocations and the immunopathogenesis of follicular lymphomas [35]. Of note, as translocations found in follicular lymphomas are generally those that provide a selective advantage, they can be considered to be the tip of the iceberg, raising the question as to whether other highly transcribed GC genes can serve as substrates for the mutation machinery. However, as of yet no other genes have been identified as substrates [36]. In conclusion, targeting of hypermutation is not ideal; besides the preferential V-region substrates other targets exist, which can be causative of oncogenic translocations.

Searching for a repair pathway involved in somatic hypermutation

As suggested previously [37], the molecular mechanism underlying somatic hypermutation is unlikely to be achieved by a single gene product but is most likely to employ pre-existing components involved in DNA modification; that is, the hypermutator probably depends on other molecules involved in DNA repair and/or DNA synthesis to complete the mutagenesis. This scenario is reminiscent of V(D)J-recombination, which utilizes the site-specific RAG recombinase but strictly depends on components of the non-homologous-DNA repair — such as DNA-PKcs (cs, catalytic subunit), Ku70 and Ku86, as well as XRCC — to ligate the coding ends and complete the recombination process [1].

Based on these considerations, the screening of patients and mouse models with deficiencies in DNA repair or DNA synthesis for their hypermutation capability appears obvious. In addition, models and candidates proposed to be involved in hypermutation can be tested in these systems. The DNA-repair mutants analysed so far all appear to be hypermutation-competent. Given space limitations and the fact that most of these studies have been reviewed extensively already [38–41], the reader is referred to these publications. However, observations made in mismatch repair (MMR)-deficient B cells are further summarized below to discuss the potential relationship between hypermutation, transcription and MMR.

Somatic hypermutation in mismatch-repair-deficient mice

Based on sequence analysis of V-regions in single GCs, genealogical trees can be derived; these suggest a rate of about one additional mutation per B cell generation [42].

However, this does not necessarily imply a replication-based mutagenic step. Actually, recent data suggest that somatic hypermutation derives from a replication-independent, error-prone, short-patch DNA synthesis [43]. So, if mutations are introduced as a mismatch, how can they escape the active MMR machinery in GC B cells [44]?

Soon after the first orthologes of genes involved in MMR were cloned from mice, the respective mouse mutants became available and provided ideal model systems to address the impact of MMR on hypermutation. In general, the frequency of hypermutation ranges between 20 and 100% of wild-type levels and an altered base-exchange pattern with a bias for G/C mutations has been observed in several independent studies [38]. Although the reduced frequency of mutations in MMR-deficient GC B cells can in part be explained by a reduced viability due to genome instability [45,46], it does not explain the altered base-exchange pattern often found in MMR-deficient mice [38]. Compared with MMR-competent mice, the frequency in mutations of A nucleotides in the coding strand is relatively decreased whereas those at G and C nucleotides are increased.

Two principal opposing models can explain this G/C bias: in the first, MMR acts nonconventionally, as a pro- and co-mutagenic 'repair' system that fixes the introduced mutation by 'repairing' the opposite, unmutated nucleotide; alternatively, MMR acts conventionally as a contra- and post-mutagenic repair system. Hypermutation preferentially creates mutations at G/C nucleotides but MMR also repairs these mutations preferentially [47]. In both cases, the net result of hypermutation is an increased GC bias if MMR is defective (for a review, see [38]). An additional alternative is that two independent mutagenic steps operate: an MMR-independent/GC-biased/hot-spot-focused step and an MMR-dependent/AT-biased step [48]. Although so far there is no direct proof for an MMR-dependent step, MMR might still act conventionally, by providing sometimes a refinement to the hypermutator's own 'creativity'. However, the high frequency of DSBs at hot-spots of hypermutating immunoglobulin genes [13••] supports an MMR-insensitive/MMR-independent/GC-biased/hot-spot focused mechanism (see below).

DNA double-strand breaks in hypermutating B cells

In human B cell tumors, deletions and duplications frequently accompany the introduction of point mutations in the HYM domain, implicating DNA strand breaks in the mechanism of somatic hypermutation [14]. In addition, nucleotide insertions into the V- but not C-region have been revealed in a hypermutation-competent B cell tumor line expressing a transfected terminal-deoxynucleotidyl transferase (TdT) gene [49]. These data suggest the presence of either single- or double-strand breaks as intermediates in somatically mutating immunoglobulin genes.

As nonhomologous DSB repair is inherently associated with the generation of the primary B cell repertoire, a potential role for nonhomologous DSB repair in secondary diversification of immunoglobulin genes can not be addressed directly in mice defective in this repair pathway. Firstly, these mice need to be reconstituted with functional IgH and IgL genes to bypass defective V(D)J recombination and allow B cells to develop and, secondly, T cells have to be provided to allow T-cell-dependent immune response.

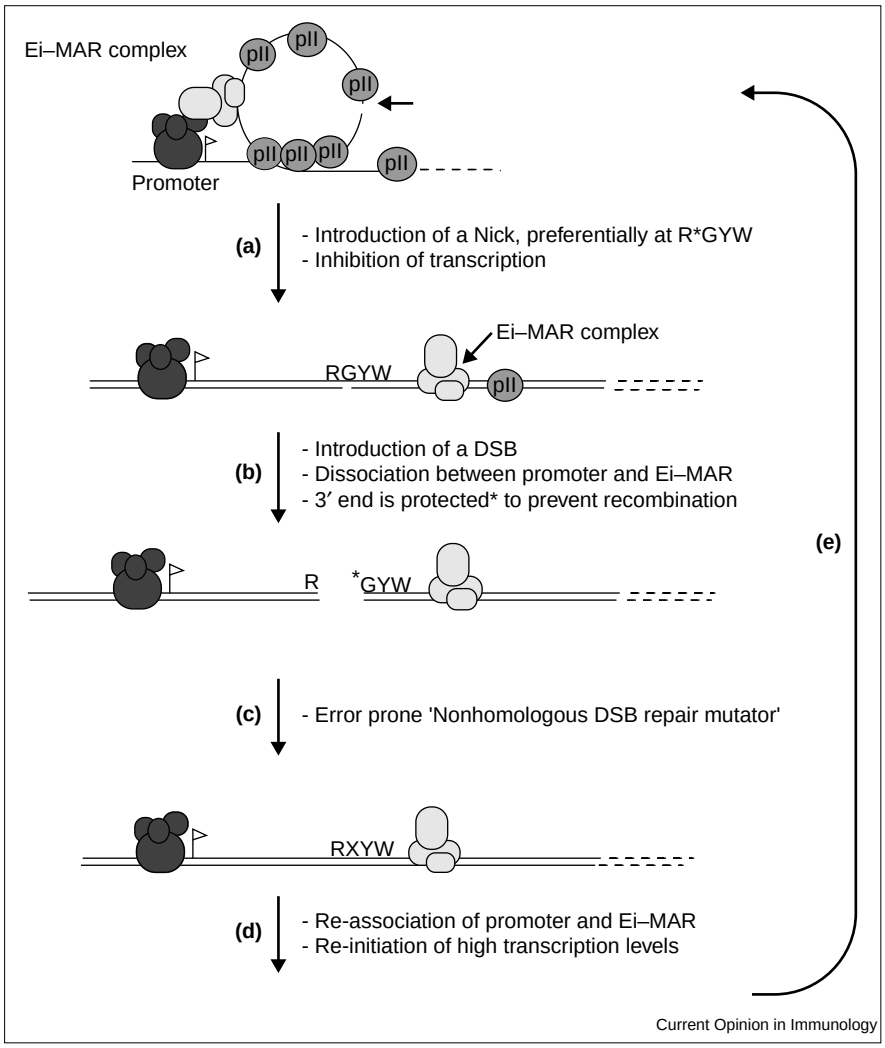
As an alternative, we directly addressed the potential occurrence of DSBs in the HYM domain of GC B cells taken *ex vivo* [13••]. The pII and pΔ IgH knockin miceserved in an ideal system to address these questions [25]. Besides a functionally rearranged endogenous IgH allele, all B cells of pII and pΔ mice carry the same nonfunctional 'passenger VDJ transgene' on the other allele. Employing a sensitive ligation-mediated PCR strategy, we demonstrated a high frequency of DSBs in and around the targeted VDJ element in B cells undergoing somatic hypermutation. The generation of these DSBs is favored by increased transcriptional activity, and their distribution follows that of point mutations in the HYM domain [13••].

Furthermore, similar to hot-spots of somatic hypermutation, 50–60% of all DSBs occur preferentially at RGYW motifs. Most remarkably, even within the RGYW motif these breaks do not occur randomly but preferentially (25 out of 36 breaks at RGYWs) 5′ of the G and R residues of the RGYW motif. This observation might relate to the previously observed differential mutability of the four residues within the RGYW motif: G > R > Y > W [17,50]. Nucleotides in codons 31^{II} (.A^{v}GC.T), 38^{III} (.AA^{v}G.CA), 65^{II} (.A^{v}GC.A), 100^{II} (.A^{v}GT.A) and $100A^{II\&III}$ (.A^{v}G^{v}C.T) (numbering according to [51]) have been described as mutational hot-spots of the V_HB1–8 family. From the total of 52 breaks, 2 breaks occurred 5′ of 38^{III}, 5 breaks 5′ of 65^{II} and 1 break 5′ of $100A^{III}$. Thus, 15% (8/52) of DSBs locate directly 5′ of three hot-spots of somatic hypermutation.

Because transcription favors the generation of DSBs, it appears advantageous to block transcription after the first DSB has been introduced. In our model, transcription is blocked by the DSB itself (Figure 2), by causing a transient dissociation of the immunoglobulin promoter from the intronic enhancer (Ei). At the same time, consecutive promoter-proximal DSBs are suppressed, deletions are prevented and an error-prone nonhomologous DSB repair pathway is initiated.

To test this model, 'IgH-promoter/V_HB1–8 inversion mice' were derived, which hereafter are referred to as pII^{inv} mice. In such mice, the IgH promoter/V_HB1–8 transcription unit is inverted such that the IgH promoter locates proximal to the Ei–MAR (matrix-attachment region) region, pointing away from the Ei–MAR region (Figure 1b). Like pII mice, the pII^{inv} mice carry the same targeted, nonfunctional V_HB1–8 IgH passenger allele under the transcriptional control of its own IgH promoter and a stop codon in the leader exon. The inversion and the premature termination codon in the leader

Figure 2



A model of DSBs as reaction intermediates in hypermutation. (a) Our data favor a model in which, early in somatic hypermutation, a nick is introduced by an as-yet-unknown endonuclease with an intrinsic preference for RGYW motifs. (b) Conversion of this nick into a DSB may be the result of a transesterfication or a second nuclease hit. As a consequence of the DSB, the Ei–MAR complex (light shading) might be uncoupled from the promoter complex (dark shading), thereby blocking transcription and subsequent deleterious DSBs. Although our PCR-based assay detects only blunt-ended DSBs, they may initially be introduced as staggered DSBs. (c) Subsequently, an error-prone repair process might act at the break sites and introduce mutation(s), here indicated by X. (e) As soon as the DSB is re-ligated by a nonhomologous DSB repair system, the enhancer–promoter complex reassembles and transcription is re-initiated. The * at the 3′ break end relates to the inefficient detection of this end by ligation-mediated (LM)-PCR, suggesting a protection of the 3′ end (i.e. a hairpin or a covalently bound protein)

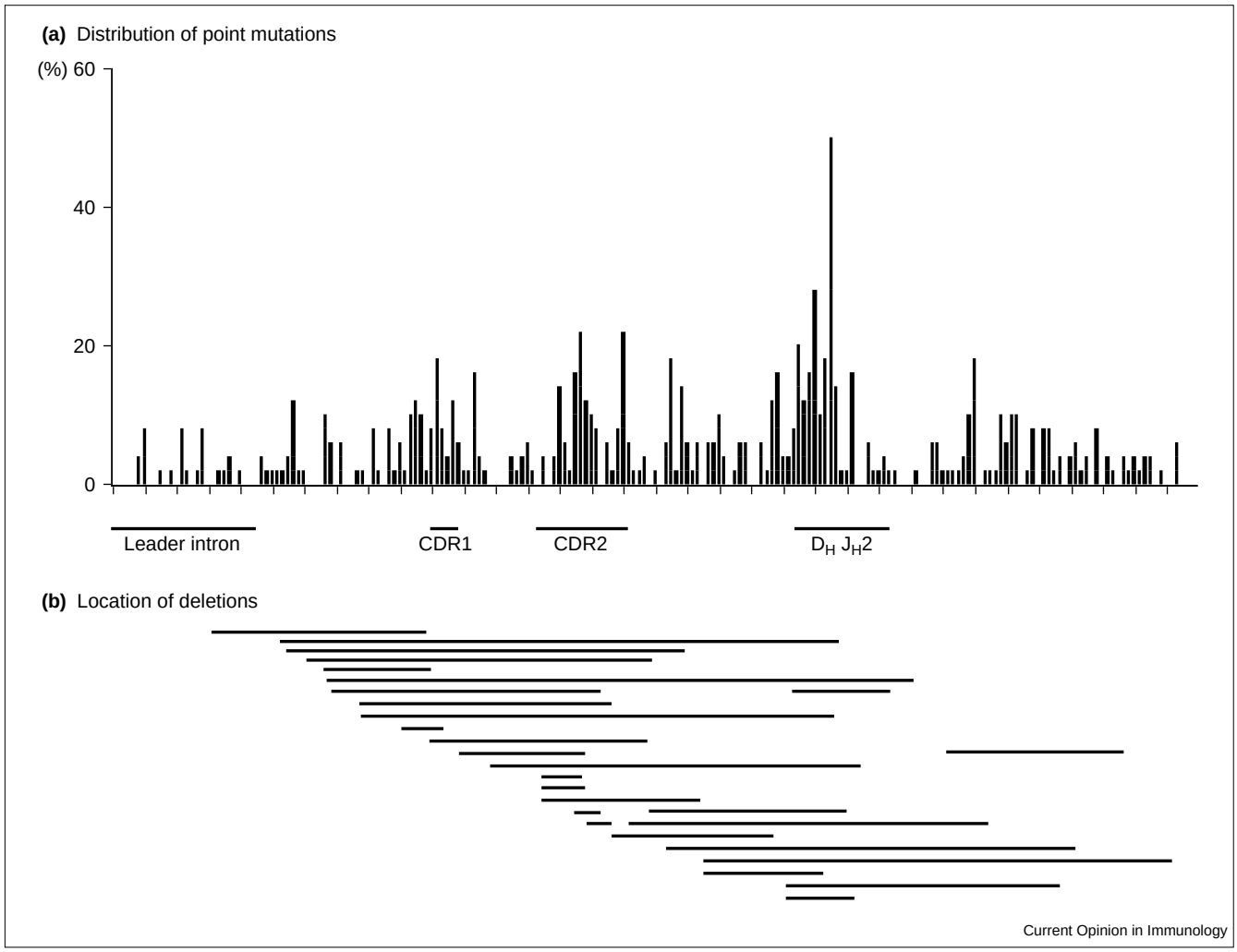
intron of the targeted $V_HB1–8$ render the inserted pII^{inv} transgene nonproductive.

The mutation frequency of the targeted $V_HB1–8$ genes of isotype-switched pII^{inv} memory B cells is normal. These results indicate that the orientation and location of the IgH promoter/$V_HB1–8$ unit relative to the Ei–MAR element is not essential for the mechanism underlying somatic mutations. Therefore, the role of the Ei–MAR in somatic hypermutation of rearranged V genes in the IgL κ locus [27] and IgH locus [13••] appears to relate primarily to transcription, that is the efficient loading of RNA polymerases at the immunoglobulin promoter.

Despite a comparable frequency of deletions in the $V_HB1–8$ gene of memory B cells of pII and pII^{inv} mice (5.7% and 7.3%, respectively), the most remarkable result of the pII^{inv} analysis is the size of deletions in pII^{inv} memory B cells (Figure 3b). Whereas deletions in pII mice are relatively small (ranging between 3 and 32 base pairs), more than 80% (28 out of 34) of the deletions in pII^{inv} mice are larger than 32 base pairs, the maximum size found in pII mice. The deletions in pII^{inv} mice often (16 out of 34 [47%]) span a region of more than 100 base pairs and range between 6 to 338 base pairs.

Whereas small deletions can be explained by exonucleases acting on the DSB, large deletions are most likely to be introduced by the ongoing loading of RNA polymerases after the introduction of a primary DSB in the pII^{inv} $V_HB1–8$ gene, thereby favoring multiple promoter-proximal DSBs prior to the repair of promoter-distal DSB(s). In accordance to the mapping of DSBs in GC B cells of pII and pΔ mice, the distribution and site preference of the deletion ends in memory B cells of pII^{inv} mice are very similar. Therefore, most mutations, and especially hot-spot mutations of somatically mutated immunoglobulin genes, are likely to be related to an error-prone, nonhomologous DSB repair pathway active in B cells of the normal GC. As predicted from our model

Figure 3



Distribution of point mutations and deletions in the targeted V_HB1–8 gene in memory B cells of pII^{inv} mice. Starting at the first base pair of the leader intron into a window of 612 nucleotides (shown on the X-axis) of V_HB1–8, the location and distribution of (a) point mutations and (b) deletions and are plotted. Point mutations are given as a percentage of mutations per codon sequenced. Bars below the histogram indicate the position of the leader intron, CDR1 and CDR2, and the D_HJ_H region.

(Figure 2), positioning the Ei–MAR upstream of the V_H gene promoter in the pII^{inv} system allows consecutive breaks, resulting in large deletions within the targeted VDJ element. Thus, the normal spatial separation between the immunoglobulin promoter and enhancer elements in rearranged immunoglobulin genes favors the maintenance of functional immunoglobulin genes and consequently the survival of B cells undergoing somatic hypermutation [13••].

These data provide the direct evidence for the existence of DSBs as reaction intermediates in somatic hypermutation. The introduction of DSBs and subsequent nonhomologous DNA end-joining are not only part of the mechanism creating the primary immunoglobulin repertoire through V(D)J recombination — or changing immunoglobulin effector functions through immunoglobulin class-switching — but also serve in the formation of the secondary immunoglobulin repertoire through somatic hypermutation. Thus, introducing DSBs in immunoglobulin genes is used repeatedly to diversify the immunoglobulin repertoire and immunoglobulin effector functions, leaving a potential for pathogenic translocations at early and late stages of B cell development.

Subversion of mismatch repair: DNA double-strand breaks and transcription

If mutations are introduced as a mismatch, there is a central question: how can mismatches, once introduced by an error-prone DNA polymerase, escape detection in a MMR-proficient environment? In this respect, the fact that a MMR deficiency has only a minor effect on somatic hypermutation is actually most remarkable [38]. In our view, there are two potential mechanisms by which Nature has accomplished this problem on the molecular level. Firstly,

most mutations are introduced at the DSB-intermediate stage of hypermutation and do not appear as a mismatch in an otherwise intact double-stranded DNA substrate. Once the DSB is ligated in an error-prone manner, the mismatch should become detectable. However, now the transcription machinery is re-activated again and acts in a mismatch-protective fashion, as mismatches in the transcription bubble are 'invisible/hidden' in front of the MMR system. In the alternative mechanism, one can postulate that the mutagenesis acting at a break-end does not generate a mismatch at all; for example, it could be envisaged as a TdT-like activity. An enzyme with such an activity is DNA polymerase μ which, based on its expression pattern and activity, might be involved in somatic hypermutation [20•].

The generation of duplications and deletions [13••,14] supports the idea that the blunt-ended DSBs detected by ligation-mediated PCR are actually derivatives of staggered breaks. Although normally one of the overhangs is filled by an error-prone DNA polymerase and the other break end presumably blunted by a nuclease or by a transesterification reaction, occassionally both staggered ends may get trimmed (leading to deletion) or filled (leading to duplication) by an error-prone DNA polymerase. The sloppiness of recently identified DNA polymerases makes these enzymes interesting candidates for an error-prone fill-in reaction [21,22]. Of note, DNA polymerase ι is the most error-prone eukaryotic polymerase. In 75% of cases, this enzyme introduces a G (rather than an A in the other 25%) opposite of template T *in vitro* [19•], suggesting that DNA polymerase ι could establish the A/T bias.

In summary, besides the fact that high levels of transcription are important for antibody production by plasma cells, high levels of transcription in GC B cells favor the generation of DSBs. These DSBs are likely intermediates of the mutation reaction [13••] and simultaneously help in the subversion of the MMR system by 'hiding' the mismatch in a transcription bubble. In this respect, the recently discovered family of error-prone DNA polymerases, with either a TdT activity, lack of proof-reading activity or capable of translesion synthesis, are obvious candidates for gene targeting in hypermutation research [19•,20•,21,22].

The activation-induced cytidine deaminase

Using cDNA subtraction, Muramatsu *et al.* [52•], in The Department of Medical Chemistry at Kyoto University, recently cloned the activation-induced cytidine deaminase (AID), which is specifically expressed in a switch-induced B lymphoma line and GC B cells. AID is a 198-residue protein with a predicted molecular mass of 24 kDa and is homologous to the cytidine deaminases, Apobec-1 and Apobec-2. The former forms the catalytic subunit of the apolipoprotein B (apoB) mRNA-editing complex and associates with the Apobec-1 complementation factor (ACF), a 64 kDa protein that contains three nonidentical RNA-binding motifs. This complex is capable of deaminating specifically the C^{6666} of the apoB mRNA into a U^{6666} [53••].

Interestingly, unlike Apobec-2 [54] — which maps to human chromosome 6p12 — both AID and Apobec-1 are encoded on chromosome 12p13 [55] and the cDNAs of AID and Apobec-1 are more similar to each other than to Apobec-2 (Figure 4a). This suggests a common AID/Apobec-1 ancestor and a conserved function of these proteins as RNA-editing enzymes. Like Apobec-1 and Apobec-2, the amino-terminal region of AID has a highly conserved cytidine deaminase domain and a leucine-rich region in the carboxy-terminal region, presumably involved in protein–protein interaction with ACF-like proteins [52•].

Interestingly, as shown in AID-deficient mice and humans, AID is required for somatic hypermutation as well as class-switch recombination but not for the formation of GCs [56••,57••]. In humans, mutations in AID are associated with the autosomal recessive form of the hyper-IgM syndrome (HIGM2) [57••]. These findings indicate that there is one key player that controls two central phenomena of the GC reaction and argue for the existence of a critical initial event controlling somatic hypermutation and class-switch recombination.

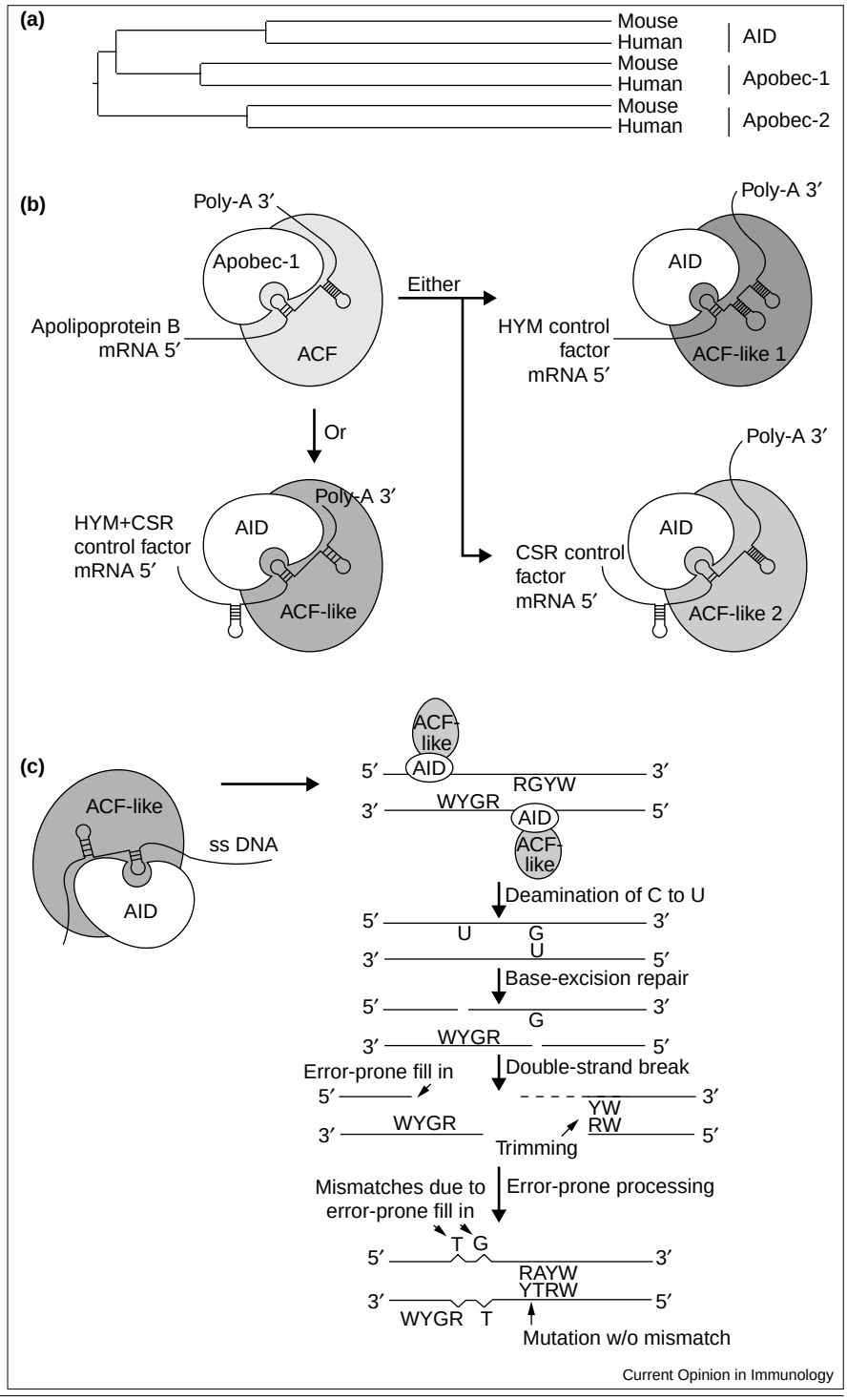
Potential functions of activation-induced cytidine deaminase

AID may function like other members belonging to this gene family — as a catalytic subunit of an RNA-editing complex [52•,53••,56••,57••], by editing specifically the RNA of the putative 'hypermutator and/or class switcher' (Figure 4b). Alternatively, as AID can deaminate deoxycytidine [52•], AID itself might be part of the hypermutator and class-switcher, and function more directly as an enzyme that induces nicks or single-nucleotide gaps (Figure 4c). Together with ACF-like factor(s), AID might specifically deaminate cytosines that occur in secondary structures of transcribed DNA.

The generation and stability of these secondary structures are likely to depend on transcription. Hydrolytic deamination of C→U in DNA is a frequent spontaneous lesion in any cell type. If not repaired, these lesions have an increased spontaneous mutation rate, causing G/C→A/T transitions in *Escherichia coli*. However, normally these lesions are efficiently repaired by base-excision repair (BER): U-DNA glycosylase hydrolyses the N-glycosylic bond that links U to the sugar-phosphate backbone (the base-exision step) and leaves an AP site (apurynic or apyrimidinic site) in the DNA duplex; the next step is the removal of the AP site by an AP-endonuclease, which most often (depending on the enzyme) hydrolyse the 5′ phospho-diester bond of the AP site, causing a strand break in the duplex DNA; the 5′-terminal deoxyribose-phosphate is then excised by a DNA deoxyribosphosphodiesterase, leaving a single-nucleotide gap that is repaired by a fill-in synthesis and ligation [58].

Based on this knowledge, one can postulate that AID, together with ACF-like component(s), can induce a

Figure 4



Features and potential functions of AID. (a) AID is a homolog of Apobec-1 and Apobec-2. This phylogenetic tree was derived by alignment of the nucleotide sequence of the coding region of human and mouse AID, Apobec-1 and Apobec-2 using the Jotun Hein method as employed in the Lasergen software package. The different stem–loop structures symbolize different mRNA species. Consistent with the chromosomal colocalisation, AID and Apobec-1 are more related to each other than to Apobec-2. Accession numbers (Genbank) of the nucleotide sequences used are as follows: human AID, AB040431; mouse AID, AF132979; human Apobec-1, AB009422; mouse Apobec-1, U21951; human Apobec-2, AF161698; mouse Apobec-2, NM009694. (b) AID is a potential RNA-editing enzyme. Based on sequence identity and functional characteristics, AID might function, like Apobec-1, as an enzymatic component of an RNA-editing complex. Together with an ACF-like RNA-binding protein, AID may edit the RNA of HYM and/or CSR control factor(s). (c) AID might function as a nick-inducing deaminase by deamination of C to U in the HYM domain or CSR of immunoglobulin genes. As shown in the figure, deamination of both strands can lead to staggered DSBs, which become the substrates of error-prone DSB repair (see text for more details). w/o, without.

more-or-less site-specific deamination of cytosine, thereby recruiting global BER and causing nicks and single-nucleotide gaps. If the AID-containing complexes deaminate Cs on both strands, the introduction of staggered DSBs is preprogrammed. As discussed already, the processing of these staggered DSBs in the hypermutation domain of immunoglobuin genes might then involve error-prone DNA polymerases such as μ and ι.

Germinal centers in the absence of activation-induced cytidine deaminase: from stock options to grandpa's mattress

AID is not required for the formation of GCs in mice and humans [56••,57••]. In fact, GCs in AID-deficient patients and mice are relatively enlarged. As a consequence, one might think that AID deficiency affects selection of B cells [59]. In our view it is more likely that enlarged GCs should

be expected in B cells with a 'defective hypermutator'. Accordingly, AID-deficient, antigen-specific B cells 'play it safe'. They do not take the risk of 'investing' in their immunoglobulin options and therefore do not win (i.e. they neither increase their affinity by up to 1000-fold nor class-switch) but also do not lose (i.e. they do not fail to make an antigen-specific immunoglobulin). AID-deficient GC B cells follow 'grandpa's mattress strategy': they save more or less every single penny (low-affinity B cells) that is deposited; they do not mutate and therefore escape the substantial cell loss that is normally a result of the GC reaction.

Conclusions

Besides the fact that high levels of transcription are important for antibody production by plasma cells, high levels of immunoglobulin transcription favor the generation of DSBs in the GC B cells and simultaneously might help in the subversion of the MMR system. DSBs are found in the HYM domain and are, like mutation hot-spots, preferentially found at RGYW motifs. These DSBs, which can be revealed as deletion-ends in memory B cells of pII^{inv} mice, are likely to be intermediates of an error-prone, nonhomologous end-joining repair pathway activated in GC B cells. In this respect, the recently discovered family of error-prone DNA polymerases, with TdT activity, lack of proof-reading activity or translesion synthesis are of special interest.

Thus, the introduction of DSBs and subsequent nonhomologous DNA end-joining are not only part of the mechanism creating the primary immunoglobulin repertoire through V(D)J recombination — or changing immunoglobulin effector functions through immunoglobulin class-switching — but also serve in the formation of the secondary immunoglobulin repertoire through somatic hypermutation. Thus, introducing DSBs in immunoglobulin genes is used generally to diversify the immunoglobulin repertoire and immunoglobulin effector functions, leaving a potential for pathogenic mutations at early and late stages of B cell development.

The finding that AID is a key player in controlling two central events of the GC reaction argues for the existence of a single critical initial event that controls somatic hypermutation and class-switch recombination. No matter whether AID will turn out to play its role indirectly as a catalytic subdomain in an RNA-editing complex(es) or more directly as nick-inducing enzyme, the identification, cloning and functional characterisation of AID already is the major breakthrough in understanding these two critical events of late B cell development.

Update

After the submission of this review, Papavasiliou and Schatz [60••] also reported on the existence of abundant DSBs in hypermutating immunoglobulin genes. Except for the model of somatic hypermutation, their data are in agreement with the data described in [13••]. The first paper to describe the use of ligation-mediated (LM)-PCR as a method to look at strand breaks in somatic hypermutation was by Lo *et al.* [61].

Acknowledgements

The authors are grateful to R Torres, G Kline and F Mc Blane for critical reading the manuscript. Our work is supported by F Hoffmann-La Roche Ltd., Basel Switzerland, which founded and supports the Basel Institute for Immunology.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Ramsden DA, van Gent DC, Gellert M: Specificity in V(D)J recombination: new lessons from biochemistry and genetics. *Curr Opin Immunol* 1997, 9:114-120.

2. Eisen HN, Siskind GW: Variations in affinities of antibodies during the immune response. *Biochemistry* 1964, 3:996-1008.

3. Weigert MG, Cesari IM, Yonkovich SJ, Cohn M: Variability in the lambda light chain sequences of mouse antibody. *Nature* 1970, 228:1045-1047.

4. Storb U: The molecular basis of somatic hypermutation of immunoglobulin genes. *Curr Opin Immunol* 1996, 8:206-214.

5. Stamatopoulos K, Kosmas C, Belessi C, Stavroyianni N, Kyriazopoulos P, Papadaki T: Molecular insights into the immunopathogenesis of follicular lymphoma. *Immunol Today* 2000, 21:298-305.

6. Borst J, Jacobs H, Brouns G: Composition and function of T-cell receptor and B-cell receptor complexes on precursor lymphocytes. *Curr Opin Immunol* 1996, 8:181-190.

7. MacLennan IC: Germinal centers. *Annu Rev Immunol* 1994, 12:117-139.

8. Rajewsky K: Clonal selection and learning in the antibody system. *Nature* 1996, 381:751-758.

9. Neuberger MS, Milstein C: Somatic hypermutation. *Curr Opin Immunol* 1995, 7:248-254.

10. McKean D, Huppi K, Bell M, Staudt L, Gerhard W, Weigert M: Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc Natl Acad Sci USA* 1984, 81:3180-3184.

11. Allen D, Cumano A, Dildrop R, Kocks C, Rajewsky K, Rajewsky N, Roes J, Sablitzky F, Siekevitz M: Timing, genetic requirements and functional consequences of somatic hypermutation during B-cell development. *Immunol Rev* 1987, 96:5-22.

12. Berek C, Berger A, Apel M: Maturation of the immune response in germinal centers. *Cell* 1991, 67:1121-1129.

13. •• Bross L, Fukita Y, McBlane F, Demolliere C, Rajewsky K, Jacobs H: DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. *Immunity* 2000, 13:in press.
Using mice described in [25], this article reports on the direct and indirect evidence for the frequent occurrence of DSBs in mutation hot-spots of hypermutating immunoglobulin heavy chain (IgH) passenger transgenes in germinal center (GC) B cells. The generation of these DSBs depends on transcription. DSBs are likely to be reaction intermediates of an error-prone nonhomologous DSB repair system activated in GC B cells.

14. Goossens T, Klein U, Küppers R: Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy chain disease. *Proc Natl Acad Sci USA* 1998, 95:2463-2468.

15. Dorner T, Foster SJ, Farner NL, Lipsky PE: Somatic hypermutation of human immunoglobulin heavy chain genes: targeting of RGYW motifs on both DNA strands. *Eur J Immunol* 1998, 28:3384-3396.

16. Milstein C, Neuberger MS, Staden R: Both DNA strands of antibody genes are hypermutation targets. *Proc Natl Acad Sci USA* 1998, 95:8791-8794.

17. Rogozin IB, Kolchanov NA: Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim Biophys Acta* 1992, 1171:11-18.

18. Jolly CJ, Wagner SD, Rada C, Klix N, Milstein C, Neuberger MS: The targeting of somatic hypermutation. *Semin Immunol* 1996, 8:159-168.

19. • Tissier A, McDonald JP, Frank EG, Woodgate R: Poliota, a remarkably error-prone human DNA polymerase. *Genes Dev* 2000, 14:1642-1650.
See annotation to [20•].

20. • Dominguez O, Ruiz JF, Lain de Lera T, Garcia-Diaz M, Gonzalez MA, Kirchhoff T, Martinez AC, Bernad A, Blanco L: DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *EMBO J* 2000, 19:1731-1742.
Together with [19•], this paper reports newly identified DNA polymerases that are interesting candidates for an error-prone DNA-synthesis-repair step in hypermutation.

21. Woodgate R: A plethora of lesion-replicating DNA polymerases. *Genes Dev* 1999, 13:2191-2195.

22. Goodman MF, Tippin B: Sloppier copier DNA polymerases involved in genome repair. *Curr Opin Genet Dev* 2000, 10:162-168.

23. Storb U, Peters A, Klotz E, Kim N, Shen HM, Hackett J, Rogerson B, Martin TE: Cis-acting sequences that affect somatic hypermutation of Ig genes. *Immunol Rev* 1998, 162:153-160.

24. Winter DB, Gearhart PJ: Somatic hypermutation: another piece in the hypermutation puzzle. *Curr Biol* 1995, 5:1345-1346.

25. Fukita Y, Jacobs H, Rajewsky K: Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity* 1998, 9:105-114.

26. Jacobs H, Puglisi A, Rajewsky K, Fukita Y: Tuning somatic hypermutation by transcription. *Curr Top Microbiol Immunol* 1999, 246:149-158.

27. Betz AG, Milstein C, Gonzalez-Fernandez A, Pannell R, Larson T, Neuberger MS: Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell* 1994, 77:239-248.

28. Azuma T, Motoyama N, Fields LE, Loh DY: Mutations of the chloramphenicol acetyl transferase transgene driven by the immunoglobulin promoter and intron enhancer. *Int Immunol* 1993, 5:121-130.

29. Yelamos J, Klix N, Goyenechea B, Lozano F, Chui YL, Gonzalez Fernandez A, Pannell R, Neuberger MS, Milstein C: Targeting of non-Ig sequences in place of the V segment by somatic hypermutation. *Nature* 1995, 376:225-229.

30. Bachl J, Olsson C: Hypermutation targets a green fluorescent protein-encoding transgene in the presence of immunoglobulin enhancers. *Eur J Immunol* 1999, 29:1383-1389.

31. Tu Y, Tornaletti S, Pfeifer GP: DNA repair domains within a human gene: selective repair of sequences near the transcription initiation site. *EMBO J* 1996, 15:675-683.

32. Storb U, Peters A, Klotz E, Kim N, Shen HM, Kage K, Rogerson B, Martin TE: Somatic hypermutation of immunoglobulin genes is linked to transcription. *Curr Top Microbiol Immunol* 1998, 229:11-19.

33. Shen HM, Peters A, Baron B, Zhu X, Storb U: Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* 1998, 280:1750-1752.

34. Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, Chaganti RS, Klein U, Küppers R, Rajewsky K, Dalla-Favera R: BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci USA* 1998, 95:11816-11821.

35. Küppers R, Klein U, Hansmann ML, Rajewsky K: Cellular origin of human B-cell lymphomas. *N Engl J Med* 1999, 341:1520-1529.

36. Shen HM, Michael N, Kim N, Storb U: The TATA binding protein, c-Myc and survivin genes are not somatically hypermutated, while Ig and BCL6 genes are hypermutated in human memory B cells. *Int Immunol* 2000, 12:1085-1093.

37. Jacobs H, Fukita Y, van der Horst GT, de Boer J, Weeda G, Essers J, de Wind N, Engelward BP, Samson L, Verbeek S *et al.*: Hypermutation of immunoglobulin genes in memory B cells of DNA repair-deficient mice. *J Exp Med* 1998, 187:1735-1743.

38. Reynaud CA, Bertocci B, Frey S, Delbos F, Quint L, Weill JC: Mismatch repair and immunoglobulin gene hypermutation: did we learn something? *Immunol Today* 1999, 20:522-527.

39. Kim N, Bozek G, Lo JC, Storb U: Different mismatch repair deficiencies all have the same effects on somatic hypermutation: intact primary mechanism accompanied by secondary modifications. *J Exp Med* 1999, 190:21-30.

40. Kim N, Storb U: The role of DNA repair in somatic hypermutation of immunoglobulin genes. *J Exp Med* 1998, 187:1729-1733.

41. Kelsoe G: V(D)J hypermutation and DNA mismatch repair: vexed by fixation. *Proc Natl Acad Sci USA* 1998, 95:6576-6577.

42. Jacob J, Kelsoe G, Rajewsky K, Weiss U: Intraclonal generation of antibody mutants in germinal centres. *Nature* 1991, 354:389-392.

43. Bertocci B, Quint L, Delbos F, Garcia C, Reynaud CA, Weill JC: Probing immunoglobulin gene hypermutation with microsatellites suggests a nonreplicative short patch DNA synthesis process. *Immunity* 1998, 9:257-265.

44. Park K, Kim J, Kim HS, Shin HS: Isolated human germinal center centroblasts have an intact mismatch repair system. *J Immunol* 1998, 161:6128-6132.

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

46. Vora KA, Tumas-Brundage KM, Lentz VM, Cranston A, Fishel R, Manser T: Severe attenuation of the B cell immune response in Msh2-deficient mice. *J Exp Med* 1999, 189:471-482.

47. Phung QH, Winter DB, Cranston A, Tarone RE, Bohr VA, Fishel R, Gearhart PJ: Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from mice deficient in the MSH2 mismatch repair protein. *J Exp Med* 1998, 187:1745-1751.

48. Rada C, Ehrenstein MR, Neuberger MS, Milstein C: Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity* 1998, 9:135-141.

49. Sale JE, Neuberger MS: TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. *Immunity* 1998, 9:859-869.

50. Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS: Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. *Proc Natl Acad Sci USA* 1993, 90:2385-2388.

51. Kabat EA, Wu TT: Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *J Immunol* 1991, 147:1709-1719.

52. • Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T: Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem* 1999, 274:18470-18476.
The first report on the identification, cloning and characterisation of AID by cDNA subtraction of cDNAs derived from switch-induced and un-induced murine B lymphoma. The homology to the catalytic subunit of RNA-editing complexes and the expression pattern of AID suggest a specific role in events confined to GC B cells.

53. •• Mehta A, Kinter MT, Sherman NE, Driscoll DM: Molecular cloning of Apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol* 2000, 20:1846-1854.
The cloning and characterisation of ACF provides important molecular insights into the role of Apobec-1 (and other family members) as a site-specific cytidine deaminase.

54. Liao W, Hong SH, Chan BH, Rudolph FB, Clark SC, Chan L: APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochem Biophys Res Commun* 1999, 260:398-404.

55. Muto T, Muramatsu M, Taniwaki M, Kinoshita K, Honjo T: Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene. *Genomics* 2000, 68:85-88.

56. •• Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T: Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 2000, 102:553-563.
See annotation to [57••].

57. •• Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A *et al.*: Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 2000, 102:565-575.
The publications [56••,57••] actually deserve to be assigned •••. They report on the unexpected finding that class-switch recombination and somatic hypermutation are controlled by AID. This strongly supports the idea that there is a single key mechanism initiating these events, which are restricted to germinal center B cells.

58. Friedberg EC, Walker GC, Siede W: *DNA Repair and Mutagenesis.* Washington DC: ASM Press; 1994.

59. Longacre A, Storb U: A novel cytidine deaminase affects antibody diversity. *Cell* 2000, 102:541-544.

60. •• Papavasiliou FN, Schatz DG: Cell-cycle-regulated DNA double strand breaks in somatic hypermutation of immunoglobulin genes. *Nature* 2000, 408:216-221.
In line with [13••], this article reports on the existence of abundant DSBs in sequences introduced in the hypermutation domain of immunoglobulin transgenes in germinal center B cells. The generation of DSBs is coupled to transcription, is enhancer dependent and is cell cycle restricted. In contrast to [13••], an error-prone homologous DSB repair model is proposed.

61. Lo AK, Ching AK, Lim PL, Chui YL: Strand breaks in immunoglobulin gene hypermutation. *Ann NY Acad Sci* 1997, 815:432-435.