THE ROYAL SOCIETY

doi 10.1098/rstb.2000.0743

Somatic hypermutation of immunoglobulin and non-immunoglobulin genes

Ursula Storb[1*], Hong Ming Shen[1], Nancy Michael[1] and Nayun Kim[2]

[1]*Department of Molecular Genetics and Cell Biology, and* [2]*Department of Biochemistry and Molecular Biology, University of Chicago, 920 East 58th Street, Chicago, IL 60637, USA*

Somatic hypermutation (SHM) of immunoglobulin (Ig) genes is a highly specific mechanism restricted to B lymphocytes during only a few cell generations. Data presented here suggest that transcription of the target genes is required, but not sufficient for SHM. Presumably, *cis*-acting elements, such as those present in the Ig enhancers, are required to target a mutator factor (MuF) to Ig and human *BCL-6* genes. It is postulated that the MuF travels with the transcribing RNA polymerase and is deposited on the target gene when the polymerase pauses. Point mutations, and rare deletions and insertions, are created by the combined actions of MuF and certain DNA polymerases. A subset of the mutations is corrected during SHM by DNA mismatch repair.

Keywords: somatic hypermutation; immunoglobulin genes; transcription; DNA repair; hot spots; cold spots

1. INTRODUCTION

Immunoglobulin (Ig) genes are expressed in B lymphocytes. Site-specific recombination events (V(D)J recombination) occurring in precursor cells (preB) lead to the expression in each B cell of a single heavy chain and a single light chain gene, each with a specific variable (V) region. Based on about 100 V genes, their combinatorial joining with a small set of different J (and D) genes, and the occurrence of small deletions and random insertions of nucleotides at the V(D)J joining sites, an initial repertoire of millions of different B cells is created, each with a different antigen binding site. Encounter of a mature B cell with an antigen that can bind to the Ig molecules expressed on the B-cell surface induces a somatic hypermutation (SHM) process which alters the variable regions of the expressed heavy and light chain genes. Since this process is coupled with a very high rate of cell divisions (approximately one cell cycle every 6 h) and since the mutation process continues for about ten or more cell generations, a clone of related B cells is created that express a variety of differently modified derivatives of the Ig heavy and light chain genes expressed by the founder cell. B cells which produce mutated Igs that have a high affinity for the immunizing antigen are selected for survival and may become long-lived memory cells.

The Ig SHM process is limited to the variable region of the Ig gene and its proximate flanks, covering a mutable stretch of 1–2 kb from the 5′-end of the transcribed region (see a review by Storb 1996). There is clear evidence that SHM is linked to transcription (reviewed by Storb *et al.* 1998*a*). Inserting an Ig promoter 5′ of the constant region of an Ig gene causes mutation of the normally unmutated constant region of the gene (Peters & Storb 1996). Removal of the Ig enhancers inhibits SHM (Betz *et al.* 1994). An Ig gene-specific promoter, however, does not seem to be required, at least in the presence of Ig enhancers (Betz *et al.* 1994; Tumas-Brundage & Manser 1997). Furthermore, the *c-Myc* gene, when translocated into an Ig locus, becomes a target for SHM; in this situation, the *c-Myc* gene is transcribed from its own promoter (Rabbits *et al.* 1984).

Based on these findings we have begun to investigate whether other genes that are expressed in B cells during SHM may be targets for this process. We found that the proto-oncogene *BCL-6* is highly mutated in normal human memory B cells (Shen *et al.* 1998). (So far, the mouse *BCL-6* gene has not been found to be mutated.) The mutations have all the hallmarks of SHM. In the study reported here we have further investigated the question of target specificity of SHM in comparing the mutation frequencies of Ig and *BCL-6* genes with those of other non-Ig genes that are expressed in B cells undergoing SHM. Furthermore, we have considered the presence of hot and cold spots for SHM and their mutability in Ig and non-Ig genes.

2. TARGETING OF SOMATIC HYPERMUTATION TO Ig AND *BCL-6* GENES

The sequences of the expressed variable regions of the V_{H4} family were determined in the memory B cells of two normal human donors. High levels of somatic point mutations were found with average mutation frequencies of 2.9×10^{-2} and 12.6×10^{-2} in the V_{H4} genes associated with μ and γ constant region genes, respectively, of donor A, and 5.6×10^{-2}, in the V_{H4} genes of IgM^{low}, IgD^{-} B cells of donor B (table 1). The constant regions were tested in donor A and were not mutated above background levels. The memory B-cell donors were 42 and 55 years old, respectively. Their γ heavy chain mutation frequencies were considerably higher than those reported for a

*Author for correspondence (stor@midway.uchicago.edu).

Table 1. *Mutations in Ig and non-Ig genes in memory and virgin B cells*
(Modified after Shen *et al*. 2000.)

gene	donor	isotypes	cells[b]	clones	base pairs	mutated bases	frequencies of mutated bases
IgV_{H4}	A	IgM	—	11	1899	53	2.9×10^{-2}
		IgG	—	17	2900	365	$12.6 \times 10^{\times 2}$
	B	IgG[a]	—	10	1600	90	5.6×10^{-2}
BCL-6[c]	A	—	IgD^- M	34	26 894	20	7.4×10^{-4}
	C	—	IgD^-IgM^- M	19	15 029	29	1.9×10^{-3}
	D	—	IgD^-IgM^- M	18	14 238	21	1.5×10^{-3}
TBP	C	—	IgD^-IgM^- M	98	82 712	3	3.6×10^{-5}
			IgD^+IgM^+ V	107	90 308	1	1.1×10^{-5}
	D	—	IgD^-IgM^- M	120	101 280	8	7.9×10^{-5}
			IgD^+IgM^+ V	70	59 080	2	3.4×10^{-5}
	D[d]	—	IgD^-IgM^- M	16	8000	0	$< 1.25 \times 10^{-4}$
c-Myc	C	—	IgD^-IgM^- M	117	99 099	3	3.0×10^{-5}
			IgD^+IgM^+ V	125	105 875	4	3.8×10^{-5}
	D	—	IgD^-IgM^- M	59	49 973	7	1.4×10^{-4}
			IgD^+IgM^+ V	54	45 738	3	6.6×10^{-5}
survivin	C	—	IgD^-IgM^- M	109	91 761	5	5.4×10^{-5}
			IgD^+IgM^+ V	130	111 957	6	5.4×10^{-5}
	D	—	IgD^-IgM^- M	131	108 072	8	7.4×10^{-5}
			IgD^+IgM^+ V	143	118 470	3	2.0×10^{-5}

[a] The analysis was done on IgM^{low}, IgD^- peripheral blood cells using a JH primer and a V_{H4} primer. Thus, some other isotypes besides IgG are probably also included.
[b] M, memory B cells; V, virgin B cells.
[c] Data from Shen *et al*. 1998.
[d] DNA from 16 clones was sequenced directly, without single-strand conformational polymorphism.

four-year-old child (Klein *et al*. 1994). This may suggest that memory B cells continue to undergo somatic hypermutation during life, although it is possible that a continuous selection process culls the B cells with less mutated Ig genes and perhaps lower affinity Igs. The mutation frequencies in the μ genes analysed from donor A were similar in the older donors to those in the four-year-old child, which presumably indicates that most of the expressed μ genes were derived from B cells that had recently emerged from their precursors and were undergoing a first immunization event.

Mutation frequencies in *BCL-6* genes in memory B cells of donor A and two additional donors, C and D, were also determined and found to be well above background mutation levels, although about ten to 100 times lower than in the Ig genes (table 1). They have all the hallmarks of SHM (Shen *et al*. 1998). Mutations of *BCL-6* in the tonsillar germinal centre B cells of normal human subjects have also been found by others (Pasqualucci *et al*. 1998; Peng *et al*. 1999). The fact that *BCL-6* mutability is lower than Ig mutability may suggest that the rate of transcription influences the rate of SHM. Ig genes are the most highly transcribed genes in B cells. *BCL-6*, although upregulated in mutating B cells, is expressed at moderate levels (Shen *et al*. 2000). A correlation between the rate of transcription and frequency of SHM has also been observed in mouse heavy chain genes that were expressed from promoters with different transcriptional efficiencies (Fukita *et al*. 1998).

It has been observed in most studies that SHM shows a particular profile of hot-spot targeting and cold-spot avoidance. Detailed analyses of large databases of somatically hypermutated mouse and human Ig genes were carried out by Wysocki and collaborators to determine the relative mutabilities of the 64 possible nucleotide triplets (Shapiro *et al*. 1999; Smith *et al*. 1996). The analyses were done with non-selectable sequences. For example, the V_H gene mutations of non-productively rearranged human Ig heavy chain genes were used to derive mutation indices (MI) for each of the 64 triplets (Shapiro *et al*. 1999; Smith *et al*. 1996) and for individual nucleotides within the triplets (Shapiro *et al*. 1999; G. Shapiro and L. Wysocki, personal communication). The values of the MI range from 0.24 to 2.62. An MI of 1.0 is assigned to a triplet that is found to be mutated at the frequency that would be expected if SHM is a random process. An MI >1.0 indicates a favoured SHM target or hot spot, and an MI < 1.0 indicates an avoided target or cold spot. The greater the difference from 1.0, the higher is the likelihood of mutation or no-mutation, respectively. Based on the findings of Wysocki and collaborators (Smith 1996;

Table 2. *Average mutation indices (MI) of mutated nucleotides in Ig and non-Ig genes in memory and virgin B cells* [a]

	Ig	*BCL-6*	*TBP*	*c-Myc*	*survivin*
donor A (M)[b]	1.28	1.34	—	0.81	—
donor B (M)	1.58	—	—	—	—
donor C (M)	—	1.31	0.57	0.84	0.68
donor C (V)	—	—	0.40	0.92	0.84
donor D (M)	—	1.39	0.97	0.93	0.93
donor D (V)	—	—	0.65	0.96	0.66

[a] The MI was calculated for all mutated nucleotides (see table 1). MIs > 1.0 are in italic.
[b] (M), memory B cells; (V), virgin B cells.

Table 3. *Intrinsic mutability of IgV_H and the non-Ig genes tested*[a]

(Table modified from Shen *et al*. 2000.)

		percentage of hottest hot spots and coldest cold spots[d]				
triplet[b]	mutation index[c]	IgV_{H4-11}	*BCL-6*	*TBP*	*c-Myc*	*survivin*
hot						
AGC	2.5	2.73	0.52	0.83	1.3	1.46
GCT	2.18	1.82	3.51	1.31	2.37	1.80
GTA	2.62	0.91	0.65	0.83	0.83	0.69
TAA	2.22	0.45	0.78	1.43	1.66	1.44
TAC	2.33	3.18	0.26	1.43	0.36	1.06
TAG	2.54	0.45	0.91	1.43	0.83	0.38
total	—	9.54	6.63	7.26	7.35	6.83
cold						
CCC	0.37	1.82	2.08	1.66	2.37	2.85
CTC	0.37	2.73	3.51	2.26	1.78	1.82
GAC	0.24	1.36	0.52	1.31	1.89	1.21
GCG	0.43	0.91	1.04	1.07	1.3	0.91
GTC	0.46	1.82	1.43	1.07	1.18	1.08
TCG	0.32	0.45	1.56	0.95	0.95	0.68
TGA	0.28	1.36	0.91	1.07	1.42	2.04
total	—	9.45	11.05	9.39	10.89	10.59
mean MI[e]		1.62	1.49	1.58	1.51	1.51

[a] Number of nucleotides evaluated in this analysis: IgV_{H4-11}, 222; *BCL-6*, 771; *TBP*, 844; *c-Myc*, 847; *survivin*, 1647.
[b] The six trinucleotides with the highest mutation index (MI) > 2.0, and the seven trinucleotides with the lowest MI < 0.5 (Shapiro *et al*. 1999).
[c] MI of the hottest and coldest triplets.
[d] Percentage of indicated triplets in the germline sequences of V_{H4-11} and the four non-Ig genes.
[e] Average intrinsic mutation index of each of the 64 triplets × percentage of each in the sequences investigated.

Shapiro 1999), we developed a computer-based analysis program that scans a given sequence and records the intrinsic MI of each triplet in single nucleotide steps (N. Michael and U. Storb, unpublished data).

To determine further whether the mutations in the *BCL-6* gene were due to SHM, we determined the MIs of the mutated nucleotides in *BCL-6* and compared them with those in the IgV_H genes (table 2). The MIs were 1.28 and 1.58 for the Ig genes and between 1.31 and 1.39 in the *BCL-6* genes. Thus, it appears that the same rules as those for Ig genes govern the targeting of *BCL-6* genes for SHM.

3. MUTATIONS IN NON-Ig GENES, OTHER THAN *BCL-6*, DO NOT TARGET SOMATIC HYPERMUTATION HOT SPOTS

In order to investigate whether expression of a gene was sufficient for targeting of SHM, three other non-Ig genes were analysed from memory and virgin B cells of donors C and D (Shen *et al*. 2000). These donors had the highest levels of *BCL-6* mutations and therefore the greatest likelihood of having other non-Ig genes mutated if they were mutable by SHM. The three genes chosen are all well expressed in mutating B cells. The TATA-binding protein (*TBP*) gene encodes a basal transcription factor that is expressed in all cells and is upregulated in germinal centre B cells (Shen *et al*. 2000). The *c-Myc* gene encodes a regulator of the cell cycle and is also upregulated in mutating B cells which proliferate highly (Shen *et al*. 2000). The *survivin* gene, involved in anti-apoptotic activity, has been shown to be expressed in lymphomas derived from germinal centre blasts (Ambrosini *et al*. 1997) as well as in normal tonsillar germinal centre B cells (Alizadeh *et al*. 2000). Its expression is also up-regulated in germinal centre B cells compared with virgin B cells (Shen *et al*. 2000).

As shown in table 1, very few mutations were found in these three genes. In memory B cells, the average mutation frequencies ranged from 3×10^{-5} to 1.4×10^{-4}. Mutations were also found in the virgin B cells. However, in contrast to the findings with Ig and *BCL-6* genes, the mutation frequencies in the memory B cells were not significantly different from the mutation frequencies in virgin B cells ($p = 0.063$ for mutations per total nucleotides analysed or $p = 0.152$ for mutated DNA clones per total clones, using the Mantel–Haenszel test).

It therefore appears that these genes are not mutated by the SHM process. This conclusion is further supported by comparing the MI of the mutated nucleotides in these three genes with those in the Ig and *BCL-6* genes (table 2). While the latter are well above 1.0, the MIs for mutated nucleotides in *TBP, c-Myc* and *survivin* are between 0.57 to 0.97. Furthermore, the average MI of all mutated nucleotides in these three genes is 0.82 in memory B cells and 0.73 in virgin B cells, and thus not significantly different. Interestingly, none of the strong hot spots present in the sequences of these genes (see below) is mutated (not shown). Presumably, the few mutations found in the *TBP, c-Myc* and *survivin* genes are

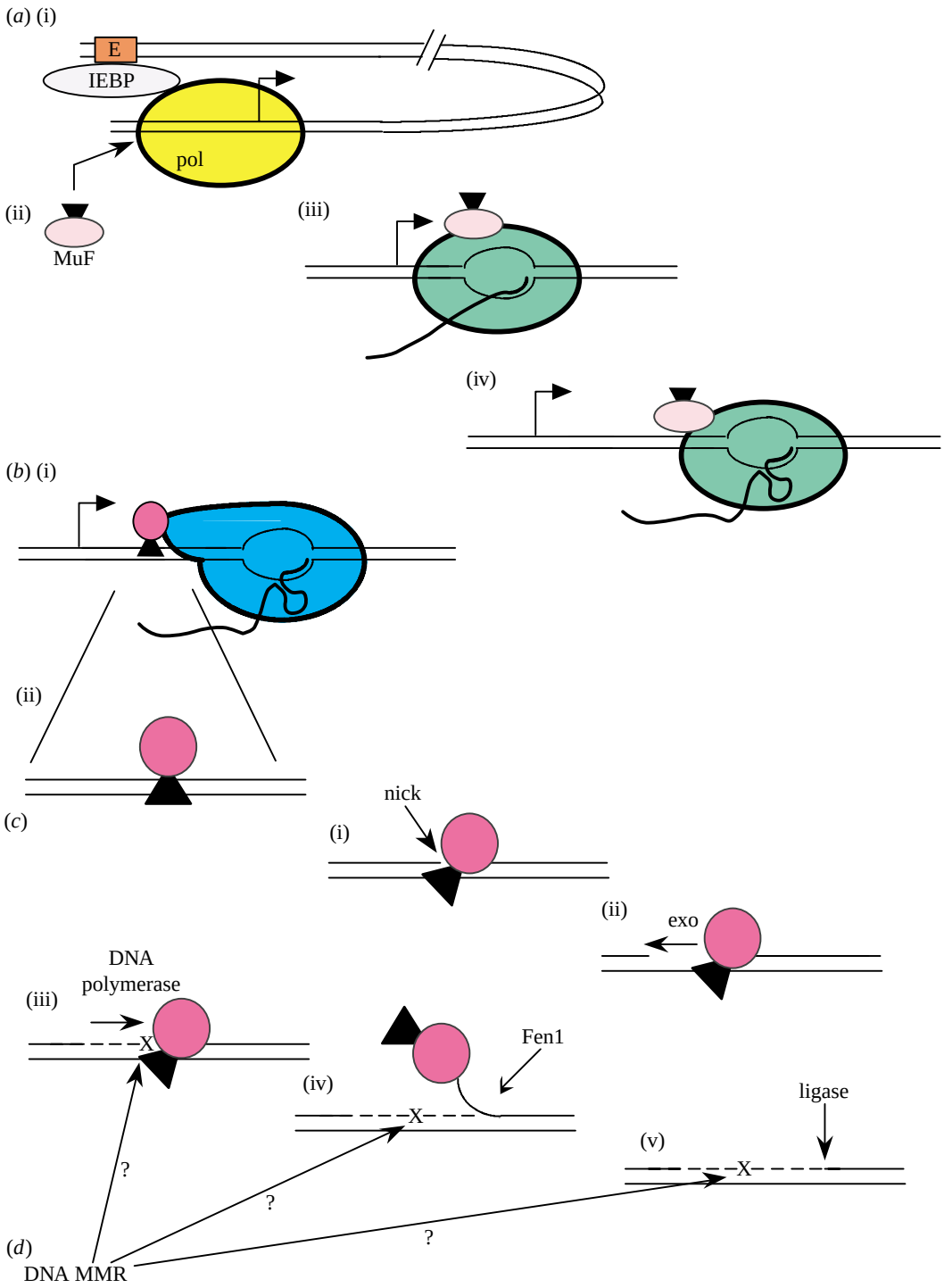

Figure 1. Model of somatic hypermutation, modified after Storb *et al.* (1998*a*).

due to PCR errors. Since their MIs, based on SHM criteria, are very different from the MIs of mutations in Ig and *BCL-6* genes, SHM hot spots are not targeted in these genes. If targeting the empirically derived hot spots is a hallmark of SHM, then the mutation pattern observed in *TBP, c-Myc* and *survivin* suggests that these mutations must have arisen by different rules than those governing SHM of Ig and *BCL-6* genes.

The lack of targeting of the non-Ig genes *TBP, c-Myc* and *survivin* is not due to an intrinsic lack of Ig gene-type mutability of these three genes compared with *BCL-6* (table 3). It can be assumed that the intrinsic MI would predict the likelihood that a sequence would be a substrate for SHM, if the primary sequence of an expressed gene and its location within 1–2 kb from the transcriptional promoter were the only requirements for targeting SHM. Considering the total NT content, the average MI of *BCL-6* is lower than those of the other three non-Ig genes. Furthermore, in *BCL-6* the percentage of the hottest hot spots is lower, and the percentage of the coldest cold spots is higher than the average of the other three non-Ig genes. Therefore, the fact that *BCL-6* is mutable by the SHM mechanism must be due to other *cis*-acting elements besides the primary mutation target

sequence. Presumably, in Ig genes these *cis*-acting elements lie within the Ig enhancers. Elements with the same function, but perhaps not necessarily the same sequence, must be present within or near the *BCL-6* gene.

4. A MODEL OF SOMATIC HYPERMUTATION REVISITED

We have proposed a model for SHM of Ig genes (Storb *et al.* 1998*b*) that we will revisit here with consideration of the findings with non-Ig genes. We will take into account the reports on novel error-prone DNA polymerases (Winter & Gearhart, this issue; Diaz *et al.*, this issue; Lawrence & Mayer, this issue; McDonald *et al.*, this issue; Ruiz *et al.*, this issue; Tippin & Goodman, this issue; Bridges, this issue; Reynaud *et al.*, this issue) and data relating to hot spots and secondary structure of the SHM target sequence (U. Storb and N. Michael, unpublished data).

SHM probably involves the activity of an unidentified mutator factor (MuF) which is only synthesized in mutating B cells and not in their precursors or progeny. Based on the need for MuF and the findings that SHM depends on transcription, SHM can be divided into the following stages (figure 1):

(i) loading of MuF on the RNA polymerase (Pol);
(ii) deposition of MuF onto the target DNA during a pause in transcription elongation;
(iii) introduction of point mutations by the combined action of MuF and DNA polymerases;
(iv) mismatch repair (MMR) of some of the mutated bases.

While transcription is apparently required for SHM and the rate of transcription appears to correlate with the rate of SHM, transcription of a gene by itself does not make it a target for SHM as shown by the findings reported here. Thus *cis*-elements in or near the Ig and *BCL-6* genes (and presumably other, not yet identified genes susceptible to SHM) probably target the postulated MuF to the correct target genes. In the Ig genes, such *cis*-elements are probably present in the Ig enhancers since an Ig enhancer appears to be sufficient to target SHM to an associated chloramphenicol acetyltransferase gene (Azuma *et al.* 1993). It may be envisioned that the MuF binds directly to such *cis*-elements, or perhaps is recruited by other factors binding to the specific *cis*-elements. For example, a specific combination of transactivating factors binding to the Ig enhancers may bind MuF and somehow alter its configuration so that it can associate with a Pol engaged in a transcription initiation complex (figure 1*a*).

The MuF may remain bound to Pol during transcript elongation until the Pol pauses and changes its configuration, at which time the MuF may be deposited on the transcribed DNA (figure 1*b*–*e*). We have postulated that certain secondary structures, such as hairpins that can form in the nascent RNA, may occasionally cause Pol to pause. Other causes may be in the primary structure of the transcribed DNA or in the need for remodelling of chromatin by displacing nucleosomes during progression of the RNA polymerase. If Pol pauses for an extended period of time, other RNA polymerases upstream may also have to pause and, if they happen to carry MuF, may deposit it as well. Such a chain reaction may be the cause of clusters of point mutations over 100–200 nucleotides observed in Ig genes that carry multiple mutations (N. Michael and U. Storb, unpublished data).

Once deposited on the DNA, MuF will induce the actual mutation event. It is not known whether mutations occur preferentially or exclusively on either the transcribed or the non-transcribed DNA strand or whether they occur randomly on either strand. In most cases of SHM there appears a strong bias of mutations from A over mutations from T when the non-transcribed strand is considered, which would indicate one of two possibilities: (i) that the mutations target A in preference to T and that the non-transcribed strand is targeted, or (ii) that the mutations target T rather than A and that the transcribed strand is targeted (Storb *et al.* 1998*c*). However, cases of SHM without A–T bias have also been reported and the question of strand-specific targeting has not been resolved. We postulate that MuF is an endonuclease which creates a single nick (figure 1*f*) (or double-strand cut; the question of 'nick' or 'cut' was discussed at The Royal Society discussion meeting; Jacobs *et al.*, this issue; Papavasilou & Schatz, this issue) but is not resolved) and that the repair of the nick (or double-strand cut) is accompanied with the introduction of point mutations. A double-strand cut is perhaps less likely, since it would either require non-homologous end joining with somehow replacing a nucleotide, or repair by homologous recombination. The latter is unlikely since there are no donors for most of the sequences, unless, of course, SHM occurs in the G2 phase of the cell cycle where an intact sister chromatid could serve as recombination donor. However, homologous recombination is generally not highly error prone as would be required to obtain the high mutation frequencies observed in SHM.

For simplicity, the model in figure 1 is based on a single-strand nick and we have arbitrarily chosen the top strand as the target (figure 1*f*). The nick could be resolved in at least two ways, assuming in both cases that MuF is covalently bound to the 5′-end of the nick. A 3′ exonuclease could create a short gap which could then be filled by a high-fidelity DNA polymerase, such as Pol δ, or any one of the low fidelity repair polymerases (Pol η, ζ, ι, κ, μ, etc.) (figure 1*g*,*h*). It is further postulated that in the case of Pol δ, when it reaches the position where MuF is covalently bound to the single strand and interacting with the intact opposite strand, a bypass DNA polymerase may have to take over (Goodman & Tippin 2000). Alternatively, one of the bypass DNA polymerases could, without an exonuclease step, directly initiate repair DNA synthesis from the nick, replacing the DNA strand with the bound MuF. In analogy to long patch base excision repair (Klungland & Lindahl 1997) we further postulate that the DNA strand with the MuF attached 5′ would be cleaved by Fen1 and the newly synthesized sequence would be ligated to the Fen1 created 5′-end (figure 1*i*,*j*). Finally, the MuF may not be covalently bound and a single nucleotide or base may be deleted at one end of the nick and replaced by an error-prone mechanism.

It is likely that the pattern of hot and cold spots specific for SHM depends on MuF or DNA exonucleases and polymerases preferences during this mutation phase of

SHM. The fact that this pattern depends on RNA polymerase pausing preferences is unlikely since these do not need to be sequence specific. Thus, perhaps the sites of deposition of MuF or its endonuclease cutting have certain sequence preferences. Furthermore, the hot spots and cold spots may be due to DNA polymerase preferences. In our understanding of the known error-prone DNA polymerases, none show the pattern found in SHM. Of course, the presence of the postulated MuF or of other factors specific to the Ig genes may alter DNA polymerase specificities in unexpected ways *in vivo*. It is also possible that there is yet another as yet unidentified DNA poymerase that is involved in SHM.

We postulate that MuF is released at stage (i). (figure 1). Perhaps MuF is recycled. If so, it cannot reassociate with an elongating RNA polymerase, only with one at the stage of transcription initiation and only after interaction with enhancer elements and/or enhancer bound transactivating factors. This would explain why SHM is restricted to the first 1–2 kb 3′ of the promoter. Presumably, nearly 100% of the depositions of MuF onto the target DNA occur within this region.

There is considerable evidence that DNA MMR has an interplay with SHM (Cascalho *et al*. 1998; Frey *et al*. 1998; Jacobs *et al*. 1998; Kim *et al*. 1999; Kim & Storb 1998; Phung *et al*. 1998; Rada *et al*. 1998; Winter *et al*. 1998). In our view, MMR is not involved in creating the mutations since in MMR-deficient mice the frequency of mutations and the mutation hot spots and cold spots are not significantly altered (Kim *et al*. 1999). The major change observed in MMR-deficient mice is a bias for mutations from G and C over mutations from A and T (however, the A > T bias on the non-transcribed strand remains, further supporting that MMR is not involved in the creation of the mutations). Thus, the currently available evidence suggests that MMR preferentially corrects certain SHM-created mutations, namely those from G or C. MMR must be recruited to the DNA sequence undergoing SHM immediately after the mutations are created so that the mismatch is retained for MMR recognition before replication can fix the mutation (figure 1*h–j*). Normally, MMR is found coupled to DNA replication. There is no compelling evidence that SHM occurs during the S phase of the cell cycle. Thus, MMR may operate independently of genomic DNA replication. If, as proposed by Jiricny (1998), the DNA strand that has the newly introduced mutation is marked for MMR in eukaryotic cells by a single-strand nick, one has to postulate that MMR interacts with the SHM complex before the DNA strand containing mutation(s) is religated.

Obviously, there are many intriguing possibilities of how SHM of Ig genes interacts with the mechanisms of chromatin remodelling, transcription and DNA repair. Fortunately, we have a battery of techniques, and cell and animal models at our disposition that promise new insights by the time the next Royal Society Discussion Meeting on SHM is convened.

We are grateful to L. Wysocki and G. Shapiro for providing unpublished analyses and T. E. Martin for critical reading of the paper. This work was supported by National Institutes of Health grants GM38649 and AI47380.

REFERENCES

Alizadeh, A. (and 30 others) 2000 Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511.

Ambrosini, G., Adida, C. & Altieri, D. 1997 A novel anti-apoptosis gene, *survivin*, expressed in cancer and lymphoma. *Nature Med.* **3**, 917–921.

Azuma, T., Motoyama, N., Fields, L. & Loh, D. 1993 Mutations of the chloramphenicol acetyl transferase transgene driven by the immunoglobulin promoter and intron enhancer. *Int. Immunol.* **5**, 121–130.

Betz, A., Milstein, C., Gonzalez-Fernandes, R., Pannell, R., Larson, T. & Neuberger, M. 1994 Elements regulating somatic hypermutation of an immunoglobulin K gene: critical role for the intron enhancer/matrix attachment region. *Cell* **77**, 239–248.

Cascalho, M., Wong, J., Steinberg, C. & Wabl, M. 1998 Mismatch repair co-opted by hypermutation. *Science* **266**, 1207–1210.

Frey, S., Bertocci, B., Delbos, F., Quint, L., Weill, J. & Reynaud, C. 1998 Mismatch repair deficiency interferes with the accumulation of mutations in chronically stimulated B cells and not with the hypermutation process. *Immunity* **9**, 127–134.

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

Goodman, M. F. & Tippin, B. 2000 Sloppier copier DNA polymerases involved in genome repair. *Curr. Opin. Genet. Dev.* **10**, 162–168.

Jacobs, H. (and 13 others) 1998 Hypermutation of immunoglobulin genes in memory B cells of DNA repair-deficient mice. *J. Exp. Med.* **187**, 1735–1743.

Jiricny, J. 1998 Replication errors: cha(lle)nging the genome. *EMBO J.* **17**, 6427–6436.

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

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

Klein, U., Kueppers, R. & Rajewsky, K. 1994 Variable region gene analysis of B cell subsets derived from a 4-year-old child: somatically mutated memory B cells accumulate in the peripheral blood already at a young age. *J. Exp. Med.* **180**, 1383–1393.

Klungland, A. & Lindahl, T. 1997 Second pathway for completion of human DNA base excision–repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.* **16**, 3341–3348.

Papavasilou, F. N. & Schatz, D. 2000 Cell-cycle-regulated DMA double-strand breaks in somatic hypermutation of immunoglobulin genes. *Nature* **408**, 216–221.

Pasqualucci, L. (and 10 others) 1998 BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc. Natl Acad. Sci. USA* **95**, 11 816–11 821.

Peng, H.-Z., Du, M., Koulis, A., Aiello, A., Dogan, A., Pan, L. & Isaacson, P. 1999 Nonimmunoglobulin gene hypermutation in germinal center B cells. *Blood* **93**, 2167–2172.

Peters, A. & Storb, U. 1996 Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* **4**, 57–65.

Phung, Q., Winter, D., Cranston, A., Tarone, R., Bohr, W., Fishel, R. & Gearhart, P. 1998 Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from

mice deficient for the MSH2 mismatch repair protein. *J. Exp. Med.* **187**, 1745–1751.

Rabbits, T., Forster, A., Hamlyn, P. & Baer, R. 1984 Effect of somatic mutation within translocated *c-myc* genes in Burkitt's lymphoma. *Nature* **309**, 592–597.

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

Shapiro, G., Aviszus, K., Ikle, D. & Wysocki, L. 1999 Predicting regional mutability in antibody V genes based solely on di- and trinucleotide sequence composition. *J. Immunol.* **163**, 259–268.

Shen, H., Peters, A., Baron, B., Zhu, X. & Storb, U. 1998 The *BCL-6* gene in normal memory B cells is mutated by the process of somatic hypermutation of immunoglobulin genes. *Science* **280**, 1750–1752.

Shen, H., Michael, N., Kim, N. & Storb, U. 2000 The TATA binding protein, c-myc, and survivin genes are not somatically hypermutated, while Ig and *BCL-6* genes are hypermutated in human memory B cells. *Int. Immunol.* **12**, 1085–1093.

Smith, D., Creadon, G., Jena, P., Portanova, J., Kotzin, B. & Wysocki, L. 1996 Di- and trinucleotide target preferences of somatic mutagenesis in normal and autoreactive B cells. *J. Immunol.* **156**, 2642–2652.

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

Storb, U., Peters, A., Klotz, E., Kim, N., Shen, H. M., Hackett, J., Rogerson, B. & Martin, T. E. 1998*a* *Cis*-acting sequences that affect somatic hypermutation of Ig genes. *Immunol. Rev.* **162**, 153–160.

Storb, U., Klotz, E., Hackett, J., Kage, K., Bozek, G. & Martin, T. E. 1998*b* A hypermutable insert in an immunoglobulin transgene contains hotspots of somatic mutation and sequences predicting highly stable structures in the RNA transcript. *J. Exp. Med.* **188**, 689–698.

Storb, U., Peters, A., Klotz, E., Kim, N., Shen, H. M., Kage, K. & Rogerson, B. 1998*c* Somatic hypermutation of immunoglobulin genes is linked to transcription. *Curr. Top. Microbiol. Immunol.* **229**, 11–19.

Tumas-Brundage, K. & Manser, T. 1997 The transcriptional promoter regulates hypermutation of the antibody heavy chain locus. *J. Exp. Med.* **185**, 239–250.

Winter, D., Phung, Q., Umar, A., Baker, S., Tarone, R., Tanaka, R., Liskay, R., Kunkel, T., Bohr, V. & Gearhart, P. 1998 Altered spectra of hypermutation in antibodies from mice deficient for the DNA mismatch repair protein PMS2. *Proc. Natl Acad. Sci. USA* **95**, 6953–6958.