Mutation of *BCL-6* Gene in Normal B Cells by the Process of Somatic Hypermutation of Ig Genes

Hong Ming Shen, Andrew Peters, Beverly Baron, Xiangdong Zhu, Ursula Storb*

Immunoglobulin (Ig) genes are hypermutated in B lymphocytes that are the precursors to memory B cells. The mutations are linked to transcription initiation, but non-Ig promoters are permissible for the mutation process; thus, other genes expressed in mutating B cells may also be subject to somatic hypermutation. Significant mutations were not observed in c-*MYC*, *S14*, or α-fetoprotein (*AFP*) genes, but *BCL-6* was highly mutated in a large proportion of memory B cells of normal individuals. The mutation pattern was similar to that of Ig genes.

Somatic hypermutation in B lymphocytes has been assumed to be restricted to Ig genes (*1*). The somatic point mutations arise between the Ig promoter and 1 to 2 kb downstream, and thus only the variable (V) region, but not the constant (C) region, of an Ig gene is affected. However, if a κ chain promoter is artificially inserted upstream of the C region, both C and V are mutated at equal frequencies (*2*). This suggests that initiation of transcription of Ig genes is required for somatic hypermutation and that the mutation domain is restricted to the 5′ end of the gene because a postulated mutator factor acts early in transcript elongation (*2*).

All Ig transgenes that were shown to mutate contain an Ig enhancer (*3*). Deletion of the enhancer impairs somatic mutation (*4*), but its replacement by a non-Ig enhancer has not been tested. Somatic hypermutation may not be Ig-specific, given that Ig genes still mutate when the Ig promoter is replaced with another promoter (*4*, *5*), non-Ig sequences can be mutated when in the context of Ig genes (*1*), and all Ig enhancers permit mutation (*6*).

To investigate this issue, we determined whether human *BCL-6*, c-MYC, ribosomal small subunit protein 14 (*S14*), and α-fetoprotein (*AFP*) genes are subject to somatic mutation (Fig. 1). Because Ig genes contain a TATA box, we selected genes driven by TATA promoters. Mutations in a translated exon may cause cell death, leading to underestimation of the rate of mutation. Therefore, genes were chosen in which the first 2 kb or so—that is, the region that is hypermutable in Ig genes—contains only untranslated exons or introns (except for the *AFP* gene). *BCL-6* and c-MYC are oncogenes that are expressed in germinal center (GC) B cells. These genes are often mutated when translocated to an Ig locus in lymphomas. *S14* is also expressed in the GC. *AFP* was chosen as a negative control because it is presumably not expressed in the GC.

Cells representing mainly memory or naive B cells were isolated from the peripheral blood of four healthy donors (*7*). DNA was prepared from isolated $CD19^+IgD^-$ cells or $CD19^+IgD^-IgM^-$ cells (mainly memory cells) and from $CD19^+IgD^+IgM^+$ B cells (mainly naive B cells) (*8*), and a polymerase chain reaction (PCR) was done to amplify either Ig heavy chain genes or the non-Ig genes (*9*). The PCR primers amplified fragments within 1.5 kb from the promoter, thus corresponding to the region of the Ig gene that is most highly mutated. The *BCL-6* gene has been reported to be mutated within the first intron in non-Hodgkin's lymphoma patients. A fragment of 790 base pairs (bp) located in the intron starting around 650 bp 3′ of the TATA box was therefore amplified. For c-MYC, the analyzed 850-bp region is often mutated in Burkitt's lymphomas and in human immunodeficiency virus–infected individuals with c-MYC translocations into the Ig locus (*10*, *11*). For both *S14* and *AFP* genes, 890-bp fragments were amplified that start around 400 bp downstream of the TATA boxes. The amplified fragments were cloned and sequenced (*12*, *13*) (Fig. 2).

Surprisingly, the *BCL-6* gene was rather highly mutated in the memory B cells from three donors, A, C, and D (Table 1). The mutation frequency varied from 6.8×10^{-4} to 1.9×10^{-3} bp, which is at least 10 times

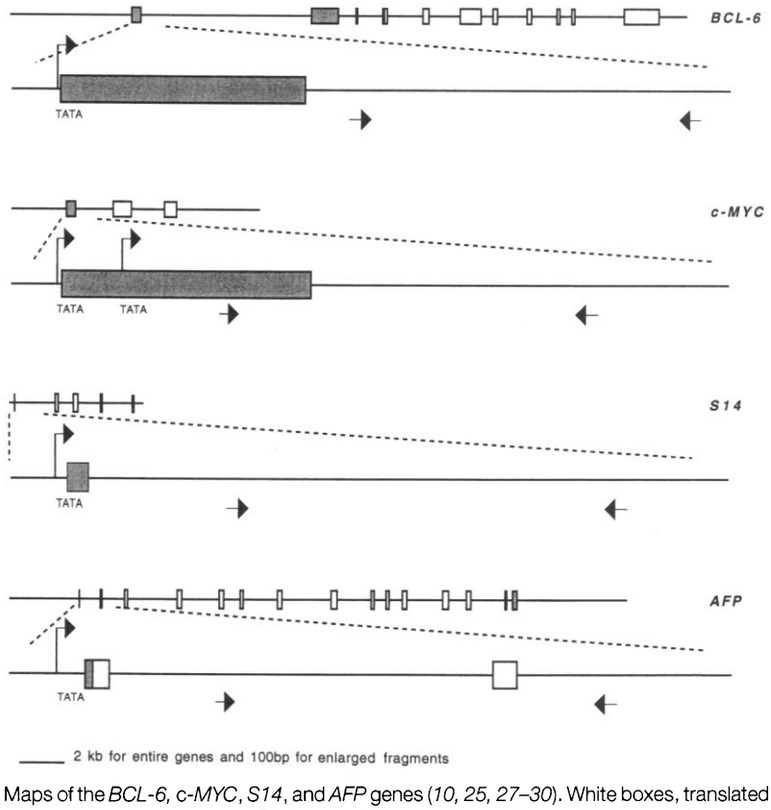

Fig. 1. Maps of the *BCL-6*, c-*MYC*, *S14*, and *AFP* genes (*10*, *25*, *27–30*). White boxes, translated exons; gray boxes, untranslated exons. (In *S14*, exons 2 to 5 are translated; in *AFP*, exon 2 is translated and exon 14 is partially translated.) Regions amplified and sequenced are indicated by arrows.

H. M. Shen and U. Storb, Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60617, USA.
A. Peters, Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60617, USA.
B. Baron, Department of Pathology, University of Chicago, Chicago, IL 60617, USA.
X. Zhu, Department of Medicine, University of Chicago, Chicago, IL 60617, USA.
*To whom correspondence should be addressed.

the mutation frequency expected as a result of PCR error (*14*) [in all four donors, some changes from GenBank sequences were observed (*15*)]. More than one-third of the amplified 790-bp sequences showed point mutations. Moreover, 9 of 26 mutated sequences had three or more mutations. For all three donors, the distribution of the somatic point mutations in *BCL-6*—mostly spread over the whole region, sometimes in small clusters—is typical for somatic hypermutation of Ig genes. Also, transition mutations were more frequent than transversions, as is observed in somatic hypermutation of Ig genes (*1*). The proportion of A's that were mutated was about twice the proportion of T's, which suggested that the mutation process has a strand bias and A/T bias, just like the Ig gene mutation process (*16*).

Although donors A, C, and D showed mutated *BCL-6* genes, the *BCL-6* sequences obtained from donor B were not mutated (Table 1). Because no IgM^- cells were seen in this donor, we isolated IgM^{dim} cells for analysis. The paucity of memory B cells in this donor's blood sample may explain the lack of mutation in *BCL-6* (see below). To avoid any possible mistake caused by a PCR artifact, we repeated the PCR reaction, cloning, and sequencing for the memory B cells of donor A. Again, similar mutation frequencies were found in this sample (Table 1).

We next determined whether mutation of *BCL-6* also occurs in naive B cells. *BCL-6* was amplified from DNA from $CD19^+IgD^+IgM^+$ naive B cells of donor A. In 15,820 bp, only one deletion of a single base pair, but no point mutations, were found (Table 1). Thus, somatic mutations in *BCL-6* appear to be restricted to memory B cells.

Because somatic hypermutation of Ig genes occurs only in the first 1 to 2 kb downstream from the promoter, we searched for mutations more 3′ in the *BCL-6* gene. No mutations were found between 2.44 and 3.24 kb downstream of the promoter in memory B cells from donor A (Table 1), which suggests that the distribution of mutations in the *BCL-6* gene is comparable to that of the Ig genes.

The finding that *BCL-6* is mutated in memory B cells prompted us to analyze other non-Ig genes, including c-MYC, *S14*, and *AFP* (Table 1). We selected the region of the c-MYC gene that is highly mutated in lymphomas with 8;14 translocations, and found no significant mutation in this region from donor A's memory B cells. Two mutations were observed in 20,000 bp, which is not significantly different from the proportion of Pfu errors (*14*). Similarly, no mutations were seen in *S14*.

To determine whether an untranscribed gene was mutable, we sequenced the *AFP* gene, which is only expressed early in development and in some tumors. Because this gene is presumably not expressed in B cells, mutations in the exon in this region would not damage B cells. No mutations were observed (Table 1).

Clearly, *BCL-6* is a target for relatively extensive somatic hypermutation. The mutations were found only in the 5′ end of the gene, not in a more distant region.

Table 1. Somatic mutation in non-Ig genes. Two female donors were 42 (A) and 55 (B) years old, and two male donors were 44 (C) and 47 (D) years old.

Gene	Donor	B cell	Base pairs	Mutations	Frequency	Total sequences	Sequences with mutation	Sequences mutated (%)
Significant mutation of BCL-6 gene 0.64 to 1.43 kb downstream of the promoter in human memory B cells								
BCL-6	A	IgD^-	18,984	13	6.8×10^{-4}	24	9	38
*BCL-6**	A	IgD^-	7,910	7	8.8×10^{-4}	10	3	30
BCL-6	B	IgD^-IgM^{lo}	15,820	0	0	20	0	0
BCL-6	C	IgD^-IgM^-	15,029	29	1.9×10^{-3}	19	8	42
BCL-6	D	IgD^-IgM^-	14,238	21	1.5×10^{-3}	18	6	33
No mutation of BCL-6 gene 0.64 to 1.43 kb downstream of the promoter in human naive B cells								
BCL-6	A	IgD^+IgM^+	15,820	0	0	20	0	0
No mutation of BCL-6 gene 2.44 to 3.24 kb downstream of the promoter in human memory B cells								
BCL-6	A	IgD^-IgM^-	13,600	0	0	17	0	0
No significant mutation of c-MYC, S14, and AFP genes in human memory B cells								
MYC	A	IgD^-	20,352	2	9.8×10^{-5}	24	2	8.3
S14	A	IgD^-	17,900	0	0	20	0	0
AFP	A	IgD^-	16,986	0	0	19	0	0

*Second PCR was carried out to confirm the mutation in the *BCL-6* gene.

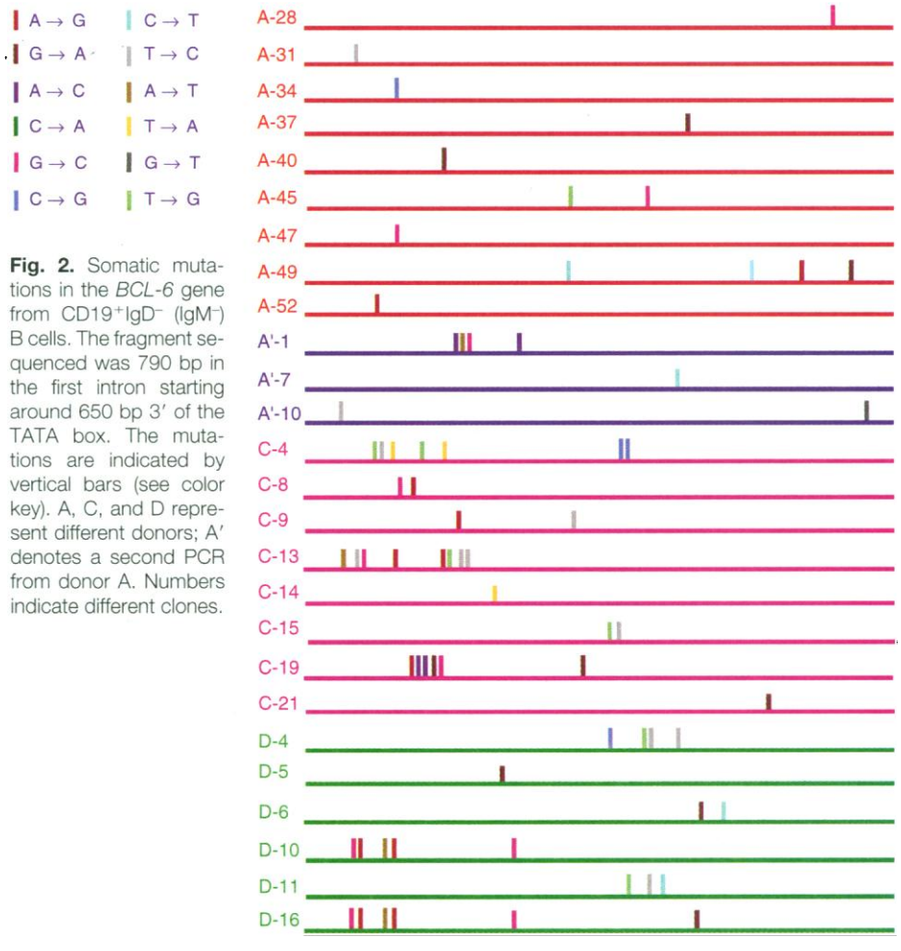

Fig. 2. Somatic mutations in the *BCL-6* gene from $CD19^+IgD^-$ (IgM^-) B cells. The fragment sequenced was 790 bp in the first intron starting around 650 bp 3′ of the TATA box. The mutations are indicated by vertical bars (see color key). A, C, and D represent different donors; A′ denotes a second PCR from donor A. Numbers indicate different clones.

They were found only in memory B cells, not in naive B cells that had not yet undergone somatic hypermutation of Ig genes. The fine specificity of the mutations with respect to transitions or transversions, A/T bias, and strand bias was the same as for Ig gene mutations. It is therefore likely that mutations in the *BCL-6* gene resulted from the same process that gives rise to somatic hypermutation of Ig genes. In the memory B cell sample of human donor A, we found a mutation frequency of 5×10^{-2} in Ig heavy chain genes (*17*). The mutation frequency was lower in the *BCL-6* genes of donors A, C, and D than in the Ig genes of donor A by one to two orders of magnitude. The difference in mutability of Ig and *BCL-6* genes may be attributable to differences in transcription rates or qualitative differences in control elements that target a postulated mutator factor (*1*).

BCL-6 gene mutations had been observed previously in B lymphomas of the GC/post-GC type (*18*). In the tumors with *BCL-6* mutations, the *BCL-6* gene is often translocated into an Ig locus. In those cases, the translocated *BCL-6* gene has come under the control of Ig enhancers, and the somatic mutations could be explained as a direct result of the mutation process that occurs at the endogenous Ig locus. This would be similar to the mutations in c-*MYC* genes translocated into the Ig locus in Burkitt's lymphomas (*19*). However, in diffuse large cell lymphomas and follicular lymphomas, nontranslocated *BCL-6* gene can also be mutated (*20*). Although translocations distant from the *BCL-6* gene may not have been detected, our study offers an explanation for the origin of these mutations. It is unlikely that in our analysis, all three randomly chosen healthy donors with *BCL-6* mutations had translocations of *BCL-6*, although this has not been ruled out. It appears more likely that the endogenous *BCL-6* gene is a real target for the somatic hypermutation process in normal B lymphocytes. We postulate that this is the case for most of the human population. It is therefore also likely that the mutations seen in presumably untranslocated *BCL-6* genes in B cell lymphomas were present in the B cell precursors before malignant transformation. The frequency of mutations of the translocated *BCL-6* genes appears to be similar to that of the untranslocated *BCL-6* genes seen in tumors (*20*) and in normal B cells (this report). It is thus possible that after the translocation event, B cells do not reenter GCs and acquire additional mutations, because the mutation frequency under the control of the Ig locus would be expected to be higher than in the normal *BCL-6* gene.

The absence of mutations in *S14* (expressed in all cells) and in c-MYC (expressed in proliferating cells, including activated B cells) (*21*, *22*) indicates that not all expressed genes become targets for somatic mutation. It is not currently possible to determine with certainty whether the mutability of *BCL-6* may be a result of higher transcription rates relative to the other non-Ig genes. However, this seems not to be true for the *S14* gene. In mouse myeloma cells, the expression of ribosomal protein genes was found to be about one-tenth that of Ig genes (*21*), whereas the rate of Ig gene transcription in myeloma cells (as well as in normal plasma cells) is about 10 times that in B cells. Thus, in B cells undergoing somatic hypermutation, the transcription rates of *S14* and Ig genes can be assumed to be similar. It is thus more likely that the *BCL-6* gene has certain cis-acting elements shared with Ig genes that are essential for targeting the mutations, and that the *S14* and c-MYC genes lack these elements. Most likely, such elements as are essential to target a postulated mutator factor (*23*) are present in the Ig enhancers. The challenge now is to define the responsible cis-acting elements, and to determine the extent of somatic hypermutation of *BCL-6* in the human population and its possible role in oncogenesis.

REFERENCES AND NOTES

1. U. Storb, in *The Biology of Germinal Centers in Lymphoid Tissues*, G. J. T. and V. K. Tsiagbe, Eds. (Landes, Austin, TX, in press).
2. A. Peters and U. Storb, *Immunity* **4**, 57 (1996).
3. U. Storb, *Curr. Opin. Immunol.* **8**, 206 (1996).
4. A. Betz *et al.*, *Cell* **77**, 239 (1994).
5. K. Tumas-Brundage and T. Manser, *J. Exp. Med.* **185**, 239 (1997).
6. E. Klotz and U. Storb, *J. Immunol.* **157**, 4458 (1996).
7. Peripheral blood mononuclear cells (PBMCs) were purified with Percoll (Sigma). $CD19^+$ B cells were isolated with Dynabeads coated with mouse monoclonal antibody to human CD19 according to the manufacturer's instructions (Dynal, Lake Success, NY). $CD19^+$ cells were stained with fluorescein isothiocyanate (FITC)–conjugated goat antibody to human IgD (Southern Biotech, Birmingham, AL), either alone or in combination with phycoerythrin-conjugated goat antibody to human IgM (Sigma). $CD19^+IgD^-$, $CD19^+IgD^-IgM^-$, and $CD19^+IgD^+IgM^+$ B cells were sorted by single- or dual-color flow cytometry (EPICS ELITE ESP; Coulter, Hialeah, FL). The purity of the cells of interest was >97%. IgD^- B cells (donor A, Table 1) are likely to contain some IgM memory B cells, which have generally fewer mutations than IgG B cells (*24*); IgM^+IgD^+ B cells (Table 1) are unlikely to have undergone somatic mutation.
8. B cells were pelleted and resuspended, and one volume of 2× DNA lysis buffer [proteinase K (1 mg/ml), 100 mM EDTA (pH 8.0), 2% SDS, and 100 mM tris (pH 8.0)] was added. The lysate was incubated at 37°C overnight. DNAs were extracted twice with phenol and then twice with phenol/chloroform/isoamyl alcohol (25:24:1). DNAs were precipitated with ethanol, pelleted, and dissolved in distilled water.
9. We used the following primers: *BCL-6* gene primers, 5′-GC*TCTAGA*CCGCTGCTCATGATCATTATTT(sense), 5′-CGG*GGTACC*TAGACACGATACTTCATCTCAT (antisense) or 5′-CCGCTGCTCATGATCATTATTT (sense), 5′-TAGACACGATACTTCATCTCAT (antisense); c-*MYC* gene primers, 5′-CCG*GTACC*CTTGCCGCATCCACGAAACTTT (sense), 5′-GC*TCTAGA*GACCCAGGTTTTA (antisense); *S14* gene primers, 5′-CGG*GGTACC*CGGGACAGACGTGGGCTCCCG (sense), GC*TCTAGA*TCTAAGGGAGAGAGAAACTGAC (antisense); *AFP* gene primers, 5′-GGATGAATGGTTTGTATGTTTC (sense), GGTTTGACTCATGAGATTTC (antisense). PCR conditions for *BCL-6*, c-*MYC*, and *AFP* were 94°C for 5 min, 57°C for 30 s, 75°C for 1 min, 1 cycle; 94°C for 30 s, 57°C for 30 s, 75°C for 1 min, 29 cycles; and 75°C for 6 min, 1 cycle. PCR conditions for *S14* were 95°C for 5 min, 67°C for 30 s, 75°C for 1 min, 1 cycle; 95°C for 30 s, 67°C for 30 s, 75°C for 1 min, 29 cycles; and 75°C for 6 min, 1 cycle.
10. R. Taub *et al.*, *Cell* **36**, 339 (1984).
11. J. Mueller, S. Janz, J. Goedert, M. Potter, C. Rabkin, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6577 (1995).
12. *BCL-6* DNA fragments from memory B cells were inserted into pBluescript KS at Kpn I and Xba I sites, and the other gene fragments and *BCL-6* from naive B cells were inserted into the Srf I site in pBluescript SK according to manufacturer's instructions (Strategene). Plasmids with the inserts were transformed into *Escherichia coli* JH3.
13. Automated sequencing was done for all amplified gene fragments. Manual sequencing was performed to confirm some mutations. Sequences were aligned with the data from GenBank and (*25*) by MacVector4.14 to find mutations.
14. The high-fidelity polymerase, Pfu, was used to amplify genomic DNA; its error rate was reported to be one mutation per 7.8×10^5 bp per cycle (*26*). Thus, for 30 cycles, roughly one mutation in 26,000 bp would be expected.
15. Compared with the GenBank sequences, many changes were seen in *BCL-6* in all four donors and in *S14* donor A.
16. U. Storb *et al.*, *Curr. Top. Microbiol. Immunol.* **229**, 11 (1998).
17. Nine rearranged V_H4 family genes were sequenced from positions 130 to 296 of V. Seven of the nine genes were mutated with 5 to 24 mutations per 167 nucleotides.
18. A. Migliazza *et al.*, *Blood* **90** (suppl. 1, pt. 1), 177a (1997).
19. T. Rabbits, A. Forster, P. Hamlyn, R. Baer, *Nature* **309**, 592 (1984).
20. A. Migliazza *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12520 (1995).
21. N. Hariharan, D. Kelley, R. Perry, *Genes Dev.* **3**, 1789 (1989).
22. J. Newton *et al.*, *Eur. J. Immunol.* **26**, 811 (1996).
23. U. Storb *et al.*, in preparation.
24. U. Klein, R. Kueppers, K. Rajewsky, *J. Exp. Med.* **180**, 1383 (1994).
25. F. Bernardin *et al.*, *Oncogene* **14**, 849 (1997).
26. J. Cline, J. Braman, H. Hogrefe, *Nucleic Acids Res.* **24**, 3546 (1996).
27. D. D. Rhoads, A. Dixit, D. J. Roufa, *Mol. Cell. Biol.* **6**, 2774 (1986).
28. P. E. M. Gibbs, R. Zielinski, C. Boyd, A. Dugaiczyk, *Biochemistry* **26**, 1332 (1987).
29. W. W. Colby, E. Y. Chen, D. H. Smith, A. D. Levinson, *Nature* **301**, 722 (1983).
30. C. Gazin *et al.*, *EMBO J.* **3**, 383 (1984).

31. Supported by NIH grant GM38649 to U.S. A.P. was partly supported by NIH training grant GM07183. We thank T. McKeithan, T. E. Martin, N. Michael, and P. Engler for critical reading of the manuscript.

11 March 1998; accepted 28 April 1998