TLR Ligation Triggers Somatic Hypermutation in Transitional B Cells Inducing the Generation of IgM Memory B Cells

Alaitz Aranburu, Sara Ceccarelli, Ezio Giorda, Rosa Lasorella, Giovanna Ballatore and Rita Carsetti

*J Immunol* published online 15 November 2010
http://www.jimmunol.org/content/early/2010/11/14/jimmunol.1002722

Supplementary Material
http://www.jimmunol.org/content/suppl/2010/11/15/jimmunol.1002722.2.DC1

Subscription
Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

Permissions
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
TLR Ligation Triggers Somatic Hypermutation in Transitional B Cells Inducing the Generation of IgM Memory B Cells

Alaitz Aranburu,* Sara Ceccarelli,* Ezio Giorda,* Rosa Lasorella,† Giovanna Ballatore, † and Rita Carsetti*

TLR9 activation by unmethylated CpG provides a homeostatic mechanism to maintain B cell memory in the absence of Ag. In this study, we demonstrate that CpG also triggers the generation of somatically mutated memory B cells from immature transitional B cells. In response to CpG, a fraction of transitional B cells proliferates and introduces somatic hypermutations in the H chain V regions. The nonproliferating pool of transitional B cells mostly maintains germline configurations. Mutations are VH specific: VH5 is the least mutated family, whereas VH1 and VH4/6 are the most mutated families. CpG stimulation also results in upregulation of VH5 transcripts in proliferating cells. Therefore, early recognition of bacterial DNA preferentially expands VH5-expressing B cells while inducing somatic hypermutations in other families. The mutation frequency, range, and type of substitutions observed in vitro are comparable to those found in memory B cells from the peripheral blood of Hyper IgM type 1 patients and the spleen of normal infants. The process triggered by TLRs may represent a first step leading to additional diversification of the germline repertoire and to the generation of memory B cells that will further refine their repertoire and specificity in the germinal centers. The Journal of Immunology, 2010, 185: 000–000.

The formation of a broad Ag receptor repertoire by DNA rearrangements and mutations is a unique property of B lymphocytes. DNA recombination is complex and follows an orderly multistep process enabling B cells to express unlimited Ag receptors. Successful V(D)J rearrangements allow B cells to start expressing IgM on their surface and to migrate to secondary lymphoid organs where they become activated by Ag encounter. Upon activation, B cells build up germinal centers (GCs) that are transient reactive lymphoid structures, where they diversify their IgRs. The molecular process that modifies Ag receptors is called somatic hypermutation (SHM) (1), and it operates by introducing nucleotide base substitutions within the Ig genes. The major player in this process was cloned 10 y ago; it was named AID (protein coded by AICDA) because of its intrinsic deamination factors belong to the base excision repair and mismatch repair pathways, whose primary function is to safeguard the integrity of the genome from mutagenesis. During the course of affinity maturation of Ig genes, these enzymes are diverted from high-fidelity DNA repair.

The defining characteristics of memory B cells are the ability to respond rapidly upon reencounter with cognate Ags and the footprints of SHM in their Ag receptors. In humans, two main types of memory B cells have been described: classical isotype-switched memory and the so-called “IgM memory” B cells (9). The former are the products of GC reactions generated during T cell-dependent immune responses. Conversely, IgM memory B cells are generated in T cell-independent reactions and do not seem to require GC formation (9, 10). Recently, these cells were suggested to be the recirculating splenic marginal zone B cells, based on surface-molecule and gene-expression analysis (11). Accordingly, IgM memory B cells are extremely reduced in splenectomized and asplenic patients (12). In normal physiology, IgM memory B cells collect in the spleen, an optimal anatomical site where blood-borne pathogens and cellular debris are filtered by resident macrophages. Thus, we hypothesized that the progenitors of IgM memory B cells may be the transitional B cells, recent bone marrow emigrants that are transported by the
blood to the spleen (13). In this article, we propose that transitional B cells diversify their Ag receptors in the presence of bacterial components, exemplified by bacterial CpG DNA in our experimental approach. We demonstrated that, upon exposure to CpG in vitro, a fraction of the transitional B cells proliferate and differentiate into IgM-secreting plasma cells or into cells phenotypically identical to IgM memory B cells. The remaining nonproliferating (NP) B cells downregulate the expression of CD24 and CD38, thus resembling mature B cells (14). Based on these observations, we asked whether the population of B cells generated in vitro by CpG, which we call natural “memory” cells, has the hallmark of memory cells (i.e., carries somatic mutations). We show that mutations are introduced as a result of CpG stimulation. Mutations are found mostly in VH sequences from proliferating (P) cells compared with NP cells. Mutation does not occur in a random fashion but follows a distinct distribution, as judged by the mutational weight on the gene families studied (VH1, VH3, VH4/6, and VH5). Specifically, VH1 is the most mutated family, and VH5 is the least mutated family. Finally, we observed that CpG preferentially induces the expansion of VH5 transcripts. In view of our results, we reason that an unknown mechanism triggered by CpG stimulation selects and expands a cohort of unmutated and uniquely rearranged VH5-expressing B cell clones, while introducing mutations in the rest of the families.

Materials and Methods
Flow-cytometry analysis and cell sorting
After donation, placental cord blood (CB) mononuclear cells (Ospedale Sant’Eugenio) were isolated by density-gradient centrifugation using Histopaque-1077 (Sigma-Aldrich, St. Louis, MO). To sort unstимulated transitional B cells, single-cell suspensions were stained with Abs against CD24 and CD38 (BD Biosciences, San Jose, CA) and sorted using FACS Vantage SE (BD Biosciences, San Jose, CA). After donation, placental cord blood (CB) mononuclear cells (Ospedale Sant’Eugenio) were isolated by density-gradient centrifugation using Histopaque-1077 (Sigma-Aldrich, St. Louis, MO). To sort unstimulated transitional B cells, single-cell suspensions were stained with Abs against CD24 and CD38 (BD Biosciences, San Jose, CA) and sorted using FACS Vantage SE (BD Biosciences, San Jose, CA). After donation, placental cord blood (CB) mononuclear cells (Ospedale Sant’Eugenio) were isolated by density-gradient centrifugation using Histopaque-1077 (Sigma-Aldrich, St. Louis, MO). To sort unstimulated transitional B cells, single-cell suspensions were stained with Abs against CD24 and CD38 (BD Biosciences, San Jose, CA) and sorted using FACS Vantage SE (BD Biosciences, San Jose, CA). After the addition of 2.5 μg/ml CpG-C oligodeoxynucleotide (Hycult Biotechnology, Uden, The Netherlands), at a cell density of 2.5 × 10^6 cells/ml in complete RPMI 1640 (InvivoGen) supplemented with 10% FBS (HyClone Laboratories, Logan, UT) (14). After 7 d in culture, cells were collected by centrifugation, washed, stained, and sorted. P and NP cells were characterized by the loss or presence of CMFDA staining, respectively. For the anti-CpG + IL-4 stimulations, PBMCs from healthy donors were stimulated for 5 d with 1.5 μg/ml anti-CD40 (PeproTech, Rocky Hill, NJ) and 100 U/ml IL-4 (Calbiochem, Darmstadt, Germany), followed by sorting of CD19^b and CD19^d cells.

CpG oligodeoxynucleotide and anti-CD40 + IL-4 stimulations
CB-derived mononuclear cells were incubated with 100 ng/ml CMFDA at a cell density of 2 × 10^6 cells/ml at 37°C for 20 min, followed by the addition of 2.5 μg/ml CpG-C oligodeoxynucleotide (Hycult Biotechnology, Uden, The Netherlands), at a cell density of 2.5 × 10^6 cells/ml in complete RPMI 1640 (InvivoGen) supplemented with 10% FBS (HyClone Laboratories, Logan, UT) (14). After 7 d in culture, cells were collected by centrifugation, washed, stained, and sorted. P and NP cells were characterized by the loss or presence of CMFDA staining, respectively. For the anti-CpG + IL-4 stimulations, PBMCs from healthy donors were stimulated for 5 d with 1.5 μg/ml anti-CD40 (PeproTech, Rocky Hill, NJ) and 100 U/ml IL-4 (Calbiochem, Darmstadt, Germany), followed by sorting of CD19^b and CD19^d cells.

RNA extraction and cDNA reaction
RNA from various cell types was isolated using the RNeasy Plus Mini Kit (Qiagen, Hilden, Germany), and cDNA reactions were made using random hexamers (1st strand cDNA synthesis kit; Roche Diagnostics, Mannheim, Germany).

Cloning and sequencing of PCR products
VH gene families and the constant CH4 region were amplified with Expand High Fidelity PCR System (Roche Diagnostics), and family-specific forward primer sequences for L-VH1, L-VH3, L-VH4/6, or L-VH5 was used with a-specific reverse primer (5′ C′ CH1). All of the primers are described in Supplemental Table I. PCR products were amplified using a DNA Engine Dyad thermal cycler (Bio-Rad, Hercules, CA), separated on agarose gels and purified using Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI), and subcloned using Zero Blunt TOPO PCR cloning kit (Invitrogen, Carlsbad, CA), following instructions provided by the manufacturer. Subsequently, minipreps were made of each individual clone (Qiagen Spin Miniprep Kit; Qiagen), and inserts were sequenced at Bio-Fab Research (Pomezia, Rome, Italy).

Mutation and VH-DH-JH gene segment analysis
Using the international ImMunoGeneTics information system (15) (http://www.imgt.org), obtained sequences were compared with germline sequences with the V-QUEST tool (16). The first nine nucleotides from each V region were excluded from the analysis because this region overlapped partially with forward primer sequences. We scrutinized every clone generated in our experiments and systematically excluded unproductive and duplicated rearrangements from our analysis (Supplemental Table II).

Semiquantitative RT-PCR
Serial dilutions of cDNAs from transitional, P, and NP cells were prepared, and their contents were normalized to the housekeeping gene Hprt1. Selected dilutions were used to amplify VH-Cμ transcripts, using the primer sequences described in Supplemental Table I. PCR products were amplified using a DNA Engine Dyad thermal cycler (Bio-Rad) and separated on agarose gels stained with GelRed (Biotium, Hayward, CA), and images were acquired using Fluor-S Multimager MAX and quantified by densitometry using Quantity One 4.1.1 (Bio-Rad). Similarly, serial dilutions of cDNAs from sorted resting and activated CD19^b and CD19^d cells were prepared and used as controls for the expression of AICDA, UNG2, POLH, MS12, and MSH6 (Supplemental Table I).

Statistical analysis
GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA) was used for statistical analyses. Two groups were compared in a 2 × 2 contingency table with the two-tailed χ² test with Yates correction; p values ≤0.05 were considered significant.

Results
CpG induces SHMs
We demonstrated previously that CpG, the ligand of TLR9, is the only known stimulus able to induce transitional B cells to proliferate and produce Abs in vitro. Because proliferation was associated with the acquisition of the phenotype typical of IgM memory B cells (14), we addressed the question of whether transitional B cells acquire somatically mutated IgRs following TLR9 stimulation. We sorted P and NP B cells generated from CB-derived transitional B cells (Supplemental Fig. 1A). To study the occurrence of somatic mutations, the VH families from P and NP cells were amplified by PCR, subcloned, and the sequences were analyzed. Although our work generated data sets of respectable size, we are aware that PCR amplification might introduce artifacts. Hence, following a similar procedure, we sorted transitional B cells derived from CB (Supplemental Fig. 1B) and subcloned their Ig VH genes to establish the lower baseline of mutation frequency (n = 100, 0.06%; Fig. 1A) (17). Likewise, the PCR error was estimated by sequencing clones that express β-actin (protein coded by ACTB) using the same method (n = 20, 0.036%). Our results demonstrated that CpG induces somatic mutations at a frequency comparable to that of IgM memory B cells in the spleen of children (18). These mutations are more frequent in the VH regions of the P population (n = 188) compared with the NP one (n = 68), which mainly displays germline configurations (Fig. 1A, Supplemental Table II). Moreover, the incidence of mutations is VH specific; the mutation frequency of VH1 (0.37%; p = 0.0006) is the highest, followed by that of VH4/6 (0.27%; p = 0.0008). Mutations in VH3 were not statistically significant (0.21%; p = 0.06) (Fig. 1A). Remarkably, the VH5 gene family, which was the most expanded in the P pool, showed the least number of mutations (0.04%; p = 0.77), coinciding with those present in VH genes from sorted immature transitional B cells (Fig. 1A). The mechanism of SHM targets the VH region of Ig genes where mutations introduce changes in the Ag-binding site. No mutations are inserted in CH regions to preserve
Ig function. It was necessary to investigate whether SHM induced in vitro by CpG also respects this rule, to exclude that the phenomenon observed in vitro was not an artifact due to a deregulation of chromatin structure and gene expression created by our culture condition. Therefore, we analyzed the Ig CH4 region of 33 clones from P cells and 28 clones from transitional CB B cells. Our results showed that this region does not acquire mutations after CpG stimulation, because the mutation frequency in the CH4 region of CB or P transitional B cells was similar to that observed in the β-actin control (0.036%) (Supplemental Table II).

The majority of mutations are confined to the framework regions (FRs), the scaffolding units of the Ig molecule that help CDRs adjust for Ag binding (Fig. 1B). Replacement (R) mutations are enriched over silent (S) mutations in VH1-expressing (36:18, R/S ratio), VH4/6-expressing (19:12, R/S ratio), and VH5-expressing (7:1, R/S ratio) P cells (Fig. 1C). In contrast, VH3-expressing P cells display mainly silent replacements (6:15, R/S ratio), specifically localized to FR1 and FR3 (Fig. 1C, data not shown). The overall R/S ratios (FR = 1.18, 47:40; CDR = 3.5, 21:6) correspond to values found in healthy individuals (FR = 1.7; CDR = 3.4) (19). Thus, we conclude that exposure to CpG prompts early transitional B cells to initiate a mutational process to modify their IgRs. The mutational machinery seems to target VH families differently, judged by the disparate mutation frequencies observed. Furthermore, FR regions seem to be preferentially targeted; CDR1 is mutated only in VH1- and VH4/6-expressing cells. Also, replacement mutations are accumulated.

**FIGURE 1.** CpG induces SHM in Ig VH genes. A, Mutation frequency displayed by transitional B cells, P cells, and NP cells. B, Mutation frequency (%) within discrete FR and CDR regions in P cells. C, The percentage of R and S mutations is shown in P cells for all families studied. Also, the distribution of R and S mutations throughout FRs and CDRs is shown in all families. The mutation frequency (%) was calculated \(\frac{\text{mutations}}{\text{nucleotides}} \times 100\) for each VH gene family studied. The results shown were collected from independent experiments generated from three diverse CB samples. *Significant difference in the mutation frequency for VH1 \((p = 0.0006)\) and VH4/6 \((p = 0.0008)\).
AICDA, UNG2, POLH, MSH2, and MSH6 transcripts in P cells

The SHM process is currently viewed as occurring in two phases (20). The first phase of mutation is defined by the deamination of C (deoxyadenosine) residues to U (deoxyuridine) residues by AID, thus generating a DNA lesion. As a result, the G (deoxyguanosine) to U mispairing (U:G) is replicated as such, giving rise to transition mutations of C-to-T type, and/or the uracil residue is removed by an uracil glycosylase belonging to the base-excision repair DNA-repair pathway. The second phase of mutation seems to be largely restricted to A:T pairs. Although the mechanism is still unknown, a number of important phase 2 contributors have been described, such as polymerase η (21). To verify whether the mutations detected in transitional B cells after CpG stimulation are introduced by the same mechanism acting in the GC, we measured the expression of the elements of the SHM machinery in P and NP B cells (Fig. 2). We used sorted CD19+ and CD19− cells as controls, which were stimulated with anti-CD40 and IL-4. As expected, AICDA and UNG2 were only detected in stimulated CD19+ B cells. The constitutively expressed DNA repair enzymes POLH, MSH2, and MSH6 were upregulated after stimulation. We then compared P and NP transitional B cells (Fig. 2). We found that AICDA and UNG2 were present in both cell types, but their expression was higher in P B cells, consistent with published data (14, 22). Conversely, POLH expression was greater in NP cells than in P cells. As for the DNA repair proteins MSH2 and MSH6, they are expressed in both cell types; however, they are slightly more abundant in NP cells in agreement with POLH expression. Therefore, CpG sustains the expression of basic DNA-repair systems and upregulates the transcription of AID and UNG2, the two main enzymes of SHM.

Transitions prevail over transversions

We also examined the type of mutations induced by CpG in more detail and found that, collectively, transition substitutions dominate over transversions in P cells (80:35) (Fig. 3A). Our observation extends also into single families, displaying distributions of 36:18 for VH1, 13:8 for VH3, 24:8 for VH4/6, and 7:1 for VH5 (Fig. 3B). In contrast, the NP population, with considerably lower mutation frequencies, showed a modest predisposition toward transitions (11:6) (Fig. 3A). There are four types of transitions, all induced by CpG stimulation: G-to-A (17/80, 21.2%), C-to-T (11/80, 13.7%), A-to-G (28/80, 35%), and T-to-C (24/80, 30%) (Fig. 3C). We believe that this finding is highly significant in view of the fact that increased mutations on RGYW (underlined residue is mutated) motifs represent the hallmark of SHM (23), as well as the site for AID deamination at G:C pairs (24) during the first phase of SHM (25). Yet, it is remarkable that transversions are less frequent in both cell types, as it would be expected during the course of a conventional phase 1 extension (25). Notwithstanding the marked differences in mutation frequencies that we observed between P and NP cells, the prevalence of mutations off G:C (P, 41% and NP, 47%) and off A:T pairs (P, 59% and NP, 53%) was comparable in both populations (Fig. 3C), as well as coincident with established percentages published by other investigators (26). In addition, the incidence of G-to-A and C-to-T transitions (P, 35% versus NP, 36.4%), as well as A-to-G and T-to-C transitions (P, 65% versus NP, 63.6%), were comparable in both populations (Fig. 3C). Collectively, transition substitutions predominate for the most part in P cells. However, it is noticeable that although absolute mutation counts are greater in the P population, the distribution of the percentage of these is similar in both cell types, which is suggestive of a common mechanism operating in both populations, albeit at diverse intensities.

RGYW/WRCY and WA/TW hotspot motifs are targeted upon CpG stimulation

We thoroughly looked for the hotspot motif RGYW (or its complementary WRCY) in our pool of mutated VH sequences and specified exact matches, as well as motifs bearing one or two mismatches from the conserved consensus sequence. Although exact match sequences were observed (5/45, 11%), the majority of mutated G:C residues were positioned within an imperfect RGYW/WRCY motif carrying one mismatch (14/45, 31%) or two mismatches (19/45, 42%) from the classical motif (Supplemental Table III). However, when generating sequence alignments to other sequence motifs described in the literature (23), the relative abundance of sequences bearing exact matches (7/45, 15.5%) and one mismatch (26/45, 57.8%) increases with the expected reduction in occurrence/unavailability of a cofactor (19, 27). This might illustrate a rudimentary form of SHM that CpG alone initiates in transitional B cells and that is subsequently perfected with other signals or stimuli.

It is now well established in the literature that mutations occurring on WA/TW (W = A or T) motifs are due to the activity of polymerase η (28), an error-prone trans-lesion DNA polymerase associated with mismatch repair proteins. Our mutation analysis on A:T base pairs shows that A residues are only slightly more mutated (36/66) than are T residues (30/66) in P cells (Supplemental Table III). In addition, half of these mutations occur on WA/TW hotspot motifs (20/36). This finding is consistent with the literature that CpG sustains the expression of the basic DNA-repair system and upregulates the transcription of AID and UNG2, the two main enzymes of SHM.

FIGURE 2. CpG stimulation activates the SHM machinery in CpG-stimulated CB cultures. A, RT-PCR in sorted control B cell subsets known to induce the expression of AICDA and UNG2 following stimulation with anti-CD40 and IL-4. The expression of other important players in SHM was upregulated upon activation. B, Semiquantitative RT-PCR in sorted P and NP B cell subsets. The expression patterns of UNG2, AICDA, POLH, MSH2, and MSH6 are markedly different in P and NP cells. Serially diluted cDNAs were used as template. cDNAs were normalized for the expression of the housekeeping gene HPT1.
motifs (35/66) that were previously described (28), as well as in other noncanonical sequence motifs, of which GAG (10/31) and CGT (8/31) were the most abundantly used, followed by CAG (6/31), GTG (3/31/GTC (3/31), and GAC motifs used once only (Supplemental Table III). Similarly, we observed A:T directed mutations in NP cells, being the mutations from A and from T comparable (5:4). As for the conservation of the WA/TW motif, the majority of substitutions occurred on conserved sites (7/9), with only two nonclassical sequence motifs, GAG and CGT, as described above for the P cells.

CpG induces VH5 family selection in P cells

When stimulated with CpG, transitional B cells undergo SHM in their rearranged Ig H chains, resulting in distinct and, apparently, nonrandom mutation burdens on individual VH families. To further explore the effect CpG had on transitional B cells, we asked whether CpG specifically influenced the usage of the preimmune VH repertoire, by selecting or disfavoring the use of certain VH families. To attain our goal, we used the same cDNAs derived from P and NP cells, as well as unstimulated CB cells, in a semi-quantitative RT-PCR assay using VH family-specific combinations.
of primers. cDNAs were normalized for HPRT1 before quantification. Importantly, amplification of VHx \((x = 1, 3, 4/6, 5)\) products from CB cells determined the lower baseline of expression for all VH families prior to CpG stimulation. Hence, results were obtained by agarose gel densitometry and depicted as ratios referring to corresponding VHx \((x = 1, 3, 4/6, 5)\) values in CB cells. Transitional B cells express abundant VH3, followed by similar expression of VH1 and VH4/6, with modest VH5 expression at the lowest end of the spectra, revealing dissimilar, yet conserved, relative-expression patterns (Figs. 4A, 5) (17). After CpG stimulation, we observed that the pool of NP cells did not upregulate their IgRs to the extent that the P counterparts did, resembling more the VH family distribution displayed by control unstimulated CB cells, albeit at moderately increased ratios (Fig. 4B). In addition, the P cells seem to have largely upregulated most VH families, with the striking expansion of VH5-expressing B cell clones. Hence, the initial repertoire of VH genes displayed by transitional B cells is modulated by CpG, giving rise to two diverse populations of cells, P and NP, each displaying a diverse distribution of amplified VH genes. The fact that CpG preferentially induces VH5-bearing B cells to expand/increase transcription and VH1 and VH4/6 B cells to mutate is, to our knowledge, a novel phenomenon lacking a defined mechanistic basis. The observation is suggestive of a crosstalk between the signaling cascades from the BCR and TLR9.

**CpG selects IgH CDR3 lengths typically used by IgM memory B cells**

To address further the role of CpG in selecting among P and NP cells, we studied the IgH CDR3 length of every unique B cell clone derived from our screening. Although amino acid sequence alone cannot be used as a predictive marker for self-reactive Abs, long IgH CDR3s have been associated with self-reactive or polyclonal Abs (29). In addition, early B cell progenitors were reported to express longer CDR3 lengths compared with peripheral B cells (30, 31). Furthermore, it was shown that as B cells progress through their developmental program, they acquire shorter CDR3 lengths (32). Consistent with these observations, our results demonstrated that immature transitional B cells, on average, consist partly of cells with CDR3 lengths >21 aa (7%) that become moderately shortened as a result of CpG stimulation in P (6%) and NP cells (2%) (Fig. 6A). Moreover, cells with CDR3 lengths that fall into the 16–20 aa range are more frequent in transitional B cells (40%) than in CpG-stimulated P (21%) and NP (32%) cells. Conversely, medium-sized CDR3s (11–15 aa) are favored in P (59%) and NP (60%) cells as opposed to transitional B cells that have not received CpG (48%) (Fig. 6A). Also, our results on the average length of CDR3s used in transitional B (15.5 aa), P (14.1 aa), and NP (14.6 aa) cells are similar to those reported by other investigators (30, 33). Overall, these observations can be applied to individual families as well, which have primarily adopted intermediate-sized CDR3s (11–15 aa) following TLR9 engagement, in contrast with transitional B cells that make equal use of 11–15 aa and 16–20 aa CDR3 ranges (Fig. 6B). B cells belonging to the VH3 family deviate slightly from this experimental observation; short CDR3s are promoted in P cells, and the use of 11–15- and 16–20-aa-long chains is similar in P and NP cells (Fig. 6B). Collectively, the CDR3 length repertoire displayed by transitional B cells is influenced greatly by CpG stimulation giving rise to, on average, B cell clones expressing IgH chains of intermediate lengths.

**Discussion**

The uniqueness of each BCR and width of the B cell repertoire are due to the DNA rearrangements at HC and LC loci. B cells leave the bone marrow at the transitional stage carrying a wide array of germline-encoded and randomly assembled Ig genes. Transitional B cells migrate with the blood to the spleen. They are not fully functional in the immune responses: they are not able to respond to inflammatory chemokines, and engagement of the BCR leads to their apoptotic cell death. These are safety measures against autoreactivity preventing self-reactive transitional B cells from further development and migration to sites rich in cytokines and stimulatory cells. In contrast to the mortal effect of BCR cross-linking, engagement of TLR9 leads to differentiation of transitional B cells. In response to CpG, a fraction of the transitional...
B cells (26%) proliferates and acquires the phenotype of memory B cells or directly matures into plasma cells. The remaining cells start to express the markers of mature-naive B cells and survive without dividing (14). Because transitional B cells are the only B cell type present at birth, we hypothesized that their ability to respond to bacterial DNA through TLR9 represents a mechanism to rapidly produce a mixture of Abs of various specificities, all encoded in the germline and probably selected by thousands of

**FIGURE 6.** CpG modulates CDR3 lengths in P and NP cells to CDR3 chains displayed by IgM memory and mature B cells. A, CDR3 lengths were classified arbitrarily into <10 aa (white), 11–15 aa, 16–20 aa, and >21 aa in sorted transitional P and NP cells. Mean lengths were also calculated for each cell type. B, Adjustments in CDR3 chain length within each VH gene family and cell type in the transition to P and NP cells. Lines depict percentages observed in transitional B cells. The results shown are pooled from independent experiments generated from three diverse CB samples.
years of coevolution with ubiquitous microbes. These Abs may be indispensible for immediate survival after birth, when the newborn, coming from the maternal sterile microenvironment, is overwhelmed by bacteria that colonize the gut, airways, and skin. Fragments of bacterial DNA, draining from the degraded microflora of the intestine, come in contact with transitional B cells, probably in the spleen, the site where these cells collect. Bacterial fragments, filtered in the marginal zone, may trigger further B cell maturation. The memory-like B cells obtained in vitro after CpG stimulation may correspond to the first IgM memory B cell pool that, later in life, provides an indispensable protection against bacterial infection.

To investigate whether the memory-like population has the hallmark of memory B-cells (i.e., carries somatic mutations), the variable regions of the Ig H chains (VH) from P and NP cells were amplified by PCR, subcloned, and the sequences were analyzed. Our results demonstrated that CpG induces somatic mutations in the VH regions of the P population, whereas the NP cells mostly maintain their Ig germline sequences. The incidence of mutations is VH specific, because mutations are more frequent in VH1 and VH4/6, whereas VH3 has fewer mutations, and VH5 is basically not mutated. Are the nucleotide exchanges that we observed bona fide somatic mutations? First of all, all mutations induced by CpG are localized to the VH region and, similar to SHMs detected in vivo, they are never introduced in the constant domains of the Ig genes. The enzymatic machinery indispensable for SHM includes the basic DNA-repair apparatus, always expressed by cells, complemented by two specific mutators, AID and UNG, which are induced in B cells in the GCs and increase the frequency of mutations at the appropriate time, in the appropriate cells, and in the appropriate gene. We demonstrated that both enzymes are upregulated by TLR9 stimulation, supporting the hypothesis that SHMs can be induced outside the GCs. This idea first arose from the studies of J.C. Weill and colleagues (10), who broke a dogma, demonstrating that patients with X-linked Hyper IgM disease type 1, who lack the CD40L and cannot build GCs, have IgM memory B cells carrying somatic mutations. A detailed analysis of mutations in these patients revealed that the frequency of mutations is lower than in memory B cells of healthy donors, with an overrepresentation (60%) of clones with a single mutation (Supplemental Fig. 2) (19). The distribution and type of mutations in these patients do not show molecular signs of Ag selection, in contrast to what is observed in normal donors. In addition, transitions on dC/dG residues dominate over transversions (19), suggesting that UNG2 may be absent or not fully functional in these patients. Comparison of the features of memory B cells of Hyper IgM patients with those of in vitro-generated memory B cells revealed interesting similarities. The frequency of somatic mutations measured in P transitional B cells stimulated in vitro with CpG was equivalent to the frequency found in memory B cells of Hyper IgM patients (Supplemental Fig. 2B). In addition, 66% of the P cells carried a single mutation, whereas clones carrying two to four or five to nine accounted for only 20% and 11% of the mutated population, respectively (Supplemental Fig. 2A). Also, transitions prevail over transversions in in vitro-mutated transitional B cells. In our case, the observed skewing toward transitions cannot be attributed to the absence of UNG2. Additional experiments are necessary to demonstrate whether UNG2 is functional. The hypervariable CDR regions play the most important role in determining Ab specificity. Single amino acid substitutions have strong effects on affinity and may lead to autoreactivity. For these reasons, CDR mutations are scrutinized and selected by Ag in the GCs. The mutations that we observed in vitro were preferentially localized to the FRs, and CDR regions were less mutated. FRs have an important structural role, but they may also contribute to promiscuous Ag binding in natural Abs (34, 35). We propose that, in the absence of a nominal Ag, somatic mutations induced by TLR engagement aim at Ab diversification, avoiding the drastic effects that may be induced targeting CDR sequences. After CpG stimulation, we observed a selection for shorter CDR3s in all VH families. Long CDR3 chains bear the risk of autoreactivity and, although present in transitional B cells, are eliminated from the mature and memory pools (30, 31). We think that this is suggestive of a coengagement between the BCR and TLR9, where synergism between receptors might dictate inclusion or exclusion of individual emerging transitional B cells into or from the memory repertoire. In conclusion, our data demonstrate that CpG causes somatic mutations in vitro, at a frequency and with features similar to that observed in patients with Hyper IgM syndrome. This suggests that TLR9 stimulation may be a T independent signal able to generate a primitive memory B cell repertoire that is not Ag selected. In agreement with our hypothesis, a low frequency of somatic mutations was also observed in the memory populations isolated from the spleens of young children (18).

In summary, we observed that early recognition of bacterial DNA induces proliferation, somatic mutation, Ab secretion, or simply survival in transitional B cells. The different effects are largely based on VH family selection (Fig. 5). CpG does not amplify or mutate VH3 significantly; furthermore, most mutations are silent. Phylogenetically, VH3 is the oldest family, conserved for >300 million years in evolution from shark to man and maybe, essentially perfect to be included only marginally modified in the NP or P repertoire. VH5-expressing transitional B cells, which are rare in the bone marrow, are also expanded without modifications. Most mutations are introduced in VH1 and VH4/6. The selective effects of TLR9 on cells carrying different VH families could be explained by the recognition of these VH genes by CpG (super-antigen, anti-DNA specificity of Ig), thus generating a BCR signal that cooperates with the TLR9 pathway. It is also possible that the synergy between the BCR and TLR signaling complexes depends on the particular activation status of selected cells (low-affinity binding of unknown autoantigens, intrinsic BCR tonic signal) (36–38). Altogether, TLR9 activation alone seems to initiate the process of SHM, an imperfect and rudimentary mechanism of Ig diversification, which has undeniably evolved to create the most sophisticated molecular device to generate high-affinity Abs of the adaptive immune system. The result of this early selection may be the transformation of transitional B cells, a short-lived population unable to respond to inflammatory chemokines, into more useful cell types. Natural Abs are immediately produced; primitive memory B cells, although imperfect and probably promiscuous, may be able to migrate to sites of inflammation and further improve their specificity or at least try to counteract infections; meanwhile adaptive memory B cells are generated in the GCs. The NP population may develop by default into mature-naive B cells. BCR signaling is the driving and selective force of B cell development in the bone marrow and in the periphery compared with the bone marrow (manuscript in preparation). We still do not know whether TLRs other than TLR9 play a role in B cell development. However, it was recently demonstrated that the lack of fundamental elements of TLR9 signaling (cascade MyD88, IRAK-4, and UNC-93B) are associated with bacterial infections, as well as with unexpected defects in tolerance.
induction, a process that has always been thought to be uniquely dependent on the BCR (40).

Acknowledgments

We thank M.M. Rosado for critical reading of the manuscript and helpful discussions.

Disclosures

The authors have no financial conflicts of interest.

References

Figure 1. Transitional B cells respond strongly to CpG stimulation

(A) After 7 days in culture proliferating (P or CMFDA+) and non-proliferating (NP or CMFDA+) CD19+ B cells were sorted, collected and total RNA was extracted for following studies. The dot plots in the upper panels show our sorting strategy. We sorted cells in the lymphocyte region (BEFORE SORT, left panel) and positioned the NP gate on the basis of CMFDA staining in cord blood maintained in culture for 7 days without CpG (middle panel). Proliferating cells, visible in CpG stimulated cultures, progressively lose CMFDA staining intensity (right panel). The dot plots in the lower panels (AFTER SORT) show the cells sorted from CpG stimulated cultures. The numbers refer to the percentage of cells in the P or NP gate in the samples re-analyzed after sorting.

(B) The sorting strategy for isolating cord blood transitional B cells and the re-analysis after sorting are shown. Transitional B cells are contained in the lymphocyte gate and express high levels of CD38 and CD24. After sorting purity was 99%.
Figure 2. Comparison between mutation data from X-linked HlgM patients and memory B cells generated in vitro as a result of CpG stimulation.

A

B

The source of all X-HlgM data is the article corresponding to reference number 19.

(A) Observed mutation range per analysed sequence is plotted, as total numbers or percentages. It is apparent the prevalence of sequences bearing a single mutation (60% vs 66%). Sequences with 2-4 and 5-9 mutations are not only less frequent in both datasets, but percentages are coincident as well.

(B) Mutation frequencies (%) in adult X-linked HlgM patients and CpG-derived proliferating memory B cells are highlighted in black. White bars depict either frequencies found in X-linked HlgM infants or in discrete VH families from in vitro generated memory B cells.
Table SI. Semi-quantitative RT-PCR primer sequences and PCR conditions

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer sequence (5'--&gt;3)</th>
<th>Annealing temperature (°C)</th>
<th>Number of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-VH3</td>
<td>Forw 5'-AAGGTGTCCAGTGTGARGTGCAG-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-VH4/6</td>
<td>Forw 5’-CCCAGATGGGTCCTGTCCACAGGTGCAG-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-VH5</td>
<td>Forw 5’-CAAGGAGTCTGGTCCAGAGGTGCAG-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3’ Cµ CH1</td>
<td>Rev 5’-GGGAATTCTCACAGGAGACGA-3’</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>CH4</td>
<td>Forw 5’-GACCATCTCCTGGAACCACGC-3’</td>
<td>68</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-TCCCTGGTCATGTCCGACAC-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AICDA</td>
<td>Forw 5’-TAGACCCTGGCCGCTGCTACC-3’</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-CAAAGGATGCGCCGAAGCTGCTGAG-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNG2</td>
<td>Forw 5’-ATCGGCCAGAAGACGCTCTA-3’</td>
<td>65</td>
<td>40*</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-ACAGCAGCTTCTCAAGGCCA-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POLΗ</td>
<td>Forw 5’-GAGTGGTGGAATAATTGGAG-3’</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-TGCCAGGACCTTATTGTCG-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSH2</td>
<td>Forw 5’-GGCGGAGGAGAAAGATGAAT-3’</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-GATGAAACTGCAACCTGACCC-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSH6</td>
<td>Forw 5’-GGCAGGAACCTCAACGGG-3’</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-GTTGTCCTACCTCCATCTCTCC-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPRT1</td>
<td>Forw 5’-CGACCGCTGGGCTGATTAGT-3’</td>
<td>60</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-CCCTGTTGACTGTCATTACAAT-3’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-ACTIN</td>
<td>Forw 5’-GCGGGGAAATCAGTGACATT-3’</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Rev 5’-GTGGACTTGGAGGAGACTCCTGGG-3’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PCR conditions: 94°C for 30 sec, Tm for 30 sec, 72°C for 45 sec (*2 min) followed by an extended elongation step of 7 min
Table SII: Summary of the cloning experiments in proliferating (P), non proliferating (NP) and transitional (T) B cells.

A

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Family</th>
<th>Total clones (unique)</th>
<th>Number of mutated clones (unique)</th>
<th>% Mutated clones</th>
<th>Total nucleotides (unique clones)</th>
<th>Single mutations (unique clones)</th>
<th>%Mutation Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>VH1</td>
<td>48</td>
<td>23</td>
<td>48</td>
<td>13392</td>
<td>54</td>
<td>0,37</td>
</tr>
<tr>
<td></td>
<td>VH3</td>
<td>35</td>
<td>10</td>
<td>28,6</td>
<td>9765</td>
<td>21</td>
<td>0,21</td>
</tr>
<tr>
<td></td>
<td>VH4/6</td>
<td>41</td>
<td>12</td>
<td>29,3</td>
<td>11439</td>
<td>32</td>
<td>0,27</td>
</tr>
<tr>
<td></td>
<td>VH5</td>
<td>64</td>
<td>8</td>
<td>12,5</td>
<td>17856</td>
<td>8</td>
<td>0,04</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>188</strong></td>
<td><strong>53</strong></td>
<td><strong>28</strong></td>
<td><strong>52452</strong></td>
<td><strong>115</strong></td>
<td><strong>0,04</strong></td>
</tr>
<tr>
<td>NP</td>
<td>VH1</td>
<td>12</td>
<td>4</td>
<td>13,8</td>
<td>4012</td>
<td>6</td>
<td>0,14</td>
</tr>
<tr>
<td></td>
<td>VH3</td>
<td>16</td>
<td>3</td>
<td>11,5</td>
<td>5002</td>
<td>6</td>
<td>0,11</td>
</tr>
<tr>
<td></td>
<td>VH4/6</td>
<td>21</td>
<td>2</td>
<td>6,9</td>
<td>5859</td>
<td>3</td>
<td>0,05</td>
</tr>
<tr>
<td></td>
<td>VH5</td>
<td>19</td>
<td>2</td>
<td>6,9</td>
<td>5301</td>
<td>2</td>
<td>0,04</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>68</strong></td>
<td><strong>11</strong></td>
<td><strong>9,8</strong></td>
<td><strong>20174</strong></td>
<td><strong>17</strong></td>
<td><strong>0,04</strong></td>
</tr>
<tr>
<td>T</td>
<td>VH1</td>
<td>17</td>
<td>3</td>
<td>11,1</td>
<td>4743</td>
<td>3</td>
<td>0,06</td>
</tr>
<tr>
<td></td>
<td>VH3</td>
<td>14</td>
<td>2</td>
<td>12,5</td>
<td>3906</td>
<td>2</td>
<td>0,05</td>
</tr>
<tr>
<td></td>
<td>VH4/6</td>
<td>27</td>
<td>3</td>
<td>8,1</td>
<td>7533</td>
<td>4</td>
<td>0,05</td>
</tr>
<tr>
<td></td>
<td>VH5</td>
<td>42</td>
<td>8</td>
<td>17,8</td>
<td>11718</td>
<td>7</td>
<td>0,08</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>100</strong></td>
<td><strong>16</strong></td>
<td><strong>12,4</strong></td>
<td><strong>27900</strong></td>
<td><strong>16</strong></td>
<td><strong>0,04</strong></td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Total clones</th>
<th>Total nucleotides</th>
<th>Single mutations</th>
<th>%Mutation Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-CH4</td>
<td>28</td>
<td>11200</td>
<td>1</td>
<td>0,009</td>
</tr>
<tr>
<td>P-CH4</td>
<td>33</td>
<td>13200</td>
<td>4</td>
<td>0,030</td>
</tr>
</tbody>
</table>

(A) Total numbers of unique clones, mutated unique clones, single mutations and nucleotides sequenced. Percentages of mutated clones and mutation frequency are also shown. The difference in mutation frequencies between P-VH1 and T-VH1 ($P = 0.0006$) and P-VH4/6 and T-VH4/6 ($P = 0.0008$) is significant.

(B) Data referring to the CH4 region of Transitional (T-CH4) and Proliferating (P-CH4) B cells. The difference in mutation frequency between P-CH4 and T-CH4 ($P = 0.48$) is not significant.
Table SIII. The SHM machinery targets both RGYW/WRCY and WA/TW hotspot motifs.

Motifs are depicted with, underlined residues to denote AID targeting and capital (conserved) and small (mismatch) letters to indicate similarity with classical RGYW/WRCY motif. The analysis with alternative SHM motifs gives a better fit. The reduction in the number of mismatches is specified. R:A/G, Y:C/T, W:A/T, D:A/G/T, H:T/C/A and N:A/G/C/T. Mutated residues are shown underlined, within established WA/TW hotspot motifs, as well as within other non-conventional motifs.