The Repertoire of Somatic Antibody Mutants Accumulating in the Memory Compartment after Primary Immunization Is Restricted through Affinity Maturation and Mirrors That Expressed in the Secondary Response

By Ursula Weiss and Klaus Rajewsky

From the Institute for Genetics, University of Cologne, D-5000 Köln 41, FRG

Summary

The anti-(4-hydroxy-3-nitro-phenyl)acetyl (NP) response is dominated by λ1 chain-bearing antibodies expressing the V_H gene V186.2 in combination with the D element DFL16.1. λ1-positive B cells were isolated from the spleens of mice immunized with NP-chicken gamma globulin 6 wk earlier. Rearranged V186.2 genes were amplified from the genomic DNA of these cells and sequenced. In cases where the rearrangement was typical for secondary anti-NP antibodies, the V_HDJ_H sequences were generally heavily mutated. The frequency and the nature of the nucleotide exchanges mirrored those of secondary response antibodies. V186.2 genes with other rearrangements and V186.2-related genes isolated concomitantly were essentially unmutated. These results demonstrate: (*a*) that somatic antibody mutants are largely restricted to a small compartment of peripheral B cells, namely, that of memory cells; (*b*) that the memory compartment is strongly selected for high affinity precursors and largely purged from antigen-binding loss mutants; and (*c*) that the repertoire of binding specificities expressed in the secondary response is established in its final form before secondary immunization.

Present knowledge about the molecular nature of the B cell memory compartment is derived from the phenotypic and genotypic analysis of secondary immune responses. Memory B cells that differentiate into secondary response plasma cells are of different clonal origin than primary response precursor cells. Their variable region (V)[1] genes are extensively somatically mutated. Collectively, they bind the immunizing antigen with higher affinity than primary response antibodies (1). However, it is not known whether the secondary response readout of B cell memory generated in response to a particular antigen reflects the overall composition or a selected subset of the memory cell pool. Considering the high rate of somatic mutation, $\sim 10^{-3}$/bp/cell division (2, 3), one must expect the generation of antibody mutants that do not bind antigen with increased affinity, including low affinity binders, loss variants, and mutants that have acquired specificity for autoantigens. In fact, experimental evidence has been obtained indicating that autoreactivity can be generated by somatic mutation (4), and that variants with drastically reduced affinities or altered antigen binding specificities do persist in the memory compartment (5). We previously identified the B cell precursor of a secondary response lymphocyte with genotypic characteristics strongly suggesting its original stimulation by the immunizing antigen and subsequent loss of immunogen binding through somatic mutation (5). Its crippling mutation had been identified independently in a hybridoma variant isolated in vitro (6). Interestingly, the loss of binding to the immunizing antigen in this variant was paralleled by the acquisition of binding specificity for a related hapten. Based on these observations, one could hypothesize that the antigen-induced hypermutation mechanism not only serves to increase the efficiency with which an organism deals with a primary immunizing agent, but in addition, acts as a means to increase the overall diversity of the antibody repertoire of peripheral B cells. A prediction resulting from this model would be a discrepancy between the molecular characteristics of the overall population of memory B cells induced by a particular antigen, and the subset expressed after secondary immunization.

With the advent of PCR technology, it has become possible to selectively draw into a molecular analysis the subset of rearranged V genes known to dominate a particular immune response. This has enabled us to study the composition of the memory compartment directly, independent of its expression triggered by secondary immunization with antigen. As a model system, we used the immune response to

[1] *Abbreviations used in this paper:* CG, chicken gamma globulin; NP, (4-hydroxy-3-nitro-phenyl)acetyl; R/S ratio, ratio of replacement to silent mutations; V genes, variable region genes.

the hapten (4-hydroxy-3-nitro-phenyl)acetyl (NP) coupled to the T cell–dependent carrier chicken gamma globulin (CG). This response has been well characterized. It is dominated by λ1-bearing antibodies expressing the H chain V gene V186.2 in association with the D segment DFL16.1. In approximately half of these antibodies, the 5′ part of DFL16.1 consists of germline sequences encoding three consecutive tyrosines (7–9). In secondary response antibodies, the 3′ part of DFL16.1 is replaced by heterogeneous N sequences (8, 9). 70% of clonally independent V186.2-encoded secondary response antibodies share a mutation at position 33 of the H chain (8, 9). This tryptophan-to-leucine exchange alone is sufficient to account for the 10-fold increase in affinity that distinguishes most secondary from primary response anti-NP antibodies (10). Additional mutations in the H chain gene as well as amino acid substitutions in the V region of the L chain contribute in most cases at best marginally to the hapten binding affinity. We have studied the composition of the memory compartment by isolating and sequencing V186.2 genes joined to DFL16.1 from λ1-expressing spleen cells of mice that had received a single immunization with NP 6 wk earlier.

Materials and Methods

Immunizations. 10-wk-old female C57BL/6 mice from our colony were primed by intraperitoneal injection of 100 μg of alum-precipitated NP-CG together with 2 × 10^9 inactivated pertussis bacteria.

Isolation of Splenic λ1-Expressing B Cells. Spleen cells of 12 mice were pooled and depleted of erythrocytes by treatment with 0.08% NH_4Cl. The cells were stained on the surface with the biotinylated anti-λ1 antibody Ls136 (11), and streptavidin-FITC. Ls136-positive cells comprised 1.8% of the population. $λ1^+$ cells were enriched to a level of 45% using a magnetic cell sorter (12), and subsequently enriched to 97% by fluorescence-activated cell sorting on a FACS 440. Dead cells and cells other than lymphocytes were excluded by the addition of propodium iodide and appropriate gating.

Preparation of Cellular DNA and PCR for Amplification. Genomic DNA of 4.5 × 10^{-6} sorted cells was purified (13). The following oligonucleotide primers were used for PCR amplification: V186.2 5′CCTGACCCAGATGTCCCTTCTTCTCCAGCAGG 3′ (this primer hybridizes to genomic DNA 5′ of the transcription start site of V186.2); intron J_H1 5′GGGTCTAGACCCGTTTCAGAATGGAATGTGCAG 3′; intron J_H2 5′GGGTCTAGAGGTGTCCCTAGTCCTTCATGACC 3′ (the J_H primers include the recognition sequence for XbaI). V186.2 genes rearranged to J_H1 and J_H2, respectively, were amplified in separate reactions. Each PCR reaction contained 10% of total purified DNA, 20 pmol of each primer, 10 mM Tris-Cl (pH 8.3), 50 mM KCl, 4 mM $MgCl_2$, 200 μM of each dNTP, 0.01% gelatine, and 7 U Taq polymerase (Perken-Elmer Corp., Hayward, CA) in a total volume of 50 μl. PCR was performed for 50 cycles: each cycle consisted of 96°C for 1.4 min and 70°C for 2 min.

Cloning of Amplified DNA and Bacterial Transformation. Amplified DNA was extracted once with phenol-chloroform, precipitated with NaAc/ethanol, digested with the restriction enzymes NcoI and XbaI, and subjected to agarose gel electrophoresis. The DNA fragment corresponding to the expected size was purified and ligated into a modified pUC19 vector which contains an additional multisite cloning cassette derived from the plasmid vector pGEM-5Zf(+) (Promega Biotec, Madison, WI) (kind gift of Steffen Jung, Institute for Genetics, Köln, FRG). DH-5α bacteria were transformed by electroporation (gene pulser; Bio-Rad Laboratories, Munich, FRG).

Colony Hybridization and DNA Sequence Determination. Bacterial colony hybridization with a ^{32}P end-labeled oligonucleotide was performed as described (14). The probe had the following sequence: 5′GTAGTAATATCTTGC 3′; it hybridized to the terminal six nucleotides of the V_H186.2 gene plus the initial nine nucleotides of DFL16.1 when joined directly to the V_H gene in reading frame 1 (15). V_H gene sequences were obtained by direct plasmid sequencing using the Sequenase TM kit (United States Biochemical Corp., Cleveland, OH).

Expression Vector Construction and Transfection into Myeloma Cells. The mutated V186.2 genes designated K3 and K11 in Fig. 3 were excised from modified pUC19 and inserted into an Ig γ1 expression vector ($pEV_Hcγ1$). The vector $pEV_Hcγ1$ was constructed by inserting the enhancer cassette of pEV_H (16) into the vector pγ1gpt (17). The resulting constructs were linearized with PvuI and transfected into J558L myeloma cells (18) by electroporation. J558L cells were grown in RPMI 1640 supplemented with 10% FCS, penicillin (100 μg/ml), streptomycin (100 μg/ml), and 2 × 10^{-5} M 2-ME. Transfected cells were seeded into 48-well plates 48 h after transfection and selected for expression of the bacterial gpt marker in supplemented RPMI containing 5 μg/ml mycophenolic acid, 250 μg/ml xanthine, and 50 μg/ml hypoxanthine. Ig-producing cell lines were identified by screening culture supernatants in an avidin-biotin ELISA (19) using goat anti–mouse IgG1 antibodies (Southern Biotechnology Associates, Birmingham, AL) to coat the plastic plates. Cell supernatants were used for serological assays.

Farr Assay. The method used for determining antibody binding affinity has been described (20). It is based on the inhibition of binding antibody to $N^{125}IP$-caproate by increasing concentrations of unlabeled NP-caproate. The concentration of hapten resulting in 50% inhibition of $N^{125}IP$-caproate binding approximates the binding constant for the hapten.

Results

Experimental Design. C57BL/6 mice were immunized intraperitoneally with a single dose of NP-CG. Six wk later spleen cells expressing λ1 L chains on their surface were isolated as described above, using a biotinylated λ1-specific mAb, and magnetic and fluorescence activated cell sorting. Genomic DNA was prepared, and V186.2 genes rearranged to J_H1 and J_H2 were amplified in separate PCRs. In most (57%) of the secondary response anti-NP antibodies that express the V_H gene V186.2, it is joined to J_H2, but rearrangements to J_H1, J_H3, and J_H4 were also seen. After digestion with appropriate restriction enzymes, the amplified DNA was cloned into a derivative of the plasmid vector pUC19. Bacterial colonies expressing V186.2 genes were identified by direct DNA sequence analysis or after screening with a radioactively labeled oligonucleotide. Since the V_H gene specificity of the amplification reaction rests on the 5′ PCR primer, and V186.2 belongs to a large family of related V_H genes (group 1 in the classification of Dildrop [reference 21]), it was not surprising to find that the majority (75%) of genes recovered by direct sequence analysis were genes closely related to V186.2 but not V186.2 itself. To facilitate the identification of rearranged

V186.2 genes derived from NP-induced memory cells, we designed an oligonucleotide homologous to the V_H-to-D junction of approximately half of the antibodies that had previously been recovered after immunization with NP (7–9). It encompasses the six most 3′ bp of V186.2 and the nine most 5′ bp of DFL16.1. We cannot rule out that the selection associated with this screening procedure introduces a bias in the scope of the analysis. However, it is unlikely that the focus on V_H-to-D junctions characteristic for rearranged genes known to participate in the response against NP should selectively exclude those genes from the analysis that encode antibodies which, through hypermutation, have lost the ability to respond to a secondary immunization with NP and/or have acquired new antigenic specificities.

The V_H Gene Repertoire of Splenic B Cells Includes Germline and Somatically Mutated V_H Genes. The nucleotide sequences of altogether 49 rearranged V_H genes were determined. They fall into two classes: genes that differ from the germline at most at two positions, and genes that are extensively mutated. The first class includes V186.2-related genes joined to heterogenous D elements, V186.2 genes with heterogenous V_H-to-D junctions, and V186.2 genes joined directly to DFL16.1 in reading frame 1. The second class consists exclusively of V186.2 genes in direct association with DFL16.1.

Most of the Rearranged V_H Genes in the Peripheral B Cell Pool of the Mouse Are Unmutated. Taking into account the available sequence information encompassing the V_H region but excluding possible nucleotide exchanges in the D region and at the V_H-D and D-J_H joins, we observe three mutations in 1,827 nucleotides derived from 8 V186.2-related genes, eight mutations in 5,258 nucleotides derived from 17 V186.2 genes with heterogenous V_H-to-D junctions (data not shown), and three mutations in 2,127 nucleotides derived from 7 V186.2 genes joined directly to DFL16.1 (Fig. 1). This corresponds to mutation frequencies of 1 in 609, 1 in 657, and 1 in 709, respectively. The 14 exchanges are distributed over 13 of the 32 V_H gene sequences present in this class. We believe that these mutations were introduced not in vivo but in vitro for the reasons outlined below.

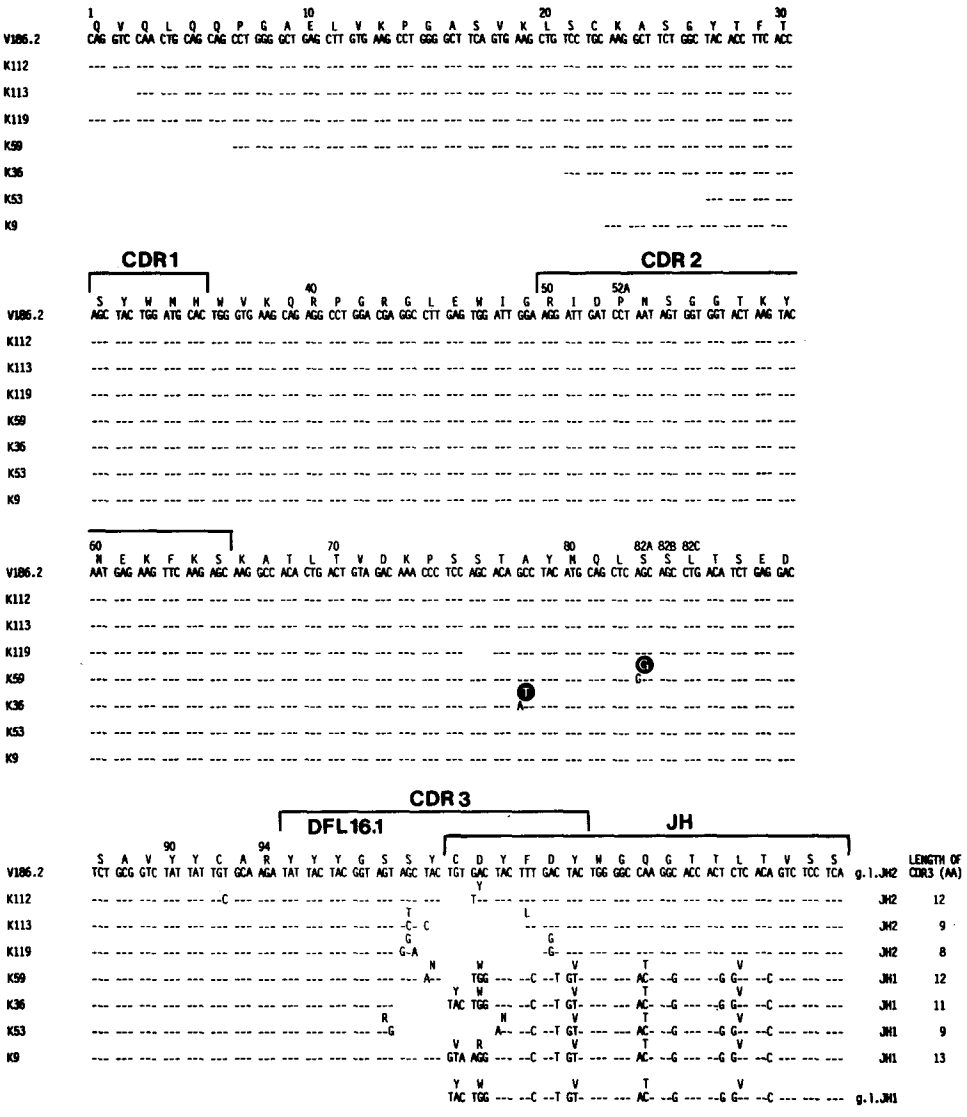

Figure 1. Sequences of the V186.2 genes that are functionally rearranged to the D element DFL16.1 in reading frame 1 and show no evidence for hypermutation. The sequences are compared with the germline V_H gene V186.2, the DFL16.1 segment, and the J_H segments J_H1 and J_H2. Nucleotides identical to the reference sequence are indicated by dashes; blanks indicate uncertainty. Blanks at the 5′ end of the D and J_H segments denote the absence of those codons from the sequenced gene. Marked in black circles are the amino acid substitutions relative to the reference sequence. The codons are numbered according to Kabat et al. (22). The amino acid length of CDR3 is indicated. *g.l.*, germline.

It has been shown that the Taq polymerase used for PCR amplification has a misincorporation rate of $\sim 10^{-4}$ (23, 24). From the number of PCR cycles our DNA isolates were subjected to before sequence analysis (i.e., 50) and the observed mutation frequency of overall 1 in 658, we calculate a misincorporation rate of 6×10^{-5} (23). This mutation rate lies somewhat below the Taq polymerase error rate described by others (23, 24). This difference could be due to slightly varied reaction conditions during PCR amplification. The mutations apparent in this set of genes resemble in their molecular nature typical Taq polymerase errors in that there is a strong bias for purine transitions (86%), most of which result in the replacement of a T-A by a C-G bp (67%). Taken together, the frequency and the nature of the mutations make it likely that they were generated in vitro, and that the peripheral B cells from which these gene sequences were derived expressed antibody receptors encoded by germline genes.

7 of the 8 V186.2-related and 16 of the 17 V186.2 genes of this essentially unmutated subset are productively rearranged to D elements differing widely in length and sequence, and are characterized by extensive junctional variability. This is depicted in Fig. 2. Because of this variability, we consider these V_H region genes representative of the genes expressed by the population of naive peripheral B cells. This sequence analysis therefore gives direct evidence for the predominant expression of germline-encoded antibodies in the peripheral B cell pool of the mouse.

The seven V186.2-DFL16.1-J_H gene combinations depicted in Fig. 1 resemble the H chain V genes of typical anti-NP antibodies in that they encode three tyrosines at the 5′ end of CDR3. However, three of the seven genes in this group have retained all or most of the germline sequence of the 3′ half of DFL16.1. In particular, downstream of the codons specifying the three tyrosines, they encode a glycine followed by two serines. In all secondary response antibodies that have been obtained from hybridomas so far, the amino acid sequence deviates from the germline-encoded gly-ser-ser sequence (8, 9). These codons are substituted by N sequences generating a truncated CDR3 with an average length of nine amino acids. Curiously, with few exceptions, all primary anti-NP antibodies studied to date do express the germline-encoded gly-ser-ser sequence (7). Thus, as judged by the sequence and length of their CDR3s, at least three of the gene sequences in this set, i.e., K112, K59, and K9, encode H chain structures that appear to be incompatible with those required for channeling the parent B cells into the memory compartment. The CDR3s specified by the other four V_H gene sequences in this set have also not been seen in secondary response antibodies. It is not clear whether these genes originate from B cells with antibody receptors that did not fit the prerequisites for memory cell precursors or else from B cells that were seeded into the peripheral immune system after termination of the antigen-stimulated recruitment of memory cell precursors into the hypermutation pathway.

V_H clone	3′end of V_H gene	D element and N additions	5′ end of J_H
Q1	AGA	TAT AGG GTT ACG ACA G	AC TGG TAC TTC
Q2	AGA	TA	C TAC TGG TAC TTC
Q3	AGA	TAC CGC GAC TCC CGT TAC TGG	TAC TTC
Q4	AGA	A	AC TGG TAC TTC
Q5	AGA	TGG AAC TAT GGT A	AC TGG TAC TTC
Q6	AGA	TTC GCT ACT ATG GTT ACA A	AC TGG TAC TTC
Q7	AGA	GAT GGT TAC TGC GGC T	AC TGG TAC TTC
Q8	AGA	TGT ATC TAT GAT GG	C TGG CAC TTC
Q9	AGA	G	GG TAC TTC
Q10	ACG	ATG GGA TGG TTA CTG C	TC
Q11	AGA	TCC TAT TT	G TAC TTC
Q12	AGA	AGT ACC CAG GGG	TAC TTC
Q13	AGA	TAT TAT GGG ACT CCC C	GG TAC TTC
Q14	AGA	TCC GTT TAC TAC GGT AGT AGC C	TT
Q15	AGA	GAG GGA TTA CGA CC	C TTT
Q16	AGA	CGG TTA ACT GGG AC	T
R1	AGA	TCG C	GG AAC TTC
R2	AGA	TGG GGX ACG GTA GTA GAT	TAC TGG TAC TTC
R3	AGA	GGA GGG AGG	TAC TGG TAC TTC
R4	AGA	GGG ACA TCG ATG ATG GCA CCA TTA TTC TGA	TAC TTC
R5	AGA	ATG ATT ACA ACG G	TGG TAC TTC
R6	AGA	GAG GGT ATT ACT ACT GTA GTA G	AC TGG TAC TTC
R7	AGA	GCG GAA TAT AAC TTC	TAC TTC

Figure 2. V_HDJ_H junctional sequences of V186.2 genes functionally rearranged to D elements other than DFL16.1 (designated Q1 to Q16), and of V186.2-related genes (designated R1 to R7). The V_H germline genes corresponding to R1 to R7 are as follows: R1, V 165.1 (reference 5); R2, V 671.5.6 (5); R3, V 24.8 (5); R4, V C1H4 (Dildrop, R., unpublished results); R5, V 24.8 (5); R6, V CH10 (reference 25); R7, V 3 (reference 26).

The Majority of Rearranged V_H Genes Likely to Encode NP Binding Specificity in Peripheral B Cells of NP-primed Mice Are Mutated. 24 V genes specifying the gene segment combination known to dominate the anti-NP response, i.e., V186.2 joined to the D element DFL16.1, were isolated in the course of this study. The majority of them, i.e., 17, carry between three and 15 substitutions (excluded are possible nucleotide exchanges in the D region and at the V_H-D and D-J_H joins) (Fig. 3). Within 5,565 nucleotides sequenced altogether, we identified 108 exchanges, corresponding to a mutation frequency of 1 in 52. We have classified these genes as somatic variants of V186.2 rather than representatives of related germline genes based on the following: (*a*) a comparison of the nucleotide sequences of the V_H genes depicted in Fig. 3, with all members of the group 1 gene family, shows that each is more homologous to V186.2 than to any other one member of this family. 63 different germline V_H genes can presently be assigned to V_H group 1 in strain C57BL/6 on the basis of nucleotide sequence homologies (Gu, H., I. Förster, W. Müller, and K. Rajewsky, manuscript in preparation). (*b*) In contrast, eight V186.2-related genes can clearly be identified as known germline genes (Fig. 2). (*c*) The tryptophan-to-leucine exchange at codon 33, which is specified by 70% of the genes depicted in Fig. 3, is typical for V186.2-encoded antibodies of the secondary response against NP. The genomic DNA of six secondary response hybridomas secreting antibodies that contain leucine at position 33 has previously been analyzed by restriction enzyme analysis and found to be consistent with the predictions of V186.2 rearrangements to different J_H segments (8).

The mutation frequency of 1 in 52 is 13-fold higher than the frequency of errors introduced by Taq polymerase in the course of PCR. Since both classes of genes in the present collection — those with few and those with many mutations — had been amplified together, it is fair to assume that they were mutated by Taq polymerase to the same extent. Therefore, the great majority ($\sim$90%) of the mutations present in the 17 genes depicted in Fig. 3 cannot be attrib-

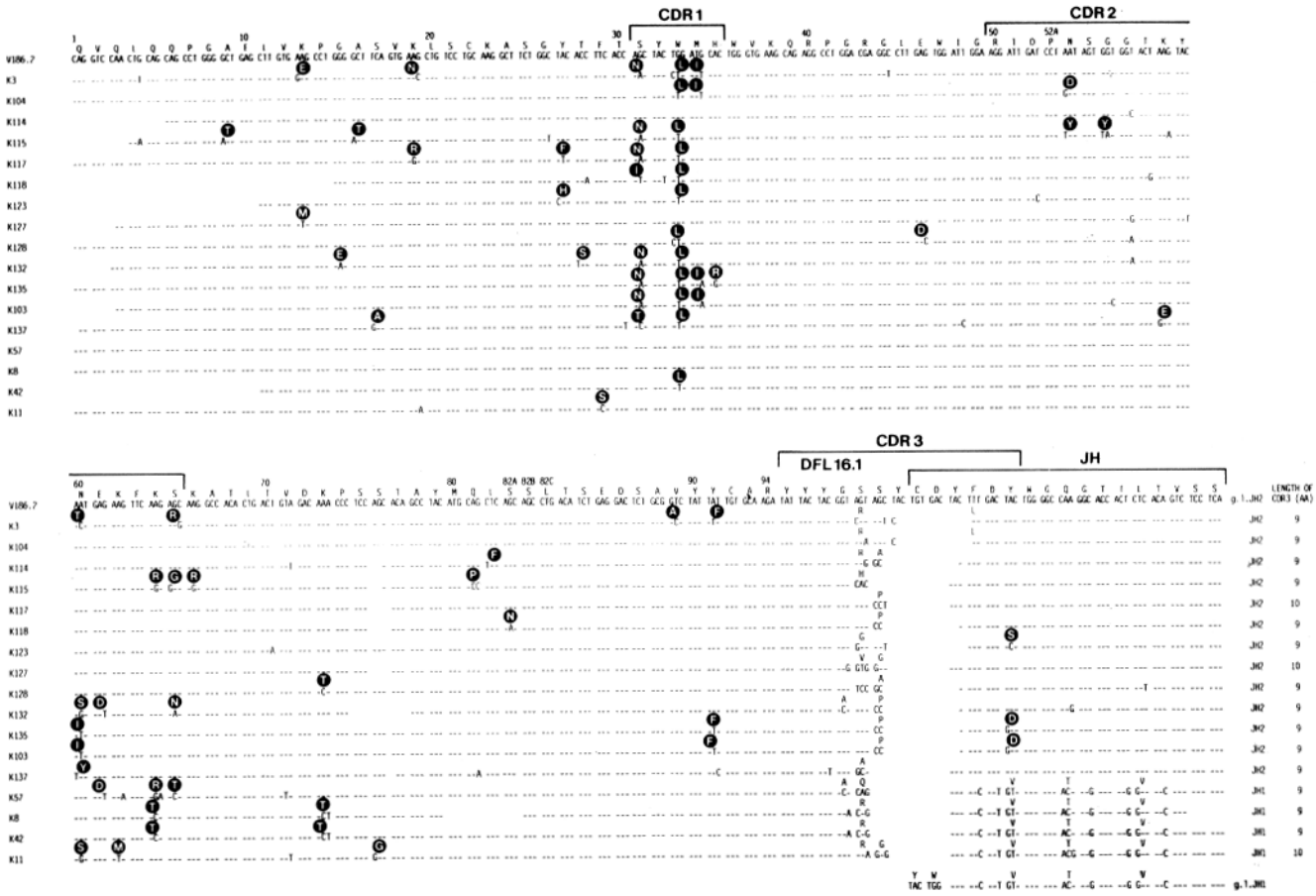

Figure 3. V region sequences of somatically mutated V186.2 genes joined to the D element DFL16.1 in reading frame 1. Replacement mutations (*filled circles*), and the amino acid length of CDR3 are indicated. For further description see legend to Fig. 1.

uted to processes that occurred in vitro, and consequently must have been generated in vivo.

Apart from the extensive mutagenesis characterizing these genes, two additional features make their derivation from memory cells likely: (*a*) the CDR3s they specify have an average length of nine amino acids as is typically seen in secondary anti-NP antibodies (8, 9); and (*b*) in all cases, the germline sequence encoding gly-ser-ser in the 3′ half of DFL16.1 has been replaced by N sequences. As pointed out above, the absence of the germline-encoded sequence gly-ser-ser distinguishes all V186.2-DFL16.1–encoded secondary response anti-NP antibodies from the great majority of the corresponding primary response antibodies (7–9).

The Pattern of Mutations Suggests the Selective Accumulation of High Affinity Mutants in the Memory Compartment. A comparison of the 17 heavily mutated V186.2 genes depicted in Fig. 3 with the genes encoding previously isolated secondary response antibodies reveals striking similarities in terms of frequency and nature of the mutations. In our present collection of genes derived from NP-primed mice, we observe an average of 6.3 exchanges per rearranged V_HDJ_H gene, spread over a range of 3 to 15. Secondary response antibodies isolated in two independent studies showed an average of 5.5 and 6.7 mutations per gene falling into a range of 2 to 12 and 3 to 12, respectively (8, 9). The ratio of replacement to silent (R/S) mutations in the genes of secondary response antibodies typically diverges from the value expected in a model of random mutagenesis in that replacement mutations are overrepresented in the complementarity-determining regions and underrepresented in the framework regions. This bias is apparent also in the V_H gene sequences depicted in Fig. 3. In the framework regions we observe an R/S ratio of 1.6:1. Taking into account the codon composition of V186.2, we would have expected an R/S ratio of 3.1:1 for random mutagenesis. The observed ratio of 1.6:1 correlates with the R/S ratio of mutations in the frameworks of V_H from 29 previously analyzed secondary anti-NP antibodies, i.e., 1.5:1 (references 8, 9), and indicates that amino acid exchanges were counterselected in the region defining the structural backbone of an antibody. In contrast, in CDR1 we observe an R/S ratio of 25:1 in the present gene collection that compares with a ratio of 21:1 in previously isolated secondary anti-NP antibody V_H genes (8, 9). The R/S ratio in CDR1 exceeds the expected value of 14:1, indicating that a fraction of the replacement mutations are positively selected in this region (see also below).

In the NP system, the secondary immune response differs from the primary one by an ~10-fold increase in affinity (27).

We have previously shown in the analysis of in vitro–engineered V186.2-DFL16.1–encoded antibodies that a single replacement in CDR1, a tryptophan-to-leucine exchange at codon 33, raises the binding affinity 10-fold (3, 10). This particular mutation is shared by 70% of all clonally independent secondary response antibodies derived from hybridomas (8, 9). Also 70%, i.e., 12 of the 17 mutated genes that we isolated from NP-primed mice and analyzed in the present study, encode this particular amino acid substitution (Fig. 3).

To confirm the association of the position 33 mutation with an increase in affinity to the hapten NP in the present V_H gene collection, we subcloned two of the V_H sequences into an expression vector carrying the γ1 C region gene. One of the two, K3, encodes nine replacement mutations, including the tryptophan-to-leucine exchange at position 33. The other, K11 encodes four amino acid replacements, and specifies the germline tryptophan at position 33. The expression vector constructs were transfected into the myeloma cell line J558L, which synthesizes germline λ1 L chains. The endogenous L chains pair with the H chains encoded by the transfected gene and antibodies are secreted. (The mutation rate for the λ1 L chain is severalfold lower than for the H chain, and studies with engineered antibodies indicate that mutations in the λ1 chain generally do not contribute to an increase in the binding affinity in the NP system [8–10].) The affinities to NP were measured by competitive inhibition of radiolabeled hapten (Table 1). As expected, the antibody formed with the H chain mutant K3 (which contains the tryptophan-to-leucine exchange at position 33) exhibits an eightfold higher affinity than the prototype primary antibody N1G9 which expresses the germline form of V_H186.2 (0.36 vs. 2.7 μM). It thus lies within the range of affinities displayed by most secondary response antibodies (0.3–1 μM). In contrast, the four amino acid exchanges in the antibody formed with the H chain mutant K11, which does not contain the position 33 mutation, do not effect an increase in affinity. A few low affinity binders have also been found among hybridoma-derived secondary response antibodies (8).

In addition to the key mutation specifying the amino acid exchange at position 33, other amino acid residues are recurrently mutated in the genes depicted in Fig. 3. Within the regions of CDR1 and CDR2, codons 31, 34, 53, 60, 61, 64, and 65 are mutated in more than one gene. We believe that most of these mutations are the result of mutational hotspots and reflect intrinsic properties of the hypermutation mechanism rather than selective forces for hapten binding affinity for the following reasons. Although an amino acid replacement can usually be brought about by an exchange in the first or second nucleotide position of a codon, in six of the seven codons considered here the same nucleotide is replaced. The restriction characterizing the target of mutagenesis is not paralleled by a similar restriction concerning the nature of the inserted amino acid. In the two codons affected the most, codon 31 (mutated in 8 of 17 genes) and codon 60 (mutated in 6 of 17 genes), the mutated nucleotide is seen replaced with any one of the three possible alternatives, leading to the insertion of three different amino acids.

Table 1. *Affinities of Somatically Recombinant Antibodies to NP*

	K_{NP}
	μM
N1G9*	2.7
aK3‡	0.36
aK11‡	12.0

Affinities of the antibodies for NP-caproate are shown; for experimental details see Materials and Methods.

* Antibody N1G9 is an unmutated primary response anti-NP IgG1 antibody carrying λ1 L chains and expressing V186.2 in combination with DFL16.2 (reference 7).

‡ aK3 and aK11 are recombinant antibodies formed by association of the germline λ1 L chain expressed in J558L cells with the H chains encoded by the expression vectors $pEV_HK3c\gamma1$ and $pEV_HK11c\gamma1$, respectively.

The center base of codon 64 which is affected less often (mutated in 4 of 17 genes) has been exchanged for two different bases specifying two different amino acids. If the amino acid exchanges were driven by selective forces, one would expect to observe a trend pointing in the opposite direction, i.e., to find a preponderance of parallel amino acid exchanges encoded by different triplets. (For example, the strongly selected leucine that replaces tryptophan at position 33 constitutes the only alternative amino acid inserted in this position in the anti-NP response, and has been seen encoded by any one of three different triplets; i.e., TTG, TTA, and CTG [reference 3, and Fig. 3]). The isoleucine replacing methionine in codon 34 is specified by two different triplets in 4 of the 17 genes depicted in Fig. 3. Previous studies with engineered antibodies, however, argue against the involvement of selective forces also in this exchange. The presence of isoleucine at position 34 in a secondary response antibody in which the position 33 mutation had been reverted to germline did not suffice to rescue the antibody from a 10-fold loss in affinity, thereby returning it to the affinity range of primary response antibodies (10). The consistent replacement of methionine by isoleucine at codon 34 may therefore result from the wobble nature of the base constituting a mutational hotspot at this position. Mutations presumably resulting from hotspot sites have also been described for other antigenic systems (28, 29). The suggestion has been made that mutational hotspots are associated with special features of the primary sequence or secondary structure of a DNA template that induce the polymerizing enzyme to pause and concomitantly relax its fidelity (30).

The recurrent mutation of particular nucleotides in codons 31, 34, 60, 61, 64, and 65 in the genes depicted in Fig. 3 is evident also in the genes of hybridoma-derived secondary response anti-NP antibodies (8, 9). Of the overall 21 different replacement mutations that fall into CDR1 and CDR2 in the present set of V_H sequences, 13 have also been identified in the NP-binding population of secondary response hybridomas.

Discussion

Predominance of Germline-encoded Specificities in the Peripheral B Cell Pool. It had been shown in earlier work that mitogen-activated peripheral B cells of the mouse, including cells of the Ly1 B subset, as well as B cells driven into primary antibody response, express germline-encoded V region genes (1, 31, 32). This suggested that the preimmune antibody repertoire in the mouse is essentially germline encoded. However, the earlier experiments invariably involved activation of subsets of B cells that might not be representative for the overall B cell population, the major fraction of which consists of long-lived, presumably selected cells (33). This problem is overcome in the present study in which V_H region genes are isolated from peripheral B cells independent of their state of activation. We find evidence for somatic hypermutation exclusively in the subset of V region genes with V_HDJ_H combinations characteristic for anti-NP antibodies. In contrast, 23 functionally rearranged genes with extensive combinatorial and junctional diversity (Fig. 2), and unrelated to the anti-NP response, are virtually free of mutations. We consider these 23 sequences a representative sample of the V genes expressed in the peripheral B cell pool and conclude that most cells in this population express germline-encoded antibody specificities. This conclusion is further supported by recent experiments in which 32 different, functionally rearranged V_H genes were cloned directly from splenic B cells of unimmunized mice and found to be virtually unmutated (Gu, H., I. Förster, W. Müller, and K. Rajewsky, manuscript in preparation).

Size of the NP-specific Memory B Cell Compartment in NP-primed Mice. Given that most peripheral B cells in the mouse express germline-encoded antibodies, memory B cells, known to express somatic antibody mutants, must represent a minor B cell subset. The consideration of known parameters of the splenic B cell repertoire in C57BL/6 mice in conjunction with the data obtained in the course of this study permit a rough estimate of the size of the NP-specific memory B cell pool 6 wk after priming. Approximately 3% of the 4×10^7 splenic B cells in C57BL/6 mice express λ1 L chains (reference 34; and this paper). We assume that 64% of these cells, as of LPS-reactive splenic B cells (35), express V_H genes of the group 1, and 1–2% of these 64% express the V_H gene V186.2. This latter estimate is based on our recent analysis of group 1 V_H genes and their expression in strain C57BL/6 through cDNA amplification and sequencing (Gu, H., et al., manuscript in preparation). The number of splenic B cells expressing λ1 L chains in association with V186.2 should then be $\sim 10^4$. When we transformed bacteria with the DNA of amplified rearranged V_H genes from λ1-positive spleen cells and sequenced randomly picked clones, we found that ~20% (4/19) of the V186.2 genes were linked to the D element DFL16.1 and were hypermutated. All four of these hypermutated genes encoded three consecutive tyrosines at the V_HD border, as is typical of NP-specific antibodies (data not shown). Based on this observation one arrives at an estimated number of 2×10^3 memory cells that bear λ1 L chains and express the H chain gene V186.2 joined to the D element DFL16.1 per spleen. From the analysis of hybridoma-derived secondary anti-NP antibodies, it is known that approximately half of the memory B cell precursors express λ1 L chains in association with V_H186.2 H chains that are joined to DFL16.1 and express three consecutive tyrosines at the beginning of CDR3. Taking this into account, the number of memory cells induced by the hapten NP in the spleen of primed mice should be on the order of $2–4 \times 10^3$. This frequency compares well with the frequency of antigen-specific memory cells that was measured late after a single immunization with the (multideterminant) protein PE. In two independent studies it was found to lie in the range of $1–5 \times 10^4$ per spleen (36, 37).

The Antigen-induced Memory B Cell Compartment Is Dominated by Antigen-selected High Affinity Binders. The somatically mutated, rearranged V186.2 genes, which we isolated from λ1 L chain-expressing spleen cells of NP-primed mice, strikingly resemble those expressed in secondary response anti-NP antibodies in terms of both rearrangement (i.e., CDR3) and frequency and characteristics of somatic mutations. Most importantly, the amino acid exchange, which represents the key mutation in the affinity maturation of the anti-NP response, is shared by the same proportion of genes (70%) in both collections. We conclude from this result that in a first approximation, the secondary anti-NP response is a direct readout of the somatic antibody mutants that have accumulated in the memory compartment.

This result supports our earlier demonstration that memory B cells recruited into the secondary response do not undergo further somatic mutation and produce a stable response of high affinity antibodies (5). Whether memory B cells can, alternatively, propagate themselves by reentering the pathway of hypermutation and selection (38) remains an open question.

The efficient and selective accumulation of cells expressing somatic antibody mutants with a high affinity for the immunizing antigen in the memory compartment has important biological implications. It is clear from the high rate of somatic mutation that memory cell generation must be accompanied by the production of antibody mutants that have changed their specificity away from the antigen. The present study demonstrates that the memory compartment is efficiently purged from cells expressing such antibodies. Apparently, antigenic selection is the key element in memory cell generation, and consequently, antigen-binding loss mutants only rarely make it to the stage of a long-lived memory cell. The persistence of such mutants in the memory compartment (5) is probably a rare event, triggered perhaps by the acquisition of a particular binding specificity.

The receptor repertoire of the B cells in the memory compartment thus focuses on the immunizing antigen, and high affinity binding to the immunizing antigen is its common denominator. However, the repertoire is also characterized by an extreme heterogeneity, due to the extent to which the antibodies are diversified by somatic mutation. Because of the high rate at which the hypermutation mechanism operates in the cells, affinity-selected mutations are almost invariably introduced in the context of other replacement mutations. Overall, this might be an efficient way for the systems to

acquire a broad range of crossreactivities as it would seem desirable in the defense against repeated infections by closely related microbes. Still, immunity to a crossreactive agent would be carried by antibodies with a better fit to the original immunogen, quite in the sense of "original antigenic sin" (39).

We are grateful to C. Berek and T. Simon for comments on the manuscript, U. Ringeisen for the art work, and G. Schmall for secretarial help.

This work was supported by the Deutsche Forschungsgemeinschaft through SFB 243 and by the Fazit Foundation. Ursula Weiss was the recipient of a Deutsche Forschungsgemeinschaft fellowship.

Address correspondence to Ursula Weiss, Institute for Genetics, University of Colonge, Weyertal 121, D-5000 Köln 41, Federal Republic of Germany.

Received for publication 25 June 1990.

References

1. Kocks, C., and K. Rajewsky. 1989. Stable expression and somatic hypermutation of antibody V regions in B-cell developmental pathways. *Annu. Rev. Immunol.* 7:537.
2. McKean, D., K. Hüppi, M. Bell, L. Gerhard, and M. Weigert. 1984. Generation of antibody diversity in the immune response of Balb/c mice to influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA.* 81:3180.
3. Allen, D., A. Cumano, R. Dildrop, C. Kocks, K. Rajewsky, N. Rajewsky, J. Roes, F. Sablitzky, and M. Siekevitz. 1987. Timing, genetic requirements, and functional consequences of somatic hypermutation during B-cell development. *Immunol. Rev.* 96:5.
4. Shlomchik, M.J., A. Marshak-Rothstein, C.B. Wolfowicz, T.L. Rothstein, and M. Weigert. 1987. The role of clonal selection and somatic mutation in autoimmunity. *Nature (Lond.).* 328:805.
5. Siekevitz, M., C. Kocks, K. Rajewsky, and R. Dildrop. 1987. Analysis of somatic mutation and class switching in naive and memory B cells generating adoptive primary and secondary responses. *Cell.* 48:757.
6. Brüggemann, M., H.-J. Müller, C. Burger, and K. Rajewsky. 1986. Idiotypic selection of an antibody mutant with changed hapten binding specificity, resulting from a point mutation in position 50 of the heavy chain. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:1561.
7. Cumano, A., and K. Rajewsky. 1985. Structure of primary anti-(4-hydroxy-3-nitrophenyl)acetyl (NP) antibodies in normal and idiotypically suppressed C57BL/6 mice. *Eur. J. Immunol.* 15:512.
8. Cumano, A., and K. Rajewsky. 1986. Clonal recruitment and somatic mutation in the generation of immunological memory to the hapten NP. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2459.
9. Blier, P.R., and A. Bothwell. 1987. A limited number of B cell lineages generates the heterogeneity of a secondary immune response. *J. Immunol.* 139:3996.
10. Allen, D., T. Simon, F. Sablitzky, K. Rajewsky, and A. Cumano. 1988. Antibody engineering for the analysis of affinity maturation of an anti-hapten response. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1995.
11. Reth, M., T. Imanishi-Kari, and K. Rajewsky. 1979. Analysis of the repertoire of anti-(4-hydroxy-3-nitrophenyl)acetyl (NP) antibodies in C57BL/6 mice by cell fusion. II. Characterization of idiotopes by monoclonal anti-idiotope antibodies. *Eur. J. Immunol.* 9:1004.
12. Miltenyl, S., W. Müller, W. Weichel, and A. Radbruch. 1990. High gradient magnetic cell separation with MACS. *Cytometry.* 11:231.
13. Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
14. Hanahan, D., and M. Meselson. 1983. Plasmid screening at high colony density. *Methods Enzymol.* 100:333.
15. Ichihara, Y., H. Hayashiba, S. Myazawa, and Y. Kurasawa. 1989. Only D_{FL16}, D_{SP2}, and D_{Q52} gene families exist in mouse immunoglobulin heavy chain diversity gene loci, of which D_{FL16} and D_{SP2} originate from the same primordial gene. *Eur. J. Immunol.* 19:1849.
16. Simon, T., and K. Rajewsky. 1988. "Enhancer-constitutive" vectors for the expression of recombinant antibodies. *Nucleic Acids Res.* 16:354.
17. Kocks, C., and K. Rajewsky. 1988. Stepwise intraclonal maturation of antibody affinity through somatic hypermutation. *Proc. Natl. Acad. Sci. USA.* 85:8206.
18. Oi, V.T., S.L. Morrison, L.A. Herzenberg, and P. Berg. 1983. Immunoglobulin gene expression in transformed lymphoid cells. *Proc. Natl. Acad. Sci. USA.* 80:825.
19. Kendall, C., I. Ionesco-Matin, and G.R. Dreezman. 1984. Utilization of the biotin-avidin system to amplify the sensitivity of the enzyme-linked immunsorbent assay. *J. Immunol. Methods.* 56:329.
20. Reth, M., G.J. Hämmerling, and K. Rajewsky. 1978. Analysis of the repertoire of anti-NP antibodies in C57BL/6 mice by cell fusion. I. Characterization of antibody families in the primary and hyperimmune response. *Eur. J. Immunol.* 8:393.
21. Dildrop, R. 1984. A new classification of mouse V_H sequences. *Immunol. Today.* 5:85.
22. Kabat, E.A., T.T. Wu, H. Bilofsky, M. Reid-Müller, and H. Perry. 1983. Sequences of Proteins of Immunological Interest. U.S. Department of Health and Human Services. U.S. Public Health Service, National Institutes of Health, Bethesda, MD. 511 pp.
23. Saiki, R.K., D.H. Gelfand, S. Stoffel, S.S. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed

enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC).* 239:487.
24. Tindall, K.R., and T.A. Kunkel. 1988. Fidelity of DNA synthesis by the Thermus Aquaticus DNA polymerase. *Biochemistry.* 27:6008.
25. Pennell, C.A., L.W. Arnold, G. Haughton, and S.H. Clarke. 1988. Restricted Ig variable region gene expression among Ly-1$^+$ B cell lymphomas. *J. Immunol.* 141:2788.
26. Bothwell, A.L.M., M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, and D. Baltimore. 1981. Heavy chain variable region contribution to the NP^b family of antibodies: somatic mutation evident in a γ2a variable region. *Cell.* 24:625.
27. Imanishi, T., and O. Mäkelä. 1975. Inheritance of antibody specificity. II. Anti-(4-hydroxy-5-bromo-3-nitrophenyl)acetyl in the mouse. *J. Exp. Med.* 141:840.
28. Berek, C., and C. Milstein. 1987. Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.* 96:23.
29. Levy, S., E. Mendel, S. Kon, Z. Avnur, and R. Levy. 1988. Mutational hot spots in Ig V region genes of human follicular lymphomas. *J. Exp. Med.* 168:475.
30. Seidman, M.M., A. Bredberg, S. Seetharam, and K.H. Kraemer. 1987. Multiple point mutations in a shuttle vector propagated in human cells. Evidence for an error-prone DNA polymerase activity. *Proc. Natl. Acad. Sci. USA.* 84:4944.
31. Manser, T. 1987. Mitogen-driven B cell proliferation and differentiation are not accompanied by hypermutation of immunoglobulin variable region genes. *J. Immunol.* 139:234.
32. Förster, I., H. Gu, and K. Rajewsky. 1988. Germline antibody V regions as determinants of clonal persistence and malignant growth in the B cell compartment. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:3693.
33. Förster, I., and K. Rajewsky. 1990. The bulk of peripheral B cell pool in mice is stable and not rapidly renewed from the bone marrow. *Proc. Natl. Acad. Sci. USA.* 87:4781.
34. Takemori, T., and K. Rajewski. 1981. Lambda chain expression at different stages of ontogeny in C57BL/6, Balb/c and SJL mice. *J. Immunol.* 11:618.
35. Dildrop, R., U. Krawinkel, E. Winter, and K. Rajewsky. 1985. V_H-gene expression in murine lipopolysaccharide blasts distributes over the nine known V_H-gene groups and may be random. *Eur. J. Immunol.* 15:1154.
36. Hayakawa, K., R Ishii, K. Yamasaki, T. Kishimoto, and R. Hardy. 1987. Isolation of high-affinity memory B cells: phycoerythrin as a probe for antigen-binding cells. *Proc. Natl. Acad. Sci. USA.* 84:1379.
37. Schittek, B., and K. Rajewsky. 1990. Maintenance of B cell memory by long-lived cells generated from proliferating precursors. *Nature (Lond.).* 346:749.
38. Berek, C., J.M. Jarvis, and C. Milstein. 1987. Activation of memory and virgin B cell clones in hyperimmune animals. *Eur. J. Immunol.* 17:1121.
39. Fazekas de St. Groth, S., and R.G. Webster. 1966. Disquisition on original antigenic sin. I. Evidence in man. *J. Exp. Med.* 124:331.