The Journal of Immunology

Delineation of Selective Influences Shaping the Mutated Expressed Human Ig Heavy Chain Repertoire

Thomas Dörner, Hans-Peter Brezinschek, Sandra J. Foster, Ruth I. Brezinschek, Nancy L. Farner and Peter E. Lipsky

J Immunol 1998; 160:2831-2841; ;
http://www.jimmunol.org/content/160/6/2831

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Delineation of Selective Influences Shaping the Mutated Expressed Human Ig Heavy Chain Repertoire[1]

Thomas Dörner, Hans-Peter Brezinschek, Sandra J. Foster, Ruth I. Brezinschek, Nancy L. Farner, and Peter E. Lipsky[2]

After Ag exposure, somatic hypermutation and subsequent selection play significant roles in shaping the peripheral B cell repertoire. However, the disparate impact of each process has not been completely delineated. To address this, the mutational patterns of a large panel of productive V_HDJ_H rearrangements of individual human B cells were analyzed and compared with those of a previously reported panel of nonproductive V_HDJ_H rearrangements. The productive V_H rearrangements exhibited a significantly lower mutational frequency and a significantly smaller number of replacement mutations than the nonproductively rearranged genes, suggesting that structural constraints of the Ig molecule and selective influences both impacted the repertoire, militating against replacement mutations. Positive selection favored a mean of four to six replacements in complementarity-determining region 1 (CDR1) and CDR2, and less than two replacements in the framework regions (FRs). In contrast, the negative impact of replacement mutations generated an increased number of silent mutations within both the CDRs and FRs of the productive repertoire accompanied by a net increase in the ratio of replacement to silent mutations in the CDRs compared with that in the FRs. Moreover, there was a negative influence on the distribution of amino acid changes resulting from mutations of highly mutable codons, such as AGY, TAY, GTA, and GCT, preferentially leading to conservative changes in the expressed Ig repertoire. The results are consistent with the conclusion that the expressed repertoire is limited, compared with the potential generated by the mutational machinery, by the dual requirements of avoiding autoreactivity and satisfying structural constraints of an intact Ig molecule. ***The Journal of Immunology,*** **1998, 160: 2831–2841.**

Ig genes are assembled during early B cell development by the recombination of V, D, and J elements. During this process, diversity is further increased by the insertion of N-nucleotides as well as by random pairing of heavy and light chains. Each of these molecular events occurs before antigenic stimulation and, if they result in a functional Ig protein, provide the preimmune repertoire from which positive and negative selections generate the expressed Ig repertoire. Exposure to autoantigens or exogenous Ag stimulates additional molecular processes, including receptor editing and somatic hypermutation, that can add to the diversity of the Ig repertoire (1, 2). Finally, positive and negative selections based on the expression of an intact Ig molecule shape the expressed B cell repertoire.

Somatic hypermutation of Ig genes is unique to B cells and serves as the basis for avidity maturation during immune responses. The distribution and nature of mutations, with an increased ratio of replacement to silent mutations in the CDRs[3] of Ig genes, have classically been assumed to be an indication of Ag-mediated selection for higher avidity. However, more recent data suggest that the molecular targeting of the mutator mechanism favors the preferential occurrence of replacement mutations in the CDRs, independent of Ag- or autoantigen-mediated selection (3, 4).

Despite the molecular targeting of the mutator mechanism, subsequent positive and negative selections of B cells expressing mutated Ig genes play a major role in shaping the expressed immune repertoire. Negative selection shapes the repertoire when autoantibodies are generated by the somatic hypermutation process (5–7), whereas positive selection for increased avidity to exogenous Ag leads to expansion of B cells containing mutations that enhance binding (8).

A recent analysis of V_HDJ_H rearrangements expressed by individual human B cells revealed evidence of a deletion of B cells containing specific mutations (3). However, this appeared to relate to a loss of B cells expressing V_HDJ_H rearrangements with mutations that altered the structural integrity of the Ig molecule and not negative selection related to putative reactivity with autoantigens. Moreover, no evidence of increased numbers of replacement mutations in the CDRs of productive V_HDJ_H rearrangements was observed, suggesting that Ag-mediated selection for higher avidity was not sufficient to influence the overall expressed immune Ig repertoire.

The current analysis was conducted to determine whether evidence of positive or negative selection of the human V_H repertoire could be detected. To accomplish this, the distribution of mutations in a large panel of productively rearranged V_H gene sequences amplified from genomic DNA of individual B cells was analyzed and compared with that found in a previously reported panel of mutated nonproductive V_HDJ_H rearrangements (3). It was reasoned that the distribution of mutations in the productively rearranged V_H genes would reflect the combined influences of the mutator mechanism and the subsequent effects of selection, whereas mutations of the nonproductive rearrangements are not

Department of Internal Medicine and Harold C. Simmons Arthritis Research Center, University of Texas Southwestern Medical Center, Dallas, TX 75235

Received for publication July 17, 1997. Accepted for publication November 24, 1997.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by National Institutes of Health Grant AI31229, Deutsche Forschungsgemeinschaft Grant Do 491/2-1 (to T.D.), and Erwin-Schrödinger stipends J0715 and J0929 (to H.-P.B.).

[2] Address correspondence and reprint requests to Dr. Peter E. Lipsky, Department of Internal Medicine, Harold C. Simmons Arthritis Research Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX. 75235–8884. E-mail address: peter.lipsky@email.swmed.edu

[3] Abbreviations used in this paper: CDR, complementarity-determining region; FR, framework region; R:S ratio, replacement mutation:silent mutation ratio.

0022-1767/98/$02.00

influenced by selection and, therefore, would indicate the molecular targeting of the mutator itself. Differences between the distribution of mutations should, therefore, be a direct indication of the influence of selection, both positive and negative. The data provide evidence that the expressed Ig heavy chain repertoire is selected by both Ag (as indicated by biases in the number and the nature of mutations in the CDRs of productively rearranged V_HDJ_H genes) as well as the likely structural constraints on the expression of a functional Ig molecule.

Materials and Methods

Rearranged V_H genes of individual B cells were specifically amplified from genomic DNA by primer extension preamplification followed by nested PCR and thereafter directly sequenced as previously described (9). Sequences were analyzed using the V Base Directory (10) to identify the respective germline genes. All 232 mutated productive V_HDJ_H rearrangements of 350 IgM^+ B cells were analyzed and compared with a previously reported panel of 37 mutated nonproductive V_HDJ_H rearrangements obtained from B cells of the same two donors and one additional donor (3, 11). Despite the previously reported evidence of negative selection of the V_H4 family and positive selection of the V_H3 family in the expressed repertoire (11), no significant difference in the distribution of V_H family members was observed when the mutated productive and nonproductive repertoires were compared (data not shown). Sequences are accessible in the EMBL, GenBank, and DDBJ Nucleotide Sequence Database (accession no. Z80363–Z80770, X87006, X87008–10, X87012–14, X87016–20, X87022–28, X87035, X87039, X87042, X87044–48, X87050–58, X87060–69, X87071–75, X87080, and X87082). Because of the nature of the internal primers used for amplification, mutations could not be detected in the first 12 codons of FR1; therefore, analysis of this region was omitted. Codon position 94 of FR3 (usually AGA or AGG) frequently exhibited evidence of exonuclease trimming and subsequent N-nucleotide addition; therefore, changes were not considered to be mutations. Shared mutations in the two clonally related sequences (BF2N2g3B07 and BF2N2g3F07) were considered only once in the analysis.

Determination of Taq polymerase fidelity and the frequency of potential sequence errors

All sequences in this study were analyzed by direct sequencing of PCR products. The maximal error rate of the amplification and sequencing technique has been documented to range from 1.2×10^{-3} to 1.7×10^{-3} mutations/bp or about 0.5 error/V_H gene segment (3). This number is in agreement with the error rate calculated from sequencing Ig cDNA from hybridomas and splenic Ab-forming cells (12), whereas a slightly lower error rate of 0.7×10^{-3} was obtained from sequencing a murine κ transgene (13). Further analysis of the error rate of the single cell PCR technique indicated that when a known Ig sequence (Vλ1 rearrangement L18F) of 238 bp was subjected to multiple preamplifications and subsequent nested amplifications ($n = 81$), no errors were detected in the resulting copies (19,261 bp). Thus, few, if any, of the nucleotide changes encountered in this analysis can be ascribed to amplification errors.

Statistical analysis

Sequences were analyzed with the χ^2 test to compare the differences in the frequencies of mutations and the R:S ratios within specific V_H regions. Where indicated the Bonferroni correction was used to correct for the possible influences of analysis of multiple variables (14). The goodness of fit χ^2 test was used to compare the nature of mutational events and the mutational events per codon to their expected mutational frequencies. These frequencies were calculated based on the random chance of a codon being mutated, calculated as 8.10% for productive and 9.00% for nonproductive V_H rearrangements. $p < 0.05$ was considered statistically significant.

The random chance that replacement mutations would lead to uncharged polar, nonpolar, basic, acidic amino acid, or silent amino acid substitutions was calculated based on the distribution found in the nonproductive rearrangements. Replacements with uncharged, nonpolar, basic, or acidic amino acids and silent mutations were predicted to be 45, 23, 16, 3, and 12%, respectively, for AGC; 27, 47, 20, 0, and 7%, respectively, for AGT; 35, 20, 15, 15, and 15%, respectively, for TAC; 21, 43, 7, 7, and 21%, respectively, for TAT; 0, 67, 0, 0, and 33%, respectively, for GTA; and 33, 53, 0, 7, and 7%, respectively, for GCT. These expected values were then used as comparators for the actual numbers of amino acid substitutions observed using the goodness of fit χ^2 test. This test was used to compare specific codon mutations leading to putative amino acid substitutions between productive and nonproductive rearrangements.

To compare the total numbers of mutational events occurring in the CDRs and FRs of certain rearranged V_H genes, the Spearman rank correlation and linear regression were applied. In this analysis, potential base pair changes introduced by exonuclease trimming and N-nucleotide addition at the 3′ end of FR3 were excluded.

Table I. *Mutational frequency of V_HDJ_H rearrangements*

	Mutational Frequency			
	Productive rearrangements		Nonproductive rearrangements[a]	
	Total	Mutated	Total	Mutated
Donor 1	1.3×10^{-2}[b]	2.0×10^{-2}[c]		
Donor 2	2.2×10^{-2}[b]	3.3×10^{-2}[c]		
Total	2.0×10^{-2}	3.1×10^{-2}[d]	2.0×10^{-2}	3.8×10^{-2}[d]

[a] Mutations in nonproductive rearrangements have been reported elsewhere (3).
[b, c] $p < 0.001$, significant difference between the mutational frequencies of donor 1 and donor 2.
[d] $p < 0.001$, significant difference between the frequencies of mutations of productive and nonproductive rearrangements.

Results

Within the productively rearranged V_HDJ_H genes amplified from 350 individual IgM^+ B cells, 232 (66.3%) contained base pair changes compared with their respective germline genes and were further analyzed. Within these genes, 42 of 67 (62.7%) mutated V_HDJ_H rearrangements were obtained from a 24-yr-old Hispanic male donor, whereas 190 of 283 (67.1%) of the mutated V_HDJ_H rearrangements were derived from a 45-yr-old Caucasian male donor. All together, 1,819 individual mutations were analyzed. The mutational frequency within the mutated sequences was 3.1×10^{-2} (1,819 mutations/58,494 bp), whereas the overall mutational

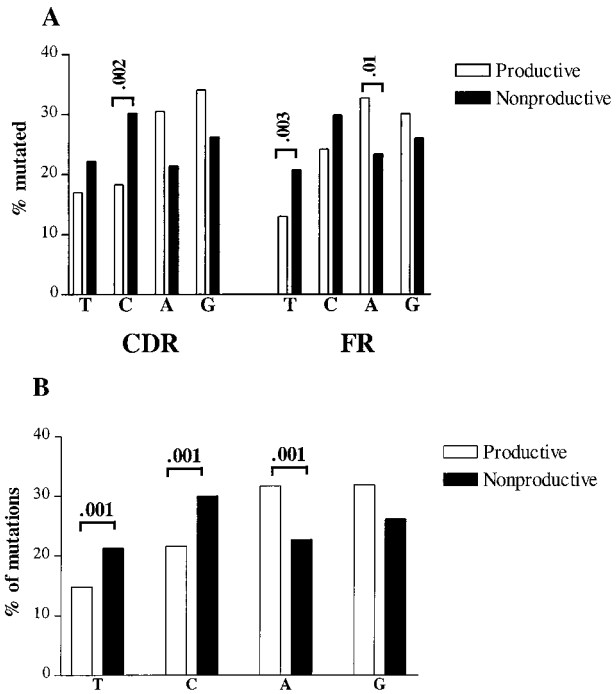

FIGURE 1. Distribution of mutated nucleotides of productive and nonproductive rearrangements in the CDRs and FRs (*A*) and in the overall V_HDJ_H rearrangements (*B*). Numbers and brackets indicate significant differences (by χ^2 test with Bonferroni correction). Mutations in nonproductive rearrangements have been reported previously (3).

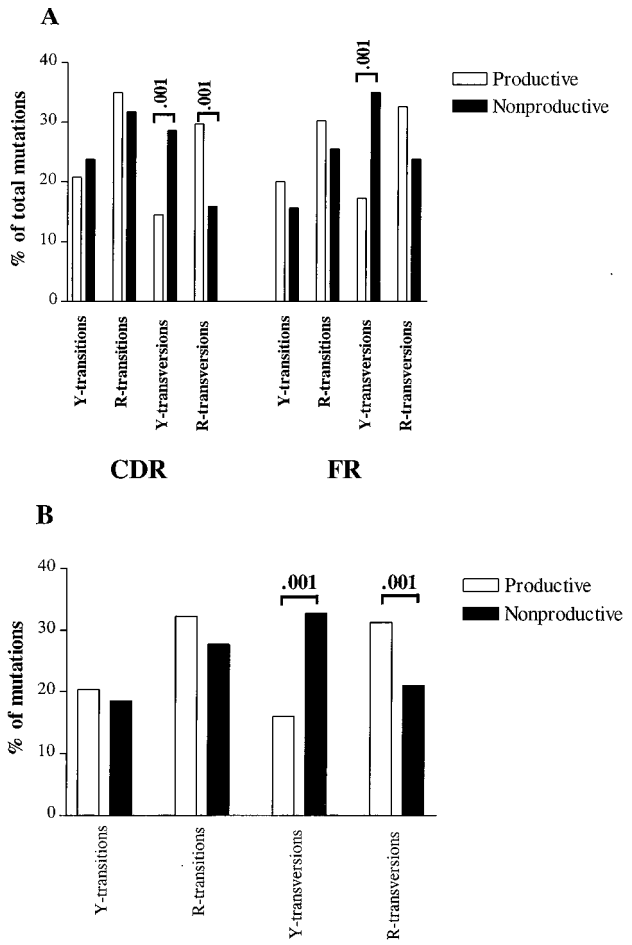

FIGURE 2. Nature of base pair substitutions in productive and nonproductive V_HDJ_H rearrangements in the CDRs and FRs (*A*) as well as in the overall rearrangements (*B*). Numbers and brackets indicate significant differences (by χ^2 test with Bonferroni correction). The nature of base pair mutations in nonproductive rearrangements has been reported previously (3).

frequency (mutated nucleotides per total base pairs analyzed) was 2.0×10^{-2} (1,819 mutations/90,776 bp). A previous mutational analysis (3) of nonproductive rearrangements with most sequences from these same donors documented 357 base pair changes in a total of 9,498 bp of mutated V_HDJ_H genes (mutational frequency, 3.8×10^{-2}) and in a total of 17,909 bp of V_HDJ_H analyzed (overall mutational frequency, 2.0×10^{-2}). The difference between the mutational frequencies of mutated V_HDJ_H rearrangements was statistically significant ($p < 0.001$, by χ^2 test), whereas the overall mutational frequencies did not differ between the productive and nonproductive repertoires (Table I).

The 42 productively rearranged V_HDJ_H genes obtained from the younger donor contained 218 nucleotide substitutions occurring within 182 codons (10,622 bp of mutated genes and 16,912 total bp), whereas the 190 mutated productively rearranged V_HDJ_H genes from the older donor contained 1601 mutations in 1,388 codons (47,872 bp of mutated genes and 73,864 total bp). Therefore, the mutational frequencies of mutated productive V_HDJ_H rearrangements were 2.0×10^{-2} (younger donor) and 3.3×10^{-2} (older donor), respectively, whereas the overall mutational frequencies were 1.3×10^{-2} and 2.2×10^{-2}, respectively. The differences between the mutational frequencies of these donors were statistically significant ($p < 0.001$, by χ^2 test).

In an attempt to delineate the impact of selection, we compared the nature of base pair changes in the productive and nonproductive repertoires. In this analysis, mutations of productive V_HDJ_H rearrangements affected purines more often (64.6% in CDRs and 62.8% in FRs), whereas mutations of T occurred with the lowest frequency (17 and 13% of all mutations in the CDRs and FRs, respectively) regardless of the V_H region analyzed (Fig. 1*A*). By contrast, in the CDRs and FRs of nonproductive rearrangements, there were no significant differences in the frequency of mutations of any nucleotide, as previously noted (3). Comparison of the affected nucleotides in the productive and nonproductive V_HDJ_H rearrangements provided evidence of the impact of selective influences. Thus, mutations of A were significantly more frequent, and mutations of T were significantly less frequent in the FRs of productive compared with nonproductive V_HDJ_H rearrangements, whereas C was mutated significantly less frequently in CDRs of productive rearrangements (Fig. 1*A*). Overall, T and C mutations appeared significantly less often, and A mutations appeared significantly more frequently in the productive compared with the nonproductive V_HDJ_H rearrangements, suggesting that their distributions were influenced by negative and positive selections, respectively (Fig. 1*B*).

Comparison of the nature of base pair mutations showed that R transversions and R transitions were most frequent within the productive repertoire, with the former comprising 64.7% of the mutations in the CDRs and the latter comprising 62.8% of the mutations in the FRs (Fig. 2*A*). In general, Y transversions occurred at the lowest frequency within the productively rearranged V_HDJ_H genes. Compared with the nonproductive repertoire, Y transversions were significantly less frequent in the CDRs and FRs of the

Table II. *Frequency of multiple mutations in the same codon*

	Mutations/Codon		
	1 event	2 events	3 events
Productive rearrangements	1345 (85.7%)[a]	201 (12.8%)*	24 (1.5%)**
Nonproductive rearrangements	279 (89.1%)	24 (7.7%)*	10 (3.2%)**

[a] Percentage of B cells with mutated V_HDJ_H rearrangements with the respective number of mutations.

*, ** Significant difference between the productive and nonproductive repertoire ($p < 0.05$).

Table III. *Ratio of replacement to silent mutations in specific regions of productive and nonproductive V_HDJ_H rearrangements*

	No. of Replacement/Silent Mutations (Ratio)		
V_H Region	Productive rearrangements ($n = 232$)	Nonproductive rearrangements ($n = 37$)[a]	p[b]
FR1	108:52 (2.1)	25:4 (6.3)	< 0.05
CDR1	159:32 (5.0)	21:3 (7.0)	NS
FR2	61:44 (1.4)	15:6 (2.5)	NS
CDR2	406:86 (4.7)	80:10 (8.0)	NS
FR3	389:233 (1.7)	120:29 (4.1)	< 0.001
CDR	565:118 (4.8)[c]	101:13 (7.8)	NS
FR	558:329 (1.7)[c]	160:39 (4.1)	< 0.001
total	1123:447 (2.5)	261:52 (5.0)	< 0.001

[a] R:S ratios found in nonproductive rearrangements have been reported elsewhere (3).

[b] χ^2 test comparing the R:S ratio of productive and nonproductive V_HDJ_H rearrangements.

[c] Significant difference between the R:S ratio in the CDRs and FRs ($p < 0.0001$, χ^2 test) of productive rearrangements. This difference was not significant in the nonproductive repertoire ($p = 0.061$).

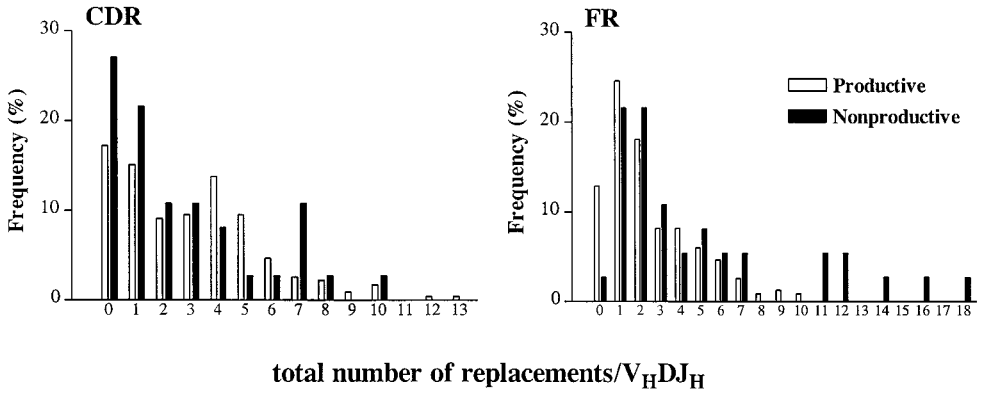

FIGURE 3. Distribution of the total number of replacement mutations in the CDRs and FRs in the productive and nonproductive V_HDJ_H rearrangements, indicating positive selection for four to six replacements in the CDRs and negative selection against more than two mutations in the FR in the productive repertoire.

productively rearranged genes, whereas R transversions were found significantly more often in the CDRs of the productive rearrangements. The remaining mutational changes did not differ significantly. Of note, Y transversions were less frequent, and R transversions were more frequent in the productive repertoire, suggesting that selective influences shaped the repertoire (Fig. 2*B*).

In a further analysis, we determined whether the occurrence of multiple mutations within the same codon differed between productive and nonproductive V_HDJ_H rearrangements. As shown in Table II, two mutational events per codon occurred at a higher frequency in the productive rearrangements ($p < 0.05$), whereas three mutations per codon were more frequent in the nonproductive repertoire ($p < 0.05$).

An increase in the R:S ratios of CDRs is often taken as an indication of antigenic selection (15–17). Our recent analysis (3) of a small number of productive V_HDJ_H rearrangements revealed that the R:S ratio in the CDRs and FRs of nonproductive rearrangements was significantly greater than that in the productive rearrangements. Although the R:S ratio in the CDRs of nonproductive rearrangements was greater than that in the productive repertoire (7.8 vs 4.8), the difference was not statistically significant ($p = 0.117$; Table III). Moreover, a significantly greater R:S ratio was found in FR1 ($p < 0.05$) and FR3 ($p < 0.001$) as well as in the total FRs ($p < 0.001$) of nonproductively rearranged V_HDJ_H genes. Within the productive rearrangements, the R:S ratio in the CDRs was significantly greater (2.8-fold) than that found in the FRs ($p < 0.001$, by χ^2 test). This difference, however, was not noted within the nonproductively rearranged genes ($p = 0.061$), suggesting that selective influences operating on the CDRs and the FRs mitigated the tendency for replacement mutations to be eliminated from the productive repertoire.

Analysis of the distribution of the total number of replacements per gene clearly showed selective influences on individual genes. There was clear selection against cells with no mutations in the CDRs, whereas the frequency of mutations in CDRs of nonproductive and productive V_HDJ_H gene mutations clearly indicated that there was selection of cells with four to six replacement mutations in the CDRs (Fig. 3). Larger numbers of replacement mutations in CDRs appear to be selected against. In the FRs, more than one mutation was clearly negatively selected, whereas B cells with no mutations in FRs were clearly favored in the productive repertoire.

Correlation of mutational events occurring in CDRs and FRs

Because of the decreased R:S ratios within the FRs of productive V_HDJ_H rearrangements, the nature of the mutational events was examined in greater detail. When all productive rearrangements were examined, a significant correlation was noted between the number of mutations in CDRs and FRs of both productive and nonproductive rearrangements (Fig. 4). Overall, the slope of the line relating CDR to FR mutations was greater in the productive than in the nonproductive repertoire. This largely reflected retention of silent mutations within FRs and deletion of replacement mutations in both CDRs and FRs within the productive repertoire.

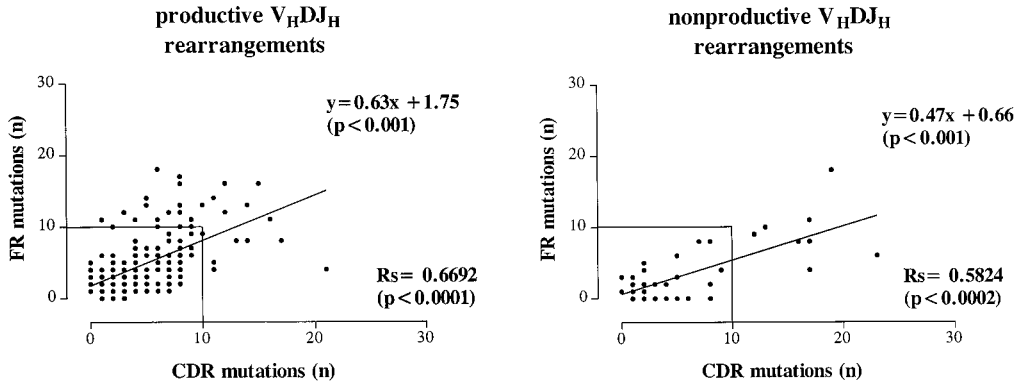

FIGURE 4. Comparison of mutational events occurring in CDRs and FRs of productive and nonproductive V_HDJ_H rearrangements. Spearman rank correlation (Rs) and linear regression were applied to compare the distribution of mutations in the productive and nonproductive V_H rearrangements.

Table IV. *Mutational frequencies of individual codons*[a]

Codon	Productive		Nonproductive[b]	
	Mutated	p^c	Mutated	p^c
AAA	0.20	<0.001	0.05	NS
AAT	0.19	<0.001	0.24	<0.025
AGA	0.14	<0.001	0.17	<0.01
AGC	0.24	<0.001	0.26	<0.001
AGT	0.16	<0.001	0.16	<0.025
GCT	0.14	<0.001	0.23	<0.001
GGT	0.15	<0.001	0.15	NS
GTA	0.19	<0.001	0.27	<0.005
TAC	0.11*	<0.001	0.19*	<0.001
ATA	0.15	<0.005	0.14	NS
AAC	0.12	<0.005	0.15	NS
CAG	0.01	<0.001	0.02	NS
CCT	0.01**	<0.001	0.07**	NS
CGC	0.01	<0.001	0.04	NS
CTC	0.03	<0.001	0.02	NS
CTG	0.03	<0.001	0.05	NS
GAC	0.03**	<0.001	0.06**	NS
GAG	0.04	<0.001	0.04	NS
GGA	0.02	<0.001	0.03	NS
GGC	0.04	<0.001	0.06	NS
GGG	0.03	<0.001	0.03	<0.025
GTC	0.02	<0.001	0.01	<0.025
GTG	0.05	<0.001	0.03	<0.025
TCC	0.02	<0.001	0.03	<0.01
TGG	0.02	<0.001	0.05	NS
TGT	0.02*	<0.001	0.07*	NS
TCT	0.01	<0.001	0.00	<0.005
CCA	0.03	<0.005	0.08	NS
TTG	0.01	<0.005	0.05	NS
TCG	0.00*	<0.005	0.05*	NS
GCC	0.05**	<0.005	0.15**	NS
CGA	0.02**	<0.005	0.13**	NS
TCA	0.08	NS	0.00	<0.05

[a] The mutational frequencies of specific codons in productive and nonproductive repertoires were compared.
[b] Mutations in nonproductive rearrangements have been reported elsewhere (3).
[c] The frequency of mutations of each codon was compared to the probability of mutation based on random chance (0.081 for productively rearranged V_HDJ_H genes and 0.090 for nonproductively rearranged V_HDJ_H genes) using the goodness of fit χ^2 test. *, ** significantly higher mutational frequency in the nonproductive compared to the productive repertoire (*$p < 0.05$, **$p < 0.01$).

Codon bias

To delineate whether certain mutated codons were more likely to be influenced by selection, the distribution of mutations in specific codons of productive rearrangements was examined and compared with that found in the nonproductive V_HDJ_H genes (Table IV). As previously noted, mutations were more frequent in some codons than in others. In general, the distribution of mutations in specific highly mutable codons, such as AGC, AGT, GCT, GGT, GTA, AAT and AGA, was comparable in productive and nonproductive rearrangements. Some differences were noted, however. For example, mutations of TAC codons were found significantly more often in the nonproductive than in the productive repertoire ($p = 0.018$, by χ^2 test). This was specifically related to the frequency of mutations in the FRs (see below). Although there were apparent differences between the frequencies of mutations of some codons, such as AAA, AGA, AGC, GCT, GTA, AAC, and AAT, in productive and nonproductive repertoires, these differences were not statistically significant. Of note, some codons that frequently resided within FRs, such as CAG, CCT, CGC, CTC, CTG, GAC, GAG, GGA, GGC, TGG, TGT, CCA, TTG, TCG, GCC, and CGA, were mutated significantly less often in the productively rearranged repertoire. On the other hand, mutations in infrequently mutated codons, such as GGG, GTC, GTG, TCC, TCT, and TCA, appeared in the nonproductive repertoire more often than expected. When the frequencies of mutations of these less frequently mutated codons were analyzed, CCT ($p < 0.002$), GAC ($p < 0.004$), TGT ($p < 0.015$), TCG ($p < 0.025$), GCC ($p < 0.001$), and CGA ($p < 0.001$) were found to be mutated significantly more often in the nonproductive than in the productive repertoires, suggesting that mutations of these codons are not tolerated in the expressed V_H repertoire.

Mutations of AGC and indications of selection

As previously noted, the triplet AGC was the most frequently mutated sequence, as either a codon or a noncoding triplet. Thus, 24% of AGC codons were mutated in the productive and 26% in the nonproductive V_HDJ_H rearrangements. As a codon, AGC contained 11.7% of all base pair changes in the productive and 10.4% of all mutated nucleotides in the nonproductive repertoire. Somewhat different from the common preference for mutations at the third position within the productive repertoire, mutations of the second position of AGC were most frequent ($p < 0.001$) in both the productive and nonproductive repertoires (Table V). Moreover, G mutations of AGC appeared to be positively selected in the CDRs of productive rearrangements ($p < 0.02$, by χ^2 test), whereas no difference was noted between the frequencies of G mutations in FRs of productive and nonproductive V_HDJ_H rearrangements ($p = 0.068$).

Table V. *Analysis of AGC mutations*

	Distribution of Mutations in AGC[a]						
				AGC in RGYW		AGC outside of RGYW	
	A	G	C	Total[b]	% mutated[c]	Total[b]	% mutated[c]
Productive							
CDR	26 (17%)[d]	93 (61%)[e]	34 (22%)	95	99	5	1
FR	7 (11%)	30 (49%)	24 (39%)	54	80[f,g]	46	20[g]
Nonproductive							
CDR	7 (29%)	10 (41%)[e]	7 (29%)	96	99	4	1
FR	2 (15%)	10 (77%)	1 (8%)	66	54[f]	34	46

[a] Total number and location of mutations within the AGC codons depending on the location within CDRs and FRs.
[b] Percentage of AGC within or outside of RGYW in germline configuration.
[c] Percentage of mutated AGC codons that were included within RGYW sequences or outside of RGYW, respectively.
[d] Percentage of mutated position within the specific V_H region.
[e] Significant difference between productive and nonproductive rearrangements ($p < 0.02$).
[f] Significant difference between mutated AGC codons within RGYW motifs of productive and nonproductive V_H rearrangements ($p < 0.05$)
[g] Significant difference between the percentage of mutated AGC in RGYW motifs and the percentage of mutated AGC outside RGYW ($p < 0.001$).

Table VI. *Comparison of the R:S ratios of highly mutated codons in the productive and nonproductive V_HDJ_H rearrangements*

	Nonproductive Rearrangements						Productive Rearrangements						p^a	
	FR			CDR			FR			CDR				
Codon	R	S	R:S	R	S	R:S	R	S	R:S	R	S	R:S	FR	CDR
AGC	16	3	5.3	11	1	11.0	44	11	4.0	107	21	5.1	NS	NS
GCT	10	1	10.0	4	0	n.d.	18	12	1.5*	35	2	17.5*	NS	NS
TAC	18	2	9.0	4	1	4.0	16	16	1.0**	43	10	4.3**	0.01	NS
GTA	4	2	2.0	0	0	n.d.	9	17	0.5	3	1	3.0	NS	NS
ATT	2	0	n.d.	3	2	1.5	2	2	1.0	4	4	1.0	NS	NS
AAT	5	0	n.d.	0	0	n.d.	3	2	1.5	17	4	4.3	NS	NS
AGT	13	1	13	1	0	n.d.	15	4	3.8	70	6	11.7	NS	NS

[a] The relation of replacement vs silent mutations was compared in the CDRs and FRs between productive and nonproductive rearrangements using the χ^2 test (n.d. indicates that the value is not defined). *, ** Within the productive rearrangements, GCT and TAC manifested a significantly lower R:S ratio in the FRs than in the CDRs ($p < 0.003$), whereas the R:S ratios of these mutated codons did not differ in the nonproductively rearranged V_H genes.

AGC is part of the putative mutationally active RGYW motif (18). Therefore, the question arose of whether the increased frequency of AGC mutations depended upon its occurrence within the RGYW motif (Table V). The majority of mutations of AGC occurred when it was incorporated in an RGYW motif, regardless of whether it occurred in the FRs or CDRs. Only 15 of 214 (7.0%) mutated AGC codons occurred outside of RGYW motifs in the productively rearranged V_H genes. Within the nonproductive repertoire, 7 of 37 (18.9%) AGC mutations occurred outside of RGYW sequences. Of note, there was a greater likelihood that AGC would be mutated when it resided in an RGYW motif in productive compared with nonproductive V_HDJ_H rearrangements ($p < 0.02$, by χ^2 test). This was particularly noteworthy in the FRs (80% of mutated AGC in RGYW in the productive vs 54% in the nonproductive repertoire). The frequency with which mutated AGC codons were found within RGYW was significantly different between the FRs of productive and nonproductive rearrangements (Table V; $p = 0.04$, by χ^2 test). Therefore, positive selection of mutated AGC codons occurring within RGYW motifs was clearly evident in the FRs of productively rearranged genes. Of note, the overall occurrence of AGC as part of an RGYW motif was not significantly different in the CDRs and FRs of the productive and nonproductive repertoires. However, there was a significantly greater likelihood that AGC was contained within an RGYW germline motif in the CDRs vs the FRs ($p < 0.001$) in both productive and nonproductive rearrangements.

Specific codon mutations and selection

A more detailed analysis provided insight into the impact of selection on mutations of specific frequently mutable codons. Thus, a comparison of the R:S ratios of frequently mutated codons demonstrated higher R:S ratios for mutated AGY motifs in the FRs of productive V_HDJ_H rearrangements (Table VI) compared with the R:S ratios for all codons in the FRs of productive rearrangements (Table III). In addition, there were significantly greater R:S ratios in mutated TAC codons in the FRs of the nonproductive V_HDJ_H rearrangements compared with those in the productive repertoire ($p < 0.01$, by χ^2 test). Moreover, a significantly greater R:S ratio was found for mutated GCT ($p < 0.001$) and TAC ($p < 0.003$) codons in CDRs compared with FRs of productively rearranged V_HDJ_H genes, but not in nonproductive rearrangements.

Specific amino acid changes and selection

To address the potential role of selection in greater detail, the nature of the amino acid changes was analyzed. These detailed analyses were conducted for the highly mutable codons AGY, TAY, GCT, and GTA (Fig. 5). The underlying assumption of this analysis was that the distribution of replacement amino acids in the productive repertoire resulting from mutational events could be predicted from the changes occurring in the nonproductive V_HDJ_H rearrangements. In this regard, the distribution of conservative and nonconservative amino acid changes in the nonproductive repertoire resulting from replacement mutations was not different from that expected from random chance (Fig. 5). Comparison of the distribution of replaced amino acids resulting from AGC mutations (AGC encodes serine, an uncharged polar amino acid) in productive and nonproductive rearrangements, however, demonstrated increased substitutions by uncharged polar amino acids ($p < 0.005$) and a decreased frequency of substitutions by basic amino acids ($p < 0.001$) in the CDRs of productive rearrangements. No differences between productive and nonproductive rearrangements were noted for replacement AGC mutations within the FRs. A similar distribution of replacement mutations of AGT (encoding serine, an uncharged polar amino acid) in productive rearrangements was noted, with an increased frequency of substitutions by uncharged polar amino acids in both CDRs and FRs ($p < 0.005$ and $p < 0.05$, respectively) and less frequent substitution by nonpolar amino acids ($p < 0.05$ in CDRs).

Analysis of the distribution of TAC (encoding tyrosine, an uncharged polar amino acid) mutations indicated a significantly higher frequency of replacement basic amino acids in the CDRs ($p < 0.005$) of productive rearrangements, whereas a lower frequency of replacement uncharged polar and acidic FR mutations ($p < 0.05$) was noted. A higher rate of silent mutations ($p < 0.001$) in the FRs of productive rearrangements was also found. Analysis of TAC mutations showed significant differences in regard to resulting stop codons within the FRs (four stop codons; $p < 0.001$) as well as in the CDRs (one stop codon; $p < 0.05$) compared with the productive repertoire. By contrast, the distribution of replacement mutations of TAT was not different between productive and nonproductive rearrangements. Replacement mutations of GTA (encodes valine, a nonpolar amino acid) resulting in nonpolar amino acids were decreased in the FRs of productive rearrangements ($p < 0.001$), whereas silent mutations appeared at a high frequency in the FRs of productive rearrangements ($p < 0.001$). Finally, GCT (encodes alanine, a nonpolar amino acid) mutations exhibited a significantly decreased frequency of uncharged polar amino acids ($p < 0.005$) as well as increased silent mutations ($p < 0.001$) in the FRs of productively rearranged V_HDJ_H genes.

Mutational cold spots

Comparison of the localization of replacement mutations between the 232 productive and the 37 nonproductive V_H rearrangements

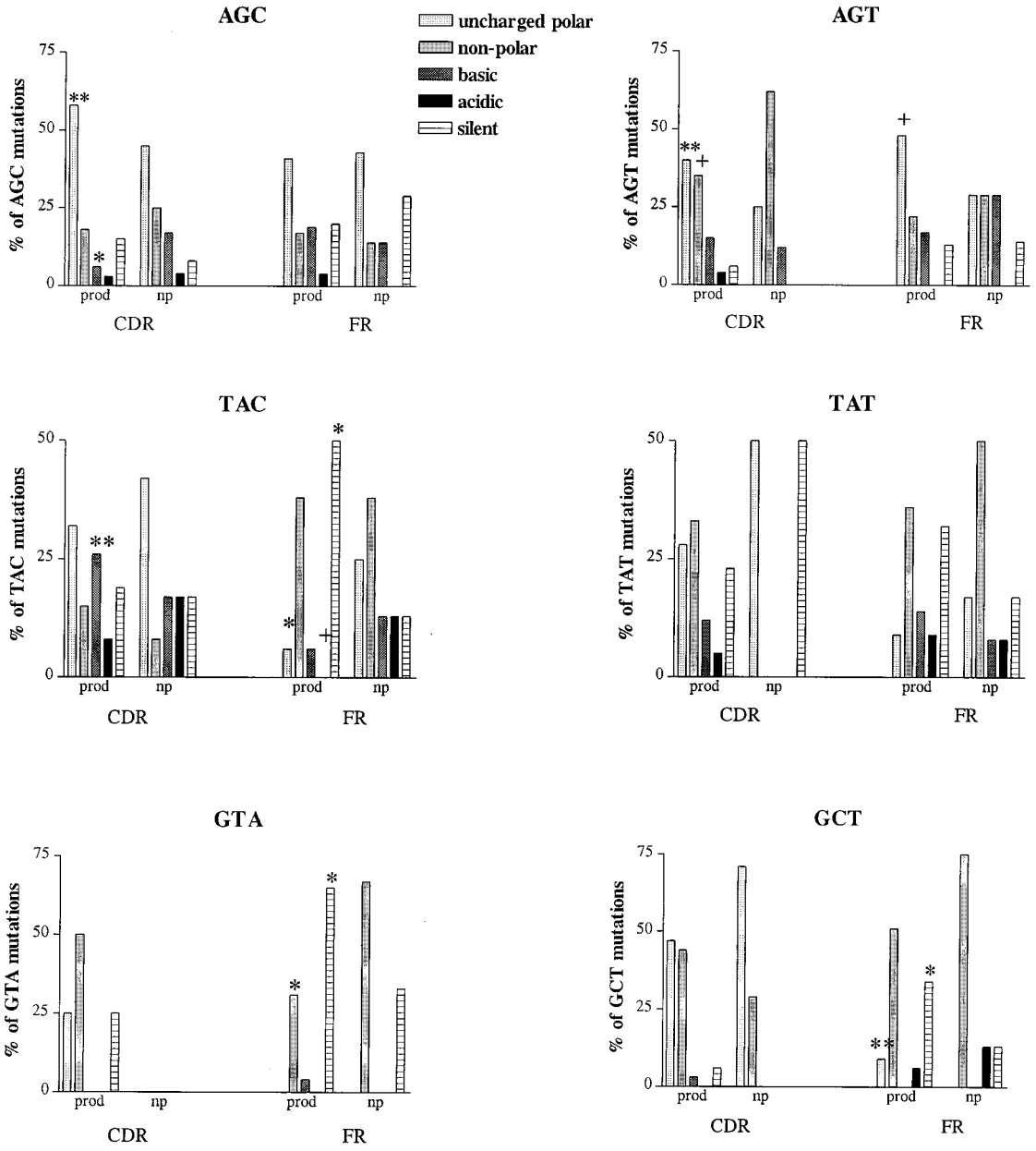

FIGURE 5. Distribution of amino acid changes resulting from mutations of AGY, TAY, GTA, and GCT codons. Statistical differences were analyzed based on the distribution found in the unselected nonproductive rearrangements by the goodness of fit χ^2 test. The underlying assumption was that mutational events within the nonproductive repertoire yield the unselected distribution of resulting amino acid changes. Random replacements with uncharged, nonpolar, basic, or acidic amino acids and silent mutations were calculated as described in *Materials and Methods*. * indicates $p < 0.001$; ** indicates $p < 0.005$; + indicates $p < 0.05$.

provided evidence for positions that could not tolerate replacement mutations. In this analysis, the R:S ratio was significantly less at positions 80 ($p < 0.025$, ATG, CTT, or CTG), 87 ($p < 0.01$, ACC or ACG), and 91 ($p < 0.025$, 100% TAC) in productive compared with nonproductive rearrangements. Whereas position 18 was not mutated in either repertoire, codon position 36 was mutated in one nonproductively rearranged V_H gene only. In the aggregate of the analyzed V_HDJ_H rearrangements, no specific position could be identified as favored by positive selection.

Delineation of the importance of the RGYW motif

As previously reported, AGY, TAC, GTA, and GCT are highly mutable codons, three of which reside within the larger RGYW domain that is thought to be a target for mutations (18). To investigate this in greater detail, mutations of RGYW were analyzed. Of note, the contributions of mutations of RGYW to the total number of mutations was not different between the productive (27.8%) and the nonproductive (23.5%) repertoires ($p = 0.35$). Whereas there was no significant difference in the frequency of mutated RGYW motifs between the CDRs (32.6%) and FRs (14.8%) of nonproductive rearrangements ($p = 0.182$, by χ^2 test), productive rearrangements exhibited a significantly higher frequency of RGYW mutations in the CDRs (36.3%) compared with the FRs (10.7%; $p < 0.001$).

Discussion

The current results clearly indicate that the expressed V_HDJ_H repertoire is negatively influenced by the loss of replacement mutations in the FRs and, to a lesser extent, in the CDRs that preferentially acquired four to six replacements. This is in accordance

with the findings of higher frequencies of first-position mutations in the nonproductive and third-position mutations in the productive repertoire (3). Moreover, selection clearly influenced the nature of amino acid replacements found in the expressed repertoire after their generation by the mutational machinery. Thus, when mutations affected frequently mutated codons, conservative changes appeared to be favored in the CDRs of the productive repertoire, and silent mutations seemed to be favored in the FRs. Overall, RGYW mutations appeared to be favored in CDRs and negatively selected in the FRs, whereas mutations of AGC incorporated in RGYW motifs were positively selected in the FRs. In contrast, TAC codons that were frequently mutated in the second position and occasionally occurred as part of RGYW sequences also appeared to be a preferential target of the mutator. However, replacement mutations of TAC in the FRs were negatively selected to a significant degree. These results strongly indicate the important role of selective influences in shaping and limiting the potential diversity of Ig receptors generated by the mutational machinery.

The results document that positive selection favors four to six replacement mutations in the CDRs, but less than two replacement mutations in FRs. This finding is most consistent with the conclusion that a limited number of replacement mutations per V_HDJ_H rearrangement provides the most favorable basis of positive Ag-mediated selection into the productive repertoire, whereas V_H rearrangements with larger numbers of replacement mutations in CDRs are less likely to be positively selected. This may reflect the greater likelihood of introducing an unacceptable amino acid change as the number of replacement mutations is increased. In contrast, none or one FR replacement mutation in the FRs per V_HDJ_H rearrangement appeared to be favored in the productive repertoire, probably indicating the loss of rearrangements with more FR mutations from the productive repertoire. This finding is likely to reflect the strong need for stability in the FR of an intact Ig molecule even though these regions can be subjected to hypermutation.

Mutational frequencies

The overall frequency of mutations was comparable between the productive and nonproductive repertoires, whereas the frequency of mutations in the mutated V_HDJ_H genes was significantly greater among the nonproductive rearrangements. This finding is consistent with the conclusion that there is preferential loss from the productive repertoire of more highly mutated V_HDJ_H rearrangements. In this context, it appears that V_HDJ_H rearrangements are limited in the number of tolerated mutations. As the number of replacement mutations exceeds six per V_HDJ_H rearrangement or six per aggregate of CDR1 and CDR2, there is a greater likelihood for the B cell to be lost from the productive repertoire. This is consistent with reports analyzing cDNA from V_H5 and V_H6 family members from peripheral B cells demonstrating a maximum of seven replacements (in one gene) in the CDRs of V_H6 (19) and a maximum of eight replacements in CDR1 and two of V_H5 family members (20). Similarly, seven replacement mutations were most frequent in V_H genes from IgM rheumatoid factor-producing hybridomas (6), whereas V_H genes from memory B cells also contained a maximum of seven replacements (21). The maximum of seven to eight replacements in a broad variety of specific Ag-selected V_H genes, including two to six replacements in the CDRs, is consistent with the conclusion that a limited number of replacement mutations provides the optimal balance to achieve positive selection for Ag binding with the minimal likelihood of the imposition of the negative impact of autoantigen binding or structural compromise of the integrity of the Ig molecule. Most replacement mutations in the FRs appear to be subjected to negative selection. As the FRs, in general, do not directly contribute to Ag binding capacity, most of the deletion of replacement mutations from FRs is likely to be related to amino acid substitutions that can alter the structural integrity of the Ig molecule.

R:S ratio

The comparison of R:S ratios between the panels of productive and nonproductive rearrangements also provided evidence for selection against replacement mutations in the FRs and to a lesser extent within the CDRs. The productively rearranged V_HDJ_H genes appear to have precisely defined R:S ratios (4.7 to 5.0 for the CDRs and 1.4 to 2.1 for the FRs). The ratio in the FRs of productive rearrangements was significantly reduced compared with that in the nonproductive repertoire. Moreover, the imprint of selection was documented by the significant difference in the R:S ratio found between the CDRs and FRs of productive rearrangements, which was not detected among the nonproductive V_HDJ_H rearrangements.

The R:S ratios of FR2 of both the productive and nonproductive repertoires exhibited the lowest R:S values, suggesting that replacement mutations within FR2 are infrequent, presumably because of structural constraints. FR2 may be particularly sensitive to replacement mutations because of its involvement in the interaction with light chains (22) and maintaining the stability of CDR1 and CDR2. Significant influences of negative selection were noted within FR1 and FR3 compared with FR2. The differences in R:S ratios in FR1 and FR2 between productive and nonproductive repertoires were somewhat different in this larger analysis than previously reported (3), but each of the previously noted trends was similar. Overall, negative selection of replacement mutations in the FRs in the productive repertoire was again noted to be significant, and this suggests that structural constraints shape the Ig repertoire, leading to well-defined R:S ratios for the FRs of productive V_HDJ_H rearrangements that are significantly lower than those of the CDRs. Although the R:S ratios in the CDRs of productively rearrangements were significantly lower than those in the nonproductive repertoire in our previous analysis ($p < 0.03$) (3), and the mean R/S value of the productively rearranged genes of the current study was lower than that of the nonproductive rearrangements (4.8 vs 7.8), the difference was not significant. Considering the observed bias for four to six replacements per CDR1 plus CDR2 in the productive repertoire, negative selection of sequences with less than four or more than six amino acid substitutions in these regions might further influence the final number of replacement mutations in the CDRs of productive rearrangements. In contrast to the FRs, significant negative selection of replacements in the CDRs was not observed. In summary, the data clearly indicate that an increase in R:S ratios in the CDRs of the expressed repertoire of Ig molecules does not relate solely to antigenic selection or necessarily to Ag-mediated selection at all.

Nature of base pair substitutions and selection

Of note, the current study provided evidence that selection influences the distribution of affected nucleotides in the CDRs and FRs. The occurrence of mutations of T in the FRs and that of C in the CDRs were decreased in the productive repertoire. A reduced frequency of mutations of T and enhanced mutations of purines has been ascribed to the mutational machinery by several reports (23–28), whereas the current study indicates marked additional influences of selection. In the current analysis, the preference for purine mutations was noted in the productive V_HDJ_H rearrangements. In the nonproductive repertoire, however, G and C were targeted for mutation. Selection subsequently decreased the frequency of C

mutations and increased the frequency of A mutations in the productive repertoire, leading to the previously noted purine bias (23–28). It is apparent from this analysis that this pattern could not be ascribed to the direct action of the mutational machinery.

The relation of elevated A and decreased T mutations is often taken as an indication for strand polarity of the mutational machinery (4, 24, 26, 28–34). The current results alter this conclusion, however, because of the apparent impact of overriding selective influences. This is of particular interest because one potential mechanism of hypermutation appears to be related to transcription (30, 31), whereas other models implicate an error-prone DNA repair (35). The current analysis of the nonproductive repertoire does not indicate that the mutational machinery exhibits strand polarity. On the contrary, the biases of A and T as well as G and C mutations found in the productive repertoire appear to arise largely from selective influences and thereby imply that apparent strand polarity is imposed by selection and is not necessarily evidence of biased mutator activity.

Delineation of the influence of negative selection on mutations in the FRs

Since it was noted that the FRs can potentially yield a considerable number of mutations with subsequent fine-tuning by selection, most of the mutations of less frequently mutated codons only appeared in the nonproductive repertoire. Importantly, replacement mutations of TAC codons were frequently eliminated from the productive repertoire, especially at position 91 ("cold spot"). Part of this is related to the introduction of stop codons, as five of the TAC mutations detected in the nonproductive repertoire would have resulted in stop codons, including four that occurred in the FRs. Moreover, mutations of CCT, GAC, TGT, TCG, GCC, and CGA codons appeared significantly more often in the nonproductive repertoire, indicating that mutations of these FR codons cannot be tolerated in the expressed repertoire. This suggests that few changes in the FRs can be tolerated by an intact Ig molecule. Almost all of the less frequently mutated codons reside within the FRs; therefore, the data suggest that the mutational machinery has evolved to avoid mutating them. When they are mutated, however, structural constraints of an intact Ig molecule probably lead to their deletion from the productive repertoire, presumably because B cells expressing such mutated Ig cannot survive.

Delineation of selection for AGC mutations in the FRs and CDRs

As previously noted (3, 24, 36), the codon AGC is the most frequently mutated codon in the productive as well as in the nonproductive repertoire. AGC appears to be highly targeted by the hypermutational machinery, whereas overall selection of mutated AGC codons was not noted. However, the predominance of G mutations within AGC was significantly enhanced by selection. The preference for G mutations within AGC of productive rearrangements has been reported by several studies (24, 26, 37, 38). Based on the current data, this relates to a combination of events in which selection increases the frequency of mutations of G within AGC in the CDRs after the G nucleotides have been preferentially targeted by the mutator. This mutational preference provides a higher potential for a replacement mutation compared with mutations of C. Of note, third-codon position mutations with a 33% chance of being silent were found more frequently in the productive than in the nonproductive repertoire, especially in the FRs (39%). The apparent positive selection for third-position mutations of AGC in the FRs probably reflects a compensatory change after loss of codons with replacement mutations. Despite this distribution, mutations of AGC resulting in conservative changes were favored by selection in the CDRs, whereas negative selection against changes to basic amino acids was also evident in the CDRs.

When AGC occurred within RGYW motifs in the FRs, it was not only the target of the mutational machinery, but also appeared to provide the substrate for significant positive selection. The vast majority of AGC in CDRs occurred in RGYW motifs, whereas fewer AGC occurred in RGYW in FRs. However, when the latter were mutated, they were markedly increased in their frequency in the productive repertoire. In contrast, there was no selection of mutated AGCs that occurred outside of RGYW motifs in the FRs. AGC codons incorporated in RGYW motifs reside in the FRs at positions 28 (V_H5), 30 (V_H2 and V_H4), 76 (V_H1 and V_H5), and 82a (V_H1, V_H4, and V_H5). The potential importance of FR mutations, especially within FR3, in altering Ag binding by influencing the conformation of the CDRs has recently been documented (39). This may account for the apparent positive selection of AGC codons within RGYW motifs in the FRs. Although selection against replacements in the FRs and against RGYW mutations in the FRs was identified, the positive selection of AGC mutations within RGYW in the FRs suggests the beneficial role of mutations in these codons for the formation of an Ig molecule with enhanced Ag binding capability or, alternatively, the lack of a negative impact on the conformation of the Ig molecule of replacement mutations in AGC within RGYW of the FRs.

Imprint of selection on frequently mutated codons

The analysis of frequently mutated codons gave further insight into the relation of the influences of mutation and selection. Thus, there was a strong tendency for AGY, TAC, GTA, and GCT mutations to be selected in the CDRs for replacements encoding conservative amino acids. Moreover, selection against replacement mutations resulting in a higher frequency of silent mutations in the FRs was found for AGY, TAC, GTA, and GCT mutations. Of note, a change of AGC can result in AGR or CGN codons encoding arginine. An enrichment of this amino acid in the CDRs is known to increase certain autoantigen binding capabilities (DNA) (5, 15, 40, 41) at specific codon positions. In this regard, AGC mutations resulted in arginine codons in 12.9% of the nonproductive rearrangements compared with 9.4% in the productive repertoire. The current data, therefore, are consistent with the conclusion that there is negative selection against AGC mutations encoding basic amino acids in the normal repertoire, possibly to avoid autoreactivity. Moreover, high tyrosine, asparagine and glutamine content of the CDR regions of some anti-DNA Abs has also been reported (41). The importance of somatic hypermutation in the induction of high affinity anti-DNA Abs has also been reported (5, 42, 43). Since autoantibodies, such as anti-DNA, are potentially pathogenic and may be generated by hypermutation, it would be anticipated that they would be negatively selected in the normal repertoire to avoid autoreactivity. Thus, it appears reasonable that selection against replacements in the CDRs, especially those that encode for charged amino acids thought to enhance anti-DNA binding, may play a significant role in shaping the repertoire in normal individuals.

Hot and cold spots and selective influences

Most of the mutational hot spots appear to be intrinsically determined by the concentration of highly mutable sequences at these positions (3) and have been driven into the CDRs by evolutionary pressures (36). Lacking a defined Ag, no indication for a specific selective hot spot could be determined. The current data, however, suggest that there are important negative selective pressures against FR replacement mutations. Specific cold spots of mutation

that do not appear to tolerate changes in the productive repertoire became apparent. The current analysis identified positions 87, 91, and 82c as having significantly lower R:S ratios in the productive V_HDJ_H rearrangements, whereas position 18 was never found to be mutated in the 444 V_H genes analyzed to date.

Position 36 (TGG) was mutated in only one nonproductive V_H rearrangement (TGG→TGC, a cysteine), but not in the expressed B cell repertoire. However, we found one mutation at position 36 in the $CD19^+$ B cells in a previous analysis (3) (TGG→TGT, a cysteine), indicating that mutations of this amino acid can be tolerated in the expressed repertoire. Since TGG was found to be mutated at position 36 in one nonproductive gene rearrangement only, the mutational machinery appears to target this sequence less frequently. This is noteworthy, because TGG is the only codon that codes for tryptophan, and each of its mutations leads to a replacement amino acid. TGG at codon position 36 is highly conserved throughout all V_H genes encoding one of the residues of the inner hydrophobic pocket of the Ig molecule (22). TGG also occurs frequently as a codon at position 47 within FR2, but was found to be mutated in only three genes within the 232 productive rearrangements. Moreover, the TGG codon also occurs within CDR1 and two of particular germline segments. Base pair changes in TGG were found in the productive repertoire in one case at position 33, in five genes at position 50, in two genes at position 52a, and in one gene at position 55 in CDR1 and CDR2, respectively. This indicated that TGG codons were mutated and selected in the CDRs (TGG was mutated to CGG encoding arginine six of eight times, of which five occurred at position 50) at a higher frequency, whereas TGG replacement mutations in FR2 occurred less frequently in the productive repertoire.

Of note, mutations of TGG (Trp→Leu) at position 33 (CDR1) in one particular murine Ab (V_H186.2) have been identified as an important determinant of the affinity to the hapten 4-hydroxy-3-nitrophenylacetate (44–46). The current results and previous reports (3, 20) indicate that mutations of TGG at either position 33 or 36 are less common than expected, suggesting that this codon is, in general, avoided by the mutational machinery. The power of Ag-specific positive selection is evident, however, in that an infrequent mutation, such as position 33 Trp→Leu, can dominate an immune response if it results in increased affinity to the immunizing hapten (44–46), even though this mutation is not essential for binding to 4-hydroxy-3-nitrophenylacetate (47).

Of note, the current data emphasize that mutation and selection do not occur randomly. Both processes play a significant role at specific locations in shaping the expressed Ig repertoire. Primarily, the nucleotide sequences themselves and their locations influence the hypermutation process, whereas subsequent selection fine-tunes the repertoire, functioning, in general, to delete a large number of mutations from the productive repertoire. The net result is a mutated repertoire with less diversity than anticipated from the action of the mutational machinery.

Acknowledgments

We thank Dr. Qin-Chang Cheng for help in determining appropriate statistical analyses.

References

1. Radic, M. Z., and M. Zouali. 1996. Receptor editing, immune diversification, and self-tolerance. *Immunity 5:505.*
2. Kelsoe, G. 1996. Life and death in germinal centers (redux). *Immunity 4:107.*
3. Dörner, T., H.-P. Brezinschek, R. I. Brezinschek, S. J. Foster, R. Domiati-Saad, and P. E. Lipsky. 1997. Analysis of the frequency and pattern of somatic mutations within non-productively rearranged human V_H genes. *J. Immunol. 158: 2779.*
4. Reynaud, C. A., C. Garcia, W. R. Hein, and J. C. Weill. 1995. Hypermutation generating the sheep immunoglobulin repertoire is an antigen-independent process. *Cell 80:115.*
5. Ray, S. K., C. Putterman, and B. Diamond. 1996. Pathogenic autoantibodies are routinely generated during the response to foreign antigen: a paradigm for autoimmune disease. *Proc. Natl. Acad. Sci. USA 93:2019.*
6. Borretzen, M., C. Chapman, J. B. Natvig, and K. M. Thompson. 1997. Differences in the mutational patterns between rheumatoid factors in health and disease are related to variable heavy chain family and germline gene usage. *Eur. J. Immunol. 27:735.*
7. Isenberg, D. A., M. A. Rahman, C. T. Ravirajan, and J. K. Kalsi. 1997. Anti-DNA antibodies: from gene usage to crystal structures. *Immunol. Today 18:149.*
8. Putterman, C., W. Limpanasithikul, M. Edelman, and B. Diamond. 1996. The double edged sword of the immune response: mutational analysis of a murine anti-pneumococcal, anti-DNA antibody. *J. Clin. Invest. 97:2251.*
9. Brezinschek, H.-P., R. I. Brezinschek, and P. E. Lipsky. 1995. Analysis of the heavy chain repertoire of human peripheral B cells using single-cell polymerase chain reaction. *J. Immunol. 155:190.*
10. Tomlinson, I. M., S. C. Williams, S. J. Corbett, J. B. L. Cox, and G. Winter. 1996. *V BASE Sequence Directory.* Medical Research Council Center for Protein Engineering, Cambridge, U.K.
11. Brezinschek, H.-P., S. J. Foster, R. I. Brezinschek, T. Dörner, R. Domiati-Saad, and P. E. Lipsky. 1997. Analysis of the human VH gene repertoire: differential effects of selection and somatic hypermutation on human peripheral $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B cells. *J. Clin. Invest. 99:2488.*
12. Jacob, J., J. Przylepa, C. Miller, and G. Kelsoe. 1993. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. III. The kinetics of V region mutation and selection in germinal center B cells. *J. Exp. Med. 178: 1293.*
13. Betz, A. G., C. Milstein, A. Gonzalez-Fernandez, R. Pannell, T. Larsson, and M. S. Neuberger. 1994. Elements regulating somatic hypermutation of an immunglobulin κ gene: critical role for the intron enhancer/matrix attachment region. *Cell 77:239.*
14. Woolson, R. F. 1987. *Statistical Methods for the Analysis of Biomedical Data.* John Wiley and Sons, New York, pp. 337.
15. Shlomchik, M. J., A. H. Aucoin, D. S. Pisetsky, and M. G. Weigert. 1987. Structure and function of anti-DNA autoantibodies derived from a single autoimmune mouse. *Proc. Natl. Acad. Sci. USA 84:9150.*
16. Quin, Y., A. Greiner, M. J. F. Trunk, B. Schmausser, M. M. Ott, and H. K. Müller-Härmelink. 1995. Somatic hypermutation in low-grade mucosa-associated lymphoid tissue-type B-cell lymphoma. *Blood 86:3528.*
17. Chapman, C., J. X. Zhou, C. Gregory, A. B. Rickinson, and F. K. Stevenson. 1996. V_H and V_L gene analysis in sporadic Burkitt's lymphoma shows somatic hypermutation, intraclonal heterogeneity, and a role for antigen selection. *Blood 88:3562.*
18. Rogozin, I. B., and N. A. Kolchanov. 1992. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim. Biophys. Acta 1171:11.*
19. Ebeling, S. B., M. E. M. Schutte, and T. Logtenberg. 1993. Peripheral human $CD5^+$ and $CD5^-$ B cells may express somatically mutated V_H5- and V_H6-encoded IgM receptors. *J. Immunol. 151:6891.*
20. Snow, R. E., C. J. Chapman, A. J. Frew, S. T. Holgate, and F. K. Stevenson. 1997. Pattern of usage and somatic hypermutation in the V_H5 gene segments of a patient with asthma: implications for IgE. *Eur. J. Immunol. 27:162.*
21. van Es, J. H., F. H. J. Gmelig Meyling, and T. Logtenberg. 1992. High frequency of somatically mutated IgM molecules in the human adult blood B cell repertoire. *Eur. J. Immunol. 22:2761.*
22. Chothia, C., and A. M. Lesk. 1987. Canonical structures for the hypervariable regions on immunoglobulins. *J. Mol. Biol. 196:901.*
23. Gonzalez-Fernandez, A., S. K. Gupta, R. Pannell, M. S. Neuberger, and C. Milstein. 1994. Somatic mutation of immunoglobulin λ chains: a segment of the major intron hypermutates as much as the complementarity-determining regions. *Proc. Natl. Acad. Sci. USA 91:12614.*
24. Betz, A. G., M. S. Neuberger, and C. Milstein. 1993. Discriminating intrinsic and antigen-selected mutational hot spots in immunoglobulin V genes. *Immunol. Today 14:405.*
25. Yelamos, J., N. Klix, B. Goyenechea, F. Lozano, Y. L. Chui, A. Gonzalez-Fernandez, R. Panell, M. S. Neuberger, and C. Milstein. 1995. Targeting of non-Ig sequences in place of V segment by somatic hypermutation. *Nature 376:225.*
26. Neuberger, M. S., and C. Milstein. 1995. Somatic hypermutation. *Curr. Opin. Immunol. 7:248.*
27. Jolly, C. J., S. D. Wagner, C. Rada, N. Klix, C. Milstein, and M. S. Neuberger. 1996. The targeting of somatic hypermutation. *Semin. Immunol. 8:159.*
28. Maizels, N. 1995. Somatic hypermutation: how many mechanisms diversify V region sequences? *Cell 83:9.*
29. Pascual, V., Y. J. Liu, A. Magalski, O. de Boutellier, J. Banchereau, and J. D. Capra. 1994. Analysis of somatic mutation in five B cell subsets of human tonsil. *J. Exp. Med. 180:329.*
30. Tumas-Brundage, K., and T. Manser. 1997. The transcriptional promotor regulates hypermutation of the antibody heavy chain locus. *J. Exp. Med. 185:239.*
31. Peters, A., and U. Storb. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity 4:57.*
32. Chang, B., and P. Casali. 1994. The CDR1 sequences of a major proportion of human germline Ig V_H genes are inherently susceptible to amino acid replacement. *Immunol. Today 15:367.*

33. Lebecque, S. G., and P. J. Gearhart. 1990. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5′boundary is near the promoter, and 3′boundary is approximately 1kb from V(D)J gene. *J. Exp. Med. 172:1717.*
34. Insel, R. A., and W. S. Varade. 1994. Bias in somatic hypermutation of human V_H genes. *Int. Immunol. 6:1437.*
35. Brenner, S., and C. Milstein. 1966. Origin of antibody variation. *Nature 211:242.*
36. Wagner, S. D., C. Milstein, and M. S. Neuberger. 1995. Codon bias targets mutation. *Nature 376:732.*
37. Reynaud, C. A., L. Quint, B. Bertocci, and J. C. Weill. 1996. What mechanism(s) drive hypermutation? *Semin. Immunol. 8:125.*
38. Goyonechea, B., and C. Milstein. 1996. Modifying the sequence of an immunoglobulin V-gene alters the resulting pattern of hypermutation. *Proc. Natl. Acad. Sci. USA 93:13979.*
39. Wedemayer, G. J., P. A. Patten, L. H. Wang, P. G. Schultz, and R. C. Stevens. 1997. Structural insights into the evolution of an antibody combining site. *Science 276:1665.*
40. Eilat, D., D. M. Webster, and A. R. Rees. 1988. V region sequences of anti-DNA and anti-RNA autoantibodies from NZB/NZW F1 mice. *J. Immunol. 141:1745.*
41. Stewart, A. K., C. Huang, A. A. Long, B. D. Stollar, and R. S. Schwartz. 1992. V_H-gene representation in autoantibodies reflects the normal human B cell repertoire. *Immunol. Rev. 128:101.*
42. Diamond, B., and M. D. Scharff. 1984. Somatic mutation of the T15 heavy chain gives rise to an antibody with autoantibody specificity. *Proc. Natl Acad. Sci. USA 81:5811.*
43. Naparstek, Y., J. Andre-Schwartz, T. Manser, L. J. Wysocki, L. Breitman, M. Gefter, and R. S. Schwartz. 1986. A single germline V_H gene segment of normal A/J mice encodes autoantibodies characteristic of systemic lupus erythematosus. *J. Exp. Med. 164:614.*
44. Rajewski, K., I. Förster, and A. Cumano. 1987. Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science 238:1088.*
45. Matsumoto, M., S. F. Lo, C. J. L. Carruthers, J. Min, S. Mariathasan, G. Huang, D. R. Plas, S. M. Martin, R. S. Geha, M. H. Nahm, and D. D. Chaplin. 1996. Affinity maturation without germinal centres in lymphotoxin-α-deficient mice. *Nature 382:462.*
46. Allen, D., A. Cumano, R. Dildrop, C. Kocks, K. Rajewski, N. Rajewski, J. Roes, F. Sablitzky, and M. Siekevietz. 1987. Timing, genetic requirements and functional consequences of somatic hypermutation during B-cell development. *Immunol. Rev. 96:5.*
47. Berek, C., and M. Ziegner. 1993. The maturation of the immune response. *Immunol. Today 14:400.*