

Alternative mechanisms of receptor editing in autoreactive B cells

Olga Kalinina^a, Colleen M. Doyle-Cooper^a, Jennifer Miksanek^a, Wenzhao Meng^b, Eline Luning Prak^b, and Martin G. Weigert^{a,1}

^aDepartment of Pathology, Gwen Knapp Center for Lupus and Immunology Research, University of Chicago, Chicago, IL 60637; and ^bDepartment of Pathology and Laboratory Medicine, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

Contributed by Martin G. Weigert, March 2, 2011 (sent for review December 22, 2010)

Pathogenic anti-DNA antibodies expressed in systemic lupus erythematosus bind DNA mainly through electrostatic interactions between the positively charged Arg residues of the antibody complementarity determining region (CDR) and the negatively charged phosphate groups of DNA. The importance of Arg in CDR3 for DNA binding has been shown in mice with transgenes coding for anti-DNA V_H regions; there is also a close correlation between arginines in CDR3 of antibodies and DNA binding. Codons for Arg can readily be formed by V(D)J rearrangement; thereby, antibodies that bind DNA are part of the preimmune repertoire. Anti-DNAs in healthy mice are regulated by receptor editing, a mechanism that replaces κ light (L) chains compatible with DNA binding with κ L chains that harbor aspartic residues. This negatively charged amino acid is thought to neutralize Arg sites in the V_H. Editing by replacement is allowed at the κ locus, because the rearranged VJ is nested between unrearranged Vs and Js. However, neither λ nor heavy (H) chain loci are organized so as to allow such second rearrangements. In this study, we analyze regulation of anti-DNA H chains in mice that lack the κ locus, κ-/κ- mice. These mice show that the endogenous preimmune repertoire does indeed include a high frequency of antibodies with Arg in their CDR3s (putative anti-DNAs) and they are associated mainly with the editor L chain λ_X. The editing mechanisms in the case of λ-expressing B cells include L chain allelic inclusion and V_H replacement.

autoimmunity | tolerance

Anti-DNA antibodies bind DNA mainly through electrostatic interactions between the positively charged Arg (R) residues of the antibody complementarity determining regions (CDRs) and the negatively charged DNA phosphodiester backbone (1). Arg enrichment has been demonstrated in the protein sequences of anti-DNA antibody heavy (H) chains (2, 3). In addition, reverse mutagenesis of an H chain Arg to a germline amino acid weakened DNA binding, whereas forward mutagenesis with additional Arg residues enhanced DNA binding (1). The predominant occurrence of Arg residues in the V_H-encoded CDRs of anti-DNA antibodies indicates a dominant role for V_H in DNA binding, and the DNA specificity of these H chains persists when paired with a wide variety of light (L) chains (4–6). There are exceptional L chains (4 Vκ L chains of 95 functional Vκ genes and 1 Vλ L chain of 4 functional Vλ L chains) of the mouse that modify or veto the DNA-binding quality of Arg-enriched H chains (4, 7, 8). These editor L chains are characterized by a high content of aspartate, a negatively charged amino acid that efficiently neutralizes the positively charged Arg residues of anti-DNA H chains (4). B cells that express an anti-DNA V_H with a noneditor L chain can undergo further L chain rearrangement. If the original L chain is replaced by an editor L chain, the B cell will be tolerized (4, 7). Similarly, MHC class I-reactive B cells (found in 3–83μ transgenic mice) that encounter autoantigen in the bone marrow continue rearrangement and change receptor specificity by replacing the transgenic 3–83 κ L chain with an endogenous κ or λ L chain (9, 10).

Here we wished to evaluate the extent to which editor L chains rescue B cells with endogenous H chains. Given the clear evidence for H chains influencing the expressed L chain repertoire, we now examine whether and to what degree L chains can also

influence the H chain repertoire. To do so, we studied mice with a normal germline H chain locus and a simplified L chain locus resulting from targeted deletion of the κ locus (11). These mice have a simple V_L repertoire with only four possible λ L chains: three noneditors, λ1, λ2, and λ3 (encoded by three possible combinations of Vλ1 or Vλ2 with available JCL segments, referred to collectively as λ1,2,3), are nearly identical and one editor, λX (VλX recombined to JCL2; Fig. S1), which is only 50% identical to the λ1,2,3 class (12). λX is one of the most acidic L chains (pI 4.62) and λ1,2,3 are among the most basic (pI 10.9) (13). λX efficiently vetoes DNA binding by the interaction of Asp with the basic amino acids of anti-DNA V_Hs. We found evidence for L chain influence on the V_H repertoire by comparing V_Hs expressed in B cells that express the editor λX chain to those that express the noneditor λ1,2,3 chains. We have analyzed the repertoire of λ-associated antibodies and found that the editing mechanisms in the case of λ-expressing B cells include L chain allelic inclusion and V_H replacement.

Results

Editing Shapes the V_H Repertoire. In DNA-binding antibody H-chain transgenic models such as 56R, editing yields a population limited to L chains that modify the specificity of an anti-DNA V_H (4). These edited antibodies either avoid binding self-antigen or bind DNA with low affinity [3H9/Vκ8 (14)]. Here, we studied the converse, namely the influence of the L chain on the endogenous V_H repertoire. To test whether B cells expressing an editor L chain have a distinct repertoire of H chains, we compared the V_H repertoire of B cells that expressed editors to those that express noneditor L chains. To quickly assess whether we would see a detectable difference in λ use that would reveal the kinds of repertoire effects of L chains that we planned to assess at a molecular level, we examined editor vs. noneditor λ L chain use in κ-/κ- mice compared with 56R κ-/κ- mice by flow cytometry. A clear alteration in the B-cell repertoire is evident based on a shift in the ratio of λX:λ1 from 9:1 in 56R κ-/κ- to 1:9 for λX:λ1,2,3 in κ-/κ- mice (Fig. 1). In this example, the 56R H chain is mainly associated with the editor L chain λX (13).

The λ chains expressed in splenic B cells of the κ-/κ- mouse were unevenly distributed between λX (~10%) and λ1,2,3 (~84%) (Fig. 1). In addition, there was a small fraction of cells (2%) in the quadrant expected for B cells that coexpress λ1,2,3 and λX. We next used monoclonal antibodies specific for each Vλ to sort single Vλ1,2,3- and VλX-expressing B cells and study their V_H repertoires in single B cells. We predicted that the λX population would be enriched in V_Hs that bind DNA.

Author contributions: O.K., C.M.D.-C., and M.G.W. designed research; O.K., C.M.D.-C., J.M., and W.M. performed research; O.K., C.M.D.-C., J.M., W.M., E.L.P., and M.G.W. analyzed data; and O.K., C.M.D.-C., and M.G.W. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the NCBI GenBank database (accession nos. JF690280–JF690665).

¹To whom correspondence should be addressed. E-mail: mweigert@bsd.uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1019389108/-DCSupplemental.

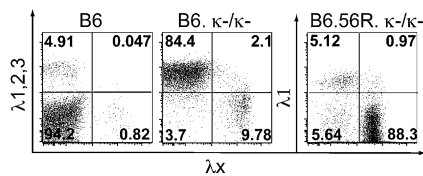


Fig. 1. L chain isotype use in B6 wild-type (Left), B6 κ -/ κ - (Center), and 56R B6 κ -/ κ - (Right) mice. Splenocytes were stained with PE/Cy7 anti-B220, Alexa-647, anti- λ X specific monoclonal antibodies (kindly provided by P.A. Cazeneuve (57)) and FITC-anti- λ 1,2,3-monoclonal antibodies for B6 wild-type and B6 κ -/ κ - and PE anti- λ 1 antibody (L22.18.2, kindly provided by U. Storb) for 56R B6 κ -/ κ - plots. The dot plots display λ 1,2,3 for Left and Central panels and λ 1 only for the Right panel (y axis) vs. λ X expression (x axis) on cells in the live, lymphoid, B220+ gate.

V_H Repertoires in λ X and λ 1,2,3 B Cells. To compare the V_H repertoires of λ X and λ 1,2,3 B cells, single-cell reverse transcription-PCR was followed by sequencing to identify the H and L chain genes in each of 303 FACS-sorted single cells (Fig. 2 and Fig. S2). The relative frequency of each V_H family was calculated as a percentage of the total number of V_H sequences within each population of cells as defined by L chain expression. The V_H repertoires were similar to each other with respect to gene family use (Fig. 2) and resembled the previously described repertoire of wild-type mice (15). The V_H families V_H J558, V_H 7183, V_H Q52, V_H 9, V_H 10, V_H 36-60, and V_H 3609 were represented in equal proportions in both populations. Further, there were several examples of the same V_H gene being rearranged in B cells expressing either the editor λ X or the noneditor L chain (λ 1).

There were exceptions: two V_H families (J606 and S107) were observed only among λ X in this survey (Fig. 2). The frequency of J606 use is significantly different in λ X vs. λ 1,2,3 B cells ($P \sim 0.0002$, see *SI Results*). Furthermore, among the 20 unique J606 sequences, 16 used the V_H J606.1 gene segment, and the remainder used V_H J606.4. Restricted V_H/V_L associations have been explained by positive selection upon the specificity (presumably anti-self) formed by the H + L chain pair (16-18), but preferential association of particular H chains with λ X could also be explained by negative selection. For example, if the combinations of V_H from J606 with λ 1,2,3 L chains were autoreactive, then B cells with these combinations would be deleted or edited.

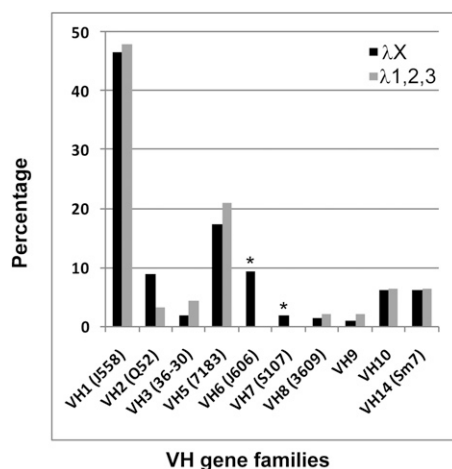


Fig. 2. V_H gene family use in the splenic B cells expressing editor (λ X, dark bars) and noneditor (λ 1,2,3, light bars) L chains from B6 κ -/ κ - mice as determined by single cell reverse transcription-PCR. V_H use is depicted as the relative frequency of each V_H gene family among all analyzed V_H sequences within each sorted population. $N = 215$ (100%) for λ X and $N = 88$ (100%) for λ 1,2,3. V_H genes that were found only in λ X antibodies are indicated with asterisk (*).

To further evaluate the influence of the L chain on the repertoire of V_H J606.1, we performed CDR3 spectratyping on V_H J606.1- J_H 2 rearrangements in the genomic DNA of five different subsets of B cells: λ X and λ 1,2,3 from B6 κ -/ κ -, λ and total splenocytes ($\sim 95\%$ κ +) from B6, and total splenocytes from B6.V κ 8 $^{+/-}$. By analyzing genomic DNA, we are able to measure the fraction of rearrangements that are in-frame, which is correlated with selection of individual V_H S (19). Thus, if the manner of V_H J606.1 selection is distinctive with different L chains, then we expect to find differences in the fraction of rearrangements that are in-frame. As shown in Fig. S3, this is indeed the case. The fraction of in-frame V_H J606.1- J_H 2 rearrangements in λ X B cells from κ -/ κ - mice was 100% (23/23 rearrangements), whereas the corresponding fraction in λ 1,2,3 B cells from κ -/ κ - mice was 73% ($P < 0.01$ by Fisher exact test exact test). The fraction of in-frame rearrangements is intermediate in B-cell populations from λ or total splenocytes from B6, and total splenocytes from B6.V κ 8 $^{+/-}$ mice.

Another example of selection for a particular V_H gene is V_H SM7.4. Although the SM7 family was represented in both the λ X and λ 1,2,3 populations, V_H gene use and CDR3 sequences were very distinct. Nine of 14 independent SM7 sequences from λ X-expressing B cells expressed the same V_H family member, SM7.4, which was not found in the sampling of λ 1,2,3 antibodies in our survey. Moreover, seven of nine SM7.4 sequences had similar CDR3s with respect to length and polarity, both of which are ordinarily highly variable properties of CDR3s (Fig. 3). CDR3 similarity could be dictated by the SM7.4 V_H sequence. The SM7.4 V_H encodes negatively charged residues Asp and Glu at CDR1 and CDR2. A strongly negatively charged CDR2 is likely to bind positively charged antigens. Indeed, the SM7.4 V_H has been shown to bind Sm (20). Short and Arg-rich CDR3s could be the result of editing in response to self-reactivity. Alternatively, a combination of V_H SM7.4 with a short, positively charged CDR3 and λ X could be positively selected by some unknown (self) antigen.

CDR3 Sequences of λ X and λ 1,2,3 Antibodies. The correlation of antibody specificity and selection is apparent from the difference in the frequency of Arg residues in the CDR3s of V_H S expressed with editor or noneditor L chain (Table 1 and Fig. S3A). There are 119 Args of 2058 V_H CDR3 residues in λ X antibodies (6% of the total amino acid content of CDR3), significantly more than the 16 Args of 875 V_H CDR3 residues (2%) in λ 1,2,3 antibodies ($P < 0.00001$; χ^2 test). The pool of known V_H CDR3 sequences allows us to compare the frequency of Arg residues in the V_H CDR3 regions of λ X- and λ 1,2,3-encoded antibodies to V_H S expressed with a variety of L chains (Table 1). There are 3,003 Args in 74,923 CDR3 residues in the Abys database (<http://www.bioinf.org.uk/abs/>), which represents an Arg content of 4% in the known mouse H chain repertoire [it was 105 Args of 3,066 residues (3.4%) in the Kabat et al. (2) compilation 20 y ago when we originally made this observation]. Thus, λ X antibodies have a significantly higher level ($P = 0.0001$; χ^2 test) of V_H CDR3 arginines compared with sequences from the total recorded mouse antibody sequences.

Mechanisms That Generate Arg Codons in CDR3 Regions. Several mechanisms can contribute to the generation and diversity of CDR3 regions, including alternative D_H reading frames (RF), D-D fusions, imprecise joining at the DJ and VD junctions, and N and P addition. We examined the contributions of these mechanisms to the formation of Arg codons in the CDR3 regions of H chains paired with λ X or λ 1,2,3 L chains in our sorted cells.

The reading frame of D_H genes can vary. Further, D_H genes can be rearranged in either transcriptional orientation, which provides three additional reading frames. However, CDR3 sequences are encoded by a preferred D_H reading frame and D_H genes are usually rearranged to J_H at their 3' recombination signal sequence (RSS) (21, 22). Here, we found that the majority of D_H genes were read in RF1: 74% and 66% for λ 1,2,3 and λ X B cells, respectively. These frequencies are similar to those reported for normal new-

λX				$\lambda 1,2,3$			$\lambda 1,2,3 + \lambda X$				
SM7.1.44	CTS	IYYGNSYFDY	WGQ	SM7.3.54	CAR	CGSSLWYFDV	WGT	SM7.2.49	CAQ	LTLFAY	WGQ
SM7.1.44	CTA	LLYYFDY	WGQ	SM7.3.54	CAR	SFYYYGSSPFAY	WGQ	SM7.3.54	CAR	GYGGRANY	WGQ R
SM7.2.49	CAR	QLRLLDY	WGQ R	SM7.3.54	CAR	GYPYAMDY	WGQ	SM7.4.63	CTR	LRY	WGQ R
SM7.2.49	CAR	PRYYGSSYYYAMDY	WGQ R	SM7.2.49	CAR	WGDY	WGQ	SM7.4.63	CTT	SYLRYYFDY	WGQ R
SM7.4.63	CTS	LRP	WGQ R	SM7.2.49	CAR	GFYYFDY	WGQ				
SM7.4.63	CTT	GYGSTHYFDY	WGQ	SM7.2.49	CAR	EDGNYEVDY	WGQ				
SM7.4.63	CTT	IRY	WGQ R								
SM7.4.63	CTT	IYH	WGQ								
SM7.4.63	CTT	LRH	WGQ R								
SM7.4.63	CTT	LRY	WGQ R								
SM7.4.63	CTT	RHR	WGQ RR								
SM7.4.63	CTT	RYH	WGQ R								
SM7.4.63	CTT	PYYGVS	WGQ								

Fig. 3. V_H family Sm7 gene use and V_H CDR3 amino acid sequences of H chains in B cells sorted on the basis of L chain expression and expressed with editor ($V\lambda X$), noneditor ($V\lambda 1,2,3$), or both ($V\lambda 1,2,3/V\lambda X$) L chains from B6 κ/κ - mice. Args in each CDR3 are highlighted in red.

born and adult mice (21, 22). Reading frame usage could account for the CDR3 Arg bias, because Arg codons are overrepresented in reading frames usage other than RF1. In fact, when their D_H gene segments are read in RF2 or RF3, 43% and more than 50% of H chains have Args in the CDR3 region in $\lambda 1,2,3$ and λX B cells, respectively, whereas in RF1 this number drops to 21% (in $\lambda 1$ cells) and 31% (in λX cells). Restricted reading frame use has been attributed to rearrangement requirements such as the translational competence of D_HJ_H rearranged genes ($D\mu$) read in RF2 and the abundance of stop codons in RF3 (23). Our data favor an alternative explanation that CDR3 sequences predominantly use the first RF, because RF1 disfavors Arg codons so that more B cells using this RF survive the selection process (24). Selection against Arg in CDR3 may result from the association between CDR3 Arg and DNA binding.

D-D fusion can generate arginines in V_H CDR3 regions. D-D fusions can contribute more reading frame choices, thereby increasing the likelihood of Arg codons. Fusions also provide additional opportunities for N and P addition at the junction of the fusion. Direct fusions, those that lack N or P addition, can form Arg codons made up of 5' and 3' nucleotides from two Ds. To identify D gene fusions, we first searched for multiple D segments in each CDR3 from our data set. We found that 8–10% of the H chains in our survey contained a D-D fusion (Fig. S4) in both λX - and $\lambda 1,2,3$ -expressing B cells. Ten of 21 H chain sequences with evidence of D-D fusion also exhibited CDR3-encoded arginine residues (Fig. S4).

$\lambda 1,2,3/\lambda X$ Double-Expression Occurs with V_H Genes That Exhibit Anti-DNA Features. Receptor editing at the κ locus is replacement of a $V\kappa J\kappa$ that contributes to autoreactivity by leapfrogging rearrangement of a nonautoreactive $V\kappa J\kappa$ or by inactivation via rearrangement of the recombination signal (RS) to the κ -locus intronic recombination signal 1 (iRS1) or RS sequences 3' of $J\kappa$ s or $V\kappa$, respectively (Fig. S1) (25, 26). The structure of the murine

λ locus, however, does not support leapfrogging or RS deletion. Therefore, editing of autoreactive B cells in the κ/κ - mouse or among λ -expressing B cells in B6 mice may require alternative mechanisms. Ongoing rearrangement stimulated by autoreactivity (9, 27) allows a B cell to express another L chain from the λ locus, but this new rearrangement does not inactivate previous autoreactive rearrangements. As such, the consequence of editing in κ/κ - mice would likely be isotypic and/or allelic inclusion, leading to the expression of two different antibodies, one of which is the original autoreactive antibody. These κ/λ or κ/κ "allelically included" B cells are thought to be inactive (28) or sequestered (29). This form of editing may be represented by the population of isotypically included $\lambda 1/\lambda X$ B cells in 3H9 and 56R κ/κ - mice (Fig. 1). In fact, the 3H9 κ/κ - and 56R κ/κ - mice have twice as many $\lambda 1,2,3/\lambda X$ double-expressing B cells as nontransgenic (nontg) κ/κ - mice (Table 2 and Fig. 1).

To obtain direct evidence that isotypically included ($\lambda 1,2,3/\lambda X$) B cells are edited cells, we sorted $\lambda 1,2,3/\lambda X$ double-positive B cells and analyzed the H chain sequences from 57 single B cells. As was shown above, H chains of B cells expressing editor (λX) and noneditor ($\lambda 1,2,3$) L chains are different with regard to Arg content in CDR3. We reasoned that if isotypically included ($\lambda 1,2,3/\lambda X$) B cells are edited, then the H chain CDR3s in these cells should resemble CDR3s of H chains typically paired with the editor λX . Alternatively, if the combination of anti-DNA H chain and $\lambda 1,2,3$ does not represent an edited population, then the H chain CDR3s of isotypically included B cells should more closely resemble the CDR3s of H chain from $\lambda 1,2,3$ B cells.

Our data indicate that isotypically included B cells express broad V_H gene family repertoire similar to that of $\lambda 1,2,3$ - or λX -only antibodies as well as to antibodies from wild-type mice (15). However, V_{H8} from J606 and S107 families, which were found only in λX B cells in our survey, were also found in B cells expressing two L chains ($\lambda 1,2,3/\lambda X$). The Arg content in CDR3 sequences of allelically included B cells ($\lambda 1,2,3/\lambda X$) was significantly higher than that in $\lambda 1,2,3$ B cells (4.5 vs. 1.8%, respectively; $P = 0.01$; χ^2 test), indicating that included B cells are likely to be edited cells.

Editing at V_H . Another potential substrate for editing in the κ/κ - mice is the H chain locus. Complete $V(D)J$ rearrangements are not nested between "functional" (12/23) RS sequences as in the κ locus, because all D_H segments are deleted during primary $V(D)J$ rearrangement. However, the V_H gene segment of most H chains can be replaced (edited) by an upstream V_H (30–35). In these cases, an upstream V_H can be rearranged to a cryptic recombination signal sequence (cRSS, or embedded heptamer), located near the 3' end of the coding region of the originally rearranged V_H gene (Fig. 4A). The original V_H region is thus replaced by a new V_H gene except for a short stretch of nucleotides (footprint), located downstream of the embedded heptamer. The recipient CDR3 and J_H are preserved, but additional nucleotides are added to the 5' end of the CDR3 from the 3' end of donor V_H gene, the footprint left over from 3' end

Table 1. Arg content in V_H CDR3 regions of λX - and $\lambda 1$ -encoded Abs and Abs from total recorded mouse B cell repertoire

	Total CDR3 amino acids	Args	% of Args
Abyss database	74,923	3,003	4
B cells expressing λX	2,058	119	5.8
B cells expressing $\lambda 1,2,3$	875	16	1.8

Numbers of total amino acids and arginine residues in V_H CDR3s were calculated through analysis of 215 and 88 V_H sequences from λX - and $\lambda 1,2,3$ -expressing B cells, respectively (CDR3 sequences are provided in Fig. S2). Numbers for total recorded mouse antibodies were obtained from Abyss database (<http://www.bioinf.org.uk/abs/>). The CDR3 regions were identified according to the convention of Kabat et al. (59), where CDR3 includes amino acids at positions H95–H102. The percentage of Arg in V_H CDR3s was calculated as the ratio of the number of Arg residues in CDR3s from all V_H sequences of each group to the total number of amino acids in CDR3s of all V_H sequences from each group multiplied by 100.

Table 2. The frequencies of different λ rearrangements in splenic B cells from κ -sufficient and κ -/ κ - mice and in LPS hybridomas from different lines of κ -/ κ - mice

λ genotype	Hybridomas			Flow cytometry data	
	Nontg. κ -/ κ -, N (%)	3H9. κ -/ κ -, N (%)	56R. κ -/ κ -, N (%)	B6 (%)	B6. κ -/ κ - (%)
λ 1,2,3	115 (92)	23 (22)	26 (29)	(4.9)	(84)
λ x	5 (4)	40 (38)	35 (39)	(0.8)	(10)
λ 1,2,3/ λ x double	5 (4)	42 (40)	29 (32)	(0.05)	(2)
Total	125	105	90	(5.75)	(96)

The frequencies of different λ rearrangements were evaluated as a percentage of the live B220+ lymphocytes by flow cytometric analysis of splenocytes from B6. κ -/ κ - and B6 mice. Shown for comparison are previously published data from hybridoma panels from 3H9. κ -/ κ -, 56R. κ -/ κ -, and nontg κ -/ κ - mice (13; data from ref. 60, copyright 1994, Rockefeller University Press). κ -/ κ - refers to homozygous κ -deficiency. Numbers of hybridomas are given; all percentage data are given in parentheses.

of recipient V_H gene and P-nucleotides and N addition (36). In addition, nucleotides of the footprint and the 5' end of CDR3 can be deleted, because the cleavage at the heptamer leaves this sequence unprotected and subject to exonuclease digestion [referred to as chopping (37)]. Here, we defined footprints as five or more consecutive nucleotides of identity with a different V_H from the upstream V_H as evidence of V_H replacement.

Among the 240 analyzed V_H sequences, 18 (7.5%) had identifiable V_H replacement footprints (Table 3 and Fig. 4B). Fifteen of the 18 (83%) V_H replacement products replaced a non-functional original V(D)J rearrangement, whereas 17% of the V_H replacement products replaced a functional original V(D)J. Assuming the latter are the result of editing, we concluded there was evidence of editing by V_H replacement in 3 of 240 (1%) examined H chains. Considering that V_H replacement footprints can be shortened or eliminated during V_H replacement, this may be an underestimation. It was shown in knockin mice homozygous for a nonproductive IgH rearrangement that, of all of the functional H chains generated through V_H replacement, only 20% of the functional replacement products contained a V_H replacement footprint (37). Moreover, V_H replacement can lead to both productive and nonproductive rearrangements; therefore, the rate of V_H replacement may be substantially higher than the number of identifiable V_H replacement products. To test if V_H replacement is the unique property of κ -/ κ - mice, we analyzed the frequency of the potential V_H replacement products in IgH sequences from the wild-type B6 κ -sufficient mouse and found it to be ~6% (9/157; Fig. S5).

Short footprint lengths combined with the fact that there are only four nucleotides makes it difficult to determine whether a footprint is actually a remnant of the recipient V_H or the result of N or P nucleotide addition. To verify that these sequences represent bona fide footprints, we compared the number of footprints in the V-D junction (N1) and in the D-J junction (N2), as was previously described by Zhang et al. (36). The rationale for this comparison is that nucleotide sequences in both V-D and D-J are created by the same mechanisms, namely N and P addition, and if these mechanisms can create sequences that look like footprints, such sequences should be created in both locations. We found that 18 of 240 H chain sequences (7.5%) contain V_H replacement footprints at the V-D junction and 4 of 240 (2%) have footprint-like sequences in the D-J region ($P = 0.002$) (Table 3). Similar percentages have been shown by Zhang and Collins (36, 38); however, this reasoning may not apply to every V_H gene (39).

Another way to test whether these sequences represent true footprints is to determine whether the donor V_H is located upstream of the putative recipient V_H . All 18 possible products of V_H replacement passed this test. For two sequences with footprints we were not able to find the possible recipient; these sequences were not included in the pool of 18 possible V_H replacement products. Therefore, we conclude that most footprints at V-D junctions are contributed by V_H replacement.

Discussion

Receptor editing is a process of ongoing rearrangement following an initial rearrangement event in developing and/or auto-

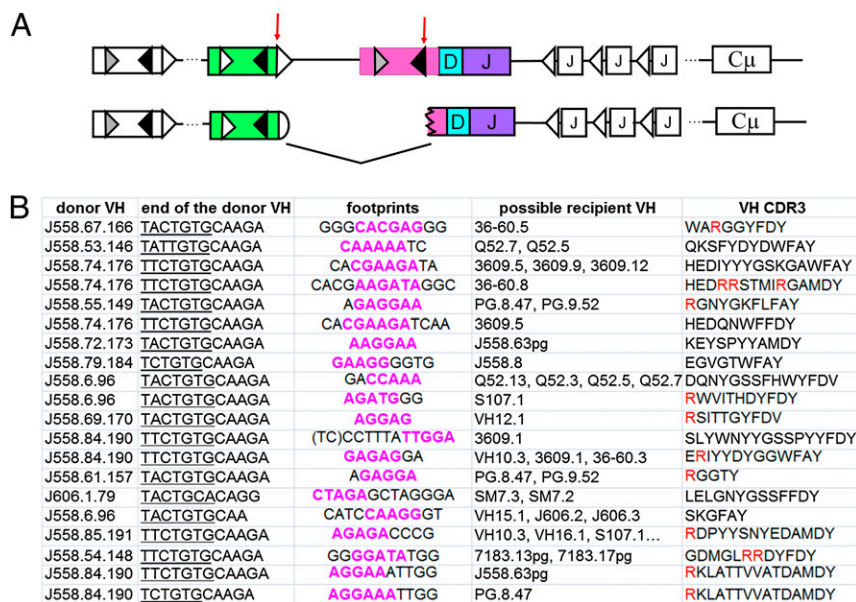


Fig. 4. V_H replacement footprints. (A) Schematic representation of V_H replacement. White triangles represent RSS sequences. Dark triangles represent the cRSS or embedded heptamer. The green region represents a rearranging V_H (donor) that replaced the initially rearranged V_H (recipient, magenta) and joined to the short stretch of nucleotides (footprint, magenta) that remained after the recipient V_H was cleaved at cRSS. (B) Potential V_H replacement products were found among H chain sequences from splenic B cells from κ -/ κ - mice. Donor V_H refers to the V_H gene segment fully represented in the complete V(D)J sequence. The nucleotide sequences shown are the end of the V_H donor sequence and the N1 nucleotides located between the last nucleotide of framework 3 and the first D nucleotide. Potential V_H replacement footprints are defined as 5 or more consecutive nucleotides of identity with the 3' end sequence of a nondonor germline V_H and are highlighted in magenta. Possible recipients are germline V_H s whose 3' end sequence contains the footprint sequence found in the V_H replacement product. The amino acids shown are from the V_H CDR3. P nucleotides are in parentheses.

Table 3. Number of potential V_H replacement footprints in V-D and footprint-like sequences in D-J regions

	Total B cells	λX	$\lambda 1$
No. of sequences analyzed	240	178	62
No. of sequences with footprints in V-D (N1)	18 (7.5%)*	14 (8%)	4 (6%)
Footprint-like sequences in D-J (N2)	4 (2%)	3 (2%)	1 (2%)
Replacements of IF V _H s	3 (17%)	2 (14%)	1 (25%)

The number of potential V_H replacement footprints is shown for total analyzed H chain sequences (only V_Hs with identified D_H regions were analyzed) and separately for λX and $\lambda 1,2,3$ antibodies. The number of footprints in the N1 region was significantly higher than the number of footprint-like sequences in the N2 region (* $P = 0.001$, χ^2 test). The number of V_H replacement products that are the result of replacing of in-frame (IF) V(D)J rearrangements is indicated for all groups.

reactive B cells. Studies of receptor editing in anti-DNA transgenic mice have shown that anti-DNA H chains are absent from the mature B-cell repertoire unless paired with an editor L chain (4). Does this observation apply to endogenous anti-DNA B-cell receptors that arise in a nontransgenic system, and if so, then how frequently do anti-DNA cells occur in a normal repertoire? To test this, we separated B cells into fractions that express either editor or noneditor L chains and compared the V_H repertoire in the two populations.

We used mice without κ L chains [κ -/ κ - (11)], which have a simple V_L repertoire consisting of one editor (λX) and three noneditor ($\lambda 1,2,3$) L chains. Assuming that B cells are formed with a wide diversity of V_H rearrangements at the pre-B-cell stage and go on to have an equal probability of association with one of the four V_L/C_L combinations as they progress through development, the V_H repertoire should be randomly distributed between $\lambda 1,2,3$ and λX . However, we found that V_H sequences differed between the two populations: the B-cell population expressing the editor L chain (λX) contained a greater proportion of cells expressing antibodies with Arg residues in the V_H CDR3 than did the population of noneditor-expressing ($\lambda 1,2,3$) B cells (40% vs. 12%, respectively). Because V_Hs that are enriched in Arg residues are likely to bind DNA, these results suggest that the endogenous V_H repertoire includes a high frequency of B cells with the potential to bind DNA. Although there is a strong correlation between the presence of Arg residues and DNA binding, we do not know whether it is a strict correlation. Ongoing experiments are testing this question. Here, we are mainly concerned with Arg frequencies and how they differ in the λX and $\lambda 1,2,3$ populations. The presence of antibodies enriched in Arg content in the λX compartment shows that B cells expressing endogenous V_H repertoires require receptor editing and that editing is not restricted to (or an artifact of) B cells expressing transgenic anti-DNA V_Hs. A similar conclusion was reached in studies of κ -macroself transgenic mice (40).

Our analysis shows that the frequency of V_Hs with Arg in CDR3 in $\lambda 1,2,3$ -expressing B cells is lower than in λX -expressing B cells. How can we account for the lower frequency of V_Hs with Arg in CDR3 in the $\lambda 1,2,3$ population? If Arg-enriched V_H + $\lambda 1,2,3$ combinations are autoreactive, as in the case of 3H9/ $\lambda 1$, cells bearing these antibodies may have been edited. However, this conclusion bears the caveat that the editing mechanism(s) used by the κ locus is not available to κ -/ κ - mice. The κ -/ κ - B cells cannot edit autoreactive receptors unless they use alternative (non- κ) mechanisms of editing.

An Alternative Mechanism of Editing: Allelic Inclusion. The only known way of editing the L chain locus in a κ -/ κ - mouse is by secondary λ rearrangement, thereby forcing a B cell to allelic and/or isotypic inclusion, exemplified by $\lambda 1,2,3/\lambda X$ -double-expressing B cells. Lymphocytes with two L chains have been described in

transgenic and normal mice (28, 29, 41–43). Among anti-DNA B cells from Ig transgenic mice, allelically included B cells usually coexpress an editor L chain, implying that these B cells must be the result of editing. In the κ -/ κ - mouse, we identified and characterized a population of B cells that coexpress $\lambda 1,2,3$ and λX . In this population, as in the λX -only population, we found a high frequency of V_Hs that have Arg in CDR3 (32%). Thus, the normal B-cell repertoire contains cells edited by isotypic inclusion.

Another Alternative Mechanism of Editing: V_H Replacement. Secondary rearrangement can also replace rearranged H chain genes. Originally discovered in B-cell lines (30, 32, 44), V_H replacement was subsequently demonstrated in Ig knockin mice (34, 45). Based on the footprints of replacement, V_H genes that have been replaced have been estimated to comprise 5–12% of the human H chain repertoire (36). Here, we found V_H replacement footprints in 7.5% of all analyzed H chain sequences in κ -/ κ - mice and a similar frequency in V_H rearrangements cloned from B6 spleen DNA. V_H replacement may play a physiological role in B-cell repertoire diversification based upon the conservation of the embedded RSS sequence in most murine and human V_H genes (34). Recently, it was reported that cattle V_H segments have conserved nucleotide sequences (CSNS) comprising 13–18 nt located at the V_H-D_H junction (46). These sequences belong neither to the rearranged V_H nor to any known cattle D_H genes. They are too long and are not GC-rich enough to be N nucleotides, nor are they palindromic. They could be V_H replacement footprints. Furthermore, cattle CSNS are rich in A nucleotides (46), a characteristic of footprints in murine and human H chains.

Despite the contradictory evidence regarding the frequency of V_H replacement as an editing mechanism, V_H replacement is not very efficient (37, 47–50). It adds charged amino acids (most often Arg) to the V_H CDR3 and, like conventional rearrangement, can generate nonproductive rearrangements. However, deletion of anti-DNA H chains by nonproductive V_H replacement represents a form of editing analogous to κ deletion and may allow rearrangement to occur at the other H chain locus (51).

Methods

Mouse Breeding. Nonautoimmune homozygous κ knockout C57BL/6 (B6 κ -/ κ -) mice (11) were purchased from GenPharm International. 56R site-directed transgenic mice (4) were crossed to homozygous κ -/ κ - mice to produce 56R/ κ -/ κ - mice. Tail DNA samples of 56R, V κ 8, and κ -/ κ - mice were genotyped as described previously (4, 52, 53). C57BL/6J (B6) mice were purchased from The Jackson Laboratory. Presence of anti-DNA and antinuclear antibodies was analyzed by ELISA (54) and ANA (BION) using sera from B6 κ -/ κ - and B6 mice ($n = 5$ mice per group). Mice used for these experiments were 6–12 wk old. All animal procedures were conducted in accordance with the guidelines and regulations of the University of Chicago IACUC.

Flow Cytometry. Splenocytes were isolated from κ -/ κ - ($n = 5$), 56R κ -/ κ - ($n = 5$), and wild-type ($n = 5$) mice, all on the B6 background. B cells expressing $\lambda 1$, $\lambda 2$, or $\lambda 3$ were identified with fluorescein isothiocyanate (FITC)-labeled anti-mouse Ig $\lambda 1,2,3$ mAb (clone R26-46, Pharmingen). $\lambda 1$ -only cells were stained with monoclonal phycoerythrin (PE)-labeled anti- $\lambda 1$ antibody [clone L22.18.2, kindly provided by U. Storb (University of Chicago, Chicago) (55, 56)]. B cells expressing λX were identified with Alexa-Fluor647-conjugated anti- λX monoclonal antibody (10C5) specific for the λX gene segment [kindly provided by P. A. Cazenave (Pasteur Institute, Paris) (57)]. For the B-cell gate, PE-Cy7-labeled CD45R/B220 antibody (RA3-6B2, 1:400, Pharmingen) was used. Fluorescence analysis was performed on a FACS Canto (BD Biosciences) with DiVa software (BD Biosciences) and analyzed with FlowJo software (Tree Star, Inc.).

Cell Sorting and H Chain Amplification. Single cell sorting of B220+ λX + cells and B220+ $\lambda 1,2,3$ + cells followed by reverse transcriptase-PCR was performed as described previously (58). V_H genes from single cells were amplified using a degenerate H chain (FR1) sense primer V_H5.3 (34) and antisense C _{μ} primers (IgMc1st 5'-GAAGGAAATGGTCTGGGGAGCAGGAA-3' and IgMc2nd 5'-GACAGGGGGCTCTCGCAGGA-3'). V_L genes were amplified with V $\lambda 1,2,3$ pri-

mers (54) and λ X primers (sense: λ X-L 5'-ATGGCCTGGACTCCTCTCTCT-3'; λ X -64 5'-TGCACCTTGAGTAGTCAGCACAGT-3' and antisense λ XR1 5'-CTGG-ACAGTGACCTTGGTT-3'; λ XR2 5'-GTTCCACCCGCGAAACATA-3'). PCR products were sequenced at the University of Chicago sequencing core. H and L chain genes were identified using IgBLAST (<http://www.ncbi.nlm.nih.gov/igblast/>) and IMGTV-quest (http://www.imgt.org/IMGT_vquest/share/textes/).

CDR3 Spectratyping. Spectratyping protocol is given in *SI Methods*.

- Radic MZ, et al. (1993) Residues that mediate DNA binding of autoimmune antibodies. *J Immunol* 150:4966–4977.
- Shlomchik M, et al. (1990) Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic mutation. *J Exp Med* 171:265–292.
- Radic MZ, et al. (1989) Structural patterns in anti-DNA antibodies from MRL/lpr mice. *Cold Spring Harb Symp Quant Biol* 54:933–946.
- Li H, Jiang Y, Prak EL, Radic M, Weigert M (2001) Editors and editing of anti-DNA receptors. *Immunity* 15:947–957.
- Ibrahim SM, Weigert M, Basu C, Erikson J, Radic MZ (1995) Light chain contribution to specificity in anti-DNA antibodies. *J Immunol* 155:3223–3233.
- Radic MZ, Ibrahim SM, Rauch J, Camper SA, Weigert M (1995) Constitutive secretion of transgene-encoded IgG2b autoantibodies leads to symptoms of autoimmune disease. *J Immunol* 155:3213–3222.
- Gay D, Saunders T, Camper S, Weigert M (1993) Receptor editing: an approach by autoreactive B cells to escape tolerance. *J Exp Med* 177:999–1008.
- Brekke KM, Garrard WT (2004) Assembly and analysis of the mouse immunoglobulin kappa gene sequence. *Immunogenetics* 56:490–505.
- Tiegs SL, Russell DM, Nemazee D (1993) Receptor editing in self-reactive bone marrow B cells. *J Exp Med* 177:1009–1020.
- Pelanda R, et al. (1997) Receptor editing in a transgenic mouse model: site, efficiency, and role in B cell tolerance and antibody diversification. *Immunity* 7:765–775.
- Chen J, et al. (1993) B cell development in mice that lack one or both immunoglobulin kappa light chain genes. *EMBO J* 12:821–830.
- Sanchez P, Marche PN, Le Guern C, Cazenave PA (1987) Structure of a third murine immunoglobulin lambda light chain variable region that is expressed in laboratory mice. *Proc Natl Acad Sci USA* 84:9185–9188.
- Doyle CM, Han J, Weigert MG, Prak ET (2006) Consequences of receptor editing at the lambda locus: multireactivity and light chain secretion. *Proc Natl Acad Sci USA* 103:11264–11269.
- Radic MZ, Mascelli MA, Erikson J, Shan H, Weigert M (1991) Ig H and L chain contributions to autoimmune specificities. *J Immunol* 146:176–182.
- Viale AC, Coutinho A, Freitas AA (1992) Differential expression of VH gene families in peripheral B cell repertoires of newborn or adult immunoglobulin H chain congenic mice. *J Exp Med* 175:1449–1456.
- Levine MH, et al. (2000) A B-cell receptor-specific selection step governs immature to mature B cell differentiation. *Proc Natl Acad Sci USA* 97:2743–2748.
- Hachemi-Rachedi S, Cumano A, Drapier AM, Cazenave PA, Sanchez P (1997) Does positive selection determine the B cell repertoire? *Eur J Immunol* 27:1069–1074.
- Hachemi-Rachedi S, Drapier AM, Cazenave PA, Sanchez P (2000) Affiliation to mature B cell repertoire and positive selection can be separated in two distinct processes. *Int Immunol* 12:385–395.
- Decker DJ, Kline GH, Hayden TA, Zaharevitz SN, Klinman NR (1995) Heavy chain V gene-specific elimination of B cells during the pre-B cell to B cell transition. *J Immunol* 154:4924–4935.
- Bloom DD, et al. (1993) V region gene analysis of anti-Sm hybridomas from MRL/Mp-lpr/lpr mice. *J Immunol* 150:1591–1610.
- Feeney AJ (1990) Lack of N regions in fetal and neonatal mouse immunoglobulin V-D-J junctional sequences. *J Exp Med* 172:1377–1390.
- Chang Y, Paige CJ, Wu GE (1992) Enumeration and characterization of DJH structures in mouse fetal liver. *EMBO J* 11:1891–1899.
- Reth MG, Alt FW (1984) Novel immunoglobulin heavy chains are produced from DJH gene segment rearrangements in lymphoid cells. *Nature* 312:418–423.
- Louzoun Y, Luning Prak E, Litwin S, Weigert M (2002) D is for different-differences between H and L chain rearrangement. *Semin Immunol* 14:239–241.
- Durdik J, Moore MW, Selsing E (1984) Novel kappa light-chain gene rearrangements in mouse lambda light chain-producing B lymphocytes. *Nature* 307:749–752.
- Moore MW, Durdik J, Persiani DM, Selsing E (1985) Deletions of kappa chain constant region genes in mouse lambda chain-producing B cells involve intrachromosomal DNA recombinations similar to V-J joining. *Proc Natl Acad Sci USA* 82:6211–6215.
- Radic MZ, Erikson J, Litwin S, Weigert M (1993) B lymphocytes may escape tolerance by revising their antigen receptors. *J Exp Med* 177:1165–1173.
- Huang H, Kearney JF, Grusby MJ, Benoist C, Mathis D (2006) Induction of tolerance in arthritogenic B cells with receptors of differing affinity for self-antigen. *Proc Natl Acad Sci USA* 103:3734–3739.
- Li Y, Li H, Weigert M (2002) Autoreactive B cells in the marginal zone that express dual receptors. *J Exp Med* 195:181–188.
- Kleinfield R, et al. (1986) Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1+ B-cell lymphoma. *Nature* 322:843–846.
- Kleinfield RW, Weigert MG (1989) Analysis of VH gene replacement events in a B cell lymphoma. *J Immunol* 142:4475–4482.
- Reth M, Gehrman P, Petrac E, Wiese P (1986) A novel VH to VHDJH joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production. *Nature* 322:840–842.
- Wilson PC, et al. (2000) Receptor revision of immunoglobulin heavy chain variable region genes in normal human B lymphocytes. *J Exp Med* 191:1881–1894.
- Chen C, Nagy Z, Prak EL, Weigert M (1995) Immunoglobulin heavy chain gene replacement: a mechanism of receptor editing. *Immunity* 3:747–755.
- Pewzner-Jung Y, et al. (1998) B cell deletion, anergy, and receptor editing in “knock in” mice targeted with a germline-encoded or somatically mutated anti-DNA heavy chain. *J Immunol* 161:4634–4645.
- Zhang Z, et al. (2003) Contribution of Vh gene replacement to the primary B cell repertoire. *Immunity* 19:21–31.
- Koralov SB, Novobrantseva TI, Königsmann J, Ehlich A, Rajewsky K (2006) Antibody repertoires generated by VH replacement and direct VH to JH joining. *Immunity* 25:43–53.
- Collins AM, et al. (2004) Partitioning of rearranged Ig genes by mutation analysis demonstrates D-D fusion and V gene replacement in the expressed human repertoire. *J Immunol* 172:340–348.
- Ohm-Laursen L, Nielsen M, Larsen SR, Barington T (2006) No evidence for the use of DIR, D-D fusions, chromosome 15 open reading frames or VH replacement in the peripheral repertoire was found on application of an improved algorithm, JointML, to 6329 human immunoglobulin H rearrangements. *Immunology* 119:265–277.
- Ait-Azzouzene D, et al. (2005) An immunoglobulin C kappa-reactive single chain antibody fusion protein induces tolerance through receptor editing in a normal polyclonal immune system. *J Exp Med* 201:817–828.
- Velez MG, et al. (2007) Ig allotypic inclusion does not prevent B cell development or response. *J Immunol* 179:1049–1057.
- Casellas R, et al. (2007) Igkappa allelic inclusion is a consequence of receptor editing. *J Exp Med* 204:153–160.
- Liu S, et al. (2005) Receptor editing can lead to allelic inclusion and development of B cells that retain antibodies reacting with high avidity autoantigens. *J Immunol* 175:5067–5076.
- Covey LR, Ferrier P, Alt FW (1990) VH to VHDJH rearrangement is mediated by the internal VH heptamer. *Int Immunol* 2:579–583.
- Taki S, Schwenk F, Rajewsky K (1995) Rearrangement of upstream DH and VH genes to a rearranged immunoglobulin variable region gene inserted into the DQ52-JH region of the immunoglobulin heavy chain locus. *Eur J Immunol* 25:1888–1896.
- Koti M, Kataeva G, Kaushik AK (2010) Novel atypical nucleotide insertions specifically at VH-DH junction generate exceptionally long CDR3H in cattle antibodies. *Mol Immunol* 47:2119–2128.
- Davila M, et al. (2007) Multiple, conserved cryptic recombination signals in VH gene segments: detection of cleavage products only in pro B cells. *J Exp Med* 204:3195–3208.
- Yunk L, Meng W, Cohen PL, Eisenberg RA, Luning Prak ET (2009) Antibodies in a heavy chain knock-in mouse exhibit characteristics of early heavy chain rearrangement. *J Immunol* 183:452–461.
- Nakajima PB, Kiefer K, Price A, Bosma GC, Bosma MJ (2009) Two distinct populations of H chain-edited B cells show differential surrogate L chain dependence. *J Immunol* 182:3583–3596.
- Lutz J, Müller W, Jäck HM (2006) VH replacement rescues progenitor B cells with two nonproductive VDJ alleles. *J Immunol* 177:7007–7014.
- Liu Y, Li L, Mohan C (2008) The role of rearrangement at the second Ig heavy chain locus in maintaining B cell tolerance to DNA. *J Immunol* 180:7721–7727.
- Prak EL, Weigert M (1995) Light chain replacement: a new model for antibody gene rearrangement. *J Exp Med* 182:541–548.
- Li Y, Louzoun Y, Weigert M (2004) Editing anti-DNA B cells by Vlambda-dx. *J Exp Med* 199:337–346.
- Witsch EJ, Cao H, Fukuyama H, Weigert M (2006) Light chain editing generates polyreactive antibodies in chronic graft-versus-host reaction. *J Exp Med* 203:1761–1772.
- Weiss S, Lehmann K, Cohn M (1983) Monoclonal antibodies to murine immunoglobulin isotypes. *Hybridoma* 2:49–54.
- Kim JY, Kurtz B, Huszar D, Storb U (1994) Crossing the SJL lambda locus into kappa-knockout mice reveals a dysfunction of the lambda 1-containing immunoglobulin receptor in B cell differentiation. *EMBO J* 13:827–834.
- Sanchez P, Nadel B, Cazenave PA (1991) V lambda-D lambda rearrangements are restricted within a V-J-C recombination unit in the mouse. *Eur J Immunol* 21:907–911.
- Wardemann H, et al. (2003) Predominant autoantibody production by early human B cell precursors. *Science* 301:1374–1377.
- Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C (1991) *Sequences of Proteins of Immunological Interest* (5th ed. Natl Inst Health, Bethesda), DHHS Publ No (NIH) 91-3242.
- Prak EL, Trounstein M, Huszar D, Weigert M (1994) Light chain editing in kappa-deficient animals: a potential mechanism of B cell tolerance. *J Exp Med* 180:1805–1815.