Chronic Lymphocytic Leukemia B Cells Express Restricted Sets of Mutated and Unmutated Antigen Receptors

Franco Fais,* Fabio Ghiotto,* Shiori Hashimoto,* Brian Sellars,* Angelo Valetto,* Steven L. Allen,* Philip Schulman,* Vincent P. Vinciguerra,* Kanti Rai,‡ Laura Z. Rassenti,§ Thomas J. Kipps,§ Guillaume Dighiero,‖ Harry W. Schroeder, Jr.,¶ Manlio Ferrarini, and Nicholas Chiorazzi***

**Department of Medicine, North Shore University Hospital and New York University School of Medicine, Manhasset, New York 11030; ‡Department of Medicine, Long Island Jewish Medical Center, New Hyde Park, New York 11040; §Department of Medicine, University of California at San Diego, La Jolla, California 92037; ‖Department of Medicine, Institut Pasteur, 75724 Paris, France; ¶Department of Medicine, The University of Alabama at Birmingham, Birmingham, Alabama 35294; and **The Division of Clinical Immunology, Istituto Nazionale per la Ricerca sul Cancro, Dipartmento di Oncologia Clinica e Sperimentale, Universita di Genova, 16132 Genova, Italy*

Abstract

To better understand the stage(s) of differentiation reached by B-type chronic lymphocytic leukemia (B-CLL) cells and to gain insight into the potential role of antigenic stimulation in the development and diversification of these cells, we analyzed the rearranged V_H genes expressed by 83 B-CLL cells (64 IgM^+ and 19 non-IgM^+). Our results confirm and extend the observations of a bias in the use of certain V_H, D, and J_H genes among B-CLL cells. In addition, they indicate that the V_H genes of $\sim$ 50% of the IgM^+ B-CLL cells and $\sim$ 75% of the non-IgM^+ B-CLL cells can exhibit somatic mutations. The presence of mutation varies according to the V_H family expressed by the B-CLL cell (V_H3 expressers displaying more mutation than V_H1 and V_H4 expressers). In addition, the extent of mutation can be sizeable with $\sim$ 32% of the IgM^+ cases and $\sim$ 68% of the non-IgM^+ cases differing by $>$ 5% from the most similar germline gene. Approximately 20% of the mutated V_H genes display replacement mutations in a pattern consistent with antigen selection. However, CDR3 characteristics (D and J_H gene use and association and HCDR3 length, composition, and charge) suggest that selection for distinct B cell receptors (BCR) occurs in many more B-CLL cells. Based on these data, we suggest three prototypic BCR, representing the V_H genes most frequently encountered in our study. These data suggest that many B-CLL cells have been previously stimulated, placing them in the "experienced" or "memory" $CD5^+$ B cell subset. (*J. Clin. Invest.* 1998. 102:1515–1525.)
Key words: antibodies • antigens • mutational analysis, DNA • receptors, antigen, B cell • hematologic neoplasms

Introduction

Over the past decade, there has been considerable interest and controversy about the immunobiology of the B cell clone that is overexpanded in B-type chronic lymphocytic leukemia (B-CLL).[1] In particular, two questions have been stressed: whether B-CLL cells have experienced antigenic stimulation and whether antigenic experience, if encountered, influences the development and diversification of these cells and their Ig variable region (V) genes.

Several groups, including our own, have attempted to address these questions by analyzing the expressed V gene repertoire of B-CLL cells. Because V gene repertoires are influenced by apparent combinatorial, expression, and selection biases (1–4) that occur at various stages of B cell differentiation and in distinct B cell subsets (5–7), it is best to compare V gene analyses with B cells of a similar subset. In this regard, as B-CLL cells almost invariably express the CD5 antigen, it has been presumed that their V gene use and mutation frequency would resemble those of human $CD5^+$ B cells (8) and their mouse homologues (9). Specifically, it has been assumed that B-CLL cells would accumulate little, if any, somatic mutation. This notion has been supported by several studies (10–17) and challenged by others (18–22).

In the present study, we have analyzed the V_H gene sequences of a large cohort of IgM-expressing and non-IgM–expressing B-CLL cases. The results of these studies confirm that V_H gene use among B-CLL cells is not random. They also indicate that these V_H genes frequently undergo somatic mutation, although mutation frequencies vary according to the V_H family expressed by the B-CLL cell. Furthermore, although only a small subset of these somatically mutated V_H gene sequences display replacement mutations in a pattern typical of antigen selection, other findings (e.g., complementarity-determining region[CDR]3 characteristics) suggest that selection for distinct subsets of surface membrane Ig receptors has occurred in many of these cells. These findings indicate that many B-CLL cells are derived from previously stimulated $CD5^+$ B cells.

Methods

CLL samples. From a large cohort of patients with clinical and laboratory features of B-CLL, 69 patients with expansions of $IgM^+/CD5^+/CD19^+$ B cells were chosen randomly for study. From the same cohort, 19 patients with expansions of $CD5^+/CD19^+$ B cells expressing IgG or IgA were also analyzed; some of these sequences were described previously (22). PBMC from these patients, obtained from anticoagulated venous blood by density gradient centrifugation (Fi-

This study was presented in part at the 7th International Workshop on Chronic Lymphocytic Leukemia, May 2–4, 1997, Crete, Greece.

Address correspondence to Nicholas Chiorazzi, North Shore University Hospital, 350 Community Drive, Manhasset, NY 11030. Phone: 516-562-1085; FAX: 516-562-1683; E-mail: nchizzi@nshs.edu

Received for publication 6 February 1998 and accepted in revised form 13 August 1998.

J. Clin. Invest.

0021-9738/98/10/1515/11 $2.00
Volume 102, Number 8, October 1998, 1515–1525
http://www.jci.org

1. *Abbreviations used in this paper:* aa, amino acid; B-CLL, B-type chronic lymphocytic leukemia; CDR, complementarity-determining region; FR, framework region; R, replacement; V, variable.

coll-Paque; Pharmacia LKB Biotechnology, Piscataway, NJ), were used either immediately or after thawing samples that had been cryopreserved with a programmable cell freezing machine (CryoMed, Mt. Clemens, MI).

cDNA prepared from these B-CLL samples were screened for expression of a dominant V_H family (representing that of the B-CLL clone) by standard PCR and, in some cases, by ELISA–PCR. In five of the IgM-expressing cases, the ELISA–PCR revealed equal expression of two different V_H families and, therefore, these samples were not analyzed further. These cases may represent B-CLL cells that lack allelic exclusion, as recently reported by Rassenti et al (23).

Preparation of RNA and cDNA synthesis. Total RNA was isolated from PBMC using Ultraspec RNA (Biotecx Laboratories, Houston, TX) according to the manufacturer's instructions. 1 μg of RNA was reverse transcribed to cDNA using 200 U of M-MLV reverse transcriptase (GIBCO BRL, Life Technologies, Grand Island, NY), 1 U of RNase inhibitor (5 Prime 3 Prime, Boulder, CO) and 20 pmol of oligo dT primer in a total volume of 20 μl. Reactions were carried out at 42°C for 1 h, heated at 65°C for 10 min to stop the reactions, and then diluted to a final volume of 100 μl.

PCR conditions for V_H family assignment and cDNA sequencing. To determine the V_H gene family used by the B-CLL cells, 2 μl of cDNA were amplified using a sense V_H family–specific framework region (FR) primer in conjunction with the appropriate antisense C_H primer (Table I). The reactions were carried out in 50 μl using 20 pmol of each primer and cycled with a 9600 GeneAmp System (Perkin Elmer Cetus, Emeryville, CA) as follows: denaturation at 94°C for 45 s; annealing at 65°C for 45 s; and extension at 72°C for 45 s. After 35 cycles, extension was continued for an additional 10 min. In certain instances, the V_H family expressed by the B-CLL cells was confirmed by ELISA–PCR as previously described (24).

V_H gene DNA sequences were determined by reamplifying 5 μl of the original cDNA using the appropriate V_H leader and C_H primers (Table I). These reactions were carried out as follows: denaturation at 94°C for 45 s; annealing at 62°C for 30 s; and extension at 72°C for 45 s. After 30–32 cycles, extension was continued for an additional 10 min. PCR products were sequenced directly after purification with Wizard PCR Preps (Promega, Madison, WI) using an automated sequenator (Applied Biosystems, Foster City, CA). In some instances where mutations were detected, an independent PCR product was generated. This product was then either sequenced directly or cloned into TA vector (Invitrogen, San Diego, CA), processed using Wizard minipreps (Promega), and sequenced using M13 forward and reverse primers.

Table I. Oligonucleotide Primers Used for V_H Gene Amplification

V_H Leader primers	
V_H1 and V_H7:	ATGGACTGGACCTGGAGG
V_H2:	CAC(GA)CTCCTGCTGCTGACCA
V_H3a:	GCTGGGTTTTCCTTGTTGC
V_H3b:	ATGGAGTT(TG)GG(AG)CTGAGCTG
V_H4:	GCTCCCAGATGGGGTCCTG
V_H5:	CTCCTCCTGGCTGTTCTCC
V_H6:	CTGTCTCCTTCCTCATCTTCC
V_H Family–specific FR1 primers	
V_H1:	GT(GA)CAGTCTGG(GA)(GC)CTGAGG
V_H2:	AAACCCACA(CG)AGACCCTCAC
V_H3:	GGTCCCT(GT)AGACTCTCCTGT
V_H4:	(CG)ACCCTGTCCCTCACCTGC
V_H5:	AAAGCCCGGGGAGTCTCTG
V_H6:	CCCTCGCAGACCCTCTCAC
V_H7:	GGTGCAATCTGGGTCTGAGTT
C_H Isotype-specific primers	
IgM:	CAGGAGAAAGTGATGGAGTCG
IgG:	GGGGAAGTAGTCCTTGACCAG
IgA:	GAGGCTCAGCGGGAAGACCTT

Analyses of V_H, D, and J_H sequences. Sequences were compared with those in the V BASE sequence directory (25) using MacVector software, version 6.0 (Eastern Kodak Co., New Haven, CT). In those instances where a > 1% deviation from a germline sequence was found in an expressed V_H gene, the algorithm of Chang and Casali (26) was used to determine whether "antigen selection" of the replacement (R) mutations had occurred, taking into account the inherent susceptibility of CDR to R mutations. Thus, the expected number of R mutations in CDR and FR was calculated using the formula $R = n \times CDR\ Rf$ (or $FR\ Rf$) $\times$ $CDRrel$ (or $FRrel$). Where n is the total number of observed mutations, Rf is the replacement frequency inherent to the CDR or FR, and $CDRrel$ and $FRrel$ are the relative sizes of the segments. A binomial probability model was used to evaluate whether the excess of R mutations in CDR or the scarcity in FR was due to chance (26).

The criteria of Corbett et al. (i.e., the requirement for 10 consecutive nucleotides of identity [27]) were used to assign members of the longer D gene families (D2 and D3); for members of the shorter D families (D1, D3, D4, D5, D6, D7), a requirement for seven consecutive nucleotides and no more than two differences was used. The DIR segments and the "minor" D segments were eliminated from these analyses as suggested (27). Reading frames of the D segments were categorized as yielding either a stop codon, a hydrophilic segment, or a hydrophobic segment (27, 28).

Analyses of HCDR3 rearrangements. HCDR3 length was determined by counting the number of amino acids (aa) between position 94 at the end of FR3 (usually two aa downstream of the conserved cystine) and position 102 at the beginning of FR4 (a conserved tryptophan in all J_H segments). Acidic and basic aa and HCDR3 charge, as defined by an estimated pI, were determined using the MacVector software programs.

Results

V_H gene use. The characteristics of the IgH chain variable region gene cDNA sequences of the 64 IgM^+ B-CLL cells and the 19 non-IgM^+ (17 IgG^+ and 2 IgA^+) B-CLL cells are listed in Tables II and III. These data indicate that the V_H genes used by these 81 leukemias were derived from each of the seven human V_H gene families in the following distribution: V_H1: 24.1%, V_H2: 1.2%, V_H3: 38.6%, V_H4: 30.1%, V_H5: 2.4%, V_H6: 2.4%, and V_H7: 1.2%. When divided into IgM^+ and non-IgM^+ categories, the V_H family distributions were different. For the IgM^+ samples, the percentages of the major families were V_H1: 28.1%, V_H3: 37.5%, and V_H4: 26.6%, whereas for the non-IgM^+ cases, they were V_H1: 10.5%, V_H3: 42.1%, and V_H4: 42.1%.

Overall, the most frequently encountered gene was V_H 4-34 (n = 15), which represented 18.1% of all the genes detected and 60% of the V_H4 genes (Tables II and III). The distribution of 4-34 among the V_H4 genes of the IgM^+ group was 52.9% and of the non-IgM^+ group, 75%. The next two most frequently found genes were V_H 3-07 (n = 10) and V_H 1-69 (n = 6), which comprised 12.1 and 7.2% of the total and 31.3 and 30.0% of their respective families. These two genes were found almost exclusively in the IgM-expressing group. Several other genes were found in multiple cases: 1-02 (n = 5), 1-18 (n = 4), 1-03 (n = 3), 3-23 (n = 4), 3-30 (n = 3), 4-39 (n = 4), and 4-59 (n = 3). Each of these was represented in both the IgM^+ and the non-IgM^+ groups, except for 4-59 which was

Table II. Molecular Genetic Characteristics of the IgM$^+$ CD5$^+$ B-CLL Cases

IgM CLL No.	Most similar germline V_H gene*	Percent V_H gene difference	Probability that R mutations occurred by chance‡		Likely D segment and reading frame*	J_H	HCDR3 length	HCDR3 sequence and charged residues‖	Number of charged residues and estimated pI in HCDR3¶			GenBank/ EMBL/DDBJ accession number
			CDR	FR					+	−	pI	
008	1-69	0.0	—	—	3-10; -philic	5b	14	VWGGSGSY YIWFDP	0	1	3.43	AF021950
011	1-02	0.0	—	—	6-19; stop	4b	12	EQWLVLE*H* YFDY	1	3	3.88	U71104
014	1-69	0.0	—	—	3-3; -philic	6c	21	*K*NDFWSGYYEG YYYYYYMDV	1	3	3.77	AF021951
017	1-69	0.0	—	—	3-3; -philic	4b	10	NYDFWSGYP Y	0	1	3.43	AF021952
020	1-58	0.0	—	—	1-26; -phobic	4b	11	PPLVGATTIG Y	0	0	5.50	AF021955
021	1-03	10.8	0.1427	§0.0004	N.A.; —	3b	14	GYIYGDYTWGTL DI	0	2	3.22	AF021956
042	1-18	0.0	—	—	3-3; -philic	4b	19	D*R*TP*R*YYDFWSGYYN *H*FDY	3	3	5.20	AF021967
047	1-18	0.1	—	—	3-3; -philic	6b	17	DYDFWSG YYPYYYGMDV	0	3	3.10	AF021968
048	1-02	2.7	0.3571	0.1412	6-13; -philic	3b	15	VDWTGYSSSWA AFDI	0	2	3.22	AF021969
051	1-69	0.0	—	—	3-3; -phobic	6c	18	VEIFGVVNLN YYYYYMDV	0	2	3.32	AF021970
063	1-46	0.0	—	—	N.A.; —	6b	21	MGNSGYSSSLGVD YYYYGMDV	0	2	3.22	AF021974
099	1-03	5.7	0.2349	§2 E-05	6-19; -phobic	4b	13	EDITVTGTG GFDY	0	3	3.16	AF021984
110	1-02	4.3	0.2406	0.1091	N.A.; —	4b	17	DL*R*VYYY*H*NLG*H* YFLDF	3	2	6.00	AF021988
112	1-69	0.0	—	—	N.A.; —	6c	20	YGGGYNLFSYQL YYYYYMDV	0	1	3.43	U71105
118	1-69	4.1	0.0997	0.1976	N.A.; —	3b	15	DAE*K*TAGTYSS AFDI	1	3	3.77	AF021992
130	1-03	0.0	—	—	5-12; -philic	6b	15	MYSGYSY YYYYGMDV	0	1	3.43	AF022000
152	1-02	0.0	—	—	N.A.; —	5b	16	GGES*R*APIVTY NWFDP	1	2	4.12	AF022007
154	1-18	2.4	0.2560	0.2135	6-19; stop	4b	11	EQWLVLS*H* FDY	1	2	4.11	AF022009
							15.5 ave.				*3.85 ave.*	
113	2-05	7.0	0.1678	§0.0051	N.A.; —	4b	15	*RRH*QGDTWSYG AFDY	3	2	7.00	AF021989
003	DP58	4.4	0.0521	§0.0367	4-23; -philic	6a	16	GDYGGNG YFYYYAMDV	0	2	3.22	AF021949
018	3-07	5.7	0.1969	§0.0040	N.A.; —	4b	7	GA YYFGY	0	0	5.50	AF021953
027	3-30.3	0.0	—	—	3.3; -philic	6b	22	GGADYDFWSGYY *H*PLE*K*GGMDV	2	3	3.99	AF021960
035	3-07	6.1	0.2059	0.1525	N.A.; —	4b	13	GCGAASC*R* YYFDY	1	1	5.95	AF021961
038	3-07	7.8	0.1850	§0.0362	N.A.; —	4b	11	DGGPPD YGMDV	0	3	3.10	AF021962
056	3-30.3	0.0	—	—	3-9; stop	6b	20	DGYEGYFDWLYN YYYYGMDV	0	4	3.06	AF021971
058	3-15	0.0	—	—	3-9; stop	4b	16	LL*R*YFDWLLSP YYFDY	1	2	4.00	AF021972
059	VH3-8	0.7	—	—	4-17; -phobic	6b	22	DPETTVTTEGYA*R*N YYYYGMDV	1	4	3.66	AF021973
065	3-73	7.3	0.1854	§0.0236	N.A.; —	3a	9	LYYD GSPNC	0	1	3.43	AF021976
066	3-48	8.1	§0.0400	§0.0486	N.A.; —	6b	13	A*R*SSSSWYND MDV	1	2	3.95	AF021977
081	3-48	4.4	0.2343	0.0861	N.A.; —	6b	13	GG YL*R*DYYGMDV	1	2	3.95	AF021979
085	3-74	2.4	§0.0028	0.1035	N.A.; —	4b	16	GAPGYD*R*SGSL YYFDY	1	2	3.95	AF021981
105	3-23	1.7	§0.0365	0.3432	2-21; -philic	4b	16	DQCGGDCP*R*LGG YFDY	1	3	3.67	AF021986
108	3-07	6.8	§0.0333	§0.0048	N.A.; —	4b	13	TLAVQEEAG YFNY	0	2	3.44	AF021987
119	3-23	7.1	0.2133	§0.0054	N.A.; —	4d	9	DGTYD YSTS	0	2	3.22	AF021993
121	3-07	9.0	0.1634	0.1292	3-22; -philic	4b	20	GVE*KH*YYDS*R*GLNWV YYFDW	3	3	5.28	AF021994
122	3-07	1.7	0.4084	0.3431	3-3; -phobic	4b	19	V*R*DP*R*WVTIFGVVIT YFDY	2	2	6.09	AF021995
123	3-11	1.0	—	—	3-22; -philic	6b	23	D*H*YYDSSGYY*HR*LG YYYYYGMDV	3	3	5.04	AF021996
126	3-07	5.4	0.1348	0.1971	N.A.; —	4d	9	AVL*RR* TF*H*I	3	0	12.48	AF021998
135	3-07	4.1	0.2356	§0.0362	N.A.; —	4b	8	V*R*FGV FDS	1	1	5.96	AF022001
138	3-33	9.7	0.1239	§0.0071	N.A.; —	4b	12	DE*R*PLGPIP FDY	1	3	3.77	AF022003
153	3-66	4.4	0.1130	§0.0330	N.A.; —	3b	11	D*R*NADGST FDI	1	3	3.67	AF022008
200	3-07	8.8	0.1928	§0.0013	N.A.; —	1	13	DLYVNMAFT*R*E *H*	2	2	5.22	AF022010
201	3-23	0.3	—	—	N.A.; —	3b	10	D*R*AVA*H* AFDI	2	2	5.11	AF022011

(Continued)

present only in the IgM$^+$ group and 3-30 which was found only in the non-IgM$^+$ group.

V_H gene mutations. A significant level of somatic mutation was found in both the IgM$^+$ and the non-IgM$^+$ groups (Tables II and III). Approximately half (51.6%; 33/64) of the IgM-expressing B-CLL cells differed by 2.0% or more from the most similar germline counterpart and 32.8% (21/64) differed by > 5.0%. Among the isotype class switched B-CLL cells, even more extensive mutation was detected with 73.7% (14/19) differing by 2.0% or more and 68.4% (13/19) differing by > 5.0%.

Mutations detected in the leukemic cells varied according

Table II (Continued)

IgM CLL No.	Most similar germline V_H gene*	Percent V_H gene difference	Probability that R mutations occurred by chance‡		Likely D segment and reading frame*	J_H	HCDR3 length	HCDR3 sequence and charged residues‖		Number of charged residues and estimated pI in HCDR3¶			GenBank/ EMBL/DDBJ accession number
			CDR	FR						+	−	pI	
							14.2 ave.					*4.61 ave.*	
002	4-34	0.0	—	—	3-3; -philic	6c	21	VTLYYDFWSGYSP	YYYYYMDV	0	2	3.22	U71103
019	4-31	0.0	—	—	4-17; -phobic	4b	9	GATVT*H*	FDY	1	1	4.96	AF021954
023	4-34	0.0	—	—	3-3; -philic	6c	17	CGFWSGYYTGIP	YYMDV	0	1	3.43	AF021957
025	4-39	0.0	—	—	3-3; -philic	6b	23	*H*PAQYAYYDFWSGYY**E**GV	YGMDV	1	3	3.77	AF021958
041	4-59	0.3	—	—	3-3; -philic	6b	24	VDPGDYDFWSGYLG*RR*	YYYYGMDV	2	4	3.91	AF021963
064	4-39	0.0	—	—	3-22; -philic	2	16	PLIYYDSSGPD	WYFDL	0	3	3.10	AF021975
067	4-34	0.0	—	—	2-15; -philic	4b	20	VFGGYCSGGSC**E**GQ**E**	YYFDY	0	3	3.24	AF021990
071	4-34	4.8	0.2753	0.1170	1-26; -philic	4b	11	*R*SGNYWG**E**	VDY	1	2	4.12	AF021991
079	4-30.2	0.0	—	—	N.A.; —	4b	11	GGWDLNY	YFDY	0	2	3.22	AF021978
083	4-59	7.2	§0.0422	§8 E-06	N.A.; —	5b	15	LST*HR*GG*R*NLD	WFDP	3	2	7.00	AF021980
093	4-34	6.6	0.2380	§0.0048	N.A.; —	1	13	GQTSSLPSG	YFLY	0	0	5.50	AF021983
125	4-34	5.1	0.0744	0.1933	1-26; -philic	4b	11	**E**GLSGSYF	VDY	0	2	3.32	AF021997
129	4-31	0.0	—	—	3-22; -philic	4b	18	L*H*YYDSSGYYPVP	YYFDY	1	2	3.94	U71106
136	4-34	6.2	§0.0014	§0.0004	N.A.; —	6b	23	G*HK*TAI*R***E**PPTIGPI	YYSYDMDV	3	3	5.28	AF022002
139	4-34	4.8	0.2753	0.1170	N.A.; —	6b	18	DFSPSPPG*H*YDA*R*ND	MDV	2	4	3.90	AF022004
141	4-34	0.0	—	—	2-2(4-b); -phobic	5b	23	GDW*R*IVVVPAAVDTAMAAN	WFDP	1	3	3.67	AF022005
147	4-59	6.5	0.0673	§0.0171	2-15; -philic	3b	17	L*HR*YCSGASCYS	DAFDI	2	2	5.11	AF022006
							17.1 ave.					*4.16 ave.*	
026	VHVMW	0.0	—	—	6-19; stop	4b	11	QQWLGGD	YFDY	0	2	3.22	AF021998
088	5-51	9.9	0.1815	0.0113	N.A.; —	4b	13	SGYYNAWYGSL	DS	0	1	3.43	AF021982
100	6-1	2.0	0.3948	0.2815	5-24; -philic	4b	11	ST*R*DGYNG	FDY	1	2	3.67	AF021985
127	6-1	8.5	0.1724	§0.0248	N.A.; —	6b	9	D*R*ADYGM	DV	1	3	3.67	AF021999

*Genes identified by two-number code with first number indicating the family, and, the second, the relative position in the locus from V_H to J_H (27, 29). N.A.: not assignable. ‡Calculated according to Chang and Casali using a binomial probability model to evaluate whether the excess of R mutations in CDR or the scarcity in FR was due to chance (26). §Denotes statistically significant difference ($P < 0.05$). ‖Amino acids on left contributed by D segment; those on right by J_H. Positively charged aa are italicized and underlined; negatively charged residues are represented in bold type. ¶Calculated from the deduced aa sequence using MacVector software, version 6.0.

to V_H family and specific V_H gene use (Tables II and III). These differences were most obvious among the IgM^+ samples. Thus, 66.7% of the V_H3 genes displayed 2.0% or more mutations, whereas only 33.3% of the V_H1 genes differed by $> 2\%$ ($P < 0.005$; Mann Whitney Test). Furthermore, 45.8% of the V_H3 genes displayed $> 5\%$ difference in contrast to only 11.1% of the V_H1 genes. This difference was more striking when the two most frequently encountered genes in these two families (3-07 and 1-69) were compared. Thus, 88.9% of the 3-07 genes differed by 2%, whereas only 16.7% of the 1–69 genes had this level of difference ($P < 0.005$; Mann Whitney test). Similarly, 77.8% of the 3-07 genes differed by $> 5\%$, whereas none of the 1-69 genes had $> 5\%$ mutations. 41% of the V_H4 genes of the IgM^+ group displayed 2% or more mutation, and the most frequently encountered gene, 4-34, was mutated in 55.6% of the cases.

Differences in the level of mutations among the various V_H families were also seen in the non-IgM^+ group, although it was less obvious due to the lower number of cases studied. Thus, 87.5% (7/8) of the V_H3- and 75% (6/8) of the V_H4-expressing cases exhibited 2% or greater difference from their germline counterparts (Tables II and III). Indeed, in every case but one (CLL No. 183), these B cells displayed $\geq 5\%$ difference. Furthermore, in 100% of the non-IgM^+ cases, the 4-34 gene was mutated. Remarkably, the V_H 4-39–expressing cases exhibited virtually no mutations in both the IgM^+ and the non-IgM^+ settings.

D segment use and reading frame. We were able to identify D genes in 50.6% (42/83) of the B-CLL cases examined (Tables II and III). Gene assignments were made more readily for the IgM-expressing B-CLL cells than for the isotype-switched, non-IgM^+ cases; indeed, in only seven of the latter cases was D segment identification possible.

Among the 35 IgM^+ cases with identifiable D segments, genes of the D3 family were found most frequently (47.6%), followed by D6 (16.7%) and D2 (14.3%) family genes. Among the non-IgM^+ cases, genes of these three D families were used in similar proportions. D3 gene use was most frequent among the V_H3-expressing cells (66.7%) and was similar among the V_H1- and V_H4-expressing cases (46.2% and 41.2%, respectively). The most frequently used D3 segment was D3-3 (previously called DXP4), which was found in 26.2% of the identifiable D genes (11/42 cases) and in 50% of those cases expressing the V_H 1-69 gene. In general, D segments were ex-

Table III. Molecular Genetic Characteristics of the Non-IgM$^+$ CD5$^+$ B-CLL Cases

Non-IgM CLL No.	Most similar germline V_H gene	Percent V_H gene difference	Probability that R mutations occurred by chance‡		Likely D segment and reading frame	J_H	HCDR3 length	HCDR3 sequence and charged residues	Number of charged residues and estimated pI in HCDR3¶			GenBank/ EMBL/DDBJ accession number
			CDR	FR					+	−	pI	
109 (γ2)	1-18	7.7	0.0685	§0.0009	N.A.; —	4b	10	GGVQVWAN DY	0	1	3.43	AF021944
158 (α)	1-02	0.0	—	—	2-2; -philic	5b	17	GYCSSTSCY*K*GY NWFDP	1	0	8.67	AF021941
005 (γ1)	3-30	8.4	0.0803	§0.0109	N.A.; —	2	11	EQA*H*DL WFFDL	1	3	3.77	AF021964
030 (γ1)	H11	6.0	§0.0171	0.2442	N.A.; —	5b	12	A*H*SP*H*GS*H* YPS	3	0	7.26	X84334
040 (γ1)	3-30	0.3	—	—	N.A.; —	6c	13	D*R*GIGGWQN YMDV	1	2	3.95	AF021965
075 (γ3)	3-33	10.2	0.1502	§0.0012	N.A.; —	6b	20	G*K*V*R*SLDWLISGS*R*S YSLDL	3	2	8.80	AF021966
078 (γ1)	3-30	5.3	§0.0077	0.0621	N.A.; —	6c	17	DG*H*DYTWGD YYYYMDV	1	4	3.51	AF021942
087 (γ1)	3-73	6.3	§0.0481	0.0011	N.A.; —	4b	11	YDNDGNY YYNY	0	2	3.22	AF078551
089 (γ1)	3-23	6.1	§0.0313	§0.0238	N.A.; —	6a	18	L*R*SSS*R*LPG*R* YY*H*YYSMDV	4	1	9.63	AF021943
111 (γ1)	3-07	6.1	0.1851	§0.0231	N.A.; —	6b	15	EQT*K*VWL*K* YYYGMDV	2	2	6.20	AF021945
							14.6 ave.				*5.79 ave.*	
001 (γ1)	4-34	5.0	0.2436	§0.0125	3-22; -philic	6c	19	WYYFDTSGYYP*R*NF YYMDV	1	2	3.95	X84333
033 (γ1)	4-34	8.0	0.2109	§0.0098	2-15; -philic	6b	22	*R*FYCSGETC*H*SSQ FYYY*H*GLDA	3	2	6.01	X84335
039 (γ3)	4-39	0.3	—	—	6-13; -philic	5b	16	S*R*GYSSSWWSS NWFDP	1	1	5.96	X84336
055 (γ1)	4-34	9.0	0.1824	§0.0010	N.A.; —	5b	15	APLGGGAGLY NWFDP	0	1	3.43	X84338
057 (γ3)	4-39	0.7	—	—	6-13; -philic	5b	17	*H*LGYSSSWYGAA NWFDP	1	1	4.96	X84339
128 (γ1)	4-34	8.1	0.1345	§0.0006	N.A.; —	4b	12	EGDGSLLNS FDY	0	3	3.16	AF021946
132 (γ2)	4-34	5.4	0.2459	0.1090	5-24; -philic	4b	12	SG*R*DAYNY YFDS	1	2	3.95	AF021947
183 (γ2)	4-34	3.1	0.2884	§0.0335	N.A.; —	6b	18	GYGDTPTI*RR* YYYYGMDV	2	2	6.09	AF021948
							16.4 ave.				*4.69 ave.*	
097 (α)	7-04.1	0.0	—	—	3-10; stop	4b	11	VQWFGEYF FDY	0	2	3.32	AF021940

See legend to Table II.

pressed in their hydrophilic reading frames (71.4%; 30/42 cases; Tables II and III).

J_H gene use. J_H use differed between the IgM$^+$ and the non-IgM$^+$ B-CLL cases (Tables II and III). Among the IgM$^+$ group, J_H4, J_H6, and J_H3 genes predominated with an overall distribution as follows: J_H1: 3.1%, J_H2: 1.6%, J_H3: 10.8%, J_H4: 46.9%, J_H5: 6.3%, and J_H6: 31.3%. In contrast, among the non-IgM$^+$ group, J_H4 and J_H3 use was less (26.3 and 0%, respectively), with an increase in J_H5 (26.3%) and J_H6 (42.1%) use.

The pattern of J_H use differed among the various V_H families and among the most frequently encountered members of the V_H1, 3, and 4 families. Thus, 88.9% of the 3-07 genes used a J_H4 gene segment compared to 16.7% of those expressing a 1-69 gene ($P < 0.05$; Fisher Exact test). Furthermore, 50% of the 1-69 genes used a J_H6 segment compared to none for 3-07 ($P < 0.05$). Finally, 33.3% of the 4-34 genes used a J_H4 segment and 46.7% used a J_H6 segment.

CDR3 length. The average HCDR3 length for all samples was 15.12 aa (Tables II and III). This value did not differ significantly between the IgM$^+$ (15.14) and non-IgM$^+$ (15.05) samples. When CDR3 length was analyzed in relation to the V_H family incorporated into the rearranged gene, some differences were seen (V_H4: 16.84; V_H1: 15.3; V_H3: 14.31). However, when CDR3 length was compared among the most frequently used genes in these families, these differences were more striking. Thus, the 3-07 gene had an average length of 12.80, considerably shorter than the composite average of 15.11 aa. In contrast, the average length among the 4-34-expressing B-CLL cells was longer (17.0 aa). Furthermore, these 4-34–expressing cells could be divided into two categories: those with CDR3 lengths longer than the average (20.11; $n = 9$) that usually contained a J_H6 or J_H5 segment, and those shorter than the average (12.33; $n = 6$) that usually contained a J_H4 segment. Finally, the CDR3 lengths among the 1-69–expressing B-CLL cells were longer than the average (16.33 versus 15.11).

CDR3 composition and charge. The HCDR3 of the IgM$^+$ group frequently contained relatively long stretches of tyrosines at their 3′ ends coded for by the J_H6 segment (Table II). Among the B cells expressing V_H1 genes, these germline residues were altered only negligibly by somatic events. In contrast, these J_H6 sequences were altered appreciably in three of seven B-CLL cells using V_H3 genes (Nos. 027, 066, and 081), and three of six using V_H4 genes (Nos. 023, 025, and 139). This was also reflected in two of the five non-IgM$^+$ B-CLL cells using V_H3 genes (Nos. 040 and 075) and all of those using V_H4 genes (Table III).

In the absence of charged residues, rearranged HCDR3 usually have estimated pI values of 5.50 (e.g., CLL Nos. 020, 018, and 093; Table II). However, due to the presence of D and J_H segments that code for negatively charged residues (e.g., aspartic acid: D), most of the IgM$^+$ B-CLL HCDR3 segments were more acidic with estimated pI values as low as 3.06 (e.g., CLL Nos. 047, 038, 056, and 064; Table II). In only rare instances did positively charged residues offset this bias and re-

sult in pI > 5.50 (6/63 for IgM$^+$; CLL Nos. 110, 113, 122, 126, 083, and 135). The IgM$^+$ B-CLL cells using V_H1 genes had the least of these residues and, therefore, as a group had the lowest average estimated pI (3.85). In contrast, those B-CLL cells using V_H3 genes had more of these positively charged residues and had an average estimated pI of 4.61. The V_H4-expressing B-CLL cells had an intermediate phenotype with an average estimated pI of 4.16. In concert with this principle, the V_H 1-69–expressing B-CLL cells had an average pI of 3.53; those expressing 4-34 an average pI of 4.37; and those using 3-07 a pI of 5.89.

Among the non-IgM$^+$ group, HCDR3 charge was more heterogeneous (Table III). In 42.1% of cases, the estimated pI exceeded 5.50. Indeed, in 14/19 instances (73.7%), at least one positively charged residue was present in the CDR3 segment and, in those cases where the pI exceeded 5.50, there was an average of 2.4 positively charged residues/CDR3. Similar to the IgM$^+$ B-CLL cells, the non-IgM$^+$ cells that expressed V_H3 genes had the highest estimated pI (5.79).

Discussion

The non-random use of gene segments, the presence and extent of mutation, and the apparent constraints in HCDR3 structure observed in these cases provide further insights into the stage of maturation and the potential role of antigen stimulation in the evolution of these leukemic cells.

Gene use. To improve the power of these comparisons, we combined our findings for IgM$^+$ B-CLL cells with those from other B-CLL studies (21, 30–38) and then compared these data with those for normal adult CD5$^+$ blood B cells (39). With this approach, it became apparent that V_H gene family use was not comparable to that in the normal repertoire; rather, there was a statistically significant increase in V_H1 gene expression and a statistically significant decrease in V_H3 expression in B-CLL (Fig. 1). J_H family use among the pooled IgM$^+$ B-CLL cells was not different from that of the normal CD5$^+$ B cell repertoire (data not shown).

V_H 4-34, 3-07, and 1-69 were the most commonly expressed V_H genes in our B-CLL patients. Although neither 4-34 nor 3-07 were identified as dominant genes in several previous surveys of CLL patients (17, 40; reviewed in reference 41), the 4-34 gene has been found in virtually all cases of cold agglutinin disease (42–44), frequently in diffuse large cell lymphoma (45) and autoimmune disorders (46), and virtually never in multiple myeloma (40, 47). Interestingly, when we combine our B-CLL cases with those in the literature (21, 30–38), the frequency of V_H 4-34 and 3-07 expression matches that in the normal adult CD5$^+$ repertoire, although there is a statistically significant overrepresentation of the V_H 1-69 gene (Fig. 2). This latter finding confirms studies of Kipps, Carson, and coworkers, who have indicated the prominence of the 1-69 gene (31) and some of its alleles (48) in B-CLL. Although several other V_H1 and V_H3 genes (e.g., 1-02, 1-18, and 3-23) were found often in the combined B-CLL repertoire, their expression did not differ significantly from the normal. Significant underrepresentations were found for the V_H 3-30, V_H 3-30.3, and V_H 3-33 genes (Fig. 2), which may be responsible for the decreased use of V_H3 genes illustrated in Fig. 1.

However, all of these comparisons need to be evaluated with caution because of the limited numbers of normal CD5$^+$ and CD5$^-$ blood B cell sequences that are presently available in the literature for comparison. Indeed, because of this relative dearth of data, we have used the V_H sequences of 144 individual CD5$^+$ blood B cells from two normal adults (39). In addition, genetic differences in gene copy numbers may be important factors in determining gene expression. This may be especially relevant as it has been reported that copies of the V_H 1-69 and 3-23 genes, which are frequently found in B-CLL cells, and the V_H 3-30 genes, which we have found underrepresented in this study, vary among individuals (49–51).

J_H segment use differed among the three most frequently encountered genes in that ~ 90% of the V_H 3-07 genes were associated with a J_H4 segment, whereas ~ 50% of the V_H 1-69 and V_H 4-34 genes were associated with a J_H6 segment. D3 family genes were found most frequently in B-CLL cells (47.6%), followed by D6 (16.7%) and D2 (14.3%) family genes. These percentages are very similar to those recently identified in 451 rearranged HCDR3 from antibodies in the databases (27). The individual D segment used most frequently in our cohort of B-CLL cells was D3-3 which was present in ~ 26% of the identifiable D segments, a frequency that is almost threefold more than that found in the 451 normal sequences mentioned above (26.2 versus 9.6%). It has been reported (48) that V_H 1-69$^+$ B-CLL cells frequently use the D3-3 (DXP4) segment. In our cases, 50% of the 1–69 genes

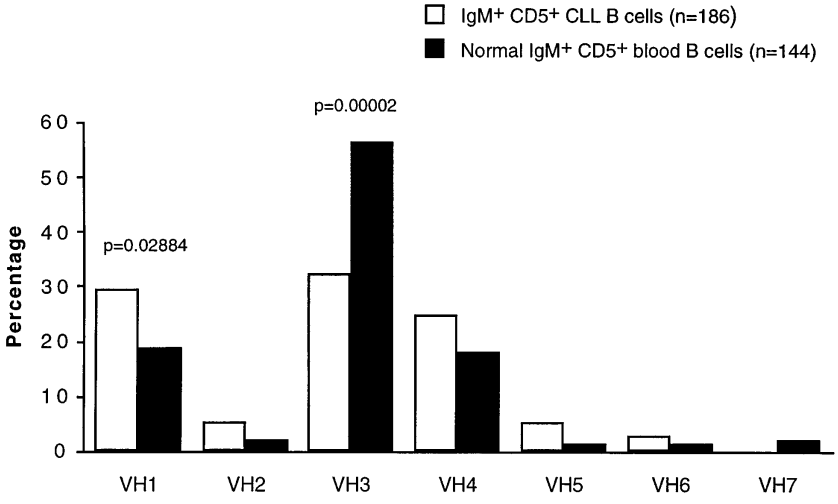

Figure 1. Comparison of V_H gene use between IgM$^+$ B-CLL cases and normal blood B cells. The IgM$^+$ B-CLL cases represent a pool of those reported in this study and those compiled by Schroeder and Dighiero (21), and available in other studies (30–38) and/or in GenBank/EMBL/DDBJ. Normal IgM$^+$ CD5$^+$ blood B cell sequences were derived from single B cells of two healthy adult males as reported by Brezinschek et al (39). Statistical comparisons were performed using the Fisher's Exact test.

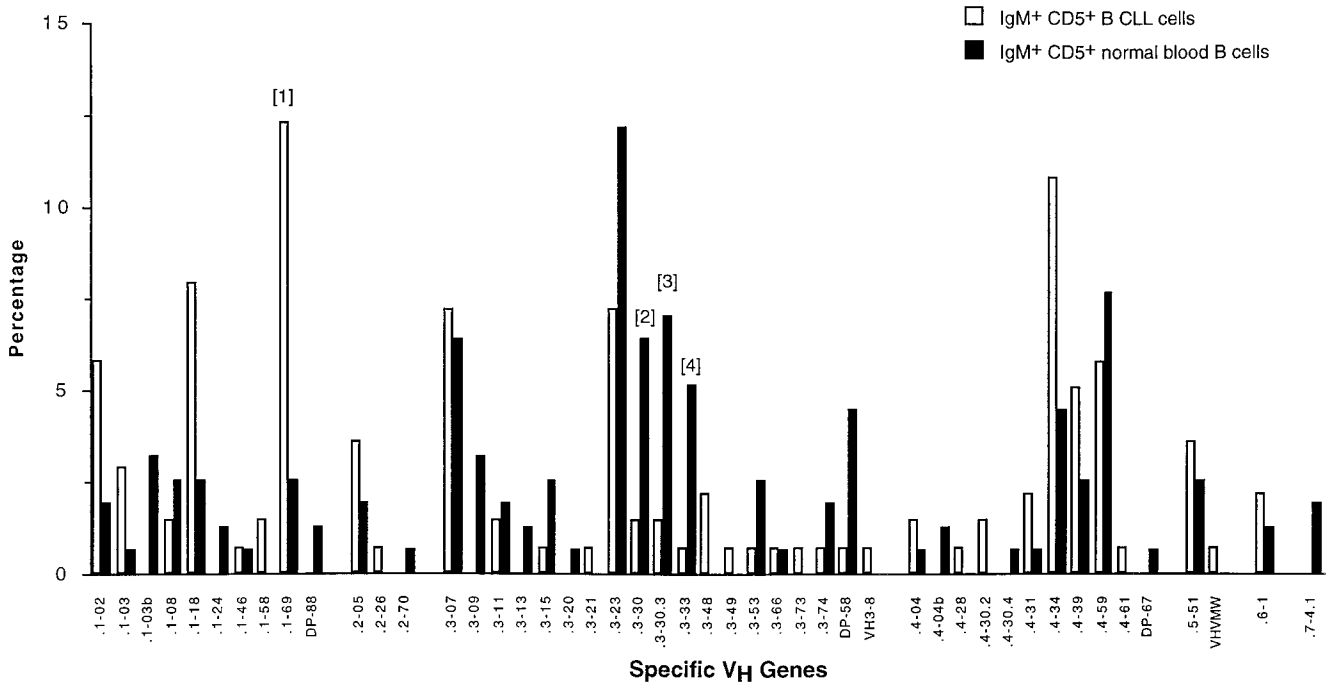

Figure 2. Comparison of specific V_H gene use between IgM^+ B-CLL cases and normal blood B cells. Statistical comparisons were performed using the Fisher's Exact test: (*1*) $P = 0.0011$; (*2*) $P = 0.0394$; (*3*) $P = 0.0230$; and (*4*) $P = 0.0397$.

were linked with this D gene. Finally, D segment expression was markedly biased in favor of hydrophilic reading frames, especially in the non-IgM^+ group. This finding is consistent with the normal adult B cell repertoire (27, 28).

Somatic mutation. Approximately 57% of the B-CLL V_H gene sequences determined in this study were mutated. Indeed, we found mutations $\geq 2\%$ in 50.8% of the IgM^+ B-CLL cells and mutations $> 5\%$ in 31.8% of these patients. An even higher percentage of the isotype-switched B-CLL cells exhibited these levels of mutation (72.2 and 66.7%, respectively). Although the percentage of IgM^+ B-CLL cells expressing mutations in the 2–5% range was similar to that reported for normal $CD5^+$ blood B cells (18.1 versus 19.1%; [39]), the percentage of B-CLL cells expressing mutations in the $> 5\%$ range was significantly different ($P < 0.000001$) from these normal $CD5^+$ B cells. Indeed, there were ten times as many cases of IgM^+ $CD5^+$ B-CLL with mutations $> 5\%$ than in the normal (31.8 versus 3.5%). Surprisingly, this frequency even exceeded that reported for circulating $CD5^-$ B cells (31.8 versus 17.6%). These latter comparisons are presented graphically in Fig. 3.

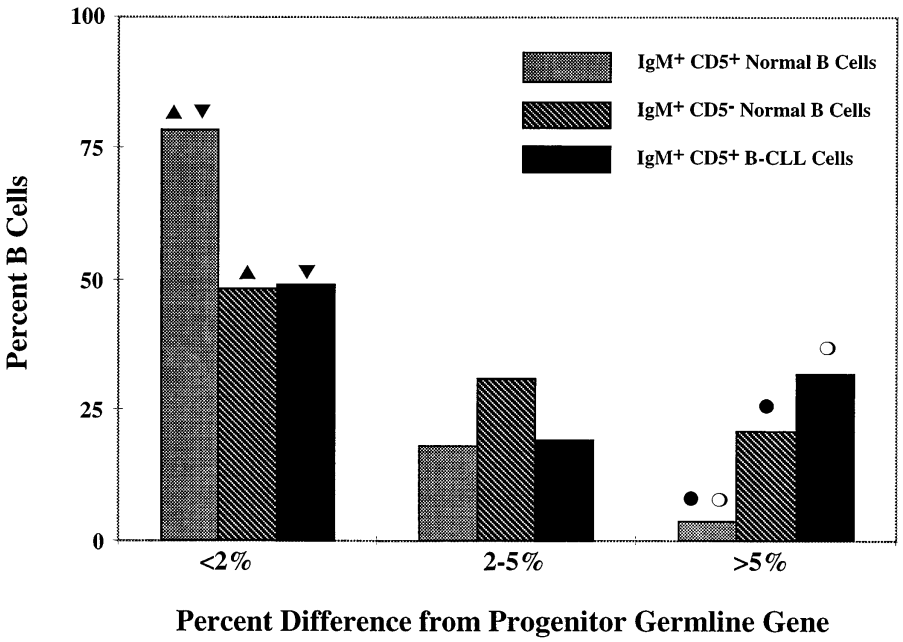

Figure 3. Comparison of extent of V_H gene mutation between normal peripheral blood IgM^+ $CD5^+$ and IgM^+ $CD5^-$ B cells, and the IgM^+ $CD5^+$ B-CLL cells in this study. The IgM^+ $CD5^+$ blood B cells ($n = 144$) and the IgM^+ $CD5^-$ B cells ($n = 206$) are from reference 39; the IgM^+ $CD5^+$ B-CLL cells ($n = 63$) are from this study. Statistically significant differences (Fisher's Exact test) were found in the $< 2\%$ group between the $CD5^+$ and the $CD5^-$ normal B cells (*regular triangle*) and between the $CD5^+$ normal cells and the B-CLL cells (*inverted triangle*) with *P* values for both < 0.00005. Similarly, within the $> 5\%$ group, statistically significant differences were found in the same two comparisons (*closed and open circles*, respectively) with *P* values for both < 0.000001.

However, these mutations did not occur uniformly among the various V_H gene families. Indeed, there was a V_H family–related hierarchy of mutation ($V_H3 > V_H4 > V_H1$). This was most obvious when comparing the 3-07 (90% of cases mutated), 4-34 (73.3%), and 1-69 (16.7%) genes. The observation that 1-69–expressing B-CLL cells have a lower frequency of mutation is supported by the study of Johnson et al. (48) that found that only 1 of 26 V_H1-69$^+$ B-CLL cells had $< 99\%$ similarity to a V_H 1-69 allele. This V_H family–related difference in mutation may explain, in part, the discrepancies about V_H gene mutation reported in other B-CLL studies, as cohorts with disproportionate numbers of either V_H1- or V_H3-expressing cases would likely yield conflicting results. A further confounding issue may be the fact that certain B-CLL cells can productively express two V_H genes, which can differ in the presence and extent of mutation (23).

B cells with receptors that have been selected by antigen often display a higher frequency of R mutations in the CDR than in the FR (52, 53). However, when assessing the possibility of antigen selection, the inherent susceptibility of the progenitor germline gene to aa replacement needs to be considered. Based on the algorithm of Chang and Casali that incorporates these considerations (26), 20% of our B-CLL cells demonstrate antigen selection of R mutations (noted by § in Tables II and III).

However, it is important to recognize that such algorithms are statistical approaches that are unable to account for all aspects of "selection biology" and, therefore, may not always accurately reflect the influences of antigen. For example, in some instances, high affinity antigen binding may depend on only one or a few critical amino acids (54–56). In these instances, the occurrence of an "advantageous" mutation would permit survival and positive selection of the mutating B cells. If, however, some of these positively selected, mutated B cells subsequently accumulate additional mutations outside of the CDR that do not alter selection by antigen, these "neutral" mutations might then result in an algorithmic determination that suggests statistical insignificance, albeit in the setting of definite biological significance. Furthermore, recent studies suggest that R mutations located outside of the CDR can enhance antigen binding (57), a fact that these algorithms by design cannot consider. Conversely, antigen selection could favor the preservation of certain residues. For example, certain portions of the FR are important for superantigen binding (58, 59) and these need to be preserved to permit superantigen-mediated selection and rescue. Indeed, in this study, 61.7% (29/47) of the cases exhibited statistically significant preservations of FR sequences (Tables II and III, marked with §). Preservation of FR integrity is consistent with the necessity for B cells to maintain an intact B cell receptor (60). Finally, these algorithms cannot take into account the structure of the H and L CDR3, the regions which have the most intimate contact with antigen.

CDR3 characteristics. We were able to define three HCDR3 categories among our B-CLL cases based on differences in length, aa composition, and charge. Each of these varied in a V_H family–related manner. First, the CDR3 length of V_H4-expressing B-CLL cells was greater than in V_H1-expressing B cells and this was in turn greater than in V_H3-expressing cells. This was most obvious when comparing the 3-07 (12.56 aa), 1-69 (16.33 aa), and 4-34 (17.0 aa) genes. The 4-34 genes, furthermore, could be divided into two categories: those with CDR3 lengths longer than average (20.11) and those shorter than average (12.33). When our IgM$^+$ cases were pooled with those available in GenBank/EMBL/DDBJ, these V_H family–related hierarchies of CDR3 length were just as apparent and, in some instances, reached statistical significance (e.g., 3-07 versus 1-69: $P < 0.05$). In most cases, the short CDR3 segments of the V_H3 group and the shorter V_H 4-34 subgroup contained a J_H4 segment, whereas the V_H1 group and the longer 4-34 subgroup contained a J_H6 or J_H5 segment.

Second, members of the V_H1 family frequently contained long stretches of tyrosines coded for by the J_H6 segment that were only minimally somatically altered. This was in contrast to most B-CLL cells using V_H3 genes and certain B-CLL cells using V_H4 genes, which had shorter CDR3 that frequently used a significantly altered J_H4 gene.

Finally, these differences in J_H gene association resulted in a V_H family–related hierarchy in charge (V_H1, V_H4, and V_H3) that, again, was most easily recognized by inspecting the V_H 1-69 (pI: 3.53), 4-34 (pI: 4.37), and 3-07 (pI: 5.89) genes of the IgM$^+$ group. Among the non-IgM$^+$ group, CDR3 charge was much less acidic, often exceeding a value of 5.50.

We believe that these V_H family– and gene-related differences in HCDR3 reflect selection for specific structural motifs that facilitate antigen binding. If so, many more than 20% of the IgM$^+$ B-CLL cells would have been selected for antigen binding. However, the stage(s) of B cell development at which these selections occur is not clear at this point.

Receptor prototypes. Taken together, these studies suggest prototypic rearranged H chain variable region genes in B-CLL cells (Fig. 4). Thus, V_H1 sequences, exemplified by the 1-69–expressing B-CLL cells, remain predominantly unmutated and frequently associated with D3 (3-3) and J_H6 segments. This results in a highly acidic HCDR3 that is slightly longer than the average. V_H3 sequences, exemplified by the 3-07–expressing B-CLLs, frequently accumulate numerous mutations and associate with D3 and J_H4 genes. These receptors have a CDR3 that is shorter and considerably less acidic than the V_H1 prototype. Finally, V_H4 sequences, exemplified by the 4-34–expressing CLLs, accumulate mutations in $\sim 50\%$ of the cases. Because these receptors use either a J_H4 or J_H6 gene, their CDR3 are either shorter, like the V_H3 group, or longer, like the V_H1 group. In addition, they tend to be slightly less acidic than the V_H1 group.

It is unclear at this juncture whether these prototypes represent features unique to B-CLL cells or are shared with B cells of the normal adult or aging repertoire. Although we have been able to collect data on 188 B-CLL cases by combining our data with those in published articles and in GenBank/EMBL/DDBJ (21, 30–38), a coordinated compilation of both B-CLL gene sequences and normal B cell subset sequences will permit more definitive answers to these questions.

Maturation stage(s) of B-CLL cells. Finally, based on the presence of somatic mutations, it seems incontrovertible that $\sim 50\%$ of IgM$^+$ B-CLL cells and $\sim 75\%$ of non-IgM$^+$ cells have been previously stimulated by antigen. Therefore, these cells should be considered "experienced" or "memory" CD5$^+$ B cell progeny. This conclusion is in line with those of Schroeder and Dighiero, who previously compared the sequences of a series of B-CLL cases (21), and with the observations that IgM$^+$ B-CLL cells can undergo intraclonal isotype class switching (61–64), a process that suggests clonal maturation.

However, what about those B-CLL cells that do not display V_H gene mutations? Should these be considered virgin B cells

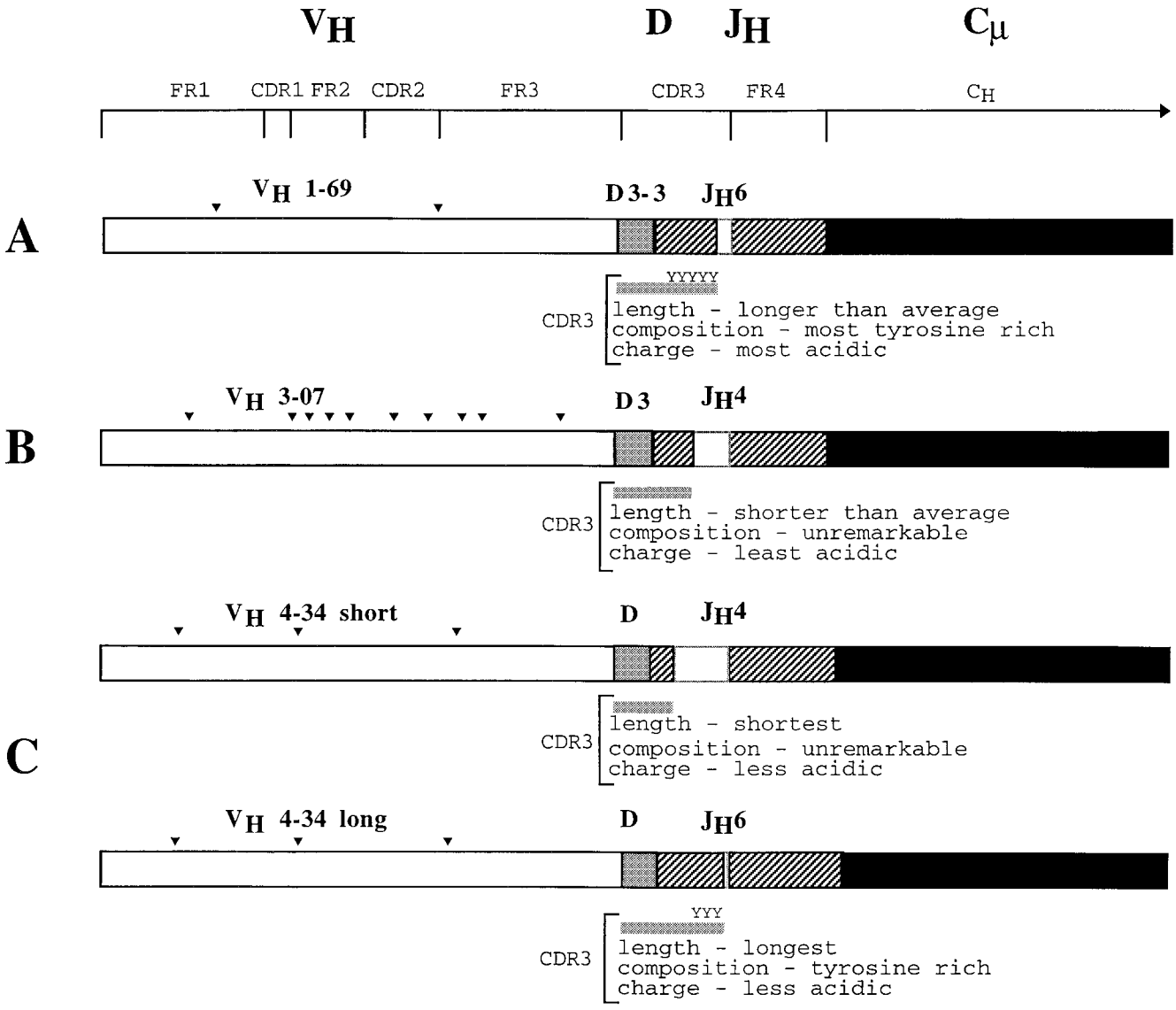

Figure 4. Prototypic variable regions of surface membrane IgM receptors in B-CLL. Schematic representations of the proposed V_H 1-69 (*A*), V_H 3-07 (*B*), and V_H 4-34 (*C*) prototypic V regions. Arrowheads represent possible mutations. See Discussion for details.

or B cells driven by antigens that cannot induce somatic mutations (e.g., autoantigens) or antigens that select for unmutated V_H sequences (e.g., superantigens) and/or distinct CDR3 characteristics (certain exo- or autoantigens)? Based on the overuse of specific V_H, D, and J_H genes, and the apparent constraints on HCDR3 structure, we favor the hypothesis that these cells also have been antigen driven and represent memory cells. However, we cannot eliminate the possibility that these receptor restrictions are developmental and represent selections that have occurred before exiting the bone marrow, thus making these virgin B cells with albeit developmentally restricted receptor structures.

Acknowledgments

We thank Cristina Sison and Martin Lesser for their help in the statistical analyses and Charles Chu and Peter K. Gregersen for their critical review of the manuscript.

These studies were supported in part by U.S. Public Health Service grants AI 10811 (N. Chiorazzi) and AI 33621 (H.W. Schroeder, Jr.), the Joseph Eletto Leukemia Research Fund, the Jean Walton Fund for Lymphoma & Myeloma Research, the Richard and Nancy Leeds Fund of the Department of Medicine of North Shore University Hospital, and the Associazione Italiana per la Ricerca sul Cancro.

References

1. Yancopoulos, G.D., S.V. Desiderio, M. Paskind, J.F. Kearney, D. Baltimore, and F.W. Alt. 1984. Preferential utilization of the most J-proximal V_H gene segments in pre-B cell lines. *Nature.* 311:727–733.

2. Perlmutter, R.M., J.F. Kearney, S.P. Chang, and L.E. Hood. 1985. Developmentally controlled expression of immunoglobulin V_H genes. *Science.* 227: 1597–1601.

3. Schroeder, H.W., Jr., J.L. Hillson, and R.M. Perlmutter. 1987. Early restriction of the human antibody repertoire. *Science.* 238:791–793.

4. Suzuki, I., L. Pfister, A. Glas, C. Nottenburg, and E.C.B. Milner. 1995. Representation of rearranged V_H gene segments in the human adult antibody repertoire. *J. Immunol.* 154:3902–3911.

5. Hillson, J.L., I.R. Oppliger, E.H. Sasso, E.C.B. Milner, and M.H. Wener.

1992. Emerging human B cell repertoire. Influence of developmental stage and interindividual variation. *J. Immunol.* 149:3741–3752.

6. Kraj, P., S.P. Rao, A.M. Glas, R.R. Hardy, E.C. Milner, and L.E. Silberstein. 1997. The human heavy chain Ig V region gene repertoire is biased at all stages of B cell ontogeny, including early pre-B cells. *J. Immunol.* 158:5824–5832.

7. Pascual, V., Y-J. Liu, A. Magalski, O. de Bouteiller, J. Banchereau, and J.D. Capra. 1994. Analysis of somatic mutations in five B cell subsets of human tonsil. *J. Exp. Med.* 180:329–339.

8. Casali, P., M.T. Kasaian, and G. Haughton. 1994. B-1 (CD5) cells. *In* Autoimmunity: Physiology and Disease. A. Coutinho and M.D. Kazatchkine, editors. John Wiley and Sons, Inc. New York. 57–88.

9. Hardy, R.R., and K. Hayakawa. 1992. Developmental origins, specificities and immunoglobulin gene biases of murine Ly-1 B cells. *Intern. Rev. Immunol.* 8:189–207.

10. Kipps, T.J., E. Tomhave, P.P. Chen, and D.A. Carson. 1988. Autoantibody-associated κ light chain variable region gene expressed in chronic lymphocytic leukemia with little or no somatic mutation. Implications for etiology and immunotherapy. *J. Exp. Med.* 167:840–852.

11. Meeker, T.C., J.C. Grimaldi, R. O'Rourke, J. Loeb, G. Juliusson, and S. Einhorn. 1988. Lack of detectable somatic hypermutation in the V region of the Ig H chain gene of a human chronic B lymphocytic leukemia. *J. Immunol.* 141: 3994–3998.

12. Pratt, L.F., L. Rassenti, J. Larrick, B. Robbins, P.M. Banks, and T.J. Kipps. 1989. Ig V region gene expression in small lymphocytic lymphoma with little or no somatic hypermutation. *J. Immunol.* 143:699–705.

13. Deane, M., and J.D. Norton. 1990. Immunoglobulin heavy chain variable region family usage is independent of tumor cell phenotype in human B lineage leukemias. *Eur. J. Immunol.* 20:2209–2217.

14. Kuppers, R., A. Gause, and K. Rajewsky. 1991. B cells of chronic lymphatic leukemia express V genes in unmutated form. *Leuk. Res.* 15:487–496.

15. Friedman, D.F., J.S. Moore, J. Erikson, J. Manz, J. Goldman, P.C. Nowell, and L.E. Silberstein. 1992. Variable region gene analysis of an isotype-switched (IgA) variant of chronic lymphocytic leukemia. *Blood.* 80:2287–2297.

16. Wagner, S.D., and L. Luzzatto. 1993. V_k gene segments rearranged in chronic lymphocytic leukemia are distributed over a large portion of the V_k locus and do not show somatic mutation. *Eur. J. Immunol.* 23:391–397.

17. Stevenson, F.K., M.B. Spellerberg, J. Treasure, C.J. Chapman, L.E. Silberstein, T.J. Hamblin, and D.B. Jones. 1993. Differential usage of an Ig heavy chain variable region gene by human B cell tumors. *Blood.* 82:224–230.

18. Cai, J., C. Humphries, A. Richardson, and P.W. Tucker. 1992. Extensive and selective mutation of a rearranged V_H5 gene in human B cell chronic lymphocytic leukemia. *J. Exp. Med.* 176:1073–1081.

19. Hashimoto, S., M. Wakai, J. Silver, and N. Chiorazzi. 1992. Biased usage of variable and constant region genes by IgG^+, $CD5^+$ human leukemic B cells. *NY Acad. Sci.* 651:477–479.

20. Ebeling, S.B., M.E.M. Schutte, and T. Logtenberg. 1993. Molecular analysis of V_H and V_L regions expressed in IgG-bearing chronic lymphocytic leukemia (CLL): further evidence that CLL is a heterogenous group of tumors. *Blood.* 82:1626–1631.

21. Schroeder, H.W., and G. Dighiero. 1994. Clues to the pathogenesis of CLL through analysis of B cell CLL antibody repertoires. *Immunol. Today.* 15: 288–294.

22. Hashimoto, S., M. Dono, M. Wakai, S.L. Allen, S.M. Lichtman, P. Schulman, V.P. Vinciguerra, M. Ferrarini, J. Silver, and N. Chiorazzi. 1995. Somatic diversification and selection of immunoglobulin heavy and light chain variable region genes in IgG^+ $CD5^+$ chronic lymphocytic leukemia B cells. *J. Exp. Med.* 181:1507–1517.

23. Rassenti, L.Z., and T.J. Kipps. 1997. Lack of allelic exclusion in B cell chronic lymphocytic leukemia. *J. Exp. Med.* 185:1435–1445.

24. Rassenti, L.Z., H. Kohsaka, and T.J. Kipps. 1995. Analysis of immunoglobulin V_H gene repertoire by an anchored PCR-ELISA. *Ann. NY Acad. Sci.* 764:463–473.

25. Tomlinson, I.M., S.C. Williams, S.J. Corbett, J.B.L. Cox, G. Winter. 1996. V BASE sequence directory. MRC Centre for Protein Engineering, Cambridge, UK.

26. 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.* 15:367–373.

27. Corbett, S.J., I.M. Tomlinson, E.L.L. Sonnhammer, D. Buck, and G. Winter. 1997. Sequence of the human immunoglobulin diversity (D) segment locus: a systematic analysis provides no evidence for the use of DIR segments, inverted D segments, "minor" D segments or D-D recombination. *J. Mol. Biol.* 270:587–597.

28. Raaphorst, F.M., C.S. Raman, J. Tami, M. Fischbach, and I. Sanz. 1997. Human Ig heavy chain CDR3 regions in adult bone marrow pre-B cells display an adult phenotype of diversity: evidence for structural selection of D_H amino acid sequences. *Int. Immunol.* 9:1503–1515.

29. Shin, E.K., F. Matsuda, H. Nagaoka, Y. Fukita, T. Imai, K. Yokoyama, E. Soeda, and T. Honjo. 1991. Physical map of the 3′ region of the human immunoglobulin heavy chain locus: clustering of autoantibody-related variable segments in one haplotype. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:3641–3645.

30. Shen A, C. Humphries, P. Tucker, and F. Blattner. 1987. Human heavy-chain variable region gene family nonrandomly rearranged in familial chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 84:8563–8567.

31. Kipps, T., E. Tomhave, L.F. Pratt, S. Duffy, P.P. Chen, and D.A. Carson. 1989. Developmentally restricted immunoglobulin heavy chain variable region gene expressed at high frequency in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 86:5913–5917.

32. Spatz, L.A., K.K. Wong, M. Williams, R. Desai, J. Golier, J.E. Berman, F.W. Alt, and N. Latov. 1990. Cloning and sequence analysis of the V_H and V_L regions of an anti-myelin/DNA antibody from a patient with peripheral neuropathy and chronic lymphocytic leukemia. *J. Immunol.* 144:2821–2828.

33. Friedman, D.F., E.A. Cho, J. Goldman, C.E. Carmak, E.C. Besa, R.R. Hardy, and L.E. Silberstein. 1991. The role of clonal selection in the pathogenesis of an autoreactive human B cell lymphoma. *J. Exp. Med.* 174:525–537.

34. Matolcsy, A., P. Casali, R.G. Nador, Y.F. Liu, and D.M. Knowles. 1997. Molecular characterization of IgA- and/or IgG-switched chronic lymphocytic leukemia B cells. *Blood.* 89:1732–1739.

35. Cherepakhin V., S.M. Baird, G.W. Meisenholder, and T.J. Kipps. 1993. Common clonal origin of chronic lymphocytic leukemia and high-grade lymphoma of Richter's syndrome. *Blood.* 82:3141–3147.

36. Korganow, A.S., T. Martin, J.C. Weber, B. Lioure, P. Lutz, A.M. Knapp, and J.L. Pasquali. 1994. Molecular analysis of rearranged V_H genes during B cell chronic lymphocytic leukemia: intraclonal stability is frequent but not constant. *Leuk. Lymphoma.* 14:55–69.

37. Ikematsu W., H. Ikematsu, T. Otsuka, S. Okamura, S. Kashiwagi, and Y. Niha. 1995. Surface phenotype and immunoglobulin heavy chain gene usage in chronic B cell leukemias. *Ann. NY Acad. Sci.* 764:492–495.

38. Foroni, L., A. Werner, M. Papaioannou, J. Yaxley, M. Attard, and A.V. Hoffbrand. 1996. The V_H usage in B cell chronic lymphocytic leukemia cells is a random event and it is not disease related. *Blood.* 88:238. (Abst.)

39. Brezinschek, H-P., S.J. Foster, R.I. Brezinschek, T. Dorner, R. Domiati-Saad, and P.E. Lipsky. 1997. Analysis of the human V_H gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral $CD5^+/IgM^+$ and $CD5^-/IgM^+$ B cells. *J. Clin. Invest.* 99:2488–2501.

40. Stevenson, F.K., M.B. Spellerberg, C.J. Chaplin, and T.J. Hamblin. 1995. Differential usage of an autoantibody-associated V_H gene, V_H4-21, by human B cell tumors. *Leuk. Lymphoma.* 16:379–384.

41. Kipps, T. 1993. Immunoglobulin genes in chronic lymphocytic leukemia. *Blood Cells.* 19:615–625.

42. Silberstein, L.E., L.C. Jefferies, J. Goldman, D. Friedman, J.S. Moore, P.C. Nowell, D. Roelcke, W. Pruzanski, J. Roudier, and G.J. Silverman. 1991. Variable region gene analysis of pathologic human autoantibodies to the related i and I red blood cell antigens. *Blood.* 78:2372–2386.

43. Pascual, V., K. Victor, D. Lelsz, M.B. Spellerberg, T.J. Hamblin, K.M. Thompson, I. Randen, J. Natvig, J.D. Capra, and F.K. Stevenson. 1991. Nucleotide sequence analysis of the V regions of two IgM cold agglutinins. Evidence that the V_H4-21 gene segment is responsible for the major cross-reactive idiotype. *J. Immunol.* 146:4385–4391.

44. Leoni, J., J. Ghiso, F. Goni, and B. Frangione. 1991. The primary structure of the Fab fragment of protein KAU, a monoclonal immunoglobulin M cold agglutinin. *J. Biol. Chem.* 266:2836–2842.

45. Hsu, F.J., and R. Levy. 1995. Preferential use of the V_H4 Ig gene family by diffuse large cell lymphoma. *Blood.* 86:3072–3082.

46. Pascual, V., and J.D. Capra. 1992. V_H4-21, a human V_H gene segment overrepresented in the autoimmune repertoire. *Arthritis Rheum.* 35:11–18.

47. Rettig, M.B., R.A. Vescio, J. Cao, C.H. Wu, J.C. Lee, E. Han, M. Der-Danielian, R. Newman, C. Hong, A.K. Lichtenstein, and J.R. Berenson. 1996. V_H gene usage in multiple myeloma: complete absence of the V_H4-21 (V_H4-34) gene. *Blood.* 87:2846–2852.

48. Johnson, T.A., L.Z. Rassenti, and T.J. Kipps. 1997. V_H1 genes expressed in B cell chronic lymphocytic leukemia exhibit distinctive molecular features. *J. Immunol.* 158:235–246.

49. Sasso, E.H., T. Johnson, and T.J. Kipps. 1996. Expression of the immunoglobulin V_H gene 51p1 is proportional to its germline gene copy number. *J. Clin Invest.* 97:2074–2080.

50. Stewart, A.K., C. Huang, B.D. Stollar, and R.R. Schwartz. 1993. High-frequency representation of a single V_H gene in the expressed human B cell repertoire. *J. Exp. Med.* 177:409–418.

51. Milner, E.C.B., W.O. Hufnagle, A.M. Glas, I. Suzuki, and C. Alexander. 1995. Polymorphism and utilization of human V_H genes. *Ann. NY Acad. Sci.* 764:50–61.

52. Jukes, T.H., and J.L. King. 1979. Evolutionary nucleotide replacements in DNA. *Nature.* 281:605–606.

53. Shlomchik, M.J., A. Marshak-Rothstein, C.B. Wolfowicz, T.L. Rothstein, and M.G. Weigert. 1987. The role of clonal selection and somatic mutation in autoimmunity. *Nature.* 328:805–811.

54. Rudikoff, S., A.M. Giusti, W.D. Cook, and M.D. Scharff. 1982. Single amino acid substitution altering antigen-binding specificity. *Proc. Natl. Acad. Sci. USA.* 79:1979–1983.

55. Jeske, D.J., J. Jarvis, C. Milstein, J.D. Capra. 1984. Junctional diversity is essential to antibody activity. *J. Immunol.* 133:1090–1092.

56. Kussie, P.H., B. Parhami-Seren, L.J. Wysocki, and M.N. Margolies.

1994. A single engineered amino acid substitution changes antibody fine specificity. *J. Immunol.* 152:146–152.

57. Katz, J.B., W. Limpanasithikul, and B. Diamond. 1994. Mutational analysis of an autoantibody: differential binding and pathogenicity. *J. Exp. Med.* 180:925–932.

58. Hillson, J., N.S. Karr, I.R. Oppliger, M. Mannik, and E.H. Sasso. 1993. The structural basis of germline-encoded V_H3 immunoglobulin binding to Staphylococcal protein A. *J. Exp. Med.* 178:331–336.

59. Silverman, G.J., R. Pires, and J.P. Bouvet. 1996. An endogenous sialoprotein and a bacterial B cell superantigen compete with their V_H family-specific binding interactions with human Igs. *J. Immunol.* 157:4496–4502.

60. Lam, K.P., R. Kuhn, and K. Rajewsky. 1997. In vivo ablation of surface immunoglobulin on mature B cells by inducible targeting results in rapid cell death. *Cell.* 90:1073–1083.

61. Dono, M., S. Hashimoto, F. Fais, V. Trejo, S.L. Allen, S.M. Lichtman, P. Schulman, V.P. Vinciguerra, B. Sellars, P.K. Gregersen, M. Ferrarini, and N. Chiorazzi. 1996. Evidence for progenitors of chronic lymphocytic leukemia B cells that undergo intraclonal differentiation and diversification. *Blood.* 87: 1586–1594.

62. Malisan, F., A-C. Fluckiger, S. Ho, C. Guret, J. Banchereau, and H. Martinez-Valdez. 1996. B-chronic lymphocytic leukemia can undergo isotype class switching in vivo and can be induced to differentiate and switch in vitro. *Blood.* 87:717–724.

63. Efremov, D.G., M. Ivanovski, F.D. Batista, G. Pozzato, and O. Burrone. 1996. IgM-producing chronic lymphocytic leukemia cells undergo immunoglobulin isotype-switching without acquiring somatic mutations. *J. Clin. Invest.* 98: 290–298.

64. Fais, F., B. Sellars, F. Ghiotto, X-J. Yan, M. Dono, S.L. Allen, D. Budman, K. Dittmar, J. Kolitz, S.M. Lichtman, et al. 1996. Examples of in vivo isotype class switching of IgM^+ chronic lymphocytic leukemia B cells. *J. Clin. Invest.* 98:1659–1666.