Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Sci Immunol. Author manuscript; available in PMC 2017 February 17.

Published in final edited form as:

Sci Immunol. 2016 July 14; 1(1): . doi:10.1126/sciimmunol.aaf6628.

VpreB serves as an invariant surrogate antigen for selecting immunoglobulin antigen-binding sites

Mohamed Khass[1,2], **Tessa Blackburn**[3], **Peter D Burrows**[3], **Mark R. Walter**[3], **Emidio Capriotti**[4,5,§], and **Harry W Schroeder Jr.**[1,3,*]

[1]Department of Medicine, University of Alabama at Birmingham, Birmingham, AL

[2]Division of Genetic Engineering, National Research Center of Egypt, Cairo, Egypt

[*]Correspondence to: Harry W Schroeder Jr., hschroeder@uabmc.edu.

[§]For information about structural modeling contact: Emidio Capriotti, emidio.capriotti@hhu.de

List of supplementary materials:

Fig. S1. Comparison of critical residues in mouse and human SLC and LC.

Fig. S2. Percentage rank order of individual amino acids by position and by reading frame in the germline sequence of the BALB/c wild-type D_H repertoire and in the D-altered ΔD-DFL and ΔD-DμFS mice.

Fig. S3. Detection of preB cells expressing the preBCR.

Fig. S4. Quantification of BrdU incorporation in developing B cells by flow cytometry.

Fig. S5. Distribution of developing B cells that have incorporated BrdU or remained quiescent in G0.

Fig. S6. Flow cytometric gates used to evaluate the cell cycle distribution of developing bone marrow B cell subsets.

Fig. S7. DH reading frame usage among in-frame heavy chains cloned from living and apoptotic preB cells.

Fig. S8. Comparison of predicted structures of preB cell receptors from apoptotic C and living C' cells.

Table S1. Statistical tests reported in upper panel of Fig. 4C

Table S2. Analysis of predicted hydrogen bond interactions in positions 101–103 for living and apoptotic cells fraction C and C'.

Table S3. Statistical tests reported in bottom panel of Fig. 4C

Table S4. Statistical tests reported in Fig. 3C

Table S5. Statistical tests reported in Fig. 3A

Table S6. Hydrogen bond interaction in the IgG Fab:antigen PDB structures in position 101–103.

Table S7. Statistical tests reported in Fig. 5C

Table S8. Hydrogen bond interaction in the IgG Fab:antigen PDB structures in position 101–103* without flanking tyrosines.

Supplemental File 1. List of the CDR-H3 sequences from WT and ΔD-DμFS cells in fractions B and D fractions used in this work.

Supplemental File 2. List of the V_H and CDR-H3 sequences from apoptotic and living C and C' fractions used in this work.

Supplemental File 3. Table of Fab: Ag crystal structures.

Supplemental File 4. Sequence of the VH regions and alignments used in the comparative modeling protocol.

Supplemental File 5. PDB files of generated mouse preBCR models. http://biofold.org/emidio/data/cdr-h3/models.tar.gz.

Supplemental File 6. Excel file containing tabulated data for Figure 1 and S5 entitled preBCR_Sup_6_Raw_Data_Fig_1_and_S_5_16-05-12.xls..

Supplemental File 7. PDF file containing the results of the statistical analysis of the data for Figure 1 and S5 entitled preBCR_Sup_7_JMP_analysis_Fig_1_and_S_5_16-05-12.

The sequences reported in the paper are available in the Supplementary Materials.

Author contributions:

Mohamed Khass designed and performed the experiments, performed the statistical analysis of the flow data, and prepared the first draft of the manuscript. Tessa Blackburn provided technical assistance in obtaining the CDR-H3 sequences. Peter Burrows provided scientific insights in preBCR biology and played a major role in the interpretation of the results and the editing of the manuscript. Mark Walter provided scientific insight and helped analyze the CDR-H3 sequences and the preBCR structures, and helped edit the manuscript. Emidio Capriotti created and designed all the bio-informatics analysis. He also performed the molecular modeling, helped analyze the IgG Fab:antigen structures, performed the statistical analysis of the models, and helped edit the paper. Harry Schroeder directed all the work, and edited the paper.

Competing interests:

The authors declare no competing interests.

Data and materials availability:

PDB files of generated mouse preBCR models are available at: http://biofold.org/emidio/data/cdr-h3/models.tar.gz.

Dr. Schroeder is the author identified to respond to readers' enquiries and requests for materials, and to coordinate the handling of any other matters arising from the published contribution, including corrections or complaints.

For information about structural modeling contact: Emidio Capriotti, emidio.capriotti@hhu.de.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[3]Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL

[4]Department of Pathology, University of Alabama at Birmingham, Birmingham, AL

[5]Institute for Mathematical Modeling of Biological Systems, Department of Biology, University of Düsseldorf, Düsseldorf, Germany

Abstract

Developmental checkpoints eliminate B cells synthesizing defective immunoglobulin heavy (HC) and light (LC) chains. The first checkpoint tests for formation of a VpreB/λ5/μHC-containing preB-cell receptor (preBCR) and predicts whether μHCs will bind conventional LCs to form membrane IgM. VpreB and λ5 also create a sensing site that interacts with μHC antigen-binding region CDR-H3, but whether it plays a role in immunoglobulin repertoire selection and function is unknown. On a position-by-position basis, we analyzed the amino acid content of CDR-H3s from H chains cloned from living and apoptotic preB cells and from IgG:Antigen structures. Using a panel of D_H gene-targeted mice, we show that progressively reducing CDR-H3 tyrosine content increasingly impairs preBCR checkpoint passage. Counting from cysteine at Framework 3 position 96, we found that VpreB particularly selects for tyrosine at CDR-H3 position 101, and that Y101 also binds antigen in IgG:Antigen structures. VpreB thus acts as an early invariant antigen. It selects for particular CDR-H3 amino acids and shapes the specificity of the IgG humoral response. This helps explain why some neutralizing antibodies against pathogens are readily produced while others are rare.

Summary

In addition to testing for the likelihood of binding to conventional light chain to create a membrane IgM, the VpreB component of the surrogate light chain serves an an invariant antigen that selects for particular amino acids at specific positions at the center of the immunoglobulin antigen binding site, thereby affecting future antigen recognition and antibody production.

Introduction

The genes that encode immunoglobulins are assembled in developing B cells by a series of gene segment rearrangement events that begin with Variable (V), Diversity (D_H), and Joining (J) gene segments at the heavy chain locus to encode a μHC. Progeny cells then rearrange a κ or λ light chain gene and the newly formed B cell expresses membrane IgM as an antigen receptor (BCR). The process of B cell development is optimized to create a highly diverse antibody repertoire. The antigen binding sites of the antibody are formed by the juxtaposition of six complementarity determining regions (CDRs), three from the heavy and three from the light chain. V and J gene segments are locked into one reading frame, but the D_H gene segments can rearrange into any one of six different reading frames (RFs), during which two rounds of non-templated N nucleotide addition can also occur. Thus, the inclusion of a D_H makes CDR-H3 the most variable component of the antigen binding site and it often plays a key role in antigen recognition.

In order to avoid production of defective H chains, developing B cells must pass through a series of quality control (QC) checkpoints that test the integrity and function of their

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

immunoglobulin (1). The first checkpoint occurs during the transition from the early (Hardy fraction C) to late (Hardy fraction D) preB cell stage (2) and tests for the ability of a nascent μHC to associate with surrogate light chain (SLC) to form a preB cell receptor (preBCR) (1). The SLC consists of two non-covalently associated proteins, the V_L homologue VpreB and the J_LC_L homologue λ5 (1). Conventional V_L contains conserved Framework Region 2 (FR2) amino acids that associate with H chain FR2 and FR4 (Fig S1) (1, 3) to form a supportive scaffold for the HCDRs. VpreB shares several of these amino acids with V_L, thus the ability of the μHC to form a preBCR predicts that it will ultimately be able to form an IgM BCR. PreB cells that fail to form a functional preBCR perish by apoptosis in the bone marrow. Unlike conventional L chains, VpreB and λ 5 are invariant, making the preBCR checkpoint quite stringent.

In addition to FR2 and FR4, VpreB associates with CDR-H3. The VpreB portion of the CDR-H3 sensing site (CDR-H3 SS) contains a set of charged or hydrophilic residues, three of which are conserved between human and mouse (R51, D57 and R101). These residues are rare or absent in conventional V_L (Fig. S1) (4). In this study, we sought to test whether the SLC could use VpreB as an invariant surrogate antigen to select for, and hence against CDR-H3s with a certain content of amino acids (5). If so, such an early selection step would limit the preimmune antibody repertoire and could help explain why some immunoglobulin antigen binding sites are rare and thus more difficult to elicit after vaccination or infection (6).

The global distribution of amino acids in the primary CDR-H3 repertoire is already known to be non-random and specifically enriched for tyrosine (7). In part, this bias reflects the nonrandom use of individual D_H reading frames (8), each of which exhibits a characteristic amino acid signature. In the preimmune repertoire, RF1 is the most frequently utilized, RF2 and RF3 are used less frequently, and the inverted (i) RFs are rarely used (9). RF1 is enriched for tyrosine (Fig. S2) (7), RF2, RF3, iRF2 and iRF3 are enriched for hydrophobic amino acids, and iRF1 is enriched for charged amino acids. RF3 tends to contain termination codons, limiting its use and thus the presence of leucine in CDR-H3.

The reduced use of RF2 encoded amino acids in mice has been attributed to the presence of an upstream ATG codon in-frame with RF2 as a translation start site (8), which results in production of a truncated protein product of DJ rearrangement termed Dμ. Although lacking the V_H component of the H chain, Dμ can associate with the SLC to create a ‘sterile’ preBCR that prevents further development. However, this blockade is not absolute and one-fifth of in-frame VDJ rearrangements in early preB cells use RF2 (10, 11).

Four RF1 CDR-H3 tyrosines (Y101-Y104), each of which can interact with CDR-H3 SS amino acids by means of hydrogen bonds, potentially stabilizing the complex, are observed in the only published preBCR crystal structure (PDB: 2H32) (4). Since RF2 does not encode such tyrosines, we postulated that the halving of RF2 usage that marks passage through the preBCR checkpoint (10, 11) might reflect reduced stability of preBCRs containing RF2 encoded amino acids in CDR-H3. Support for this view comes from studies in human where Dμ protein is not made, yet RF2 is still underrepresented in CDR-H3 (7); and from

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

enrichment for CDR-H3s with RF2 encoded amino acids among immunoglobulins from a λ5, and thus preBCR-deficient, patient (12).

To test whether amino acids encoded by RF2 are subject to selection by the preBCR checkpoint, we took advantage of our panel of BALB/c mice with homozygous gene-targeted D_H alleles (13). The ΔD-DFL D_H allele contains a single wild-type DFL16.1 gene segment, which is the most frequently used D_H. The ΔD-DμFS D_H allele contains a single, frameshifted DFL16.1 that favors the use of RF2 since the ATG Dμ start codon is in frame with RF1 and RF3 still contains a stop codon (Fig. S2). Twenty percent of mature B cells in wild-type mice use RF2, while 30% of ΔD-DFL and 60% of ΔD-DμFS B cells use this reading frame (14). Thus, B cells in our panel create Ig CDR-H3s with a progressive increase in use of RF2 (13), which allowed us to test whether increased use of RF2 encoded amino acids influenced successful passage through the preBCR checkpoint.

Results

Increasing the use of RF2 encoded amino acids reduces preBCR expression

We used the scheme of Hardy and Hayakawa (2) to identify progenitor B cell subsets in bone marrow. Fraction B contains proB cells that express SLC proteins and are undergoing H chain rearrangement; Fraction C contains early preB cells that express μHC and SLC; Fraction C' contains larger early preB cells where successful expression of preBCR is associated with cell division; and Fraction D contains late preB cells that no longer express SLC and are undergoing L chain rearrangement.

We examined preBCR formation, cell cycling, and apoptosis in fractions B through D from WT, ΔD-DFL, ΔD-DμFS, and λ5-deficient (λ5KO) mice, using the latter as a control for the absence of a preBCR. Among these mice, increased incorporation of RF2-encoded amino acids was associated with diminished preBCR formation, reduced cell division, and enhanced apoptosis; each a hallmark of failure to pass the preBCR QC checkpoint.

To assess preBCR expression, we used μHC-specific polyclonal antibodies and a monoclonal antibody that only recognizes a completely formed preBCR and not its separate components (15). Because cell surface preBCR expression is limited, intracellular staining techniques were used (Fig. S3). When compared to WT mice (n=8), there was a major decrease in the frequency of preBCR expressing preB cells among ΔD-DμFS fractions C and C' ($p=0.003$ and $p<0.0001$, n=8, t test) (Fig. 1A). The ΔD-DFL preB cells were intermediate between WT and ΔD-DμFS ($p=0.03$ and $p=0.0003$, n=8, t test). in Fraction C and C', respectively), which correlated with the more frequent use of RF2 with this gene segment when compared to the other D_H (14).

Increasing the use of RF2 encoded amino acids decreases preB cell proliferation

When compared to wild-type littermate controls, we previously found that while ΔD-DFL and ΔD-DμFS mice had normal absolute numbers of early preB cells [Fraction C (C+C')], there was a 17% and 40% reduction in the absolute number of late preB cells, respectively (13). To assess cell division, we evaluated BrdU incorporation after intraperitoneal injection (Fig. S4). Increased use of RF2 correlated with a decrease in BrdU incorporation (Fig. S5A).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In all four Hardy fractions studied, WT mice incorporated the most BrdU. ΔD-DFL cells were similar to WT or intermediate between WT and ΔD-DμFS, depending on the fraction analyzed. Cell division was significantly reduced among ΔD-DμFS B lineage cells. Indeed, BrdU incorporation among Fractions B, C and C' cells was similar to that observed in λ5-deficient cells. The $BrdU^+$ cells in Fraction D likely represent preB cells that proliferated when they were passing through Fraction C', since pre-B cells can progress through the small pre-B cell stage (from large cycling pre-B cells to immature B cells) in approximately 5 hours and the mice in our studies were exposed to BrdU for 12 hours (16).

To refine the analysis of cell division, we used 7AAD and BrdU staining to examine cell cycling (Fig S6). The hierarchy of cells in S-phase (Fig. 1B) correlated with total BrdU incorporation (Fig. S5A); more WT than ΔD-DμFS or λ5KO preB cells were replicating. The average frequency of replicating ΔD-DFL preB cells tracked with WT but individual mice demonstrated a greater variance than those from the other three strains. The hierarchy of G0 quiescent cells was complementary. ΔD-DμFS and λ5KO preB cells were more quiescent than WT, and ΔD-DFL preB cells were either similar to WT or intermediate between WT and ΔD-DμFS (Fig. S5B).

Increasing the use of RF2 encoded amino acids increases preB cell apoptosis

Since apoptosis is a mark of failure to pass the preBCR QC checkpoint, we used staining with Annexin V to quantify this process in the different mouse strains. Among proB cells, which have not yet reached the preBCR checkpoint, the percentage of Annexin V^+ cells was similar (Fig. 1C). Among both early and late preB cells, however, more ΔD-DμFS than wild-type cells were undergoing apoptosis. For ΔD-DFL, apoptosis was similar to wild-type in early preB cells and then increased among late preB cells, lying intermediate between wild-type and ΔD-DμFS cells. This suggests that the use of RF2 encoded amino acids in CDR-H3 is disfavored for preBCR formation and stability.

Selection for CDR-H3 Y101in mice forced to preferentially rearrange into RF2

A key prediction of our hypothesis is that passage through the preBCR checkpoint would be associated with changes in the prevalence of particular amino acids within CDR-H3. CDR-H3 can be divided into a base and a loop, the latter encoded by the D, by N nucleotides, and by the amino terminus of the J. It rests on the surface of the V domain and can assume multiple conformations. The carboxyl terminus of the loop is enriched for J-encoded amino acids and exhibits less sequence diversity. Thus, we focused on the amino acids in the N terminal end of CDR-H3, which is primarily encoded by N addition and D_H.

Counting from cysteine at position 96 [Protein Data Bank (PDB) numbering], we examined amino acids 99 to 103 among a large database of previously published sequences from wild-type and D-altered mice (13) (Supplemental File 1). Even though CDR-H3 is of variable length, position 101 tends to be centrally located in mouse immunoglobulins. Approximately two-thirds of the amino acids at 101 were D-derived, one quarter created by N addition, and one-tenth were J-encoded.

To assess the contribution of individual amino acids at positions 99–103 pre- and post-preBCR QC selection, we subtracted the fraction of usage per amino acid in fraction B

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(proB) from fraction D (late preB) cells (Fig. 2A). No major fractional changes in positional amino acid distribution were observed among WT (Fig. 2A). Thus, CDR-H3 amino acid prevalence in wild-type B cells passing though the preBCR QC (Fig. 2B) appeared to maintain the amino acid preferences initially created by preferential use of D_H RF1 (Fig. S2).

Unlike WT, in ΔD-DμFS a number of striking differences were observed after preBCR QC (Fig. 2C and D). These included selection for tyrosine (RF1) at 101 and 102 and against threonine (RF2) and leucine (RF3) at 102. We also observed a decreasing fraction of phenylalanine in positions 101 and 102. This suggested that preBCR selection can be influenced by specific amino acid usage at particular CDR-H3 positions.

RF2 and RF3 encoded hydrophobic amino acids at CDR-H3 position 101 predispose to apoptosis

To better reveal the CDR-H3 molecular determinants for preBCR selection in WT, we pooled bone marrow from twenty 8 week old BALB/c mice, stained cells from Hardy Fractions C, C', and D with Annexin V and PI (Fig. S7A), isolated the double positive, late apoptotic cell fraction, and used RT-PCR to clone and sequence CDR-H3s in cells that had failed preBCR QC (Supplemental File 2). We compared this set of CDR-H3s to those previously isolated from live cells of the same fractions (17) and found that apoptotic cells were more likely to use amino acids encoded by RF2 or RF3 (Fig. S7B).

A comparison of CDR-H3 sequences in living and apoptotic Fraction C early preB cells revealed that apoptotic cells exhibited increased use of glycine (RF1) at position 100. The presence of glycine precludes the formation of side chain interactions. Living cells demonstrated increased use of tyrosine and glycine (RF1) at 101; decreased use of leucine (RF3) at 101; and decreased threonine (RF2) at 102 and 103 (Fig. 3A, D). Tyrosine provides a bulky, flexible side chain that permits changes in conformation as well as a terminal hydroxyl group that permits hydrogen bonding. Threonine has a hydroxyl group, but lacks the bulky side chain. Leucine is hydrophobic and does not engage in hydrogen bonding with side-chain atoms. When compared to Fraction C' apoptotic cells, living Fraction C' cells exhibited an increase in tyrosine (RF1) and a decrease in glycine (RF1) and leucine (RF3) at position 100; and an increase in tyrosine (RF1) alone at 101 (Fig. 3B, D). These differences were magnified when compared to Fraction C apoptotic cells (Fig. 3C, D). These findings supported the hypothesis that amino acids encoded by RF2 and RF3 were disfavored and amino acids encoded by RF1, especially tyrosine, were favored for successful preBCR formation.

CDR-H3 Y101 has a beneficial effect on passage through the preBCR quality control checkpoint

To gain further insight into the mechanism by which highly prevalent amino acids at specific CDR-H3 positions might be selected, we examined the crystal structure of the preB receptor (PDB:2H32, Fig. 4A). This is a human preBCR structure and contains serine at position 59. We modeled the murine preBCR CDR-H3 SS by replacing human S59 with murine H59. CDR-H3 in this structure contains tyrosines at positions 100 through 103, all of which

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

interact with the VpreB CDR-H3 SS. Of these, Y101 forms the most extensive hydrogen bond network and van der Waals contacts, with its hydroxyl group positioned for interactions with VpreB R51, D57, and R101. The side chains of the three conserved SLC CDR-H3 SS residues (Fig. S1) form a pocket that can bind tyrosine at CDR-H3 position 101 (Fig. 4A), and the main chain hydroxyl group of VpreB R101 can form a hydrogen bond with CDR-H3 Y103. Finally, analysis of the original human pre-B cell receptor complex and the modeled murine CDR-H3 SS predicts that CDR-H3 Y102 may form a hydrogen bond with mouse VpreB H59.

Our experimental data support a role for invariant VpreB in selecting for μHC that contain a tyrosine at position 101. To test the structural and functional consequences of this selection, we generated and then analyzed molecular models of preBCR structures that used CDR-H3s cloned from the living and apoptotic preB cells (Fig. 4B,C, D). We estimated the contribution of tyrosine at CDR-H3 positions 101, 102 and 103 to preBCR formation in sequences from living C and C' preB cells that either lacked or contained tyrosine (Fig. 4C, top panel). We compared the probability of interaction (fraction of models with at least one hydrogen bond) with the VpreB CDR-H3 SS (see Supplementary Materials). To avoid the potential confounding factor of runs of tyrosines, we focused this site-specific analysis on those sequences with tyrosine at the position of interest that lacked flanking tyrosines (1 left, 1 right). At positions 101 and 103, the distribution of the probability of interaction with the CDR-H3 SS when tyrosine is present (black) or not (white) achieved significance ($p < 0.01$ and $p = 0.03$, respectively; Kolmogorov-Smirnov; Table S1) (Fig. 4C, top panel).

We then grouped the sequences from WT living and apoptotic C and C' preB cells into those with either < 5% or ≥ 5% of the models predicting one or more hydrogen bonds between amino acids 101–103 and the CDR-H3 SS (Table S2). CDR-H3 position 103 showed the highest fraction of interacting sequences (Fig. 4C, top panel), possibly due its ability to form a hydrogen bond with the main chain of VpreB R101 (Fig. 4A). However, the most significant differences between the modeled preBCR structures from apoptotic C and those from living C and C' sequences were at position 101 ($p= 0.04$ and 0.01, respectively; one-tail Fisher's exact test; Table S3) (Fig. 4C, bottom panel). Thus, while the presence of a tyrosine at either 101 or 103 significantly increases the probability of forming hydrogen bonds with VpreB, the major differences between the apoptotic and living cells reflected the interaction of the CDR-H3 SS with CDR-H3 position 101.

The importance of position 101 is illustrated in a comparison of the models of a living C' sequence that uses the same V_H7183-10 gene segment as an apoptotic C sequence (Fig. 4D). The CDR-H3s from these sequences differ in length by only one amino acid. When superimposed, their structures show a Root Mean Square Deviation (RMSD) of only 1.5 Å for the modeled structures as a whole, and 3.5 Å for the CDR-H3 region (Fig. S8). CDR-H3 Y101 in the modeled living cell sequence is adjacent to VpreB R51, D57, and R101; and Y102 is adjacent to VpreB H59 (Fig. 4B, D). Conversely, CDR-H3 L101 in the modelled apoptotic cell sequence is distant from VpreB even though CDR-H3 R102 is adjacent to VpreB R101 (Fig. 4D).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Further support for the beneficial effect of Y101 on passage through the preBCR QC derives from the distribution of amino acids at CDR-H3 101 in the sequences from living versus the apoptotic cells (Fig. 3). Y101 proved more common in living C' sequences than in apoptotic C ($p=0.005$; one-tail Fisher's exact test; Table S4) (Fig. 3C). Conversely, L101, which is frequently encoded by RF3, is more common in apoptotic C than in living C ($p= 0.02$, one-tail Fisher's exact test; Table S5) (Fig. 3A) [Details regarding the comparative modeling protocol used for predicting the structures and their structural comparisons are reported in the supplementary material.]

In IgG Fab:antigen crystal structures, CDR-H3 Y101 often interacts with antigen

Having shown that CDR-H3 tyrosines, in particular Y101 are selected for and interact with VpreB during early B cell development, we next investigated the role of these tyrosines in the interaction of the Fab with antigen, using the IMGT database to identify a set of 133 non-redundant murine antigen:Fab crystal structures in the Protein Data Bank (Supplemental File 3) (http://www.pdb.org). A comparison of the distribution of amino acids at positions 99 through 103 revealed that G99 was preferred in the IgG Fabs rather than H99, which was more prevalent in the living progenitor B cells (Figs. 2B, 3D, and 5A). In positions 100–103, however, both the preimmune living progenitor B cells and the antigen selected IgGs demonstrated the same preference for G100, Y101, Y102 and Y103 (Fig. 5A).

We examined the IgG:antigen crystal structures to determine which CDR-H3 residues used a hydrogen bond to engage with antigen (Fig 5B) or L chain. We found that the 101–103 tripeptide (101,102,103) was more likely to engage antigen (65% of cases) than L chain (42%) (Table S6). As above, we then considered only the structures with a tyrosine at 101, 102 or 103 that lacked flanking tyrosines. At positions 101 and 103, structures that use tyrosine (black bar) were significantly more likely to engage antigen with a hydrogen bond than structures with residues other than tyrosine (white bar) ($p = 0.008$, one-tail Fisher's exact test; Table S7) (Fig. 5C). No significant preferences for specific amino acid composition were found for structures binding L chain, which is more likely to engage CDR-H3 at position 103 than position 101 (Table S8).

Discussion

μHCs in late preB cells have access to an array of diverse L chains. Thus, the only time during B cell development when the L chain partner is invariant occurs at the preBCR QC checkpoint. Here we provide evidence that this checkpoint not only selects for cells expressing μHC able to bind to a conventional L chain, but also selects for the use of particular amino acids at specific positions at the center of the antigen binding site, where they can subsequently influence antigen recognition and antibody production. This selection process appears to begin just after the μHC is created, in the endoplasmic reticulum (ER) of the early preB cell. Presence of a CDR-H3 that cannot bind effectively to the CDR-H3 SS of the SLC would result in retention of the μHC by ER chaperones and failure to form a preBCR, resulting in decreased cell proliferation and increased cell loss by apoptosis.

Although not identical, the level of sequence similarity (>70%) between human and mouse VpreB suggests that the selection mechanisms described in mouse may also be used in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

human to regulate CDR-H3 amino acid content and reading frame usage (4). We have shown that changing of the pattern of D_H reading frame usage, and thus the prevalence of tyrosine in CDR-H3, in mice alters patterns of epitope recognition after vaccination with HIV gp140 (6). It is possible that the binding instability of SLC to μHCs containing CDR-H3s with amino acids commonly encoded by non-tyrosine enriched reading frames may help explain the difficulty experienced by HIV patients in generating neutralizing antibodies that require CDR-H3 amino acids typically encoded by RF2 or RF3.

We have previously shown that natural selection of germline D_H sequence creates a bias in global CDR-H3 amino acid content that is established at the time of VDJ rearrangement prior to antigen exposure (17). The presence of other tyrosines around position 101 and their possible interactions with VpreB support the hypothesis that the binding of CDR-H3 to the SLC is a stochastic process in which specific amino acids increase the chance of forming hydrogen bonds that stabilizes the preBCR.

Although among individual structures the likelihood of interactions between the antigen and individual CDR-H3 residues can be influenced by the length of the CDR-H3 loop, our observations support the global hypothesis that CDR-H3 residue 101 and 103 often play an important role, especially in the presence of tyrosine, in binding both the VpreB in the preBCR (Fig. 4D) and in binding the antigen in the IgG:antigen complex (Fig. 5C). High throughput sequencing of H chains from developing preB cells from the WT and D-altered mice would enable enhanced comparisons of CDR-H3s of similar length.

Both natural selection of D_H germline sequence and passage through the preBCR checkpoint promote the inclusion of a tyrosine at CDR-H3 position 101. However, our results do not exclude alternative interaction patterns in which amino acids with similar chemico-physical properties can form stable preBCR. In our view, the CDR-H3 regions from apoptotic cells correspond to amino acids combinations with a low probability of forming hydrogen bonds with VpreB R51/D57/R101 component of the CDR-H3 sensing site and thus creating a functional preBCR with an H chain antigen binding site optimized for binding antigen. An analysis of repertoire selection and responses to antigenic challenges in mice where sequence of the CDR-H3 SS has been mutated would provide confirmation of this interpretation.

In summary, the sequence of the D_H has a positional effect, selecting for particular amino acids at specific locations within the antigen binding site (Fig 2) (6). In this work, we show that selection of CDR-H3 by the VpreB portion of SLC reinforces the positional bias in amino acid content, which is consistent with the bias in antigen-binding IgG. Our findings suggest that the molecular mechanism underlying the selection of CDR-H3 Y101 is the stabilization of the preBCR by the creation of hydrogen bonds between Y101 and VpreB R51/D57/R101. Analysis of the crystal structures of IgG Fab:antigen complexes indicates that the CDR-H3 residues selected by VpreB are ultimately more likely to form hydrogen bonds with antigen than L chain. The combination of natural selection of D_H sequence and somatic selection by invariant surrogate light chain favors the presence of tyrosine at key positions in the antigen binding site, thus shaping the antigen specificity of the primary (IgM) and secondary (IgG) humoral response.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Materials and Methods

Study design

We sought to determine whether preBCR expression, progenitor B cell division, and progenitor B cell apoptosis was influenced by the amino acid sequence of immunoglobulin CDR-H3 during B cell development by analyzing a panel of BALB/c mouse strains that had been gene targeted in their D_H locus and thus expressed a gradient of altered amino acid sequence. We then performed a combined sequence structural analysis of H chains cloned from living and apoptotic progenitor B cells to test for positional patterns of amino acid preference and compared the results to IgG Fab:antigen structures that were publically available.

All data obtained from the mice is presented, including outliers. We used six or more mice per group. The genotypes were not blinded and randomization was not performed. All of the modeled H chains cloned from early preB cells were unique. To select a representative set of structures we queried the IMGT database (http://www.imgt.org/) searching for *Mus musculus* Fab with protein antigen and resolution ≤3.5 Å. From this initial group of structures we selected the PDBs with heavy chains sequence similar to our set of VH 7183 family members. The sequence similarity was detected using the BLAST algorithm (18) using an e-value threshold of 10–3. Redundant PDBs were removed selecting only the representative structures of each cluster defined using a 95% sequence similarity threshold (ftp://resources.rcsb.org/sequence/clusters/clusters95.txt). The final set of 133 Fab:antigen PDB structures is provided in Supplemental File 3.

Mice

All experiments and animal procedures were performed using protocols approved by the University of Alabama Institutional Animal Care and Use Committee (IACUC). BALB/c mice were bred under specific pathogen free conditions in the UAB vivarium. The ΔD-DFL and ΔD-DμFS D-altered mice were generated on a BALB/c background as described previously (14, 19). λ5KO mice, originally on a C57BL/6 background (20), were backcrossed for 22 generations onto BALB/c.

Flow cytometry analysis of bone marrow B cell subsets

Single cell suspensions were prepared from femurs of 8 week old mice as previously described (7). Cells were washed, stained, and analyzed on an LSRII (BD Biosciences) using the following labeled antibodies: anti-B220 (PB), anti-CD19 (APC-Cy7), anti-CD43 (Biotin), anti-BP-1 (PE), and anti-IgM (APC) from BD Pharmingen; anti-AA4.1 (PE-Cy7) from eBioscience; and anti-IgD (Biotin) from Southern Biotech. Propidium Iodide (PI) was used to identify dead cells. The scheme of Hardy (2) was used to identify different B cell subsets in the BM. Fractions C and C' were distinguished based on cell size.

Intracellular evaluation of binding of μHC to SLC

PreB cell subsets in the BM, as identified by the expression of relevant surface markers, were fixed using a Cytofix/Cytoperrm kit [BD Pharmingen] and then stained for intracellular preBCR using a pre BCR mAb (clone SL156) [BD Pharmingen] and goat anti-mouse μHC

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(FITC) (Southern Biotech) (Fig 1A). Successful formation of preBCR was defined by dual staining with both antibodies, as described by Kawano et al. (15) (Fig. S3).

Bromodeoxyuridine labeling

Bromodeoxyuridine (BrdU) incorporation was used to assess B cell proliferation. BrdU was administered by intraperitoneal injection (1 mg/mouse) twice daily at 12 hour intervals. Twelve hours after the last injection, mice were sacrificed, BM cells were washed, fixed and BrdU positive B cells were detected with antibodies specific for BrdU using the BrdU-flow kit [BD Biosciences] according to the manufacturer's protocol (Fig. S4).

Cell cycle analysis

Bone marrow B cell subsets were analyzed for cell cycle progression using BrdU and 7AAD staining according to the BrdU-flow kit [BD Biosciences]. Cell cycle stages were identified as: G0-phase, quiescent cells (BrdU negative and 7AAD negative), G1-phase, cells preparing for cycling (BrdU positive and 7AAD negative), S-phase, proliferating cells (BrdU positive and 7AAD positive) and G2-M-phase cells in process of finishing cycling (BrdU intermediate and 7AAD positive) (Figs 1B, S4, 5 & 6)

Evaluating the extent of apoptosis in different B cell subsets

The frequency of apoptosis in different B cell subsets in the BM was determined based on Propidium Iodide (PI) and Annexin V staining using the FITC Annexin V Apoptosis Detection Kit II [BD Biosciences]. Late apoptotic cells that express both Annexin V and PI were sorted for CDR-H3 cDNA construction, sequencing (as described below) and structural modeling as they represent a true indicator of fully apoptotic cells (Figs. 1C & S7).

Flow cytometric sorting, RNA preparation, RT-PCR and sequencing

Living and late apoptotic preB cells were sorted based on expression of surface markers (2) using the FACSAria cell sorter [BD Biosciences]. Total RNA was isolated and converted to cDNA using V_H7183 specific VDJCμ RT-PCR primers. Clones positive for an insert of the appropriate size were sequenced and analyzed as previously described (14, 21, 22). A list of the unique, in-frame sequences used for analysis in this work is provided in Supplemental File 2.

Dataset for structural modeling

Total living and apoptotic cells were collected without bias. Sequence/structure-function relationships were determined using potential differences in the predicted structures obtained from living and apoptotic cells. We examined 31 and 39 unique sequences from apoptotic C and C', and 83 and 36 sequences from living C and C', respectively. The 189 CDR-H3 sequences are reported in Supplemental File 2.

Sequence analysis of the CDR-H3 regions

CDR-H3 was defined as the sequence starting immediately after the cysteine at the end of FR3 at position 96 and then through, but not including, the tryptophan that begins FR4. At positions 99 through 103, we calculated the frequency of the amino acids in individual

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

developing B cell subsets. The residues are grouped in three main categories according to their hydrophobicity as assessed in the Kyte-Doolittle scale as modified by Eisenberg (23, 24). Under this scheme arginine, lysine, asparagine, aspartic acid, glutamine, glutamic acid and histidine are considered charged and shown in red; proline, tryptophan, serine, threonine and glycine are considered neutral or hydrophilic and shown in green; and alanine, methionine, cysteine, phenylalanine, leucine, valine, and isoleucine are considered hydrophobic and shown in blue. Tyrosine, which is the most abundant residue in our dataset (~21%), is normally included among the hydrophilic, neutral fraction. For this work, tyrosine was counted separately and is shown in black for emphasis. The amino acid composition in the different positions was displayed using WebLogo (http://weblogo.berkeley.edu/) (25).

Structural analysis of the predicted preB cell receptor models

To study the structural and functional features of the CDR-H3 region, we predicted the structure of the mouse pre-B cell receptor using homology modeling (26) using the structure of the human pre-B cell receptor (PDB: 2H32) as a template (4). The structure of this protein complex is a trimer containing VpreB, λ5 and immunoglobulin μHC.

All the murine preBCR models were then generated using the sequences of the mouse Vpre1 (SwissProt: VPRE1_MOUSE) and λ5 (SwissProt: IGLL1_MOUSE). Each model of the pre-B cell receptor complex differs in the sequence of the Ig heavy chain, which is assembled using the different V_H 7183 family members with unique CDR-H3s. Global alignments were calculated using *align* (27) using default parameters without end gap penalties. The sequence identities between human and mouse sequences are 77% for VpreB, 65% for λ5 and in the range of 40–58% for μHC depending on the specific V_H183 family member (see alignments in Supplemental File 4).

Mouse pre-B protein structures were predicted using MODELLER (28). We generated 50 models for each sequence in our dataset. These were used to predict the interactions between the CDR-H3 regions, VpreB and μ5 in the pre-B cell receptor complex. An example of a comparison between two similar predicted structures of μHC sequences from living C' (cyan) and apoptotic C (magenta) cells is shown in Fig. 4B,D and Fig. S8. We analyzed the region surrounding CDR-H3 residues 101 through 103 and their possible interaction with VpreB residues. A hydrogen bond was predicted considering a distance threshold of 3.5 Å between two electronegative atoms (N, O). For each CDR-H3 sequence we calculated the probability of interacting models (IM) in a given position as the fraction of models in which a hydrogen bond was detected. To evaluate the importance of tyrosine in residues 101–103 and to reduce the bias of consecutive tryrosines, we evaluated a subset of sequences with and without tyrosine in the central position without flanking tyrosines. The analysis of this subset of sequences from living C and C' cells is reported in the **top panel** of Fig. 4C. We then focused only on position 101 and considered the fraction of sequences forming a hydrogen bond (HB) with CDR-H3 SS in at least 5% of the models (blue bar, HB≥5%) compared with the fraction of sequences that engaged in hydrogen bond in less than 5% of the models (red bar, HB<5%), across living and apoptotic fraction C and C' sequences (Fig. 4C, bottom panel). The significance in the difference of the predicted hydrogen bonds in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

position 101 was estimated by one-tail Fisher's exact test. All the PDB files of the predicted models are available in Supplemental File 5.

Structural analysis of the Fab:antigen PDB structures

Finally, we performed an analysis of murine Fab:antigen PDB structures. The structures were analyzed calculating the hydrogen bonds using a distance threshold of 3.5 Å between two electronegative atoms (N, O). To estimate the importance of tyrosine in the formation of hydrogen bonds in positions 101–103 we compared the fraction interacting and non-interacting structures with and without tyrosine in a given position. The bias for consecutive tyrosines was reduced considering the subset of sequences without flanking tyrosines around the central position. The result of this analysis is reported in Fig. 5C.

Statistical analyses

Differences between populations or groups were assessed by Student's t-test, one- tailed; Fisher's exact test, two tailed; the Levene test for the homogeneity of variance; the Kolmogorov-Smirnov test, or the non-parametric Mann-Whitney test. A p-value (p) of less than 0.05 was considered significant. JMP version 12 (SAS Institute Inc.) was used for analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank S. Louis Bridges, Harry W Schroeder III, and Robert R. Rich for helpful discussions. We thank Leticia Watkins, Cun Ren Liu and Yingxin Zhuang for expert technical assistance.

Funding:

The present study was supported by the US National Institutes of Health (NIH) AI048115, AI049342, AI090902, AI090742, AI097629, and AI117703. Flow cytometry was supported by NIH AR48311 and AI027767. Modeling studies were supported by funds provided by the Department of Pathology, University of Alabama at Birmingham and computational resources of the Geneifx - Genome Informatics Service.

References

1. Melchers F. Checkpoints that control B cell development. J Clin Invest. 2015:1–8. [PubMed: 25654544]
2. Hardy RR, Hayakawa K. B cell development pathways. Ann Rev Immunol. 2001; 19:595–621. [PubMed: 11244048]
3. Narciso JE, Uy ID, Cabang AB, Chavez JF, Pablo JL, Padilla-Concepcion GP, Padlan EA. Analysis of the antibody structure based on high-resolution crystallographic studies. Nat Biotechnol. 2011; 28:435–447.
4. Bankovich AJ, Raunser S, Juo ZS, Walz T, Davis MM, Garcia KC. Structural insight into pre-B cell receptor function. Science. 2007; 316:291–294. [PubMed: 17431183]
5. Kim ST, Shirai H, Nakajima N, Higo J, Nakamura H. Enhanced conformational diversity search of CDR-H3 in antibodies: role of the first CDR-H3 residue. Proteins. 1999; 37:683–696. [PubMed: 10651282]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

6. YK Wang P, Parks R, Silva-Sanchez A, Alam SM, Verkoczy L, Liao HX, Zhuang Y, Burrows PD, Levinson M, Elgavish A, Cui X, Haynes BF, Schroeder HW Jr. HIV-1 gp140 epitope recognition is influenced by immunoglobulin DH gene segment sequence. Immunogenet. 2016; 68:145–155.
7. Ivanov, II., Link, JM., Ippolito, GC., Schroeder, HW, Jr. The Antibodies. Zanetti, M., Capra, JD., editors. Vol. 7. London: Taylor and Francis Group; 2002. p. 43-67.
8. Gu H, Kitamura D, Rajewsky K. DH reading frame bias: evolutionary selection, antigen selection or both? Evolutionary selection. Immunol Today. 1991; 12:420–421. [PubMed: 1786075]
9. AL Silva-Sanchez CR, Vale AM, Khass M, Kapoor P, Elgavish A, Ivanov II, Ippolito GC, Schoeb T, Burrows PD, Schroeder HW Jr. 2015
10. Zemlin M, Schelonka RL, Ippolito GC, Zemlin C, Zhuang Y, Gartland GL, Nitschke L, Pelkonen J, Rajewsky K, Schroeder HW Jr. Regulation of repertoire development through genetic control of D H reading frame preference. J Immunol. 2008; 181:8416–8424. [PubMed: 19050259]
11. Khass M, Buckley K, Kapoor P, Schelonka RL, Watkins LS, Zhuang Y, Schroeder HW Jr. Recirculating bone marrow B cells in C57BL/6 mice are more tolerant of highly hydrophobic and highly charged CDR-H3s than those in BALB/c mice. Eur J Immunol. 2013; 43:629–640. [PubMed: 23225217]
12. Minegishi Y, Conley ME. Negative selection at the pre-BCR checkpoint elicited by human mu heavy chains with unusual CDR3 regions. Immunity. 2001; 14:631–641. [PubMed: 11371364]
13. Schroeder HW Jr, Zemlin M, Khass M, Nguyen HH, Schelonka RL. Genetic control of DH reading frame and its effect on B-cell development and antigen-specifc antibody production. Crit Rev Immunol. 2010; 30:327–344. [PubMed: 20666706]
14. Schelonka RL, Ivanov II, Jung D, Ippolito GC, Nitschke L, Zhuang Y, Gartland GL, Pelkonen J, Alt FW, Rajewsky K, Schroeder HW Jr. A single DH gene segment is sufficient for B cell development and immune function. J Immunol. 2005; 175:6624–6632. [PubMed: 16272317]
15. Kawano Y, Yoshikawa S, Minegishi Y, Karasuyama H. Selection of stereotyped VH81X-{micro}H chains via pre-B cell receptor early in ontogeny and their conservation in adults by marginal zone B cells. Int Immunol. 2005; 17:857–867. [PubMed: 15908445]
16. Casellas R, Shih TA, Kleinewietfeld M, Rakonjac J, Nemazee D, Rajewsky K, Nussenzweig MC. Contribution of receptor editing to the antibody repertoire. Science. 2001; 291:1541–1544. [PubMed: 11222858]
17. Ivanov II, Schelonka RL, Zhuang Y, Gartland GL, Zemlin M, Schroeder HW Jr. Development of the expressed immunoglobulin CDR-H3 repertoire is marked by focusing of constraints in length, amino acid utilization, and charge that are first established in early B cell progenitors. J Immunol. 2005; 174:7773–7780. [PubMed: 15944280]
18. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25:3389–3402. [PubMed: 9254694]
19. Schelonka RL, Zemlin M, Kobayashi R, Szalai A, Ippolito GC, Zhuang Y, Gartland GL, Fujihashi K, Rajewsky K, Schroeder HW Jr. Preferential use of D H reading frame 2 alters B cell development and antigen-specific antibody production. J Immunol. 2008; 181:8409–8415. [PubMed: 19050258]
20. Kitamura D, Kudo A, Schaal S, Muller W, Melchers F, Rajewsky K. A critical role of lambda 5 protein in B cell development. Cell. 1992; 69:823–831. [PubMed: 1591779]
21. Ivanov II, Schelonka RL, Zhuang Y, Gartland GL, Zemlin M, Schroeder HW Jr. Development of the expressed Ig CDR-H3 repertoire is marked by focusing of constraints in length, amino acid use, and charge that are first established in early B cell progenitors. J Immunol. 2005; 174:7773–7780. [PubMed: 15944280]
22. Ippolito GC, Schelonka RL, Zemlin M, Ivanov II, Kobayashi R, Zemlin C, Gartland GL, Nitschke L, Pelkonen J, Fujihashi K, Rajewsky K, Schroeder HW Jr. Forced usage of positively charged amino acids in immunoglobulin CDR-H3 impairs B cell development and antibody production. J Exp Med. 2006; 203:1567–1578. [PubMed: 16754718]
23. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982; 157:105–132. [PubMed: 7108955]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

24. Eisenberg D. Three-dimensional structure of membrane and surface proteins. Annu Rev Biochem. 1984; 53:595–623. [PubMed: 6383201]

25. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004; 14:1188–1190. [PubMed: 15173120]

26. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct. 2000; 29:291–325. [PubMed: 10940251]

27. Huang X, Miller M. A time-efficient, linear-space local similarity algorithm. Adv. Appl. Math. 1991; 12:337–357.

28. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using Modeller. Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]. 2006 Chapter 5, Unit 5 6.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

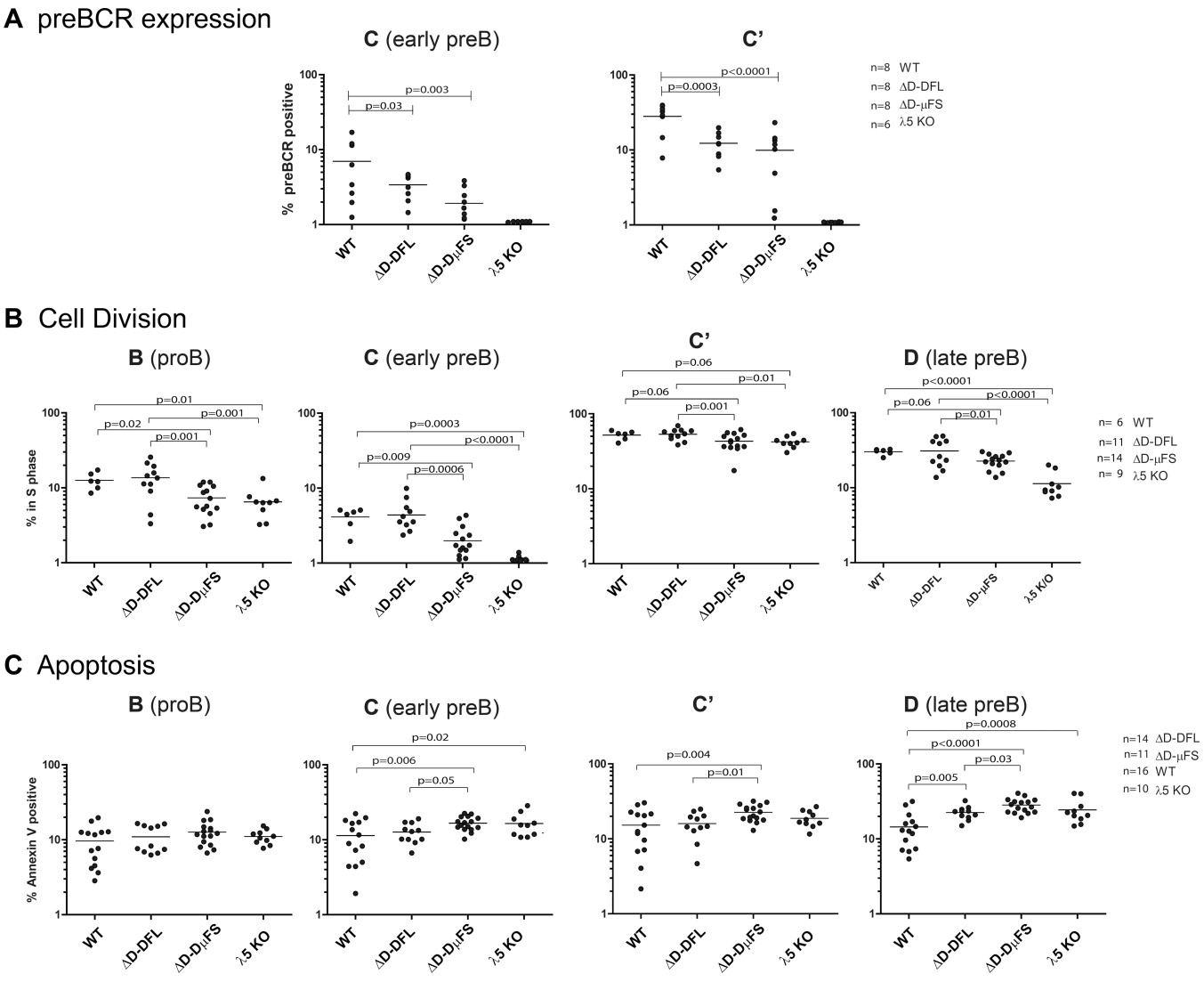

Fig. 1. Increasing the use of RF2-encoded amino acids in CDR-H3 diminishes preBCR formation, decreases cell cycling, and enhances apoptosis

Among progenitor B lineage cells: **(A)** Percent of preBCR positive cells. The cells were intracellularly stained with μ HC-specific polyclonal antibodies and with the SL156 monoclonal preBCR-specific antibody that detects a conformational epitope on the complete preBCR. Cells stained with both antibodies were counted as pre BCR^+. **(B)** Percent of cells in S-phase as determined by BrdU incorporation. **(C)** Percent of cells undergoing apoptosis as determined by Annexin V staining. Each dot indicates a measurement from an individual mouse. Dots are grouped by mouse strain (WT, ΔD-DFL, ΔD-DμFS, and λ5) and B cell developmental stage.

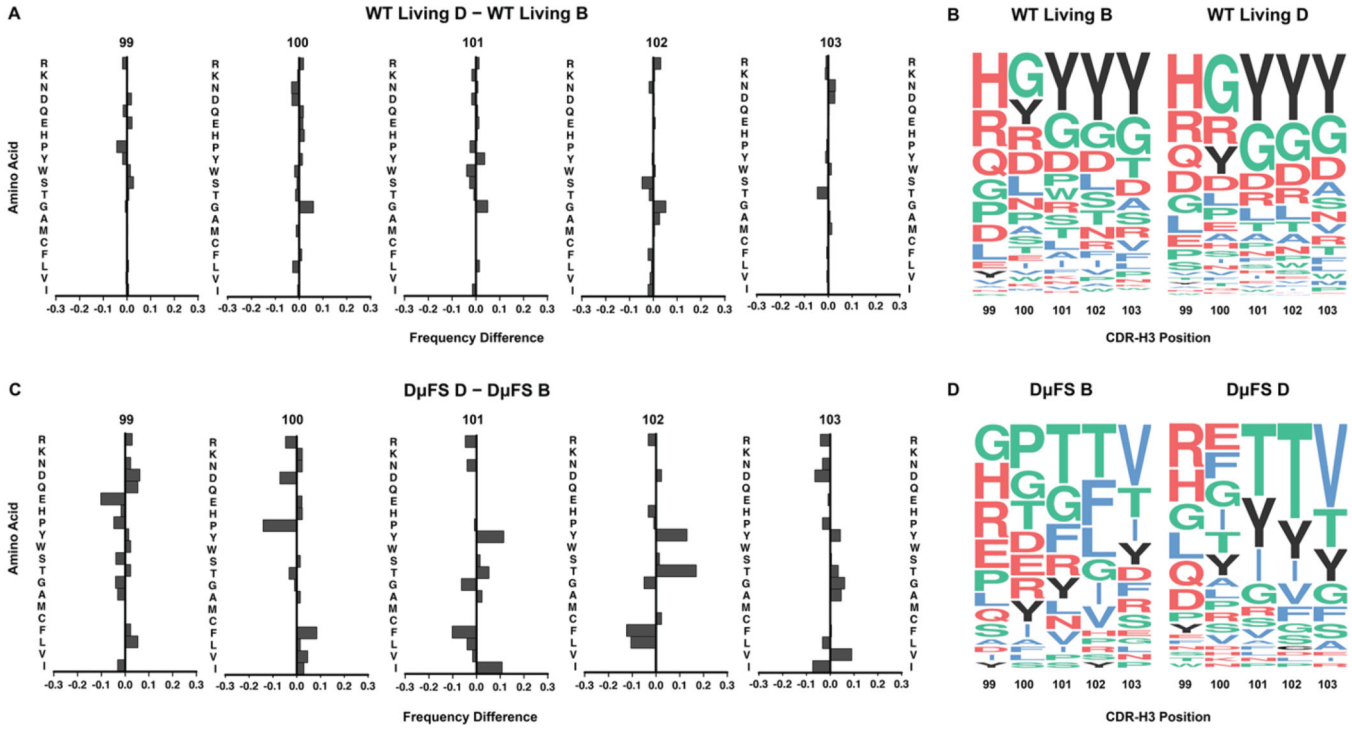

Fig. 2. CDR-H3 amino acid usage by position in proB (Fraction B) versus late preB (Fraction D) cells

(**A, C**) Fractional difference in the percent usage of individual amino acids between WT (A) and ΔD-DμFS (C) CDR-H3s at positions 99 to 103. Amino acids are arranged by hydrophobicity in the order of R, K, N, D, Q, E, H, P, Y, W, S, T, G, A, M, C, F, L, V, and I. (**B, D**) CDR-H3 amino acid usage by percentage rank order. The list of CDR-H3 sequences from WT and ΔD-DμFS cells in fractions B and D is reported in Supplemental File 1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

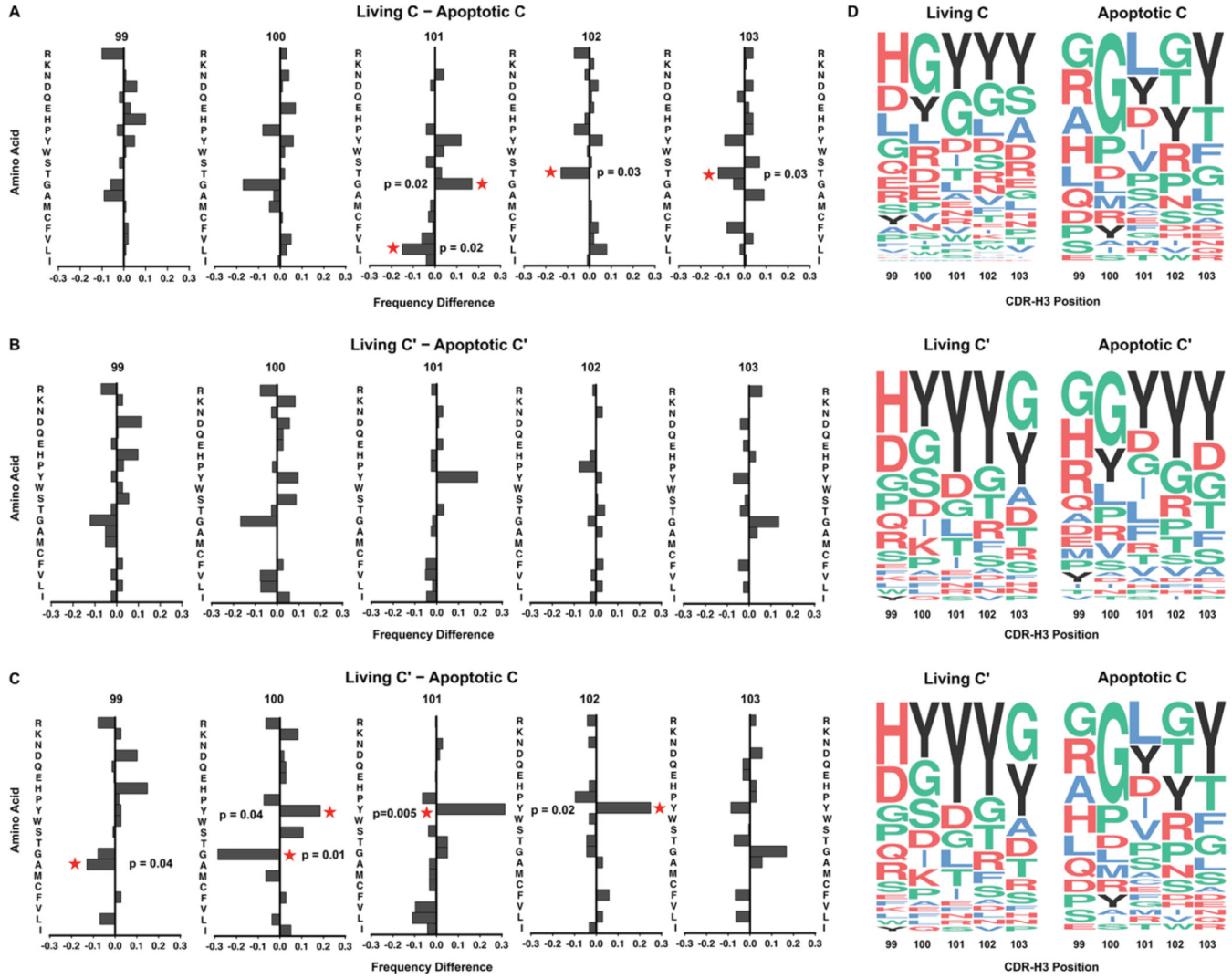

Fig. 3. CDR-H3 amino acid usage by position from wild-type living and apoptotic early preB cells

(A-C) Fractional difference in the percent usage of individual amino acids between CDR-H3 sequences cloned from apoptotic and living fraction C and C' preB cells at positions 99 through 103. Amino acids are arranged by hydrophobicity as in Figure 2. **(D)** CDR-H3 amino acid usage by percentage rank order. The list of CDR-H3 sequences from living and apoptotic cells in fractions C and C' is reported in Supplemental File 2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

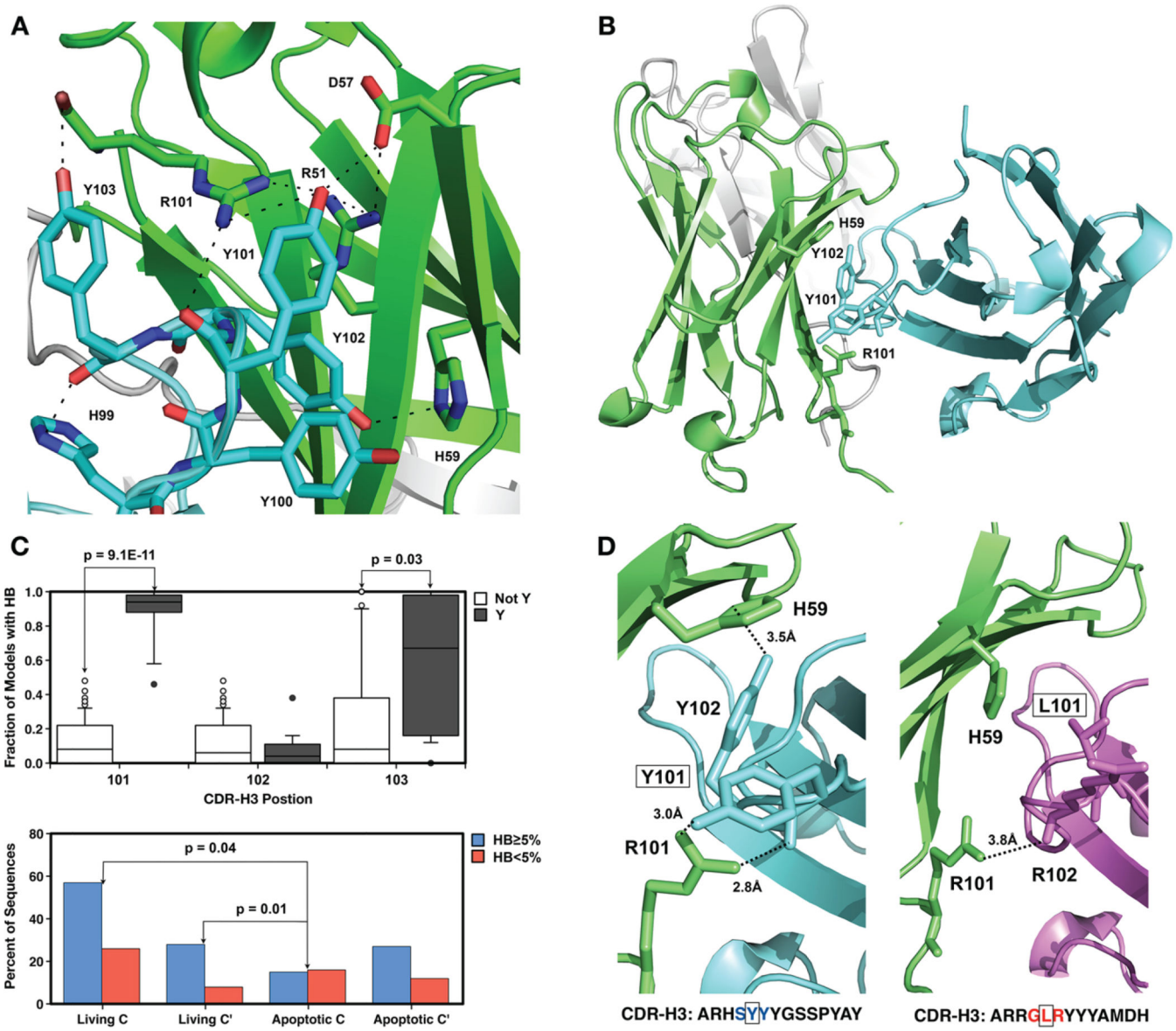

Fig. 4. Analysis of predicted preB cell receptor complex structures

(A) Structure of the murine preBCR obtained by modeling the H59 sidechain, which is a serine in the human VpreB structure (PDB: 2H32). VpreB is in green and CDR-H3 residues 99–103 are in cyan. **(B)** Example of a predicted structure of a mouse preB cell receptor complex from a wild-type living C' sequence. VpreB is in green, λ5 in gray, and the H chain V domain in cyan. **(C)** Analysis of the preBCR models from CDR-H3 sequences in living C and C' cells. The top panel shows the fraction of interacting models (IM) for the subset of structures with (black, Y) and without (white, Not Y) tyrosine in the indicated CDR-H3 position for those sequences without a tyrosine flanking that position. In the bottom panel, we compared the percentage of sequences from living and apoptotic C/C' cells forming a hydrogen bond (HB) with CDR-H3 SS at position 101 only in at least 5% of the models (blue bar, HB≥5%), with sequences which engage in hydrogen bond with position 101 in less than 5% of the models (red bar, HB<5%). **(D)** Comparison of the predicted structures of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

two almost identical VDJCμ sequences from living C' and apoptotic C cells that use the same V_H7183-10 gene segment and differ only in CDR-H3. The predicted interactions between the SLC CDR-H3 SS (green) and the CDR-H3 region from a living C' cell (left, cyan) and an apoptotic C' cell (right, magenta).

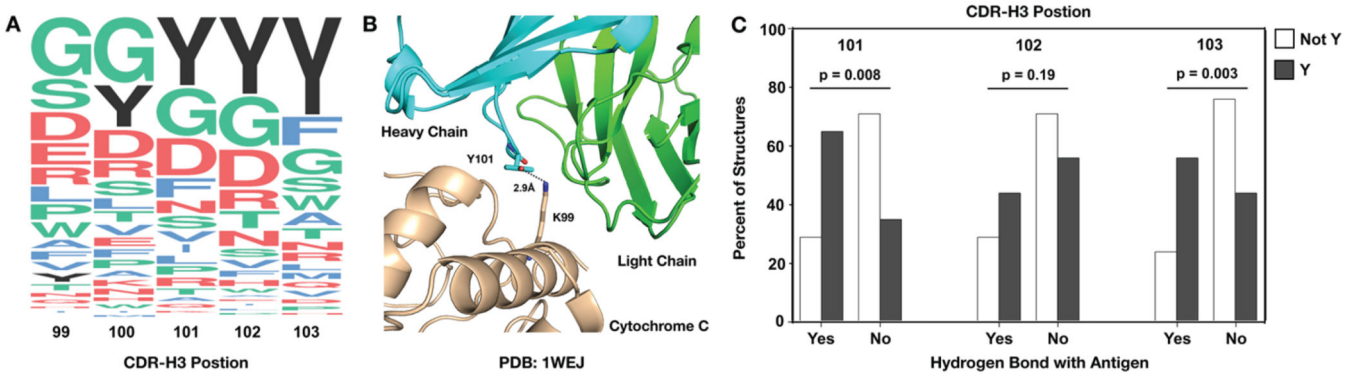

Fig. 5. Analysis of the CDR-H3 region in murine IgG Fab:antigen PDB structures

(A) Amino acid composition of CHR-H3 regions at position 99–103 in the 133 analyzed structures. **(B)** Example of CDR-H3 Y101 interacting with antigen (Cytochrome C) in PDB: 1WEJ. **(C)** Fraction of structures with (black, Y) and without (white, Not Y) tyrosine in CDR-H3 positions 101–103 that form hydrogen bonds with antigen among those sequences were the tyrosine in the selected position lacks a flanking (1 left, 1 right) tyrosine. The list of Fab:antigen PDB structures used in this analysis is reported in Supplemental File 3. The procedure used for selection of the Fab:antigen PDB structures is described in Supplementary Materials.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript