

Susceptibility Loci in C57BL/6 sle1, sle2 and sle3 Contain Genes that Alter Peripheral Selection of the CDR-H3 Sequences Enriched for Arginine

Mohamed Khass^{1,2}, Peter D Burrows³ and Harry W Schroeder^{1*}

¹Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA

²Genetic Engineering and Biotechnology division, National Research Center, Cairo, Egypt

³Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL, USA

*Corresponding author: Harry W Schroeder, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA, Tel: +205-382-2613; E-mail: hschroeder@uabmc.edu

Received date: February 18, 2018; Accepted date: March 09, 2018; Published date: March 14, 2018

Copyright: ©2018 Khass M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium.

Abstract

Systemic lupus erythematosus (SLE) is a multifactorial autoimmune disease characterized by deposition of ds-DNA binding autoantibodies in various body organs. These antibodies result from failure to control the composition of the B cell repertoire. Development of optimum B cell repertoire depends on the amino acid composition and the physicochemical characteristics at the center of the antigen binding site, the third complementarity determining region heavy chain (CDR-H3). Repertoire control involves positive selection for hydrophilic amino acids such as tyrosine and negative selection of hydrophobic and charged amino acids, specifically those containing arginine within the CDR-H3. Anti-dsDNA antibodies present in SLE patients exist in healthy individuals but at low levels, since dsDNA-specific B cells are deleted from the repertoire, but amplified in SLE patients. These antibodies contain arginine residues in CDR-H3, especially at positions 99-102, where they are positioned to bind negatively charged phosphate groups on the DNA backbone. Three genomic intervals, namely sle1 on chromosome 1, sle2 on chromosome 4, and sle3 on chromosome 7, were found to be associated with SLE susceptibility. We hypothesized that development of ds-DNA binding antibodies in SLE might result from failure to control CDR-H3 amino acid composition. We proposed that the SLE congenic loci might have unique effects in allowing survival/expansion of B cells expressing these auto-reactive antibodies. Our strategy was to change the composition of CDR-H3 by altering the germline composition of the DH gene segments. We created a Δ D-iD altered allele enriched for arginine while depleted of tyrosine at positions 99-102. We then monitored the influence of different SLE loci on the development and maintenance of B cells bearing CDR-H3 arginine. These findings support our hypothesis that peripheral B cell selection is altered by the presence of sle congenic alleles, allowing passage of B cells able to produce autoreactive antibodies binding ds-DNA. These findings may help in developing therapeutics to suppress autoimmunity in SLE.

Keywords: Review; Systemic lupus erythematosus; B cell repertoire; Third complementary determining region heavy chain; CDR-H3; Diversity gene segments DH

Mini Review

SLE is a systemic autoimmune disorder characterized by the production of a range of self-reactive antibodies [1,2], especially antibodies binding double stranded DNA, which can deposit in critical organs such as kidney, giving rise to many clinical manifestations of the disease. A better understanding of the mechanisms that control the composition of the antibody repertoire is thus critical to our understanding of the pathogenesis and treatment of SLE and other autoimmune diseases.

The mechanisms that are used to produce immunoglobulin diversity are extraordinarily powerful. Diversification begins with combinatorial V, D, J rearrangement. In the mouse, there are approximately 90 kappa variable gene segments (V κ) and four joining J κ gene segments, as well as three V λ and three J λ gene segments. Together, these give rise to approximately 370 different light (L) chain combinations. The addition of a diversity (D) gene segment in the heavy (H) chain locus geometrically enhances H chain combinatorial diversity. Each DH can undergo rearrangement by either deletion or inversion, thus encoding six different peptide sequences per gene segment. Thus, the

approximately 180 VH gene segments, 13 DH gene segments and four JH gene segments can produce more than 56,000 combinations. Then there is variation in the site of gene segment joining due to variation in exonucleolytic nibbling of the terminus of the recombining gene segments. At the ends, there can be variable addition of germline encoded palindromic sequence (P junctions) and, most powerfully, random addition of non-templated N nucleotides. Every three N nucleotides added increases the potential diversity of the antibody repertoire 20 fold. N addition is rare in L chains, but common in H chains. This further places the burden of diversity in the developing repertoire on the H chain. At a later stage in B cell development, typically following exposure to antigen and T cell help, each variable domain can undergo somatic hypermutation. It is only at this point that the L chain can attempt to engender the diversity created early in B cell development in the H chain by combinatorial and junctional diversification mechanisms. This incredible and almost astronomical potential for diversity has led to a common perception that the immunoglobulin (Ig) repertoire can be treated as random in its composition [3,4].

Diversity in the H chain and L chain repertoires is asymmetrically distributed. Each chain contains four sequence intervals of relatively conserved sequence, termed framework regions that are separated by four more diverse intervals, termed complementarity determining regions, or CDRs. These CDRs are juxtaposed to form the antigen

binding site, as classically defined. CDRs 1 and 2 form the outside of the antigen binding site, and CDR3 of the L chain forms the base. The third CDR of the H chain, CDR-H3, lies as the center of the antigen binding site where it often plays a key role in defining the binding activity of the antibody. Of these six CDRs, CDR-H3 is the most variable because it contains the termini of the V and J gene segments, the D gene segment in its entirety, and the effects of terminal nibbling, P junctions and N addition. As a result of its diversity and its central position, CDR-H3 often plays a critical role in defining the specificity and affinity of its antibody.

Each B cell, as it develops encounters self or non-self-antigens as it passes through a sequential series of quality control checkpoints. The assumption of a random repertoire is often coupled to the view that the repertoire is shaped primarily by the individual reactivity of immunoglobulin towards these antigens. Central checkpoints occur in the fetal liver and in the bone marrow as developing B cells build their immunoglobulins, and peripheral checkpoints are found in the spleen and other lymphoid organs as more mature B cells circulate through the body [5].

Close inspection of the amino acid composition of the developing CDR-H3 repertoire has led to a different view of the forces shaping the B cell repertoire in normal development and disease. Instead of the repertoire being selected individually as a result of reactivities against individual self and non-self-antigens, there is now evidence that the repertoire is being selected categorically as a result of the physico-chemical properties (e.g., tendency to form salt-bridges due to charge or a tendency to form hydrophobic interactions) of the epitopes on the self and non-self-antigens encountered during the process of development. This process of categorical selection leads to evolutionary pressure to control the germ line sequence of the various gene segments [3,4]. Evolutionary pressure is the product of natural selection for reproductive fitness. This view predicts that the presence or absence of specific categories of immunoglobulin sequence can influence the likelihood of protection against pathogen challenge and/or susceptibility to the production of disease-causing auto reactive antibodies.

Evidence in favor of natural selection of germ line immunoglobulin sequence can be found in the sequences of the various D gene segments. Each of the six DH reading frames exhibits a characteristic amino acid signature. One of the six reading frames tends to be enriched for neutral amino acids, including most notably glycine and tyrosine. Four of the six reading frames tend to be enriched for hydrophobic amino acids. The sixth and last reading frame tends to be enriched for charged amino acids, most notably arginine.

If the repertoire were to be truly random, then each of these reading frames would have a similar probability of expression. However, in vivo there is a clear preference for the use of the neutral reading frame (RF1). Hydrophobic RF2 and, to a lesser extent, hydrophobic RF3 are used less commonly; whereas the two hydrophobic RFs by inversion and the charged RF by inversion are used only rarely [6]. Reading frame control is heavily dependent on evolutionary selection of DH sequence. For example, use of RF3 is uncommon because it often contains a termination codon. Use of RF2 is inhibited by the presence of an ATG start site upstream of the DH in RF2 that enables production of a truncated D μ protein that helps block B cell development at the preB cell stage. And, use of RF1 is promoted by microhomology between the 3' terminus of the DH and the 5' terminus of the JH, which facilitates recombination into that RF.

Categorical selection of the germline repertoire by evolutionary pressure is insufficient to provide final control of the repertoire because N addition enables the inclusion of amino acids that are disfavored at the germline level. However, inspection of the CDR-H3 repertoire as it progresses through developmental checkpoints reveals evidence of continued categorical selection at developmental quality control checkpoints. I.e., somatic selection of the repertoire is also being exerted categorically. These categories include CDR-H3 lengths, average hydrophobicity, and even specific amino acids at particular structural positions [4,7-9]. Adding further complexity to this process is the observation that certain categories of CDR-H3s appear to be more or less tolerated in specific anatomical or developmental subsets, such as the peritoneal cavity, the splenic follicles, the marginal zone, and recirculating mature B cells, including those found in the bone marrow.

These two mechanisms, categorical natural selection of germline Ig sequence and categorical somatic selection during checkpoint passage, appear to work in tandem to control the diversity of the repertoire. To test the functional effects of violating these categorical controls, we previously generated a panel of mice where we had altered the DH locus to force alterations in the pattern of use of individual RFs [4]. We used gene targeting methods to delete all but one of the functional DH gene segments in the DH locus, and then altered the sequence of the remaining DH. The Δ D-DFL strain retains one normal DH. The Δ D-D μ FS strain is frameshifted to promote use of hydrophobic RF2 over neutral RF1. And, the Δ D-iD strain uses inverted RF1 sequence that is enriched for use of arginine in place of normal RF1 sequence that is enriched for tyrosine (Figure 1). As a result of these changes, the preimmune IgM repertoire demonstrated enrichment for hydrophobic amino acids in CDR-H3 in the Δ D-D μ FS mice, and for arginine in CDR-H3 in the Δ D-iD mice. Further, we observed major alterations in B cell numbers in the bone marrow and in the periphery, progressively altered patterns of antibody production in general and in specific to individual antigens as the repertoire deviated from normal, and enhanced susceptibility to infection after pathogen challenge. In particular, we observed that categorical changes in the composition of the CDR-H3 repertoire led to changes in epitope recognition [10].

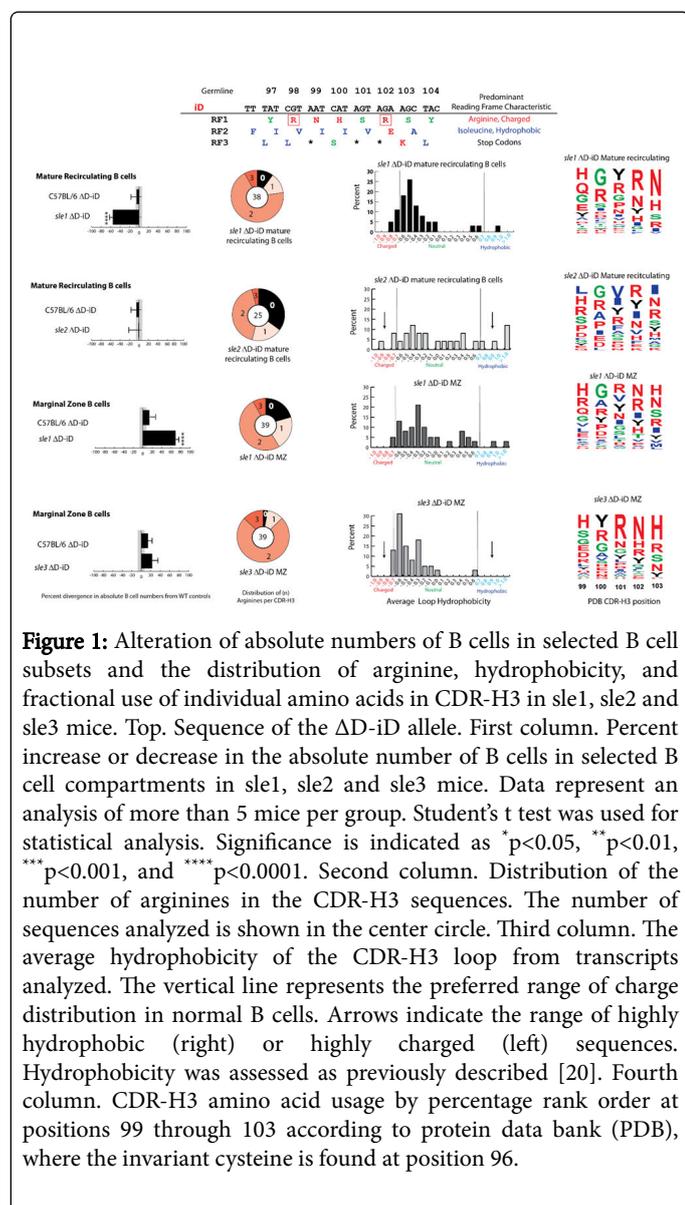
The deviations in B cell number reflected the ameliorating effects of passage through developmental checkpoints centrally and in the periphery [11] to adjust the repertoire. However, the effects of the change in DH sequence could not be erased, and the repertoires remained altered. Most pertinent to our current work, we found that with the passage of time, mice both homozygous and heterozygous for the Δ D-iD DH allele, which is enriched for arginine, began to produce dsDNA binding IgG antibodies in BALB/c mice, which are viewed as normally resistant to autoantibody production [12].

These considerations led us to the hypothesis that one set of mechanisms that leads to the production of dsDNA binding antibodies in lupus-prone individuals is a failure to properly regulate the CDR-H3 antigen receptor repertoire. This hypothesis would suggest that the range of epitope reactivities in patients with autoimmune disease in general would differ from non-disease susceptible individuals and thus manifest the ability to respond to a range of antigens that other people might find more difficult to elicit. This could help explain, for example, why some SLE patients can more quickly and easily produce broadly reactive HIV antibodies [13].

To test this hypothesis, we compared selection of the antibody repertoire in mice that were more susceptible to producing dsDNA binding antibodies, such as MRL and C57BL/6 [14-16], to those that

were less susceptible, such as C3H and BALB/c [17,18]. In these studies, we found evidence that categorical selection against charged, arginine-enriched CDR-H3s was impaired in the susceptible strains.

The NZM2410 mouse is a New Zealand Black/White-derived inbred strain that develops early-onset lupus nephritis [14]. Backcrossing the NZM2410 genome onto C57BL/6 led to the identification of three novel genomic intervals, *sle1* on chromosome 1, *sle2* on chromosome 4, and *sle3* on chromosome 7, that are associated with susceptibility to lupus [19]. In the congenic strain B6.NZMc1, the *sle1* locus is associated with potentiating a strong, spontaneous humoral response to H2A/H2B/DNA subnucleosomes. In the B6.NZMc4 strain, *sle2* leads to B-cell hyperactivity, elevated levels of B1a cells in the spleen and peritoneal cavity, and increased total serum IgM. In the congenic strain B6.NZMc7, *sle3* promotes an increase in activated CD4 T cells, decreased susceptibility to apoptosis, and production of low titers of antinuclear antibodies. Triple congenic C57BL/6 *sle1,2,3* mice approach the autoimmune disease phenotype of the parental NZM2410 strain, including high ANA titers.



As noted above, the Δ D-iD DH allele promotes the inclusion of arginine in CDR-H3 in place of tyrosine, especially at CDR-H3 position 101 (Figure 1) [4,18]. To test the hypothesis that the inheritance of lupus-susceptibility genes could generally influence categorical selection of the Ig repertoire, we backcrossed the Δ D-iD allele onto C57BL/6 congenic for *sle1*, *sle2* or *sle3* loci. We then examined B cell and repertoire development by assessing absolute B cell numbers and CDR-H3 sequences from selected B cell subsets, and evaluating serum anti ds-DNA antibodies.

We found that the introduction of the Δ D-iD DH allele into C57BL/6 *sle1* congenic mice led to a reduction in the absolute number of mature recirculating B cells and an increase in marginal zone B cell numbers. In these cells, a highly charged, an arginine-enriched CDR-H3 repertoire was maintained. In contrast, introduction of *sle2* normalized mature B cell numbers with a partial normalization of CDR-H3 hydrophobicity and length, resulting in enhanced use of hydrophobic amino acids at position 101. *sle3*/ Δ D-iD increased the number of MZ cells and maintained preference for arginine at CDR-H3 position 101. Twelve month old Δ D-iD DH C57BL/6 showed increased titers of IgM and IgG dsDNA binding antibodies when either *sle1* or *sle3* was present.

Our previous studies found evidence that the antibody repertoire could be regulated categorically by the presence or absence of amino acids with specific physicochemical properties in CDR-H3, which forms the center of the antigen binding site. We provided evidence that categorical selection begins with natural selection of germline immunoglobulin sequence by evolution, likely as a result of changes in reproductive fitness. We previously found that categorical changes in the composition of the immunoglobulin repertoire were associated with changes in patterns of epitope recognition. We observed that the introduction of somatic variability into the germline repertoire by mechanisms such as N addition resulted in the creation of diversity that was outside of the bounds set by natural selection.

We found evidence that passage through quality control checkpoints during B cell development could be used to ameliorate the effects of violating the boundaries placed by evolution on germline immunoglobulin sequence. This led us to the hypothesis that non-immunoglobulin genes that can influence checkpoint passage could categorically alter the immunoglobulin repertoire and thus alter normal patterns of self-non-self recognition and thus the production of autoreactive antibodies.

Our findings in *sle* mice support the hypothesis that *sle* congenic alleles may act, in part, to influence the process of categorical selection of the antibody repertoire [9]. A corollary of this hypothesis is that some lupus-susceptibility alleles may have general effects on repertoire development that could alter patterns of antibody production against multiple antigens, not just the pathognomonic reactivities that define a specific clinical entity. Conversely, our findings suggest that pharmacologic interventions directed at the mechanisms that regulate quality control checkpoint passage in B cell development may be able to alter the range of antigen binding site selection to allow, or prevent, specific categories of antigen binding sites to be produced. Such interventions might be helpful as a means to elicit or modify vaccine responses, manipulate or modulate the immune response to infectious agents, and suppress development of autoimmunity.

References

1. Sang A, Zheng YY, Morel L (2014) Contributions of B cells to lupus pathogenesis. *Mol Immunol* 62: 329-338.
2. Kaul A, Gordon C, Crow MK, Touma Z, Urowitz MB, et al. (2016) Systemic lupus erythematosus. *Nat Rev Dis Primers* 2: 16039.
3. Kirkham PM, Mortari F, Newton JA, Schroeder HW Jr (1992) Immunoglobulin VH clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO Journal* 11: 603-609.
4. Schroeder HW Jr, Zemlin M, Khass M, Nguyen HH, Schelonka RL (2010) Genetic control of DH reading frame and its effect on B-cell development and antigen-specific antibody production. *Crit Rev Immunol* 30: 327-344.
5. Karrar S, Cunninghame Graham DS (2017) Abnormal B-cell development in Systemic Lupus Erythematosus: what the genetics tell us. *Arthritis Rheumatol*.
6. Gu H, Kitamura D, Rajewsky K (1991) DH reading frame bias: evolutionary selection, antigen selection or both? *Evolutionary selection. Immunol Today* 12: 420-421.
7. Verkoczy L, Chen Y, Zhang J, Bouton-Verville H, Newman A, et al. (2013) Induction of HIV-1 broad neutralizing antibodies in 2F5 knock-in mice: selection against membrane proximal external region-associated autoreactivity limits T-dependent responses. *J Immunol* 191: 2538-2550.
8. Meffre E, Milili M, Blanco-Betancourt C, Antunes H, Nussenzweig MC (2001) Immunoglobulin heavy chain expression shapes the B cell receptor repertoire in human B cell development. *J Clin Invest* 108: 879-886.
9. Khass M, Blackburn T, Burrows PD, Walter MR, Capriotti E (2016) VpreB serves as an invariant surrogate antigen for selecting immunoglobulin antigen-binding sites. *Sci Immunol* 1: aaf6628.
10. Wang Y, Kapoor P, Parks R, Silva-Sanchez A, Alam SM, et al. (2016) HIV-1 gp140 epitope recognition is influenced by immunoglobulin DH gene segment sequence. *Immunogenet* 68: 145-155.
11. Khass M, Blackburn T, Elgavish A, Burrows PD, Schroeder HW Jr (2018) In the Absence of Central pre-B Cell Receptor Selection, Peripheral Selection Attempts to Optimize the Antibody Repertoire by Enriching for CDR-H3 Y101. *Front Immunol* 9: 120.
12. Keenan RA, De Riva A, Corleis B, Hepburn L, Licence S, et al. (2008) Censoring of autoreactive B cell development by the pre-B cell receptor. *Science* 321: 696-699.
13. Bonsignori M (2014) Will studies in individuals with systemic lupus erythematosus be the key to future HIV vaccine design? *Expert review of vaccines* 13: 1271-1273.
14. Bolland S, Yim YS, Tus K, Wakeland EK, Ravetch JV (2002) Genetic modifiers of systemic lupus erythematosus in FcγRIIB(-/-) mice. *J Exp Med* 195: 1167-1174.
15. Zemlin M, Ippolito GC, Zemlin C, Link J, Monestier M (2005) Adult lupus-prone MRL/MpJ2+ mice express a primary antibody repertoire that differs in CDR-H3 length distribution and hydrophobicity from that expressed in the C3H parental strain. *Mol Immunol* 42: 789-798.
16. Khass M, Buckley K, Kapoor P, Schelonka RL, Watkins LS, et al. (2013) 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* 43: 629-640.
17. Izui S, Kelley VE, Masuda K, Yoshida H, Roths JB (1984) Induction of various autoantibodies by mutant gene *lpr* in several strains of mice. *J Immunol* 133: 227-233.
18. Silva-Sanchez A, Liu CR, Vale AM, Khass M, Kapoor P, et al. (2015) Violation of an evolutionarily conserved immunoglobulin diversity gene sequence preference promotes production of dsDNA-specific IgG antibodies. *PLoS One* 10: e0118171.
19. Wakeland EK, Morel L, Mohan C, Yui M (1997) Genetic dissection of lupus nephritis in murine models of SLE. *J Clin Immunol* 17: 272-281.
20. Ivanov II, Schelonka RL, Zhuang Y, Gartland GL, Zemlin M (2005) 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* 174: 7773-7780.