Published in final edited form as:

Semin Immunol. 2010 December ; 22(6): 346–352. doi:10.1016/j.smim.2010.08.002.

Epigenetic and 3-dimensional regulation of V(D)J rearrangement of immunoglobulin genes

Stephanie C. Degner-Leisso and Ann J. Feeney*

The Scripps Research Institute, Department of Immunology and Microbial Science, IMM-22, 10550 North Torrey Pines Rd., La Jolla, CA 92037

Abstract

V(D)J recombination is a crucial component of the adaptive immune response, allowing for the production of a diverse antigen receptor repertoire (Ig and TCR). This review will focus on how epigenetic regulation and 3-dimensional (3D) interactions may control V(D)J recombination at Ig loci. The interplay between transcription factors and post-translational modifications at the *Igh*, *Igk*, and *Igl* loci will be highlighted. Furthermore, we propose that the spatial organization and epigenetic boundaries of each Ig loci before and during V(D)J recombination may be influenced in part by the CTCF/cohesin complex. Taken together, the many epigenetic and 3-D layers of control ensure that Ig loci are only rearranged at appropriate stages of B-cell development.

Keywords

epigenetics; immunoglobulin genes; locus contraction; histone post-translational modifications; V(D)J recombination; repertoire

1. Introduction

The process of V(D)J recombination utilizes the many germline V, D and J genes to generate a highly diverse repertoire of immunoglobulin (Ig) heavy and light chains, which combine to form the complete Ig receptors that are placed on the surface of each B lymphocyte. Thus, the millions of B lymphocytes in the body are armed with a great diversity of B cell receptors to combat a wide variety of pathogens. In each precursor B cell, a different set of V, D and J genes recombine to form the exons for the heavy and light chains of the antibody molecule. Between the combinatorial diversity afforded by the large number of V, D and J genes that are chosen for recombination in a semi-random manner, and the junctional diversity provided by deletion of a small number of nucleotides from the coding ends and the addition of N nucleotides, the vast majority of B cell receptors that are created in the adult bone marrow are unique. T cells also create a highly diverse repertoire of receptors using V, D and J genes from the T cell receptor loci that recombine in the same fashion as Ig genes.

The V(D)J recombination process is highly regulated [1]. Firstly, although both T and B cells use the same RAG1 and RAG2 recombinase, T cell receptor genes are never recombined in B

© 2010 Elsevier Ltd. All rights reserved.

*to whom correspondence should be addressed: phone: 1-858-784-2979, fax: 1-858-784-9190, feeney@scripps.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

cell precursors, and T cell precursors do not rearrange any B cell receptor light or heavy chain genes, with the exception of limited rearrangement of D_H and J_H genes in T cells. Since D-J rearrangement begins in precursor lymphocytes before commitment to the B cell lineage, the limited rearrangement of DJ observed in thymocytes may partially occur at this earlier stage.

Secondly, rearrangement occurs in a precise order [2]. D_H to J_H rearrangements always occur on both alleles before V_H to DJ_H rearrangement begins on one allele. Furthermore, the *Igh* locus always completes rearrangement before light chain rearrangement begins. Once the pro-B cells have successfully rearranged a heavy chain gene, the heavy chain protein binds to lambda 5 and VpreB that together form the pre-BCR. The pro-B cell then undergoes several rounds of proliferation as a result of signaling through the pre-BCR. Once the proliferation ceases, the pre-B cells revert to small size, no further rearrangements take place at the *Igh* locus, and the cells begin rearrangement of the *Igk* V and J genes. For the pre-B cells, there is a third stage of ordered rearrangement. Ig molecules can use either kappa or lambda light chains, and *Igk* rearrangement precedes *Igl* rearrangement [3]. This same ordered rearrangement process occurs for $\alpha\beta$ T cells also, with the beta chain undergoing $D\beta$ - $J\beta$ rearrangement, followed by $V\beta$ to $DJ\beta$. The TCR β protein binds to pre-TCR α , the thymocytes proliferate for a few rounds, then the *TCR α* locus begins to undergo rearrangement.

This lineage specificity and precisely ordered developmental stage specificity was discovered over 25 years ago by Alt and colleagues, and the data led them to propose the “accessibility hypothesis” which states that only the particular part of the locus which is undergoing rearrangement is accessible to the RAG recombinase [1]. At that time, the only indication of the accessibility of a region was the presence of so-called germline or sterile transcripts, which are non-coding transcripts from unrearranged genes just preceding rearrangement at that part of the locus [4,5]. Whether these germline transcripts were merely a by-product of the fact that that region was now accessible, or whether the transcription played a role in making the region accessible for rearrangement was unclear for a long time. Evidence is accumulating that the transcription is essential, at least in the J portion of the loci [6]. Now, 25 years later, we know quite a bit more about this precise control of accessibility of the V(D)J recombination mechanism. In this review, we will present highlights of our lab’s research in this area.

2. Dynamic landscape of chromatin modifications associated with Ig and TCR loci during B and T cell differentiation

In order to determine how the landscape of histone posttranslational modifications changes during lymphocyte development, and between the T and B lineage precursors, we sorted pro-B cells, pre-B cells and immature B cells, as well as double negative and double positive thymocytes [7]. These cells were surveyed by chromatin immunoprecipitation (ChIP) with antibodies against 4 common histone posttranslational modifications associated with active genes: H3K4me2, H3K4me3, H3K79me2 and AcH3. Real time PCR was used to assess the status of the histones associated with representative V, D and J genes from the *Igh*, *Igk*, *Igl* and *TCR β* loci in rearranging or non-rearranging loci.

The first observation, not surprisingly, is that there was lineage and stage specificity to the histone modifications, in that loci that were not undergoing rearrangement did not bear any (or low levels) of these histone modifications. However, the pattern of each of these 4 histone modifications was different, and it was a surprise that that many loci, most notably the V loci, did not have high levels of any of these modifications, with the notable exception of H3K4me2, which was found associated with V, D and J genes. This histone modification had been studied previously in cell lines, and was shown to be present on genes “poised” for recombination [8]. We found that H3 acetylation was far higher on J genes than on V or D genes in all loci, and again with a lineage and stage specific pattern (Fig. 1). This pattern was previously

observed in cell lines [8]. Most dramatic was H3K4me3 which we observed to be almost exclusively associated with J genes of rearranging loci, and with nearby, but not distant, D genes (Fig. 1). This observation is of great biological importance, since it had been shown that RAG2 can bind to H3K4me3 [9,10]. We hypothesized that if only the J genes bear H3K4me3, this could facilitate ordered rearrangement, since RAG2 would preferentially bind to the J genes [7]. Recently, Schatz and colleagues have elegantly demonstrated by ChIP that indeed RAG2 does bind to J genes and nearby D genes exclusively, with a pattern genome-wide which mirrors that of H3K4me3 [11]. Thus, the pattern of H3K4me3 on rearranging genes could well enforce the ordered rearrangement of DJ before V to DJ, and of heavy chain rearrangement but not light chain rearrangement at the pro-B cell stage. Since histone methyltransferases can travel with the RNA polymerase II complex, it is likely that the germline transcription which is observed before genes begin to rearrange could be responsible for adding the H3K4me3 modification on histones associated with J gene during the process of germline transcription, thus making a binding site for RAG2. Direct demonstration of the necessity of germline transcription for induction of accessibility, increase in H3K4me3, and rearrangement has been nicely provided by Abarrategui and Krangel who inserted a transcriptional terminator into the midst of the $J\alpha$ locus [6]. This resulted in greatly decreased recombination of the $J\alpha$ genes downstream of the terminator.

3. Epigenetic control of kappa and lambda light chain rearrangement

As mentioned above, kappa rearrangement precedes lambda rearrangement. We investigated the epigenetic status of kappa and lambda genes throughout B cell development by ChIP, using primers for V κ 1 (family-specific), J κ 1, J κ 5, V λ 1, J λ 1 [7]. J κ 1 and J κ 2 are often used in the original kappa rearrangements, and J κ 5 is enriched in cells that have undergone multiple rearrangements, either due to initial non-productive rearrangements or due to receptor editing [12-14]. V λ 1 is the most frequently observed rearranging gene of the 3 V λ genes and it predominantly rearranges to J λ 1. Since λ rearrangement occurs after kappa rearrangement, the presence of λ light chain genes is a hallmark of a cell that has undergone receptor editing [13].

We observed that the V genes showed little histone modification compared to the J genes except for H3K4me2 which was present on all of the genes analyzed in pre-B cells. The other three modifications were clearly present on the J κ gene in pre-B cells, but only at a very low level on J λ 1. However, the extent of these active modifications increased on J λ in the immature B cell stage. Although the pre-B cell stage is the stage where most light chain rearrangement occurs, once a cell makes a productive light chain, it pairs with the existing heavy chain protein creating a receptor which is expressed on the surface of the immature B cell. If this surface immunoglobulin reacts with a self antigen, the immature B cell receives a receptor editing signal telling it that this is an autoreactive receptor, and that the cell should continue light chain rearrangement and try again to make a different light chain, in hopes that this new light chain will not make an autoreactive receptor with the cell's heavy chain. The cell reverts to a pre-B cell phenotype as it first downregulates then completely eliminates the surface Ig, and deletes the autoreactive light chain gene with additional rearrangements of an upstream V to a downstream J, or RS rearrangement which ends rearrangement at the kappa locus [15,16]. It has been shown that ~50% of receptors initially made are autoreactive and must undergo receptor editing [17]. Thus, the late onset of histone modifications at J λ correlate with the fact that it rearranges later in B cell development. Furthermore, the fact that higher levels of histone modifications are observed at the immature B cell stage than at the pre-B cell stage suggests that indeed lambda rearrangements often occur as a response to receptor editing signals.

4. Pre-BCR signaling is required for induction of histone modifications on light chain genes

The histone modifications appear early on the J κ genes. After cells successfully rearrange the heavy chain gene, the heavy chain associates with the surrogate light chain components $\lambda 5$ and V pre-B, and together comprise the pre-BCR. This pre-BCR gives a signal to the pro-B cell to stop heavy chain rearrangement and to undergo several rounds of proliferation. After this proliferative burst, the cells revert to being small cells, and light chain rearrangement is initiated. We sorted large and small pre-B cells, and found that even in the large pre-B cells, histone modifications were already present on the J κ genes to a level ~ half of that of small pre-B cells, and germline transcription of the kappa locus has already begun [7]. No modifications were seen on the J λ gene other than a low level of H3K4me2. In order to see if the pre-BCR signal was necessary to initiate the histone modifications, we analyzed CD19⁺ cells from the bone marrow of two strains of mice (μ MT and *mb1*^{-/-}) that can rearrange *Igh* genes and make a heavy chain protein, but due to signaling defects, cannot successfully transmit a pre-BCR signal. We compared these to CD19⁺ cells from the bone marrow of RAG mice bearing a rearranged heavy chain gene (RAG^{-/-} μ ⁺). Due to the absence of RAG1, these CD19⁺ cells cannot rearrange light or heavy chain genes, but the heavy chain transgene pairs with the surrogate light chain components and thus successfully transmits a pre-BCR signal to the cells. Only the CD19⁺ cells from the RAG^{-/-} μ ⁺ mice showed any H3K4me2, H3K4me3, H3K79me2 or acetylated H3 on the κ or λ J genes. Similarly, it was the only strain of the 3 we analyzed that had J κ C κ germline transcription of the *Ig κ* locus. Thus, there is a requirement for pre-BCR signaling to initiate the histone posttranslational modifications and germline transcription on the light chain genes. We further showed that the kappa intronic enhancer (iEk) was required for histone modifications and for germline transcription by analysis of pre-B cells from mice with a germline deletion of iEk.

5. Transcription factor control of accessibility

In collaboration with Kees Murre, we showed that ectopic expression of E2A and RAG1/RAG2 in a human non-lymphoid cell line induced V(D)J rearrangement of kappa genes, but no lambda rearrangement was induced [18]. Conversely, when we ectopically expressed EBF with RAG1/RAG2, we induced rearrangement of lambda genes, but not of any kappa genes. Further investigation revealed that E2A induced a specific subset of V κ genes, the V κ I family, but did not induce rearrangement of the interspersed V κ II or V κ III genes [19]. This indicated that E2A did not make the entire kappa locus accessible for rearrangement, but rather that it made individual genes accessible, presumably by binding to sites near the promoter or RSS of the V gene. Since E2A has been demonstrated to be able to recruit histone acetyltransferases [20], we hypothesized that E2A induced accessibility of the V κ I genes to undergo rearrangement by inducing histone acetylation of these genes. Indeed, ectopic expression of E2A induced acetylation of histones associated with V κ I genes, but not of other V κ families nor of V λ genes. Likewise, EBF induced histone acetylation exclusively of V λ genes, the genes that are induced to undergo rearrangement after EBF ectopic expression. Thus, our hypothesis is that the way that the transcription factors can induce rearrangement of specific subsets of V genes is by binding near V genes and recruiting histone modifying enzymes that will acetylate or methylate the histones, or chromatin remodeling enzymes, rendering the V gene accessible for recombination. The data further suggest that each locus will have V genes that are differentially susceptible to induction of accessibility by various transcription factors, and that it may take the concerted effort of several transcription factors to make a diverse antibody repertoire.

6. Epigenetic status correlates with rearrangement frequency of V_H genes

In our global analysis of the dynamic changes in histone posttranslational modification during lymphocyte development, we found that acetylation of histone H3 was very low on V genes [7]. However, if one ignores the fact that the extent of modification of H3 was much higher on J_H genes, and just focuses on the V_H genes, then one can see that the extent of acetylation mirrors rearrangement potential. For example, we have studied the V_H S107 family extensively. The V1 gene rearranges far more than the other two members of the family, and the V13 gene hardly ever rearranges [21]. The V1 gene has an RSS which is closer to consensus than the other genes, and indeed it rearranges better in a plasmid-based recombination substrate [22]. However, V11 and V13 rearrange at very different frequencies, yet they have identical RSS. Analysis of the acetylation status, however, showed a gradient of enrichment of H3Ac that mirrored the rearrangement frequency, and the reciprocal pattern was observed for the repressive modification H3K9me2 [23,24]. Likewise, the extent of acetylation of 81X gene is much higher than that of the other V_H 7183 genes, and it rearranges at a very high frequency [23]. Thus, the epigenetic status of these V_H genes reflects their rearrangement frequency.

7. Complex regulation of D_H-proximal vs. D_H-distal V_H genes at the *Igh* locus

The V_H locus contains ~100 functional V_H genes spread over 2.5 Mb, whereas the ~13 D_H genes are located in a 86 kb region. The rearrangement of the most D_H-proximal V_H family, V_H7183, is under very different regulation from that of the distal half of the V_H locus, which contains the very large V_HJ558 family and also the V_H3609 genes. It has been reported that mice deficient in the transcription factors Pax5 or YY1, the histone methyltransferase Ezh2, and the cytokine receptor IL7R α chain, all share the same phenotype of rearranging the genes in the D_H-proximal V_H7183 family at almost wild type levels, while barely rearranging the D_H-distal V_HJ558 and V_H3609 families at all [25-28]. A gradient of rearrangement of the smaller V_H gene families in between these large V_H families is observed. These 4 strains were all surveyed by family-specific semi-quantitative PCR. The V_HQ52 gene family is intermixed with the V_H7183 genes in the proximal portion of the locus, yet all of these studies showed that the V_HQ52 family rearranged somewhat less frequently than the V_H7183 family. This may be due to the high frequency of rearrangement of the most 3' functional V_H7183 gene, 81X. Individual V_H gene usage was only determined for the Pax5 KO pro-B cells, and there the V_H7183 and V_HQ52 genes in the D_H-proximal half of the V_H7183/V_HQ52 region were slightly overrepresented and the ones furthest from the D regions were slightly underrepresented in rearrangements [29]. Thus, there was a real gradient of rearrangement frequency from one end of the V_H locus to the other end in all of these deficient strains of mice.

The Ezh2-deficient pro-B cells were of particular interest to our studies on epigenetic regulation of V_H gene rearrangement since this is the histone methyltransferase that methylates lysine 27 of histone 3, a repressive modification. We therefore wanted to determine if any V_H genes were methylated on H3K27. ChIP analysis revealed that the proximal V_H7183 family genes did have H3K27 methylation, but not the distal V_HJ558 genes (Fig. 1), and ChIP-chip analysis of the entire *Igh* locus confirmed that H3K27me2 was strictly localized on the V_H7183/V_HQ52 portion of the locus [30]. Since absence of Ezh2 results in an inability to rearrange distal V_H genes, we presume that the presence of H3K27me2 on proximal V_H genes may aid the ability of distal V_H genes to undergo rearrangement, perhaps by making the proximal V_H genes less accessible. Since fetal B cells preferentially rearrange proximal V_H genes, we asked if the *Igh* locus from fetal mice had H3K27me3 on proximal V_H genes. If this repressive modification on proximal V_H genes aided distal V_H gene rearrangement, then it might be possible that the absence of H3K27me3 on proximal V_H genes would bias rearrangement to that region. Indeed, we found no H3K27me3 on proximal V_H genes in fetal pro-B cells [30].

In order to see if there was any intersection of the IL7 signaling pathway or the Pax5 pathway in the Ezh2-deficiency, we investigated whether H3K27me3 was present on pro-B cells from these mutant strains. IL7-deficient or IL7 transgenic (therefore overexpressing IL7) pro-B cells had normal levels of H3K27me3. However, Pax5^{-/-} pro-B cells lacked H3K27me3. We had previously demonstrated that Pax5 bound to V_H genes both by EMSA and by ChIP [31], so it is possible that Pax5 might bind to the V_H locus and recruit Ezh2 to modify the V genes. However, since we found Pax5 binding to distal V_H genes as well as to the proximal V_H genes, this may not be the mechanism by which these two pathways interact, although the binding of Pax5 to proximal V_H genes was much stronger than the binding to the distal V_H genes [31]. Nonetheless, these two pathways appear to intersect.

We found another histone posttranslational modification that was present at relatively high levels on V_H genes. H3K36me2 is present on activated genes in general, and we found a reciprocal pattern of H3K36me2 and H3K27me2 by ChIP-chip throughout the V_H locus [30]. The V_HJ558/V_H3609 distal half of the V_H locus had high levels of H3K36me2, and the level tapered off in the middle of the V_H locus. Low levels were observed in the V_H7183/V_HQ52 proximal V_H genes. We found that IL7 signaling increased the level of H3K36me2, while Pax5 signaling did not affect the level at all. Since IL7 signaling had previously been demonstrated to raise the level of acetylation of H3 on distal V_H genes [32], it appears that IL7 signaling promotes two histone posttranslational modifications associated with active, and thus accessible, genes specifically on the distal V_H genes: AcH3 and H3K36me2.

8. Organization of receptor loci: need for locus contraction for efficient rearrangement

All of the large antigen receptor loci (*Igh*, *Igk*, *TCRβ*, *TCRα/δ*) are organized in a similar manner. The V gene portion of the locus covers the vast majority of the locus for all 4 of these loci (Fig. 2). Although the Ig loci only have 4-5 J genes, the TCR loci, especially *TCRα/δ*, have many more J genes, yet the proportion of the locus which is covered by the J genes is very low in all the loci. The reason for this is likely to be related to the fact that each locus has to undergo locus contraction at the time for rearrangement, and that in the uncontracted conformation, most V genes are too far from the (D)J regions to undergo rearrangement. This would aid in developmental stage-specificity of recombination. Indeed, contraction of the *Igh* locus exclusively in pro-B cells and of *Igk* exclusively in small B cells has been demonstrated by 3D-FISH [29]. Similarly, stage-specific contraction of the *TCRβ* and *TCRα/δ* loci have been documented by Skok et al. [33]. Detailed 3D-FISH studies of locus contraction at the *Igh* locus have been performed by Jhunjhunwala et al. [34]. Using 12 10-kb probes spread across the *Igh* locus, and using trilateration techniques, they were able to obtain a detailed picture of the changes in the 3-dimensional structure of the *Igh* locus as it progresses from pre-pro-B cells to pro-B cells. They demonstrated that the *Igh* locus forms a multi-loop rosette-like structure in the pro-B cells, with the D_H-J_H part of the locus in close proximity to both distal and proximal parts of the V_H part of the locus, while the pre-pro-B cells have the D_H-J_H part of the locus far away from the V_H genes, and in their own separate compartment. Thus, the contraction process serves to bring all of the V genes spread over these enormous megabase distances from the J genes into proximity with the J genes so that a highly diverse repertoire of Ig or TCR can be assembled making full use of the large number of V genes.

9. Role of CTCF in locus contraction

What is mediating the contraction of the receptor loci and the formation of the rosette-like structure? A good candidate is CTCF, a ubiquitously expressed vertebrate zinc-finger protein. CTCF has been shown to be involved in looping and long-range chromosomal interactions at many complex loci, which may be the basis for its known insulator and enhancer blocking

activities [35-37]. Since the studies of Jhunjhunwala et al have shown that the *Igh* locus forms a multi-loop structure during recombination, we hypothesized that CTCF may well facilitate the looping of the *Igh* locus at the pro-B cell stage. In order for this hypothesis to be true, there must be many CTCF binding sites throughout the V_H portion of the *Igh* locus that could form the bases of the multi-loop structure. To test this hypothesis, we performed ChIP-Chip studies to locate sites of CTCF binding throughout the *Igh* locus in primary pro-B cells from Rag^{-/-} mice as well as pre-B cells from RAG^{-/-}μ⁺ mice [38]. If CTCF were the developmental stage-specific factor facilitating *Igh* locus contraction and looping, we expected to see more CTCF binding sites at the *Igh* locus in pro-B cells than pre-B cells, however the CTCF binding pattern was largely similar in both. We also tested CTCF enrichment at selected CTCF sites in MEFs, pre-pro-B, and thymocytes by quantitative PCR and found the levels of CTCF binding were similar across all cell and tissue types. Since CTCF binding was not limited to the stage at which the loci were undergoing locus contraction, then CTCF alone could not be responsible for the developmental stage-specific contraction. Cohesin had been reported to colocalize with many CTCF sites in three genome-wide studies [39-41], so we tested a subset of CTCF sites for cohesin binding using ChIP and quantitative PCR for one of the cohesin subunits, Rad21. Indeed, cohesin displayed a more stage specific binding pattern than CTCF, with increased binding at selected sites in the *Igh* locus during the pro-B cell stage.

We also used ChIP-Chip to investigate binding of CTCF in the *Igκ* and *Igλ* loci and found that CTCF binding in these loci was also present among different cell types, although the binding pattern was more stage-specific than at the *Igh* locus [38]. Like the *Igh* locus, binding of Rad21 to the tested CTCF sites showed lineage and stage specificity. Both the *Igh* and *Igκ* loci have many CTCF binding sites throughout the V portion of the locus, and thus are well suited to serve as bases to facilitate looping and contraction of the loci (Fig.2). In contrast, the small *Igλ* locus, consisting of two small clusters of V-D-J-C genes, only had four CTCF binding sites, each set of sites bracketing the V-D-J-C clusters (Fig. 2).

Other than in the many CTCF binding sites in the V region portion of the *Igκ* and *Igh* loci, there were only 1-2 other regions of binding, respectively. In the *Igh* locus, we confirmed the cluster of CTCF binding sites just downstream of the 3'RR region reported by Birshtein and colleagues [42]. The only other sites outside of the V_H locus were a pair of adjacent sites just upstream of the most 5' D_H gene, DFL16.1. We hypothesized that this pair of sites may insulate or bracket the 3' region of the *Igh* locus containing the 3'RR, constant region genes, Eμ, J_H and D_H genes from the V_H portion of the *Igh* locus. If one wanted to have a boundary that would exclude the V_H genes during the stage of D-J_H rearrangement, these two CTCF binding regions would be in the ideal locations. A loop formed between those two regions of CTCF sites would create a domain excluding the V_H portion of the locus from the D_H-J_H-C_H portion of the locus. This hypothesis is consistent with 3D-FISH studies performed by the Murre lab [34]. The h5 probe used in their study is located very near the CTCF sites that are upstream of DFL16.1. In pre-pro-B cells, this h5 probe and the h2 probe, located upstream of the 3'RR cluster of CTCF sites, were in close proximity to each other, but the V_H probes were all positioned far away. However, in pro-B cells the probe near the D_H region was in close proximity with V_H probes and especially close to one of the distal V_H probes (probe h10). In further support of this hypothesis, our preliminary 3C studies indicate that loops are formed between the 3'RR cluster of CTCF sites and the DFL16.1-flanking CTCF sites (unpublished data). We hypothesize that these CTCF sites form a loop that excludes the V_H portion during D_H-J_H rearrangement, and that a new or additional set of loops and interactions would be formed during V_H to DJ_H rearrangement to bring the V_H portion of the locus in proximity with the DJ_H gene segment to create a vast repertoire of VD_H combinations (Fig. 3).

Also in support of this model is an experiment reported by Schlissel and colleagues in 2007 [43]. They targeted a V_H gene segment just 5' of DFL16.1, but inside of the CTCF sites which

flank DFL16.1, and this relocated V_H gene rearranged at the time of D-J rearrangement, thus disrupting normal ordered rearrangement (Fig. 4). Lineage restriction was also disrupted, since this V_H gene rearranged in thymocytes, which only undergo D-J rearrangement, but never V_H to DJ H . These results suggested that the D_H -J H genes and V_H genes reside in functionally separate compartments of the *Igh* locus. Recently, the pair of CTCF sites just upstream of DFL16.1 was found to possess enhancer-blocking activities *in vitro* and there was a sharp decrease in antisense transcription just upstream of these two CTCF sites [44].

In the V_H portion of the *Igh* locus we reported over 50 CTCF binding sites. The CTCF binding sites in the proximal region (V_H 7183/Q52 gene families) were all within a close distance 3' of the recombination signal sequence. In the distal V_H portion of the *Igh* locus all of the CTCF binding sites were intergenic or located farther upstream of the V_H coding regions. The spacing of these CTCF sites throughout the V_H portion of the *Igh* locus is ideal for the formation of the multi-loop 3-D structure during the pro-B stage of B-cell development. In collaboration with Drs. Cornelis Murre and Roy Riblet, we have obtained preliminary results from 3D-FISH studies indicating that there is decrease in *Igh* locus contraction in pro-B cells that have been transduced with shRNA retroviruses knocking down CTCF compared to pro-B cells transduced with control retroviruses (unpublished data).

Another interesting site revealed in our ChIP-Chip studies was the only CTCF binding site in the *Igk* locus outside of the V_k gene region. It was located within a region called Sis, located between the V_k genes and J k genes. Deletion of a 3.7 kb region containing Sis, which also included the CTCF binding sites, by Garrard and colleagues resulted in a decrease in V_k to J k rearrangement and a loss of recruitment of *Igk* alleles to heterochromatin [45]. ChIP assays revealed Ikaros binding in the Sis region as well as the J k C k region downstream of Sis. Since the Sis deletion resulted in the loss of the Sis Ikaros sites but not the nearby J k C k Ikaros sites, these results suggested that Ikaros binding alone was not sufficient to recruit the *Igk* locus to heterochromatin. We hypothesized that the CTCF site near Sis helps to facilitate the recruitment of the *Igk* locus to heterochromatin [38].

There are also many CTCF and Rad21 binding sites throughout the TCR loci (unpublished data). The CTCF/Rad21 binding sites in the *TCR α / δ* locus are especially intriguing. They demarcate some of the boundaries between regions with different epigenetic profiles [46]. Most interestingly, there are many more CTCF/Rad21 sites occupied in the 5' half of the locus in double negative thymocytes compared to double positive thymocytes. These CTCF/Rad21 sites could potentially result in locus contraction to allow distal $V\delta$ genes to gain access to the DJ δ genes. In contrast, *TCR α* rearrangement proceeds from proximal to distal $V\alpha$ genes, and the RAG^{-/-} mice with TCR β transgene that we used would be poised for proximal $V\alpha$ gene rearrangement. However, conditional CTCF knockdown in thymocytes does not appear to reduce TCR assembly [47]. Further studies are necessary to resolve this issue.

The Ig and TCR loci harbor many CTCF/Rad21 binding sites that may function differently depending on their location. We speculate that subsets of CTCF/Rad21 sites may serve different functions depending on their locations throughout the loci and interacting protein partners [48-52].

Conclusions

We have come a long way since the early 1980's in understanding the molecular mechanisms that control the lineage-specific and stage-specific regulation of accessibility of Ig and TCR loci, but there are still many unanswered questions. We know that transcription factors such as Pax5 and YY1 are essential for distal V_H gene rearrangement and for IgH locus compaction, but how are they working? How do transcription factors control the preferential rearrangement of certain V genes? We know that some epigenetic marks can have significant consequences,

such as the presence of H3K4me3 on J genes. The H3K4me3 forms recombination centers by recruiting RAG to J genes, and thus helps enforce ordered rearrangements utilizing J genes first. What influence do the other epigenetic differences on V, D and J genes play in regulating V(D)J recombination? How does Ezh2 aid in *Igh* locus contraction? Is it by methylating H3K27 on proximal V genes, and if so, how does that result in locus contraction? Is the CTCF and cohesin complex the major player in locus contraction and looping of receptor loci? Do all loci have similar rules, or are we going to find different transcription factors or proteins influencing rearrangement at different stages of differentiation? It is clear that we are just scratching the surface of the molecular details of the regulation of accessibility for recombination, and about the intersection of these various pathways to control and regulate this highly important function in lymphocyte development.

Acknowledgments

The work presented here was supported by NIH grants R01 AI029672, AI052313, and AI082918 to AJF, and F32 AI084418 to SCD. We thank all present and former members of the Feeney lab for the contributions to the research summarize in this review.

Abbreviations

AcH3	acetylated-histone H3
BCR	B-cell receptor
CTCF	CCCTC-binding factor
ChIP	chromatin immunoprecipitation
E	enhancer
H3K4me2	histone H3 dimethylated at lysine 4
H3K4me3	histone H3 trimethylated at lysine 4
H3K27me3	histone H3 trimethylated at lysine 27
H3K79me2	histone H3 dimethylated at lysine 79
Ig	immunoglobulin
MEF	mouse embryonic fibroblast
PTM	posttranslational modification
RAG	recombination-activating genes
RSS	recombination signal sequences
TCR	T-cell receptor
3C	chromosome conformation capture
3-D	3-dimensional
3-D FISH	3D Fluorescence In Situ Hybridization
3'RR	3' regulatory region

References

- [1]. Yancopoulos GD, Alt FW. Regulation of the assembly and expression of variable-region genes. *Ann Rev Immunol* 1986;4:339–368. [PubMed: 3085692]

- [2]. Alt FW, Yancopoulos GD, Blackwell TK, Wood C, Thomas E, Boss M, et al. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J* 1984;3:1209–1219. [PubMed: 6086308]
- [3]. Engel H, Rolink A, Weiss S. B cells are programmed to activate kappa and lambda for rearrangement at consecutive developmental stages. *Eur J Immunol* 1999;29:2167–2176. [PubMed: 10427979]
- [4]. Van Ness BG, Weigert M, Coleclough C, Mather EL, Kelley DW, Perry RP. Transcription of the unrearranged mouse C_k locus: Sequence of the initiation region and comparison of activity with a rearranged V_k - C_k gene. *Cell* 1981;27:593–602. [PubMed: 6101210]
- [5]. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged V_H gene segments. *Cell* 1985;40:271–281. [PubMed: 2578321]
- [6]. Abarrategui I, Krangel MS. Regulation of T cell receptor-alpha gene recombination by transcription. *Nat Immunol* 2006;7:1109–1115. [PubMed: 16936730]
- [7]. Xu CR, Feeney AJ. The epigenetic profile of immunoglobulin genes is dynamically regulated during B cell differentiation and is modulated by pre-BCR signaling. *J Immunol* 2009;182:1362–1369. [PubMed: 19155482]
- [8]. Morshead KB, Ciccone DN, Taverna SD, Allis CD, Oettinger MA. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. *Proc Natl Acad Sci U S A* 2003;100:11577–11582. [PubMed: 14500909]
- [9]. Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* 2007;27:561–571. [PubMed: 17936034]
- [10]. Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, Ivanov D, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 2007;450:1106–1110. [PubMed: 18033247]
- [11]. Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell* 2010;141:419–431. [PubMed: 20398922]
- [12]. Wood DL, Coleclough C. Different joining region J elements of the murine k immunoglobulin light chain locus are used at markedly different frequencies. *Proc Natl Acad Sci USA* 1984;81:4756–4760. [PubMed: 6431410]
- [13]. Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. *J Exp Med* 1993;177:1009–1020. [PubMed: 8459201]
- [14]. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: An approach by autoreactive B cells to escape tolerance. *J Exp Med* 1993;177:999–1008. [PubMed: 8459227]
- [15]. Tze LE, Schram BR, Lam KP, Hogquist KA, Hippen KL, Liu J, et al. Basal immunoglobulin signaling actively maintains developmental stage in immature B cells. *PLoS Biol* 2005;3:e82. [PubMed: 15752064]
- [16]. Vela JL, Ait-Azzouzene D, Duong BH, Ota T, Nemazee D. Rearrangement of mouse immunoglobulin kappa deleting element recombining sequence promotes immune tolerance and lambda B cell production. *Immunity* 2008;28:161–170. [PubMed: 18261939]
- [17]. Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. Predominant autoantibody production by early human B cell precursors. *Science* 2003;301:1374–1377. [PubMed: 12920303]
- [18]. Romanow WJ, Langerak AW, Goebel P, Wolvers-Tettero IL, van Dongen JJ, Feeney AJ, et al. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. *Mol Cell* 2000;5:343–353. [PubMed: 10882075]
- [19]. Goebel P, Janney N, Valenzuela JR, Romanow WJ, Murre C, Feeney AJ. Localized gene-specific induction of accessibility to V(D)J recombination induced by E2A and early B cell factor in nonlymphoid cells. *J Exp Med* 2001;194:645–656. [PubMed: 11535632]
- [20]. Massari ME, Grant PA, Pray-Grant MG, Berger SL, Workman JL, Murre C. A conserved motif present in a class of helix-loop-helix proteins activates transcription by direct recruitment of the SAGA complex. *Mol Cell* 1999;4:63–73. [PubMed: 10445028]
- [21]. Love VA, Lugo G, Merz D, Feeney AJ. Individual promoters vary in strength, but the frequency of rearrangement of those V_H genes does not correlate with promoter strength nor enhancer independence. *Molecular Immunol* 2000;37:29–39.

- [22]. Montalbano A, Ogwaro KM, Tang A, Matthews AG, Larijani M, Oettinger MA, et al. V(D)J recombination frequencies can be profoundly affected by changes in the spacer sequence. *J Immunol* 2003;171:5296–5304. [PubMed: 14607931]
- [23]. Espinoza CR, Feeney AJ. The extent of histone acetylation correlates with the differential rearrangement frequency of individual V_H genes in pro-B cells. *J Immunol* 2005;175:6668–6675. [PubMed: 16272322]
- [24]. Espinoza CR, Feeney AJ. Chromatin accessibility and epigenetic modifications differ between frequently and infrequently rearranging V_H genes. *Mol Immunol* 2007;44:2675–2685. [PubMed: 17218014]
- [25]. Hesslein DG, Pflugh DL, Chowdhury D, Bothwell AL, Sen R, Schatz DG. Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. *Genes Dev* 2003;17:37–42. [PubMed: 12514097]
- [26]. Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, et al. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol* 2003;4:124–131. [PubMed: 12496962]
- [27]. Corcoran AE, Riddell A, Krooshoop D, Venkitaraman AR. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature* 1998;391:904–907. [PubMed: 9495344]
- [28]. Liu H, Schmidt-Supprian M, Shi Y, Hobeika E, Barteneva N, Jumaa H, et al. Yin Yang 1 is a critical regulator of B-cell development. *Genes Dev* 2007;21:1179–1189. [PubMed: 17504937]
- [29]. Roldan E, Fuxa M, Chong W, Martinez D, Novatchkova M, Busslinger M, et al. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat Immunol* 2005;6:31–41. [PubMed: 15580273]
- [30]. Xu CR, Schaffer L, Head SR, Feeney AJ. Reciprocal patterns of methylation of H3K36 and H3K27 on proximal vs. distal IgVH genes are modulated by IL-7 and Pax5. *Proc Natl Acad Sci U S A* 2008;105:8685–8690. [PubMed: 18562282]
- [31]. Zhang Z, Espinoza CR, Yu Z, Stephan R, He T, Williams GS, et al. Transcription factor Pax5 (BSAP) transactivates the RAG-mediated V(H)-to-DJ(H) rearrangement of immunoglobulin genes. *Nat Immunol* 2006;7:616–624. [PubMed: 16680144]
- [32]. Chowdhury D, Sen R. Stepwise activation of the immunoglobulin mu heavy chain gene locus. *Embo J* 2001;20:6394–6403. [PubMed: 11707410]
- [33]. Skok JA, Gisler R, Novatchkova M, Farmer D, de Laat W, Busslinger M. Reversible contraction by looping of the Tcra and Terb loci in rearranging thymocytes. *Nat Immunol* 2007;8:378–387. [PubMed: 17334367]
- [34]. Jhunjhunwala S, van Zelm MC, Peak MM, Cutchin S, Riblet R, van Dongen JJ, et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell* 2008;133:265–279. [PubMed: 18423198]
- [35]. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 1999;98:387–396. [PubMed: 10458613]
- [36]. Gaszner M, Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nature reviews* 2006;7:703–713.
- [37]. Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell* 2009;137:1194–1211. [PubMed: 19563753]
- [38]. Degner SC, Wong TP, Jankevicius G, Feeney AJ. Cutting Edge: Developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol* 2009;182:44–48. [PubMed: 19109133]
- [39]. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 2008;132:422–433. [PubMed: 18237772]
- [40]. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 2008;451:796–801. [PubMed: 18235444]
- [41]. Rubio ED, Reiss DJ, Welch PL, Distech CM, Filippova GN, Baliga NS, et al. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* 2008;105:8309–8314. [PubMed: 18550811]

- [42]. Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, Li F, et al. Chromatin architecture near a potential 3' end of the igh locus involves modular regulation of histone modifications during B-Cell development and in vivo occupancy at CTCF sites. *Mol Cell Biol* 2005;25:1511–1525. [PubMed: 15684400]
- [43]. Bates JG, Cado D, Nolla H, Schlissel MS. Chromosomal position of a VH gene segment determines its activation and inactivation as a substrate for V(D)J recombination. *J Exp Med* 2007;204:3247–3256. [PubMed: 18056289]
- [44]. Featherstone K, Wood AL, Bowen AJ, Corcoran AE. The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. *J Biol Chem* 2010;285:9327–9338. [PubMed: 20100833]
- [45]. Liu Z, Widlak P, Zou Y, Xiao F, Oh M, Li S, et al. A recombination silencer that specifies heterochromatin positioning and ikaros association in the immunoglobulin kappa locus. *Immunity* 2006;24:405–415. [PubMed: 16618599]
- [46]. Hawwari A, Krangel MS. Regulation of TCR delta and alpha repertoires by local and long-distance control of variable gene segment chromatin structure. *J Exp Med* 2005;202:467–472. [PubMed: 16087716]
- [47]. Heath H, Ribeiro de Almeida C, Sleutels F, Dingjan G, van de Nobelen S, Jonkers I, et al. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *Embo J* 2008;27:2839–2850. [PubMed: 18923423]
- [48]. Defossez PA, Kelly KF, Filion GJ, Perez-Torrado R, Magdinier F, Menoni H, et al. The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. *J Biol Chem* 2005;280:43017–43023. [PubMed: 16230345]
- [49]. Ishihara K, Oshimura M, Nakao M. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol Cell* 2006;23:733–742. [PubMed: 16949368]
- [50]. Chernukhin I, Shamsuddin S, Kang SY, Bergstrom R, Kwon YW, Yu W, et al. CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol Cell Biol* 2007;27:1631–1648. [PubMed: 17210645]
- [51]. Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT. Identification of a Ctfc cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* 2007;25:43–56. [PubMed: 17218270]
- [52]. Weth O, Weth C, Bartkuhn M, Leers J, Uhle F, Renkawitz R. Modular insulators: genome wide search for composite CTCF/thyroid hormone receptor binding sites. *PLoS ONE* 2010;5:e10119. [PubMed: 20404925]

Fig 1.

Histone posttranslational modifications on Ig genes. ChIP was performed on RAG^{-/-} pro-B cells for H3K4me3, acetylated H3, and H3K27me3. H3K4me3 and AcH3 were predominantly on J genes, but not on V or D genes to any large extent [7]. H3K27me3, the histone modification catalyzed by Ezh2, was present on proximal V_H genes but not on distal V_H genes [30]. In the absence of Ezh2, distal V_H genes are unable to undergo V(D)J rearrangement.

Fig. 2. Genomic structure of germline immunoglobulin loci and locations of CTCF sites. Enhancers are depicted by yellow circles, V genes are depicted by blue boxes, and constant regions are represented by red squares. Genomic locations of CTCF sites in $RAG^{-/-}$ pro-B cells (*IgH* locus) and $\mu^{+}RAG^{-/-}$ pre-B cells (*IgK* and *IgL* loci) are depicted by black lines below the schematics of each loci [38].

Fig. 3.

Proposed model of the role of CTCF in V(D)J recombination. Cartoon of germline genomic structure of the *Igh* locus with CTCF sites (green boxes) located throughout the V_H locus, 5' of DFL16.1, and in the 3'RR (top portion of schematic). Loops between CTCF sites located 5' of DFL16.1 and CTCF sites clustered in the 3'RR may form a loop to exclude the V_H portion of the locus during the stage of D_H - J_H recombination (middle portion of schematic). During V_H to D_H - J_H recombination, CTCF/cohesin may facilitate the compaction of the *Igh* locus and the formation of a multi-loop structure, giving all V_H genes access to the D_H - J_H gene segment (lower portion of schematic). Grey rings depict cohesin rings. Yellow circles represent enhancers.

Fig. 4.

CTCF sites near DFL16.1 may form a boundary to insulate the *D_H-C_H-J_H* region of the *Igh* locus from the *V_H* portion of the locus. Middle portion of schematic highlights region of the *Igh* locus located between the CTCF sites 5' of DFL16.1 and CTCF sites in the 3'RR. Lower portion of schematic depicts the targeted insertion of the *V_H* gene segment (blue box) into the region 5' of DFL16.1 in studies by Bates et al [43]. The targeted insertion was just 3' of the two CTCF sites 5' of DFL16.1. Green boxes depict CTCF sites. Yellow circles represent enhancers.