New insights emerge as antibody repertoire diversification meets chromosome conformation [version 1; peer review: 3 approved]

Amy L. Kenter¹, Ann J. Feeney²

¹Department of Microbiology and Immunology, University of Illinois College of Medicine, Chicago, IL, 60612-7344, USA
²Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, 92037, USA

Abstract
Vast repertoires of unique antigen receptors are created in developing lymphocytes. The antigen receptor loci contain many variable (V), diversity (D), and joining (J) gene segments that are arrayed across very large genomic expanses and are joined to form variable-region exons. This process creates the potential for an organism to respond to large numbers of different pathogens. Here, we consider the underlying molecular mechanisms that favor some V genes for recombination prior to selection of the final antigen receptor repertoire. We discuss chromatin structures that form in antigen receptor loci to permit spatial proximity among the V, D, and J gene segments and how these relate to the generation of antigen receptor diversity.

Keywords
antibody genes, B cells, V(D)J recombination
Corresponding author: Amy L. Kenter (star1@uic.edu)

Author roles: Kenter AL: Conceptualization, Writing – Original Draft Preparation; Feeney AJ: Conceptualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by NIH R01AI121286 and R21AI137865 to ALK and R56AI19092, R211137867, and R01AI082918 to AJF.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Kenter AL and Feeney AJ. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Kenter AL and Feeney AJ. New insights emerge as antibody repertoire diversification meets chromosome conformation [version 1; peer review: 3 approved] F1000Research 2019, 8(F1000 Faculty Rev):347 (https://doi.org/10.12688/f1000research.17358.1)

First published: 28 Mar 2019, 8(F1000 Faculty Rev):347 (https://doi.org/10.12688/f1000research.17358.1)
Introduction

In vertebrates, the adaptive immune response is capable of recognizing pathogens using antigen-specific receptors expressed on B and T lymphocytes. The B-cell receptor (BCR) is composed of two identical immunoglobulin (Ig) heavy chains (IgH) and two identical light chains (Igκ or Igλ). There are two lineages of T cells that are distinguished by the type of T-cell receptor (TCR) expressed. TCRαβ is encoded by the Tcra and Tcrb loci, whereas TCRγδ is encoded by the Tcrg and Tcrd loci. Antigen receptors are composed of variable (V) and constant (C) regions. The organization of the Igh and Igκ loci are schematically depicted (Figure 1A and Figure 2A). Igh variable-region exons are produced by the joining of one each of the many variable (V), diversity (D), and joining (J) gene segments, whereas Igκ and Igλ are created by joining one each of the V and J gene segments, all by V(D)J recombination during lymphocyte development (Figure 1B). V(D)J recombination is a stepwise process during which D_J→J_H recombination occurs first.

Figure 1. Overview of the Igh locus. The Igh locus spans 2.9 Mb and contains about 100 V_H gene segments. (A) (Upper panel) Schematic diagram of the Igh locus showing the V_s, D_s, J_s, and C_s exons and regulatory elements (not to scale). The V_H7183 and V_HQ52 families—blue and red bars, respectively (lower panel)—are located at the D_J→J_H-proximal end of the locus. Each D_J→J_H-proximal V_H gene segment is paired with a recombination signal sequence (not shown) and a CTCF-binding element (CBE) (purple triangles). The CBE associated with the V_H81X segment is non-functional (gray triangle). CBE orientation is indicated by the direction of the triangle. V_H gene segment names indicate their position along the locus. V_H81X (V_H5-2) is the original name of the second gene segment relative to intergenic control region 1 (IGCR1) and is used because it is well known by this nomenclature. The intermediate V_H segments include the V_HS107 family along with nine smaller V_H families. At the 5′ end of the locus, the interspersed distal V_H segments are composed of the V_HJ558 and V_H3609 families. Regulatory elements include intronic E_μ and 3′E_α super-enhancers and IGCR1, which is composed of two divergent CBEs. A cluster of at least nine CBEs is located at the 3′ boundary of the Igh locus and is adjacent to 3′E_α. The 3′CBEs and 3′E_α are referred to as the 3′ regulatory region (3′RR). Sites I, II, and III (red circles) engage in exceptionally long-range looping interactions and may mediate locus compaction. Sub-topologically associating domain (Sub-TADs) A, B, and C are indicated. (B) Diagram of the stepwise process of V(D)J recombination. D-J rearrangement precedes V-DJ recombination. (C) A schematic of the Igh TAD in pro-B cells that is subdivided into three sub-TADS A, B, and C. Looping interactions between Eμ:3′Eα and Eμ:IGCR1 (black arcs), Sites I and II, Sites II and III, Sites I to III (red arcs), Site I-FρOSIa, and Site II-FρOSTb (blue arcs) were detected and are not described here in detail.
followed by V<sub>μ</sub>-to-D,I<sub>μ</sub> rearrangement. This process depends on the lymphocyte-specific V(D)J recombinase, RAG1/2, which recognizes recombination signal sequences (RSSs) that flank all V, D, and J gene segments<sup>3</sup>. During V(D)J recombination, two RSSs adjacent to V, D, or J gene segments partner such that cleavage and rejoining occur. RAG1 contains endonuclease activity and targets the RSS, and RAG2 is recruited to the epigenetically modified histone 3 when it is trimethylated on lysine 4<sup>3</sup>. Within each antigen receptor locus, the RAG recombinase concentrates in the recombination center (RC) that is focused to the J segment–containing domain. Double-strand DNA breaks are generated at RSSs by RAG1/2, and the V, D, and J exons are joined together through non-homologous end joining<sup>4</sup>. Antigen receptor gene rearrangement is tightly regulated during lymphocyte development; in turn, lymphocyte development is strictly dependent on V(D)J recombination<sup>5,6</sup>. The composition and complexity of antigen receptor repertoires depend on the number of V, D, and J gene segments and the degree to which those segments are available for rearrangement. However, V gene usage in the pre-selected Igh repertoire is only quasi-random since it has been shown that V genes rearrange at very different intrinsic frequencies<sup>7–14</sup>. No one factor, or combination of factors, could fully account for unequal V gene usage in studies considering V germline transcript levels, transcription factor (TF) binding, RSS quality, and the distribution of a variety of epigenetic marks<sup>2,5,7,11,14</sup>. Hence, the mechanisms underlying V gene rearrangement frequencies remain to be determined. Antigen receptor loci are quite large spanning 0.67 Mb-3.0 Mb and containing up to 100 functional V genes. Chromatin conformational changes in antigen receptor loci are important determinants for long-distance V(D)J recombination events<sup>6</sup>. Developmental stage-specific contraction of Ig and TCR loci promotes proximity of J-distal V genes with (D-)J segments and generally is thought to facilitate recombination but this has not been formally proven<sup>16–19</sup>. Two currently unresolved questions in the formation of the antigen receptor repertoires are (1) what is the molecular basis for locus contraction that is hypothesized to support V->DJ recombination over exceptionally long genomic distances and (2) what underlies the unequal rearrangement potential of individual V genes? Here, we focus on the murine Igh and Igκ loci to address these questions. The molecular principles resulting from these studies may be generally applicable to all antigen receptor loci.

**Figure 2. Three-dimensional conformation of the Igκ locus.** The Igκ locus spans 3.2-Mb topologically associating domain (TAD) and contains about 120 functional Vκ gene segments. (A) Schematic diagram of the Igκ locus showing the Vs, Js, and C exons and regulatory elements (not to scale). Regulatory elements include intronic Ex (iEx), 3′Eκ, Ed, and E88 elements. Contracting element for recombination (Cer) and silencer in the intervening sequence (Sis) are located between the V and J domains and are composed of CTCF-binding elements (CBEs) (purple triangles). The orientation of each CBE is indicated. The Igκ locus is subdivided into five sub-TADs (A–E) as indicated. (B) Sub-TAD structure of the Igκ locus as determined by Hi-C<sup>2</sup>. Each loop represents a sub-TAD that is labeled A–E. The regulatory region containing the Jκ genes, the three distal enhancers, and the constant region are in gray. (C) Deletion of E88 results in untethering of sub-TADs C and D from the regulatory region.
transcription\textsuperscript{16,20–29}. Additionally, the Igh locus undergoes large-scale locus contraction during development that is detected by using three-dimensional (3D) DNA fluorescence in situ hybridization (FISH) methods\textsuperscript{24,25}.

The early observations by Kosak revealed two fundamental findings regarding the disposition of the Igh locus in the nucleus\textsuperscript{31}. First, the Igh locus is located at the nuclear periphery in non-B cells and relocates to the nuclear center at the pro-B cell stage\textsuperscript{34} through a process that requires active dislocation from the nuclear lamina\textsuperscript{35}. Second, the Igh locus is in an extended conformation in non-B cells and lymphoid progenitors, whereas both Igh alleles are contracted in pro-B cells, a developmental stage coincident with V(D)J recombination\textsuperscript{34,35}. These pioneering studies have led to the recognition that all of the large antigen receptor loci undergo developmentally regulated conformational changes before rearrangement at that locus\textsuperscript{34,35,27–30}. The contracted Igh locus in pro-B cells undergoes decontraction at the pre-B cell stage of development to prevent a second round of V\textsubscript{H}-D\textsubscript{J\textsc{H}} rearrangement on the second Igh allele, presumably allowing allelic exclusion\textsuperscript{36}. Degrees of locus compaction have been inferred from the relationship of inter-probe nuclear distances derived from 3D DNA FISH versus genomic distances and these measurements have limited resolution (100–1000 nm). Consequently, it has been difficult to identify DNA elements that mediate locus contraction.

Igh locus contraction depends on the TFs Pax5, Ikaros, and YY1\textsuperscript{31,32}. Loss of Igh locus compaction is correlated with preferential usage of the most D\textsubscript{H}-proximal V\textsubscript{H} genes\textsuperscript{25,31,33}, indicating that spatial access to the more distally positioned V\textsubscript{H} gene segments has been lost. Although deletion of any of these TFs reduces distal V\textsubscript{H} rearrangement, chromatin accessibility remains unchanged\textsuperscript{33,34,35}. Low-level transcription over V\textsubscript{H} genes and intergenic regions occurs as the locus is preparing to undergo rearrangement\textsuperscript{34–37}. The highest level of non-coding RNA (ncRNA) in the Igh locus is found at elements called Pax5-activated intergenic repeats (PAIRs) and these ncRNAs are dependent upon the presence of Pax5 and YY1\textsuperscript{15,38}. Although the function of TFs in locus contraction remains speculative, PAIR elements have been suggested to induce long-range chromatin looping by relocating to transcription factories where they associate with the 3’ proximal Eq-J\textsubscript{H}-D\textsubscript{H} domain\textsuperscript{39}. However, the molecular mechanism that mediates locus contraction remains unclear.

The Igh locus is conformationally distinct in pro-B cells

Eukaryotic chromosomes are organized into higher-order spatial configurations of multiple-length scales as determined by using high-resolution chromosome conformation capture (3C)-based approaches and microscopy-based methods, including 3D DNA FISH and live cell imaging\textsuperscript{30–32}. For example, insulators and enhancers often engage in physical interactions with their target promoters\textsuperscript{33–35}, indicating that regulatory elements can control distant gene expression through direct long-range molecular contact. However, not all long-range chromatin interactions are directed toward regulating gene expression. For example, intra-chromosomal interactions are required to regulate V(D)J recombination and Ig class-switch recombination (CSR)\textsuperscript{36,37,32}. In CSR, the constant (C\textsubscript{H})-region exons encoding IgM are substituted with a downstream C\textsubscript{H} gene such that IgM is no longer produced and instead IgG, IgE, or IgA is made in conjunction with the original recombined variable-region exons. CSR is dependent on 3D chromatin architecture mediated by long-range intra-chromosomal interactions between distantly located transcriptional elements\textsuperscript{35–38}. During V(D)J recombination, antigen receptor genes undergo ordered rearrangement with D\textsubscript{H}-to-J\textsubscript{H} joining preceding V\textsubscript{H}-to-D\textsubscript{H} recombination\textsuperscript{1}. To produce a fully representative Ig repertoire, it is essential that the distal V\textsubscript{H} genes achieve spatial proximity with the RC and D\textsubscript{H}J\textsubscript{H} domain. Murre and colleagues have shown that Igh locus topology is best described as a series of three large chromatin loops joined by linkers in pre-pro-B cells but that these loops have intermingled and provide equal access of the D\textsubscript{H}-distal and -proximal V\textsubscript{H} gene segments with rearranged 3’ D\textsubscript{H}J\textsubscript{H} in pro-B cells\textsuperscript{37}. The time interval for D\textsubscript{H}J\textsubscript{H} to gain proximity with a V\textsubscript{H} gene segment is on the order of minutes, and spatial confinement of topological domains largely regulates first-passage times for chromatin interactions in vivo\textsuperscript{38}. Although it is clear that Igh locus conformation is structured, the DNA elements that anchor chromatin looping in support of V(D)J recombination remain largely undefined.

The Igh and Ig\textsubscript{\kappa} loci are configured as topologically associating domains

Topologically associating domains (TADs) are megabase sized and represent regions of high-frequency self-interacting chromatin contacts as defined in 3C-based studies\textsuperscript{50,60}. The organization of interphase chromatin is largely conserved between cell types, especially with regard to TAD boundaries\textsuperscript{54,58,81}. Strikingly, the Igh locus is contained within a 2.9-Mb TAD in pro-B cells\textsuperscript{1}.

The murine Igh TAD is partitioned into two highly structured sub-TADs A and C—corresponding to the D\textsubscript{H}-proximal and D\textsubscript{H}-distal V\textsubscript{H} gene families, respectively—and flank a less structured sub-TAD B that includes the intermediate V\textsubscript{H} gene segments\textsuperscript{1} (Figure 1C). Sub-TADs are zones within a TAD in which chromatin contacts are more frequent than with sites outside the sub-domain, and contacts can be tissue-specific and can contribute to the overall architectural structure of the TAD\textsuperscript{9,62,63}.

V genes can be subdivided into V families based on sequence relatedness and this reflects gene duplication and divergence of primordial V genes. The correspondence of sub-TAD structure with the murine V\textsubscript{H} gene family distribution profile is striking (Figure 1C). In the murine Igh locus, V\textsubscript{H} families tend to be clustered, but in most other antigen receptor loci and in other species, the members of individual V\textsubscript{H} families generally are interspersed.

The Ig\textsubscript{\kappa} locus is also contained within a 3.5-Mb TAD which is subdivided into five sub-domains\textsuperscript{1} (Figure 2A, B). However, because Vk gene families are interspersed across the locus, there is no correspondence between Vk families and sub-TAD structure. One unique feature of the Ig\textsubscript{\kappa} locus is that about one
third of Vκ genes are present in the reverse orientation such that they rearrange to Jκ genes by inversion as opposed to the predominant deleterional rearrangement found at other antigen receptor loci. However, there is no correlation between sub-TAD structure and the inversive or deleterional orientation of Vκ genes. The conservation of Iγκ TAD and sub-TAD structure has not been examined in different cell types.

**Igh TAD conformation is sculpted by developmentally specific chromatin looping**

TADs can be thought of as scaffolds for constitutive architectural interactions. Nevertheless, interactions within TADs may vary significantly between cell types or developmental stages and for private enhancer–promoter contacts. Igh sub-TADs A, B, and C become juxtaposed in pro-B cells via megabase-scale chromatin looping but these contacts are absent in non-B cells. The loop anchors located in sub-TAD A and C are termed sites I, II, and III (Figure 1C). Our FISH studies indicated that sites I, II, and III participate in three-way physical contacts in about 32% of pro-B cells and in less than 5% of non-B cells and may functionally contribute to the proximal Vκ47 domain with the RC/Dκ/Jκ region to facilitate efficient access of all Vκ genes for recombination1.

The structure of sub-TAD A is worthy of additional consideration as it contains the Eμ and 3′ε enhancers, the RC, located in the Jκ-Dκ domain, intergenic control region 1 (IGCR1) (an insulator which will be discussed in detail in sections below), and the proximal Vκ47 genes (Figure 1A). Sub-TAD A becomes modified in pro-B as compared with non-B cells. In non-B cells, Igh sub-TAD A encompasses the proximal Vκ47 genes spanning from site I to Eμ (Figure 1C). In pro-B cells, sub-TAD A becomes subsumed within a larger topological fold that extends from site I to the 3′ε enhancer (Figure 1C). The higher-order chromatin structure in pro-B cells may have significant implications for Dκ-proximal Vκ47 gene usage during V(D)J recombination.

Several earlier observations have shown that Dκ-proximal Vκ47 genes are regulated differently from the rest of the Vκ47 genes. Although distal Vκ47 gene recombination is reduced in Pax5-/-, YY1-/-, and Ezh2-deficient pro-B cells, Dκ-proximal Vκ47 genes recombine normally25,31-33. Thus, the localization of the Dκ-proximal Vκ47 genes within the same conformational sub-TAD as the RC/Dκ/Jκ region distinguishes them from distal Vκ47 genes that lie within sub-TADs B and C.

Pax5 organizes sub-TAD C that spans the distal Vκ-Jκ558 gene family. Site III within sub-TAD C fails to associate with sites I and II in Pax5-deficient pro-B cells, thus providing a possible explanation for reduced Vκ-Jκ558 rearrangements in Pax5-deficient pro-B cells. Notably, 14 PAIR elements that were proposed to mediate locus compaction via Pax5 are all situated within sub-TAD C, and PAIR motifs 10 and 11 overlap with site III.38 PAIR elements are bound by the TFs Pax5, E2A, and CTCF (CCCTC-binding factor) in pro-B cells38. It is not known whether transcriptional activity at PAIR elements regulates chromatin looping. Our studies provide a potential molecular definition of locus contraction by identifying loop anchor sites that are key mediators of this process.

**CTCF mediates insulator function at TAD boundaries**

TAD boundaries are frequently enriched for CTCF binding and CTCF-binding elements (CBEs)45,55,60,68. CTCF is a ubiquitously expressed zinc-finger protein that binds DNA, functions as an insulator in vertebrates96, and plays a key role in chromatin looping45,50,53,72. There is an observed inward or convergent orientation of CBEs flanking TADs63,70,73. Insulators were originally defined as genomic elements that act as a barrier to position effects caused by the spreading of chromatin marks and they block enhancer activity57,59. Although loci situated within TADs are relatively insulated from loci outside the domain, these same elements readily interact with other loci within the same domain. CRISPR/Cas9-mediated rearrangements of TAD boundaries and regulatory elements facilitate or prevent looping interactions with distal regulatory elements76-79. Acute depletion of CTCF leads to loss of loop domains and impaired regulation of nearby genes through loss of enhancer insulation59.

High-resolution in situ Hi-C studies demonstrated that mammalian genomes are partitioned into context domains45. Context domains with end points that anchor a loop are referred to as loop domains49,50. TADs are most frequently loop domains but not all loop domains are TADs. In the context of V(D)J recombination, RAG recombinase activity was shown to be confined to loop domains that are defined by convergent CTCF-bound elements. RAG primarily initiates double-stranded breaks (DSBs) at RSSs within the antigen receptor loci. However, RAG can also initiate low-frequency DSBs at off-target sites that have sequence similarity to RSSs and cause chromosomal rearrangements and translocations82,90-92. Notably, when RAG was experimentally directed to chromosomal domains outside of antigen receptor loci, off-target DSBs were confined within loop domains and deletion of convergent CBEs extended the range of RAG activity83.

**CTCF partners with cohesin to mediate chromatin looping**

CTCF-based long-range looping interactions are dependent on co-binding with cohesin84,85. The cohesin complex is thought to form a ring around two CTCF proteins bound to DNA83,86. Different combinations of architectural proteins may mediate context-specific genomic organization57,67. Promoter–enhancer interactions are disrupted in embryonic stem cells86 and in thymocytes88 when cohesin is depleted. There is a rich CTCF-cohesin landscape in the Iγκ locus. One hundred thirty-two sites are bound by CTCF and cohesin and the majority of these are located at a distance of 1 to 32 kb from Vκ genes segments in the Iγκ locus90,91. Strikingly, all of the rearranging Dκ-proximal Vκ genes are closely paired with CBEs that are located within 68 base pairs (bp) of the RSS (Figure 1A)90. However, CBEs in the non-rearranging Dκ-proximal Vκ genes are located more than 1 kb from the RSS in the two most Dκ-proximal Vκ genes families. As described below, close proximity to the adjacent CBE has functional consequences for these Vκ genes90,91. In addition, a cluster of nine CBEs marks the 3′ boundary of the Iγκ TAD92, and two CBEs located within IGCR1 mark the boundary between the Dκ-Jκ domain and the Dκ-proximal Vκ genes (Figure 1B)93,94. Similarly, the Tcrb and Tcrr loci have CBEs located between the V and J gene segments95,96. In the Iγκ
locus, CBEs, termed contracting element for recombination (Cer) and silencer in the intervening sequence (Sis), are located between the V and J gene segments, and many CBEs are found throughout the Vκ domain (Figure 2A)\(^{91,97,98}\).

In the Igh and TCRβ loci, all bound CTCF sites in the V exon domains upstream of the D-J-C-regions are oriented toward them, and the CBEs in D-J-C regions of those loci are oriented toward the V exons. In contrast, the other two large antigen receptor loci (TCRα/δ and Igκ) have more complex patterns with the bound CTCF sites in the V gene portion of each locus found in both orientations\(^{91}\). A role for the CTCF-cohesin complex in Igh locus looping has been suggested by shRNA knockdown studies in pro-B cells demonstrating that the Igh locus is less contracted after CTCF is knocked down\(^{99}\).

**A convergence of loop extrusion and directional RAG tracking?**

TADs and loop domains have been implicated in regulating gene expression in mammalian cells\(^{40,86}\), with convergent CBEs in a large subset of cases\(^{21,73,101}\). It has been proposed that TADs can be formed by the loop extrusion activity of cohesin (Figure 3A)\(^{100,102}\). When cohesin is bound to chromatin, it forms a progressively larger loop until it encounters an obstacle formed by another cohesin or boundary protein including CTCF (Figure 3B). The association of CTCF with widely separated convergent CBEs may involve cohesin that is halted upon arriving at convergent CTCF-bound loop anchors\(^{40,101,102}\). It has been proposed that loop extrusion may also facilitate close-range contacts between regulatory elements, including promoters and enhancers, by bringing them into molecular contact\(^{40}\) (Figure 3C). Promoter–enhancer interactions may preferentially occur within chromatin domains that are insulated by extrusion blocking factors.

In a situation strikingly analogous to convergent CBE-mediated loop formation, RAG-dependent recombination involves interactions between distant convergent RSSs with the exception of inverted RSSs in some antigen receptor loci. The Alt group has shown that RAG off-target activity within CTCF loop domains spanning 2 Mb depends on orientation-specific RSSs\(^83\). It was inferred from DNA sequencing data that RAG can travel directionally from a physiological or ectopically introduced RC within a convergent CBE-based loop domain of megabase size\(^83\). Long-range directional exploration by RAG can be blocked by an encounter with cohesin-bound convergent CBE pairs and possibly by other impediments that create chromatin sub-domains within TADs\(^82,104\). Alt and colleagues proposed that RAG complexes bind one RSS and then track along the chromatin fiber in a linear fashion to the next convergent RSS\(^82,104\). Several different topological machine models have been postulated to explain directional cis-guided long-range looping interactions\(^40\). It remains unclear whether RAG tracking occurs via loop extrusion or by a mechanistically different activity.

**IGCR1 is an insulator that partitions the D\(_H\)J\(_H\) domain from V\(_H\) genes**

CTCF has been implicated as a mediator of transcriptional insulation through its ability to participate in chromatin looping\(^71\). The striking number and organization of CBEs across antigen receptor loci have led to the proposal for a role of CBE in V(D)J recombination\(^71,99\). The Igh sub-TAD A contains several important looping contacts, including Eμ-IGCR looping interactions in pro-B cells (Figure 1C)\(^93,106\). IGCR1 contains a

---

**Figure 3. Loop extrusion as a topologically associating domain generating machine.** (A) The chromatin fiber extrudes over time through an extruding factor (possibly cohesin; yellow cylinders). (B) A boundary element (possibly CTCF; green cube) can block loop extrusion when the CTCF-binding element is in the proper orientation. It has been proposed that CTCF can block extrusion by one of the cohesin extruding motors while the second motor will be unobstructed and continue to extrude the loop\(^{102}\). (C) Regulatory elements may come into close molecular contact by the process of loop extrusion. These interactions will occur only within a topologically associating domain and in the presence of extrusion blocking elements.
pair of divergent CBEs that demarcate the boundary of the RC/ $D_{91}^G$ domain and function as an insulator that prevents $D_{91}^G$-to-$V_{H1}$ joining prior to $D_{91}^G$ rearrangements\cite{140, 144}. However, the relationship of CTCF-anchored chromatin looping for $V_{H1}$-IGCR1 and antigen receptor rearrangement frequency remains unclear.

$V_{H1}$ CBEs are convergently oriented with respect to the upstream IGCR1 CBE, and $3'$ CBEs are convergently oriented relative to the downstream IGCR1 CBE\cite{143} (Figure 1A). Although CBE-dependent $Eμ$:IGCR1 looping is prominent in pro-B cells, it is striking that $Eμ$-$V_{H1}$81X contacts are largely undetectable in the wild-type context, indicating that the RC located between $Eμ$ and IGCR1 is sequestered away from all $V_{H1}$ genes\cite{148, 149}. The functional $D_{91}^G$-proximal $V_{H1}$ genes are very closely paired with CBEs\cite{146}. For example, $V_{H1}$81X is the first functional proximal $V_{H1}$ gene located about 100 kb from IGCR1 and it is immediately adjacent to a CBE (Figure 1A). Two new studies have shown that when the IGCR1 CBEs are deleted, $Eμ$-$V_{H1}$81X contacts are newly observed, indicating that IGCR1 CBEs prevent looping interactions between the $Eμ$-RC/$D_{91}^G$ domain and the proximal $V_{H1}$ genes\cite{148, 149}. Strikingly, IGCR1-$V_{H1}$81X interactions are dependent on the $V_{H1}$81X CBE, as shown by deletion of the $V_{H1}$81X-flanking CBE.

Interestingly, the Igκ locus contains the Cer/Sis CBEs in the V-J intervening region (Figure 2A). Deletion or inversion of Cer leads to preferential usage of $κ$-proximal $V_{κ}$ genes\cite{148, 149}, highlighting the importance of convergent CTCF-mediated long-range interactions that facilitate spatial proximity of the distal $V_{κ}$ with the J segments. Cer/Sis and IGCR1 are similarly located between the $V_{κ}$ genes and the (D)J genes, and both are involved in mediating chromatin looping. Cer also functions as a transcriptional insulator\cite{148}.

**Proximal $V_{H1}$ gene rearrangement frequencies are determined by CTCF looping**

To begin, one might expect that the $V_{H1}$5-1 gene segment would be highly used in $V_{H1}$5-to-$D_{91}^G$ rearrangements since it is most proximal to the RC/$D_{91}^G$ domain (Figure 1A). However, despite being paired with a highly conserved RSS, $V_{H1}$5-1 is not used in V(D)J recombination. In contrast, $V_{H1}$81X ($V_{H1}$5-2), the next $V_{H1}$ gene segment along the genome, is the most frequently used in V(D)J recombination. The question of why $V_{H1}$81X and not $V_{H1}$5-1 is used is long-standing. Two groups have explored the relationship between CTCF-mediated chromatin looping and proximal $V_{H1}$ gene usage during V(D)J recombination\cite{148, 149}.

CBEs adjacent to the functional $D_{91}^G$-proximal $V_{H1}$ genes are found within 68 bp downstream of the RSSs\cite{146}. Mutagenesis analyses have revealed that proximal $V_{H1}$ CBEs dramatically influence the frequency of V(D)J rearrangement of that $V_{H1}$ gene\cite{148, 149}. Mutation of the CBE associated with $V_{H1}$81X ($V_{H1}$5-2) (Figure 1A) greatly reduced both looping with IGCR1 and its rearrangement frequency and boosted the rearrangement frequency of the next most upstream $V_{H1}$ gene, $V_{H1}$2-2\cite{148}. Genomic editing of the non-functional $V_{H1}$5-1 CBE into a functional motif turns this non-rearranging $V_{H1}$ gene into the most frequently rearranging gene\cite{149} (Figure 1A). Thus, as discussed below, CBE quality and chromatin looping between IGCR1 and the $D_{91}^G$-proximal VH gene segments are significant factors determining $V_{H1}$ gene usage in V(DJ) recombination.

The antigen receptor loci have a much higher density of CTCF sites than the genome overall, making CTCF/cohesin a candidate for forming multiple long-range loops within these loci\cite{150, 151}. Although it is clear that TAD boundaries are usually formed between convergent CBEs\cite{152, 153}, relatively little is known regarding the CBE orientation dependence in anchoring chromosome loops within the V domains of Ig loci. All of the bound CBEs in the $V_{H1}$ domain are oriented toward the $3'$ regulatory region (3'RR) and a single CBE within IGCR1 (Figure 1A). If CTCF-mediated looping occurs only between convergent CBE, one would predict that the orientation of motifs adjacent to proximal $V_{H1}$ genes will be critically required for looping and V(DJ) rearrangement. However, when the $V_{H1}$81X CBE was inverted, usage of $V_{H1}$81X in V(DJ) rearrangement was only modestly decreased\cite{149}, indicating that the orientation specificity inside the $V_{H1}$ sub-TADs is not strictly required.

Together, these studies demonstrate that the proximal $V_{H1}$ gene CBE’s quality determines looping efficiency with IGCR1 and determines that $V_{H1}$ gene’s recombination efficiency. It is noteworthy that most $V_{H1}$ and all $Vκ$ genes do not have any CTCF sites in close proximity, in contrast with the location of CBEs for the proximal $V_{H1}$ genes\cite{150, 151}. Thus, for the majority of $Vκ$ genes, CBE-mediated looping with IGCR1 may have a less straightforward impact on $V_{H1}$ gene rearrangement frequency.

**Vκ rearrangement frequency is determined by enhancer E88**

In addition to long-range loops mediated by CTCF, other long-range loops can be enhancer-mediated. The Igκ locus is encompassed within a TAD that is subdivided into at least five sub-TADs A–E based on Hi-C studies (Figure 2A)\cite{154}. We identified a novel enhancer element, E88, which is located close to the boundary separating sub-TADs C and D which becomes active at the pro-B cell stage prior to V-J rearrangement (Figure 2A)\cite{154}. E88 is the major site of interaction with $κ$Ex detected by 4C analyses in pro-B cells. In pre-B cells, the stage at which V-J recombination occurs, E88 continues to interact strongly with $κ$Ex and also contacts many more sites throughout the locus (Figure 2B). Strikingly, deletion of E88 results in significant changes in long-range looping interactions and in reduction in rearrangement levels of adjacent $Vκ$ genes (Figure 2C). Its deletion also results in a modest but consistent reduction of rearrangement of almost all $Vκ$ genes in a 1.5 Mb region surrounding E88 that corresponds to sub-TADs C and D (Figure 2C). Most $Vκ$ genes that are upstream and downstream of sub-TADs C and D—located in sub-TADs A and B and sub-TADs E, respectively—were modestly increased in relative rearrangement frequency\cite{154}. Thus, our studies revealed the novel concept that $Vκ$ rearrangement is regulated in a domain-specific manner and suggest that sub-TAD structure has functional ramifications.
Future questions

Chromatin conformation is now recognized as an important feature regulating gene expression and recombination. Although locus contraction has been a recognized feature of antigen receptor loci for more than 15 years, its underlying molecular mechanism remains largely undefined. Recent studies have provided new insights regarding the conformation of chromatin, TAD and sub-TAD structure, and CTCF-cohesin-mediated looping with V(D)J recombination. These studies revealed that the conformational organization of the IgH, Igκ, and TCRTαβ loci has significant implications for locus contraction and likely influences skewed V gene usage that together affects the composition of the pre-selected repertoires. Going forward, studies focused on the relationship of CTCF- and promoter-enhancer-mediated chromatin looping with locus contraction are likely to provide new insights. Studies designed to clarify the relationship of CTCF-dependent looping and Dij-distal V gene rearrangement will be important. It is likely that new enhancers, similar to Igκ E88, will be characterized. The influence of individual enhancers on the frequency of individual V gene usage during initial repertoire formation will be important. The emergence of extremely high-resolution DNA FISH is likely to provide additional insights into locus conformation. Finally, studies that determine the extent to which the pre-selected repertoire determines the shape of the peripheral repertoire will yield new insights.

Abbreviations

3C, chromosome conformation capture; 3D, three-dimensional; bp, base pairs; C, constant; CBE, CTCF-binding element; Cer, contracting element for recombination; CSR, class-switch recombination; CTCF, CCCTC-binding factor; D, diversity; DSB, double-stranded break; FISH, fluorescence in situ hybridization; Ig, immunoglobulin; IgκR1, intergenic control region 1; IgH, immunoglobulin heavy chain; J, joining; ncRNA, non-coding RNA; PAR, Pax5-activated intergenic repeat; RC, recombination center; RSS, recombination signal sequence; Sis, silencer in the intervening sequence; TAD, topologically associating domain; TCR, T-cell receptor; TF, transcription factor; V, variable

Grant information

This work was supported by NIH RO1 AI121286 and R21 AI137865 to ALK and R56 AI19092, R21137867, and R01 AI082918 to AJF.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Montefiori L, Wuerffel R, Roqueiro D, et al.: Extremely Long-Range Chromatin Loops Link Topological Domains to Facilitate a Diverse Antibody Repertoire. Cell Rep. 2016; 14(4): 896–906. Published Abstract | Publisher Full Text | Free Full Text
2. Barajas-Mora EM, Kleinman E, Xu J, et al.: A B-Cell-Specific Enhancer Orchestrates Nuclear Architecture to Generate a Diverse Antigen Receptor Repertoire. Mol Cell. 2010; 78(1): 48–62.e45. Published Abstract | Publisher Full Text
3. Teng G, Schatz DG: Regulation and Evolution of the RAG Recombinase. Adv Immunol. 2015; 128: 1–39. Published Abstract | Publisher Full Text
4. Ali F, Zhang Y, Meng FL, et al.: Mechanisms of programmed DNA lesions and genomic instability in the immune system. Cell. 2013; 152(3): 417–429. Published Abstract | Publisher Full Text
5. Krangel MS: Mechanics of T cell receptor gene rearrangement. Curr Opin Immunol. 2009; 21(2): 133–139. Published Abstract | Publisher Full Text | Free Full Text
6. Bosser G, Mansson P, Murre C: Chromatin topology and the regulation of antigen receptor assembly. Annu Rev Immunol. 2012; 30: 337–356. Published Abstract | Publisher Full Text | Free Full Text
7. Choi NM, Loguercio S, Verma-Gaur J, et al.: Deep sequencing of the murine IgH repertoire reveals complex regulation of nonrandom V gene rearrangement frequencies. J Immunol. 2013; 191(5): 2393–2402. Published Abstract | Publisher Full Text | Free Full Text
8. Bolland DJ, Kooby H, Wood AL, et al.: Two Mutually Exclusive Local Chromatin States Drive Efficient V(D)J Recombination. Cell Rep. 2016; 15(11): 2475–2487. Published Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation
9. Kleinman E, Loguercio S, Feeney AJ: Epigenetic Enhancer Marks and Transcription Factor Binding Influence Vκ Gene Rearrangement in Pre-B Cells and Pro-B Cells. Front Immunol. 2018; 9: 2074. Published Abstract | Publisher Full Text | Free Full Text
10. Aoki-Ota M, Tokumaru A, Ota T, et al.: Skewed primary Igκ repertoire and V-J joining in C57BL/6 mice: implications for recombination accessibility and receptor editing. J Immunol. 2012; 188(5): 2305–2315. Published Abstract | Publisher Full Text | Free Full Text
11. Matheson LS, Bolland DJ, Chovanec P, et al.: Local Chromatin Features including PUL1 and IKAROS Binding and H3K4 Methylation Shape the Repertoire of Immunoglobulin Kappa Genes Chosen for V(D)J Recombination. Front Immunol. 2017; 8: 1550. Published Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation
12. Williams GS, Martinez A, Montabano A, et al.: Unequal Vκ gene rearrangement frequency within the large Vκ7183 gene family is not due to recombination signal sequence variation, and mapping of the genes shows a bias of rearrangement based on chromosomal location. J Immunol. 2001; 167(1): 257–263. Published Abstract | Publisher Full Text
13. Feeney AJ, Goebel P, Espinoza CR: Many levels of control of V gene rearrangement frequency. Immunol Rev. 2004; 200: 44–56. Published Abstract | Publisher Full Text
14. Feeney AJ, Lugo G, Escuro G: Human cord blood kappa repertoire. J Immunol. 1997; 158(8): 3761–3768. Published Abstract
15. Degner-Leisso SC, Feeney AJ: Epigenetic and 3-dimensional regulation of V(D)J rearrangement of immunoglobulin genes. Semin Immunol. 2010; 22(6): 346–352. Published Abstract | Publisher Full Text | Free Full Text
16. Kume N, G, Gen R: Chromatin Interactions in the Control of Immunoglobulin Heavy Chain Gene Assembly. Adv Immunol. 2015; 128: 41–92. Published Abstract | Publisher Full Text
17. Ebert A, Hill L, Bussinger M: Spatial Regulation of V(D)J Recombination at Antigen Receptor Loci. Adv Immunol. 2015; 128: 93–121. Published Abstract | Publisher Full Text
18. Carico Z, Krangel MS: Chromatin Dynamics and the Development of the TCR and TCRα Repertoires. Adv Immunol. 2015; 128: 307–361. Published Abstract | Publisher Full Text
19. Majumder K, Bassing CH, Ottz EM: Regulation of Trb β Gene Assembly by Genetic, Epigenetic, and Topological Mechanisms. Adv Immunol. 2015; 128: 273–306. Published Abstract | Publisher Full Text
20. Feasterstone K, Wood AL, Bowen AJ, et al.: The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. J Biol Chem. 2010; 285(13): 9327–9338. Published Abstract | Publisher Full Text | Free Full Text
21. Garrett FE, Emelyanov AV, Sepulveda MA, et al.: Chromatin architecture near a potential 3’ end of the Igκ locus involves modular regulation of histone modifications during B-Cell development and in vivo occupancy at CTCF sites. Mol Cell Biol. 2005; 25(4): 1511–1525. Published Abstract | Publisher Full Text | Free Full Text
22. Maes J, Chappaz S, Cavelier P, et al.: Activation of V(D)J recombination at the Igh chain JH locus occurs within a 6-klbase chromatin domain and is associated with nucleosomal remodeling. J Immunol. 2006; 176(9): 5409–5417. PubMed Abstract | Publisher Full Text | Free Full Text

23. Chowdhury D, Sen R: Stepwise activation of the immunoglobulin mu heavy chain gene locus. EMBO J. 2001; 20(22): 6394–6403. PubMed Abstract | Publisher Full Text | Free Full Text

24. Kosak ST, Skok JA, Medina KL, et al.: Subnuclear compartmentalization of immunoglobulin mu in pro-B cell differentiation. Science. 2002; 296(5565): 158–162. PubMed Abstract | Publisher Full Text | F1000 Recommendation

25. Fuxa M, Skok J, Souabni A, et al.: Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev. 2004; 18(4): 411–422. PubMed Abstract | Publisher Full Text | F1000 Recommendation

26. Zullo JM, Demarco IA, Piqué-Regí R, et al.: DNA-sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. Cell. 2012; 148(7): 1473–1487. PubMed Abstract | Publisher Full Text | F1000 Recommendation

27. Sayegh CE, Juhnrunwala S, Ribelet R, et al.: Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. Genes Dev. 2005; 19(9): 933–947. PubMed Abstract | Publisher Full Text | Free Full Text

28. Roldán E, Fuxa M, Chowdhury D, et al.: Locus ‘deconstruction’ and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat Immunol. 2006; 7(July 1): 31–41. PubMed Abstract | Publisher Full Text | Free Full Text

29. Chen L, Carico Z, Shin HY, et al.: A discrete chromatin loop in the mouse Tcrα-Tcrδ locus shapes the TCRI and TCRII repertoires. Nat Immunol. 2015; 16(10): 1085–1093. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

30. Liu H, Schmidt-Suppian M, Shi Y, et al.: Yin Yang 1 is a critical regulator of B-cell development. Genes Dev. 2007; 21(10): 1179–1189. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

31. Reynaud D, Demarco IA, Reddy KL, et al.: Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. Nat Immunol. 2008; 9(July 8): 927–936. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

32. Hesslein DG, Pfuhl DL, Chowdhury D, et al.: Pax5 is required for recombination of transcribed, acetylated, 5′IgV gene segments. Genes Dev. 2003; 17(1): 37–42. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

33. Yancopoulos GD, DePinho RA, Zimmerman KA, et al.: Secondary genomic rearrangement events in pre-B cells: VH/JH replacement by a LINE-1 sequence and directed class switching. Ikaros. Nat Immunol. 2000; 1(9): 927–936. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

34. Bolland DJ, Wood AL, Johnston CM, et al.: Antisense intergenic transcription in B cell recombination. Nat Immunol. 2004; 5(5): 630–637. PubMed Abstract | Publisher Full Text | F1000 Recommendation

35. Bolland DJ, Wood AL, Afshar R, et al.: Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer EγL. Mol Cell Biol. 2007; 27(15): 5255–5263. PubMed Abstract | Publisher Full Text | Free Full Text

36. Verma-Gaur J, Torkamani A, Schaffer L, et al.: Comprehensive mapping of long range interactions reveals folding principles of the human genome. Science. 2009; 326(5950): 289–293. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

37. Hao SS, Huntley MH, Durand NC, et al.: A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014; 159(7): 1665–1680. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

38. Shachar S, Voss TC, Pegoraro G, et al.: Identification of Gene Positioning Factors Using High-Throughput Imaging Mapping. Cell. 2015; 162(4): 911–923. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

39. Tang Z, Luo OJ, Li X, et al.: CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell. 2015; 163(7): 1611–1627. PubMed Abstract | Publisher Full Text | Free Full Text

40. Carter D, Chakalova L, Osborne CS, et al.: Long-range chromatin regulatory interactions in vivo. Nat Genet. 2002; 32(4): 623–626. PubMed Abstract | Publisher Full Text | F1000 Recommendation

41. Li G, Ruan X, Auerbach RK, et al.: Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell. 2012; 148(1–2): 84–98. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

42. Kumar S, Wuerffel R, Achour I, et al.: Flexorable ordering of antibody class switch and V(D)J joining during B-cell ontogeny. Genes Dev. 2013; 27(22): 2439–2444. PubMed Abstract | Publisher Full Text | Free Full Text

43. Feldman S, Achour I, Wuerffel R, et al.: Constraints contributed by chromatin looping limit recombination targeting during Ig class switch recombination. J Immunol. 2015; 194(5): 2380–2388. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

44. Wuerffel R, Wang L, Grigera F, et al.: S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. Nat Immunol. 2017; 18(5): 555–563. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

45. Sanyal A, Lajice BR, Jain G, et al.: The long-range interaction landscape of gene promoters. Nature. 2012; 489(7414): 109–113. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

46. Sellars M, Reina-San-Martin B, Kastner P, et al.: Ikaros controls isotype selection during immunoglobulin class switch recombination. J Exp Med. 2009; 206(5): 1073–1087. PubMed Abstract | Publisher Full Text | Free Full Text

47. Eisert A, Buchholz U, Schröter S, et al.: Analysis of hundreds of primary cells reveals nuclear organization of the genome. J Exp Med. 2013; 207(6): 1625–1640. PubMed Abstract | Publisher Full Text | Free Full Text

48. Phillips-Cremins JE, Sauria ME, Sanyal A, et al.: Long-range 3D nuclear organization of developmental lineages. Cell. 2017; 169(2): 321–333. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

49. Nora EP, Lajice BR, Schulz EG, et al.: Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012; 485(7398): 381–385. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

50. Dixon JR, Selvaraj S, Yue F, et al.: Identification of gene positioning factors using high-resolution capture Hi-C. Nature. 2012; 489(7414): 376–382. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

51. Hanssen LLP, Kassouf MT, Oudelaar AM, et al.: Topological biology: domain organization of the genome. Cell. 2013; 153(6): 1281–1295. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

52. Phillips-Cremins JE, Saucier ME, Phillips-Cremins JE, et al.: Molecular maps of the organization of nuclear-nuclear lamina interactions during differentiation. Mol Cell. 2013; 52(4): 631–642. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

53. Weinreb S, Ben-Yehuda D, Greenberg H, et al.: Chromatin fiber structure and the effect of external forces on long-range association. Nature. 2012; 489(7414): 383–388. PubMed Abstract | Publisher Full Text | Free Full Text | F1000 Recommendation

54. Hughes JR, Roberts N, McGowan S, et al.: Analysis of hundreds of...
Open Peer Review

Current Referee Status: ✔ ✔ ✔

Editorial Note on the Review Process
F1000 Faculty Reviews are commissioned from members of the prestigious F1000 Faculty and are edited as a service to readers. In order to make these reviews as comprehensive and accessible as possible, the referees provide input before publication and only the final, revised version is published. The referees who approved the final version are listed with their names and affiliations but without their reports on earlier versions (any comments will already have been addressed in the published version).

The referees who approved this article are:

Version 1

1 Yehudit Bergman Department of Developmental Biology and Cancer Research, Hebrew University of Jerusalem, Jerusalem, Israel
   Competing Interests: No competing interests were disclosed.

2 Harry W Schroeder Department of Genetics, University of Alabama at Birmingham, Birmingham, Alabama, USA
   Competing Interests: No competing interests were disclosed.

3 Craig H Bassing Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, 19104, USA
   Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com