Published in final edited form as:

Trends Genet. 2016 May ; 32(5): 253–255. doi:10.1016/j.tig.2016.03.003.

The Ties that Bind (the *Igh* Locus)

Michael S. Krangel^{1,*}

¹Department of Immunology, Duke University School of Medicine, Durham, NC 27708, USA

Abstract

Immunoglobulin heavy-chain locus V(D)J recombination requires a 3D chromatin organization which permits widely distributed variable (V) gene segments to contact distant diversity (D) and joining (J) gene segments. A recent study has identified key nodes in the locus interactome, paving the way for new molecular insights into how the locus is configured for recombination.

Recent years have seen major advances in our understanding of the 3D organization of the eukaryotic genome, and of the significance of genome organization for genome function [1]. To a large extent, this has been made possible by the application of chromosome conformation capture (3C)-based technologies, which use chemical crosslinking, ligation, and either PCR or deep-sequencing, to report the frequencies of interactions between DNA fragments that may be widely separated in linear chromatin [2]. These studies have shown that the genome is organized into a series of discrete and relatively stable megabase-scale topologically associating domains (TADs), which help to segregate the genome into functional units, either active or repressive, depending on cell type. Long-distance interactions within TADs are frequent and dynamic, physically connecting enhancers to their target promoters and providing the structural framework for gene expression. CTCF and cohesin have been implicated as major players in loop organization, tethering the bases of loops that create TAD boundaries and similarly contributing to fine-structure looping within TADs [3].

Adaptive immunity in jawed vertebrates depends on the ability of T and B lymphocytes to create highly diverse and clonally distributed antigen receptor (AgR) repertoires [4]. This is accomplished through a process of somatic DNA recombination which, in individual lymphocyte precursors, assembles a unique variable exon from the many V, D, and J gene segments that are distributed across AgR loci. V(D)J recombination is initiated by the lymphocyte-specific recombinase activating gene (RAG) proteins, which bind to the recombination signal sequences that flank V, D, and J gene segments and create DNA double-strand breaks. RAG proteins bind abundantly to relatively compact portions of AgR loci containing D and J gene segments, the so-called 'recombination centers' (RCs) (Figure 1), which are established by enhancer activation of D and J segment promoters [4]. Because V segments are typically distributed across several megabases of DNA, and can be distant from the RC, V segment recombination requires V segments to be brought into proximity of the RC in nuclear space. Moreover, for repertoires to be appropriately diverse, in a

*Correspondence: krang001@mc.duke.edu (M.S. Krangel).

population of developing lymphocytes, V segments located anywhere in the V array must have an opportunity to interact with the RC. Given both the complexity of AgR locus gene regulation by widely distributed enhancer and promoter elements, and the need to juxtapose distant gene segments for recombination and repertoire development, the 3D organization of antigen receptor loci has attracted tremendous interest.

The best studied of all AgR loci is the immunoglobulin heavy-chain (*Igh*) locus, which spans nearly 3 Mb on mouse chromosome 12 (Figure 1). 3D fluorescence *in situ* hybridization (3D-FISH) has shown that the locus exists in a relatively extended conformation in pre-pro-B cells, but adopts a more contracted conformation in pro-B cells, in which long-range V_H-to-DJ_H recombination occurs. At the population level, this conformational change is thought to provide all V_H segments the opportunity to contact D_H and J_H segments [5]. Nevertheless, how this happens at a molecular level is only partly understood.

The transcriptional and architectural proteins CTCF, Pax5, and YY1 have been implicated as important mediators of long-distance interactions at the *Igh* locus. Based on 3C and circular chromosome conformation capture (4C), CTCF binding to the intergenic control region 1 (IGCR1) and to the 3' regulatory region (3'RR), which flank the D_H-J_H-C_H portion of the locus, as well as YY1 binding to enhancer E_μ, were shown to create a three-way network of interactions that organizes the 3' portion of the locus [6–9]. CTCF also binds to many sites scattered across the V_H portion of the locus, and was shown to organize relatively local interactomes within this region [6]. Pax5 binds together with CTCF to a series of transcribed Pax5-activated intergenic repeat (PAIR) elements scattered among the distal V_H segments (Figure 1). YY1 also binds to some PAIR elements and to other sites in the V_H array [8]. Notably, deficiency in Pax5 or YY1, or elimination of E_μ, was shown to disrupt locus contraction [6,9]. Because deficiency in Pax5 or YY1 also disrupts long-distance 4C interactions between E_μ and distant V_H sites, these interactions were suggested to mediate, or at least reflect, the contracted *Igh* conformation [6,8,9]. Notably, the structural organization discerned in these studies is biologically significant because IGCR1 deletion, Pax5 deficiency, and YY1 deficiency each generated proximally biased V_H repertoires.

Montefiori *et al.* [10] have recently shed new light on *Igh* locus structure by addressing a key limitation of all prior structural studies, namely, that interactomes were interrogated from only a limited number of investigator-selected viewpoints. Rather than relying on 3C and 4C, Montefiori *et al.* [10] used unbiased approaches, leveraging available Hi-C data and performing chromosome conformation capture carbon copy (5C).

This allowed the first comprehensive view of intralocus interactions at high resolution. The authors showed that the *Igh* locus is wholly contained within a pro-B cell-specific TAD-like structure, and within this region they defined three subdomains in pro-B cells: A, extending 0.7 Mb from the 3'RR through the proximal 7183 family V_H segments; B, a poorly structured 0.5 Mb region containing intermediate S107 V_H segments; and C, spanning the entire 1.2 Mb array of J558 V_H segments (Figure 1). Further, they defined key nodes in the interactome, most notably site (I) at the 5' end of domain A, which not only contributes to the domain A interactome but also interacts over very long distances with multiple sites

bordering and within domains B and C (Figure 1). Moreover, by 3D-FISH they showed that three key nodes (I, II, III) colocalize in pro-B cells, consistent with their central role in locus contraction. Colocalization was independent of E μ , but depended on Pax5, which, presumably by binding to PAIR elements, helps to organize domain C.

An important conclusion from this work is that *Igh* contraction in pro-B cells is a multi-layered process, with Pax5 regulating the conformation of the C subdomain. E μ and YY1 presumably configure subdomain A and, together with Pax5, contribute to interactions between subdomains. Looking to the future, the unprecedented view of the *Igh* interactome provided by this work should pave the way for molecular analyses of the newly identified nodes and reveal how they are tied together to facilitate V_H-to-DJ_H recombination in pro-B cells.

Acknowledgments

The author is supported by National Institutes of Health grants R37 GM41052 and AI49934, and thanks D. Dauphars and L. Chen for their comments on this article.

References

1. Gibcus JH, Dekker J. The hierarchy of the 3D genome. *Mol. Cell.* 2013; 49:773–782. [PubMed: 23473598]
2. de Wit E, de Laat W. A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* 2012; 26:11–24. [PubMed: 22215806]
3. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* 2014; 15:234–246. [PubMed: 24614316]
4. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* 2011; 11:251–263. [PubMed: 21394103]
5. Jhunjhunwala S, et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell.* 2008; 133:265–279. [PubMed: 18423198]
6. Guo C, et al. Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell.* 2011; 147:332–343. [PubMed: 21982154]
7. Guo C, et al. CTCF-binding elements mediate control of V(D)J recombination. *Nature.* 2011; 477:424–430. [PubMed: 21909113]
8. Medvedovic J, et al. Flexible long-range loops in the V_H gene region of the *Igh* locus facilitate the generation of a diverse antibody repertoire. *Immunity.* 2013; 39:229–244. [PubMed: 23973221]
9. Gerasimova T, et al. A structural hierarchy mediated by multiple nuclear factors establishes IgH locus conformation. *Genes. Dev.* 2015; 29:1683–1695. [PubMed: 26302788]
10. Montefiori L, et al. Extremely long-range chromatin loops link topological domains to facilitate a diverse anti body repertoire. *Cell Rep.* 2016; 14:896–906. [PubMed: 26804913]

Trends in Genetics

Figure 1. Chromatin Domains and Long-Range Interactions at the Immunoglobulin Heavy-Chain (*Igh*) Locus

V(D)J recombination at the *Igh* locus requires a locus conformation which allows distant V_H gene segments to be brought into proximity of D_H and J_H gene segments in the RC. The 2.9 Mb *Igh* locus is contained within a single TAD-like structure in pro-B cells as determined by Hi-C (light blue). Three sub-TADS defined by 5C (A, B, C; medium blue) roughly segregate the locus into domains containing distinct families of V_H gene segments. Note that subdomain A is smaller in non-B cells because it includes E_μ but does not extend to the 3'RR. Red circles denote major (large circles) and minor (small circles) nodes of pro-B cell-specific interactions newly described in this study. Pro-B cell-specific interactions are indicated by red and orange arcs, with orange arcs identifying the subset of interactions shown by 3D-FISH to be dependent upon transcription factor Pax5. Scattered among the distal V_HJ558 gene segments are 14 PAIR elements. For simplicity, interactions described in previous studies are not depicted. The left end of the locus is referred to as the 3' end based on the transcriptional orientation of gene segments (from right to left). Abbreviations: 3D-FISH, 3D fluorescence *in situ* hybridization; 3'RR, 3' regulatory region; 5C, chromosome conformation capture carbon copy; E_μ , enhancer μ ; IGCR1, intergenic control region 1; PAIR, Pax5-activated intergenic repeat; RC, recombination center; TAD, topologically associating domain.