

CTCF-cohesin complex: architect of chromatin structure regulates V(D)J rearrangement

Ann J Feeney¹, Jiyoti Verma-Gaur¹

¹The Scripps Research Institute, Department of Immunology and Microbial Science, IMM22, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA

Cell Research (2012) 22:280–282. doi:10.1038/cr.2011.188; published online 22 November 2011

The CTCF/cohesin complex regulates higher order chromatin structure by creating long-range chromatin loops and by insulating neighboring genes from each other. The lymphocyte antigen receptor loci have large numbers of CTCF/cohesin binding sites, and recent studies demonstrate that the CTCF/cohesin complex plays several important roles in regulating the process of V(D)J recombination at these megabase-sized receptor loci.

The diverse antibody and T cell receptor (TCR) repertoires are created through the highly regulated process of V(D)J recombination. These lymphocyte receptor loci are comprised of many V (variable), D (diversity) and J (joining) gene segments. For example, in the immunoglobulin heavy chain locus (*Igh*), there are ~100 functional V_H genes spread over 2.4 Mb, 13 D_H genes and 4 J_H genes. In pro-B cells, the first rearrangement is a D_H to J_H rearrangement on both alleles, followed by V_H to DJ_H rearrangement. After heavy chain rearrangement is complete, several rounds of proliferation occur, followed by rearrangement in the light chain locus during the pre-B cell stage. A similar precise order of events, linked

with sequential T cell differentiation stages, occurs for the rearrangement of TCR genes.

In addition to the specific order of rearrangement, V(D)J rearrangement is also regulated in lineage-specific manner. Although the same RAG recombinase recombines both TCR and *Ig* genes, the rearrangements only occur in the correct lineage of cells, other than some D_H - J_H rearrangement in thymocytes. This lineage and developmental stage specificity of V(D)J rearrangement suggests precise control over the accessibility of various portions of the *Ig* and TCR loci such that only the gene segments that should recombine at that developmental stage will be accessible to the RAG recombinase. As demonstrated many years ago, this accessibility is manifest by the presence of germline transcription from unrearranged genes at the stage when that particular group of gene segments is poised to undergo rearrangement [1].

Given the enormous size of the receptor loci, the question arises as to how a single DJ rearrangement can have approximately equal access to V genes spread over a 2–3 Mb region, since we know that V genes throughout the loci are utilized. Early clues to this issue came from three dimensional (3D) FISH studies that demonstrated that the large receptor loci compact at the appropriate stage for rearrangement, and that the

loci extend again after rearrangement is completed [2, 3]. The identification of the proteins responsible for these large-scale 3D structural changes in lymphocyte receptor loci has been area of great interest. For the *Igh* locus, the transcription factors YY1 and Pax5 have been shown to be required for locus compaction [3]. Deficiency of these proteins also leads to reduced rearrangement of distal V_H J558 gene family, suggesting that locus compaction is necessary to bring distal V_H genes in proximity to the DJ_H rearrangement. However, the manner in which these transcription factors effect the change in the 3D structure of the locus is still unclear.

In addition to transcription factors, the involvement of proteins that regulate higher order chromatin structure and nuclear architecture has been speculated [4]. One such protein is CTCF, an 11-zinc finger protein associated with all vertebrate insulators, and it has been demonstrated to produce large-scale looping within β -globin and other loci, often with the aid of cohesin [5]. Cohesin is a complex of 4 subunits, which forms a ring around sister chromatids, holding them together during mitosis. ChIP-chip studies have demonstrated that cohesin is bound to a subset of CTCF sites genome-wide [6]. Indeed, the binding of CTCF and cohesin at many sites in the *Igh* locus has also been

Correspondence: Ann J Feeney
Tel: +1-858-784-2979
E-mail: feeney@scripps.edu

demonstrated, and knockdown of CTCF led to reduced *Igh* locus compaction [4, 7]. In addition to the many CTCF sites throughout the V_H part of the locus, only 2 other regions of CTCF binding are present in the *Igh* locus: one set of 2 sites just upstream of the most 5'-functional D_H gene (DFL16.1), and a set of 9 sites downstream of the most distal 3'-enhancer [4, 7]. These sites flank the region containing all the functional D_H and J_H genes and the intronic enhancer $E\mu$. Thus, it was proposed that these sites create a loop or domain in which the first step of V(D)J rearrangement, that of D_H to J_H , would take place [4, 8, 9], and this was recently demonstrated by chromosome conformation capture (3C) [7]. This domain would not include the V_H genes, and thus would help enforce ordered rearrangement.

In a recent issue of *Nature*, Alt and colleagues tested the role of the two closely spaced CTCF sites upstream of DFL16.1 by making a 4.1 kb germline deletion of them, or mutating them in the germline [10]. Both sets of mice showed similar phenotypes. The level of rearrangement of the proximal V_H 7183 and V_H Q52 families was greatly increased, whereas the rearrangement frequency of the distal V_H J558 family was reduced. Sequencing revealed that > 90% of the V_H 7183 rearrangements were to the most 3'-functional V_H gene, 81X. Likewise, > 85% of the V_H Q52 rearrangements were to the most 3'-functional V_H Q52 gene, Q52.2.4. Importantly, ordered rearrangement was also disrupted, in that V_H 81X now underwent rearrangement to a D_H gene that had not previously rearranged to a J_H gene. In agreement with the greatly heightened rearrangement of these two V_H genes, Guo *et al.* demonstrated that the level of germline transcription for these proximal V_H families was significantly increased. In addition, deletion of these CTCF sites resulted in loss of lineage specificity, with rearrangement of proximal V_H genes in deficient thymocytes. These findings

support the hypothesis that interaction of CTCF sites upstream of DFL16.1 with those flanking the 3' RR region by loop formation creates a separate functional domain containing D_H and J_H regions along with $E\mu$, and that this loop excludes the V_H region. Thus, it appears that this 4 kb region with two CTCF sites regulates and restrains the activity of the D_H -proximal V_H genes that are located ~100 kb upstream. In its absence, upregulated proximal V_H gene germline transcription, upregulated proximal V_H gene rearrangement, and inappropriate rearrangement to unrearranged D genes, or in the wrong lineage (thymocytes), are observed.

Another recent paper in *Nature* also demonstrated the role of the CTCF/cohesin complex in VDJ rearrangement. Merckenschlager and colleagues utilized mice in which the cohesin component Rad21 was conditionally deleted in the CD4⁺CD8⁺ double positive (DP) stage of thymic differentiation, the stage at which TCR α rearrangement takes place [11]. Since cells cannot survive cell division without cohesin, the authors cleverly chose to analyze DP thymocytes, which do not undergo cell division during TCR α rearrangement, thereby studying a cell division-independent role for cohesin. Firstly, they mapped cohesin (Rad21) binding in TCR α locus by ChIP-sequencing in DP thymocytes, and their analysis showed that cohesin was abundant at many key positions in the TCR α locus. The TCR α locus is different in structure from all other receptor loci, in that there are 61 *Ja* gene segments, with a similar number of *V α* gene segments. The reason for this is that T cells often undergo multiple rounds of rearrangements before they successfully pass the positive selection step of differentiation. This successive series of rearrangements is regulated by having the initial rearrangements occurring between one of a cluster of *V α* -proximal *Ja* genes and one of a cluster of *Ja*-proximal *V α* genes. Subsequent rearrangements utilize more upstream

V α and downstream *Ja* genes. The various clusters of *Ja* genes each have their own germline promoter that regulates transcription of ~10 downstream *Ja* genes. The first round of rearrangement occurred normally in these conditional Rad21 deletion mice, since Rad21 was not fully deleted initially. However, Rad21 was deleted for all subsequent rounds of rearrangement, and these rearrangements were substantially impaired. Seitan *et al.* showed that normally the TEA promoter upstream of the *Ja* genes forms a long-range loop with the downstream enhancer *E α* , presumably via the cohesin bound to both. In agreement with this hypothesis, the Rad21-deleted thymocytes had greatly reduced TEA-*E α* interactions. The authors demonstrated significantly decreased germline transcription through the middle and distal *Ja* genes and decreased levels of H3K4me3 on those genes. This histone modification is of significance for V(D)J rearrangement since RAG2 is recruited to H3K4me3. Hence, rearrangement to the middle and distal *V α* genes was also reduced. RNA-seq demonstrated that transcription of the downstream gene *Dad1* was increased while germline transcription of *C α* was decreased, presumably due to the loss of insulator function of the CTCF/cohesin sites demarcating those domains, and loss of TEA-*E α* promoter-enhancer interactions. Therefore, cohesin controls TCR α locus rearrangement at multiple levels.

Together these two recent papers in *Nature* further strengthen the hypothesis that the CTCF/cohesin complex plays an important role in the 3D chromatin structure within the *Igh* and TCR α loci by creating domains and insulation boundaries. This structural and functional compartmentalization within the large antigen receptor loci is critical to maintain the appropriate accessibility of gene segments in the complex process of V(D)J recombination. The evidence that the CTCF/cohesin complex regulates germline transcription and his-

tone modifications further provides an example of locus-specific roles of this ubiquitous complex. Thus, the highly ordered, lineage-specific process of V(D)J recombination has yet another layer of regulation imposed by long-range 3D structures.

References

- 1 Yancopoulos GD, Alt FW. Regulation of the assembly and expression of variable-region genes. *Ann Rev Immunol* 1986; **4**:339-368.
- 2 Roldan E, Fuxa M, Chong W, *et al.* Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat Immunol* 2005; **6**:31-41.
- 3 Fuxa M, Skok J, Souabni A, Salvagiotto G, Roldan E, Busslinger M. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev* 2004; **18**:411-422.
- 4 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.
- 5 Splinter E, Heath H, Kooren J, *et al.* CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* 2006; **20**:2349-2354.
- 6 Parelho V, Hadjur S, Spivakov M, *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 2008; **132**:422-433.
- 7 Degner SC, Verma-Gaur J, Wong TP, *et al.* CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci USA* 2011; **108**:9566-9571.
- 8 Degner-Leisso SC, Feeney AJ. Epigenetic and 3-dimensional regulation of V(D)J rearrangement of immunoglobulin genes. *Semin Immunol* 2010; **22**:346-352.
- 9 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.
- 10 Guo C, Yoon HS, Franklin A, *et al.* CTCF-binding elements mediate control of V(D)J recombination. *Nature* 2011; **477**:424-430.
- 11 Seitan VC, Hao B, Tachibana-Konwalski K, *et al.* A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* 2011; **476**:467-471.