DR RHYS ALLAN (Orcid ID : 0000-0003-0906-2980)

Article type : News and Commentary

ICB NC article

Strapline: RAG scanning during D-J recombination

Kinks in the chain: examining RAG scanning during V(D)J recombination

Timothy M Johanson[1,2] & Rhys S Allan[1,2]

[1]The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia.

[2]Department of Medical Biology, The University of Melbourne, Parkville, VIC 3010, Australia.

Joint Corresponding Authors: johanson@wehi.edu.au, rallan@wehi.edu.au

Running head: RAG scanning during D-J recombination

A diverse antigen receptor repertoire is key to adaptive immunity. This diversity relies on the recombination of three pools of gene segments within all antigen receptor loci [1-4]. These pools are known as variable (V), diversity (D) and joining (J). Each antigen receptor loci contains a different number and arrangement of the gene segments. For example, the mouse immunoglobulin heavy chain (*Igh*) locus contains 113 V_H segments, 9 D_H segments and 4 J_H segments covering a vast 2.4 million base pairs (Figure 1), while T cell receptor beta (*Tcrb*) contains 35 V_b, 2 D_b and 14 J_b segments covering a smaller 670 thousand base pairs [5].

Figure 1 about here

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/IMCB.12293

Mechanistically, antigen receptor loci recombination involves successive rounds of double stranded DNA break creation followed by repair and removal of the intervening DNA. In the case of *Igh* the first stage cuts and links two D_H and J_H segments, before this new couple is combined with a V_H segment [6]. The result is a (nearly) randomly generated antigen receptor gene.

This series of recombination events is facilitated by the Recombination Activating Genes (RAG) complex that binds, digests and facilitates repair of the DNA as part of a complex known as a recombination centre (RC) [7-11]. While the mechanism of RAG-mediated cleavage and repair are well understood, how the complex is able to find segments of DNA that can be separated by millions of base pairs is less clear.

A clue as to how RAG may achieve this remarkable frisk of vast loci comes from another complex that has recently been demonstrated to scan vast genomic distances in search of select regions – the cohesin complex. As such, the cohesin complex binds a region of DNA bound by CCCTC-binding factor (CTCF). While one subunit associates with CTCF and anchors the complex in place the other begins drawing the DNA strand through its hoop structure. The extruded DNA then forms a three-dimensional DNA structure anchored by the cohesin complex. The extrusion continues until a second CTCF bound region is drawn into the cohesin complex. Importantly, the orientation of the pair of CTCF bound sequences is critical. Only CTCF sites with convergent orientations will stop the process. Interestingly, RAG mediated V(D)J recombination also relies on the orientation of recombination signal sequences (RSS), that flank all V, D and J gene segments [1].

A number of recent papers have suggested that the RAG complex may function in a similar fashion to the cohesin complex, scanning and extruding antigen receptor loci for convergently orientated RSS [12-15]. However, a recent paper in *Nature* by Zhang *et al.* [16], provides the clearest evidence to date that RAG linearly scans the *Igh* loci during V(D)J recombination.

In a series of increasingly dramatic CRISPR-mediated modifications to the *Igh* locus, Zhang *et al.* show that the orientation of RSS is important in D-J recombination and that RAG scanning can be impeded by dCas9 binding.

First, Zhang *et al.* simply alter the orientation of the two RSS flanking the naturally most utilised D gene segment (DFL16.1). Quantifying recombination frequency using high-throughput genome-wide

translocation sequencing of V(D)J recombination (HTGTS V(D)J seq) they show that the convergent orientation of the flanking RSS to the RSS at the J segment (within the RAG containing recombination centre) is important for appropriate recombination. This suggests that RAG scans or searches the *Igh* locus in a linear fashion upstream from the RC. This conclusion is bolstered by similar results when the RSS flanking seven of the eight other D segments between DFL16.1 and the RC are inverted simultaneously.

Next, Zhang *et al.* used CRISPR to remove all J segments from the *Igh* locus. Interestingly, this forces the establishment of a RC at the closest D segment (DQ52) in place of the absent Js. Critically for the further examination of RAG scanning, this new RC scans not only upstream D segments but also downstream into hitherto unscanned regions. It is this downstream scanning, and subsequent aberrant recombination, that the authors quantify and manipulate to provide further evidence that RAG searches DNA linearly. For example, if dCas9 is bound to the Sy1 region downstream of the RC at DQ52 recombination between DQ52 and a further downstream region (Sy2) is dramatically reduced. This suggests that the presence of dCas9 on the DNA impedes the extrusion of DNA by the RAG complex.

In an elegant molecular flourish, the authors then quantify the level of DNA-DNA interaction between the RC at DQ52 and nearby regions using 3C-HTGTS. With no dCas9 at Sy1 the 3C-HTGTS shows clear peaks of interaction between the RC and Sy2 and other downstream regions, suggesting the RAG complex rapidly scans the intervening regions before stopping, as expected, at transcribed (Sy2) or CTCF bound (CBE) regions. Strikingly, the binding of dCas9 at Sy1 creates a new peak of interaction between this region and the RC, clearly demonstrating the obstruction of the RAG complex by DNA bound dCas9.

Collectively, these experiments provide the clearest evidence to date that RAG, as part of the RC, searches or scans DNA in a linear fashion for a partner sequence to the bound RSS. However, there are still many questions to be answered. For example, is the RAG complex exerting force on the DNA to drive scanning? If not, which components of the RC are? Furthermore, why is the blocking of the RAG complex by transcription or dCas9 not absolute and what are the factors that influence the magnitude of obstruction? Are these factors important in other antigen receptor loci recombination? Finally, while all the outlined experiments suggest that RAG scans the genome, none definitively show this as the mechanism. Similar to experiments performed with the condensin

complex [17], incontrovertible evidence for RAG scanning can only be obtained by isolating RAG or the RC complex and filming it extruding DNA *in vitro*.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

1 Sen R. Access Guide to Antigen Receptor Genes. *J Immunol* 2017; **199**: 3-4.

2 Ebert A, Hill L, Busslinger M. in *Advances in Immunology* Vol. 128 (ed Cornelis Murre) pp. 93-121. (Academic Press, 2015).

3 Aiden EL, Casellas R. Somatic Rearrangement in B Cells: It's (Mostly) Nuclear Physics. *Cell* 2015; **162**: 708-711

4 Chaumeil J, Skok JA. The role of CTCF in regulating V(D)J recombination. *Curr Opin Immunol* 2012; **24**: 153-159.

5 Proudhon C, Hao B, Raviram R, Chaumeil J, Skok J.A. Long-Range Regulation of V(D)J Recombination. *Adv Immunol* 2015; **128**: 123-182.

6 Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* 2013; **152**, 417-429.

7 Desiderio, S. Temporal and spatial regulatory functions of the V(D)J recombinase. *Semin Immunol* 2010; **22**: 362-369.

8 Helmink BA, Sleckman BP. The response to and repair of RAG-mediated DNA double-strand breaks. *Annu Rev Immunol* 2012; **30**: 175-202.

9 Ji Y, Resch W, Corbett E, *et al.* The *in vivo* pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell* 2010; **141**: 419-431.

10 Matthews AG, Oettinger MA. RAG: a recombinase diversified. *Nat Immunol* 2009; **10**: 817-821.

11 Puebla-Osorio N, Zhu C. DNA damage and repair during lymphoid development: antigen receptor diversity, genomic integrity and lymphomagenesis. *Immunol Res* 2008; **41**: 103-122.

12 Hu JZ, Zhang Y, Zhao LJ, *et al.* Chromosomal Loop Domains Direct the Recombination of Antigen Receptor Genes. *Cell* 2015; **163**: 947-959.

13 Jain S, Ba Z, Zhang Y, Dai HQ, Alt FW. CTCF-Binding Elements Mediate Accessibility of RAG Substrates During Chromatin Scanning. *Cell* 2018; **174**: 102-116 e114.

14 Qiu X, Kumari G, Gerasimova T, *et al.* Sequential Enhancer Sequestration Dysregulates Recombination Center Formation at the IgH Locus. *Mol Cell* 2018; **70**: 21-33 e26.

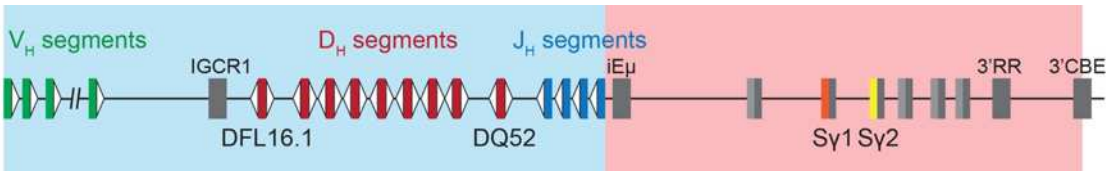
15 Zhao L, Frock RL, Du Z, *et al.* Orientation-specific RAG activity in chromosomal loop domains contributes to Tcrd V(D)J recombination during T cell development. *J Exp Med* 2016; **213**: 1921-1936.

16 Zhang Y, Zhang X, Ba Z, *et al.* The fundamental role of chromatin loop extrusion in physiological V(D)J recombination. *Nature* 2019; in press ADD VOL no. and PAGE NOS AT PROOF STAGE doi:10.1038/s41586-019-1547-y.

17 Ganji M, Shaltiel IA, Bisht S, *et al.* Real-time imaging of DNA loop extrusion by condensin. *Science* 2018; **360**: 102-105.

Figure 1 caption

Figure 1. The *Igh* locus. Schematic of the murine *Igh* locus (not to scale) showing relative position of V_H (green), D_H (red) and J_H (blue) gene segment pools, including D_H segments DFL16.1, DQ52 and constant regions promoters Sy1 (orange) and Sy2 (yellow). Blue shading represents the region normally scanned by the RAG complex. Pink shading represents the region scanned by the RAG complex when the J_H region is removed by CRISPR.



imcb_12293_f1.png