Published in final edited form as:

Nat Struct Mol Biol. 2018 August ; 25(8): 648–649. doi:10.1038/s41594-018-0107-8.

A new twist on V(D)J recombination

Fred Dyda¹, Phoebe A. Rice^{2,*}

¹Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, US National Institutes of Health, Bethesda, MD, USA.

²Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL, USA.

Abstract

A series of new cryo-EM structures reveals a surprising twist in how the RAG complex initiates V(D)J recombination. The initial complex with substrate DNA adopts two conformations: in one, the DNA is relatively undistorted but the scissile phosphate is far from the active site, and in the other the DNA is partially melted and unwound by half a turn, which allows the scissile phosphate to dock into the active site. Similar pre-catalysis DNA melting may occur with other DDE recombinases, for which equivalent complexes with uncleaved substrate DNA are not yet available.

Vertebrate genomes are not large enough to encode an immune system capable of recognizing all possible threats. Instead, diversity in immunoglobulin and T cell receptor genesis generated by DNA rearrangements and mutations, beginning with the assembly of these genes from pools of interchangeable parts, via V(D)J recombination^{1,2}. In this process, a V (variable) region-encoding DNA segment is joined with a D (diversity) and a J (joining) region-encoding segment. V(D)J recombination is initiated by the RAG complex, which pairs the DNA segments to be joined together and cleaves at the junction between the protein-coding and noncoding regions. The liberated coding regions are then ligated together by the non-homologous end joining pathway for double-strand break (DSB) repair. A series of recent structures sheds new light on how the RAG complex works^{3–6}.

The RAG complex contains a heterotetramer of two proteins, RAG1 and RAG2^{1,2}. The multidomain RAG1 subunits contain a catalytic domain that belongs to the DDE family of DNA transposases and retroviral integrases and has an RNase H- like fold⁷. Although structurally variable in detail, all DDE enzymes use a conserved cluster of acidic residues to bind two Mg²⁺ ions that catalyze phosphotransfer. The RAG complex generates DSBs in a two-step process that requires impressive DNA gymnastics (Fig. 1). After the first strand is hydrolytically nicked, the leaving 3' -OH stays in the active site, but the newly freed phosphate group is replaced by another phosphate from the opposite strand. This second scissile phosphate group is then attacked by the 3' -OH still lurking in the active site,

* price@uchicago.edu.

Competing interests

The authors declare no competing interests.

creating a hairpin on the end of the coding region. All of this action happens in a single DDE active site.

V(D)J recombination most likely arose through domestication of a DNA transposon^{8,9}. The recombination signal sequences (RSSs) recognized by the RAG complex are equivalent to the transposase binding sites found at the ends of DNA transposons, and many structural features of RAG1 are reminiscent of the hAT superfamily of DNA transposases, as is the mechanism of DSB formation¹⁰. However, after the DSBs are formed, transposases have another task, as they have to integrate into a target site, whereas RAGs have evolved not to integrate, as transposition can be dangerous to cells.

Significant advances in the structural biology of the RAG complex have been made in recent years. The first breakthroughs were a crystal structure of the protein heterotetramer, followed by cryo-EM structures with DNA, describing reaction intermediates^{3,4}. These structures showed how asymmetric DNA binding by the RAG complex ensures that it appropriately pairs two different RSSs, but details of the initial strand cleavage remained elusive. More recently, Kim et al.⁵ captured the initial assembly with intact DNA strands. Comparison with structures of the intermediate poised to form hairpins created a structural puzzle, as it became clear that the coding part of the DNA duplex must somehow rotate ~180° around its axis while remaining bound within the assembly.

Ru et al.⁶ now offer a fascinating solution through a series of cryo-EM structures that show the complete melting of two base pairs of the RSS adjacent to the cleavage site (part of the so-called heptamer sequence). This melting allows the required rotation and delivers the scissile phosphate for the nicking reaction to the active site after a 20-Ångstrom-long journey from the position previously seen in the intact DNA state. Substantial RAG1 domain movements contribute as well, especially movements of the so-called insertion domain, which is an all-helical topological interruption of the standard RNase H fold. One particular feature of the insertion domain, the α 15- α 16 loop, appears to play a key role in DNA melting. A similar insertion domain is found in some DNA transposases, and the equivalent loop has been seen to be an important player in the reaction sequence¹⁰. The results are a testament to the power of cryo-EM, which allows the elucidation of minor assembly states that are present in solution, although this requires—as was done here—a substantial effort both in data collection and iterative image classification.

The melting seen in these new structures of the RAG complex may be relevant to many DNA transposases. Although structures of several DNA transposases have been determined with DNA in intermediate and/or product states, there are currently none available with uncleaved substrate DNA, most likely because these complexes tend to be the least stable. However, biochemical experiments showed that the last few base pairs are highly distorted before the initial cleavage, and at least for Mu transposase, this occurs in a supercoiling-dependent manner (for example, see refs^{11–13}).

The fact that unwinding of DNA is required before the first chemical step in V(D)J recombination raises the possibility that this could be exploited as an additional, unforeseen mechanism for fine-tuning control of the reaction. Constraints on the ability of the DNA

within the coding region to twirl freely could inhibit the reaction, and removal of a nucleosome could release negative supercoiling and could activate the reaction. Interaction of the RAG complex with chromatin is a rapidly evolving field^{14,15}, and it will be interesting to see how these new structural results fit into the larger picture of in vivo chromatin dynamics.

References

1. Schatz DG & Swanson PCV Annu. Rev. Genet 45, 167–202 (2011). [PubMed: 21854230]
2. Gellert M Annu. Rev. Biochem 71, 101–132 (2002). [PubMed: 12045092]
3. Kim M-S, Lapkouski M, Yang W & Gellert M Nature 518, 507–511 (2015). [PubMed: 25707801]
4. Ru H et al. Cell 163, 1138–1152 (2015). [PubMed: 26548953]
5. Kim M-S et al. Mol. Cell 70, 358–370 (2018). [PubMed: 29628308]
6. Ru H et al. Nat. Struct. Mol. Biol 10.1038/s41594-018-0098-5 (2018).
7. Hickman AB & Dyda F Chem. Rev 116, 12758–12784 (2016). [PubMed: 27187082]
8. Huang S et al. Cell 166, 102–114 (2016). [PubMed: 27293192]
9. Kapitonov VV & Jurka J PLoS Biol. 3, e181 (2005). [PubMed: 15898832]
10. Hickman AB et al. Cell 158, 353–367 (2014). [PubMed: 25036632]
11. Kobryn K, Watson MA, Allison RG & Chaconas G Mol. Cell 10, 659–669 (2002). [PubMed: 12408832]
12. Wang Z, Namgoong SY, Zhang X & Harshey RM J. Biol. Chem 271, 9619–9626 (1996). [PubMed: 8621637]
13. Crellin P & Chalmers R EMBO J. 20, 3882–3891 (2001). [PubMed: 11447129]
14. Jain S, Ba Z, Zhang Y, Dai H-Q & Alt FW Cell 174, 102–116 (2018). [PubMed: 29804837]
15. Pulivarthy SR et al. Proc. Natl. Acad. Sci. USA 113, E6427–E6436 (2016). [PubMed: 27698124]

Fig. 1 |. DNA gymnastics in V(D)J recombination.

Three structures of the RAG complex are shown with the proteins displayed as semitransparent smoothed surfaces for simplicity and with close-ups of just one half of the complex underneath. One RAG1–RAG2 pair is green, the other is blue, and the insertion domain found within the RAG1 catalytic domain is a darker shade. Gray spheres mark Ca^{2+} ions in the active sites, orange spheres mark the scissile phosphate for the first chemical step, yellow spheres mark the scissile phosphate for the second chemical step, and smaller red spheres mark the 3'-OH that is the leaving group in the first step and the nucleophile in the second step. The middle structure captures a surprising new state in which the DNA is unwound by $\sim 180^\circ$ before the first chemical step. The figure was made using PDB [6DBT](#), [6DBL](#) and [6DBI](#).