

Research Summary for David B. Roth, MD, PhD

Summary: David Roth is interested in the molecular mechanisms underlying the development and function of the immune system, particularly regulatory mechanisms that maintain genomic stability during antigen receptor diversification, genetic lesions in this mechanism that lead to immunodeficiency, and aberrant V(D)J recombination events that lead to leukemias and lymphomas.

Introduction

Most of us tend to think of genomes as static entities, and consider alterations in DNA only when speaking of mutations or damage that must be repaired. But genomes are surprisingly plastic. Chromosomal breakage, a dangerous event that has triggered the evolution of numerous double strand break (DSB) repair pathways, is actually necessary for immune system development. B and T lymphocytes recognize antigens through receptor proteins whose genes are assembled from different DNA segments by a process known as V(D)J recombination. This gene-shuffling entails introducing DSBs in millions of lymphocyte precursors each day. We are interested in the mechanisms that preserve genomic stability in the face of all this cutting and pasting. Understanding the mechanisms involved in V(D)J recombination will provide insight into general DNA repair processes and the mechanisms underlying chromosomal rearrangements that cause leukemias and lymphomas.

More than just a nuclease

V(D)J recombination has traditionally been thought of as occurring in two stages. First, the lymphoid-specific recombinase proteins RAG1 and RAG2 cleave DNA at specific sites. The broken DNA ends must then be processed and joined with the help of proteins that repair DNA double-strand breaks in all cell types (the non-homologous end-joining or NHEJ factors).

We have shown that the RAG proteins, formerly thought to merely cleave DNA, are involved in nearly every step of recombination. RAG-1 and RAG-2 initiate recombination by creating a DSB between recombination signal sequences (RSSs) and the adjacent V, D, and J coding segments. RAG-1 and RAG-2 form a complex that nicks one strand of DNA between the RSS and the coding segment. The newly formed hydroxyl group then creates a DSB by attacking the opposite strand, forming a hairpin coding end and a signal end. The coding ends join to form the rearranged receptor gene; signal ends join to form a reciprocal product, a signal joint, which has no known immunological function.

To understand how the RAGs work, we needed to know something about their structure. No structural information is available, however, and sequence comparisons with other recombinases have been little help. So we took a comprehensive approach, mutating all the conserved acidic residues in both proteins to find the active site. Our pioneering large-scale mutagenesis project identified the parts of the recombinase that cleave DNA and placed the RAG proteins in the retroviral integrase superfamily of recombinases. We have since refined the map of the active site of RAG-1, discovered regions of the proteins important for hairpin formation, and demonstrated that hairpin-formation requires a distorted DNA intermediate. We have also shown that RAG-2 is involved in recognizing or cleaving these distorted DNA intermediates.

Our studies have provided the field a fairly detailed functional map of RAG-1 and the first clue to the specific roles of RAG-2.

New genome guardians?

We and others have shown that NHEJ proteins such as DNA-PKcs and Ku are important for resolving the DSBs generated during V(D)J recombination. (NHEJ factors have been called genome guardians because they reduce the incidence of chromosome rearrangements that can cause cancer.) Whether the RAG proteins help process DNA ends had remained an open question. We discovered several RAG-1 and RAG-2 mutations that allow cleavage but prevent proper joining of broken DNA ends, producing junctions with the very same defects produced by NHEJ mutations. Some of the RAG mutations block joining of both coding and signal ends (like Ku mutations), whereas others specifically block coding joint formation (like DNA-PKcs mutations). Our data provided conclusive evidence that the RAG proteins play critical roles in joining both coding and signal ends. They also raised the intriguing possibility that the RAG proteins might help prevent aberrant handling of broken ends just like NHEJ factors and thereby serve as genome guardians. In fact, several of the residues altered in our RAG joining mutants are also mutated in patients with inherited immunodeficiency syndromes. Without joining, the antigen receptor repertoire cannot be made; although lymphomas have been reported in some of these immunodeficient patients, most likely die from infections or receive bone marrow transplants before lymphoid neoplasms can develop.

Transposition and genomic stability

Recent work has shown that the RAG proteins can act as a transposase, integrating a linear fragment of DNA that terminates in signal ends into target DNA molecules. Because transposition could obviously lead to chromosomal rearrangements, we are seeking to understand how transposition occurs (or is prevented from occurring) *in vivo*. We reasoned that the cell needs to regulate transposition without impairing V(D)J recombination. One logical way to do this would be to control the selection of the target DNA molecule, since this step is not part of the normal recombination pathway. We therefore examined "target commitment", a functional association between the RAG-RSS complex and target DNA that is resistant to the addition of a second, competing target. We found that, contrary to expectation, the RAGs are a promiscuous transposase that can form productive complexes with target DNA both before and after RSS cleavage. The ability of the RAGs to commit to target *prior* to cleavage could create a preference for transposition into nearby targets, such as Ig and TCR loci, and direct transposition events into safe regions of the genome. This process might also explain why RAG-mediated transposition events have remained undetected *in vivo* despite determined searches.

Recent work in our lab, however, suggests that RAG-mediated transposition may have been under our noses all along. Our early *in vivo* studies led us to propose that the RAG proteins are capable of fusing a signal end to a hairpin coding end, thereby forming a curious product known as a hybrid joint. In 1998 the Gellert lab at the NIH verified our prediction by showing that purified RAG proteins can form hybrid joints by this mechanism *in vitro*. These data were widely interpreted as showing that all four ends (coding and signal ends) remain within a RAG-DNA post-cleavage complex that facilitates hybrid joint formation.

In a recent series of studies, however, we showed, for the first time, that the RAG-RSS complex can efficiently capture an *exogenous* hairpin from solution and use it as a target in a true transposition reaction. We also demonstrated that the RAG proteins target other forms of distorted DNA, though hairpins are strongly preferred. These are exciting findings: there has never before been any suggestion that the RAG proteins would prefer particular targets. Our results implied that RAGs should be able to form hybrid joints by capturing exogenous hairpins, i.e., by transposition. We verified this prediction experimentally, showing that hybrid joint formation does *not* occur in a stable post-cleavage complex, as had been thought. Our data strongly suggest that RAG-mediated transposition has already been detected *in vivo*—in the form of hybrid joints.

Why would the RAG proteins show such a propensity for transposing into hairpin ends and forming hybrid joints and related products termed “open-and-shut” joints, especially since neither have any known immunological function? Preferential transposition into the initial cleavage products, hairpin ends, provides an elegant mechanism for converting dangerous RAG-RSS complexes, which are capable of transposition, into harmless byproducts. Failure of such a mechanism could explain why some common lymphoid tumors manifest recurrent translocation breakpoints that contain distorted DNA. We suggest that the RAG proteins target transposition to these chromosomal locations and perform all necessary cleavage and joining events on both partner chromosomes.

Signal joints: unsafe at any speed

Signal joints have long been considered inert, dead-end products that safeguard the genome from rearrangements that occur through illegitimate joining of signal ends or transposition. We questioned this assumption and found that signal joints are not at all inert: they are readily cleaved *in vivo* and form excellent substrates for transposition. Furthermore, we discovered that the RAG proteins readily create double-strand breaks at signal joints by an entirely novel, nick-nick mechanism. Instead of serving as a safety mechanism, signal ends might actually pose a serious threat to genomic stability. Our work not only shatters the reigning dogma but raises a fundamental biological question: how are these ends prevented from engaging in inappropriate rearrangements or transposition? We are currently investigating this important problem.