

Jessica M. Jones
Martin Gellert

The taming of a transposon: V(D)J recombination and the immune system

Authors' addresses

Jessica M. Jones¹, Martin Gellert²,

¹Department of Biochemistry and Molecular Biology, Georgetown University Medical Center, Washington DC, USA.

²Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA.

Correspondence to:

Martin Gellert

Laboratory of Molecular Biology
National Institute of Diabetes and Digestive
and Kidney Diseases

National Institutes of Health

Building 5/Room 241

Bethesda, MD 20892-0560

USA

Tel.: +1 301 496 5888

Fax: +1 301 496 0201

E-mail: gellert@helix.nih.gov

Summary: The genes that encode immunoglobulins and T-cell receptors must be assembled from the multiple variable (V), joining (J), and sometimes diversity (D) gene segments present in the germline loci. This process of V(D)J recombination is the major source of the immense diversity of the immune repertoire of jawed vertebrates. The recombinase that initiates the process, recombination-activating genes 1 (RAG1) and RAG2, belongs to a large family that includes transposases and retroviral integrases. RAG1/2 cleaves the DNA adjacent to the gene segments to be recombined, and the segments are then joined together by DNA repair factors. A decade of biochemical research on RAG1/2 has revealed many similarities to transposition, culminating with the observation that RAG1/2 can carry out transpositional strand transfer. Here, we discuss the parallels between V(D)J recombination and transposition, focusing specifically on the assembly of the recombination nucleoprotein complex, the mechanism of cleavage, the disassembly of post-cleavage complexes, and aberrant reactions carried out by the recombinase that do not result in successful locus rearrangement and may be deleterious to the organism. This work highlights the considerable diversity of transposition systems and their relation to V(D)J recombination.

Introduction

The extraordinary observation that the genes encoding antigen receptors must be assembled from an array of gene segments (1), unlike all other known protein coding regions in vertebrates, was made more palatable by a similarly extraordinary assertion, made decades earlier. This finding was that genetic elements of the *activator/dissociation* (Ac/Ds) family in maize had a propensity to move among the chromosomes (2), in seeming defiance of all previously held doctrine. Since then, many mobile genetic elements have been described, carrying out a variety of transposition and site-specific recombination reactions. V(D)J recombination, the process that assembles antigen receptors, did not at first fit comfortably into this group, but extensive biochemical and genetic work during the intervening years has greatly clarified its relation to these diverse systems. The study of V(D)J recombination has been continually

Immunological Reviews 2004

Vol. 200: 233–248

Printed in Denmark. All rights reserved

Copyright © Blackwell Munksgaard 2004

Immunological Reviews

0105-2896

informed by research into other systems, and it is now apparent that V(D)J recombination shares a particularly close relationship with the family of transposons that includes Ac/Ds. These similarities support the hypothesis that V(D)J recombination evolved from a primordial transposon.

In the germline, immunoglobulin (Ig) and T-cell receptor (TCR) loci are composed of multiple variable (V), joining (J), and in some cases diversity (D) gene coding segments. Each of these segments is flanked by a recombination signal sequence (RSS), which is required to direct its rearrangement. An individual RSS comprises heptamer and nonamer motifs with the consensus sequences CACAGTG and ACAAAAACC. These motifs are separated by a relatively non-conserved spacer either 12 or 23 bases in length. Recombination at the V(D)J loci takes place only between gene segments flanked by RSSs with different spacer lengths, a phenomenon known as the 12/23 rule (1). The loci are arranged so that maintenance of the 12/23 rule ensures appropriate rearrangement (e.g. D to J but not J to J).

At the biochemical level, recombination can be divided into two phases: DNA cleavage and joining. Cleavage is carried out by a lymphoid-specific recombinase called RAG1/2, composed of the RAG1 and RAG2 proteins (3, 4). Initially, RAG1/2 nicks the DNA directly adjacent to an RSS, leaving a free 3' hydroxyl (OH) on the end of the flanking coding segment (5, 6) (Fig. 1). To fully cleave the DNA, the 3' OH freed by nicking attacks the phosphodiester backbone on the opposite strand in a direct transesterification reaction (6, 7). This activity leaves a hairpin on the end of the coding flank (the 'coding' end) and a blunt cut RSS (the 'signal' end). Nicking can take place at an isolated RSS, at least *in vitro*, while under appropriate conditions, transesterification requires a 12/23 RSS pair (8). RAG1/2 assembles a pair of RSSs into a synaptic complex in which the coupled cleavage of both RSSs is accomplished (9). After coupled cleavage *in vitro*, it is possible to isolate a complex including all four cleaved ends (9, 10). However, coding ends are present in this complex at lower levels than signal ends (9, 10), suggesting that they are only loosely bound. This view is consistent with what is known regarding the subsequent steps in recombination.

The fate of the two types of cleaved DNA ends is quite different. In the cell, the coding ends are quickly opened, processed and joined together to form coding joints (11). Genetic evidence indicates that this joining is accomplished by the general non-homologous end-joining (NHEJ) DNA repair apparatus (12, 13), including the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and its associated Ku heterodimer, DNA ligase IV and its accessory protein

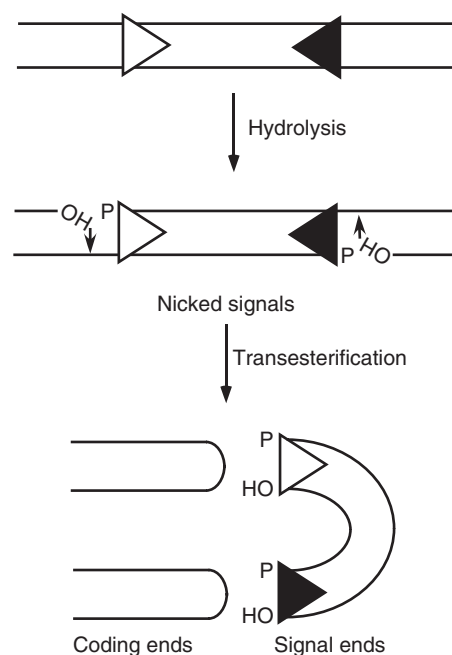


Fig. 1. DNA cleavage by the recombination-activating gene proteins.

Recombination signal sequences (RSSs) are denoted by triangles. In the hydrolysis step, a nick is made at the 5' end of the RSS, leaving a 3' hydroxyl (OH) on the coding flank. In the transesterification step, the OH group attacks the opposite strand to produce a hairpin coding end and a blunt signal end. Hairpin formation occurs in the context of a synaptic complex including a 12/23 RSS pair.

XRCC4, as well as the Artemis protein (14). Both genetic data and the known biochemical activities of these proteins suggest the following pathway. Artemis, an endonuclease whose activity is regulated by DNA-PKcs (15), nicks the coding end hairpins. The ends may be processed by a variety of factors including exonucleases and terminal deoxynucleotidyl transferase. The ligase IV/XRCC4 complex then joins the ends together (13, 16), probably with the assistance of Ku (17). Gaps in the joined DNA may be filled by repair polymerases either before or after first-strand ligation. This process introduces a high degree of variability into the coding joints, which further increases the diversity of the immune repertoire.

Unlike coding ends, signal ends appear to persist in the cell for an extended period (11, 18). When they are joined together, it is almost always without the addition or loss of bases (19). Joining requires Ku, XRCC4, and ligase IV, but is partially independent of DNA-PKcs (20) and completely independent of Artemis (14). The perfect head-to-head nature of the signal joints indicates that the ends are protected from additional DNA processing factors. *In vitro*, the signal ends remain very tightly bound to the RAG1/2 complex after cleavage (21, 22), and this complex sequesters the ends from a system of purified joining factors including Ku, XRCC4, and

ligase IV (22). The presence of this complex *in vivo* may explain why the ends remain unmodified. In theory, this stable post-cleavage complex must be disassembled or remodeled to allow access to the ends, as is discussed in greater detail below.

Parallels between V(D)J recombination and transposition are apparent at all levels. For the purposes of this discussion, mobile genetic elements, including transposons, insertion elements, and retroviral integration systems, as a group will be referred to as transposons, unless a specific element is being described. All transposons require *cis*-acting DNA sequences and *trans*-acting proteins (transposases) for mobilization (23). The *cis*-acting sequences are usually either perfect or imperfect inverted repeats at the transposon ends that serve as binding sites for the transposase. In the simplest cases, the entire element may consist of nothing more than a transposase gene, its control elements, and the flanking binding sites. More complex transposons include multiple genes required for transposition (and sometimes unrelated genes), and they may have multiple transposase binding sites at their ends, as well as internal sequences that are required for efficient transposition activity. Analysis of various eukaryotic genomes indicates that a surprisingly high percentage of DNA may originate from transposons or other mobile elements (24). However, many of these elements are no longer active. Some transposons have lost their integral transposase and must rely on a transposase provided by another element for mobilization. In other cases, transposons produce an active transposase but have lost their ability to mobilize because of inactivation of their ends through deletion or mutation.

Similarities between V(D)J recombination and transposition include (i) the DNA layout of the recombination loci, (ii) the process of RAG1/2 nucleoprotein complex assembly and cleavage, (iii) the existence of stable nucleoprotein intermediates that must be remodeled or disassembled (possibly with assistance from other factors), and (iv) the array of side reactions carried out by RAG1/2, with the most significant being its ability to carry out transpositional strand transfer.

Recombination locus structure

The first indication of a connection between V(D)J recombination and transposition came from an examination of the structure of the recombination loci. In the opposite orientation, recombining RSSs are most often found in an inverted configuration (25) (Fig. 2). In this 'deletional' configuration, signal joints and the DNA between the RSSs are lost from the chromosome in deletion circles. At the simplest level, a pair of RSSs in the deletional arrangement approximates a pair of

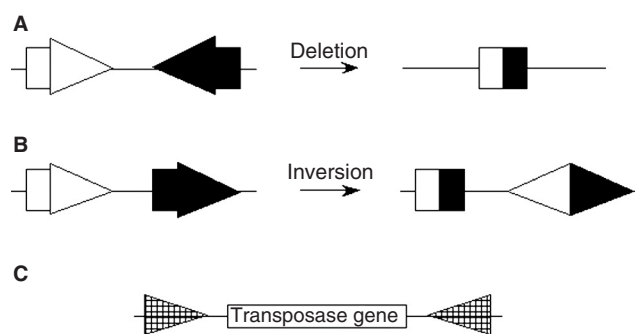


Fig. 2. Deletion and inversion during V(D)J recombination.

Recombination signal sequence (RSSs) and transposon ends are denoted by triangles; coding flanks are denoted by rectangles. Recombination between RSSs in opposite orientations results in deletion of intervening DNA (A), while recombination between RSSs in the same orientation results in inversion (B). Only products that remain in the chromosome after V(D)J recombination are shown. A pair of RSSs in the deletional configuration resembles a pair of transposon ends (C).

transposon ends, with the notable difference that recombining RSSs may be separated by hundreds of kilobases, and no mobile transposon of this length has ever been identified. In the less common direct repeat or 'inversional' arrangement, signal joints are retained in the chromosome. In both arrangements, coding joints remain in the chromosome.

The complex arrangement of RSSs most closely resembles two types of transposable elements in plants. The first are the CACT elements, so-called because the sequence of the first several bases of the transposon end. In CACT elements, the inverted repeats on either end adjoin a myriad of perfect and imperfect internal repeats in both orientations, stretching over several hundred bases (26, 27). Progressive deletion of these sequences substantially reduces transposition. These ends look like mini-V(D)J loci, but are not truly equivalent, in that only the inverted repeats at the very end of the element are actually used for cleavage. Why internal perfect repeats are not used and how these many binding sites contribute to mobilization are not clear. One model is that they increase the local concentration of transposase near the element's ends and assist in assembly of the active nucleoprotein complex. This mechanism may be common for transposons that contain multiple transposase binding sites.

The V(D)J loci also resemble a subset of *Ds* alleles in maize that include inverted and tandem double *Ds* end sequences, and these alleles provide a hypothetical model for the expansion of a single RSS pair from a primordial transposon into the current recombination loci. The double *Ds* ends often flank large (up to 45 kB) chromosomal DNA duplications (28). They can result from insertion of a *Ds* element very close to or within another *Ds* in the opposite orientation. The use of ends from nested or adjacent elements either intrachromosomally or between sister

chromatids can lead to large duplications and deletions. The V(D)J loci also resemble the *Ds* elements in that neither carries its own transposase/recombinase. Intra-element insertion could explain transposase gene inactivation, leading to the current situation in which the transposase must be supplied in *trans* from another element.

The core RAG1/2 recombinase

The delineation of active 'core' domains of RAG1 and RAG2, consisting of amino acids 384–1008 out of 1040 for RAG1 and amino acids 1–383 out of 527 for RAG2, has facilitated the study of their biochemical properties (29–32). The RAG1 and RAG2 cores have been shown to possess the minimal functions necessary for rearrangement of extra-chromosomal V(D)J recombination substrates in the cell (29–31), although at a somewhat lower level than the full-length proteins (33–35). Core RAG1 has been shown to support coding joint formation at the Ig heavy-chain (IgH) locus in cultured mouse cells (34), and a RAG1^{-/-} mouse with a knock-in of the RAG1 core undergoes lymphocyte development (36), although in both cases there appears to be a reduction in the overall level of V(D)J recombination. The RAG2 core, on the other hand, does not support B-lymphocyte development in the mouse because of a specific blockage in V to DJ rearrangement at the IgH locus (37). In human patients, mutations in the RAG1 and RAG2 core regions have been associated with severe combined immunodeficiency disease (38). N-terminal truncations of the non-core region of RAG1 as well as a certain point mutation in a highly conserved motif in this region can also lead to a B⁻/T⁺ form of immune deficiency, suggesting that the non-core region is essential for normal lymphocyte development in humans (38). Nevertheless, nearly all of what is known regarding RAG1/2 biochemistry has been learned using the core species. The experiments discussed below were performed using the core proteins, unless otherwise mentioned. Any known differences in the activities of the core and full-length proteins are noted.

DNA binding to a single RSS

The ability to specifically bind RSS DNA is intrinsic to the RAG1/2 recombinase. Gel mobility shift assay analyses have identified RAG1/2 complexes bound to a single RSS as well as synaptic complexes in which RAG1/2 bind a 12/23 RSS pair (9, 22, 39). These nucleoprotein complexes are active for DNA cleavage (40), indicating that they represent legitimate intermediates in the cleavage pathway. RAG1 is minimally

competent to bind RSS DNA (41, 42), but with relatively low affinity and specificity. Footprinting and crosslinking analyses have revealed that a dimer of RAG1 binds primarily to the nonamer region of the RSS (43–45). The addition of one or two RAG2 protomers to this complex makes binding much more specific and extends the footprint to include the heptamer (43–46). RAG2 includes a number of conserved basic residues which are required to support stable RAG1/2-RSS complex formation (47). RAG2 can be chemically crosslinked to DNA in the heptamer/coding DNA flank, suggesting that it also contacts the DNA (48), but RAG2 by itself has not been shown to possess DNA-binding activity.

The pathway to assembly of a RAG1/2 complex on a single RSS is not entirely clear. RAG1 is a dimer in solution (49), and pre-incubation of RAG1 with RAG2 in the absence of DNA increases the initial rate of cleavage (43) indicating that RAG1/RAG2 binding can occur in the absence of DNA. It is also possible for RAG1₂ or RAG1₂/RAG2 complexes already bound to DNA to attract additional RAG2 protomers. This scenario has been observed with sequential assembly of RAG1/RAG2 complexes including one and two RSSs (50). In the cell, turnover of full-length RAG1 and RAG2 is regulated independently with RAG2 being present primarily during the G1 phase of the cell cycle and RAG1 present throughout (51). It is therefore possible that RAG1 in the cell may be bound to DNA in the absence of RAG2 during much of the cell cycle. However, RAG1's relatively low affinity for DNA in the absence of RAG2 suggests that this is not the case. In addition, there may be other regulatory mechanisms that prevent RAG1–DNA interaction during the S, G2, and M phases.

One additional protein is required for optimal RAG1/2-RSS binding. This is a high-mobility group (HMG) 1 or 2 DNA-binding protein (52). HMG appears to play multiple roles in complex assembly *in vitro*. It is particularly important for binding to the 23 RSS (52), where it binds within the nonamer-proximal spacer region (53). HMG can also be incorporated into RAG1/2–12 RSS complexes (54). In addition to promoting specific RAG1/2–RSS interaction, HMG prevents the assembly of higher order RAG1/2–DNA complexes that do not appear to represent legitimate cleavage intermediates (J. M. Jones and M. Gellert, unpublished observations). HMG has been shown to assist in assembly of a RAG1/2 complex on chromatinized RSS substrates that have been acetylated and remodeled so as to resemble 'open' recombination loci (55). However, the role of HMG proteins in recombination in the cell is not easy to test because of the presence of multiple HMG family genes.

The requirement for additional structural proteins for the assembly of a stable transposase nucleoprotein complex is

common. The transposing bacteriophage Mu requires the host DNA-binding/bending protein integration host factor (IHF) and the histone-like protein HU (56). Mu ends include multiple binding sites for the phage transposase (MuA), with asymmetric spacing among the sites on each end, and MuA also binds to an internal activating sequence that assists in assembly of a stable synaptic complex. IHF and HU bind between the MuA-binding sites and bend the DNA, thus stimulating interaction between the MuA protomers and assembly of the transposase tetramer which is active for DNA cleavage (57). As with RAG1/2 binding, these accessory factors may be more important for binding to one end than to the other. While HU cannot substitute for HMG in *in vitro* RAG1/2 biochemical systems, HMG and HU appear to play similar roles in recombinase/transposase DNA interaction.

Synaptic complex assembly

A synaptic complex including two transposon ends held together in a nucleoprotein complex is a universal intermediate in transposition (58), and such a complex is also required for V(D)J recombination. While many individual steps may be involved in synaptic complex assembly, two basic pathways can be imagined (Fig. 3). In the first, the 'single-complex' pathway, a complex capable of binding both DNA ends assembles on a single end, and the second end enters the complex free of additional transposase. In the second, the 'double-complex' pathway, a transposase complex assembles on each end, and the two ends are brought together by protein-protein interactions. Until recently, it was assumed that the double-complex pathway is predominant, but it is now clear that

both pathways are used by various systems. In the case of V(D)J recombination, only the single-complex pathway leads to the assembly of active synaptic complexes.

Several lines of evidence indicate that the RAG1/2 complex capable of binding both RSSs must first assemble on a single RSS, and the second RSS enters as naked DNA. First, pre-binding of RAG1/2 to both RSSs in isolation inhibits their ability to assemble a complex that is active for cleavage (59). This indicates that RAG1/2 complexes formed on each single RSS do not represent the equivalent of half synaptic complexes, and cannot come together to form a synaptic complex. This result was confirmed by direct observation of synaptic complex assembly by gel mobility shift. Mundy *et al.* (50) made similar observations while studying sequential assembly of RAG1 and RAG2 protomers on a 12 RSS followed by the addition of a 23-RSS partner. In these experiments, two distinct RAG1/2 complexes bound to a single RSS were observed, one including a single RAG2 protomer and the second including two RAG2 protomers. The acquisition of a 23 RSS by the larger complex resulted in synaptic complex assembly. The synaptic complex also contained two RAG2 protomers, and additional lines of evidence indicated that it had the same complement of RAG1 protomers as the complex on a single RSS. These data demonstrate that no additional RAG1 or RAG2 protomers entered the complex with the acquisition of the 23 RSS. We have confirmed in this laboratory that the second RSS in a synaptic complex can enter free of additional RAG1/2 using competition experiments in which specific competitor was used to bind all free RAG1/2 in solution (59).

While the stoichiometry of RAG1/2 required to carry out cleavage of a single RSS has been fairly well established, the stoichiometry of the synaptic complex is still a matter of debate. RAG1 exists as a dimer in solution and when bound to a single RSS (40, 49, 60). With the addition of either one or two RAG2 protomers, this complex is competent to carry out nicking of a single RSS end in the presence of magnesium, or both nicking and transesterification in the presence of manganese, which relaxes the requirement for a pair of RSSs (10, 40, 43, 49). The Swanson group (54) has found that a complex of a single RAG1 dimer and two RAG2 protomers is competent to bind to a pair of RSSs and carry out cleavage in the presence of magnesium. However, two other laboratories have presented evidence that additional protomers of RAG1 are required to carry out coupled cleavage (50, 60). New approaches would be useful in resolution of this debate. One promising tool is electrospray mass spectrometry, which has been demonstrated to be gentle enough to measure the mass of intact, non-covalently associated macromolecular complexes (61). In theory, RAG1/2-RSS complexes assembled in solution could be

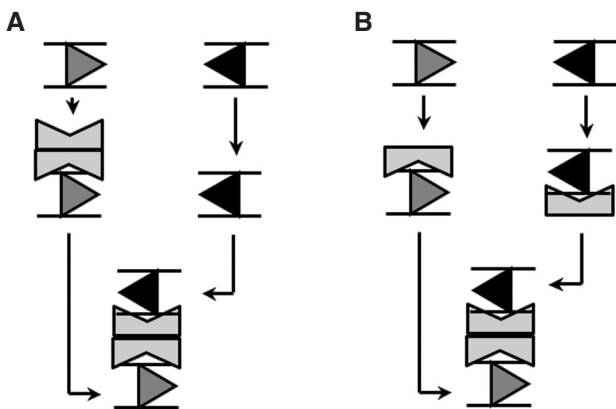


Fig. 3. Pathways for assembly of a synaptic complex. Recombination signal sequence (RSSs) are denoted by triangles, and RAG1/2 complexes are denoted by hatched shapes; high-mobility group proteins are not shown. The stoichiometry of proteins in the complex is not indicated. (A) Single-complex pathway. (B) Double-complex pathway. (see text).

analyzed directly by this method, and the composition of various complexes deduced from their total masses vs. the known molecular weights of the various components.

Synaptic complex assembly can initiate on an RSS with either a 12- or 23-bp spacer (59). However, only initial assembly on a 12-RSS leads to strict compliance with the 12/23 rule (59). After initial assembly on a 23-RSS, the nucleoprotein complex displays only a 6-fold preference for a 12-RSS partner relative to a second 23-RSS. After initial assembly on a 12-RSS, the preference for a 23-RSS is nearly absolute. Based on these data, we have developed a hypothetical model for maintenance of the 12/23 rule (Fig. 4). The model depicts RAG1/2 in solution as including protein components sufficient to bind two RSSs, but it is also possible that the assembly of multiple components on a single RSS occurs in a stepwise manner. RAG1/2 is both flexible, so that either RSS can initially be bound by either site, and asymmetric, such that the heptamer- and nonamer-binding regions in one RSS-binding site are closer together than those at the opposite site. The more narrowly and widely bipartite sites would bind preferentially to 12- and 23-RSSs, respectively. When a 12 RSS is bound by one of the RAG1/2-RSS binding sites, the complex becomes locked because the 12 bp spacer cannot be expanded to accommodate a 23-bp spacer binding site. This mechanism ensures that the second binding

site is held in a conformation that can only bind a 23-RSS; thus, this site will bind a 23-RSS even in the presence of excess free 12-RSS. If a 23-RSS is the first to be bound, the complex is not locked. Increased bending within the 23-bp spacer could allow the complex to maintain its flexibility by bringing the heptamer- and nonamer-binding regions in contact with the 23-RSS into closer proximity. The second site could then bind to a 23-RSS.

The near invariance of the 12/23 rule in recombination of the chromosomal loci suggests that initial assembly on a 12-RSS may occur there as well. The RAG1/2 proteins themselves do not seem to show a preference for binding to either a 12- or 23-RSS provided that HMG is also present (52, 59). However, this case may differ when RSSs are bound by chromatin. RSSs bound by core histones are resistant to cleavage by RAG1/2 (55). On a 12 RSS, this block can be alleviated by histone acetylation and remodeling with the swi/snf complex (62). However, 23 RSSs bound by core histones are resistant to cleavage even after such treatment. It should be noted that these experiments were carried out under conditions where RSS cleavage does not require assembly of a synaptic complex, so it is not known how assembly of this complex will influence accessibility of the 23 RSS. If assembly on a 12 RSS is the rule, then complex formation would have to initiate variously at V, D, or J segments depending on the locus.

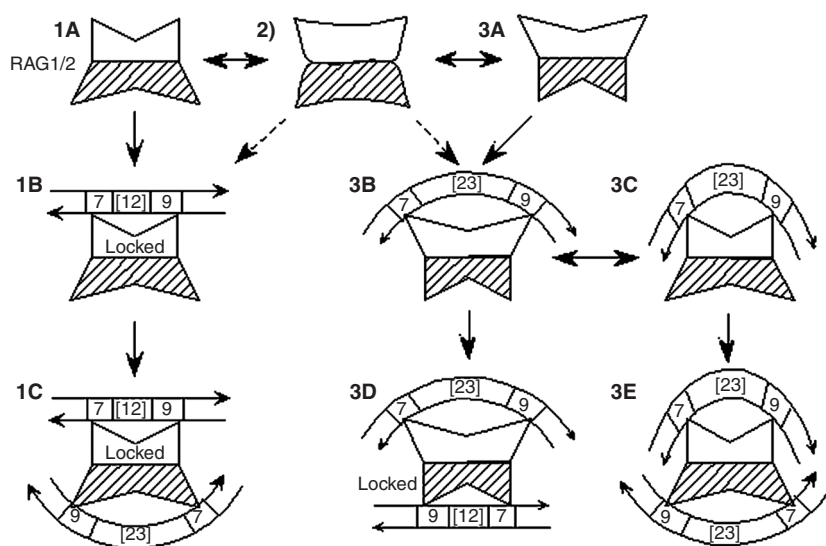


Fig. 4. Model for assembly of the synaptic complex. RAG1/2 in solution is depicted as including all its components necessary for binding to two recombination signal sequences (RSSs) (1A and 3A); alternatively, a bivalent complex could assemble after binding of a monovalent complex to a single RSS (not shown). High-mobility group proteins are not shown. Heptamer (7), nonamer (9), and spacer regions are indicated. In the bivalent RAG1/2 complex, the heptamer- and nonamer-binding domains within one RSS-binding site are optimally arranged to bind a 12-RSS (1A, white binding site), while

these domains are farther apart in the second RSS-binding site, which can only bind to a 23-RSS (3A, white binding site). These conformations may interchange rapidly in solution (2). Initial binding to a 12-RSS locks that binding site because of the fixed length of the 12-bp spacer (1B); the second site must then be occupied by a 23-RSS (1C). Initial binding to a 23-RSS (3B) does not lock the complex because of the relative flexibility of the 23-bp spacer (3C). The second RSS to enter the complex can be either a 12-RSS (3D) or a 23-RSS (3E).

Synaptic complex assembly in other systems

Obligate assembly on a single end is not universal among transposons. In the case of phage Mu, MuA complexes can preassemble on two Mu right end oligonucleotides which can then form a synaptic complex (63). Under certain conditions, this complex is active for transpositional strand transfer, although normally transposition requires the presence of both the left and right end. Mu can follow the double-complex pathway for synaptic complex assembly in which the equivalent of a half complex forms on each end, and these complexes are then legitimate intermediates in the assembly pathway (Fig. 3). The rapid exchange of MuA protomers bound to individual binding sites on the Mu ends prior to assembly of the stable synaptic complex makes it difficult to assess whether the single-complex pathway may also be used under some circumstances.

It has also been proposed that Tn5 transposase assembles by the double-complex pathway (64). Tn5 transposase is a monomer in solution (65), and the synaptic complex includes a dimer of transposase (66). These observations can most easily be explained by the double-complex pathway for synaptic complex assembly. However, no monomer of full-length transposase bound to DNA has been detected (65, 67), and the experiments necessary to rule out the single-complex pathway have not been performed. It remains possible that either pathway could occur or that one pathway is obligate.

In the case of the bacterial insertion element IS911, there is some evidence for the single-complex pathway. In experiments using a truncated transposase and a pair of IS911 ends, the first nucleoprotein complex to be observed is the synaptic complex (68). The complex of transposase with a single end is only observed at 10-fold higher transposase concentrations. The authors speculate that the two complexes include a similar protein composition but differ only in their DNA content. If true, this observation would favor a synaptic complex assembly pathway similar to that of RAG1/2, in which the transposase complex assembles on one end, and the second end enters as naked DNA. However, the complex of the IS911 transposase with a single end would be relatively unstable.

Biochemical similarities between V(D)J cleavage and transposition

Proteins associated with the mobilization of DNA elements fall into two classes, based on whether they form covalent protein-DNA intermediates (23). Proteins that form covalent intermediates rely on an active site serine or tyrosine residue, while the other class usually possesses a catalytic triad of

three acidic residues (DDE or in some cases DDD or DED, henceforth the DDE motif) (69–72). These residues coordinate the required divalent metal cation cofactor, which is most likely to be magnesium under physiological conditions. Other metals including manganese, cobalt, and iron can substitute for magnesium *in vitro*, but they may not fully support all biochemical steps or may loosen the specificity (8, 73). In many cases, calcium can support DNA binding (73, 74), but not all subsequent biochemical steps. Crystal structures of some transposases with zinc in the metal-binding pocket have also been solved (70, 75). The large variation in the size of these cations attests to a high degree of plasticity in the active sites of the transposases.

RAG1/2 does not form a covalent intermediate with DNA, which suggests that it may belong to the DDE class of transposases. Three conserved acidic residues that are required for cleavage have been identified in RAG1 (76–78). Mutation of either D600, D708 (76–78), or E962 (76, 77) abolishes recombination of extra-chromosomal substrates *in vivo* as well as RSS cleavage by the purified protein, without decreasing DNA binding. Mutation of D600 or D708 also eliminates iron-induced cleavage of RAG1 (77), confirming their role in metal binding. The role of E962 is less clear. Mutation of E962 does not affect iron-induced cleavage of RAG1, raising the possibility that it is not directly involved in metal binding. In addition, the spacing between D708 and E962 is far greater than in other DDE motifs, and E962 may in fact be located in a separate protein domain from the D600 and D708 (W. Yang, personal communication). It has been confirmed that during cleavage all three of these residues are contributed by a single RAG1 protomer (60), eliminating the possibility of domain swapping.

DDE motif proteins accomplish a variety of tasks using a remarkably limited biochemical tool kit. Essentially, it is the goal of a transposase to move the transposon from a starting position in the donor DNA to a new position in the target (23). While pathways differ, these proteins accomplish this task using similar active sites that can perform two types of phosphoryl transfer reactions: nicking and direct transesterification. Both reactions appear to occur by a one-step in-line substitution mechanism (79, 80). DNA nicking is carried out using water as a nucleophile. The appropriate metal cofactor is particularly important for this step, as certain metals greatly lower the pK_a of water making it a potent nucleophile at neutral pH. Attack on target DNA or 'strand transfer' takes place by direct transesterification in which the transferred strand becomes covalently attached to the target strand. Under normal conditions, both ends of the transposon attack

simultaneously on opposite strands of the target, staggered by a spacing characteristic of the individual transposon. Depending on the transposase, multiple nicking and transesterification reactions may take place prior to strand transfer, if the transposon is completely excised from the donor site in a 'cut-and-paste' pathway (81). Transesterification reactions are sometimes less stringent than nicking in their metal cofactor requirements. For example, MuA transposase can carry out transesterification but not nicking in the presence of calcium (63). This ability is probably because the 3' OH is already a potent nucleophile and needs only to be held in the appropriate orientation in the active site to carry out attack.

Like many transposases, RAG1/2 displays some flexibility in its use of divalent metal cofactor, although magnesium is presumed to be the cofactor used in the cell. Cleavage *in vitro* in the presence of this cation most closely mimics cleavage in the cell, specifically in its requirement for a 12/23-RSS pair (8). RAG1/2 can nick a single RSS in magnesium, manganese (5, 6), or iron (77), but not in calcium. Experiments in which the substrate is tethered and prevented from forming synaptic complexes confirm that nicking truly occurs at an isolated RSS (82) and not for example in the context of a 12/12 or 23/23 complex. In the presence of manganese, both nicking and hairpin formation can take place without assembly of a synaptic complex (10, 40). Calcium can also support hairpin formation on pre-nicked substrates in the context of a synaptic complex (J. M. Jones and M. Gellert, unpublished observations) but not on an isolated RSS (39). Iron can support hairpin formation (77).

A common theme among transposases is that the same site nicks the transferred strand and carries out strand transfer (58, 83). Furthermore, the active site that carries out these reactions has been shown to be contributed in *trans*, at least in the systems in which this question has been asked (66, 83, 84). In other words, the protomer that contacts the binding site on one end of the element actually performs biochemistry on the other end. This explains why in many but not all cases a complete synaptic complex must be assembled before any reaction can take place.

Under conditions that support cleavage of a single RSS (i.e. in the presence of manganese), the RAG1/2 complex that carries out nicking remains bound to the DNA, even in the presence of excess competitor (39), and goes on to carry out transesterification (hairpin formation). Within this complex, the same active site that carries out nicking also performs hairpin formation (54, 60). The question of whether active sites for cleavage by RAG1/2 are contributed in *trans* or in *cis* has only been addressed using conditions that support cleavage of

a single RSS. In these experiments, RAG1 constructs with mutations in the active site were co-purified with constructs that had mutations in the region required for binding to the RSS nonamer (54). When the two mutations were present on opposite protomers, the resulting heterodimer was capable of carrying out nicking of a single RSS in the presence of magnesium. When manganese was added, the same heterodimer could carry out transesterification. This finding indicates that the RAG1 protomer bound to the nonamer at a single RSS does not carry out cleavage biochemistry of that RSS. While it appears to be the equivalent of cleavage in *trans*, it leaves open the issue as to whether cleavage within a synaptic complex occurs in *trans*. This issue could be clarified by determination of the structure of the RAG1/2–12/23 RSS synaptic complex as well as definitive determination of its stoichiometry.

Diversity among transposases

While the above discussion highlights some of the general themes unifying transposases and their similarity to RAG1/2, the remarkable diversity of this class of proteins should not be ignored. For example, there is no rule governing the stoichiometry of the transposase complex or the pathway for assembling an active complex. The variability of these proteins is also apparent both in the choice of strand that is initially nicked and the target of the first transesterification step (Fig. 5). Several eukaryotic transposons, such as the Tc/mariner and hAT families (85, 86), initiate cleavage by nicking the non-transferred strand, and the mechanisms of these transposases appear to be most similar to that observed for RAG1/2.

The hairpins in V(D)J recombination are often opened off-center. Following repair of the resulting overhang, a characteristic footprint may be left in the coding joint, with palindromic or P nucleotides where the hairpin was opened (25). Such palindromic repair footprints are also seen in the donor DNA after excision of hAT family transposases (86, 87), suggesting that hairpins on the flanking DNA are also intermediates in these events. In both cases, it is indicated that the first nick takes place on the non-transferred strand, and the complementary strand is attacked in the subsequent transesterification [Fig. 5(1B)]. The nature of the footprint for hAT family members indicates the presence of a single unpaired nucleotide at the hairpin tip (86, 87). Transesterification to form the hairpin occurs diagonally, one base removed from the nick on the opposite strand, in contrast to V(D)J recombination, where the nucleophilic attack occurs on the phosphodiester bond directly opposite to the nick. Initial biochemical analysis of Hermes, the first member of this class whose activity has been

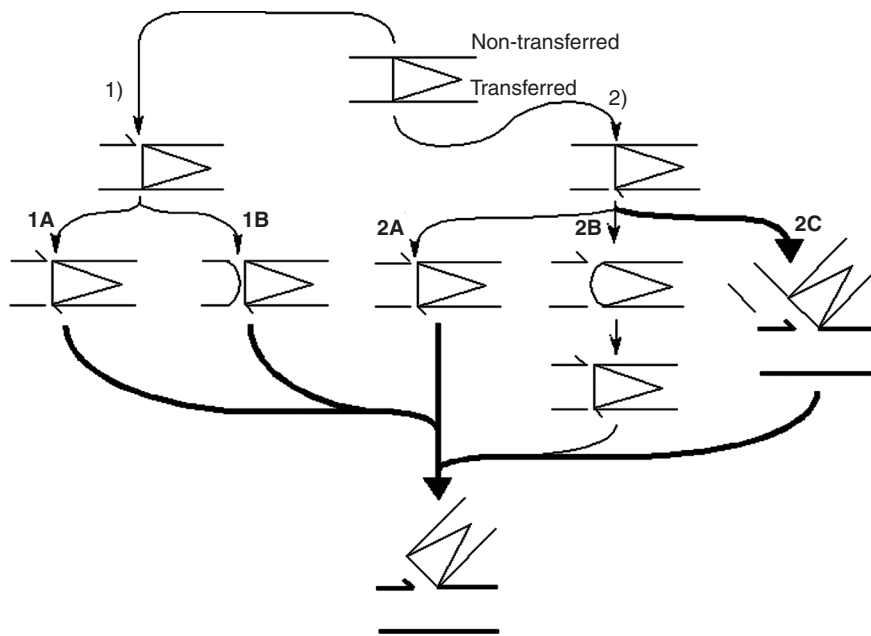


Fig. 5. Diversity among transposases. The transposon end is depicted as a triangle with the top strand running 5'-to 3' left to right; only one end is shown. 3' hydroxyl groups are depicted by half arrows. Donor/transposon DNA and target DNA are depicted by thin and thick lines, respectively. In each pathway, the strand transfer step is depicted by a bold arrow. Cleavage from the donor DNA can begin with nicking of either the non-transferred (1) or transferred (2) strands. To complete excision from the donor DNA for cut-and

paste-transposition, the opposite strand can either be nicked (1A and 2A) or subjected to attack via direct transesterification (1B and 2B). Hairpins on the transposon ends must be opened by nicking prior to strand transfer (2B). For non-cut-and-paste transposition, the transposon can carry out strand transfer immediately after nicking the transferred strand (2C), and flanking donor DNA can then be removed by nicking the non-transferred strand (2C) or by replication (not shown).

examined *in vitro*, has confirmed that such hairpins are formed during cleavage (N. Craig, personal communication).

Mos1, a member of the Tc/mariner family, excises its DNA by sequential nicking of the non-transferred and transferred strands without using a hairpin intermediate (85) [Fig. 5(1A)]. These reactions are carried out by a single polypeptide, although the stoichiometry of the active transposase complex is not known. Synaptic complex assembly is not required for both steps of cleavage, and in this way, Mos1 differs from most transposases but resembles RAG1/2. Unlike RAG1/2, however, nicking may be an obligate step prior to synaptic complex assembly, as such assembly does not occur in the presence of calcium alone and is greatly reduced by incubation at low temperature. It is not yet known whether the same active site carries out nicking of both strands. Different promoters may nick the two strands, with the nicking of the transferred strand and strand transfer being carried out by a single active site (on each end), in keeping with other transposases.

In most other systems, the transferred strand is the first to be nicked [Fig. 5(2A–C)]. For bacteriophage Mu and the eukaryotic retroviruses (reviewed in 88, 89), this nicking reaction is followed immediately by attack on target DNA to create a

branched intermediate [Fig. 5(2C)]. By simply nicking the non-transferred strand of the transposon at the branch junction, the branched intermediate can be resolved with minimal DNA synthesis, as is the case for retroviral integration and Mu simple insertion. In the case of replicative transposition for bacteriophage Mu, the transposon has devised a means of recruiting the host replicative machinery on one branched end, and in this way, it creates additional copies of itself with every round of transposition (not shown). IS911 uses a variation on this theme, in which the nicked strand attacks the opposite end of the transposon (not shown), creating a figure eight intermediate that may be resolved by simple or replicative means.

The bacterial transposons Tn5 and Tn10 employ hairpin intermediates, but the first nick is on the transferred strand, so that the ensuing transesterification attacks the phosphodiester backbone of the complementary strand of the transposon (66, 90) [Fig. 5(2B)]. This activity completely excises the transposon from its original location and leaves hairpins on the transposon ends. The transposase then nicks the hairpins and carries out strand transfer. For Tn7, an alternative mechanism is used in which the non-transferred strand is nicked by a second, non-DDE motif protein (TnsA) (91, 92) in a complex

with the DDE transposase (TnsB) (92). In this case, excision of the element does not involve a hairpin [Fig. 5(2A)].

The first transposases to be characterized biochemically all initiated cleavage by nicking the transferred strand. This finding raised a question as to why some transposases initiate cleavage with the nicking of the non-transferred strand. In theory, any transposon that nicks the transferred strand prior to the non-transferred strand has the potential to perform strand transfer prior to complete excision of the element. This activity has been observed for Tn7, where mutation of TnsA causes a switch from cut-and-paste to replicative transposition (93). Some transposons, such as phage Mu, normally make use of replicative transposition (89), but they have evolved highly sophisticated means of assembling host proteins to resolve their branched transposition intermediates. In the absence of these mechanisms, creation of such an intermediate may be a dead end. In addition, replicative transposition can be very injurious to the host. This is not a problem for a phage which has its own means of infecting a new host. For a transposon that must coexist with its host, replicative transposition could be disastrous.

Additional reactions carried out by RAG1/2

RAG1/2 can carry out a wide variety of additional DNA processing reactions *in vitro*, some of which appear to have

biological relevance. Most significant for this discussion is the demonstration *in vitro* that the complex of RAG1/2 with cleaved signal ends can carry out transpositional strand transfer (94, 95). During transposition by RAG1/2, the 3' OH on a cleaved signal end attacks the phosphodiester backbone of a strand of target DNA of non-specific sequence, generating a branched transposition intermediate (Fig. 6). Both cleaved signal ends can attack opposite strands of the target in a concerted manner, with the insertion sites staggered by four or five bases. Strand transfer is supported by magnesium, manganese, and calcium, as is the case for hairpin formation. After strand transfer, the free 3' OH at the branch junction can go on to attack the opposite target strand to form a hairpin at the target site (96) (Fig. 7). In theory, repair of this intermediate by host factors would lead to chromosomal translocation. After extensive searching and the development of numerous experimental systems, only two examples of apparent RAG1/2-mediated transposition on the chromosome have been documented (97); this observation was made in human T cells, thus in the presence of full-length RAG1 and RAG2 proteins. RAG1/2-mediated transposition does not appear to be a very common event. Aggressive transposition within a lymphoid cell by the post-cleavage signal end complex (SEC) would be disastrous for the cell and potentially for the whole organism. Despite its rarity, it is entirely possible that certain oncogenic chromosomal translocations which juxtapose the powerful Ig and TCR

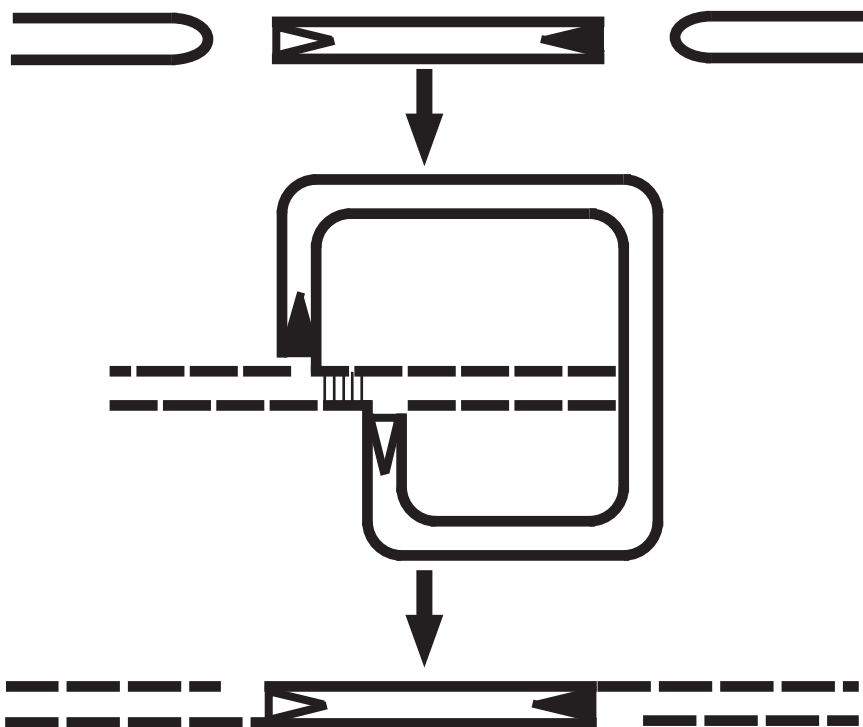


Fig. 6. Two-ended transposition by recombination-activating genes. In the top line, cleavage has liberated the two recombination signal sequence (RSS) ends that will be used to attack target DNA (dashed line). The reaction can proceed in a coupled manner, in which the 12- and 23-RSS ends attack opposite strands of the target, staggered by four or five bases, as is depicted here. Alternatively, a single end can be inserted as is shown in the next figure. In either case, both a 12- and 23-RSS are required.

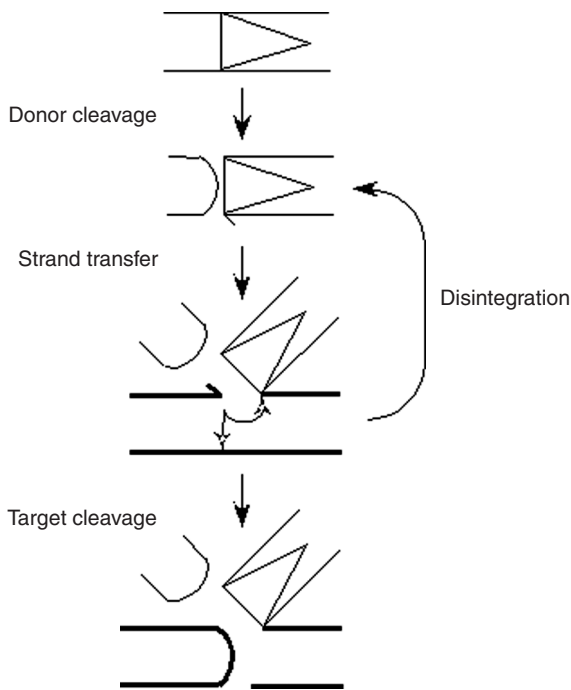


Fig. 7. Disintegration vs. target cleavage. After cleavage from the donor and one-ended transposition, the 3' hydroxyl (OH) (half arrow) at the branch junction can attack the opposite strand of the target to complete target cleavage. Joining of the coding end with the cleaved target end (not shown) could result in chromosomal translocation. Alternatively, the 3' OH can attack the recombination signal sequence/target junction, resulting in disintegration from the target. Water can also be used as a nucleophile in the disintegration reaction (not shown).

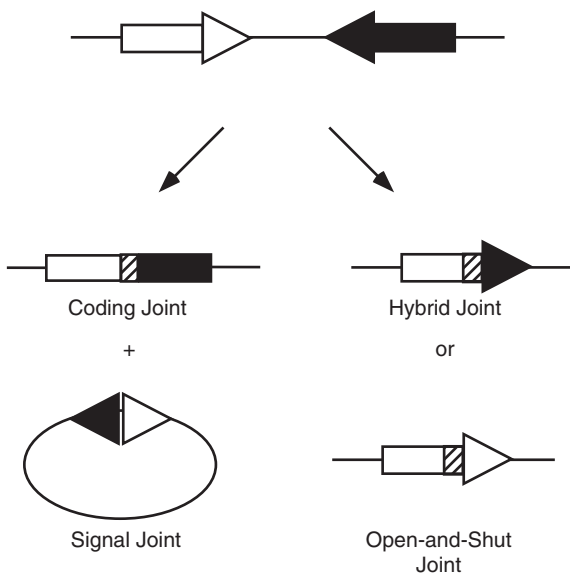


Fig. 8. Hybrid and open-and-shut joints. Recombination signal sequences (RSSs) are denoted by triangles, and coding flanks are denoted by rectangles. After cleavage, joining of one RSS to the coding flank of its partner generates a hybrid joint. Joining of an RSS to its original coding flank produces an open-and-shut joint, which can only be recognized if the junctional sequence is altered, for instance by asymmetric opening of the coding hairpin. These altered junctions are depicted by hatched boxes.

enhancers with proto-oncogenes are the result of RAG1/2-mediated transposition (95).

Reduction of transpositional strand transfer was probably an important development during the adaptation of the primordial RAG1/2 recombinase/transposase for use by the immune system. The C-terminal non-core region of RAG2 appears to play a role in reducing transpositional strand transfer *in vitro* (98–100), and regulation by GTP appears to have a similar effect (98). Downregulation of RAG2 at the G1/S transition, which also requires determinants in its C terminus, could further help to reduce transposition. Chromatin structure is a likely impediment to transposition, and during S phase, the chromatin would be in a more open conformation. Some transposons are known to target the lagging strand of active replication forks (101), so downregulating the transposase during this phase would virtually eliminate transposition. Finally, core RAG1/2 can resolve transposition intermediates in a manner that does not lead to DNA rearrangement. In the disintegration reaction (Fig. 7), the free 3' OH at the branch junction attacks the junction between RSS and target DNA, repairing the site of attack and regenerating the cleaved RSS end (96). At physiological magnesium concentrations, this reaction is strongly favored over the translocation pathway for resolution. It is currently unclear to what extent disintegration contributes to the low level of transposition observed *in vivo*.

At least one side reaction is known to occur with reasonable frequency in the cell. After cleaving the DNA, RAG1/2 can catalyze the reverse reaction in which the signal ends attack the coding end hairpins by direct transesterification (102–104). The products of such reversals are called open-and-shut (OS) or hybrid joints (HJ), depending on whether the signal end attacks the coding end to which it was originally attached or the opposite end, respectively. The existence of these products has been interpreted as support for the model in which all four ends remain in a complex after cleavage, although the formation of both types of joints suggests that such a complex would have to have a great deal of flexibility. There has been no formal demonstration that OS and HJ are generated from the same complex that resulted from cleavage or whether the coding ends may be lost from the complex and then recaptured if they are not rapidly joined together. OS and HJ may represent a special form of transposition in which the coding ends are targeted; there is evidence that hairpins and stem loops are preferred transposition targets (105). DNA hairpins and stem loops are relatively common in the cell, and it is not clear why these are not more frequently the target of RAG1/2-mediated

transposition. Higher order chromatin structure, which may be present at the recombining loci, could prevent the rapid diffusion of released ends away from the post-cleavage complex (106), having the dual effects of making the recapture of coding ends relatively favorable and decreasing the likelihood of capturing random target DNA.

RAG1/2 has the ability to nick DNA hairpins under certain conditions such as elevated pH (107, 108). It was originally proposed that this activity may be responsible for opening the hairpins at the ends of coding DNA. However, with the recognition that Artemis is a hairpin endonuclease (15), this suggestion seems less likely. The RAG1/2 hairpin-nicking activity may be an aberrant form of hairpin-targeted transposition or OS/HJ formation, the high pH making water a better nucleophile. Such plasticity with regards to nucleophile has been observed with the disintegration reaction, in which water can act in place of the junctional 3' OH (96). The same may also be true for the RAG1/2 flap endonuclease activity observed *in vitro* (109), which would represent a form of non-targeted transposition using water as the nucleophile. Of course, these observations do not rule out the possibility that the various processing reactions play a role in V(D)J recombination in the cell. It is possible, for example, that the very rare coding joints formed in Artemis and DNA-PKcs negative cells are the result of hairpin opening by the RAG1/2 recombinase. A RAG1 mutant that is competent for cleavage *in vitro* but not for hairpin opening is also associated with reduced recombination of model substrates in the cell (110), which has been interpreted as evidence for an essential role for RAG1 in hairpin opening. However, this mutant is also incompetent for signal joint formation, which can occur without opening of the coding end hairpins. Another possible interpretation of these results is that the mutant changes the nature of the post-cleavage complex (110), perhaps making it so hyper-stable that the ends cannot be released and joined together. This rigid RAG1 configuration may be less likely to perform the aberrant hairpin opening reaction *in vitro*.

Disassembly of the RAG1/2 post-cleavage complex

After cleaving a pair of RSSs, RAG1/2 remains very tightly bound to the RSS ends (9, 21, 22). This binding is only true when cleavage takes place in a synaptic complex and not after cleavage on a single end (39). *In vitro*, this stable RSS or SEC sequesters the RSS ends from functional interaction with mammalian end-joining factors including Ku, XRCC4, and ligase IV (22). These factors were unable to join the cleaved ends unless

they had been artificially deproteinized. In some cases, a complex including cleaved RSS ends, RAG1/2, and various joining factors has been observed (21). However, because the starting substrate in these reactions was a linear piece of DNA, it is not clear whether the joining factors entered the complex after cleavage or whether they may have become bound internally on the DNA prior to cleavage. The Ku heterodimer forms a structure that completely encircles the DNA strand (111), and it requires a free end to enter or exit DNA (112). Once it threads onto a piece of DNA, it can be trapped by factors binding to the ends.

Certain evidence indicates that the SEC also exists in the cell. V(D)J recombination normally takes place during the G1 phase of the cell cycle, where NHEJ is most active. Nevertheless, signal joints are not seen until the G1/S transition (11). This observation is most easily explained by the persistence of the stable SEC that prevents joining of the ends. It has also been suggested that signal ends are joined together immediately after cleavage, but they are re-cleaved by RAG1/2 (113). This cycle would continue until the G1/S transition, at which point RAG2 is downregulated. Cleavage of signal joints proceeds by nicking of each strand in a mechanism reminiscent of Mos1 rather than through a hairpin intermediate (113). Extra-chromosomal substrates, including a signal joint, can be re-cleaved by RAG1/2 proteins in the cell (113), but whether this reaction takes place in the cell during normal recombination of the chromosomal loci is less clear. *In vitro*, the SEC is resistant to treatment with specific RSS competitor (22) indicating that RAG1/2 does not dissociate from the complex. As the SEC prevents the ends from being joined together, it is unlikely that signal joints can be formed without remodeling of the complex. If this does not occur until the G1/S transition, the cycle of joining and re-cleavage may never get started. Of course, the two models for the persistence of signal ends are not mutually exclusive. The factors that remodel the SEC could be present throughout G1, in which case the ends could be joined and re-cleaved repeatedly.

Product binding energy is believed to be the main driving force for transposition in other systems (80), so the presence of a very stable SEC is consistent with what has been found for other transposases. This view raises a question as to how this highly stable post-cleavage complex is disassembled. This topic has been most thoroughly addressed in the bacteriophage Mu transposition system. The MuA tetramer that carries out strand transfer remains very tightly bound to the branched DNA product in a 'strand transfer complex' (STC1) (114). Initial remodeling of STC1 is carried out by the ClpX molecular chaperone (115–117), a member of the Clp/HSP100

ATPase family. On its own, ClpX can use the energy from ATP hydrolysis to unfold proteins containing one of its recognition sequences (118). In the context of the associated ClpP peptidase component, the unfolded polypeptides are degraded (118). MuA includes a ClpX-recognition peptide on its C-terminus (119). Recognition by ClpX of a single MuA protomer within the STC1 leads to destabilization of the entire complex to form STC2 (120). The protomer bound by ClpX is unfolded and dissociates from the DNA, and it can be degraded if ClpP is also present. However, the bulk of the MuA remains associated with the STC2, but in a conformation that can be easily disrupted (115).

Several *in vitro* assays have been developed for analysis of remodeling of the RAG1/2 SEC. RAG2 includes a peptide on its C-terminus that is intriguingly similar to a ClpX-recognition peptide. However, an exhaustive screen of mammalian chaperones to identify components necessary to remodel the SEC has so far proven fruitless, even when the SECs are formed using core RAG1 and full-length RAG2 (unpublished observations). In addition, phosphorylation of a conserved site in the RAG2 C-terminus that is necessary for its cyclic degradation (121) is not sufficient to destabilize the SEC (unpublished observations).

The N-terminal non-core region of RAG1 has been shown to stimulate signal joint formation in assays using extra-chromosomal V(D)J substrates (33–35), suggesting that it may contain determinants that promote remodeling. This region includes a zinc-binding domain of the RING configuration that is conserved in RAG1 proteins from all vertebrate classes. A point mutation that disrupts a conserved cysteine residue in the RING finger has been shown to cause Omenn's syndrome (38), a rare form of B⁻/T⁺ immune deficiency. This finding indicates that an intact RAG1 RING finger is essential for normal B-lymphocyte development in humans. In model systems, mutation or deletion of the RING finger has been shown to reduce signal joint formation (30, 34, 35).

The RAG1 RING finger has been shown recently to possess ubiquitin ligase activity (122, 123). Ubiquitin is a 76 amino acid protein that has been highly conserved throughout eukaryotic evolution. Conjugation of ubiquitin to internal lysine residues on various target proteins can lead to their destruction by the 26S proteasome or to modification of their activities (124). Like ClpX/ClpP, the 26S proteasome is a multi-component complex with both chaperone and peptidase activities. Instead of recognizing an integral peptide on target proteins, the 26S proteasome recognizes proteins conjugated to ubiquitin. Ubiquitin conjugation occurs through a multi-step cascade including the ubiquitin-activating enzyme

(E1), one of several ubiquitin-conjugating enzymes (E2s), and a ubiquitin ligase (E3). Much of the specificity of the system is contributed by the ubiquitin ligase component, which helps to bring together the E2 and the target protein. Many ubiquitin ligases also undergo ubiquitylation themselves, which can modulate their activities and/or expression levels.

We have demonstrated that full-length RAG1 undergoes ubiquitylation in cultured cells (123). Working in a cell-free system using purified proteins, we have also found that a fragment of RAG1 spanning amino acids 218–389 acts as a ubiquitin ligase and promotes its own ubiquitylation at a single, highly conserved lysine residue (K233) (123). This activity is best promoted by a specific E2 enzyme, UbcH3/CDC34. CDC34 is the E2 enzyme that promotes the G1/S transition through targeted ubiquitylation of various cell-cycle components (125). We hypothesize that it may also promote remodeling of the SEC at the G1/S transition by supporting auto-ubiquitylation of RAG1, either at K233 or another lysine residue. Using a different model system, the Sadofsky lab (122) has also identified ubiquitin ligase activity of the RAG1 RING, although supported by different E2 enzymes. Much additional work is required to determine which E2 enzyme is responsible for supporting RAG1 ubiquitin ligase activity *in vivo*, and how this activity is integrated with RAG1's role in V(D)J recombination and lymphocyte development. It should be reiterated that core RAG1/2 can carry out V(D)J recombination in cultured cells, and Omenn's syndrome patients with mutations in the RING finger can carry out V(D)J recombination in T lymphocytes. This finding indicates that the role of RAG1 E3 activity in recombination may be important but is not completely indispensable.

Concluding remarks

Remarkable progress has been made in the decade since establishment of the first biochemical system for the study of V(D)J recombination. Even as the details of core RAG1/2 biochemistry are clarified, many exciting new areas of study remain. One of these is the contribution of the non-core regions of RAG1 and RAG2. Another is the role of chromatin in regulation of the recombination process. Still another is elucidation of how aberrant rearrangement can lead to oncogenic translocations. An attentive eye cast toward other transposition systems will continue to benefit scientists studying V(D)J recombination, as breakthroughs in those fields will likely illuminate ours.

References

1. Tonegawa S. Somatic generation of antibody diversity. *Nature* 1983;**302**:575–581.
2. McClintock B. Chromosome organization and genic expression in maize. *Cold Spring Harb Symp Quant Biol* 1951;**16**:13–47.
3. Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 1990;**248**:1517–1523.
4. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. *Cell* 1989;**59**:1035–1048.
5. van Gent DC, McBlane JF, Ramsden DA, Sadofsky MJ, Hesse JE, Gellert M. Initiation of V(D)J recombination in a cell-free system. *Cell* 1995;**81**:925–934.
6. McBlane JF, et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* 1995;**83**:387–395.
7. van Gent DC, Mizuuchi K, Gellert M. Similarities between initiation of V(D)J recombination and retroviral integration. *Science* 1996;**271**:1592–1594.
8. van Gent DC, Ramsden DA, Gellert M. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. *Cell* 1996;**85**:107–113.
9. Hiom K, Gellert M. Assembly of a 12/23 paired signal complex: a critical control point in V(D)J recombination. *Mol Cell* 1998;**1**:1011–1019.
10. Bailin T, Mo X, Sadofsky MJ. A RAG1 and RAG2 tetramer complex is active in cleavage in V(D)J recombination. *Mol Cell Biol* 1999;**19**:4664–4671.
11. Ramsden CA, Gellert M. Formation and resolution of double-strand break intermediates in V(D)J rearrangement. *Genes Dev* 1995;**9**:2409–2420.
12. Taccioli GE, Rathbun GA, Oltz EM, Stamato TD, Jeggo PA, Alt F. Impairment of V(D)J recombination in double-strand break repair mutants. *Science* 1993;**260**:207–210.
13. Jeggo PA. DNA breakage and repair. *Adv Genet* 1998;**38**:185–218.
14. Moshous D, et al. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell* 2001;**105**:177–186.
15. Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 2002;**108**:781–794.
16. Grawunder U, et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* 1997;**388**:492–495.
17. Ramsden DA, Gellert M. Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *EMBO J* 1998;**17**:609–614.
18. Roth DB, Zhu C, Gellert M. Characterization of broken DNA molecules associated with V(D)J recombination. *Proc Natl Acad Sci USA* 1993;**90**:10788–10792.
19. Lieber MR, Hesse JE, Mizuuchi K, Gellert M. Studies of V(D)J recombination with extrachromosomal substrates. *Curr Top Microbiol Immunol* 1988;**137**:94–99.
20. Kulesza P, Lieber MR. DNA-PK is essential only for coding joint formation in V(D)J recombination. *Nucleic Acids Res* 1998;**26**:3944–3948.
21. Agrawal A, Schatz DG. RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell* 1997;**89**:43–53.
22. Jones JM, Gellert M. Intermediates in V(D)J recombination: a stable RAG1/2 complex sequesters cleaved RSS ends. *Proc Natl Acad Sci USA* 2001;**98**:12926–12931.
23. Craig NL. Mobile DNA: an Introduction. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, eds. *Mobile DNA II*. Washington, DC: ASM Press, 2001: 3–11.
24. Deininger PL, Roy-Engel AM. Mobile elements in animal and plant genomes. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, eds. *Mobile DNA II*. Washington, DC: ASM Press, 2001: 1071–1092.
25. Gellert MV. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem* 2002;**71**:101–132.
26. Grant SR, Hardenack S, Trentmann S, Saedler H. Functional cis-element sequence requirements for suppression of gene expression by the TnpA protein of the *Zea mays* transposon *En/Spm*. *Mol Gen Genet* 1993;**241**:153–160.
27. Masson P, Strem M, Fedoroff N. Genetic and molecular analysis of the *Spm*-dependent *a-m2* alleles of the maize *a* locus. *Genetics* 1987;**177**:117–137.
28. Kunze R. The maize transposable element *Activator* (Ac). *Curr Top Microbiol Immunol* 1993;**204**:161–194.
29. Sadofsky MJ, Hesse JE, Gellert M. Definition of a core region of RAG-2 that is functional in V(D)J recombination. *Nucleic Acids Res* 1994;**22**:1805–1809.
30. Sadofsky MJ, Hesse JE, McBlane JF, Gellert M. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Res* 1993;**21**:5644–5650.
31. Cuomo CA, Oettinger MA. Analysis of regions of RAG-2 important for V(D)J recombination. *Nucleic Acids Res* 1994;**22**:1810–1814.
32. Silver DP, Spanopoulou E, Mulligan RC, Baltimore D. Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. *Proc Natl Acad Sci USA* 1993;**90**:6100–6104.
33. Steen SB, Han JO, Mundy C, Oettinger MA, Roth DB. Roles of the “dispensable” portions of RAG-1 and RAG-2 in V(D)J recombination. *Mol Cell Biol* 1999;**19**:3010–3017.
34. Roman CA, Cherry SR, Baltimore D. Complementation of V(D)J recombination deficiency in RAG-1(-/-) B cells reveals a requirement for novel elements in the N-terminus of RAG-1. *Immunity* 1997;**7**:13–24.
35. McMahan CJ, Difilippantonio MJ, Rao N, Spanopoulou E, Schatz DG. A basic motif in the N-terminal region of RAG1 enhances V(D)J recombination activity. *Mol Cell Biol* 1997;**17**:4544–4552.
36. Dudley DD, et al. Impaired V(D)J recombination and lymphocyte development in core RAG1-expressing mice. *J Exp Med* 2003;**198**:1439–1450.
37. Kirch SA, Rathbun GA, Oettinger MA. Dual role of RAG2 in V(D)J recombination: catalysis and regulation of ordered Ig gene assembly. *EMBO J* 1998;**17**:4881–4886.
38. Villa A, et al. V(D)J recombination defects in lymphocytes due to RAG mutations: severe immunodeficiency with a spectrum of clinical presentations. *Blood* 2001;**97**:81–88.
39. Hiom K, Gellert M. A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. *Cell* 1997;**88**:65–72.
40. Swanson PC. The DDE motif in RAG-1 is contributed in trans to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. *Mol Cell Biol* 2001;**21**:449–458.
41. Spanopoulou E, Zaitseva F, Wang FH, Santagata S, Baltimore D, Panayotou G. The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. *Cell* 1996;**87**:263–276.
42. Difilippantonio MJ, McMahan CJ, Eastman QM, Spanopoulou E, Schatz DG. RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. *Cell* 1996;**87**:253–262.
43. Swanson PC, Desiderio S. RAG-2 promotes heptamer occupancy by RAG-1 in the assembly of a V(D)J initiation complex. *Mol Cell Biol* 1999;**19**:3674–3683.
44. Akamatsu Y, Oettinger MA. Distinct roles of RAG1 and RAG2 in binding the V(D)J recombination signal sequences. *Mol Cell Biol* 1998;**18**:4670–4678.
45. Nagawa F, et al. Footprint analysis of the RAG protein recombination signal sequence complex for V(D)J type recombination. *Mol Cell Biol* 1998;**18**:655–663.

46. Mo X, Bailin T, Sadofsky MJ. RAG1 and RAG2 cooperate in specific binding to the recombination signal sequence in vitro. *J Biol Chem* 1999;**274**:7025–7031.
47. Fugmann SD, Schatz DG. Identification of basic residues in RAG2 critical for DNA binding by the RAG1-RAG2 complex. *Mol Cell* 2001;**8**:899–910.
48. Mo X, Bailin T, Sadofsky MJ. A C-terminal region of RAG1 contacts the coding DNA during V(D)J recombination. *Mol Cell Biol* 2001;**21**:2038–2047.
49. Rodgers KK, Villey IJ, Ptaszek L, Corbett E, Schatz DG, Coleman JE. A dimer of the lymphoid protein RAG1 recognizes the recombination signal sequence and the complex stably incorporates the high mobility group protein HMG2. *Nucleic Acids Res* 1999;**27**:2938–2946.
50. Mundy CL, Patenge N, Matthews AG, Oettinger MA. Assembly of the RAG1/RAG2 synaptic complex. *Mol Cell Biol* 2002;**22**:69–77.
51. Desiderio S, Lin WC, Li Z. The cell cycle and V(D)J recombination. *Curr Top Microbiol Immunol* 1996;**217**:45–59.
52. van Gent DC, Hiom K, Paull TT, Gellert M. Stimulation of V(D)J cleavage by high mobility group proteins. *EMBO J* 1997;**16**:2665–2670.
53. Mo X, Bailin T, Noggle S, Sadofsky MJ. A highly ordered structure in V(D)J recombination cleavage complexes is facilitated by HMG1. *Nucleic Acids Res* 2000;**28**:1228–1236.
54. Swanson PC. A RAG-1/RAG-2 tetramer supports 12/23-regulated synapsis, cleavage, and transposition of V(D)J recombination signals. *Mol Cell Biol* 2002;**22**:7790–7801.
55. Kwon J, Imbalzano AN, Matthews A, Oettinger MA. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. *Mol Cell* 1998;**2**:829–839.
56. Craigie R, Arndt-Jovin DJ, Mizuuchi K. A defined system for the DNA strand-transfer reaction at the initiation of bacteriophage Mu transposition: protein and DNA substrate requirements. *Proc Natl Acad Sci USA* 1985;**82**:7570–7574.
57. Lavoie BD, Chan BS, Allison RG, Chaconas G. Structural aspects of a higher order nucleoprotein complex: induction of an altered DNA structure at the Mu-host junction of the Mu type 1 transpososome. *EMBO J* 1991;**10**:3051–3059.
58. Rice PA, Baker TA. Comparative architecture of transposase and integrase complexes. *Nat Struct Biol* 2001;**8**:302–307.
59. Jones JM, Gellert M. Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. *EMBO J* 2002;**21**:4162–4171.
60. Landree MA, Kale SB, Roth DB. Functional organization of single and paired V(D)J cleavage complexes. *Mol Cell Biol* 2001;**21**:4256–4264.
61. Miranker AD. Mass spectrometry of proteins of known mass. *Proc Natl Acad Sci USA* 2000;**97**:14025–14027.
62. Kwon J, Morshead KB, Guyon JR, Kingston RE, Oettinger MA. Histone acetylation and hSWI/SNF remodeling act in concert to stimulate V(D)J cleavage of nucleosomal DNA. *Mol Cell* 2000;**6**:1037–1048.
63. Savilahti H, Rice PA, Mizuuchi K. The phage Mu transpososome core: DNA requirements for assembly and function. *EMBO J* 1995;**14**:4893–4903.
64. Bhasin A, Goryshin IY, Steiniger-White M, York D, Reznikoff WS. Characterization of a Tn5 pre-cleavage synaptic complex. *J Mol Biol* 2000;**302**:49–63.
65. York D, Reznikoff WS. Purification and biochemical analyses of a monomeric form of Tn5 transposase. *Nucleic Acids Res* 1996;**24**:3790–3796.
66. Davies DR, Goryshin IY, Reznikoff WS, Rayment I. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science* 2000;**289**:77–84.
67. Weinrach MD, Mahnke-Braam L, Reznikoff WS. A functional analysis of the Tn5 transposase. Identification of domains required for DNA binding and multimerization. *EMBO J* 1994;**24**:166–177.
68. Normand C, Duval-Valentin G, Haren L, Chandler M. The terminal inverted repeats of IS911: requirements for synaptic complex assembly and activity. *J Mol Biol* 2001;**38**:853–871.
69. Tobianski DM, Buchner JM, Thiel WH, Gernert KM, Karls AC. Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-specific recombinases are required for catalysis of DNA inversion by Piv. *Mol Microbiol* 2001;**39**:641–651.
70. Rice P, Craigie R, Davies DR. Retroviral integrases and their cousins. *Curr Opin Struct Biol* 1996;**6**:76–83.
71. Polard P, Chandler M. Bacterial transposases and retroviral integrases. *Mol Microbiol* 1995;**15**:13–23.
72. Kulkosky J, Skalka AM. Molecular mechanism of retroviral DNA integration. *Pharmacol Ther* 1994;**61**:185–1203.
73. Hazuda DJ, Felock PJ, Hastings JC, Pramanik B, Wolfe AL. Differential divalent cation requirements uncouple the assembly and catalytic reactions of human immunodeficiency virus type 1 integrase. *J Virol* 1997;**71**:7005–7011.
74. Mizuuchi M, Baker TA, Mizuuchi K. Assembly of the active form of the transposase-Mu DNA complex: a critical control point in Mu transposition. *Cell* 1992;**70**:303–311.
75. Wlodawer A. Crystal structures of catalytic core domains of retroviral integrases and role of divalent cations in enzymatic activity. *Adv Virus Res* 1999;**52**:335–350.
76. Landree MA, Wibbenmeyer JA, Roth DB. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. *Genes Dev* 1999;**13**:3059–3069.
77. Kim DR, Dai Y, Mundy CL, Yang W, Oettinger MA. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. *Genes Dev* 1999;**13**:3070–3080.
78. Fugmann SD, Villey IJ, Ptaszek LM, Schatz DG. Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. *Mol Cell* 2000;**5**:97–107.
79. Engelman A, Mizuuchi K, Craigie R. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* 1991;**67**:1211–1221.
80. Mizuuchi K, Baker TA. Chemical mechanisms for mobilizing DNA. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, eds. *Mobile DNA II*. Washington, DC: ASM Press, 2001: 12–23.
81. Craig NL. Unity in transposition reactions. *Science* 1995;**270**:253–254.
82. Yu K, Lieber MR. The nicking step in V(D)J recombination is independent of synapsis: implications for the immune repertoire. *Mol Cell Biol* 2000;**20**:7914–7921.
83. Namgoong SY, Harshey RM. The same two monomers within a MuA tetramer provide the DDE domains for the strand cleavage and strand transfer steps of transposition. *EMBO J* 1998;**17**:3775–3785.
84. Jiang H, Yang JY, Harshey RM. Criss-crossed interactions between the enhancer and the att sites of phage Mu during DNA transposition. *EMBO J* 1999;**18**:3845–3855.
85. Dawson A, Finnegan DJ. Excision of the *Drosophila* mariner transposon Mos1. Comparison with bacterial transposition and V(D)J recombination. *Mol Cell* 2003;**11**:225–235.
86. Colot V, Haedens V, Rossignol JL. Extensive, nonrandom diversity of excision footprints generated by Ds-like transposon Ascot-1 suggests new parallels with V(D)J recombination. *Mol Cell Biol* 1998;**18**:4337–4346.
87. Coen ES, Carpenter R, Martin C. Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. *Cell* 1986;**47**:285–296.
88. Craigie R. Retroviral DNA integration. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, eds. *Mobile DNA II*. Washington, DC: ASM Press, 2001: 613–630.

89. Chaconas G, Harshey RM. Transposition of phage Mu DNA. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, eds. *Mobile DNA II*. Washington, DC: ASM Press, 2001: 384–402.
90. Kennedy AK, Guhathakurta A, Kleckner N, Haniford DB. Tn10 transposition via a DNA hairpin intermediate. *Cell* 1998;**95**:125–134.
91. Hickman AB, Li Y, Mathew SV, May EW, Craig NL, Dyda F. Unexpected structural diversity in DNA recombination: the restriction endonuclease connection. *Mol Cell* 2000;**5**:1025–1034.
92. Sarnovsky RJ, May EW, Craig NL. The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. *EMBO J* 1996;**15**:6348–6361.
93. May EW, Craig NL. Switching from cut-and-paste to replicative Tn7 transposition. *Science* 1996;**272**:401–404.
94. Agrawal A, Eastman QM, Schatz DG. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* 1998;**394**:744–751.
95. Hiom K, Melek M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* 1998;**94**:463–470.
96. Melek M, Gellert M. RAG1/2-mediated resolution of transposition intermediates: two pathways and possible consequences. *Cell* 2000;**101**:625–633.
97. Messier TL, O'Neill JP, Hou SM, Nicklas JA, Finette BA. In vivo transposition mediated by V(D)J recombinase in human T lymphocytes. *EMBO J* 2003;**22**:1381–1388.
98. Tsai CL, Schatz DG. Regulation of RAG1/RAG2-mediated transposition by GTP and the C-terminal region of RAG2. *EMBO J* 2003;**22**:1922–1930.
99. Elkin SK, Matthews A, Oettinger M. The C-terminal portion of RAG2 protects against transposition in vitro. *EMBO J* 2003;**22**:1931–1938.
100. Swanson PC, Volkmer D, Wang L. Full-length RAG-2, and not full-length RAG-1, specifically suppresses RAG-mediated transposition, but not hybrid joint formation or disintegration. *J Biol Chem* 2004;**279**:4034–4044.
101. Peters JE, Craig NL. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. *Genes Dev* 2001;**15**:737–747.
102. Melek M, Gellert M, van Gent DC. Rejoining of DNA by the RAG1 and RAG2 proteins. *Science* 1998;**280**:301–303.
103. Lewis SM, Hesse JE, Mizuuchi K, Gellert M. Novel strand exchanges in V(D)J recombination. *Cell* 1988;**55**:1099–1107.
104. Morzycka-Wroblewska E, Lee FE, Desiderio SV. Unusual immunoglobulin gene rearrangement leads to replacement of recombinational signal sequences. *Science* 1988;**242**:261–263.
105. Lee GS, Neiditch MB, Sinden RR, Roth DB. Targeted transposition by the V(D)J recombinase. *Mol Cell Biol* 2002;**22**:2068–2077.
106. Johnston PJ, MacPhail SH, Stamato TD, Kirchgessner CU, Olive PL. Higher-order chromatin structure-dependent repair of DNA double-strand breaks: involvement of the V(D)J recombination double-strand break repair pathway. *Radiat Res* 1998;**149**:455–462.
107. Shockett PE, Schatz DG. DNA hairpin opening mediated by the RAG1 and RAG2 proteins. *Mol Cell Biol* 1999;**19**:4159–4166.
108. Besmer E, et al. Hairpin coding end opening is mediated by RAG1 and RAG2 proteins. *Mol Cell* 1998;**2**:817–828.
109. Santagata S, et al. The RAG1/RAG2 complex constitutes a 3' flap endonuclease: implications for junctional diversity in V(D)J and transpositional recombination. *Mol Cell* 1999;**4**:935–947.
110. Yarnell Schultz H, Landree MA, Qiu JX, Kale SB, Roth DB. Joining-deficient RAG1 mutants block V(D)J recombination in vivo and hairpin opening in vitro. *Mol Cell* 2001;**7**:65–75.
111. Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 2001;**412**:607–614.
112. Paillard S, Strauss F. Analysis of the mechanism of interaction of simian Ku protein with DNA. *Nucleic Acids Res* 1991;**19**:5619–5624.
113. Neiditch MB, Lee GS, Huye LE, Brandt VL, Roth DB. The V(D)J recombinase efficiently cleaves and transposes signal joints. *Mol Cell* 2002;**9**:871–878.
114. Surette MG, Buch SJ, Chaconas G. Transpososomes: stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA. *Cell* 1987;**49**:253–262.
115. Krukltis R, Welty DJ, Nakai H. ClpX protein of *Escherichia coli* activates bacteriophage Mu transposase in the strand transfer complex for initiation of Mu DNA synthesis. *EMBO J* 1996;**15**:935–944.
116. Levchenko I, Luo L, Baker TA. Disassembly of the Mu transposase tetramer by the ClpX chaperone. *Genes Dev* 1995;**9**:2399–2408.
117. Mhammedi-Alaoui A, Pato M, Gama M-J, Toussaint A. A new component of bacteriophage Mu replicative transposition machinery: the *Escherichia coli* ClpX protein. *Mol Microbiol* 1994;**11**:1109–1116.
118. Wawrzynow A, et al. The ClpX heat-shock protein of *Escherichia coli*, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. *EMBO J* 1995;**14**:1867–1877.
119. Levchenko I, Yamauchi M, Baker TA. ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway. *Genes Dev* 1997;**11**:1561–1572.
120. Burton BM, Baker TA. Mu transpososome architecture ensures that unfolding by ClpX or proteolysis by ClpXP remodels but does not destroy the complex. *Chem Biol* 2003;**10**:463–472.
121. Lin WC, Desiderio S. Cell cycle regulation of V(D)J recombination-activating protein RAG-2. *Proc Natl Acad Sci USA* 1994;**91**:2733–2737.
122. Yurchenko V, Xue Z, Sadofsky M. The RAG1 N-terminal domain is an E3 ubiquitin ligase. *Genes Dev* 2003;**17**:581–585.
123. Jones JM, Gellert M. Auto-ubiquitylation of the V(D)J recombinase protein RAG1. *Proc Natl Acad Sci USA* 2003;**100**:15446–15451.
124. Pickart CM. Mechanisms underlying ubiquitination. *Annu Rev Biochem* 2001;**70**:503–533.
125. Deshaies RJ. SCF and cullin/RING H2-based ubiquitin ligases. *Annu Rev Cell Dev Biol* 1999;**15**:435–467.