The V(D)J Recombinase Efficiently Cleaves and Transposes Signal Joints

Matthew B. Neiditch,1 Gregory S. Lee,1 Leslie E. Huye,1 Vicky L. Brandt,2 and David B. Roth1,3,4
1Department of Immunology
2Department of Pediatrics
3Howard Hughes Medical Institute
Baylor College of Medicine
Houston, Texas 77030

Summary

V(D)J recombination generates two types of products: coding joints, which constitute the rearranged variable regions of antigen receptor genes, and signal joints, which often form on immunologically irrelevant, excised circular molecules that are lost during cell division. It has been widely believed that signal joints simply convert reactive broken DNA ends into safe, inert products. Yet two curious in vivo observations made us question this assumption: signal ends are far more abundant than coding ends, and signal joints form only after RAG expression is downregulated. In fact, we find that signal joints are not at all inert; they are cleaved quite efficiently in vivo and in vitro by a nick-nick mechanism and form an excellent substrate for RAG-mediated transposition in vitro, possibly explaining how genomic stability in lymphocytes may be compromised.

Introduction

During lymphocyte development, T cell receptor (TCR) and immunoglobulin (Ig) genes are assembled from separate coding DNA segments by a process called V(D)J recombination (schematized in Figure 1A). Adjacent to these coding segments are short recombination signal sequences (RSS) consisting of conserved heptamer and nonamer elements separated by either 12 or 23 nucleotides of spacer DNA. Efficient recombination requires two RSS (one of each type), which are recognized and bound by the recombinase proteins, RAG1 and RAG2, and brought together into a synaptic complex (Eastman et al., 1996; Steen et al., 1996; Hiom and Gellert, 1998). The RAG proteins then nick one strand of DNA precisely between the RSS and the coding segment, generating a free 3’ OH that is used to attack the opposite strand. This transsesterification reaction results in a coding end that terminates in a hairpin and a signal end terminating in a flush double-strand break (Figure 1A) (McBlane et al., 1995). After this cleavage step, the RAG proteins remain associated with the broken ends in the form of a postcleavage complex (Agrawal and Schatz, 1997; Hiom and Gellert, 1998). Ultimately, the two coding ends are joined imprecisely to form the coding joint (which creates the rearranged antigen receptor gene), and the two signal ends join perfectly, without loss of nucleotides, to form the reciprocal product, a signal joint (Figure 1A). Joining is dependent on the participation of several double-strand break repair factors, including Ku, DNA-PKcs, XRCC4, DNA ligase IV, and Artemis (Roth et al., 1992a; Taccioli et al., 1993; Li et al., 1995; Nussenzweig et al., 1996; Zhu et al., 1996; Gu et al., 1997; Frank et al., 1998; Grawunder et al., 1998; Moshous et al., 2001). Recent work has shown that the RAG proteins also play essential roles in joining both coding and signal ends (Qiu et al., 2001; Schultz et al., 2001).

Curiously, signal ends are much more abundant than coding ends in thymocytes and bone marrow cells of wild-type mice (Roth et al., 1992b; Schlissel et al., 1993; Zhu and Roth, 1995). It has been proposed that signal ends have substantially longer lifetimes than coding ends (Ramsden and Gellert, 1995; Zhu and Roth, 1995; Livak and Schatz, 1996), although it is not clear why this would be the case, particularly since coding joint formation involves more steps (and more DNA repair factors) than formation of signal joints. Nevertheless, this hypothesis has received some support from studies of a V(D)J recombination-inducible B cell line and by analysis of normal mouse thymocytes, which showed that signal ends persist until RAG expression is downregulated, allowing signal joint formation (Ramsden and Gellert, 1995; Livak and Schatz, 1996). These in vivo data, along with the observation that the RAG proteins bind tightly to signal ends in vitro, have led to the widely accepted view that the RAG proteins remain stably bound to the signal ends after cleavage and must be removed to allow joining (Ramsden and Gellert, 1995; Livak and Schatz, 1996; Zhu et al., 1996; Agrawal and Schatz, 1997). There is, however, an alternative explanation that has not been explored: signal joints (which, after all, contain an appropriate pair of RSS) might be recleaved by the RAG proteins.

Although the immunological role played by coding joints is obvious, the purpose served by signal joints is less clear. At some antigen receptor loci, the RSS are arranged in an orientation that necessitates inversional recombination. Under these circumstances, signal joint formation is required to restore chromosomal integrity. At many loci, however, recombination proceeds by a deletional mechanism in which signal joints are generated on extrachromosomal circles that are eventually lost during cell division. In these cases, the signal joint itself is of no apparent value to the cell—yet it is formed nonetheless. It has been suggested that the important function of signal joint formation in this situation is to convert potentially reactive signal ends (which might undergo illegitimate joining events) into safe, inert products (Hiom et al., 1998; Roth and Craig, 1998). The need for such a function became especially apparent with the discovery that signal ends, when complexed with the RAG proteins, are capable of transposition in vitro (Agrawal et al., 1998; Hiom et al., 1998), a reaction which could obviously compromise genomic stability. Formation of a signal joint would seem to provide an important mechanism for preventing such events.

The foregoing discussion raises a question of crucial importance: are signal joints truly inert in vivo, or are...
they susceptible to recleavage by the RAG proteins? Others have shown that signal joints can provide a single RSS for generating rare secondary rearrangements involving a second RSS located elsewhere (Lewis et al., 1985; Hesse et al., 1987). This is not surprising, as there is no reason why one of the two RSS present in a signal joint should not be available to the recombination machinery. Cleavage involving only the two RSS present in a signal joint, however, is another matter. Because these two RSS are immediately adjacent to each other, without intervening DNA, the prevailing view holds that steric constraints would render cleavage events that are dependent upon synapsis of these two signals impossible. Indeed, recombination is strongly inhibited if the distance between the RSS is decreased below a critical length (40 nt) (Lewis and Hesse, 1991; Sheehan and Lieber, 1993). Nevertheless, the question of whether signal joints are good substrates for RAG-mediated cleavage has not been directly addressed. Here we show that signal joints are efficiently cleaved in vivo and in vitro by a novel mechanism that does not require synapsis or hairpin formation and that they serve as efficient substrates for transposition in vitro. These data subvert the reigning assumption that signal joints are safe, dead-end reaction products and raise the surprising possibility that they may contribute to genomic instability in lymphocytes.

Results

Signal Joints Are Cleaved In Vivo

To determine whether signal joints are a substrate for cleavage in cells, we recovered a rearranged product of an in vivo V(D)J recombination reaction that bears a perfect signal joint. This molecule, pSJ (Figure 1C), was employed as a substrate for in vivo recombination using a well-established transient transfection assay system (Steen et al., 1996, 1997) in which expression vectors encoding truncated “core” RAG1 and RAG2 are cotransfected into fibroblasts along with a plasmid recombination substrate. Double-stranded cleavage at the 12-RSS was assayed using a standard semiquantitative ligation mediated-PCR (LM-PCR) assay (Roth et al., 1993). As expected, efficient RAG-dependent cleavage was observed using an unrearranged substrate (pJH289) containing both a 12- and 23-RSS (Figure 1B, lanes 3–6). Less efficient cleavage was detected using a substrate (p12) containing only a 12-RSS (lane 2), as described previously in extrachromosomal substrates (Steen et al., 1997; Han et al., 1999) and at a chromosomal TCR locus (Nakajima and Bosma, 1997). Cleavage of pSJ was remarkably efficient (lane 1), producing levels of signal ends substantially higher than those derived from p12 and within 2- to 3-fold of pJH289 (similar results were obtained from LM-PCR analysis of both the 12- and 23-RSS ends and with full-length RAG proteins in several independent transfections; data not shown). These data demonstrate that signal joints are efficiently cleaved in vivo.

Signal Joint Cleavage Proceeds In cis by a Nick-Nick Mechanism

Although introduction of a single-strand nick between the RSS and the coding segment can occur at a single RSS, a double-strand break typically involves hairpin formation, which under normal conditions depends upon synapsis of a 12/23 RSS pair (Hiom and Gellert, 1998; Yu and Lieber, 2000). A signal joint contains the requisite 12/23 pair, but the fact that the two RSS are adjacent, without any intervening DNA, prohibits intramolecular synapsis. Nevertheless, our in vivo data demonstrate that signal joints are efficiently cleaved to form signal ends terminating in flush double-strand breaks. We considered two possible explanations. Although the structure of a signal joint prohibits synapsis in cis, synapsis with a second RSS provided in trans would allow cleavage to proceed via the normal mechanism, producing one hairpin end and one blunt end (Figure 2A). Alternatively, cleavage could occur in cis by a nonstandard mechanism that bypasses the requirement for synapsis. Since nicking occurs efficiently at a single RSS in the absence of synapsis (McBlane et al., 1995; van Gent et al., 1996), we hypothesized that double-strand breaks could be formed by a novel “nick-nick” mechanism (Fig-
V(D)J Signal Joint Cleavage In Vivo and In Vitro

Figure 2. Mechanisms of Signal Joint Cleavage
(A) Signal joint cleavage could occur via a synaptic complex-dependent mechanism where nick and hairpin formation depend on the presence of a 12- and 23-RSS situated in trans. Step (1), binding of RAG proteins to input substrates forming a synaptic complex in trans; step (2), single-strand nicked intermediates; and step (3), hairpin coding and blunt signal end products.

(B) Synaptic complex-independent signal joint cleavage. Signal joint cleavage could occur via a nick-nick pathway. RAG proteins could cleave a single signal joint by nicking 5' of each RSS, which could occur either simultaneously or sequentially. Step (1), binding of RAG proteins to input substrates in cis; step (2), single-strand nicked intermediate; and step (3), blunt signal end products. Curved arrows indicate the position of nicking.

Figure 3. Signal Joint Cleavage In Vitro Proceeds In cis
Native polyacrylamide gel analysis of in vitro plasmid cleavage reactions. Dialysis buffer was substituted for purified RAG proteins where indicated (–). 208 and 232 bp cleavage products were not resolved from one another under the native gel conditions.

We next examined the products of signal joint cleavage by two dimensional native/denaturing gel electrophoresis, which readily reveals hairpins because they migrate at characteristic positions above the diagonal (hairpins run at twice their native size in the denaturing dimension). Cleavage of signal joints in vitro yielded uncleaved, cleaved, and nicked products, but no hairpins were detected even on long exposures (Figure 4A). These data demonstrate that signal joint cleavage proceeds by a hairpin-independent mechanism in vitro and are consistent with a nick-nick model. Furthermore, these results argue against the possibility that the signal joint substrates undergo synopsis under our in vitro cleavage conditions, as the standard cleavage mechanism should form hairpins from synaptic complexes composed of signal joints.

Analysis of cleavage of the same signal joint substrate in vivo revealed that the vast majority of products result from double-strand cleavage; nicked species were also observed below the diagonal (Figure 4B). Therefore, it appears that most signal joint cleavage in vivo proceeds via a nick-nick mechanism. A trace amount of hairpin product was also evident. This is expected, because both single-signal and trans cleavage of p12 and p23 have been observed previously in vivo (Han et al., 1999), and there is no reason that the 12- and 23-RSS present in a signal joint should not be available for trans cleavage by the normal cleavage mechanism.

To substantiate our finding that signal joint cleavage
Figure 4. Signal Joint Cleavage Occurs Predominantly via a Nick-Nick Mechanism In Vivo and In Vitro

(A) Two-dimensional native (N)/alkaline (D) gel electrophoresis of in vitro signal joint cleavage with purified proteins. For clarity, a lighter exposure of the radiolabeled marker lane is also included.

(B) Two-dimensional native (N)/alkaline (D) gel of in vivo signal joint cleavage. The uncleaved product is 440 nt. The 208 and 232 nt cleavage and nicked products and the 416 and 464 nt hairpinned products were not resolved from one another.

occurs without hairpin formation in vivo, we employed two RAG1 mutants that can nick but are incapable of hairpin formation. These mutants, K890A and R894A (Huye et al., 2002), are severely (at least 100-fold) defective for cleavage of standard plasmid substrates (Figure 5, lanes 5–6)—yet they efficiently cleaved signal joints (Figure 5, lanes 7–8). These data provide strong evidence that signal joint cleavage in vivo proceeds by a nick-nick mechanism.

Signal Joints Are Good Substrates for Transposition

Having established that signal joints can be efficiently cleaved both in vivo and in vitro, we next evaluated the ability of signal joints to participate in RAG-mediated transposition. Radiolabeled 12- and 23-RSS oligonucleotides were used as substrates for transposition using a protocol (Figure 6A), described previously, which measures transposition into an unlabeled supercoiled plasmid target (Hiom et al., 1998; Neiditch et al., 2001). As expected, most transposition products derived from the 12–23 RSS donors were linear (Figure 6B, lane 1), reflecting concerted integration of both a 12- and 23-RSS into both strands of the target plasmid (Hiom et al., 1998); no transposition was observed in the absence of the RAG proteins (lanes 2 and 4). Radiolabeled oligonucleotides containing signal joints also transposed efficiently (lane 3), giving levels of products only slightly lower than those generated by the control substrates. Again, most products were linearized, indicating that both ends produced by signal joint cleavage integrated into the target. These data demonstrate that signal joints are good substrates not only for cleavage, but also for transposition. The generation of predominantly linear products, which result from double-ended transposition, with very few nicked circular products resulting from single-ended transposition provides further evidence that the cleavage mechanism does not generate hairpin signal ends, because such ends should not be capable of integrating into the target. Finally, these data demonstrate that cleavage via the nick-nick pathway, which may not proceed via standard synaptic complex intermediates, is compatible with efficient transposition, at least in vitro.

Discussion

Efficient Signal Joint Cleavage by a Novel Mechanism

The majority of V(D)J recombination events in lymphocyte precursors produce extrachromosomal circles containing signal joints that serve no direct immunological purpose. Because unrepaired signal ends pose a danger to both the cell and the organism, the prevailing view has been that signal joint formation serves as a safety mechanism to preserve genomic integrity. This assumption now proves to be unwarranted. Our results clearly demonstrate that signal joints are far from inert products. Indeed, not only do signal joints undergo efficient cleavage both in vivo and in vitro, but cleavage in vitro proceeds by a nick-nick mechanism that produces blunt signal ends capable of transposition.

Several lines of evidence indicate that cleavage in vivo proceeds predominantly by the same mechanism. First, two-dimensional gels reveal that most cleavage produces double-strand breaks, and only a very small fraction of these products terminates in hairpins. Second, signal joints are abundant in vivo in thymocytes.
and, as discussed below, are likely to give rise to the high levels of the corresponding signal ends noted in these cells (Roth et al., 1992b). Careful examination of the ends by a variety of techniques has failed to reveal any evidence for hairpins at signal ends, even in cells bearing mutations that block hairpin opening (Roth et al., 1992a, 1992b, 1993; Zhu and Roth, 1995; D.B.R., unpublished data). Third, whereas P nucleotide insertions derived from hairpin opening are commonly observed at coding joints (Lafaille et al., 1989; Meier and Lewis, 1993), P nucleotides have not been described at signal joints, suggesting that signal joints are not cleaved by a nick-hairpin mechanism and then rejoined. Finally, mutant RAG1 proteins with specific defects in hairpin formation efficiently cleave signal joints. Together, these data show that the efficient signal joint cleavage we have observed in vivo occurs mainly via a nick-nick mechanism that leaves both signal ends blunt and potentially available for transposition events. This novel pathway is remarkable in that it apparently bypasses both synopsis and hairpin formation, two hallmarks of the standard cleavage mechanism.

Signal Joint Cleavage: Implications for Metabolism of Signal Ends

Our data provide a straightforward explanation for several puzzling observations. First, signal ends are much more abundant than coding ends in normal lymphocyte precursors, even though both kinds of ends are created by the same cleavage event. We and others had suggested that signal ends might be joined much more slowly than coding ends (Roth et al., 1992a, 1992b; Ramsden and Gellert, 1995; Zhu and Roth, 1995; Livak and Schatz, 1996), although it was not clear why signal joint formation would proceed with slower kinetics. After all, both reactions depend on the same set of joining factors (with additional requirements for coding joint formation), and signal joint formation actually requires fewer steps (coding joint formation entails hairpin opening and end processing). Signal joint recleavage provides a much more parsimonious explanation for the relative over-abundance of signal ends.

The second curiosity is that signal joints do not accumulate to appreciable levels until after the RAG proteins are downregulated (Ramsden and Gellert, 1995; Livak and Schatz, 1996). It has been suggested that the stable complexes formed by the RAG proteins on signal ends might block signal joint formation (Ramsden and Gellert, 1995; Livak and Schatz, 1996; Zhu et al., 1996; Agrawal and Schatz, 1997), but this model does not explain how the stable postcleavage complexes are removed from the ends after downregulation of RAG expression. Moreover, we now know that the RAG proteins actually play critical roles in both signal and coding joint formation (Qiu et al., 2001; Schultz et al., 2001; Huye et al., 2002). Signal joint recleavage provides a simple explanation for the observation that coding joints accumulate in the presence of high levels of RAG proteins while signal joints do not.

A third puzzle is presented by the differential effects of mutations that block joining on levels of coding and signal ends. Analysis of lymphocyte precursors from normal and mutant mice has revealed that mutations that block coding joint formation substantially increase levels of coding ends (>100-fold), but signal end levels are not appreciably affected by mutations that block signal joint formation (Roth et al., 1992a; Zhu and Roth, 1995; Zhu et al., 1996). This conundrum can be elegantly resolved by signal joint recleavage; in normal lymphocyte precursors, levels of signal ends depend upon the rate of both joining and recleavage. If recleavage is relatively rapid, decreasing the rate of joining will not substantially affect the levels of signal ends. By contrast, levels of coding ends are quite low in normal cells, and recleavage does not contribute to the steady-state level of coding ends, which we suggest is determined by the rate of joining.

Signal Joints and the Biology of V(D)J Recombination

Our data demonstrate that signal joints are not inert products and thus may not play as much of a role in
preventing aberrant rearrangements as previously thought. Why, then, are they formed? Signal joint formation is necessary to restore chromosomal integrity in the case of inversional recombination. The only difference between gene segments that undergo inversional and deletional recombination lies in the orientation of the RSS, and there is no evidence that the recombinase keeps track of this information as the reaction proceeds. We therefore propose that the capability to form signal joints is “hard-wired” into the recombination mechanism.

Inversional recombination produces signal joints that are retained in the chromosome. Efficient recleavage of these joints would have several undesirable consequences. We suggest that recleavage of chromosomal signal joints is discouraged by the same mechanisms that operate to ensure allelic exclusion: control of their accessibility to the recombinase. Although the same recombination machinery carries out rearrangement of all TCR and Ig loci, V(D)J recombination occurs at appropriate loci in a carefully ordered fashion that is controlled by accessibility of the substrate (Sleckman et al., 1998; Roth and Roth, 2000). These same mechanisms should prevent recleavage of signal joints formed by recombination once chromosomal loci are “closed.” It is not clear whether recleavage of signal joints on extrachromosomal excision circles is regulated in the same way; excision circles, which can persist in naïve T lymphocytes (Douek et al., 1998), may constitute a source of signal ends—and, perhaps, aberrant rearrangements—in cells any time RAG proteins are expressed.

Implications for Genomic Stability
It is not uncommon in biology for the resolution of one puzzle to create another. The fact that the potential for transposition in vivo is greater than previously thought makes it even more remarkable that transposition has not yet been detected in living cells. Perhaps we have not been looking in the right place; recent work suggests that there may be a bias toward transposition into the antigen receptor loci themselves (Neiditch et al., 2001) or into distorted DNA structures (Lee et al., 2002). While the data presented here provide a plausible explanation for the high levels of signal ends observed in immature lymphocytes, the apparent low frequency of transposition of these abundant RSS ends in vivo remains unexplained and may reflect the presence of stringent control mechanisms to prevent this dangerous reaction. Clearly it is necessary to dissect the mechanism of transposition in greater detail in order to understand its regulation.

Experimental Procedures

Recombination Substrates
pSJ was produced by in vivo recombination of a previously described recombinase substrate pJH289 (Lewis et al., 1988). The pSJ signal joint contains no nucleotide insertions or deletions, as confirmed by DNA sequencing. p12 (pGL3) and p23 (pGL4) are 12- or 23-RSS-only substrates created from pJH289 by deletion of either the 23- or 12-RSS, respectively.

Transfections
Chinese hamster ovary fibroblasts (RMP41) were transiently transfected with 2.1 μg truncated RAG1 (pMS127) (Sadofsky et al., 1993), 1.8 μg truncated RAG2 (pMS216) (Sadofsky et al., 1994), and 5 μg recombination substrate (p12, p23, p12/23, or pSJ) using the Fugene-6 transfection reagent (Roche) as previously described (Landree et al., 1999). RAG1 mutants K909A and R894A were generated in the previously described plasmid expression vector pEBG-1.JN. Transfections containing mutant RAG1 proteins were compared to transfections with the wt RAG1 (pEBG-1.JN) and wt RAG2 (pEBG-2.JC) expression vectors previously described (Spanopoulou et al., 1996). Full-length RAG1 (pJH548) and RAG2 (pJH549) expression vectors were previously described (Sadofsky et al., 1993) DNA was harvested 48 hr after initial transfection as described (Landree et al., 1999) and resuspended in 50 μl of TE. Transfections with pSJ and RAG proteins were repeated at least 10 times, accompanied by p12, p23, or p12/23 controls, and yielded consistent results.

In Vivo Signal End Assays
Signal ends were detected from transfections by ligating one-fiftieth of the harvested DNA to the adaptor created by annealed oligonucleotides DR19 and DR20 as previously described (Roth et al., 1993). PCR was then performed using the DR20 and DR99 primers. PCR products were electrophoresed through 6% polyacrylamide gels at 110 volts for 60 min, transferred to nylon membranes, and detected by hybridization to the radiolabeled probe DR69 and visualized with a PhosphorImager.

In Vitro Plasmid Cleavage Assays
10 μl preincubations were incubated at 37°C for 20 min and contained 150 ng of each wild-type truncated core GST-RAG1 and truncated core GST-RAG2 (Spanopoulou et al., 1996), purified as previously described (Spanopoulou et al., 1996; Sawchuk et al., 1997; Schultz et al., 2001), and 20 ng of the indicated recombination substrate(s) (p12, p23, pJH289, or pSJ) in 50 mM K+-HEPES (pH 8.0), 26 mM KCl, 4 mM NaCl, 1 mM DTT, 5 mM CaCl2, 1 μg BSA, and 200 ng HMGl (Schultz et al., 2001). Preincubations were then spiked with 1 μl of 55 mM MgCl2 and incubated at 37°C for an additional 2.5 hr. Reactions were stopped by addition of 11 μl stop buffer (100 mM Tris [pH 8.0], 10 mM EDTA, 0.2% sodium dodecyl sulfate [SDS], 0.35 mg proteinase K/ml) and continued incubation at 37°C. Reactions were then depurinated by phenol/chloroform extraction and DNA was concentrated by ETOH precipitation. DNA was resuspended in TE and digested with 2.5 U PvuII for 1.5 hr at 37°C. The digests were then separated by electrophoresis through 6% native polyacrylamide gels (80 volts, 145 min). DNA was transferred to nylon membranes, and products were detected by hybridization to a randomly primed probe created from the PvuII fragment containing the signal joint from pSJ.

Two-Dimensional Native/Alkaline Gel Electrophoresis
In vitro and in vivo signal joint cleavage reactions were digested with PvuII for 1.5 and 2 hr at 37°C, respectively. Samples were then electrophoresed through a 2% GTG SeaKem agarose gel (FMC BioProducts) under native conditions for 475 volt · hrs. The gel was stained with Coomassie blue R-250 and scanned with a PhosphorImager. All TCR and Ig loci, V(D)J recombination occurs at appropriate loci in a carefully ordered fashion that is controlled by accessibility of the substrate (Sleckman et al., 1998; Roth and Roth, 2000). These same mechanisms should prevent recleavage of signal joints formed by recombination once chromosomal loci are “closed.” It is not clear whether recleavage of signal joints on extrachromosomal excision circles is regulated in the same way; excision circles, which can persist in naïve T lymphocytes (Douek et al., 1998), may constitute a source of signal ends—and, perhaps, aberrant rearrangements—in cells any time RAG proteins are expressed.
gels (FMC BioProducts), dried, and visualized by Phosphorimager analysis as previously described (Neiditch et al., 2001).

Acknowledgments

We thank H. Yarnall Schultz and J. Yang for valuable comments on the manuscript; H. Gilbert for valuable input regarding reaction kinetics; K. Freeman and J. Yang for valuable scientific input; and J.L. Neiditch and J.P. Neiditch for helpful discussions and unwavering support. D.B.R. is an Assistant Investigator of the Howard Hughes Medical Institute. This work was supported by a grant from the National Institutes of Health (AI-36420).

Received: November 21, 2001
Revised: February 6, 2002

References


