Cell, Vol. 109, 1–4, April 5, 2002, Copyright ©2002 by Cell Press

Does Artemis End the Hunt for the Hairpin-Opening Activity in V(D)J Recombination?

Minireview

Mark S. Schlissel[1]
Division of Immunology
Department of Molecular and Cell Biology
University of California
Berkeley, California 94720

When complexed with the catalytic subunit of DNA-dependent protein kinase (DNA-PK$_{cs}$), the recently discovered dsDNA break repair protein named Artemis acquires the ability to open hairpin DNA molecules in vitro. Both genetic and biochemical data point toward a physiological role for this complex as the elusive hairpin-opening activity in V(D)J recombination.

Scientists often take great pleasure in the parsimony of nature—how structural motifs, specific proteins, or even entire biochemical pathways can be mixed and matched to serve a variety of biological functions. A striking example of this is found in V(D)J recombination, which uses site-specific DNA recombination to generate an enormous and diverse repertoire of antigen receptor genes within developing lymphocytes of the vertebrate immune system (reviewed in Fugmann et al., 2000). V(D)J recombination utilizes two lymphocyte-specific proteins, RAG1 and RAG2, to recognize conserved recombination signal sequences (RSSs) flanking a pair of rearranging immunoglobulin (Ig) or T cell receptor (TCR) gene segments and to introduce two dsDNA breaks (Figure 1). These broken DNA ends are then repaired to form coding and signal joints by the nonhomologous end-joining (NHEJ) machinery expressed in all cells. The NHEJ pathway protects the genome from the deleterious effects of ionizing radiation and other insults which produce dsDNA breaks (reviewed in Jeggo, 1998). Prior studies of dsDNA break repair in mammalian cell lines have provided insights critical to our understanding of the mechanism of V(D)J recombination, and more recently, studies of V(D)J recombination have returned the favor, leading to a deeper understanding of NHEJ. An example of such synergy is a paper in a recent issue of *Cell* by Ma et al. (2002), which describes the interaction between two proteins involved in both dsDNA break repair and V(D)J recombination, Artemis and DNA-dependent protein kinase (DNA-PK). This minireview will briefly describe how the V(D)J recombinase generates specific dsDNA breaks, and then focus on recent mechanistic studies of how RAG1, RAG2, and NHEJ proteins collaborate in their productive repair.

The Lymphocyte-Specific RAG Proteins Initiate V(D)J Recombination by Generating a Pair of dsDNA Breaks

RAG1 and RAG2 form a multimeric complex which recognizes RSSs and introduces a nick precisely at the coding segment-RSS border (Fugmann et al., 2000). The RAGs then activate the 3′OH on the coding segment side of the break to attack across the DNA duplex and hydrolyze the opposing phosphodiester bond, thus generating two DNA ends—a blunt and 5′ phosphorylated signal end, and a covalently closed hairpin coding end (Figure 1). The ability of the recombinase to recognize RSSs is affected by aspects of chromatin structure that alter the accessibility of the DNA substrate in a developmentally regulated fashion. Thus, coexpression of RAG1 and RAG2 is sufficient to activate V(D)J recombination on a cotransfected episomal rearrangement reporter construct, but not within chromosomal Ig and TCR loci, in nonlymphoid cells. In addition, specific features of individual RSSs themselves contribute to the regulation of recombinase targeting during lymphoid development.

Promiscuously Expressed dsDNA Break Repair Proteins Are Required for the Joining Steps in V(D)J Recombination

The first inkling that components of the dsDNA break repair machinery might be involved in the formation of joints during V(D)J recombination came from an analysis of the murine autosomal recessive *scid* mutation. *Scid* (severe combined immunodeficiency disease) mice have a defect in lymphocyte development which results in the near absence of mature B and T lymphocytes and a consequent profound immunodeficiency (reviewed in Bosma and Carroll, 1991). Fibroblasts from *scid* mice display a several-fold increase in sensitivity to ionizing radiation, indicating that the mutant phenotype is not confined to lymphoid cells. At a molecular level, *scid* lymphoid progenitors start the process of V(D)J recombination by the appropriate generation of dsDNA breaks at RSSs, but fail to efficiently form coding joints; surprisingly, signal joints are formed with a nearly normal efficiency. Detailed studies comparing recombination intermediates in wild-type and *scid* mice revealed an accumulation of hairpin coding ends in *scid* mutant lymphoid progenitors, leading to the hypothesis that the *scid* gene product was involved, either directly or indirectly, in opening DNA hairpins (Roth et al., 1992; see below). The rare coding joints which do form in *scid* lymphocytes are characterized by large deletions indicating that other DNA repair systems can partially rescue this NHEJ defect.

Given the dual phenotype of the *scid* mutation, investigators proceeded to ask whether other DNA repair mutants affected V(D)J recombination. In one particularly insightful set of experiments, Taccioli et al. (1993) cotransfected expression vectors encoding RAG1 and RAG2 along with a V(D)J recombination reporter construct into a set of Chinese hamster ovary (CHO) cell mutants with well characterized defects in dsDNA break repair. These experiments led to the identification and subsequent molecular characterization of Ku70, Ku80, DNA-PKcs, XRCC4, and DNA ligase IV as being involved in both V(D)J recombination in lymphocytes and dsDNA break repair in all cell types examined (reviewed in Jeggo, 1998). Ku70 binds to DNA ends and forms a heterodimer with Ku80. Ku80 binds to DNA-PKcs (forming the so-called DNA-PK holoenzyme), presumably localizing it to the sites of DNA damage. DNA-PKcs, shown to be the product of the *scid* gene, is a 469 kDa serine/

[1] Correspondence: mss@uclink4.berkeley.edu

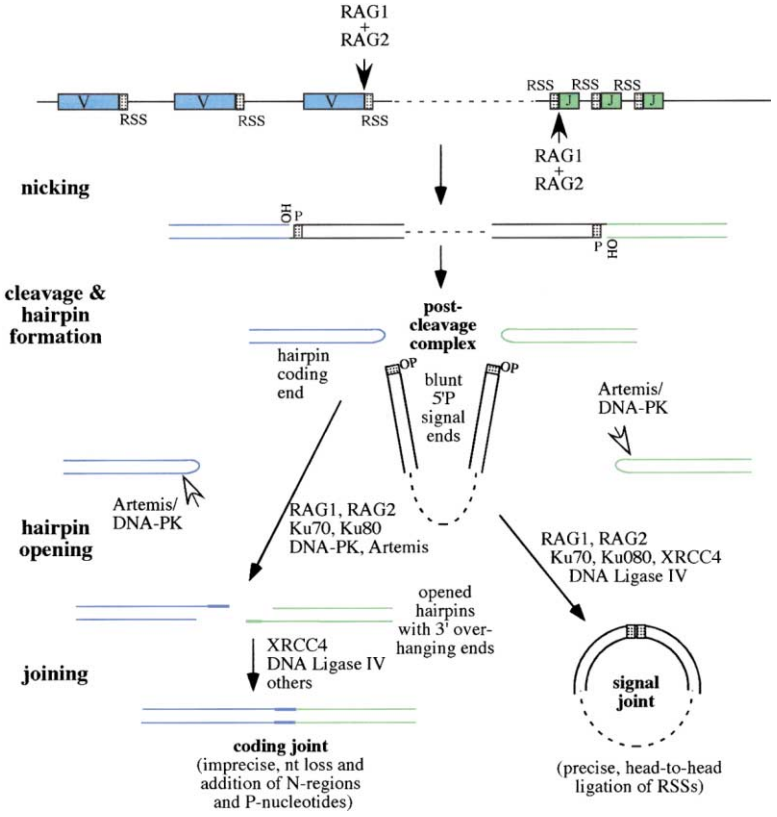

Figure 1. The V(D)J Recombination Reaction Pathway

Ig and TCR loci consist of tandem arrays of rearranging gene segments (V in blue and J in green), each flanked by a recombination signal sequence (RSS, hatched box). Early steps in the reaction (nicking, cleavage, and hairpin formation) are catalyzed by the lymphocyte-specific proteins RAG1 and RAG2, while the later steps (hairpin opening and joining) require widely expressed dsDNA break repair proteins in addition to the RAGs. The "post-cleavage complex" contains four DNA ends along with a series of associated protein factors. The open arrowheads indicate the positions at which a complex of Artemis and DNA-PK$_{cs}$ opens hairpin ends prior to coding joint formation. The Nbs1/Mre11/Rad50 complex is not pictured but may be involved in stabilizing the post-cleavage complex.

threonine kinase with homology to PI kinases. XRCC4 associates with DNA ligase IV and stimulates its ligase activity. Mutations in DNA-PKcs result in defective coding joint but preserved signal joint formation, while mutations in each of the other NHEJ proteins are associated with defects in both signal and coding joint formation. The role of these proteins in V(D)J recombination was confirmed by targeted gene disruption in mice. Null mutations in genes encoding DNA-PKcs, Ku70, Ku80, XRCC4, and DNA ligase IV all result in *scid*-like phenotypes, defective dsDNA break repair, and, in some cases, other nonlymphoid developmental defects.

Recently, the gene responsible for a novel, human form of scid (called radiation-sensitive or RS-SCID) was cloned (Moshous et al., 2001). Fibroblasts from RS-SCID patients show increased sensitivity to ionizing radiation and a defect in V(D)J recombination that is quite similar to that of murine *scid* cells—a specific block in coding joint formation (Nicolas et al., 1998). These cells contain normal Ku70, Ku80, DNA-PKcs, XRCC4, and DNA ligase IV genes, thus ruling out mutations in these known components of the NHEJ pathway. *Artemis* (named after the Greek goddess of the hunt, who is also the protector of children), the gene responsible for RS-SCID, was identified by positional cloning and encodes a 78 kDa protein with homology to metallo-β-lactamases.

The major effort in this field now is to determine the precise roles of these dsDNA break repair proteins and the RAG proteins in the post-cleavage steps of V(D)J recombination. These steps include opening of the coding-end hairpins, coding-end processing, and joint formation. While signal joints are simply precise, head-to-head fusions of signal ends, coding joints are much more diverse (Fugmann et al., 2000). They display variable nucleotide loss or addition. The additions are of two types—N regions and P nucleotides. N regions are nontemplated nucleotides added by the lymphoid-specific enzyme terminal deoxynucleotidyl transferase (TdT). P nucleotides are short (usually 1 to 4 nucleotides) palindromic repeats of sequences initially present at one or the other coding end sequence. The rare coding joints formed in *scid* lympocytes often contain much longer runs of P nucleotides. Several workers have proposed that P nucleotides are generated by repair of overhanging ends generated during hairpin opening. Structural studies of open coding ends in DNA purified from lymphoid progenitors active in V(D)J recombination showed that they most often contain 2 to 7 nucleotide 3′ overhangs (Livak and Schatz, 1997; Schlissel, 1998).

The Hunt for Hairpin Opening Activity

One report suggested that a complex of NHEJ proteins (Nbs1, Mre11, and Rad50) might be the hairpin-opening activity in V(D)J recombination (Paull and Gellert, 1999). *Mre11* and *Rad50* are mammalian homologs of yeast genes which play important roles in dsDNA break repair and homologous recombination. *Nbs1* is the gene responsible for the human disease Nijmegen breakage syndrome which is characterized by radiation sensitivity and chromosomal fragility. This protein complex is capable of opening synthetic DNA hairpins in vitro, but the reaction only occurs in the presence of Mn^{2+}, a nonphysiological divalent cation known to relax nuclease specificity. Despite the fact that hairpin opening by this complex produces a short 3′ overhanging end, it is

unlikely to be the predominant hairpin-opening activity in V(D)J recombination since coding joint formation appears normal in lymphocytes from NBS patients. The Nbs1/Mre11/Rad50 complex is likely to play some role in V(D)J recombination since Nbs1 localizes to sites of dsDNA breakage in T cell progenitors undergoing V(D)J recombination.

Although DNA-PKcs (*scid*) mutant lymphoid cells generate but fail to efficiently open and further process coding ends, DNA-PK itself is unlikely to be the hairpin nuclease. This is because DNA-PK_{cs} lacks homology to known nucleases and the purified protein fails to display nuclease activity in vitro. Several groups have shown, however, that the RAG proteins themselves have the ability to open DNA hairpins under certain in vitro conditions. One reflection of this hairpin-opening activity is the ability of RAG1 and RAG2 to catalyze the reversal of the hairpin formation reaction. In this reaction, the 3′ OH of a blunt DNA duplex (a signal end, for example) acts as the nucleophile which attacks and breaks a phosphodiester bond near the tip of a hairpin duplex (Melek et al., 1998), leading to the formation of so-called open-shut and hybrid joints. Furthermore, in the presence of Mn^{2+}, RAG1 and RAG2 can open synthetic DNA hairpins to give either blunt or short 5′ overhanging DNA ends (Besmer et al., 1998; Shockett and Schatz, 1999). The RAG complex can also nick strands near the ends of nonhairpin DNA duplexes. Interestingly, the RAG proteins can open hairpins in the presence of Mg^{2+} only when they are generated by RAG-mediated RSS cleavage. This may have to do with the potential colocalization of hairpin coding ends and RAG proteins in a post-cleavage DNA-protein complex. There is biochemical evidence that both signal and coding ends do form such complexes after cleavage (Fugmann et al., 2000). These hairpin opening and endonuclease activities resemble the reversal of the RAG-mediated cleavage activity described above, with the nucleophile being a water molecule rather than the 3′ OH of a nicked or broken DNA strand. The outstanding question, however, is whether this hairpin opening activity reflects a physiological function of the RAG proteins with relevance to V(D)J recombination in vivo, or a biochemical activity of purified, recombinant proteins of little relevance to the biology of this process.

Genetic approaches have been used to decipher the mechanistic roles of the RAG proteins in the various steps of the V(D)J recombination reaction, including hairpin opening. Early studies identified the minimal essential (or core) domains of each protein using a transient cotransfection assay in nonlymphoid cells. It is important to note that due to the insolubility of full-length RAG1 and RAG2, all biochemical studies to date of V(D)J recombination have utilized these truncated "core" versions of the RAG proteins. While core RAG1 (amino acids 384–1008) and core RAG2 (amino acids 1–387) display recombination activity which is about 50% that of wild-type in transfection assays, recent data suggest that the non-core domains may play a role in suppressing the ability of the RAGs to catalyze hybrid joining and perhaps DNA transposition (Sekiguchi et al., 2001). Thus, it is possible that the ability of core RAG1 and core RAG2 to open DNA hairpins in vitro does not represent a function of the full-length proteins in vivo. Nearly every point mutation in either RAG1 or RAG2 which interferes with the RSS cleavage activity of the recombinase also inhibits its ability to open DNA hairpins in vitro. Importantly, however, one point mutation in RAG1 and two point mutations in RAG2 maintain normal RSS cleavage activity but display severely impaired hairpin opening in vitro (Yarnell-Schultz et al., 2001; Qiu et al., 2001). While these mutants catalyze only very inefficient coding joint formation in vivo, their ability to form hybrid joints is surprisingly unimpaired. In addition, the rare coding joints which can be recovered from cells expressing these RAG mutants lack the large deletions and long P nucleotide insertions characteristic of *scid* lympocytes. Further complicating the interpretation of this mutant phenotype, however, is the fact that the in vivo activity of all three mutants was only studied in the context of the core, rather than full-length, RAG1 and RAG2. Thus, it remains possible that full-length RAG proteins carrying these mutations might form coding joints with near normal efficiency.

Are RAG Proteins the Hairpin-Opening Nuclease?

Additional evidence has accumulated both for and against the hypothesis that RAG proteins are the physiologic hairpin-opening activity in V(D)J recombination. As noted above, the RAG proteins remain bound to RSS and coding ends after in vitro cleavage, leading to the possibility that both coding and signal ends might be held in a post-cleavage complex in vivo (Fugmann et al., 2000). Furthermore, continued presence of RAG1 and RAG2 is required for coding joint formation in a variety of different in vitro systems (Leu et al., 1997; Ramsden et al., 1997; Weis-Garcia et al., 1997). Thus, the RAG proteins are present on the scene, possess the biochemical capacity to open hairpins, and are required for some post-cleavage aspect of V(D)J recombination. In addition, various mutations in RAG1 and RAG2 affect coding joint formation to a greater extent than RSS cleavage as noted above. Arguing against an enzymatic role for the RAGs in hairpin opening, however, are the phenotypes of DNA-PK_{cs}- and Artemis-deficient lymphoid cells. In both cases, RAG proteins are present and RSS cleavage occurs, yet hairpin opening fails to occur. It was proposed, however, that DNA-PK might have to phosphorylate the RAG proteins in order to activate their hairpin-opening activity (although there is no evidence for such a requirement in vitro), thus explaining the *scid* phenotype. This is not likely to be the case since mutation of all four potential DNA-PK phosphorylation sites in RAG1 and RAG2 did not interfere with coding joint formation in transfected cells (Lin et al., 1999). Furthermore, the structural characteristics of hairpins opened by the core RAGs in vitro (short 5′ overhanging or blunt ends) differ from those observed in open coding ends in lymphoid progenitor cell genomic DNA ex vivo (2 to 7 nucleotide 3′ overhanging ends; Livak and Schatz, 1997; Schlissel, 1998). Thus, there is significant evidence leading to the conclusion that the RAG proteins might not be the predominant enzymatic activity responsible for hairpin opening.

Artemis and DNA-PK_{cs} Form a Protein Complex Which Can Open Hairpins In Vitro

If the RAGs are not responsible for hairpin opening, than what factor(s) is? The Ku proteins and DNA-PKcs have been extensively studied and lack nuclease activity.

Given the observation that both DNA-PKcs mutant (*scid*) murine cells and *Artemis* mutant human cells show similar defects in dsDNA break repair and coding joint formation, Lieber and colleagues decided to test the hypothesis that these two proteins were part of a larger multiprotein complex involved in both processes (Ma et al., 2002). They discovered that DNA-PK_{cs} and Artemis form a complex both in vitro and in vivo, and that Artemis is a substrate for the DNA-PK_{cs} kinase activity. Given the structural homology of Artemis to metallo-β-lactamases (a group of enzymes which use water as the nucleophile to break covalent bonds), these investigators went on to test Artemis for nuclease activity in vitro. They found that recombinant Artemis has a 5′ to 3′ exonuclease activity in the presence of Mg^{2+}, but lacks any nuclease activity in the presence of Mn^{2+}. Remarkably, when purified DNA-PK_{cs} was added, Artemis acquired endonuclease activity, allowing it to cleave DNA hairpins as well as 5′ and 3′ overhanging ends. The observations that addition of DNA-PK_{cs} alters the specificity of the Artemis-associated nuclease and that an Artemis point mutant lacks any nuclease activity makes it very unlikely that a contaminating nuclease is responsible for these results. These investigators went on to show that DNA-PK_{cs} must phosphorylate and remain in a complex with Artemis to fully activate this hairpin opening activity. Hairpin opening by the Artemis-DNA-PK_{cs} complex could occur on coding ends generated by RAG-mediated RSS cleavage and presumably still contained within a post-cleavage complex. Importantly, these opened hairpins contain 3′ overhanging ends very similar in structure to those observed in purified DNA from lymphoid progenitors. The one critical experiment missing from this otherwise compelling study, however, is an examination of recombination reaction intermediates in Artemis-deficient cells. If the Artemis-DNA-PK_{cs} complex is the hairpin-opening nuclease, then hairpin coding ends should accumulate in Artemis-deficient cells as they do in DNA-PK_{cs}-deficient cells.

Remaining Questions

Despite strong evidence suggesting that the Artemis/DNA-PK_{cs} complex is the hairpin-opening nuclease in V(D)J recombination, a number of important questions remain. First, what role do the Ku proteins play in the post-cleavage steps of V(D)J recombination? Ku-deficient cells fail to form signal or coding joints. Although previous studies have suggested that association with the Ku complex increases the kinase activity of DNA-PK_{cs}, Ma et al. show that Ku does not effect the hairpin-opening activity of the Artemis/DNA-PK_{cs} complex in vitro. It remains possible, however, that the Ku complex may serve to recruit Artemis/DNA-PK_{cs} to the post-cleavage complex in vivo and that this recruitment activity is not necessary in vitro. Second, if the RAGs themselves are not responsible for hairpin opening, what is their precise role in the joining phase of V(D)J recombination and how might one account for the three RAG point mutants which selectively affect hairpin opening in vitro and coding joint formation in vivo? One possibility is that the RAGs are involved in maintaining the integrity of the post-cleavage complex so that broken-ended DNA molecules don't escape the complex. This is particularly important given the ability of RAG proteins to catalyze the transposition of RSS-ended DNA fragments in vitro (reviewed in Fugmann et al., 2000). The Nbs1/Mre11/Rad50 complex may also play a role in post-cleavage complex stability given the phenotype of Nbs1-deficient human cells. These various possibilities could be tested by analyzing V(D)J recombination in RAG mutant knockin and Nbs1 knockout mice. Third, detailed structural information on the post-cleavage complex will be critical to understand its function. What are its components and what are their structural and functional relationships to each other? Finally, what is the role of Artemis, with or without the collaboration of DNA-PK_{cs}, in dsDNA break repair? While DNA hairpins are not thought to be generated as a consequence of environmentally mediated DNA damage or its repair, the Artemis/DNA-PK_{cs} endonuclease activity might serve to process 3′ overhanging ends, which are intermediates in some pathways of dsDNA break repair. While final answers to these questions will require much additional work, one can predict with certainty that the synergy between investigators studying V(D)J recombination and dsDNA break repair will continue to produce novel insights of relevance to both communities.

Selected Reading

Besmer, E., Mansilla-Soto, J., Cassard, S., Sawchuk, D.J., Brown, G., Sadofsky, M., Lewis, S.M., Nussenzweig, M.C., and Cortes, P. (1998). Mol. Cell *2*, 817–828.

Bosma, M.J., and Carroll, A.M. (1991). Annu. Rev. Immunol. *9*, 323–350.

Fugmann, S.D., Lee, A.I., Shockett, P.E., Villey, I.J., and Schatz, D.G. (2000). Annu. Rev. Immunol. *18*, 495–527.

Jeggo, P.A. (1998). Adv. Genet. *38*, 185–218.

Leu, T.M.J., Eastman, Q.M., and Schatz, D.G. (1997). Immunity *7*, 303–313.

Lin, J.M., Landree, M.A., and Roth, D.B. (1999). Mol. Immunol. *36*, 1263–1269.

Livak, F., and Schatz, D.G. (1997). J. Mol. Biol. *267*, 1–9.

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M.R. (2002). Cell *108*, 781–794.

Melek, M., Gellert, M., and van Gent, D.C. (1998). Science *280*, 301–303.

Moshous, D., Callebaut, I., de Chasseval, R., Corneo, B., Cavazzana-Calvo, M., Le Deist, F., Tezcan, I., Sanal, O., Bertrand, Y., Philippe, N., et al. (2001). Cell *105*, 177–186.

Nicolas, N., Moshous, D., Cavazzana-Calvo, M., Papadopoulo, D., de Chasseval, R., Le Deist, F., Fischer, A., and de Villartay, J.P. (1998). J. Exp. Med. *188*, 627–634.

Paull, T.T., and Gellert, M. (1999). Genes Dev. *13*, 1276–1288.

Qiu, J.X., Kale, S.B., Yarnell Schultz, H., and Roth, D.B. (2001). Mol. Cell *7*, 77–87.

Ramsden, D.A., Paull, T.T., and Gellert, M. (1997). Nature *388*, 488–491.

Roth, D.B., Menetski, J.P., Nakajima, P.B., Bosma, M.J., and Gellert, M. (1992). Cell *70*, 1–9.

Schlissel, M.S. (1998). Mol. Cell. Biol. *18*, 2029–2037.

Sekiguchi, J.A., Whitlow, S., and Alt, F.W. (2001). Mol. Cell *8*, 1383–1390.

Shockett, P.E., and Schatz, D.G. (1999). Mol. Cell. Biol. *19*, 4159–4166.

Taccioli, G.E., Rathbun, G., Oltz, E., Stamato, T., Jeggo, P.A., and Alt, F.W. (1993). Science *260*, 207–210.

Weis-Garcia, F., Besmer, E., Sawchuk, D.J., Yu, W., Hu, Y., Cassard, S., Nussenzweig, M.C., and Cortes, P. (1997). Mol. Cell. Biol. *17*, 6379–6385.

Yarnell Schultz, H., Landree, M.A., Qiu, J.X., Kale, S.B., and Roth, D.B. (2001). Mol. Cell *7*, 65–75.