12. K. Barry, J. I. Stiles, D. F. Pietras, I. Melnick, F. Sherman, *Mol. Cell. Biol.* **7**, 632 (1987).
13. M. Nonet, C. Scafe, J. Sexton, R. Young, *ibid.*, p. 1602.
14. K. S. Zaret and F. Sherman, *Cell* **28**, 563 (1982).
15. G. Edwalds-Gilbert, J. Prescott, E. Falck-Pedersen, *Mol. Cell. Biol.* **13**, 3472 (1993).
16. L. Minvielle-Sebastia, B. Winsor, N. Bonneaud, F. Lacroute, *ibid.* **11**, 3075 (1991).
17. N. Amrani, M. Minet, M. Le Gouar, F. Lacroute, F. Wyers, *ibid.* **17**, 3674 (1997); L. Minvielle-Sebastia, P. J. Preker, T. Wiederkehr, Y. Strahm, W. Keller, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7897 (1997).
18. A. Sachs and J. A. Deardorff, *Cell* **70**, 961 (1992).
19. S. McCracken *et al.*, *Nature* **385**, 357 (1997); J.-C. Dantonel, K. G. K. Murthy, J. L. Manley, L. Tora, *Nature* **389**, 399 (1997).
20. The six contiguous M13 clones spanning the 3′ end of the *CYC1* gene were constructed from PCR products with pGCYC as a template. The 172-bp *GAL* probe, used to detect transcripts upstream of the *GAL* promoter, was generated with pGCYC as a PCR template. The tRNA probe was generated with genomic DNA as a PCR template. This 225-bp fragment spans the *SUP11* gene coding for Tyr-tRNA located on chromosome 6. The actin *ACT1* probe was generated as a 567-bp fragment encompassing the region 277 to 844 bp 3′ of the *ACT1* ORF start site. All fragments were cloned into M13mp19 (RF) restricted with Hinc II. The M13 control probe has no insert.
21. The wild-type Rna15p and temperature-sensitive Rna15-1p proteins were synthesized in vitro in rabbit reticulocyte lysates with [^{35}S]methionine. After incubation for 10 min at 45°C, they were assayed for their ability to bind to agarose-bound poly(U) for 15 min, as described [M. S. Swanson and G. Dreyfuss, *Mol. Cell. Biol.* **8**, 2237 (1988)]. After extensive washes, the bound material was released by boiling with SDS–polyacrylamide gel electrophoresis (PAGE) loading buffer and loaded on 10% denaturing protein gels. After electrophoresis, the gels were fixed, dried, and subjected to autoradiography.
22. We thank S. Barabino, S. Butler, F. Lacroute, A. Sachs, F. Sherman, and R. Young for various yeast strains, P. J. Preker for the *pap1* mutants, and J. Mellor for plasmids. We thank D. Pritlove and members of the N.J.P. lab for critical reading of the manuscript and for their help and advice. Supported by a Wellcome Programme Grant (032773) to N.J.P. and grants by the Kantons of Basel, the Swiss National Science Foundation, and the European Union (through the Bundesamt für Bildung und Wissenschaft, Bern) to W.K.

1 October 1997; accepted 20 February 1998

Rejoining of DNA by the RAG1 and RAG2 Proteins

Meni Melek, Martin Gellert,* Dik C. van Gent

Assembly of immunoglobulin and T cell receptor genes from separate gene segments [V(D)J recombination] begins with DNA double-strand breakage by the RAG1 and RAG2 proteins, acting at a pair of recombination signal sequences (RSSs). Here, the RAG proteins are shown to reverse the cleavage reaction by joining an RSS to a broken coding sequence end. These "hybrid joints" have also been found in lymphoid cells, even when the normal pathway of DNA double-strand break repair is inactive, and can now be explained by this activity of the RAG proteins.

An essential step in the development of lymphoid cells is the recombinational joining of gene segments to form the functional immunoglobulin and T cell receptor genes (*1*, *2*). This V(D)J recombination process is initiated by the RAG1 and RAG2 proteins (*3*, *4*), which act together to cleave DNA between the RSSs and the adjoining coding segments (*5*). The resulting double-strand breaks (DSBs) have blunt-cut RSS ends (signal ends) and coding ends that are covalently sealed into DNA hairpins. Cleavage is stimulated by the small, DNA-bending high-mobility group proteins HMG1 or HMG2 (*6*) but does not require a high-energy cofactor such as adenosine triphosphate.

Subsequently, pairs of RSS ends (signal ends) and coding ends are coupled to form signal joints and coding joints (Fig. 1). These joining reactions require the presence of several ubiquitous DSB repair factors, such as XRCC4 protein and the DNA-dependent protein kinase, composed of the catalytic subunit (DNA-PK$_{CS}$) and the heterodimeric Ku protein (*7*). Signal joints usually have a precise head-to-head linkage of the RSSs, whereas coding joints frequently have lost several nucleotides or have acquired additional nontemplated (N) nucleotides inserted by terminal deoxynucleotidyl transferase (*2*). Added nucleotides complementary to the end of the coding sequence (P nucleotides) are also observed and are readily explained as resulting from asymmetric opening of the hairpin structure formed in the first step (*1*).

However, there is an alternative outcome of V(D)J recombination: A signal end can become joined to the coding end that originally flanked the other RSS, forming a "hybrid joint" (Fig. 1). Hybrid joints can account for as many as one-fifth of the total recombinants in plasmid substrates (*8*) and have been detected at lower frequencies in the antigen receptor loci (*9*). Rejoining of the original pair of signal and coding ends, to form an "open-and-shut joint," has also been observed (*8*, *10*).

Surprisingly high numbers of hybrid joints have been found in cells that are defective in DSB joining, because of mutations in the *Ku86* or *DNA-PK*$_{CS}$ genes (*11*, *12*). These results imply either that the RAG proteins themselves are able to catalyze hybrid joint formation, as has been proposed (*11*, *12*), or that a different DSB repair pathway is involved in this joining reaction.

Using a sensitive polymerase chain reaction (PCR) assay, we showed that the RAG proteins generated hybrid joints in the absence of other joining factors. After incubation of a plasmid substrate (Fig. 2A) with the RAG1, RAG2, and HMG1 proteins in the presence of Mg^{2+}, we amplified the products by PCR, using various sets of primers to identify signal joints, coding joints, and hybrid joints.

Although we did not detect any signal or coding joints (*13*), PCR products consistent with hybrid joint formation were readily detected (Fig. 2). We assayed for a product containing the 12-RSS joined to the region that originally flanked the 23-RSS and amplified the predicted product of about 241 base pairs (bp). The DNA contained a Bam HI site but not a Sal I or Sca I site (Fig. 2B), as expected for this hybrid joint. Similarly, the reciprocal product (a junction of the 23-RSS to the coding flank originally attached to the 12-RSS) was observed with the appropriate PCR primers. This 205-bp product contained sites for Sal I and Sca I but not for Bam HI (Fig. 2C), in keeping with its expected structure. We also obtained hybrid joint products using several other DNA substrates containing different sequences flanking the RSSs or with the 12-RSS in the opposite orientation (*13*), indicating that the reaction does not depend on one particular configuration around the RSSs. This joining reaction, like RAG-mediated cleavage (*5*), did not require an exogenous energy source, suggesting that the energy of a phosphodiester bond is used to generate the new bond at the RSS border. Quantitative PCR (by dilution of rearranged substrate into the reaction mixture) revealed that the RAG proteins converted about 0.5% of the input DNA to hybrid joints.

The structure of the hybrid joints linking the 23-RSS to the 12–coding flank was analyzed in more detail by cloning and sequencing (see legend to Fig. 2). Most joints resulted from exact joining of the 23-RSS to the region originally flanking the 12-RSS, without addition or loss of nucleotides. How-

M. Melek and M. Gellert, Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892–0540, USA.
D. C. van Gent, Department of Cell Biology and Genetics, Erasmus University Rotterdam, Post Office Box 1738, 3000 DR Rotterdam, Netherlands.

*To whom correspondence should be addressed. E-mail: gellert@helix.nih.gov

ever, about one-third of the joints had acquired one or two additional nucleotides, which can be identified as P nucleotides, indicating that a hairpin in the coding flank of the 12-RSS was an intermediate in this reaction. Because the attacking group would have to be the 3′-OH from the 23-RSS, a DSB at this signal must also have been present.

If the hybrid junctions detected in the PCR reaction arise by reversal of hairpin formation, they would only be joined on one strand (connecting the 3′-OH of the broken RSS to the 5′-P of the coding strand). This prediction was confirmed by primer extension of one strand followed by a second round of primer extension of the other strand (Fig. 3A). A hybrid joint was only detected when a primer complementary to the expected continuous strand was used in the first primer extension reaction (Fig. 3B, compare lanes 3 and 4).

The experiments described above show that the RAG proteins can generate hybrid joints in a reaction that appears to be a chemical reversal of hairpin formation. A more precise reversal of cleavage would rejoin the flanking sequence to the same RSS, resulting in the formation of an open-and-shut joint. Such a reaction can only be easily observed if the sequence at the junction is changed in the process, that is, if nucleotides are added or deleted. We designed a substrate with a Hae III restriction site at the 23-RSS border that can be used for selection against the unmodified substrate, whereas addition of one or more P nucleotides would result in a new Hha I restriction site (Fig. 4A). Open-and-shut joints were detected in a PCR assay after digestion with Hae III to remove unaltered substrate (*14*). About 50% of all products had acquired a Hha I site, indicative of P nucleotide insertion (Fig. 4B, lane 2). Thus, the RAG proteins are also capable of forming open-and-shut joints.

Formation of open-and-shut joints in vivo does not require the presence of two canonical RSSs (*10*). This was also true in the reaction with purified proteins. Reaction of a substrate lacking the 12-RSS produced a similar amount of open-and-shut joints at the 23-RSS (Fig. 4B, lane 4). Additionally, removal of most of the plasmid sequence did not decrease the yield of open-and-shut joints (Fig. 4B, lane 6).

At least two lines of evidence have suggested that the RAG proteins have functions after the initial cleavage reaction. First, in a staged in vitro cleavage and joining reaction, the RAG proteins must be present during the joining step for coding joint formation (*15*). Second, a "postcleavage complex" has

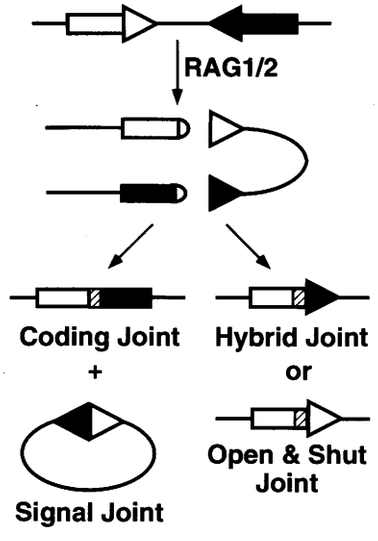

Fig. 1. Products of V(D)J recombination. Cleavage by the RAG proteins at 12- and 23-spacer RSS (white and black triangles, respectively) results in the formation of hairpins at the ends of coding sequence (boxes) and blunt ends at the signal sequence. The normal products of recombination are coding joints and signal joints (left). Sometimes junctions are formed with the coding segment of one element joined to the signal sequence of another (hybrid joint, right). Alternatively, the same pair of coding and signal sequences can be rejoined after cleavage (open-and-shut joint). Nucleotide addition or deletion can be observed in coding joints, hybrid joints, or open-and-shut joints (hatched area) but not usually in signal joints.

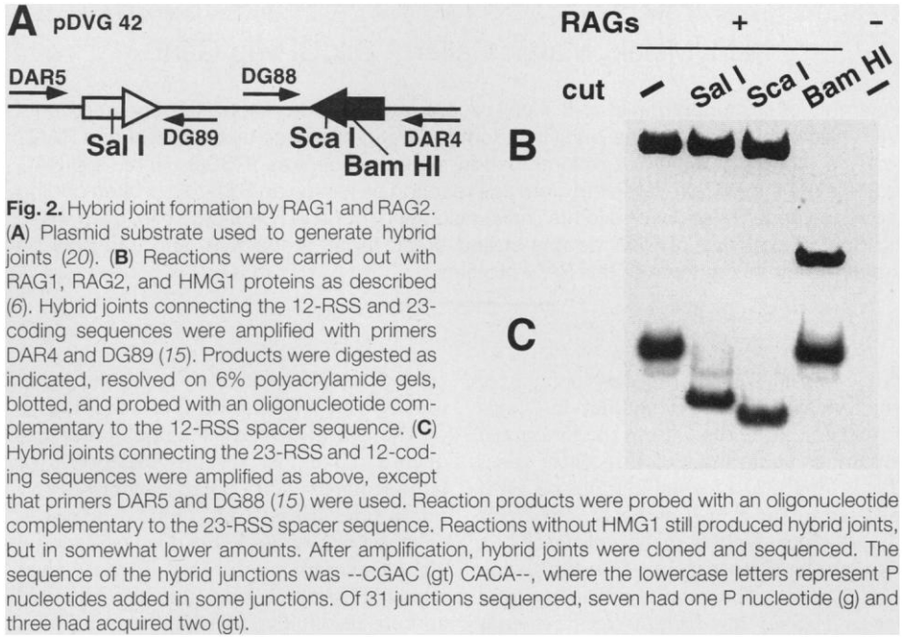

Fig. 2. Hybrid joint formation by RAG1 and RAG2. (**A**) Plasmid substrate used to generate hybrid joints (*20*). (**B**) Reactions were carried out with RAG1, RAG2, and HMG1 proteins as described (*6*). Hybrid joints connecting the 12-RSS and 23-coding sequences were amplified with primers DAR4 and DG89 (*15*). Products were digested as indicated, resolved on 6% polyacrylamide gels, blotted, and probed with an oligonucleotide complementary to the 12-RSS spacer sequence. (**C**) Hybrid joints connecting the 23-RSS and 12-coding sequences were amplified as above, except that primers DAR5 and DG88 (*15*) were used. Reaction products were probed with an oligonucleotide complementary to the 23-RSS spacer sequence. Reactions without HMG1 still produced hybrid joints, but in somewhat lower amounts. After amplification, hybrid joints were cloned and sequenced. The sequence of the hybrid junctions was --CGAC (gt) CACA--, where the lowercase letters represent P nucleotides added in some junctions. Of 31 junctions sequenced, seven had one P nucleotide (g) and three had acquired two (gt).

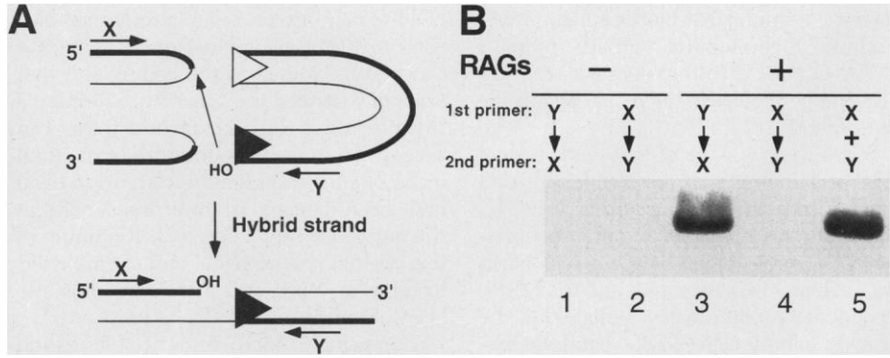

Fig. 3. Formation of one continuous strand in hybrid joints. (**A**) Schematic representation of primer extension assay on hybrid joints. The heavy and light lines distinguish the two strands of the substrate DNA. (**B**) Hybrid strand connecting the 23-RSS and 12-coding sequences. Plasmid pMS326 (*21*) was reacted with RAG1 and RAG2 proteins as described in Fig. 2. Primer extension (15 cycles) of the RAG-reacted DNA with a complement to one strand of the hybrid joint was first performed with either primer X [DAR5 of (*15*)] or Y (*22*) or both, as indicated in (B). The second primer (when used) was then added and extended (10 cycles). Products were resolved on a 6% TBE acrylamide gel and transferred to a nylon membrane. Blots were probed with an oligonucleotide spanning the predicted hybrid joint strand.

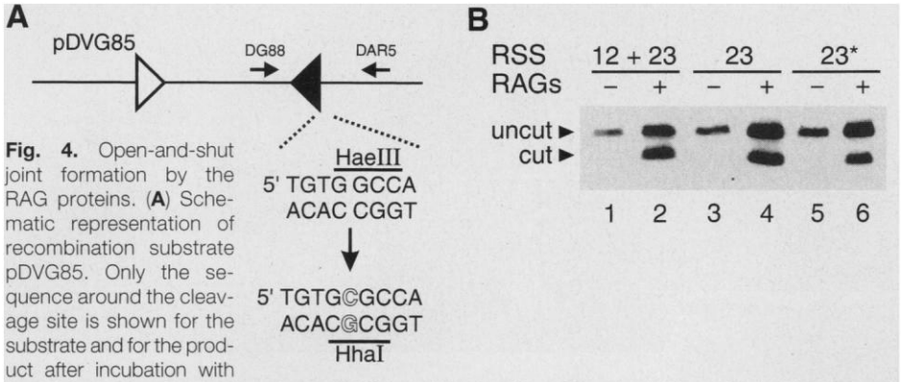

Fig. 4. Open-and-shut joint formation by the RAG proteins. (**A**) Schematic representation of recombination substrate pDVG85. Only the sequence around the cleavage site is shown for the substrate and for the product after incubation with the RAG proteins, Hae III digestion, and PCR amplification. Upon P nucleotide insertion (outlined letters), a Hha I restriction site is created at the RSS border (independent of the number of P nucleotides). (**B**) Southern (DNA) blot analysis of Hha I–digested PCR products from reactions containing pDVG85 (lanes 1 and 2), pDVG86 (which has only the 23-RSS) (lanes 3 and 4), or pDVG85 digested with Pvu II and Bst XI to separate the 23-RSS from the bulk of the DNA (lanes 5 and 6, marked 23*). RAG1 and RAG2 proteins were present in lanes 2, 4, and 6. Hha I–sensitive DNA appears only upon reaction with the RAG proteins. [Hha I–resistant DNA in these lanes may be due to junctions where the coding flank has been shortened, by attack on the other side of the hairpin, rather than extended. A small proportion of PCR products insensitive to Hha I in the negative control (lane 1) is probably due to incomplete digestion with Hae III or misincorporation in the PCR reaction (or both).]

been identified, in which the RAG proteins probably hold together the cleaved DNA signal ends (*16*). Up to this point, it has been unclear whether the RAG proteins may have a catalytic function in a joining reaction, in addition to their architectural role in a postcleavage complex. Our experiments show that they can at least play a catalytic role in the formation of nonstandard V(D)J joints. These reactions are similar to the "disintegration" reaction carried out by retroviral integrases: In both cases, the same proteins can carry out a transesterification reaction in either the forward or backward direction (*17*). In addition to the analogies observed previously in the nicking and transesterification reactions (*18*), this provides further evidence for an evolutionary link between V(D)J recombination and transposition or retroviral integration.

Although some hybrid joints found in vivo have undergone base loss or addition (*8*) and thus cannot be fully explained by a reversal of the RAG-mediated cleavage reaction, the hybrid and open-and-shut joints described here are very similar to those found in RAG-transfected fibroblasts: The junctions are either precise or contain one or a few P nucleotides (*12*). Even mutations in the *Ku86* or *DNA-PK*$_{CS}$ genes did not alter the structures (or frequency) of hybrid joints in these cells, which argues that these joints are formed independently of DNA-PK. The results described here suggest that they are made by the RAG proteins. Hybrid joints with precise junctions or a few P nucleotides have also been detected in lymphoid cells from *Ku86*$^{-/-}$ mice (*11*).

Within cells, most RAG-mediated cleavage occurs at pairs of RSSs, but the few percent of cleavages at a single RSS (*19*) would leave an unrepaired DSB in the chromosome. The open-and-shut joints made by the RAG proteins could initiate reclosure of isolated DSBs and help prevent chromosomal rearrangements, thereby contributing to genome stability. This reaction may explain why disruption of the *Ku86* gene does not result in increased chromosomal instability and carcinogenesis in B and T cell precursors.

REFERENCES AND NOTES

1. M. Gellert, *Adv. Immunol.* **64**, 39 (1997).
2. S. M. Lewis, *ibid.* **56**, 27 (1994).
3. D. G. Schatz, M. A. Oettinger, D. Baltimore, *Cell* **59**, 1035 (1989).
4. M. A. Oettinger, D. G. Schatz, C. Gorka, D. Baltimore, *Science* **248**, 1517 (1990).
5. J. F. McBlane *et al.*, *Cell* **83**, 387 (1995).
6. D. C. van Gent, K. Hiom, T. T. Paull, M. Gellert, *EMBO J.* **16**, 2265 (1997).
7. D. T. Weaver, *Adv. Immunol.* **58**, 29 (1995).
8. S. M. Lewis, J. E. Hesse, K. Mizuuchi, M. Gellert, *Cell* **55**, 1099 (1988).
9. E. Morzycka-Wroblewska, F. E. H. Lee, S. V. Desiderio, *Science* **242**, 261 (1988).
10. S. M. Lewis and J. E. Hesse, *EMBO J.* **10**, 3631 (1991).
11. M. A. Bogue, C. Wang, C. Zhu, D. B. Roth, *Immunity* **7**, 37 (1997).
12. J. O. Han, S. B. Steen, D. B. Roth, *Mol. Cell. Biol.* **17**, 2226 (1997).
13. M. Melek, M. Gellert, D. C. van Gent, data not shown.
14. Open-and-shut joints were detected in DNA derived from cleavage reactions with RAG1, RAG2, and HMG1 proteins, in the presence of Mg^{2+} (*6*). DNA (5 ng) was digested with Hae III, followed by 10 cycles of PCR amplification with primers DG88 and DAR5 (*15*). One-tenth of the amplified product was digested with Hae III (to remove any Hae III–sensitive material that survived the initial digestion), followed by 20 cycles of PCR with the same primers. Subsequently, PCR products were digested with Hha I, separated on a 6% polyacrylamide gel in tris-borate-EDTA (TBE) buffer, and blotted onto Hybond N+. Blots were hybridized with oligonucleotide NEB1233, and products were visualized by autoradiography.
15. D. A. Ramsden, T. T. Paull, M. Gellert, *Nature* **388**, 488 (1997).
16. A. Agrawal and D. G. Schatz, *Cell* **89**, 43 (1997).
17. S. A. Chow, K. A. Vincent, V. Ellison, P. O. Brown, *Science* **255**, 723 (1992).
18. D. C. van Gent, K. Mizuuchi, M. Gellert, *ibid.* **271**, 1592 (1996).
19. S. B. Steen, L. Gomelsky, S. L. Speidel, D. B. Roth, *EMBO J.* **16**, 2656 (1997).
20. D. C. van Gent, D. A. Ramsden, M. Gellert, *Cell* **85**, 107 (1996).
21. M. Sadofsky, J. E. Hesse, D. C. van Gent, M. Gellert, *Genes Dev.* **9**, 2193 (1995).
22. The sequence of primer Y is 5′-GAGAACTTTG-GAATCCAGTCCCTCTTCC-3′.

23. We thank our departmental colleagues for valuable discussions. The research of D.C.v.G. was made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

25 November 1997; accepted 12 February 1998