Cell, Vol. S116, S103–S106, January 23, 2004 Copyright ©2004 by Cell Press

Uncovering the V(D)J Recombinase

Commentary

David G. Schatz[1,*] and David Baltimore[2]
[1]Howard Hughes Medical Institute
Section of Immunobiology
Yale Medical School
300 Cedar Street
Box 208011
New Haven, Connecticut 06520
[2]President's Office
California Institute of Technology
Pasadena, California 91125

As any good high school student can tell you, normal developmental processes give rise to individuals in which each cell contains an identical copy of the genome present in the original, fertilized egg. This comfortable image of a static and inviolate genetic blueprint was shattered in the late 1970s, however, when Susumu Tonegawa and colleagues showed that the genes that encode immunoglobulin heavy and light chains have a different structure in embryonic cells from that found in tumors derived from B cells (see accompanying Commentary by Tonegawa, 2004). Over the next decade, it would emerge that developing B lymphocytes assemble immunoglobulin genes from widely scattered gene segments, using a somatic DNA rearrangement process known as V(D)J recombination (Figure 1), and that developing T cells play the same trick to assemble functional T cell receptor genes.

The discovery of a new process, such as V(D)J recombination, challenges us to transform the phenomenology of the molecular genetics into mechanistic, biochemical understanding. We could easily guess that V(D)J recombination was catalyzed by a "recombinase" enzymatic machinery, but what were the components of this machine, how did it work, and what were its evolutionary origins?

There was no lack of interest in finding the V(D)J recombinase. Throughout the 1980s, the primary approach taken was that of traditional biochemistry. Investigators sought (and found, and published) a variety of intriguing, but ultimately disappointing, nuclease and DNA binding activities. Others attempted to reconstitute the entire reaction in cell-free, crude extracts, while yet others took a molecular approach, using subtractive hybridization to search for genes expressed specifically in cells with an active V(D)J recombination system.

It was in early 1985, in this turbulent atmosphere of lofty and frustrated aspirations, that one of us (D.G.S.) began his graduate studies in the laboratory of the other (D.B.) in the Whitehead Institute at MIT. The ensuing effort to find the V(D)J recombinase was imbued with extraordinary good luck and the naivete and optimism of a graduate student who had yet to experience his first major setback.

The first stroke of good fortune was that a previous graduate student in the lab, Susanna Lewis, had just months before described a novel method of selecting for cells with the capability to perform V(D)J recombination (Lewis et al., 1984). Based on Lewis's concept, a retroviral substrate was developed in which V(D)J recombination would trigger inversion of a drug resistance gene, allowing cells with recombination activity to be selected for by their expression of drug resistance. As might be expected, the substrate underwent V(D)J recombination in a subset of lymphocyte cell lines, but not in nonlymphoid 3T3 fibroblasts (Schatz and Baltimore, 1988). When these results were presented at our weekly group meeting, Yoav Citri, a postdoctoral fellow, suggested that it would be interesting to explore methods for activating V(D)J recombination in the 3T3 fibroblast cells harboring the recombination substrate. This pivotal suggestion changed the course of the research and that of several careers.

Triggering V(D)J recombination in fibroblasts would almost certainly require these cells to express one or more lymphocyte-specific genes—the crucial phrase being "one *or more*." If just one gene were required, standard methodologies for complementation by gene transfer might succeed if the gene were reasonably sized; but if more than one gene were needed, such an approach would be little more than wishful thinking. And, given the complexities of the V(D)J recombination reaction, it did indeed seem fanciful to believe that a single novel gene could induce fibroblasts to perform this reaction.

Nevertheless, a brave graduate student, over the conservative advice of his paternalistic mentor, undertook the gene transfer/complementation approach. The standard tool one would use now would be a cDNA expression library prepared from immature lymphocytes. The cDNAs from such a library would be introduced into 3T3 fibroblasts containing the recombination substrate, and the cells would then be selected for expression of the drug resistance gene. As became obvious after the fact, such an approach would never have succeeded because, indeed, there were two proteins involved, RAG1 and RAG2, and the chance of introducing two different cDNAs into the same cell is vanishingly small. Good fortune intervened, however, in that the methodology of genomic transfection was in extensive use in the neighboring laboratory of Robert Weinberg. In this method, genomic DNA (in the form of large [~100 kb] fragments) is transfected into a recipient cell line, and the transfectants are then analyzed for a new activity or property, in this case, the ability to perform V(D)J recombination. Surprisingly, when genomic DNA was transfected into fibroblasts containing the recombination substrate, several clones were isolated that stably manifested the capacity to perform V(D)J recombination (Schatz and Baltimore, 1988).

This result seemed to indicate that one gene (or, more accurately, one genomic locus) could trigger V(D)J recombination in nonlymphoid cells. The challenge then became to find that gene (locus). This turned out to be nontrivial and might not have happened at all were it not for Marjorie Oettinger, a new graduate student who

*Correspondence: david.schatz@yale.edu

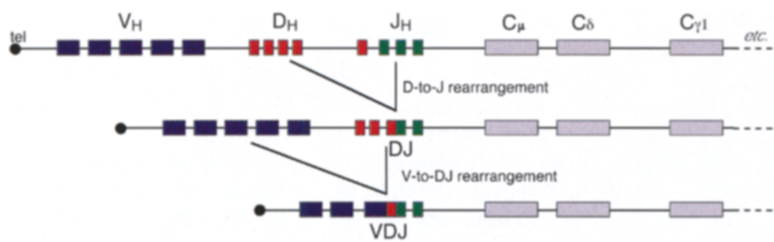

Figure 1. V(D)J Recombination and Assembly of the Ig Heavy Chain Locus

The Ig heavy chain locus is assembled in two steps, beginning with joining of a D and a J gene segment, followed by joining of a V gene segment to DJ. The chromosomal regions between D and J and between V and DJ are deleted from the chromosome, although in certain antigen receptor loci, some recombination events occur by inversion. After V(D)J recombination, transcription initiates at a promoter upstream of V and the VDJ exon is spliced to the downstream $C\mu$ constant region exons (which are depicted as a single rectangle for simplicity), allowing expression of the Ig μ heavy chain polypeptide. Splicing can also occur to $C\delta$ exons to direct expression of Ig δ, or after class switch recombination (which deletes $C\mu$ and $C\delta$), $C\gamma$, $C\epsilon$, and $C\alpha$ exons can be utilized.

threw her considerable energy, ideas, and wry wit into the effort. Although the butt of many jokes, the Tweedledee/Tweedledum team of Oettinger and Schatz was able to pursue several cloning strategies simultaneous, and perhaps more important, to maintain a sense of optimism in the face of results, and an atmosphere, that were often less than encouraging. This atmosphere is well illustrated by a comment made to D.G.S. in the summer of 1988 by a postdoctoral fellow in a nearby lab: "Oh, so you're the person who published that artifact in *Cell*" (referring to the activation of V(D)J recombination in fibroblasts; Schatz and Baltimore, 1988). To put it bluntly, many people were of the opinion that we were chasing a nonexistent phantom gene.

In the face of this skepticism, the temptation was to resort to humor, and there was no shortage of this in room 317 of the Whitehead Institute. A memorable example took place one evening when the "early" member of the tag-team (Schatz, arrival at lab 7 am, departure 7 pm) set up two capillary transfer Southern blots (his signature technique) and left for the day, turning things over to the "late" member of the team (Oettinger, arrival 11 am, departure 11 pm on a good night). By early the next morning, the two Southern transfers had apparently given birth to a third; closer inspection revealed, however, that the newcomer contained a peanut butter and jelly sandwich, instead of an agarose gel, nestled between sponges, layers of Whatman paper, and paper towels. Not only were the results of this blot of dubious value, but there was the serious question of what Schatz would eat for lunch that day.

The key challenge faced by Tweedledee and Tweedledum was to find the relevant gene (locus) amongst the morass of the genome. Of the numerous approaches taken, success came from the use of an oligonucleotide "tag" that was ligated to bulk genomic DNA and then used to trace the fate of transfected DNA through successive rounds of transfection and selection. Southern blotting with the oligonucleotide as probe revealed that the tag had become closely linked to the genetic locus of interest and could then be used as the starting point for a genomic "walk" to isolate the gene.

This turned out to be far more difficult than expected. The first challenge encountered was that fragments containing the oligonucleotide tag could not be found in multiple genomic libraries, and hence appeared to be unclonable. Only when a library was plated on a highly recombination-defective strain of bacteria was a single clone obtained. The genomic walk then began, but it moved forward with agonizing slowness because each step landed in a new, noncontiguous portion of the genome. Eventually, five different genomic libraries were required to allow these breakpoints to be spanned. The walk ended abruptly within a transcription unit whose mRNA had the satisfying property of being expressed in the appropriate lymphocyte cell lines and recombinationally active transfectants. The gene was cloned and named *RAG1*, the Recombination Activating Gene-1 (Schatz et al., 1989).

This might have been the end of the story but it was not, because the isolated gene was disturbingly ineffective. We expected expression of the *RAG1* cDNA in fibroblasts to result in very efficient activation of V(D)J recombination, but instead, the cDNA expression vector was no more efficient in activating recombination than was bulk genomic DNA. This observation led to several months of intense puzzlement, frantic experimentation, and thence to a number of wacky explanations for this unexpected result. One of the far-fetched ideas was that a second gene, closely linked to *RAG1*, was required for efficient activation of recombination. Wacky or not, a search for nearby transcription units rapidly led to the identification of *RAG2*, a gene that resided only a few kilobases away from *RAG1* (Figure 2A; Oettinger et al., 1990). Further experiments demonstrated that *RAG1* and *RAG2* were almost invariably expressed together and that their coexpression resulted in the expected potent activation of V(D)J recombination (Oettinger et al., 1990).

At this point in the story, both students left MIT with PhD degrees, jettisoned their plans to finish the medical degrees they had begun at Harvard, and each started his or her own laboratory.

The discovery of *RAG1* and *RAG2* provided the impetus for significant advances in many directions: understanding of the mechanism of V(D)J recombination, insights into lymphocyte development, a deepened understanding of mechanisms of DNA repair, and elucidation of the origin of certain human immunodeficiencies. The first order of business was to determine the basic function of the RAG1 and RAG2 proteins: did they participate directly in V(D)J recombination or did they instead play an indirect role, activating other genes/factors that were the direct players? Although it took five years, elegant biochemistry from the Gellert and Oettinger laboratories eventually demonstrated that RAG1 and RAG2 constitute the sequence-specific DNA endonuclease that initiates V(D)J recombination (McBlane et al., 1995; van Gent et al., 1995). Studies in a number of labs subsequently provided insight into how RAG1/

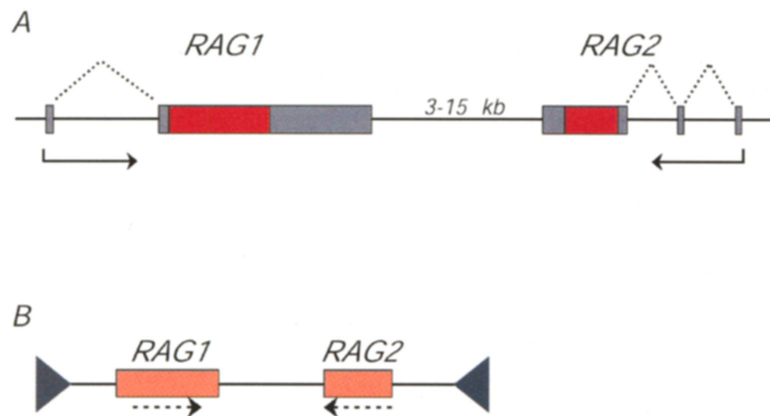

Figure 2. The *RAG* Genomic Locus and Hypothetical *RAG* Transposon

(A) The *RAG* genomic locus is remarkably compact. *RAG1* and *RAG2* are nearest neighbors and convergently transcribed (arrows) in all vertebrate species examined to date. The genes are separated by 3–15 kb depending on the species, and the open reading frames (red rectangles) of both RAG1 and RAG2 are each contained in a single large exon in almost all species (including human and mouse). Dashed lines indicate splicing of 5′ untranslated exons to the main exon.

(B) The hypothetical *RAG* transposon. It is plausible that 400–500 million years ago, at the time that jawed vertebrates first evolved, a transposable element existed containing the *RAG1* and *RAG2* open reading frames (rectangles) flanked by recombination signal sequences (RSSs; triangles). Dashed arrows indicate the presumed direction of transcription of the genes. Derived from Thompson (1995).

RAG2 bind and cleave DNA. Curiously, almost all discrete domains and activities identified to date reside within RAG1, while RAG2 remains an essential but largely enigmatic participant (Fugmann et al., 2000; van Gent et al., 1996).

Within two years of the identification of these genes, mice deficient in either *RAG1* or *RAG2* were generated and found to exhibit a complete block in V(D)J recombination (Mombaerts et al., 1992; Shinkai et al., 1992; Spanopoulou et al., 1994). Because lymphocyte development is dependent on successful assembly of antigen receptor genes, these mice exhibit a complete absence of mature B and T cells, making them ideal "empty" hosts into which lymphocyte populations or hematopoietic precursors can be transferred for in vivo analyses.

In retrospect, it is clear that cloning of *RAG1* and *RAG2* by transfecting genomic DNA was successful only because the two genes lie so close to one another (Figure 2A) and hence behaved as a single genetic locus in the assay. Was this merely a wildly unlikely, and lucky, occurrence, or is there some deeper significance to the fact that *RAG1* and *RAG2* are nearest neighbors and have a very compact organization? A satisfying answer to this question came with the discovery that RAG1 and RAG2 together constitute a transposase and catalyze a reaction with extensive mechanistic similarities to the well-studied bacterial transposases (Agrawal et al., 1998; Hiom et al., 1998). Based on this and other findings, it is reasonable to think that, early in evolution, *RAG1* and *RAG2* existed as components of a transposable element (the "RAG transposon;" Figure 2B) that contained all of the information necessary to jump from genome to genome and hence to propagate itself (Agrawal et al., 1998; Thompson, 1995). This idea, if correct, has profound implications for the evolution of adaptive immunity. It has long been speculated that antigen receptor genes came to be split into component gene segments when a transposon landed in an ancestral receptor gene and split it into two parts. The introduction of the RAG transposon into the germline of an early vertebrate ancestor could explain why all jawed vertebrates contain the *RAG* genes (immediately adjacent to one another), perform V(D)J recombination, and make use of an adaptive immune system made up of B and T lymphocytes, while no lower species examined thus far contains any trace of these entities. It is entirely possible that the existence of jawed vertebrates, which are highly vulnerable to infection, was dependent on the chance occurrence of an interspecies gene transfer providing the necessary wherewithal to counter infections through the complicated but effective process of an adaptive immune system based on gene fragment recombination. We know that all lower species share with vertebrates many other mechanisms of countering microbial invasion (the mechanisms of innate immunity). But we also know that RAG-deficient mice (as well as humans, see below) are highly susceptible to bacterial and viral pathogens, indicating the critical importance of the adaptive immune system to vertebrates.

Finally, recent studies have provided a clear link between the *RAG* genes and human disease. Two distinct forms of severe combined immunodeficiency are caused by germline mutations that result in a partial or total inactivation of either *RAG1* or *RAG2* (Schwarz et al., 1996; Villa et al., 1998). In addition, there is growing evidence that errors made during the process of V(D)J recombination are the cause of chromosomal translocations found in certain lymphoid tumors (Mills et al., 2003).

Beginning then with Tonegawa's discovery that the vertebrate genome is not inviolate but is instead molded into an astonishing array of different configurations in each of us, we have moved to an appreciation of the enzymes that generate these configurations, and thence to a growing, but still incomplete, knowledge of the mechanisms by which they act. Major challenges for the future include determining the three-dimensional structure of the RAG proteins, both individually and in complex with DNA, understanding the dynamic changes that occur in the protein-protein and protein-DNA contacts during the course of the V(D)J recombination reaction, and unraveling the secrets of the nucleus and chromatin structure that dictate the exquisite developmental regulation of antigen receptor gene assembly.

References

Agrawal, A., Eastman, Q.M., and Schatz, D.G. (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature *394*, 744–751.

Fugmann, S.D., Lee, A.I., Shockett, P.E., Villey, I.J., and Schatz, D.G. (2000). The RAG proteins and V(D)J recombination: complexes, ends, and transposition. Annu. Rev. Immunol. *18*, 495–527.

Hiom, K., Melek, M., and Gellert, M. (1998). DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell *94*, 463–470.

Lewis, S., Gifford, A., and Baltimore, D. (1984). Joining of Vκ to Jκ gene segments in a retroviral vector introduced into lymphoid cells. Nature *308*, 425–428.

McBlane, J.F., van Gent, D.C., Ramsden, D.A., Romeo, C., Cuomo, C.A., Gellert, M., and Oettinger, M.A. (1995). Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell *83*, 387–395.

Mills, K.D., Ferguson, D.O., and Alt, F.W. (2003). The role of DNA breaks in genomic instability and tumorigenesis. Immunol. Rev. *194*, 77–95.

Mombaerts, P., Iacomini, J., Johnson, R.S., Herrup, K., Tonegawa, S., and Papaioannou, V.E. (1992). RAG-1-deficient mice have no mature B and T lymphocytes. Cell *68*, 869–877.

Oettinger, M.A., Schatz, D.G., Gorka, C., and Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science *248*, 1517–1523.

Schatz, D.G., and Baltimore, D. (1988). Stable expression of immunoglobulin gene V(D)J recombinase activity by gene transfer into 3T3 fibroblasts. Cell *53*, 107–115.

Schatz, D.G., Oettinger, M.A., and Baltimore, D. (1989). The V(D)J recombination activating gene (RAG-1). Cell *59*, 1035–1048.

Schwarz, K., Gauss, G.H., Ludwig, L., Pannicke, U., Li, Z., Lindner, D., Friedrich, W., Seger, R.A., Hansenhagge, T.E., Desiderio, S., et al. (1996). RAG mutations in human B cell-negative SCID. Science *274*, 97–99.

Shinkai, Y., Rathbun, G., Kong-Peng, L., Oltz, E.M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A.M., and Alt, F.W. (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell *68*, 855–867.

Spanopoulou, E., Roman, C.A.J., Corcoran, L.M., Schlissel, M.S., Silver, D.P., Nemazee, D., Nussenzweig, M.C., Shinton, S.A., Hardy, R.R., and Baltimore, D. (1994). Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. Genes Dev. *8*, 1030–1042.

Thompson, C.B. (1995). New insights into V(D)J recombination and its role in the evolution of the immune system. Immunity *3*, 531–539.

Tonegawa, S. (2004). That great time in Basel. Cell *S116*, this issue, S99–S101.

van Gent, D.C., McBlane, J.F., Ramsden, D.A., Sadofsky, M.J., Hesse, J.E., and Gellert, M. (1995). Initiation of V(D)J recombination in a cell-free system. Cell *81*, 925–934.

van Gent, D.C., Mizuuchi, K., and Gellert, M. (1996). Similarities between initiation of V(D)J recombination and retroviral integration. Science *271*, 1592–1594.

Villa, A., Santigata, S., Bozzi, F., Giliani, S., Frattini, A., Imberti, L., Gatta, L.B., Ochs, H.D., Schwarz, K., Notarangelo, L.D., et al. (1998). Partial V(D)J recombination activity leads to Omenn syndrome. Cell *93*, 885–896.

Cell, Vol. 15, 1–14, September, 1978, Copyright ©1978 by Cell Press

A Complete Immunoglobulin Gene Is Created by Somatic Recombination

Christine Brack, Minoru Hirama,
Rita Lenhard-Schuller and Susumu Tonegawa
Basel Institute for Immunology
Grenzacherstrasse 487
Postfach, 4005 Basel 5
Switzerland

Summary

Using a pCRI plasmid containing an enzymatically synthesized, full-length DNA transcript of immunoglobulin λ chain mRNA as the hybridization probe in the Southern gel blotting experiments (Southern, 1975), we identified three DNA fragments of 8.6, 4.8 and 3.5 kb in Eco RI-digested total DNA from BLAB/c mouse embryos. A fourth fragment of 7.4 kb was found in addition to these three fragments in similarly digested total DNA from a λ chain-secreting myeloma (HOPC 2020). We have cloned the four DNA fragments in an EK-2 phage vector, λ_{WES}, and characterized them with respect to size, type of λ gene sequences contained and position of these sequences in the fragments, using agarose gel electrophoresis, the gel blotting technique and electron microscopic R loop mapping. The embryonic DNA clones Ig 99λ, Ig 25λ and Ig 13λ contain one copy each of $V_{\lambda I}$, $C_{\lambda I}$ and $V_{\lambda II}$ sequences, respectively, while the myeloma DNA clone Ig 303λ contains one copy each of $V_{\lambda I}$ and $C_{\lambda I}$ sequences that are separated by a 1.2 kb nontranslated DNA segment. Ig 25_{λ} was also shown to contain a DNA segment of approximately 40 base pairs (bp) (J sequence) that lies 1.2 kb away from the $C_{\lambda I}$ sequence and is homologous to the V-C junction region of a λ_I mRNA. Heteroduplex analysis of the three λ_I DNA clones revealed that Ig 303λ DNA is composed of two parts, one of which is entirely homologous to one end of Ig 99λ, and the other to one end of Ig 25λ DNA. The sequence arrangement observed in the cloned DNA is the same as that in the corresponding cellular DNA. This was shown by identifying certain restriction enzyme cleavage sites on the cloned DNAs and demonstrating the presence of these sites in the total cellular DNA by the gel blotting technique. The site of the homology switch is at the boundary of the V sequence and the 1.2 kb nontranslated DNA segment, and corresponds to the position of the J sequence on the Ig 25λ DNA. We consider the above experimental results the most direct evidence for somatic rearrangement in immunoglobulin genes. We discuss the significance of these findings for the origin of genes in the evolution of higher organisms and in cell differentiation.

Introduction

Are DNA sequences in the cells of higher organisms rearranged during normal cell differentiation? The restriction enzyme mapping in mouse DNAs strongly suggested that this is the case in lymphocytes for the immunoglobulin genes. Very different patterns of hybridization were obtained when κ light chain mRNA from a κ chain-producing myeloma was hybridized with mouse embryo DNA or with homologous myeloma DNA, both of which had been digested with the restriction endonuclease Bam HI and fractionated by agarose gel electrophoresis (Hozumi and Tonegawa, 1976). Suggestive evidence for a similar DNA rearrangement involving λ chain genes was obtained when DNAs from embryo and from a λ chain-producing myeloma were compared (Tonegawa et al., 1976). Furthermore, a pattern change was detected only for those immunoglobulin genes that are active in a given myeloma cell (Tonegawa et al., 1977a, 1977b). These results indicated that in embryo cells, the DNA sequences coding for the amino terminal half (V region) and for the carboxy terminal half (C region) are separate, and that the two sequences are brought to proximity during the differentiation of B (bone marrow-derived) lymphocytes. An alternative, improbable interpretation was also considered: the pattern difference might result from mutations or base modifications in the enzyme cleavage sites (Hozumi and Tonegawa 1976).

Two subtypes of mouse λ chains are known, λ_I and λ_{II}, which are characterized by the specific C region sequences $C_{\lambda I}$ and $C_{\lambda II}$. Amino acid sequence studies have so far established seven different $V_{\lambda I}$ regions and one $V_{\lambda II}$ region (Weigert et al., 1970; Dugan et al., 1973). Statistical considerations suggest that the mouse is capable of synthesizing many more than seven different V regions of the λ_I subtype (Tonegawa, 1978). The two λ subtype chains each seem to be encoded in a pair of DNA segments, one for the V and the other for the C, that lie in separate sections of the embryo DNA. Restriction enzyme mapping of total cellular DNA combined with hybridization kinetics strongly suggested that there are no more than few copies (and there is probably only one copy) per haploid genome of the DNA segment encoding each of the four λ chain regions $V_{\lambda I}$, $C_{\lambda I}$, $V_{\lambda I}$, $V_{\lambda II}$ and $C_{\lambda II}$ (Tonegawa et al., 1976). Hence the multiple $V_{\lambda I}$ regions observed in myelomas must have been generated by a somatic process.

To obtain more direct evidence for gene rearrangement, we have isolated, by in vitro recombination with phage λ DNA, DNA fragments from both embryo and myeloma cells that carry part or

Cell, Vol. 59, 1035–1048, December 22, 1989, Copyright ©1989 by Cell Press

The V(D)J Recombination Activating Gene, RAG-1

David G. Schatz, Marjorie A. Oettinger, and David Baltimore
Whitehead Institute for Biomedical Research
Nine Cambridge Center
Cambridge, Massachusetts 02142
and Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Summary

The RAG-1 (recombination activating gene-1) genomic locus, which activates V(D)J recombination when introduced into NIH 3T3 fibroblasts, was isolated by serial genomic transfections of oligonucleotide-tagged DNA. A genomic walk spanning 55 kb yielded a RAG-1 genomic probe that detects a single 6.6–7.0 kb mRNA species in transfectants and pre-B and pre-T cells. RAG-1 genomic and cDNA clones were biologically active when introduced into NIH 3T3 cells. Nucleotide sequencing of human and mouse RAG-1 cDNA clones predicts 119 kd proteins of 1043 and 1040 amino acids, respectively, with 90% sequence identity. RAG-1 has been conserved between species that carry out V(D)J recombination, and its pattern of expression correlates exactly with the pattern of expression of V(D)J recombinase activity. RAG-1 may activate V(D)J recombination indirectly, or it may encode the V(D)J recombinase itself.

Introduction

The enormous number of genes required to encode the subunits of the immunoglobulin (Ig) and T cell receptor (TCR) molecules of B and T cells are generated combinatorially in a process known as V(D)J recombination, so called for the variable (V), diversity (D), and joining (J) gene segments used in the recombination process. V(D)J recombination is known to assemble seven different loci in developing lymphocytes: μ, κ, and λ in B cells, and α, β, γ, and δ in T cells (for reviews see Blackwell and Alt, 1988; Davis and Bjorkman, 1988; Raulet, 1989). The B and T cell lineages appear to share a common V(D)J recombinase (see below) whose expression is tightly regulated during development; recombination of endogenous loci or exogenously introduced substrates occurs only in cells representing the early stages of B and T cell development and does not occur in nonlymphoid cells (Lieber et al., 1987; Schatz and Baltimore, 1988). V(D)J recombination is also regulated at the level of substrate availability, since loci normally recombined in one cell type are never fully recombined in the other. In some instances, it has been possible to correlate the recombination of a locus with its transcription and increased sensitivity to DNAase I, suggesting that the "accessibility" of a locus to the recombinase may be modulated by its transcription (Yancopoulos and Alt, 1985; Blackwell et al., 1986; Yancopoulos et al., 1986; Schlissel and Baltimore, 1989).

The *cis*-acting DNA sequences that are necessary and sufficient for V(D)J recombination consist of a palindromic heptamer and AT-rich nonamer separated by a spacer of either 12 or 23 bp (Tonegawa, 1983; Hesse et al., 1989). Spacer length defines two classes of such recombination signal sequences (RSSs); efficient recombination occurs only when one RSS of each class is involved (the "12–23 joining rule"). RSSs lie directly adjacent to all recombinationally competent V, D, and J gene segments, and their sequences are conserved between the recombining loci and between all species known to carry out V(D)J recombination (Litman et al., 1985a, 1985b; Reynaud et al., 1985; Schwager et al., 1988). RSSs from different loci and species appear to be functionally interchangeable because they direct proper V(D)J recombination in a variety of host pre-B and pre-T cells (Yancopoulos et al., 1986; Bucchini et al., 1987; Goodhardt et al., 1987; Lieber et al., 1987). The sequence conservation and functional equivalence of RSSs together provide strong evidence for a single, evolutionarily conserved V(D)J recombinase.

Remarkably little is known, however, about the *trans*-acting factors that participate in V(D)J recombination; no factor has yet been demonstrated to be an essential component of the enzyme. V(D)J recombination could be expected to require several distinct enzymatic activities: sequence-specific DNA recognition, endonucleolytic cleavage between the RSS and gene segment, and ligation of the cleaved ends. In addition, because nucleotides are frequently lost and/or added at coding segment junctions, both exonucleolytic and DNA polymerase activities should be associated with the recombinational machinery. It is not known how many factors participate in V(D)J recombination, nor is it known how many are lymphoid specific. We will use the term V(D)J recombinase to refer only to the essential component(s) of the enzyme whose activity is restricted to recombinationally active B and T cells. Thus far, only the enzyme terminal deoxynucleotidyl transferase (TdT) has been implicated in V(D)J recombination, where it is likely to add non-germline-encoded nucleotides (N regions) to the coding junction (Landau et al., 1987). TdT is not, however, required for V(D)J recombination. Some of the other activities expected to be associated with the recombinase have been detected in nuclear extracts of lymphoid cells—heptamer binding (Aguilera et al., 1987), nonamer binding (Halligan and Desiderio, 1987), and endonucleolytic cleavage (Desiderio and Baltimore, 1984; Hope et al., 1986; Kataoka et al., 1984)—but the relationship of these factors to V(D)J recombination remains unclear. Finally, a factor critical in V(D)J recombination may be contributed by the gene that is disrupted in the severe combined immunodeficient (*scid*) mouse (Bosma et al., 1983), which is characterized primarily by a defect in recombination of Ig and TCR gene segments (Schuler et al., 1986).