Targeting Retroviral Integration

Frederic Bushman

Retroviral vectors are currently the most widely used method of gene transfer in gene therapy, in part because the retroviral integration system precisely links the therapeutic sequences to the host chromosome with neat covalent bonds. DNA integration, however, comes with a price. Retroviral integration systems show little target site specificity, and haphazard insertions into a patient's chromosomes can create problems. Integration near cellular proto-oncogenes may lead to ectopic gene activation and cancer, not a desirable outcome of gene therapy (1). Retroviral integration can also inactivate genes (2). Recent findings on site-specific DNA integration by retroviruses and retrotransposons suggest a strategy for retrovirus-based gene therapy that avoids such insertional mutagenesis: Tethering the integration machinery to a specific target sequence can control the location of DNA integration, potentially making it possible to direct integration of therapeutic sequences to innocuous chromosomal sites.

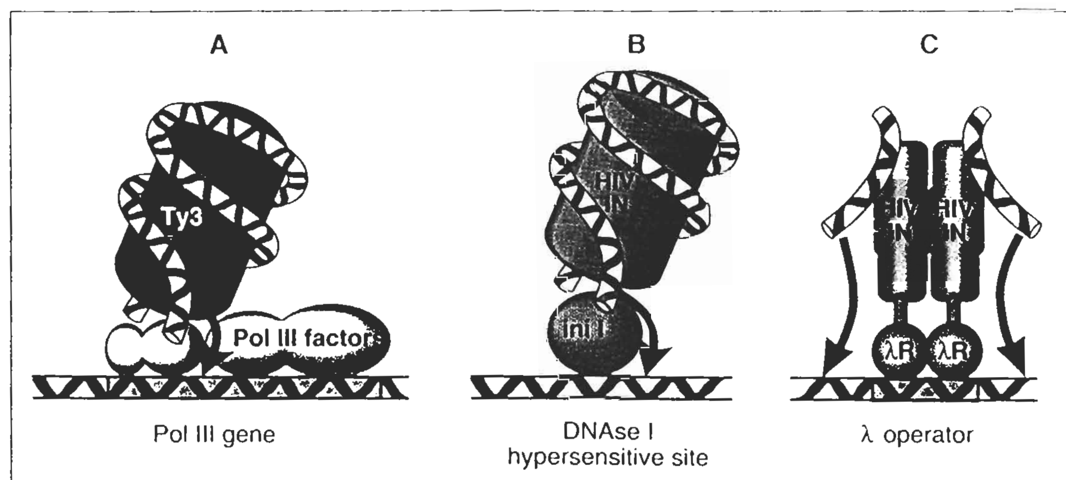
In a report on page 1488 of this week's *Science*, Kirchner and co-workers begin to define the molecular elements necessary for integration of the retrovirus-like Ty retrotransposons of yeast cells which, unlike the retroviruses, are highly selective in their choice of integration sites. The Ty genome is transcribed, reverse transcribed, and integrated all within a single cell by reverse transcriptase and integrase enzymes homologous to their retroviral counterparts. This mode of growth imposes unique constraints. Because yeast cells are often haploid, a retrotransposition event that disrupts a required cellular gene can mean suicide for both the Ty element and the host. Probably for this reason, the Ty elements display extreme biases in target site selection. Ty3 tends to integrate strictly into the five base pairs at the start site of transcription of polymerase (Pol) III–transcribed genes (3). Ty1 also usually integrates upstream of Pol III–transcription units, but over a region of several hundred base pairs (4). At first this may seem like an odd choice of integration sites, but the Ty insertions have very little effect on transcription, probably because the recognition sites for the sequence-specific Pol III factors TFIIIB and TFIIIC lie within the transcribed region.

How do the Ty elements select these innocuous target sites from the background of total genomic DNA? Previously, Sandmeyer and co-workers have shown in vivo that intact DNA sites for the transcription factors TFIIIB and TFIIIC were necessary for selective integration of Ty3 (3). In the new work, this group has established an in vitro system that recapitulates selective Ty3 integration (5). Ty3 viruslike particles from yeast cells induced to express Ty3 can direct selective integration in vitro when presented with integration targets containing assembled Pol III transcription units. Only those fractions containing TFIIIB and TFIIIC, but not RNA Pol III, are necessary for selective integration, indicating that bound factors and not transcription itself are required.

The current models for the mechanism of integration invoke tethering of the Ty3 integration machinery to the target DNA by protein-protein contacts with the Pol III transcription apparatus (see figure). Such a model fits neatly with ideas from the site-specific recombination field. Many site-specific recombination enzymes are composed of two domains, one that directs sequence-specific DNA binding at the recombination site and a second domain that catalyzes DNA breaking and joining. In Ty3, the two functions are apparently contributed by different proteins, a point that may soon be confirmed with the new in vitro system.

The mechanism of integration in yeast cells may be quite similar to that in retroviruses, and tethering may also be involved in normal integration of human immunodeficiency virus (HIV) DNA. In a recent report published in *Science*, Kalpana and co-workers used the popular yeast two hybrid system to identify a protein that binds to the HIV-1 integrase protein (6), the viral-encoded enzyme responsible for forming the initial covalent connection between viral DNA and target DNA. Binding of this protein, Ini1 (integrase interacting protein), to integrase was specific both in vivo and in vitro. The Ini1 protein also bound DNA, and in integration reactions in vitro containing purified integrase, Ini1 stimulated integration under certain conditions. This latter result must be interpreted with some caution, because nonspecific DNA-binding proteins can stimulate integration in vitro (7) and the new results do not yet confirm that Ini1 stimulates integration by tethering integrase to the target.

Nevertheless, the sequence of the Ini1 protein is consistent with an important role in vivo. Portions of Ini1 are up to 55% identical to yeast SNF5, a component of a multi-protein complex proposed to be involved in activation of gene expression (8). That the SNF5 complex can remodel chromatin so as to promote transcription factor



Models for selective integration by tethering. (A) An interaction between Pol III transcription factors and the Ty3 viruslike particle results in selective integration at the start site of transcription of Pol III genes. **(B)** An interaction between HIV-1 integrase and Ini1 may promote integration into DNA at DNAse I-hypersensitive sites. **(C)** A fusion of HIV-1 integrase to the DNA-binding domain of λ repressor promotes integration near λ operators in vitro. In the reactions in (C), the viral DNA end is modeled by a short oligonucleotide matching the sequence of one end of the unintegrated retroviral DNA. Heavy arrows indicate the covalent attachment of retrovirus or retroelement DNA to target DNA.

The author is in the Infectious Disease Laboratory, Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92037 USA

binding (9) fits with previous work indicating that retroviral integration acceptor sites tend to be enriched near deoxyribonuclease (DNAse) I–hypersensitive sites in chromatin. These findings support a model in which tethering of the HIV integration machinery to host DNA via a protein-protein interaction with Ini1 promotes integration in open chromatin.

Tethering can be exploited to control integration; tethering an integrase enzyme to a chosen target DNA can suffice to direct localized integration (10). HIV integrase was modified by fusion to a sequence-specific DNA-binding domain, that of phage λ repressor. The modified protein (λR-IN) was tested for integration site selection in reactions in vitro containing target DNAs with or without λ repressor binding sites. Integration by λR-IN was strongly favored in those target DNAs containing repressor binding sites, whereas wild-type integrase showed no such bias. Prebinding of λ repressor blocked selective integration, indicating that binding to the λ repressor sites was required. The λR-IN protein directed integration primarily to the same face of the B-DNA helix as the operators, as expected if the λR-IN protein bound to repressor sites captures target DNA by looping out the intervening sequences. These findings bolster the idea that tethering can be sufficient to direct selective integration and establish that specific binding can be engineered into an integrase without disruption of its function (at least under the conditions examined).

Can integration site selection by retroviruses be controlled in vivo by fusion of a sequence-specific DNA-binding domain to integrase? If so, then not only might therapeutic genes be integrated at innocuous chromosomal sites, but retroviral integration might be used to deliberately disrupt harmful sequences such as activated oncogenes or integrated viral genomes. It might even be possible to modify retroviruses to serve as attenuated retroviral vaccines by using integrase fusions to allay fears of insertional mutagenesis upon vaccination.

Whether fusions of integrase to sequence-specific DNA-binding domains can be used to meet these goals in practice remains to be seen. For this idea to work, high-titer stocks of these modified viruses or vector derivatives must be readily obtainable, discrimination between specific and nonspecific sites in vivo must be adequate, and useful binding specificities must be identified and built into integrase fusions. Although some of these hurdles may be difficult to leap, the fact that the Ty elements somehow manage to target integration encourages optimism for modified retroviruses.

References

1. B Shiramizu B G Herndier M S McGrath *Cancer Res* **54** 2069 (1994)
2. J M Coffin in *Virology* B N Fields and D M Knipe Eds (Raven New York 1990) vol 2 pp 1437-1500
3. D L Chalker and S B Sandmeyer *Genes Dev* **6**, 117 (1992)
4. H Ji *et al* *Cell* **73** 1 (1993)
5. J Kirchner C M Connolly S B Sandmeyer *Science* **267**, 1488 (1995)
6. G V Kalpana S Marmon W Wang G R Crabtree, S P Goff *ibid* **266** 2002 (1994)
7. F D Bushman T Fujiwara R Craigie *ibid* **249**, 1555 (1990)
8. M Carlson and B C Laurent *Curr Opin Cell Biol* **6**, 396 (1994)
9. J Côté, J Quinn J L Workman C L Peterson *Science* **265** 53 (1994)
10. F D Bushman *Proc Natl Acad Sci U S A* **91** 9233 (1994)