Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 4695–4699, June 1991
Biochemistry

A covalent complex between retroviral integrase and nicked substrate DNA

(covalent linkage/retroviral DNA integration)

MICHAEL KATZMAN*, JOSEPH P. G. MACK†, ANNA MARIE SKALKA‡, AND JONATHAN LEIS§¶

*Departments of Medicine and Microbiology and Immunology, The Milton S. Hershey Medical Center, Pennsylvania State University College of Medicine, Hershey, PA 17033; †Crystallography Laboratory, National Cancer Institute–Frederick Cancer Research and Development Center, Frederick, MD 21702; ‡Institute for Cancer Research, Fox-Chase Cancer Center, Philadelphia, PA 19111; and §Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, OH 44106

Communicated by Harland G. Wood, February 11, 1991

ABSTRACT **Purified retroviral integrase (IN) from avian sarcoma–leukosis viruses can appropriately process the termini of linear viral DNA, cleave host DNA in a sequence-independent manner, and catalyze integrative recombination; an exogenous source of energy is not required for these reactions. Using DNA substrates containing radioactive phosphate groups, we demonstrate that IN becomes covalently joined to the new 5′ phosphate ends of DNA produced at sites of cleavage. Most of the phosphodiester linkages between IN and DNA involve serine, but some involve threonine. Computer-assisted alignment of 80 retroviral and retrotransposon IN sequences identified one serine that is conserved in all of these proteins and three less-conserved threonine residues. These results identify candidate active-site residues and provide support for the participation of a covalent IN–DNA intermediate in retroviral integration.**

Integration of a double-stranded DNA copy of the retroviral genome into host-cell chromosomal DNA requires cis-acting sequences at the ends of viral DNA and results in loss of 2 terminal viral base pairs (bp) that follow the invariant C-A dinucleotide. Although there is no sequence specificity for host integration sites, integration produces a characteristic duplication of flanking host base pairs (1, 2). These alterations of viral and host DNA suggest that the process of integration requires at least two nuclease activities with different specificities—i.e., sequence-specific cleavage of viral DNA that removes precisely two nucleotides from each terminus and sequence-independent staggered cleavage of host DNA at the site of insertion of the provirus. The *pol* gene-encoded integrase (IN) protein, shown by genetic experiments to be required for integration (1, 2), can perform both types of endonucleolytic cleavages necessary for retroviral integration (3–5) and can catalyze joining of viral long terminal repeat (LTR) sequences to target DNA *in vitro* (6–8). Because results from cell-free assays showed that a source of exogenous energy is not required for joining viral to host DNA (6, 7, 9, 10), we examined our previously described oligodeoxynucleotide assay system for evidence of a protein–DNA intermediate that might conserve energy from a DNA-cleavage event and utilize it for a DNA-joining reaction. We report here that avian sarcoma–leukosis virus IN forms a covalent complex involving serine, and to a lesser extent threonine, linked to 5′ phosphate ends produced at sites of DNA cleavage.

MATERIALS AND METHODS

Purification of IN. IN was purified to homogeneity from avian myeloblastosis virus (AMV) and from Rous sarcoma virus (RSV) sequences expressed in bacteria as described (3, 4). AMV IN was able to nick 23 pmol of substrate DNA per μg of protein with 2 mM Mn^{2+} and 2.5 pmol of DNA per μg of protein with 3 mM Mg^{2+} (4).

Oligodeoxynucleotide Substrates. Double-stranded oligodeoxynucleotides with sequences corresponding to the termini of the linear form of RSV DNA (Fig. 1 *Upper*) were prepared as described (4). The strand in each substrate that contains the conserved C-A dinucleotide nicking site was either 5′-end-labeled with [γ-^{32}P]ATP (3000 Ci/mmol; 1 Ci = 37 GBq) catalyzed by T4 polynucleotide kinase (4) or internally labeled at the 3′ end by DNA polymerase I (Klenow fragment) repair synthesis with [α-^{32}P]dTTP (800 Ci/mmol) such that two radioactive phosphates were incorporated. For simplicity, substrates are referred to as 5′- or 3′-labeled U5 or U3 duplexes.

Reaction Conditions. Standard reactions contained 1.5 pmol of labeled duplex oligodeoxynucleotides (10^6–10^7 cpm/pmol) in a final volume of 10 μl containing 25 mM Tris·HCl (pH 8.3), 10 mM 2-mercaptoethanol, 3 mM $MgCl_2$ or 2 mM $MnCl_2$, and 50 ng (1.5 pmol) of IN. After incubation at 37°C for 90 min, reactions were adjusted to 2% SDS, 7 M urea, 4 mM dithiocrythritol, 20 mM Tris·HCl (pH 6.8), and 0.004% bromophenol blue. Samples were heated at 100°C for 10 min before gel analysis.

Gel Analysis. Samples were analyzed on discontinuous 12.5% polyacrylamide/SDS gels (acrylamide to methylenebisacrylamide ratio, 59:1) with 3.5% polyacrylamide/SDS stacking gels (acrylamide to methylenebisacrylamide ratio, 38:1) by the method of Laemmli (11). Electrophoresis was at 20 W until the bromophenol blue dye migrated 12 cm. Gels were then cut horizontally 2 cm above the dye, and the upper portion was stained with Coomassie brilliant blue and dried. Autoradiographs were prepared with Kodak X-Omat film, sometimes with DuPont Cronex Lightning Plus intensifying screens, for 1–3 days.

Phospho Amino Acid Analysis. Radioactive protein–DNA complexes were excised from wet or dried gels, hydrated for 45 min, and subjected to trypsin digestion, acid hydrolysis, and thin-layer electrophoresis as described (12), except that acid hydrolysis was conducted in 4 M HCl at 110°C for 30 min. As a control, substrate DNA was processed similarly.

Protein Sequence Analysis. Sequences included in the alignment of retroviral and retrotransposon IN proteins were identified by Khan *et al.* (13). Protein sequences were deduced from the DNA sequence or extracted directly from the University of Wisconsin Genetics Computer Group or Los Alamos human immunodeficiency virus data bases and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IN, retroviral integrase; RSV, Rous sarcoma virus; AMV, avian myeloblastosis virus; LTR, long terminal repeat.
¶To whom reprint requests should be addressed.

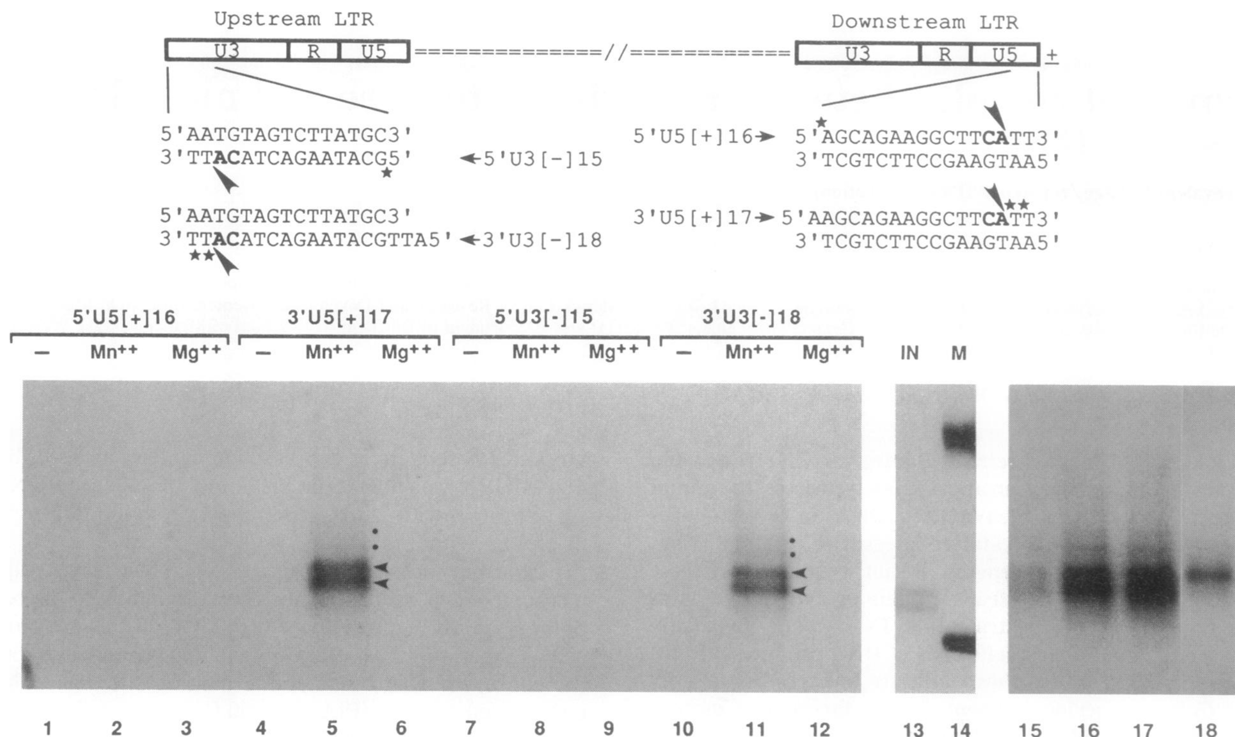

FIG. 1. (*Upper*) Oligodeoxynucleotide substrates. The linear form of viral DNA is shown schematically at the top; each LTR consists of U3, R, and U5 regions. Oligomers were named for their viral end (U3 or U5), DNA strand (plus or minus), and length. Names are shown for ^{32}P-labeled strands; prefixes indicate the relative position of radiolabels (stars). Arrowheads after the conserved C-A dinucleotide (boldface) denote the major sites of *in vitro* nicking by IN. (*Lower*) Demonstration of complexes between IN and DNA. Lanes: 1–12, labeled oligodeoxynucleotide duplexes as indicated were incubated with IN for 90 min in the absence of divalent cation (lanes 1, 4, 7, and 10) or with 2 mM Mn^{2+} (lanes 2, 5, 8, and 11) or with 3 mM Mg^{2+} (lanes 3, 6, 9, and 12), denatured, and analyzed by SDS/PAGE as described (the upper portion of the autoradiogram is shown); 13–14, Coomassie blue-stained gel, properly positioned with respect to the autoradiogram, showing the AMV IN migrating at 33–34 kDa (lane 13) and molecular mass (m) markers 45-kDa ovalbumin and 31-kDa carbonic anhydrase (lane 14); 15–17, as in lane 11 except incubations were for 15 min (lane 15), 90 min (lane 16), or 180 min (lane 17) and the autoradiogram was exposed twice as long; 18, as in lane 11 except that virion-derived IN was replaced by bacteria-derived IN.

aligned by the method of Hein (14). The sequence for Ty3-2 was manually shifted to align better with Ser-85 of RSV.

RESULTS

Formation of Complexes Between IN and DNA. We have described (4) an endonuclease assay that uses linear duplex oligodeoxynucleotides that represent the viral DNA termini as substrates for purified avian sarcoma–leukosis virus IN. More recent experiments have shown that some of the specifically cleaved products were joined to other oligodeoxynucleotides that apparently had served as surrogates for target-host DNA (6, 7). No evidence was obtained in these experiments for formation of stable complexes between IN and the 5'-labeled cleavage products, all of which contained new 3' OH ends. To examine whether IN may have interacted with the unlabeled cleavage products that had new 5' P ends, we prepared LTR substrates that were labeled near the 3' ends (Fig. 1 *Upper*, lower substrates).

When the 3'-labeled U3 DNA duplex was incubated with IN in the presence of Mn^{2+}, denatured, and analyzed by SDS/PAGE followed by autoradiography (Fig. 1 *Lower*, lane 11), strong radioactive bands were observed to migrate as a doublet (arrowheads) with relative molecular mass of 34–35 kDa. IN purified from virions migrates under these gel conditions as a 33- to 34-kDa doublet (Fig. 1 *Lower*, lane 13). Mn^{2+}-dependent digestion of the same DNA substrate by RSV IN purified from our bacterial expression system yielded a single radioactive band at 35 kDa (Fig. 1 *Lower*, lane 18). RSV IN produced in bacteria migrates under these gel conditions as a single 34-kDa band (3). Thus, the positions of these labeled bands are consistent with retardation of either IN preparation by the molecular mass of short oligomers produced by cleavage near the 3' end of this DNA substrate under Mn^{2+}-dependent conditions (4). We could not detect similar bands after digestion in the presence of Mg^{2+} when using this preparation of AMV IN (Fig. 1 *Lower*, lane 12). However, small amounts were observed after prolonged incubation with higher concentrations of Mg^{2+} and bacterial-derived IN (data not shown). Similar bands were not detected in the absence of cation (Fig. 1 *Lower*, lane 10) or when 5'-labeled U3 duplex substrates were used (lanes 7–9). Analogous results were obtained with U5 duplexes as substrates for IN (Fig. 1 *Lower*, lanes 1–6). These bands represent protein–DNA complexes because they were not detected when samples were treated with proteinase K or when IN was omitted from the reactions (data not shown). The amount of these complexes accumulated in parallel to the extent of cleavage under these conditions, reaching a plateau after 90 min (Fig. 1 *Lower*, lanes 15–17).

The postcleavage complexes described above could be eluted from gels and detected after a second round of denaturation and electrophoresis. They also were observed when 7 M urea was included in the gel. Once formed, labeled DNA could not be displaced from the complexes by addition of a 100-fold excess of unlabeled blunt-ended substrate or 10 mM EDTA. Complexes were also precipitable by the combination of SDS and K^+ as described by Trask *et al.* (15).

Further inspection of the autoradiogram shown in Fig. 1 *Lower* reveals additional bands that migrated as a doublet (dots in Fig. 1 *Lower*) at *ca.* 37–38 kDa in lanes 5 and 11 and were only evident after Mn^{2+}-dependent digestion of substrates labeled at the 3' end. These less intense bands may

Table 1. Analysis of DNA bound to IN

Size of postcleavage complexes, kDa	Relative abundance of oligodeoxynucleotides of the indicated lengths bound to IN										
	1	2	3	4	5	6	7	8	9	10	11
34–35	–	0.5	1.0	0.6	–	–	–	–	–	–	–
37–38	–	–	0.3	–	–	0.2	0.5	0.4	1.0	0.3	–

Complexes were isolated from SDS-containing gels, eluted in 0.2 M NH_4HCO_3 containing trypsin (140 μg/ml), lyophilized, and exposed to 400 mM NaOH at 37°C for 60 min. After neutralization with HCl, samples were analyzed by denaturing 20% PAGE, autoradiography, and densitometry. For each complex, results were normalized to the strongest signal in the lane. –, Values < 0.2.

represent a minor set of postcleavage complexes that arise from joining of IN to longer oligomers produced by cleavage at sites 6–10 nucleotides from the 3′ ends of the LTR sequences (4).

Comparison of the radioactive signals from postcleavage complexes and from unbound oligodeoxynucleotides indicated that only *ca.* 1% of the DNA was involved in protein–DNA complexes despite cleavage of >50% of the substrate under Mn^{2+}-dependent conditions (data not shown). Since reactions contained equimolar amounts of IN and substrate molecules, *ca.* 1% of IN (<1 ng) was trapped in the protein–DNA complexes. To confirm the size of the DNA bound to IN, the major and minor sets of postcleavage complexes were eluted from protein gels and treated with alkali to disrupt complexes at the protein–DNA linkage. DNA released from IN by this treatment was analyzed by using denaturing 20% PAGE (Table 1). As expected, the major postcleavage complexes released primarily short oligomers, whereas the minor postcleavage complexes released predominantly longer products consistent with known cleavage sites (4).

Postcleavage Complexes Have Phosphoester Linkages to Serine and Threonine. To determine the site on the DNA that was bound to IN, we tested the susceptibility to calf intestine phosphatase (CIP) of the radiolabel in postcleavage complexes that were formed by 15 min of limited digestion by IN; under these conditions most nicks occur 2 nucleotides from the 3′ end of the DNA at the site of one of the two ^{32}P groups (4). We observed that none of the label was removed by CIP treatment (data not shown). Although steric hindrance could not be excluded as an explanation for resistance to CIP, these results suggested that IN was associated with the 5′ phosphate end of DNA cleavage products.

To indicate what amino acids of IN were bound to DNA, we examined the stability of postcleavage complexes to incubation at 37°C for 60 min in a pH range from 1 to 13. Complexes were stable from pH 1 to 11, partially hydrolyzed at pH 12, and completely hydrolyzed at pH 13 (Fig. 2). The last result was not due to hydrolysis of either the DNA or the protein moiety (Table 1 and data not shown). These results make phosphoamide or thiol phosphate bonds unlikely and are most consistent with a phosphoester linkage to serine or threonine (17–19). An acyl phosphate linkage was also possible, since the complexes were partially sensitive to 1 M HCl (40% remained after 60 min). However, this type of bond was unlikely since the complex was completely resistant to 0.2 M hydroxylamine at pH 7.6 (data not shown).

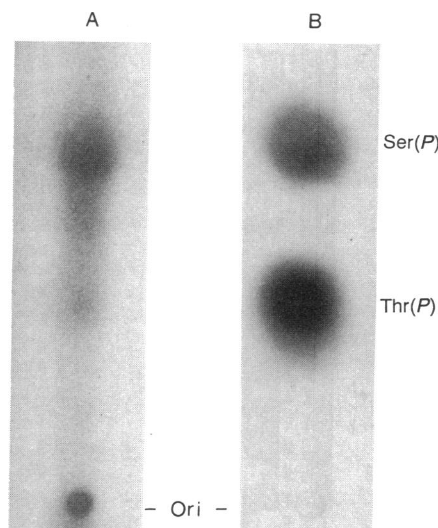

FIG. 3. Phospho amino acid analysis. (*A*) Postcleavage complexes were excised from dried gels and subjected to trypsin digestion and acid hydrolysis as described. Nonradioactive phosphoserine [Ser(*P*)] and phosphothreonine [Thr(*P*)] were added to the processed sample and resolved by thin-layer electrophoresis. An autoradiogram representative of four experiments is shown. (*B*) As in *A* except that phospho amino acids were detected by ninhydrin staining. Ori, origin.

Phospho amino acid analysis was performed to demonstrate directly transfer of a radioactive phosphate from DNA to amino acid residues of IN. Postcleavage complexes were excised from gels and subjected to trypsin digestion followed by limited acid hydrolysis. The major radioactive product migrated precisely with a phosphoserine standard analyzed by thin layer electrophoresis (Fig. 3); a minor product migrated with phosphothreonine. The background smear of radioactivity detected on the autoradiogram in Fig. 3 probably reflects incomplete hydrolysis of the radioactive DNA, as indicated by similar treatment of substrate DNA (data not shown).

Alignment of IN Sequences Reveals a Limited Number of Conserved Serine or Threonine. A previous alignment of five retroviral IN protein sequences showed five conserved serines or threonines (20). A more extensive alignment of 80 retroviral and retrotransposon sequences, of which 25 are shown in Fig. 4, revealed that there is only one conserved hydroxyl-containing residue that is present in all of these proteins. This site, containing either serine or threonine, aligns with Ser-85 of IN from RSV. The region surrounding this residue, from position 76 to 86, has conservative replacements at more than half of the positions, suggesting that this region has a conserved secondary structure and function. A second site, aligning with Thr-66 in the RSV sequence, is conserved in all of the retroviral sequences but not in the retrotransposon sequences. An additional site, usually a threonine residue, is

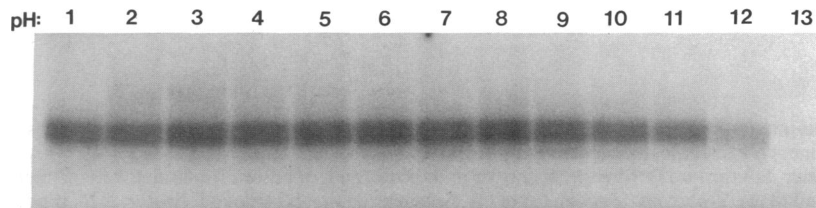

FIG. 2. Susceptibility of postcleavage complexes to varying pH. Double-stranded 3′ U3[−]18 was digested by IN in the presence of Mn^{2+}, reactions were stopped by addition of 10 mM EDTA, and buffers (16) at the indicated pH were added. After incubation for 60 min at 37°C, samples were neutralized and processed. Identical results were obtained with double-stranded 3′ U5[+]17.

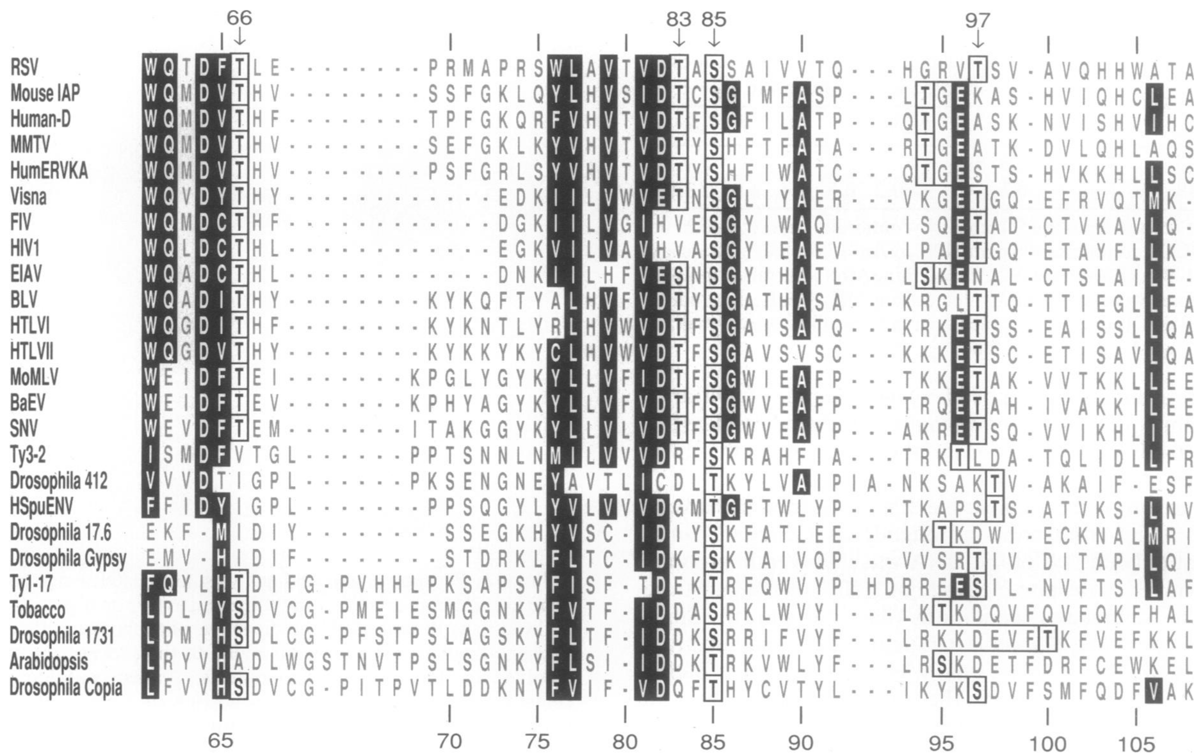

FIG. 4. Alignment of retroviral and retrotransposon IN sequences. Retroviruses and retrotransposons are listed by phylogenetic relatedness. Single-letter amino acid designations are numbered and aligned relative to residues 61–108 of the 286-amino acid RSV IN. Residue numbers are indicated at the top and bottom. Selected serine and threonine residues are boxed. Amino acids are highlighted if a single amino acid is present in at least 50% of sequences, a group of similar amino acids (ST, DE, RK, QN, IVLMC, or YFW) is present in 75% of sequences, or a larger group (IVLMCYFW) is present in 90% of sequences. IAP, intercisternal A particles; human D, human D particles; MMTV, mouse mammary tumor virus; HumERVKA, human endogenous virus; FIV and HIV, feline and human immunodeficiency viruses; EIAV, equine infectious anemia virus; BLV, bovine leukemia virus; HTLV I and II, human T-cell leukemia viruses I and II; Mo-MLV, Moloney murine leukemia virus; BaEV, baboon endogenous virus; SNV, spleen necrosis virus; HSpuENV, human spuma virus.

present in all of the sequences but only imperfectly aligns with Thr-97 of RSV. We also note a threonine in most of the retroviral sequences that aligns with position 83 of RSV. No conserved tyrosines were identified.

DISCUSSION

The ability of IN to precisely nick oligodeoxynucleotide substrates 3′ to the conserved C-A dinucleotide two nucleotides from the LTR ends has obvious relevance to the processing reaction that occurs at the termini of viral DNA integration precursors (4). Additional sites nicked by avian sarcoma–leukosis virus IN in the presence of Mn^{2+} may reflect the ability of IN to cleave host DNA during retroviral integration in a sequence-independent manner. Oligodeoxynucleotide substrates also undergo joining reactions *in vitro* that are dependent upon IN, without any requirement for

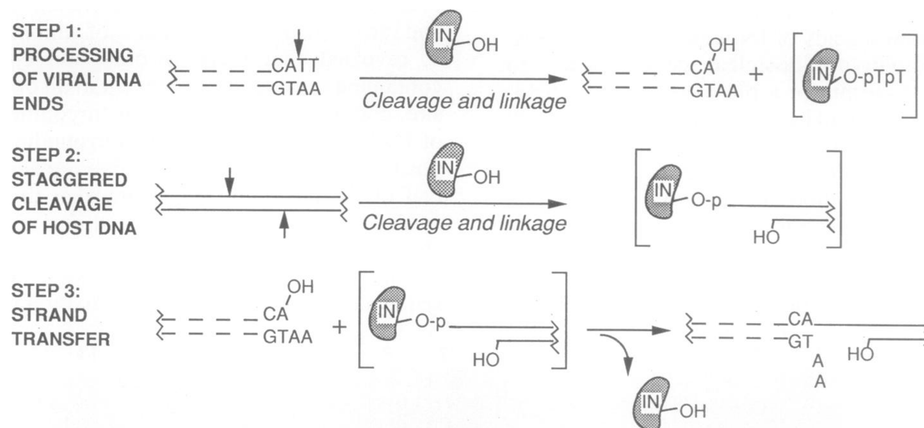

FIG. 5. Model of retroviral integration. Two possible covalent IN–DNA species are shown in brackets in steps 1 and 2. In step 1, the termini of linear viral DNA (dashed lines) are specifically processed by IN (arrow) to produce a recessed C-A-3′ OH. Only one DNA terminus is shown, and IN is depicted as a monomer, but the two DNA termini may be juxtaposed and processed simultaneously. Either a serine or threonine residue (the OH group on IN) may bind to the 5′ phosphate group (p) of the dinucleotide that is removed. In step 2, IN produces staggered nicks (arrows) in host DNA (solid lines) that are independent of DNA sequence. Either a serine or threonine residue of IN may bind to the new 5′ P ends of DNA. In step 3, IN transfers bound host DNA to the C-A end of a processed viral DNA terminus, utilizing energy stored in the covalent protein–DNA intermediate. Similar reactions would occur at each viral–host junction. Steps 2 and 3 may be concerted. Removal of the AA dinucleotide and repair synthesis would yield the integrated provirus.

exogenous energy (6, 7), suggesting that energy stored in a covalent protein–DNA intermediate might be used in the DNA-joining step. We have now detected complexes between IN and nicked DNA that were resistant to strong denaturation conditions and increased in amount as digestion proceeded. These complexes were only detected when label had been incorporated into sequences at the 3′ end of DNA and were readily detected when Mn^{2+} was the divalent cation; small amounts were detected in the presence of Mg^{2+} when the bacterial-derived IN was used. This difference could be due to the inefficiency of the Mg^{2+}-dependent reaction, typically 10% of Mn^{2+}-containing reactions (4). We note that *in vitro* integrative recombination occurs in the presence of Mg^{2+}, but at low efficiency (6). The complexes shown in Fig. 1 *Lower* involve covalent phosphoester bonds between serine, and to a lesser extent threonine residues, of IN and 5′ *P* ends of DNA produced at sites of DNA cleavage. Although we have not determined experimentally the position of IN residues linked to DNA, we have used computer-assisted analysis of 80 IN proteins to identify conserved serine or threonine residues. One of these, corresponding to Ser-85 of RSV, is conserved in all of the retroviral and retrotransposon IN sequences and thus is an attractive candidate for the putative catalytic site. Threonine at positions 66, 83, and 97 of RSV, which are less well conserved, may also be involved in these reactions.

Preliminary attempts to isolate postcleavage complexes in such a way that enzymatic activity of IN is retained have not been successful. We estimate that only 1% of IN and nicked DNA were present in postcleavage complexes. This is *ca.* 2% of cleavage and 10% of the joining reaction that is observed in this system (6, 7). Detection of small amounts of intermediate has also been reported in studies of the Int protein of bacteriophage λ, which covalently binds via tyrosine to 3′ *P* ends produced by its nicking activity (21, 22). Most DNA nicking/joining enzymes, the majority of which are topoisomerases, utilize tyrosine for transient covalent linkages to nicked DNA (19, 23). However, both the $\gamma\delta$ resolvase from the Tn*3* family of transposable elements and the invertase Gin of phage Mu bind via serine to 5′ *P* groups created by their nicking activities (24, 25).

Linkage of two amino acids of IN to cleaved DNA was unexpected. Interpretation of this result is complicated by the dual roles of the oligodeoxynucleotides in these reactions; they are utilized as surrogates both for specifically cleaved viral LTRs and nonspecifically cleaved host-target DNA (6, 7). It may be that there are two sites on IN that can participate in cleavage/covalent-linkage reactions, one that recognizes and nicks viral DNA and a second that acts on host DNA (steps 1 and 2 in Fig. 5). Linkages involving two active sites have been described previously for the ϕX174 gene *A* protein; in that system, 2 different tyrosine residues are joined to 5′ *P* groups (26). Alternatively, IN may have only one site that is involved in a cleavage/linkage reaction, followed by intramolecular or intermolecular transfer of bound DNA to a second amino acid, thus preparing a single active site on each protein molecule for a second cleavage reaction. It is also possible that transfer of host DNA between protein sites is an integral part of the DNA-joining reaction (step 3, Fig. 5). Site-directed mutagenesis of these conserved amino acid residues may help to distinguish between these possibilities.

Recent studies with murine leukemia virus (7, 27) and ASLV (6) indicate that energy for the DNA-joining reaction between viral and host DNA probably comes from cleavage of phosphodiester bonds in host DNA, since precleaved LTRs can participate in integrative recombination. In the system described here, where the oligodeoxynucleotides serve both as donor and target, it is the variably sized oligomer cleavage products from the 3′ ends of viral LTRs that have been trapped linked to IN. We propose that a similar linkage reaction occurs at host-DNA cleavage sites (Fig. 5, step 2), as suggested by the Mn^{2+}-dependent joining of IN to various sites on these substrates. Further experiments are required to verify that covalent complexes form as a result of cleavage of host DNA. The results presented here, however, indicate that some step in the process of integration proceeds via covalent linkage between IN and DNA.

We thank Jolinda Traugh and Polygena Tuazon of the University of California, Riverside, for performing the phospho amino acid analysis. The bacterially produced IN was provided by George Merkel of Fox–Chase Cancer Center. We would like to thank Michael Gribskov of the National Cancer Institute's Frederick Cancer Research Development Center (NCI-FCRDC) in Frederick, MD, for helpful discussions and the NCI-FCRDC Advanced Scientific Computing Lab for computing time and staff support. This work was supported in part by Public Health Service Grants CA 38046 to J.L. and CA 49042, CA 06927, and RR 05539 to A.M.S. from the National Cancer Institute; by Cancer Contract N01 C01 74101 to A. Wlodawer; by Cancer Research Center Grant P30 CA 4703 to J.L.; by an appropriation from the Commonwealth of Pennsylvania to A.M.S.; and by American Cancer Society Physician's Research Training Fellowship (PRTF-77) and an Infectious Diseases Society of America Ortho Pharmaceutical Corporation Young Investigator Award in Infectious Diseases to M.K.

1. Skalka, A. M. (1988) in *Genetic Recombination*, eds. Kucherlapati, R. & Smith, G. R. (Am. Soc. Microbiol., Washington), pp. 701–724.
2. Varmus, H. E. & Brown, P. (1989) in *Mobile DNA*, eds. Berg, D. & Howe, M. (Am. Soc. Microbiol., Washington), pp. 53–108.
3. Terry, R., Soltis, D. A., Katzman, M., Cobrinik, D., Leis, J. & Skalka, A. M. (1988) *J. Virol.* **62**, 2358–2365.
4. Katzman, M., Katz, R. A., Skalka, A. M. & Leis, J. (1989) *J. Virol.* **63**, 5319–5327.
5. Grandgenett, D. P., Vora, A. C. & Schiff, R. D. (1978) *Virology* **89**, 119–132.
6. Katz, R. A., Merkel, G., Kulkosky, J., Leis, J. & Skalka, A. M. (1990) *Cell* **63**, 87–95.
7. Craigie, R., Fujiwara, T. & Bushman, F. (1990) *Cell* **62**, 829–837.
8. Bushman, F. D., Fujiwara, T. & Craigie, R. (1990) *Science* **249**, 1555–1558.
9. Brown, P. O., Bowerman, B., Varmus, H. E. & Bishop, J. M. (1987) *Cell* **49**, 347–356.
10. Ellison, V., Abrams, H., Roe, T., Lifson, J. & Brown, P. (1990) *J. Virol.* **64**, 2711–2715.
11. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
12. Pendergast, A. M. & Traugh, J. A. (1985) *J. Biol. Chem.* **21**, 11769–11774.
13. Khan, E., Mack, J. P. G., Katz, R. A., Kulkosky, J. & Skalka, A. M. (1990) *Nucleic Acids Res.*, in press.
14. Hein, J. (1990) *Methods Enzymol.* **183**, 626–645.
15. Trask, D. K., DiDonato, J. A. & Muller, M. T. (1984) *EMBO J.* **3**, 671–676.
16. Bates, R. G. & Paablo, M. (1975) in *Handbook of Biochemistry and Molecular Biology: Physical and Chemical Data*, ed. Fasman, G. D. (CRC, Boca Raton, FL), pp. 353–366.
17. Pigiet, V. & Conley, R. R. (1978) *J. Biol. Chem.* **253**, 1910–1920.
18. Shabarova, Z. A. (1970) in *Progress in Nucleic Acid Research and Molecular Biology*, eds. Davidson, J. N. & Cohn, W. E. (Academic, New York), Vol. 10, pp. 145–182.
19. Champoux, J. J. (1981) *J. Biol. Chem.* **256**, 4805–4809.
20. Johnson, M. S., McClure, M. A., Feng, D.-F., Gray, J. & Doolittle, R. F. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7648–7652.
21. Craig, N. L. & Nash, H. A. (1983) *Cell* **35**, 795–803.
22. Pargellis, C. A., Nunes-Duby, S. E., de Vargas, L. M. & Landy, A. (1988) *J. Biol. Chem.* **263**, 7678–7685.
23. Wang, J. C. (1985) *Annu. Rev. Biochem.* **54**, 665–697.
24. Reed, R. R. & Moser, C. D. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 245–249.
25. Klippel, A., Mertens, G., Patschinsky, T. & Kahmann, R. (1988) *EMBO J.* **7**, 1229–1237.
26. Roth, M. J., Brown, D. R. & Hurwitz, J. (1984) *J. Biol. Chem.* **259**, 10556–10568.
27. Brown, P. O., Bowerman, B., Varmus, H. E. & Bishop, J. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2525–2529.