Requirement for a conserved serine in both processing and joining activities of retroviral integrase

(retroviral DNA integration/*in vitro* integration/integrase mutations)

RICHARD A. KATZ*, JOSEPH P. G. MACK†, GEORGE MERKEL*, JOSEPH KULKOSKY*, ZHENG GE‡,
JONATHAN LEIS‡, AND ANNA MARIE SKALKA*§

*Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111; †Crystallography Laboratory, National Cancer Institute–Frederick Cancer Research and Development Center, Frederick, MD 21702; and ‡Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, OH 44106

ABSTRACT Retroviruses encode a protein, the integrase (IN), that is required for insertion of the viral DNA into the host cell chromosome. IN alone can carry out the integration reaction *in vitro*. The reaction involves endonucleolytic cleavage near the 3' ends of both viral DNA strands (the processing step), followed by joining of these new viral DNA ends to host DNA (the joining step). Based on their evolutionary conservation, we have previously identified at least 11 amino acid residues of IN that may be essential for the reaction. Here we report that even conservative replacements of one of these residues, an invariant serine, produce severe reductions in both the processing and joining activities of Rous sarcoma virus IN *in vitro*. Replacement of the analogous serine of the type 1 human immunodeficiency virus IN had similar effects on processing activity. These results suggest that this single conserved serine is a component of the active site and that one active site is used for both processing and joining. Replacement of this serine with certain amino acids resulted in a loss or reduction in DNA binding activities, while other replacements at this position appeared to affect later steps in catalysis. All of the defective Rous sarcoma virus INs were able to compete with the wild-type protein, which supports a model in which IN functions in a multimeric complex.

The retroviral replication cycle includes reverse transcription of viral genomic RNA and integration of the linear duplex DNA product into the host cell chromosome (1, 2). The integration reaction is site specific with respect to viral DNA and largely sequence independent with respect to host DNA. A virion protein, the integrase (IN), encoded in the retroviral *pol* gene is required for integration *in vivo* (3, 4) and is sufficient for integration *in vitro* (5, 6). The *in vitro* reaction requires only IN, divalent metal ions, and substrate DNA that contains cis-acting sequences that mimic the recombination sites at the ends of the linear viral DNA. *In vivo* and *in vitro* experiments have revealed a two-step reaction (Fig. 1A). In the first step, the *processing* reaction, the IN endonuclease activity introduces nicks, usually two nucleotides from the 3' end of each viral DNA strand, to expose the conserved CA dinucleotides (11, 12). In the second step, the *joining* reaction, the two processed 3'-hydroxyl ends of viral DNA are joined to staggered 5'-phosphate ends of host DNA (5, 6). The initial product is a "gapped intermediate" (13, 14), which may be repaired by cell-mediated mechanisms.

Models for IN function must account for the site-specific nicking of viral DNA during processing and the sequence-independent recognition of host DNA during joining. These properties suggest that IN may either have one active site with flexibility in substrate recognition or separate active
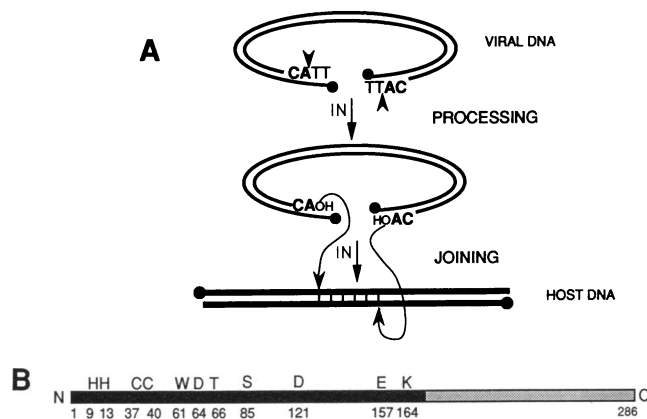

FIG. 1. (A) Illustration of the processing and joining steps in retroviral DNA integration. See text for details. The 5'-phosphate ends of each strand are indicated by a closed circle. (B) Linear representation of retroviral IN protein showing conserved amino acids identified by alignment of deduced amino acid sequences from ≈70 retroviral INs (7–9). The expanded data base of retroviral IN sequences now available has allowed a refinement of the alignments originally described by Johnson *et al.* (10). The numbering shown is for the RSV IN, and amino acids are indicated in the single-letter code. The more highly conserved N-terminal region is shown in black.

sites. Amino acid sequence alignments have identified at least 11 evolutionarily conserved amino acid residues of the retroviral IN protein that are initial candidates for participation in DNA substrate recognition and catalysis in the processing and joining reactions (Fig. 1B; refs. 7–10). Deduced protein sequence alignments have been used to predict the active site residues of retroviral RNase H (15) and protease (16). In this paper we evaluate the requirement for the single conserved Ser and the single conserved Thr of IN. These residues correspond to Ser-85 and Thr-66 of Rous sarcoma virus (RSV) IN (7). Since the processing and joining activities can be assayed separately *in vitro*, we also asked whether these conserved amino acids participate exclusively in either one of the two reactions.

MATERIALS AND METHODS

Bacterial Expression and Purification of RSV and Type 1 Human Immunodeficiency Virus (HIV-1) IN Fusion Proteins. We have previously described the expression in *Escherichia coli*, and purification, of a *Staphylococcus aureus* protein

---

Abbreviations: RSV, Rous sarcoma virus; IN, integrase; HIV-1, type 1 human immunodeficiency virus; MBP, maltose binding protein.
§To whom reprint requests should be addressed.

A–RSV IN fusion (8). The amount of recovered active protein, measured by both silver staining of SDS/polyacrylamide gels and Bio-Rad protein assays, was ≈60 μg from 30 ml of culture.

For HIV-1 IN a similar bacterial fusion protein system was constructed by using the vector pMAL-c (New England Biolabs), which contains the coding region for maltose binding protein (MBP). Details of the construction of this expression clone and affinity purification of the fusion protein by binding to amylose resin are provided elsewhere (9). The yield was ≈200 μg of protein from 30 ml of culture. All of the substituted RSV and HIV-1 fusion proteins were stably expressed in *E. coli* and were isolated in a soluble form.

Oligonucleotide-Directed Mutagenesis. Our procedures for mutagenesis have been described (8).

Assays for Processing and Joining. Processing reactions were carried out as described (6, 11). For RSV IN, the substrate was a 15-mer oligodeoxynucleotide duplex that mimics the U3 end of linear RSV DNA (Fig. 2C). The minus strand was labeled at its 5′ end by using [γ-^{32}P]ATP and polynucleotide kinase. Reaction mixtures (50 μl) contained 3 mM $MnCl_2$, 20 mM Tris·HCl (pH 8.0), 2 mM 2-mercaptoethanol, 3 pmol of protein A–IN fusion protein, and 1 pmol of DNA substrate. For assaying the HIV-1 IN–MBP fusion, 20 pmol of protein was incubated with 1 pmol of 25-mer duplex substrate representing the U5 end of HIV-1 DNA. Incubation was at 37°C in a 15-μl reaction volume containing 10 mM Tris·HCl (pH 7.5), 50 mM NaCl, 10 mM $MnCl_2$, and 6%

dimethyl sulfoxide. RSV IN joining activity was measured independently of processing by using a DNA substrate duplex in which the labeled strand lacked the two terminal residues (5, 6). Processing and joining products were detected on 20% acrylamide gels containing urea, as described (6, 11), and quantified using a radioisotope scanning system (AMBIS).

RESULTS

Effects of Amino Acid Replacements on RSV and HIV-1 IN Processing Activity. RSV IN–protein A fusions (8) containing Ser-85 amino acid replacements were analyzed for processing activity (ref. 11; Fig. 2C). The protein A–wild type IN fusion cleaves preferentially at the −2 position, producing the expected processed product (denoted PP) (Fig. 2 A and B, lanes 4). Cleavage at −3 represents a secondary IN cleavage site (11); such incorrectly cleaved molecules do not efficiently join to target DNA (6). The small amount of −1 product, seen also in the protein A controls (lanes 2 and 3), probably reflects the presence of contaminating bacterial nucleases. All Ser-85 replacements [with Ala (S85A), Cys (S85C), Gly (S85G), Asn (S85N), or Thr (S85T)] reduced the amount of the −2 product (Fig. 2A, lanes 5–8; data not shown for S85N) to the background level observed with the protein A controls (Fig. 2A, lanes 2 and 3). Prolonged incubations gave similar results (Fig. 2B, lanes 5–8). In some cases, preferential cleavages at −1, −3, and at other sites were
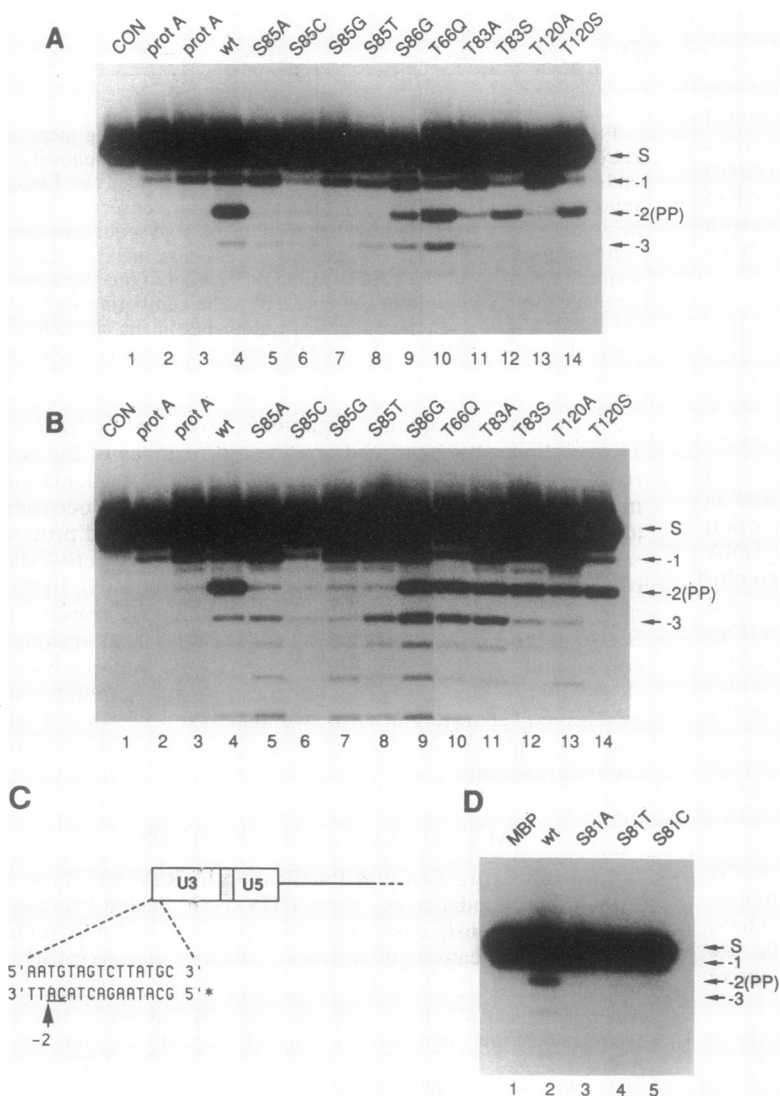

FIG. 2. Effect of amino acid replacements on IN processing. (*A* and *B*) Protein A–RSV IN fusions were assayed for processing by using a duplex oligodeoxynucleotide substrate that mimics the U3 end of RSV viral DNA (see *Materials and Methods*; refs. 6 and 11). (*C*) Sequence of the RSV substrate and illustration of the strategy. The IN processing site is indicated by an arrow, the strand that will be cleaved by IN is labeled at the 5′ end, indicated by asterisk, and the conserved CA is underlined. Products were analyzed by electrophoresis in 20% acrylamide/urea gels. The IN fusion protein included in each assay is identified above the lane by its amino acid replacement. wt, Wild type; S, position of the substrate strand; prot A, protein A. Cleavage sites are identified by the distance from the 3′ end (−1, −2, −3) of the strand that is cleaved. The −2 band corresponds to the expected processed product (PP). For analyses with RSV IN, two separate preparations of protein A, purified in a similar manner as the fusion proteins, were used as controls (lanes 2 and 3). Lane 1, no protein added. (*A*) One-hour reactions. (*B*) Four-hour reactions. (*D*) Wild-type and Ser-81-substituted HIV-1 IN–MBP fusion proteins were assayed by using a 25-mer oligodeoxynucleotide duplex substrate that mimics the U5 end of HIV-1 DNA (9). Strategy and gel conditions were similar to that used for RSV IN (*A–C*). Lane 1, MBP purified in parallel (negative control); lane 2, wild-type MBP–HIV-1 IN; lanes 3–5, Ser-81 replacements as indicated.

Genetics: Katz *et al.*

Proc. Natl. Acad. Sci. USA 89 (1992) 6743

observed after prolonged incubation with the Ser-85 replacements, indicating some residual nonprocessing IN endonuclease activity. While highly conservative chemical substitutions of the invariant Ser-85 (S85T and S85C) inactivated the normal processing activity of IN, substitution of the adjacent nonconserved Ser (S86G) resulted in only a modest ($\approx$2-fold) reduction (Fig. 2*A*, lane 9). An S86A substitution behaved as wild type (data not shown).

T66Q and T66A replacements showed wild-type processing activities (Fig. 2 *A* and *B*, lanes 10; data not shown for T66A). We conclude therefore that this conserved residue is not essential for processing. RSV Thr-83, Thr-97, and Thr-120 show significant, although not strict, evolutionary conservation (7), and amino acid replacements were also introduced at these positions for comparison. Replacement of Thr-83 or Thr-120 with Ser (T83S and T120S) had little effect (Fig. 2 *A* and *B*, lanes 12 and 14). Replacement with Ala (T83A and T120A) was more deleterious (Fig. 2*A*, lanes 11 and 13). However, with extended incubation, these substitutions also showed significant processing activity (Fig. 2*B*, lanes 11 and 13). RSV T97A and T97S replacements retained high levels of processing activity even in the shorter incubation times (data not shown).

Ser-81 of HIV-1 IN is the invariant Ser that corresponds to RSV S85 (Fig. 1*B*). If the evolutionary conservation of this residue signifies a common critical function, then HIV-1 Ser-81 should also be essential for IN activity. We have expressed HIV-1 IN in bacteria as a fusion protein with MBP (9). The wild-type fusion protein produced the expected -2 cleavage of a duplex HIV-1 DNA substrate, which is indicative of normal processing activity (Fig. 2*D*, lane 2). Replacement of Ser-81 with Ala (S81A), Thr (S81T), or Cys (S81C) reduced activity to background levels.

Effects of Amino Acid Replacements on RSV IN Joining Activity. The joining activity of IN can be measured independently of processing by using oligodeoxynucleotide du-

B

5'AATGTAGTCTTATGCAAT 3'
3' ₋ACATCAGAATACGTTA 5'*

5'AATGTAGTCTTATGCAAT 3'
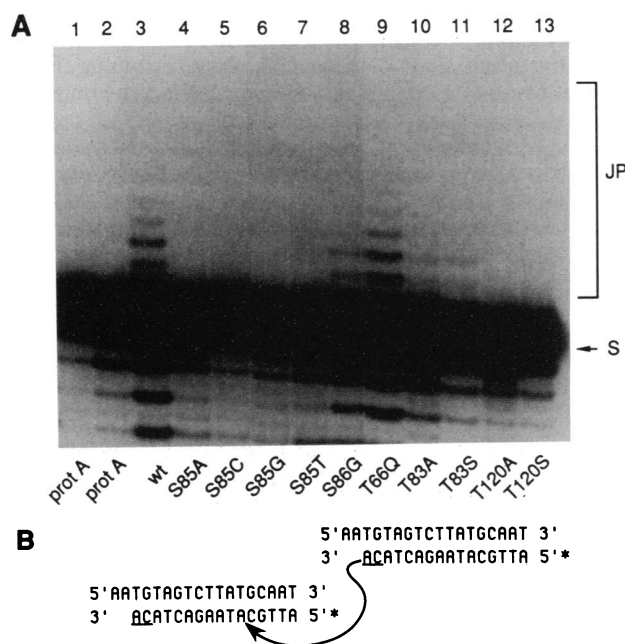3' ACATCAGAATACGTTA 5'*

FIG. 3. Effect of amino acid replacements on RSV IN joining activity. (*A*) Reactions were carried out as described in Fig. 2, except that the substrates mimic processed ends. Incubation was for 5 hr. JP, joined products; S, substrate strand; wt, wild type. (*B*) Diagram illustrating joining reactions. The 3' recessed ends of substrates (see Fig. 2*C*) can join to multiple sites on either strand of substrate DNAs. For simplicity, one joining site is shown. Other symbols are as in Fig. 2*C*. prot A, protein A.

plex substrates that mimic processed viral DNA ends (refs. 5 and 6; Fig. 3*B*). Incubation of this substrate with the wild-type protein A–IN fusion results in the formation of a ladder of longer, joined products (denoted JP), which is indicative of the reaction that joins the preprocessed ends to any one of many sites on the surrogate target DNA (Fig. 3, lane 3). No joining was detected with any of the S85 substitutions by this assay, even if incubation was extended to 5 hr (lanes 4–7). Using a more sensitive joining assay that employs circular plasmid DNA as a target (5), we were able to detect some activity with the S85T and S85C mutant proteins (data not shown). The T66Q substitution had little effect on joining. All of the other substitutions showed effects on joining that were similar to those observed with processing.

Effects of RSV IN Replacements on Sequence-Independent DNA Binding. IN exhibits sequence-independent DNA binding, which presumably reflects the ability of retroviruses to integrate randomly into host DNA. We evaluated the effects of amino acid substitutions on this binding activity by using a poly[d(A-T)] substrate (refs. 8 and 17; Table 1). The S85C, S85N, and S85T replacements did not significantly affect DNA binding as compared to the wild-type IN fusion protein. These results suggest that such conservative substitutions do not cause improper folding of IN proteins. However, since we, as others (18), have not yet been able to discriminate between binding to viral DNA ends and host DNA, we cannot exclude the possibility that Ser-85 substitutions affect specific recognition of viral DNA. The S85A and S85G substitutions did have more significant effects on sequence-independent binding; in fact, no binding above background was detected with S85G. Gly and Ala substitutions at neighboring residues (S83A, S86G, and S86A) had no significant effect on DNA binding. Also, all of the Thr replacements retained wild-type, or near wild-type, binding activity.

Ser-85-Substituted RSV IN Is Able to Compete in the Processing Reaction. We used a second approach to verify that the RSV Ser-85 substitutions did not induce gross structural alterations that might have global effects on IN functions. Our strategy was to determine whether the substituted proteins were able to compete with wild-type IN in the processing reaction. In these experiments, the amount of U3 duplex substrate was limiting, and the protein A–mutant IN fusions were present in an $\approx$2-fold molar excess over nonfused bacterially produced IN (17). Results showed that addition of S85A, S85C, or S85T IN fusion proteins reduced the processing activity of wild-type IN by $\approx$5-fold (Fig. 4*A*;

Table 1. DNA binding activities of protein A–IN fusions

Protein	DNA binding, %	Protein	DNA binding, %
Wild type	100	T66Q	126
S85A	24	T83A	99
S85C	84	T83S	90
S85G	3	T97A	120
S85N	79	T97S	136
S85T	43	T120A	40
S86A	125	T120S	87
S86G	105	Protein A	4
T66A	114	Protein A	6

DNA binding was measured by retention of ^{3}H-labeled poly[d(A-T)] on nitrocellulose filters (17). Reaction volumes were 100 μl and contained $\approx$1 μg (20 pmol) of fusion protein and 75 pmol of ^{3}H-labeled poly[d(A-T)]. The reaction buffer was 10 mM Tris·HCl (pH 8.0), 5 mM dithiothreitol, 10 mM KCl, and 10 mM $MgCl_2$. Incubations were carried out for 5 min at 37°C. Under these conditions the amount of DNA bound was dependent on protein concentration. With wild-type protein, $\approx$50% of the DNA was bound. Background counts obtained without added protein (2–3%) were subtracted from each sample. Results of representative experiments are shown.
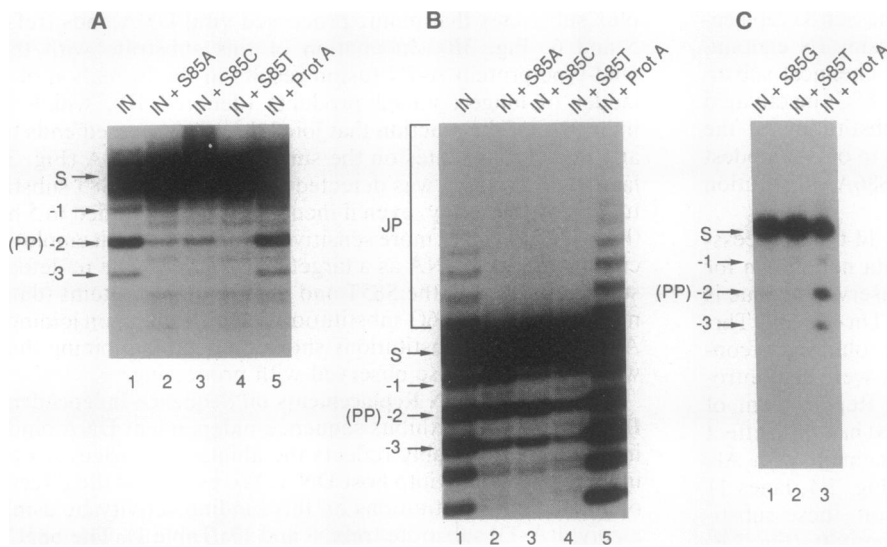
FIG. 4.   Competition with RSV IN Ser-85 replacements. Processing reactions were carried out as in Fig. 2. Each processing competition reaction (100 μl) contained 2 pmol of bacterially produced wild-type RSV IN (nonfusion), 4 pmol of the indicated protein A–IN fusion, and 1 pmol of blunt-end U3 DNA substrate. DNA was added first, followed by the competitor fusion protein indicated above each lane. This mixture was incubated on ice for 10 min. The wild-type IN was then added, followed by MnCl₂. Reactions were then incubated at 37°C for 1 or 30 min. (*A*) One-minute reaction. (*B*) Thirty-minute reaction. Joined products (JP) are indicated. (*C*) Competition by S85G. PP, expected processed product; Prot A, protein A.

lane 1 versus lanes 2–4), which suggests that these substituted proteins could bind the viral DNA substrate. No competition was detected with protein A alone (lane 5). Extended incubations (Fig. 4*B*) also showed a significant reduction in joining activity. Interestingly, the S85G protein, which did not bind to poly[d(A-T)] (Table 1), was also able to compete with wild-type IN in the processing assay (Fig. 4*C*). Such competition may reflect protein–protein interactions between wild-type and S85G IN molecules; this interpretation is consistent with the observation that IN can form dimers (19) and suggests that the wild-type/mutant heteromultimers are nonfunctional (see *Discussion*).

Role of RSV Ser-85 in DNA Linkage. We addressed a possible role of RSV Ser-85 in providing a hydroxyl oxygen for nucleophilic attack on host or viral DNA phosphates. Such a role was proposed based on the detection of a covalent RSV IN–DNA complex, produced during the *in vitro* reaction, that involved <1% of IN molecules (7). In this complex, the 5′-phosphate of cleaved DNA is linked to a hydroxyl group, primarily of Ser, but also of Thr. One or both of these linkages were proposed to represent a short-lived intermediate step in the integration reaction (7). Since Ser-85 and Thr-66 of RSV correspond to the only conserved Ser or Thr in IN and a catalytic residue would be expected to be conserved, RSV Ser-85 and/or Thr-66 were proposed as candidates for DNA linkage in the intermediate complex.

To test this model, we assayed for formation of the covalent complex by using several of the RSV Ser-85 replacements. The viral DNA substrate was a duplex similar to that used for the processing reaction except that the 3′ end, rather than the 5′ end, of the target strand was labeled (Fig. 5*B*) to allow detection of the covalent linkage of cleaved 5′ ends to IN. The radioactive complexes were detected by protein gel electrophoresis (Fig. 5*A*). No complexes of 60 kDa were detected with protein A alone, nor were any detected corresponding to the size of protein A, ≈25 kDa (the 25-kDa region is not shown in this figure). The protein A–wild-type IN fusion produced a covalent complex of ≈60 kDa (lane 2), confirming results using the nonfusion protein (7). We observed a doublet band, both radioactive and stained (data not shown), which thus appears to be a result of heterogeneity in the protein preparations, presumably at the N terminus within the protein A portion. The T66A replacement showed an extent of labeling similar to wild type, and we therefore conclude that this residue cannot be a major site for covalent attachment of DNA. Complex formation was somewhat low, but clearly detectable, with the S85T, S85N, and S85C replacements. The reduced activity is not surprising, given the severe enzymatic defects observed with these

proteins. However, these results seem inconsistent with the proposal that Ser-85 is a major site for covalent attachment. Complex formation was barely detectable with the S85A and S85G replacements (lanes 5 and 7); thus, efficient DNA binding appears to be required for this activity.

DISCUSSION

The identification of amino acid residues of the retroviral IN that may be components of its active site is an important step in elucidation of the mechanism of integration. Side chains of active site residues are expected to contact the DNA substrates, coordinate the required metal ion(s), or participate directly in catalysis. The analyses reported here show that the single evolutionary conserved Ser corresponding to RSV Ser-85 and HIV-1 Ser-81 is essential for viral DNA processing activity *in vitro*. Chemically conservative (Cys or Thr) or neutral (Ala) replacements resulted in at least a 10- to 30-fold reduction in activity in both systems. RSV IN joining activity was shown to be inhibited to a similar degree. In contrast to
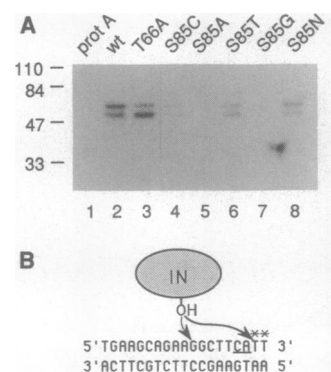

FIG. 5.   Assay for covalent complex formation between RSV IN and cleaved substrate DNA (7). (*A*) Reaction mixtures contained ≈170 ng (3 pmol) of RSV IN–protein A fusion and 1 pmol of substrate duplex labeled at the 3′ end of the donor strand by repair synthesis using [α-³²P]dTTP (7). Incubations were for 90 min. Radiolabeled complexes were detected by electrophoresis on 15% protein gels, and size markers (in kDa) are indicated. (*B*) Postulated scheme for generation of the covalent complex through nucleophilic attack of DNA phosphates by an RSV IN Ser or Thr hydroxyl oxygen. The 5′-phosphate ends of DNA become joined to a Ser or Thr hydroxyl on the protein. The thicker arrow indicates attack at the processing site; the thin arrow indicates attack at random sites on the substrate DNA, as previously discussed (7). The asterisks indicated ³²P label at the 3′ end of the viral target strand. wt, Wild type; prot A, protein A.

Genetics: Katz *et al.*

Proc. Natl. Acad. Sci. USA 89 (1992)   6745

the conserved Ser, the conserved Thr (RSV Thr-66) was not required for RSV IN processing or joining activities. We speculate that this residue is important for a function that is not detected *in vitro*.

The RSV S85C, S85N, and S85T replacements retained from ≈40% to 85% DNA binding activity as compared to wild type (Table 1). Also, these proteins were able to compete with wild-type IN in both the processing and joining reactions under conditions of limiting substrate (Fig. 4). These results suggest that (*i*) these replacements primarily affect catalysis rather than substrate binding and (*ii*) proper protein folding of the substituted proteins is maintained. However, the RSV S85A and S85G replacements, which cannot provide hydrogen bonds at position 85, showed more significantly reduced or nondetectable DNA binding, respectively (Table 1). Ser, Asn, and Thr hydrogen bond interactions with DNA bases and phosphates have been described (20), and one interpretation of our results is that the Asn, Thr, and Cys side chains may partially compensate for Ser in DNA binding. If so, the drastic decrease in activity may indicate that the DNA substrates are not positioned for cleavage properly by these replacement side chains. However, we cannot rule out the possibility that the Ala and Gly replacements at position 85 may lead to loss of DNA binding through structural perturbations at the active site, rather than by eliminating direct contacts with DNA.

The mechanism of catalysis by IN is unknown, but data supporting two different models have been presented (7, 21). Both models view processing and joining as similar reactions, which involve nucleophilic attack on DNA phosphates by a hydroxyl group oxygen. This view is supported by our observation that all replacements described here and elsewhere (9) affected both reactions to similar extents. In processing, the ultimate nucleophile for cleavage of viral DNA is presumably provided by water, since pTpT dinucleotide (or equivalent) is the major product. In joining, the ultimate nucleophile is the 3′-hydroxyl oxygen at the processed end of viral DNA, which is joined to the 5′-phosphate of cleaved host DNA. As described above, the covalent intermediate model proposed for IN postulates that a conserved hydroxyl oxygen (RSV Ser-85 or Thr-66) on the protein serves as the initial nucleophile for both the processing and joining reactions; the resulting intermediate would then be attacked by the ultimate nucleophilic oxygen. We have shown here that the conserved Ser is essential for processing and joining, and this is consistent with a critical role for this residue. Since RSV S85A and S85G show a drastic reduction in formation of covalent complexes, efficient binding of IN to DNA seems to be required for this activity. However, wild-type levels of a covalent complex are observed with the RSV Thr-66 substitutions, and significant amounts of complex can still be detected with several RSV Ser-85 replacements, even though the efficiency of both the processing and joining reactions is drastically reduced in the latter case (Fig. 5). Thus, neither of these conserved residues appears to be the major site for covalent linkage. Other preliminary results (Z.G. and J.L., unpublished results) suggest that the covalent linkage reaction can be separated from the processing and joining reactions and thus may not represent an intermediate step. Vink *et al.* (22) have shown that the hydroxyl oxygen of free Ser is also able to act as a nucleophile for the processing reaction with HIV-1 IN, suggesting that IN is able to position a target DNA phosphate for cleavage by many nucleophiles. It may be, as these investigators suggest, that such an adventitious reaction accounts for the observed covalent RSV IN–DNA complex. The second model for the catalytic mechanism has been proposed by Engelman *et al.* (21) from stereochemical analysis of the cleaved DNA phosphates produced by HIV-1 IN. This model proposes that processing and joining occur via direct attack of DNA phosphates by the ultimate nucleophiles, rather than through

covalent intermediate steps. At present, this appears to be the most likely mechanism; however, the molecular details of these reactions have yet to be elucidated.

Most existing models hypothesize that IN functions as at least a dimer, with each monomer recognizing one end of linear viral DNA and together catalyzing the coordinated insertion of the two ends at a target site on host DNA. One obvious question is whether an IN monomer uses the same or different active sites to catalyze reactions that involve viral and host DNA. In the case of all replacements examined here, the effects on joining paralleled the effects on processing. These observations suggest a model in which viral and host DNA co-occupy, or sequentially occupy, a single active site. The fact that the S85G replacement, for which DNA binding is undetectable, can act as a trans-dominant negative competitor of IN suggests that IN may function in a multimeric complex in which each subunit contributes to activity. In agreement with this interpretation, our recent kinetic and sedimentation analyses of RSV IN indicate that the minimal functional unit for both the processing and joining of oligodeoxynucleotide substrates is a dimer (23).

1. Brown, P. O. (1990) *Curr. Top. Microbiol. Immunol.* **157**, 19–48.
2. Kulkosky, J. & Skalka, A. M. (1990) *J. AIDS* **3**, 839–851.
3. Donehower, L. A. & Varmus, H. E. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6461–6465.
4. Schwartzberg, P., Colicelli, J. & Goff, S. P. (1984) *Cell* **37**, 1043–1052.
5. Craigie, R., Fujiwara, T. & Bushman, F. (1990) *Cell* **62**, 829–837.
6. Katz, R. A., Merkel, G., Kulkosky, J., Leis, J. & Skalka, A. M. (1990) *Cell* **63**, 87–95.
7. Katzman, M., Mack, J. P. G., Skalka, A. M. & Leis, J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 4695–4699.
8. Khan, E., Mack, J. P. G., Katz, R. A., Kulkosky, J. & Skalka, A. M. (1991) *Nucleic Acids Res.* **19**, 851–860.
9. Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P. G. & Skalka, A. M. (1992) *Mol. Cell. Biol.* **12**, 2331–2338.
10. Johnson, M. S., McClure, M. A., Feng, D.-F., Gray, J. & Doolittle, R. F. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7648–7652.
11. Katzman, M., Katz, R. A., Skalka, A. M. & Leis, J. (1989) *J. Virol.* **63**, 5319–5327.
12. Roth, M. J., Schwartzberg, P. L. & Goff, S. P. (1989) *Cell* **58**, 47–54.
13. Brown, P. O., Bowerman, B., Varmus, H. & Bishop, J. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2525–2529.
14. Fujiwara, T. & Mizuuchi, K. (1988) *Cell* **54**, 497–504.
15. Davies, J. F., II, Hostomska, Z., Hostomsky, Z., Jordan, S. R. & Matthews, D. A. (1991) *Science* **252**, 88–95.
16. Miller, M., Jaskólski, M., Rao, J. K. M., Leis, J. & Wlodawer, A. (1989) *Nature (London)* **337**, 576–579.
17. Terry, R., Soltis, D., Katzman, M., Cobrinik, D., Leis, J. & Skalka, A. M. (1988) *J. Virol.* **62**, 2358–2365.
18. van Gent, D. C., Elgersma, Y., Bolk, M. W. J., Vink, C. & Plasterk, R. H. A. (1991) *Nucleic Acids Res.* **19**, 3821–3827.
19. Misra, T. K., Grandgenett, D. P. & Parsons, J. T. (1982) *J. Virol.* **44**, 330–343.
20. Harrison, S. C. & Aggarwal, A. K. (1990) *Annu. Rev. Biochem.* **59**, 933–969.
21. Engelman, A., Mizuuchi, K. & Craigie, R. (1991) *Cell* **67**, 1211–1221.
22. Vink, C., Yeheskiely, E., van der Marel, G. A., van Boom, J. H. & Plasterk, R. H. A. (1991) *Nucleic Acids Res.* **19**, 6691–6698.
23. Jones, K. S., Coleman, J., Merkel, G. W., Laue, T. M. & Skalka, A. M. (1992) *J. Biol. Chem.*, in press.