Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 1843–1847, March 1994
Biochemistry

Expression and partial purification of enzymatically active recombinant Ty1 integrase in *Saccharomyces cerevisiae*

SHARON P. MOORE AND DAVID J. GARFINKEL

National Cancer Institute–Frederick Cancer Research and Development Center, Advanced BioScience Laboratories–Basic Research Program, Laboratory of Eukaryotic Gene Expression, P.O. Box B, Building 539, Frederick, MD 21702-1201

Communicated by George F. Vande Woude, November 19, 1993

ABSTRACT **Integration of the *Saccharomyces cerevisiae* retrotransposon Ty1 into the genome requires Ty1 integrase (IN). Apparent functions of Ty1 IN are target-site determination, cleavage, and joining of donor strands. To further study the mechanism of Ty1 integration, an IN expression plasmid has been constructed for use in yeast. The recombinant IN coding sequence differs from mature Ty1 IN associated with Ty1 virus-like particles only in that it has several additional N-terminal amino acid codons. Inclusion of a polyhistidine tag facilitates purification of recombinant IN by metal chelate chromatography. Recombinant Ty1 IN is active in an *in vitro* assay with short double-stranded oligonucleotide substrates and has biochemical properties similar to those observed with Ty1 virus-like particles. The full-length Ty1 IN produced in yeast should be useful for further biochemical, genetic, and structural analyses of Ty1 integration and for comparative analyses with retroviral IN proteins.**

Ty1 is a retrotransposon found in *Saccharomyces cerevisiae*. The Ty1 life cycle resembles that of retroviruses in many respects, except that Ty1 transposition is not infectious (1, 2). The functional and structural organization of the Ty1 genome is also similar to that of retroviruses. Ty1 elements contain two long terminal repeats (LTRs) flanking a central coding region. There are two partially overlapping open reading frames, *TYA1* and *TYB1*, which are analogous to the retroviral *gag* and *pol* genes, respectively. *TYA1* encodes nucleocapsid proteins that assemble as Ty1 virus-like particles (VLPs). *TYB1* encodes the following catalytic proteins: protease, integrase (IN), and reverse transcriptase/RNase H required for protein maturation, integration, and replication, respectively.

During the Ty1 transposition process, intracellular VLPs mature and collect in the cytoplasm (3–5). The maturation of Ty1 proteins is dependent on proteolytic processing of the TYA1 (gag) and TYA1–TYB1 (gag–pol) precursor proteins by Ty1 protease in the maturing VLP (6, 7). The copurification of Ty1 VLPs, Ty1 proteins and nucleic acids (4), a $tRNA^{Met}$ required for priming reverse transcription (8), a reverse transcriptase activity capable of synthesizing Ty1 DNA (4, 5), and an IN activity that catalyzes Ty1 integration events *in vitro* (9, 10) strongly suggests that the Ty1 VLP is an essential transposition intermediate.

Here we describe the expression, partial purification, and initial biochemical characterization of recombinant Ty1 IN. We have synthesized the IN coding region by polymerase chain reaction (PCR) and cloned this segment into a yeast expression vector. Sequences encoding an initiator Met, a polyhistidine tag, and an enterokinase cleavage site have been added to the coding sequence to facilitate expression and purification of Ty1 IN. Recombinant Ty1 IN is active in an *in vitro* assay with short double-stranded oligonucleotide substrates. The recombinant and Ty1 VLP-associated IN proteins show similar biochemical properties.

MATERIALS AND METHODS

Plasmids and Strains. The galactose-inducible expression vector pRDK249 and strain RDKY1293 (11) were kindly provided by R. Kolodner (Harvard Medical School, Boston). The IN coding region was amplified from the transposition-competent element Ty1-H3*HIS3* (12). Strain DG1377 is a transformant of strain RDKY1293 (*MATα ura3-52 trp1 leu2-Δ1 his3-Δ200 pep4::HIS3 prb1-Δ1.6R CAN1^R GAL*), containing the Ty1 IN expression plasmid pGTy1-IN.

Construction of Plasmid pGTy1-IN. The IN coding region of Ty1-H3*HIS3* was amplified by PCR (13) using the high-fidelity thermostable DNA polymerase Vent according to the supplier's specifications (New England Biolabs) and cloned into the *GAL10*-based expression plasmid pRDK249. The 5′ amplification primer contained the following sequences: a *Xho* I cleavage site, six adenosine residues upstream of the initiator ATG and the Ser codon TCT downstream of the ATG, a hexahistidine tag $(CATCAC)_3$ (14), the coding sequence for the enterokinase cleavage site NNNK (GACGACGACGACAAA) (15), and the coding sequence for the N-terminal 14 amino acids of Ty1 IN NVHYSESTRKYPYP [AATGTCCATACAAGTGAAAGTACACGCAAATATCCTTATCCT; the 5′ end of this sequence starts at position 2041 of Ty1-H3 (16)]. The 3′ amplification primer consisted of a *Hin*dIII site, an ochre stop codon, and the coding sequence for the putative C-terminal 6 amino acids of Ty1 IN AILHIR (TGCAATCAGGTGAATTCG; the 5′ end of the complementary strand starts at position 3945 of Ty1-H3). The N termini of Ty1 IN and reverse transcriptase/RNase H was determined by protein microsequencing and will be reported elsewhere. After 20 cycles of amplification, the Ty1 IN coding region and plasmid pRDK249 were digested with *Xho* I and *Hin*dIII, gel-purified, and cloned using standard procedures (17). Recombinant plasmids were introduced into yeast by lithium acetate transformation (18).

Cell Growth. Strain DG1377 was grown and protein expression was induced with galactose essentially as described by Johnson and Kolodner (11). Cells were harvested by centrifugation and washed once with 20 mM Tris·HCl, pH 7.5/150 mM NaCl. The final cell pellet was resuspended in ≈10 ml of washing buffer, frozen in a dry ice-ethanol bath, and stored at −70°C.

Purification of Recombinant Ty1 IN and Ty1 VLPs. Cells were thawed, pelleted, and vigorously resuspended in twice their storage volume with buffer A [20 mM Tris·HCl, pH 7.5/10% (vol/vol) glycerol/10 mM 2-mercaptoethanol/150 mM NaCl/10 mM $NaHSO_3$/2 mM benzamidine/1 mM phenylmethylsulfonyl fluoride/4 μM pepstatin A/2 μM leupep-

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, integrase; VLP, virus-like particle; LTR, long terminal repeat; Ni-NTA, nickel nitrilotriacetate; HIV, human immunodeficiency virus.

tin]. The cell suspension was transferred to Corex (Corning) glass tubes and acid-washed glass beads (425–600 μm in diameter) were added at a ratio of 1:1.3, beads to cell suspension. Cells were lysed by shaking on a Vortex mixer for 15 min, in 1-min bursts, at 0–4°C. The suspension was centrifuged for 1 h at 72,500 × g at 4°C. The pellet was reextracted with buffer A/1 M NaCl and stirred for 2 h at 4°C. The resulting lysate was centrifuged for 1 h at 105,000 × g at 4°C. The supernatant liquid (90 ml) was dialyzed against 2 liters of buffer B (20 mM sodium phosphate, pH 7.5/10% glycerol)/0.8 M NaCl, followed by one change against 2 liters of buffer B/0.65 M NaCl and one change against 2 liters of buffer B/0.5 M NaCl. The dialyzed fraction was batch-bound to 5 ml of nickel nitrilotriacetate (Ni-NTA) resin [Invitrogen or Qiagen (Chatsworth, CA)] with rocking at 4°C overnight. The Ni-NTA resin was washed using buffer B/0.5 M NaCl until the A_{280} was ≤0.01. The Ni-NTA resin was then washed with buffer B/0.5 M NaCl/60 mM imidazole until the A_{280} was ≤0.01, followed by two washes with buffer B/0.5 M NaCl without imidazole. The washed Ni-NTA resin was poured into a C10/10 column (Pharmacia). Recombinant Ty1 IN was eluted in buffer B/0.5 M NaCl with 80 mM imidazole. Peak fractions were concentrated and the buffer was exchanged by Centricon 30 filtration (Amicon). The fractions were stored in buffer C (20 mM Tris·HCl, pH 7.5/10 mM dithiothreitol/1 mM EDTA/10% glycerol/300 mM KCl) at −70°C. Ty1 VLPs were isolated from strain GRY458 (6) by using established methods (9).

Assay for IN Activity. We modified the oligonucleotide integration assay to detect Ty1 IN activity (19, 20). Complementary oligonucleotides representing the terminal 30 nt of either U3 or U5 of the Ty1 H3 LTR were used as model substrates for the integration assay. A heterologous (non-LTR) sequence oligonucleotide was used as a control for substrate specificity (Fig. 1). T4 polynucleotide kinase (New England Biolabs) was used to label the 5′ end of one of the oligonucleotides with [γ-^{32}P]ATP (5000 Ci/mmol; 1 Ci = 37 GBq; Amersham) prior to annealing. The 5′ end chosen for labeling represented sequences that are internal to the end of the LTR in native Ty1 elements. Standard assay conditions consisted of 10 mM Tris·HCl (pH 7.4), 1 mM $MnCl_2$, 0.1 mM dithiothreitol, 50 mM KCl, 1.6% glycerol, 5% (wt/vol) PEG (M_r = 8000), 150 ng of Ty1 IN or 2.4 μg of Ty1 VLPs, and 0.2 pmol of labeled double-stranded substrate in a 20-μl reaction volume. After a 1-h incubation at 30°C, the reaction was stopped by incubating with proteinase K (Boehringer Mannheim) for 15 min at 37°C. The reaction products were boiled in sequencing-gel loading buffer and visualized by autoradiography after electrophoresis through a 20% polyacrylamide/8 M urea gel.

Immunoblot Analysis. Protein samples were examined for recombinant Ty1 IN by immunoblot analysis using anti-TYB1 antiserum B2 as described (6). Cross-reactivity was detected using the ECL detection system (Amersham). Protein concentrations were determined by the dye-binding technique of Bradford (21), using commercially available reagents (Bio-Rad). Coelectrophoresis of protein markers (Bio-Rad) was performed to provide molecular mass standards.

RESULTS

Expression and Purification of Recombinant Ty1 IN. Like retroviral IN proteins, mature Ty1 IN is produced from the 190-kDa TYA1 (gag)–TYB1 (pol) precursor by Ty1 protease-mediated cleavage (6). We used the DNA sequence from the transpositionally competent retrotransposon Ty1 H3 corresponding to the N-terminal and C-terminal residues of mature Ty1 IN protein to synthesize a full-length IN segment by PCR. Several codons were added to the N terminus of the recombinant IN protein coding sequence to enhance expression and simplify purification. To optimize translation initiation in yeast (22), several adenosine residues were placed upstream of the initiator Met codon, and the Ser codon TCT was placed immediately downstream of the Met codon. Six His codons were added to facilitate purification by nickel chelate chromatography. An enterokinase cleavage site was included on the C-terminal side of the polyhistidine tag to provide the option of removing the upstream sequences not present in mature Ty1 IN.

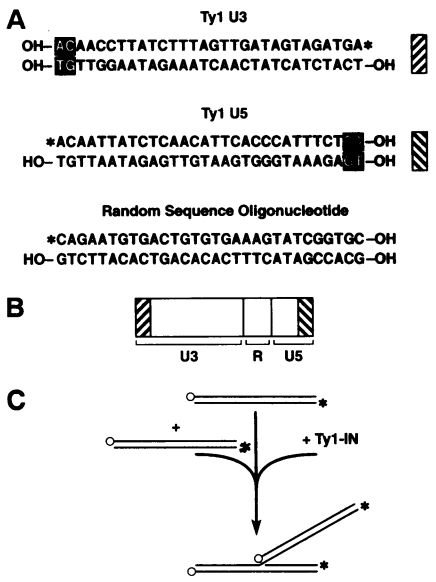

FIG. 1. (*A*) Sequence of the 30-bp oligonucleotides corresponding to the U3 and U5 segments of the Ty1 H3 LTR that were used in the integration assay and of the non-LTR oligonucleotide used as a substrate control. The position of the U3 and U5 oligonucleotides within the LTR is indicated by the hatched boxes. The Ty1 H3 sequence is taken from Boeke *et al.* (16). The conserved terminal dinucleotides are indicated by the darkened boxes. The position of the 5′ ^{32}P label is indicated by asterisks. (*B*) Position of the U3 and U5 oligonucleotides in a complete 338-bp Ty1 LTR. U3, unique 3′ sequence; U5, unique 5′ sequence; R, repeated sequence. (*C*) Diagram of the formation of Ty1-IN-catalyzed products from oligonucleotides shown in *A*. The unlabeled 5′ ends are represented by circles; the ^{32}P-labeled 5′ ends are represented by asterisks. Two types of labeled product are expected when Ty1 IN proteins are incubated with substrates labeled at the appropriate 5′ end with ^{32}P and separated by gel electrophoresis. Labeled products larger than the original 30 nt substrate represent the joining of a donor to a recipient strand. Labeled products shorter than the original 30 nt substrate represent the unjoined portion of the target strand.

Production of recombinant Ty1 IN was initially determined by immunoblot analysis using B2 antiserum, which cross-reacts with Ty1 IN (6) (Fig. 2). Very little Ty1 IN protein was detected (Fig. 2, lane 1) when strain DG1377 was grown in the repressing carbon source glucose (4, 23). When induced with galactose, strain DG1377 produced a recombinant Ty1 IN (Fig. 2, lane 2) close to the predicted molecular mass of the Ty1 VLP-associated IN (lane 3).

We took advantage of two properties of the recombinant Ty1 IN in the protein purification. (*i*) The solubility of unpurified Ty1 IN resembles that of recombinant retroviral IN proteins, which are usually insoluble in NaCl concentrations lower than 0.3–0.5 M (24–26). Therefore, when yeast cells were disrupted in low salt, most of the recombinant Ty1 IN was present in the pellet after centrifugation (Fig. 3). This pellet was resuspended in 1 M NaCl, stirred, and recentrifuged. The high salt treatment solubilized at least half of the recombinant Ty1 IN and centrifugation removed additional cellular debris. (*ii*) Recombinant IN was purified by nickel

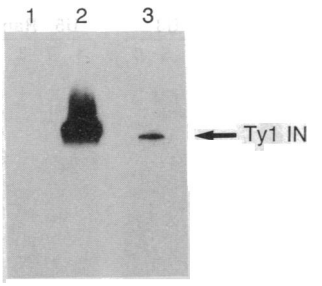

FIG. 2. Induction of recombinant Ty1 IN. Immunoblots were prepared from total protein extracted from strain DG1377 grown in synthetic complete medium with either glucose (lane 1, 10 μg of protein) or galactose (lane 2, 10 μg of protein). Also shown is IN from purified Ty1 VLPs (lane 3). Proteins were separated on an SDS/10% polyacrylamide gel, electroblotted onto an Immobilon membrane, and incubated with anti-TYB1 antiserum B2, which cross-reacts with mature Ty1 IN from VLPs (6). Immune complexes were detected using enhanced chemiluminescence.

chelate chromatography. Recombinant Ty1 IN was eluted from the Ni-NTA column as a sharp peak in 80 mM imidazole. Approximately 70% of the protein in this peak migrated at a position consistent with the molecular mass expected for recombinant Ty1 IN. This method yielded 3.5 μg of IN per g (wet weight) of cells. Based on densitometric scanning of the stained gel, we estimate 170-fold purification from the 1 M NaCl extract to the peak fraction eluted from the Ni-NTA column. Immunoblot analysis of column fractions showed that the majority of the protein reacting with the B2 antiserum migrated as a discrete band and that this band corresponded to the major protein observed by Coomassie blue staining.

Enzymatically Active Recombinant Ty1 IN. To detect recombinant Ty1 IN activity, we tested column fractions for the ability to mediate strand joining using model substrates corresponding to the first 30 nt of the U3 segment of the Ty1 H3 LTR (Fig. 1 *A* and *B*). Although this assay (Fig. 1*C*) does not duplicate the complete integration reaction, it has become

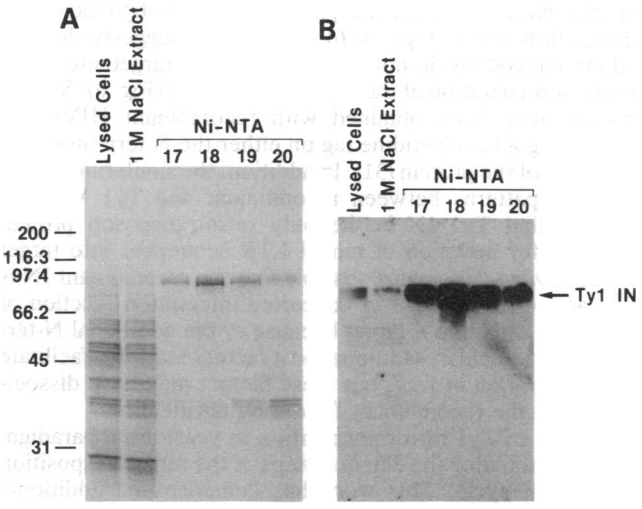

FIG. 3. Purification of recombinant Ty1 IN from strain DG1377. Lanes: lysed cells, crude protein extracts; 1 M NaCl extract, salt extract of the lysed cellular pellet; Ni-NTA 17–20, fractions 17–20 eluted from a Ni-NTA metal-chelate column with 80 mM imidazole that have recombinant Ty1 IN protein and activity. Numbers on the left represent molecular mass estimates in kDa. (*A*) Coomassie blue-stained SDS/10% polyacrylamide gel containing protein extracts from strain DG1377 expressing pGTy1-IN. (*B*) Immunoblot analysis of the gel shown in *A* using antibodies to Ty1 IN protein, as described in Fig. 2.

widely accepted as an indicator for IN activity from a variety of retroviruses (19, 20, 26–31).

IN activity was detected using purified Ty1 VLPs (Fig. 4) or the column fractions containing recombinant Ty1 IN as determined by the oligonucleotide integration assay using the Ty1 U3 oligonucleotide (Fig. 1). The unreacted substrate produced the darkest band (Fig. 4) with bands representing products >30 nt above the unreacted substrate and with bands representing <30 nt below the unreacted substrate. Although the quantity of reaction products obtained with Ty1 VLPs was less than that observed with recombinant Ty1 IN, their banding patterns were identical and reproducible. We did not detect any prominent bands that were 2 nt shorter than the starting substrate (also refer to Figs. 5 and 6). In collateral experiments, recombinant Ty1 IN catalyzed the insertion of a complete linear Ty1 LTR into a heterologous plasmid target. Southern blot hybridization analysis suggested that this product contained one end of the LTR inserted into the plasmid (data not shown), much like the products observed in the oligonucleotide integration assay. We have not detected concerted integration of both ends of a Ty1 LTR substrate using recombinant Ty1 IN (data not shown). A low level of two-ended integration activity has been observed with recombinant human immunodeficiency virus (HIV) IN produced in insect cells (30).

Requirements for Integration of Oligonucleotide Substrates. We assessed two parameters of the oligonucleotide integration assay using recombinant Ty1 IN and Ty1 VLPs. These included the metal cofactor requirement and choice of oligonucleotide substrates. As previously noted, retroviral IN-mediated reactions show a preference for Mn^{2+} over Mg^{2+} and an absolute requirement for a divalent cation (19, 27, 28, 30–33). Therefore, we compared the efficiency of the Ty1-IN-catalyzed reaction with these cations (Fig. 5). A concentration of 1 mM Mn^{2+} (Fig. 5, lanes 1–6) or 1 mM Mg^{2+} (Fig. 5, lanes 7–12) was used in each reaction mixture and recombinant IN (fraction 17; see Fig. 3) was added in increasing amounts (Fig. 5, lanes 1 and 7 contain no IN; lanes 3–6 and lanes 9–12 contain increasing amounts of IN). A constant amount of Ty1 VLPs was also assayed with Mn^{2+} (Fig. 5, lane 2) and Mg^{2+} (Fig. 5, lane 8). Mn^{2+} at 1 mM stimulated the reaction ≈2.5-fold more than 1 mM Mg^{2+} as determined by densitometric scans of the reaction products >30 nt. The number of products <30 nt was also increased in the presence of Mn^{2+} compared to that observed with Mg^{2+}. However, the integration pattern is essentially identical whether Mn^{2+} or Mg^{2+} is used in reactions containing recombinant Ty1 IN or VLPs. There is also an absolute requirement for a cofactor in the reaction, since either omitting the cation or substituting EDTA for the divalent cation resulted in complete inhibition (data not shown).

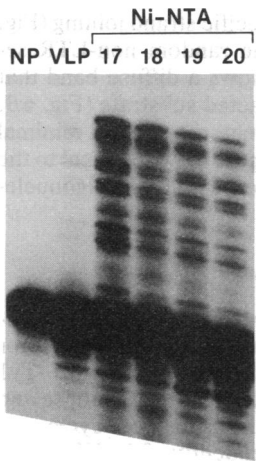

FIG. 4. DNA strand-joining activity using the oligonucleotide integration assay with a ^{32}P-labeled U3 oligonucleotide as described in Fig. 1. Lanes: NP, no protein added to the reaction mixture; VLP, IN activity present in Ty1 VLPs; Ni-NTA 17–20, recombinant Ty1 IN protein in fractions 17–20 from the column described in Fig. 3. Shorter autoradiographic exposures did not reveal any prominent bands that were 2 nt smaller than the unreacted substrate (data not shown).

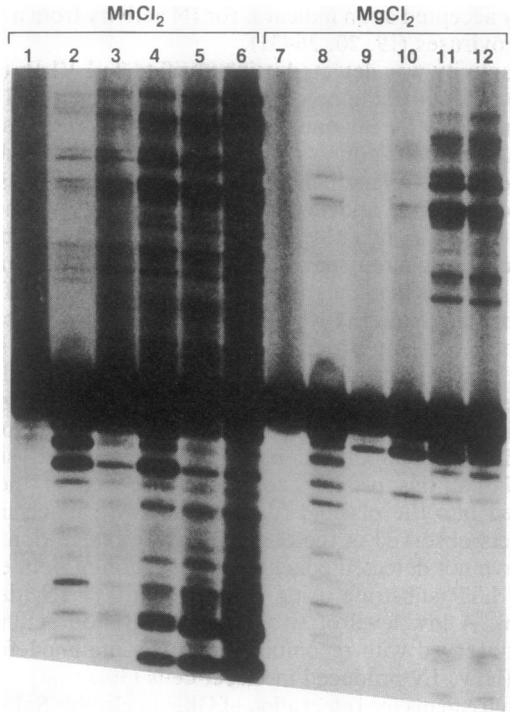

FIG. 5. Metal cofactor requirements of Ty1 IN. Integration assays were performed with either 1 mM $MnCl_2$ or 1 mM $MgCl_2$. Lanes: 1 and 7, no protein; 2 and 8, 2.4 μg of Ty1 VLPs. The remaining lanes are a titration of recombinant Ty1 IN as follows. Lanes: 3 and 9, 22 ng; 4 and 10, 45 ng; 5 and 11, 89 ng; 6 and 12, 147 ng. Shorter exposures did not reveal any prominent bands that were 2 nt smaller than the untreated substrate (data not shown).

Studies of recombinant HIV-1 IN activity *in vitro* have shown that the U5 end of the LTR is usually more active in assays of LTR cleavage and strand joining than the U3 end (26, 28, 30, 31). However, Vink *et al.* (27) have found that U3 and U5 oligonucleotides are equally active in the integration assay with recombinant HIV-1 IN. To characterize the substrate requirements of Ty1 IN by using the oligonucleotide integration assay, we determined whether there was a preference for the U3 or U5 end of the Ty1 LTR (Fig. 6*A*). When Ty1 U3 and U5 oligonucleotides of identical length and specific activity were used in separate integration assays with either Ty1 VLPs (Fig. 6*A*, lanes 2 and 5) or recombinant Ty1 IN (fraction 17; Fig. 6*A*, lanes 3 and 6), the level of U3 and U5 reactivity was similar and dependent on IN (Fig. 6*A*, lanes 1 and 4). The recombinant Ty1 IN activity was also specific for Ty1-H3 oligonucleotide substrates (Fig. 6*B*). In parallel reactions, recombinant Ty1 IN (fraction 18; see Fig. 3) incubated with U5 yielded the expected product consisting of discrete bands suggestive of base-specific strand joining (Fig. 6*B*, lanes 1 and 2). In contrast, the random non-LTR sequence oligonucleotide (Fig. 1*A*) shows a diffuse band that migrates more slowly than the unreacted substrate (Fig. 6*B*, lanes 3 and 4). Although this band may represent a minimal level of strand joining, it does not appear to be identical to the specific integration products observed with LTR oligonucleotides.

DISCUSSION

Integration of the retroviral-like retrotransposon Ty1 requires the IN protein. This protein is apparently needed for target site selection, nicking, and strand joining of Ty1 cDNA. To better understand the role of Ty1 IN in these and possibly other events in the Ty1 transposition cycle, we expressed the protein in yeast independently of the VLP (Fig. 2), developed a procedure to partially purify soluble IN (Fig. 3), showed that the recombinant protein was enzymatically active (Fig. 4), and have started to characterize the properties of this protein (Figs. 5 and 6).

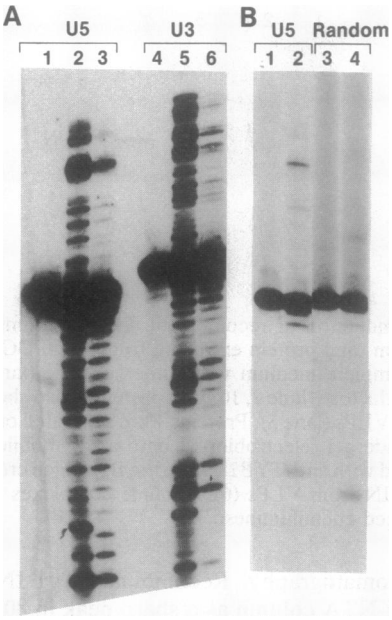

FIG. 6. Relative activities of ^{32}P-labeled oligonucleotides corresponding to U3, U5, and a random sequence oligonucleotide (Fig. 1*A*). (*A*) U3 and U5 oligonucleotides were labeled to the same specific activity. Lanes: 1 and 4, no protein added; 2 and 5, recombinant Ty1 IN; 3 and 6, Ty1 VLPs. Shorter exposures did not reveal any prominent bands that were smaller than the unreacted substrate (data not shown). (*B*) Relative activities of recombinant Ty1 IN for ^{32}P-labeled U5 and a random sequence oligonucleotide. Lanes: 1 and 3, no protein added; 2 and 4, recombinant Ty1 IN.

The recombinant Ty1 IN protein coding sequence contains an initiator Met and 12 additional amino acid codons that are not present in the mature Ty1 IN coding sequence. However, the recombinant and VLP-associated IN proteins produced a similar pattern of reaction products in the oligonucleotide integration assay (Figs. 4–6). This result suggests that the additional codons do not dramatically alter target-site selectivity or recognition of the U3 and U5 donors (Fig. 6). Similar results have been obtained with recombinant HIV-1 IN containing a hexahistidine tag on either the N terminus or C terminus of the protein (31). In addition, the similarities in the reaction patterns between recombinant and Ty1 VLP IN suggest that Ty1 IN is the only retrotransposon protein required for insertion of model LTR sequences into target DNA *in vitro*. It remains possible that the recombinant IN is less active or deficient in concerted integration reaction of two LTR ends into a target because of the additional N-terminal amino acids. Additional host factors may also facilitate Ty1 integration *in vivo*, but these factors may have dissociated from the recombinant IN during purification.

Studies of Ty1 retrotransposition in yeast are a paradigm for understanding the detailed steps of the retrotransposition replication cycle. This work has demonstrated additional similarities and differences between Ty1 and retroviral integration. Recombinant retroviral IN-mediated reactions usually show a preference for Mn^{2+} over Mg^{2+} and an absolute requirement for either cation. Our results indicate that Mn^{2+} also stimulates both recombinant Ty1 IN and VLP-associated IN activity more than does Mg^{2+} by ≈2.5-fold and that either cation is required for IN activity (Fig. 5). Similar integration patterns are also observed with either Mn^{2+} or Mg^{2+}. Furthermore, the moderate stimulation observed with Mn^{2+} in our experiments is comparable to the results ob-

tained with *in vitro* integration of endogenous Ty1 DNA carried out by VLPs (9).

In most studies using HIV-1 IN, the U5 end of the LTR is more active than the U3 end in the oligonucleotide integration assay (26, 28, 31). However, Vink *et al.* (27) find U3 and U5 oligonucleotides to be equally active in the integration assay with recombinant HIV-1 IN. Our results also suggest that Ty1 IN utilizes either U3 and U5 substrates for integration with equal efficiencies (Fig. 6). Sherman *et al.* (34) have shown that the terminal dinucleotide of U5 is cleaved more efficiently than that of U3 with purified HIV-1 IN. If Ty1 IN does not require this activity prior to strand joining, the rate-limiting step that determines the suitability of the substrates is bypassed. Although the U3 and U5 integration patterns are different, the U3 and U5 integration patterns are substrate-specific since the same patterns were observed with Ty1 VLPs or with recombinant IN. The oligonucleotide-specific integration patterns observed with U3 and U5 substrates may reflect target-site preferences.

Two major differences exist between retroviral and Ty1 integration. (*i*) Ty1 integration apparently does not require the removal of sequences from the 3′ ends of the linear cDNA (10). Our demonstration of integration of flush-ended oligonucleotide substrates by recombinant Ty1 IN supports this observation. We did not observe any products that can be correlated to the specific 3′ processing of the U3 (Fig. 4)- or U5 (Fig. 6*A*)-labeled substrates. However, the substrates used in our reactions do not contain any additional nucleotides that could be processed without altering the essential (10) highly conserved U3 and U5 terminal dinucleotide (Fig. 1*A*). Furthermore, adding nucleotides to the end of the U5 substrate so that it would more closely resemble the unprocessed retroviral U5 resulted in molecules that were poor substrates for Ty1 IN (data not shown). (*ii*) Ty1 IN has an extended C-terminal domain of unknown function that contains almost half the coding sequence when compared to that of retroviral IN proteins (J. Mack and D.J.G., unpublished results). In retroviruses, the function of this IN domain is also poorly understood, although it is required for retroviral IN activity and can bind DNA nonspecifically (35–38).

Progress on Ty1 integration will now depend upon a more detailed understanding of the biochemical and structural properties of Ty1 IN. The expression and purification of soluble enzymatically active Ty1 IN described here should greatly facilitate such an approach. Further purification and characterization of recombinant Ty1 IN will permit structural studies as well as identify possible host cofactors required for IN activity. Since we have chosen to express Ty1 IN in yeast, these hypothesized host factors may copurify with IN or may be identified genetically.

We are especially grateful to M. Powers (Program Resources, Inc.) for oligonucleotide synthesis; to D. Higgins for advice on Ni-NTA chromatography; to D. L. Court, E. Barsov, T. J. Burkett, D. V. Nissley, L. Rinckel, and J. N. Strathern for helpful comments on the manuscript; and to R. Frederickson and J. Hopkins for help in preparing the figures and manuscript, respectively. This research was sponsored by the National Cancer Institute, Department of Health and Human Services under Contract NO1-CO-74101 with Advanced BioScience Laboratories.

1. Boeke, J. D. & Sandmeyer, S. B. (1991) in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis, and Energetics*, eds. Pringle, J., Jones, E. & Broach, J. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 1, pp. 193–261.
2. Garfinkel, D. (1992) in *The Retroviridae*, ed. Levy, J. A. (Plenum, New York), Vol. 1, pp. 107–158.
3. Adams, S. E., Mellor, J., Gull, K., Sim, R. B., Tuite, M. F., Kingsman, S. M. & Kingsman, A. J. (1987) *Cell* **49**, 111–119.
4. Garfinkel, D. J., Boeke, J. D. & Fink, G. R. (1985) *Cell* **42**, 507–517.
5. Mellor, J., Malim, M. H., Gull, K., Tuite, M. F., McCready, S. M., Dibbayawan, T., Kingsman, S. M. & Kingsman, A. J. (1985) *Nature (London)* **318**, 583–586.
6. Garfinkel, D. J., Hedge, A.-M., Youngren, S. D. & Copeland, T. D. (1991) *J. Virol.* **65**, 4573–4581.
7. Youngren, S. D., Boeke, J. D., Sanders, N. J. & Garfinkel, D. J. (1988) *Mol. Cell. Biol.* **8**, 1421–1431.
8. Chapman, K. B., Bystrom, A. S. & Boeke, J. D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3236–3240.
9. Eichinger, D. J. & Boeke, J. D. (1988) *Cell* **54**, 955–966.
10. Eichinger, D. J. & Boeke, J. D. (1990) *Genes Dev.* **4**, 324–330.
11. Johnson, A. & Kolodner, R. (1991) *J. Biol. Chem.* **266**, 14046–14054.
12. Garfinkel, D. J., Mastrangelo, M. F., Sanders, N. J., Shafer, B. K. & Strathern, J. N. (1988) *Genetics* **120**, 95–108.
13. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51**, 263–273.
14. Hoffmann, A. & Roeder, R. G. (1991) *Nucleic Acids Res.* **19**, 6337–6338.
15. Maroux, S., Baratti, J. & Desnuelle, P. (1971) *J. Biol. Chem.* **246**, 5031–5039.
16. Boeke, J. D., Eichinger, D., Castrillon, D. & Fink, G. R. (1988) *Mol. Cell. Biol.* **8**, 1431–1442.
17. Sambrook, J., Fritsch, E. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
18. Gietz, D., St. Jean, A., Woods, R. & Schiestl, R. (1992) *Nucleic Acids Res.* **20**, 1425.
19. Craigie, R., Fujiwara, T. & Bushman, F. (1990) *Cell* **62**, 829–837.
20. Katz, R. A., Merkel, G., Kulkosky, J., Leis, J. & Skalka, A. M. (1990) *Cell* **63**, 87–95.
21. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254.
22. Hamilton, R., Watanabe, C. & deBoer, H. (1987) *Nucleic Acids Res.* **15**, 3581–3593.
23. Curcio, M. J. & Garfinkel, D. J. (1992) *Mol. Cell. Biol.* **12**, 2813–2825.
24. Terry, R., Soltis, D. A., Katzman, M., Cobrinik, D., Leis, J. & Skalka, A. M. (1988) *J. Virol.* **62**, 2358–2365.
25. Sherman, P. & Fyfe, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5119–5123.
26. Leavitt, A., Rose, R. & Varmus, H. (1992) *J. Virol.* **66**, 2359–2368.
27. Vink, C., van Gent, D., Elgersma, Y. & Plasterk, R. (1991) *J. Virol.* **65**, 4636–4644.
28. LaFemina, R., Callahan, P. & Cordingley, M. (1991) *J. Virol.* **65**, 5624–5630.
29. Craigie, R., Mizuuchi, K., Bushman, F. D. & Engelman, A. (1991) *Nucleic Acids Res.* **19**, 2729–2734.
30. Bushman, F. D. & Craigie, R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1339–1343.
31. Drelich, M., Wilhelm, R. & Mous, J. (1992) *Virology* **188**, 459–468.
32. van Gent, D., Elgersma, Y., Bolk, M., Vink, C. & Plasterk, R. (1991) *Nucleic Acids Res.* **19**, 3821–3827.
33. Katzman, M., Katz, R. A., Skalka, A. M. & Leis, J. (1989) *J. Virol.* **63**, 5319–5327.
34. Sherman, P., Dickson, M. & Fyfe, J. (1992) *J. Virol.* **66**, 3593–3601.
35. Woerner, A. M., Klutch, M., Levin, J. G. & Marcus-Sekura, C. J. (1992) *AIDS Res. Hum. Retroviruses* **8**, 297–304.
36. Mumm, S. R. & Grandgenett, D. P. (1991) *J. Virol.* **65**, 1160–1167.
37. Vink, C., Groeneger, A. M. O. & Plasterk, H. A. (1993) *Nucleic Acids Res.* **21**, 1419–1425.
38. Woerner, A. M. & Marcus-Sekura, C. J. (1993) *Nucleic Acids Res.* **21**, 3507–3511.