SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Proteins. TFP was expressed in bacteria Rosetta (DE3) pLysS (EMD Biosciences. San Diego, CA) and purified essentially as described previously (1). TFP was first purified through a His60 Ni Superflow column (Clontech. Mountain View, CA). The 200 mM imidazole eluate containing 500 mM NaCl was diluted in Buffer H [40 mM Na-HEPES, pH 7.8, 0.2 mM EDTA, 10 mM 2-mercaptoethanol, 20% (vol/vol) glycerol, 0.5 mM PMSF] to achieve 200 mM NaCl. This preparation was loaded onto a 2-mL DEAE Sephadex A-25 column (GE Healthcare), which was washed with 10 mL buffer H with 200 mM NaCl. Flow through and wash fractions were combined and loaded onto a 2-mL SP Sepharose column (GE Healthcare). After the SP column was washed with 5 bed volumes of buffer H with 200 mM NaCl, TFP was eluted with buffer H with 500 mM NaCl.

The IN-coding region from pET3a-IN (2) was subcloned into pET29a (EMD Biosciences, San Diego, CA) under control of the *lac* promoter and tagged at the carboxyl-terminal end with a 6xHis tag, generating pKN2412 (K. Nguyen and S.B. Sandmeyer, unpublished data). Mutations were introduced sequentially using primer pairs XQ2194/XQ2195 and XQ2196/XQ2197 to create mutant (D225E/E261D) with an inactive catalytic site (pXQ3134). PKN2412 was transformed into Rosetta (DE3) pLysS. Transformants were grown at 37 ^{o}C to OD_{600}=0.6, then grown for at 24^{o}C for additional 3 h in the presence of 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cell lysate in buffer IN [20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM dithiothreitol (DTT), 1 mM PMSF] was first purified through a His60 Ni Superflow column, and then the fraction eluted with buffer IN supplemented with 200mM imidazole was loaded directly onto a 2-mL DEAE Sephadex A-25 column pre-equilibrated with the same buffer. The column was washed with 5 bed volumes of buffer IN and eluted with buffer IN containing 500 mM NaCl. Purified protein was dialyzed in storage buffer [20 mM Tris-HCl, pH8.0, 500 mM NaCl, 1 mM DTT, 20% glycerol].

Plasmids. Plasmids pLY1855 (3), pXQ2889 (+IR) and pXQ3673 (-IR) were targets for Ty3 integration *in vitro*. Ty3-1 on a genomic *Eco*RI/*Hin*dIII fragment in plasmid pEUTy3-1 (4) was digested with *Xho*I which cleaves within both LTRs 14 nts from the downstream end, religated, subcloned into pBluescript creating pLY1386 containing a single composite LTR (L. Yieh and S.B. Sandmeyer, unpublished data). Plasmid pLY1386 was digested with *Xho*I releasing a fragment containing flanking sequence and a single LTR truncated at position 326 of the 340-bp Ty3 LTR. This fragment was cloned into pET29a creating pXQ2889. Plasmid pXQ3673 was generated by site-directed mutagenesis of the Ty3 LTR inverted repeat in pXQ2889 using primers XQ3480 and XQ3481. Positive PCR control plasmids containing Ty3 LTR insertions upstream of the *SNR6* gene (pDLC370 and pLY1842) were described previously ((5) for pDLC370, (3) for pLY1842) or constructed similarly in pCR2.1 (Life Technologies, Inc., Carlsbad, CA). PXQ3659 which contains an insertion of the downstream U5 end of the Ty3 LTR at position bp -1 relative to the l-U6 TATA transcription start site was constructed by amplifying a fragment generated with primers HH1707 and 242 (see below) from template pLY1855 and cloning it into pCR2.1. Plasmids pXQ3660 and pXQ3661 were constructed similarly and contained PCR fragments representing Ty3 LTRs with U5 ends at -5 and -226 relative to the r-U6 TSS, respectively. The pXQ3691 plasmid containing insertion of U5 DNA substrate at 6 bp upstream of LTR IR containing a fragment generated with primers HH1707 and XQ2876 and pXQ2889.

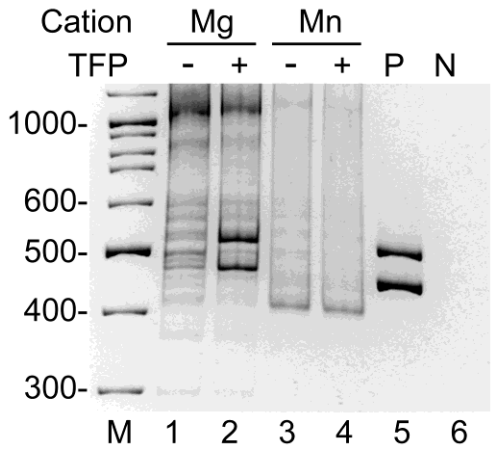
Protein DNA binding assays. Oligos XQ1800 and XQ1801 were annealed to form duplex oligonucleotides with an overhanging end (5'-GTAT), which was filled in by polymerization in the presence of [^{32}P]dATP. Reactions included 800 ng of duplex DNA, 2.5 µL [α^{32}P]dATP (25 µCi) [α^{32}P]dATP (3,000 Ci/mmol, PerkinElmer), 0.25 mM dGTP, dCTP and dTTP, and 5 units of Klenow (New England Biolabs) in 1x Buffer in a total volume of 50 µl. Reactions were incubated at 37^{o}C for 30 min. The reaction product was cleaned on Micro Bio-Spin 30 columns (BioRad Inc. Richmond, CA) according to manufacturer's manual. As previously described (6), protein-DNA complexes were formed for 30 min at 23°C in 19 µl of reaction buffer [40 mM Tris pH 8, 70 mM NaCl, 7 mM $MgCl_2$, 3 mM DTT, 100 µg/ml bovine serum albumin (BSA), 5 µg/ml poly(dG-dC)·poly(dG-dC), 6-8% (v/v) glycerol]

containing 6 fmol of DNA and 250 fmol of TFP. Increasing concentrations of IN were added in a 1 µl volume and incubated for an additional 20 min. Samples were incubated with 20 µg/ml poly(dA-dT)·poly(dA-dT) nonspecific competitor for 5 min before electrophoresis on 4% polyacrylamide gel [20 mM Tris pH 8.0, 0.5 mM EDTA, 2.5 mM $MgCl_2$, 0.5 mM DTT, 4% glycerol]. Running buffer was the same as above except that DTT and glycerol were omitted. Probe bands were detected by phosphorimaging with a Bio-Rad Molecular FX Imaging system (Bio-Rad, Richmond, CA).

In vitro integration and strand transfer assays. *In vitro* integration using VLPs was performed as described previously (7). Either TFIIIB or TFP were mixed with target plasmids on ice for 30 min before VLPs were added, and samples were incubated at 16 ^{o}C for 15 min. DNA substrates for the strand-transfer assay were prepared by annealing equimolar amounts of oligonucleotides (e.g. plus strand, U5 end substrate with underlined primer-complementary sequence 5'GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCGTAATACAACACC3'/ GGTGTTGTATTACGGGCTCGAGGGTTTGGACCTTTGAGCCAATGC) (Supplemental Table 2) in STE buffer (10 mM Tris, pH7.5, 1 mM EDTA, 50 mM NaCl). Strand-transfer reactions were performed in buffer R (20 mM HEPES pH7.5, 70 mM NaCl, 0.1% NP-40, 7.5% DMSO, 5 mM DTT) supplied with various amount of $MgCl_2$ or $MnCl_2$ as cofactors. Generally, samples contained 50 fmole of target plasmids, 250 fmole of duplex DNA, 250 fmole of TFP, and 1000 fmole of IN in a total volume of 40 µL. Reactions were incubated at 24 ^{o}C for 1 h, and DNA products were extracted as described previously (7). Reactions were incubated at 24 ^{o}C and were stopped after 1 h with the addition of SDS and EDTA, pH 8.0, to final concentrations of 0.2% and 20 mM, respectively. Reaction products were extracted with phenol/chloroform, and DNA was ethanol precipitated and resuspended in H_2O.

PCR was performed to amplify fragments diagnostic of strand transfer. One tenth of the DNA products were combined with primers 242 and 411, which anneal within the *SNR6* gene and at the downstream end of the internal domain of Ty3, respectively (3). PCR was performed for 40 cycles (94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s) in a 20-µl reaction containing 0.5 units of Choice Taq polymerase (Denville Scientific, Metuchen, NJ), 0.3 mM dNTPs, and 0.3 µM primers. In the PCR reactions monitoring strand-transfer products of duplex DNA substrates, primer HH1707, which anneals at the first half of the DNA substrates, was substituted for primer 411 but conditions were otherwise similar. Alternatively, primer M13R was substituted for primer 242 to map the strand-transfer sites on the other strand of pLY1855. Four nanograms of positive control plasmids (Supplemental Table 1) were used as templates in control reactions. As a control for DNA recovery, primers 679 and 680 (3) were used to amplify the β-lactamase gene carried by the target plasmid pLY1855. Primers XQ2876 and XQ2877 were used to monitor DNA recovery of target plasmids pXQ2889 and pXQ3673. PCR conditions were same as above except that the reactions were carried out for 25 cycles. PCR products were resolved by electrophoresis on non-denaturing 8% polyacrylamide gel or 1.5% agarose gel and visualized by staining with ethidium bromide. In order to determine strand transfer sites, DNA fragments were extracted from the gel, cloned into pCR2.1 and sequenced (Genewiz, Inc., La Jolla, CA).

SUPPLEMENTAL FIGURE 1. *In vitro* strand transfer mediated by Ty3 VLPs using Mg^{2+} or Mn^{2+} as metal cofactor. Reactions contained VLPs as a source of Ty3 cDNA and IN, TFP ($Brf1_{1-382}$-TBP_{61-240}-$Brf1_{439-596}$), and target plasmid pLY1855. As described in Fig. 1 and Experimental Procedures, strand-transfer products were detected in PCR primed from within Ty3 and *SNR6* (primers 411 and 242, respectively). Either 10 mM $MgCl_2$ (lanes 1 and 2) or 10 mM $MnCl_2$ (lanes 3 and 4) was included in the reactions as metal cofactor, with or without TFP as indicated. Strand-transfer reactions with test reactants (lanes 1-4), positive control (P) containing plasmids with a Ty3 LTR at leftward (pLY1842) or rightward TSS (pDLC370) (lane 5), or negative control (N) containing water alone (lane 6) were used to template PCR using primers 411 and 242. PCR products were separated on a nondenaturing 8% polyacrylamide gel.



SUPPLEMENTAL TABLE 1. Plasmids used in this study.

Plasmids	Description	Reference
pLY1855	*SNR6* target plasmid, bidirectional TATA box	(3)
pET21d-Brf1n-TBPc-Brf1c	Brf1n-TBPc-Brf1c triple-fusion protein (TFP) with C-terminal His-tag	(1)
pET3a-IN	pET3a, recoded Ty3 IN (CODA)	(2)
pET29a	C-terminal His-tag expression vector	EMD Biosciences (San Diego, CA)
pKN2412	pET29a, Ty3 IN from pET3a-IN, C-terminal His-tag	Nguyen and Sandmeyer, unpublished data
pXQ3134	pKN2412, IN D225E/E261D double mutation	This study
pLY1386	pBluescript, Ty3 LTR on *Eco*R I/*Hin*d III fragment	Yieh and Sandmeyer, unpublished data
pXQ2889	pET29a, truncated Ty3 LTR, *Xho* I fragment from pLY1386	This study
pXQ3673	pXQ2889, IR mutant	This study
pLY1842	Cloned 411/242 amplicon of Ty3 l-*SNR6* insertion	(3)
pDLC370	Cloned 411/242 amplicon of Ty3 r-*SNR6* insertion	(8)
pCR2.1	T/A cloning vector	Life Technologies, Inc. (Carlsbad, CA)
pXQ3659	pCR2.1, HH1707/242 amplified l-*SNR6* insertion	This study
pXQ3660	pCR2.1, HH1707/242 amplified r-*SNR6* insertion	This study
pXQ3661	pCR2.1, HH1707/242 amplified -226 IR insertion (Mn^{2+}) in pLY1855, control	This study
pXQ3691	pCR2.1, HH1707/XQ2876 amplified -6 IR insertion (Mn^{2+}) in pXQ2889, control	This study

SUPPLEMENTAL TABLE 2. Oligonucleotides used in this study.

Oligo	Sequence (5'-3')	Description
411	CGAAACACAAGACAACCC	Forward primer, anneals to Ty3 DNA
242	GGAACTGCTGATCATCTCT	Reverse primer, anneals to pLY1855 (*SNR6* at +67)
M13R	CAGGAAACAGCTATGA	Alternative primer, anneals to pLY1855
679	ACTCCCCGTCGTGTAGATAACTACG	Forward primer, control PCR
680	AAGCCATACCAAACGACGAGC	Reverse primer, control PCR
HH1704	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAACA	Plus-strand oligo for substrate mimicking pre-processed Ty3 U5 LTR terminus, underlined region same as HH1707
HH1705	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAACACC	Plus-strand oligo for substrate mimicking non-processed Ty3 U5 LTR terminus
HH1706	GGTGTTGTATTACGGGCTCGAGGGTTTGGACCT TTGAGCCAATGC	Minus-strand oligo for substrate mimicking Ty3 U5 LTR terminus
HH1707	GCATTGGCTCAAAGGTCCAAACC	Primer, anneals to substrate DNA
HH1854	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAATA	Plus-strand oligo for pre-processed U5 LTR substrate, CA to TA mutant
HH1855	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAATACC	Plus-strand oligo for non-processed U5 LTR substrate, CA to TA mutant
HH1856	GGTATTGTATTACGGGCTCGAGGGTTTGGACCT TTGAGCCAATGC	Minus-strand oligo for U5 LTR substrate, CA to TA mutant
HH1857	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAACG	Plus-strand oligo for pre-processed U5 LTR substrate, CA to CG mutant
HH1858	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAACGCC	Plus-strand oligo for non-processed U5 LTR substrate, CA to CG mutant
HH1859	GGCGTTGTATTACGGGCTCGAGGGTTTGGACCT TTGAGCCAATGC	Minus-strand oligo for U5 LTR substrate, CA to CG mutant
HH1860	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAATG	Plus-strand oligo for pre-processed U5 LTR substrate, CA to TG mutant
HH1861	GCATTGGCTCAAAGGTCCAAACCCTCGAGCCCG TAATACAATGCC	Plus-strand oligo for non-processed U5 LTR substrate, CA to TG mutant
HH1862	GGCATTGTATTACGGGCTCGAGGGTTTGGACCT TTGAGCCAATGC	Minus-strand oligo for U5 LTR substrate, CA to TG mutant
HH1798	GCATTGGCTCAAAGGTCCAAACCGCATTGGCTC AAAGGTCCAAAC	Plus-strand oligo for substrate with random sequence
HH1799	GTTTGGACCTTTGAGCCAATGCGGTTTGGACCT TTGAGCCAATGC	Minus-strand oligo for substrate with random sequence
XQ2194	CGCGTACCATCACCTCTGAgCGTGACGTTCGTAT GACTGC	IN D225E mutation forward primer
XQ2195	GCAGTCATACGAACGTCACGcTCAGAGGTGATG GTACGCG	IN D225E mutation reverse primer
XQ2196	AGACTGATGGTCAATCCGAcCGTACCATTCAGA CCCTGAA	IN D261E mutation forward primer
XQ2197	TTCAGGGTCTGAATGGTACGGTCGGATTGACCA TCAGTCT	IN D261E mutation reverse primer
XQ1800	GTATTTCGTCCACTATTTTCGGCTACTATATATA AATGTTTTTTTCGCAACTATGTG	Plus-strand oligo for TATA box DNA probe (1)

Oligo	Sequence (5'-3')	Description
XQ1801	CACATAGTTGCGAAAAAAACATTTATATATAGT AGCCGAAAATAGTGGACGAA	Minus-strand oligo for TATA box DNA probe (1)
XQ3480	TGACATATCTCATTTTGAGATACGTGACTTTCA GAAAATCAAGCTCGTAT	LTR IR mutation forward
XQ3481	ATACGAGCTTGATTTTCTGAAAGTCACGTATCT CAAAATGAGATATGTCA	LTR IR mutation reverse
XQ2876	CAGTCTTAGGGAAGTAACCAGTTGTC	Primer, anneals to Ty3 LTR
XQ3135	GGGTGACGTATTGTCATACTG	Primer, anneals to Ty3 LTR

REFERENCES

1. Kassavetis, G. A., Soragni, E., Driscoll, R., and Geiduschek, E. P. (2005) *Proc Natl Acad Sci U S A* **102**, 15406-15411
2. Larsen, L. S., Wassman, C. D., Hatfield, G. W., and Lathrop, R. H. (2008) *Int J Bioinform Res Appl* **4**, 324-336
3. Yieh, L., Kassavetis, G., Geiduschek, E. P., and Sandmeyer, S. B. (2000) *J Biol Chem* **275**, 29800-29807
4. Kinsey, P. T., and Sandmeyer, S. B. (1995) *Genetics* **139**, 81-94
5. Chalker, D. L., and Sandmeyer, S. B. (1992) *Genes Dev* **6**, 117-128
6. Kassavetis, G. A., Letts, G. A., and Geiduschek, E. P. (2001) *EMBO J* **20**, 2823-2834
7. Yieh, L., Hatzis, H., Kassavetis, G., and Sandmeyer, S. B. (2002) *J Biol Chem* **277**, 25920-25928
8. Chalker, D. L., and Sandmeyer, S. B. (1993) *Proc Natl Acad Sci U S A* **90**, 4927-4931