

Targeting of the Yeast Ty5 Retrotransposon to Silent Chromatin Is Mediated by Interactions between Integrase and Sir4p†

WEIWU XIE,¹ XIAOWU GAI,¹ YUNXIA ZHU,¹ DAVID C. ZAPPULLA,²
ROLF STERNGLANZ,² AND DANIEL F. VOYTAS^{1*}

Department of Zoology and Genetics, Iowa State University, Ames, Iowa 50011-3260,¹ and Department of Biochemistry and Cell Biology, State University of New York, Stony Brook, New York 11794-5215²

Received 12 March 2001/Returned for modification 8 May 2001/Accepted 30 June 2001

The Ty5 retrotransposons of *Saccharomyces cerevisiae* integrate preferentially into regions of silent chromatin at the telomeres and silent mating loci (*HMR* and *HML*). We define a Ty5-encoded targeting domain that spans 6 amino acid residues near the C terminus of integrase (LXSSXP). The targeting domain establishes silent chromatin when it is tethered to a weakened *HMR*-E silencer, and it disrupts telomeric silencing when it is overexpressed. As determined by both yeast two-hybrid and in vitro binding assays, the targeting domain interacts with the C terminus of Sir4p, a structural component of silent chromatin. This interaction is abrogated by mutations in the targeting domain that disrupt integration into silent chromatin, suggesting that recognition of Sir4p by the targeting domain is the primary determinant in Ty5 target specificity.

The long terminal repeat (LTR) retrotransposons are a large and ubiquitous class of mobile genetic elements. Like their cousins the retroviruses, they replicate by reverse transcribing an element mRNA and then integrating the cDNA product into their host's chromosomes. LTR retrotransposons are typically abundant components of nuclear genomes, constituting a few percentages of the *Saccharomyces cerevisiae* genome to over 50% of the genomes of some plants such as maize (31, 45). As the genome sequencing projects progress, it is apparent that most retrotransposons are not randomly distributed on chromosomes. In *Drosophila melanogaster* and *Arabidopsis thaliana*, for example, retrotransposons are highly enriched in pericentromeric heterochromatin (19, 44). This nonrandom distribution may be the result of preferential integration to these sites. It has been suggested that the low gene density of heterochromatin may offer a safe haven for transposition, which ensures persistence of retrotransposons by avoiding the harmful consequences of mutations that might occur if integration were random (5). Because repetitive sequences can form heterochromatin in some species, the accumulation of retrotransposons in certain regions of the genome may, in turn, contribute to the formation of chromatin domains (18).

How is it that retrotransposons identify certain chromosomal regions during integration? One model suggests that the integration apparatus recognizes specific chromatin states or DNA-bound protein complexes (7). This interaction tethers the integration machinery to target sites and results in the observed target site biases. This model is best supported by studies of the *S. cerevisiae* retrotransposons. Over 90% of native Ty1, Ty2, Ty3, and Ty4 insertions are located upstream of genes transcribed by RNA polymerase III (RNAP III) (31).

These regions are often gene poor and, like heterochromatin, may provide a safe haven for transposition within the streamlined *S. cerevisiae* genome (5). For Ty1 and Ty3, the association with sites of RNAP III transcription is due to targeted integration. Targeting requires assembly of the RNAP III transcription complex, and promoter mutations in target genes that prevent transcription complex assembly render them inefficient targets (8, 9, 13). In vitro targeted-transposition assays have been developed for Ty3 in which binding of TFIIB and TFIIC to tRNA gene templates is sufficient for targeting (32). The critical factors within these complexes appear to be the TATA binding protein and Brf (also called TFIIB70) (50). These data support the model that targeting results when the retrotransposon preintegration complex recognizes specific DNA-bound proteins.

In contrast to the other *S. cerevisiae* retrotransposons, native Ty5 elements are not located at sites of RNAP III transcription. Rather, like retrotransposons in many other organisms, Ty5 insertions are predominantly found within the heterochromatin-like domains of the *S. cerevisiae* genome, such as at the telomeres and silent mating loci (*HMR* and *HML*) (57). The chromatin at these sites is referred to as silent chromatin, because it represses transcription of genes located in these regions. Silent chromatin is made up of a large number of proteins that assemble at specific DNA sequences (reviewed in references 35 and 36). Several proteins or protein complexes bind the E and I silencers that flank *HMR* and *HML*, including the origin recognition complex (ORC), the transcription factor Abf1, and the repressor activator protein Rap1p. These proteins also bind to sequences near the telomeres: ORC and Abf1 bind to the subtelomeric X repeat, and Rap1p binds to the telomeric repeat sequences (TG₁₋₃). These DNA-bound proteins recruit additional components of silent chromatin, including the well-studied Sir proteins. Sir2p is a histone deacetylase (21, 33, 47), and Sir3p and Sir4p are considered structural components of silent chromatin. All three proteins interact with each other, and they nucleate at the silencers and spread outward along the chromosome (17, 40).

* Corresponding author. Mailing address: 2208 Molecular Biology Bldg., Iowa State University, Ames, IA 50011. Phone: (515) 294-1963. Fax: (515) 294-7155. E-mail: voytas@iastate.edu.

† This is Journal Paper J-19154 of the Iowa Agriculture and Home Economics Experiment Station, Ames, project 3383.

Ty5 integrates preferentially into regions of silent chromatin. Over 95% of de novo Ty5 transposition events occur within a 3-kb window on either side of the *HMR* silencers or the subtelomeric X repeat (54, 55). Silent chromatin is required for this target choice, because mutations in *HMR-E* that prevent its assembly abolish targeting to this locus (56). Targeting decreases by approximately 50% in *sir2Δ* strains, whereas it is virtually abolished in *sir3Δ* and *sir4Δ* strains (53). An allele of *SIR4* (*sir4-42*) causes a dramatic change in the chromosomal distribution of the Sir complex (29). In *sir4-42* strains, Sir3p and Sir4p move from the telomeres and silent mating loci to the ribosomal DNA (rDNA) (30). This change in Sir protein distribution is related to mother cell aging, and *sir4-42* strains are long-lived (29). Ty5 target specificity changes with the chromosomal distribution of the Sir complex in *sir4-42* strains, and over 25% of the insertions occur within the rDNA (53). These results suggest that the Sir complex, particularly Sir3p and Sir4p, determines Ty5 target choice.

Ty5-encoded proteins are also important for target site selection. A Ty5 missense mutation decreases targeting more than 20-fold and provides the first direct evidence that retroelements encode their own targeting determinants (14). In this paper, we further define Ty5-encoded factors required for targeting and describe a short targeting domain (TD) near the integrase C terminus (INC). This TD interacts with silent chromatin because, when tethered to a defective *HMR-E* silencer, reporter genes at *HMR* are transcriptionally silenced in a Sir-dependent fashion. Overexpression of the TD disrupts telomeric silencing, likely by titrating away critical silencing components. We show that the TD interacts with Sir4p in both yeast two-hybrid and in vitro binding assays. Sir4p, therefore, appears to be the primary host determinant mediating Ty5 target specificity, and the interaction between the Ty5-encoded TD and Sir4p appears to determine target choice.

MATERIALS AND METHODS

Mutagenesis of Ty5 elements. A *BspEI-PflMI* fragment of Ty5 (Fig. 1) was mutagenized by PCR, using two different protocols to minimize mutation biases. The first used the nucleoside triphosphate analogue dPTP, which pairs with both A and G and thereby increases the mutation spectrum (51). A typical 20- μ l reaction mixture included 5 ng of template DNA (pXW27, a plasmid containing the *BspEI-PflMI* fragment), 2.5 U of *Taq* polymerase, 2 μ l of 10 \times buffer, 0.5 μ l of the universal and reverse primers (20 pM), 1.6 μ l of 25 mM MgCl₂, 4 μ l of deoxynucleoside triphosphates (dNTPs) (2.5 mM each), and 2.5 μ l of dPTP (400 μ M). The PCR was carried out for five cycles as follows: 92°C for 1 min, 50°C for 1.5 min, and 72°C for 5 min. An aliquot of the reaction mixture (0.5 μ l) was then used for PCR amplification without dPTP (3). The second mutagenesis method used Mn²⁺ (rather than Mg²⁺) and biased amounts of dNTPs (46). A typical 50- μ l reaction mixture included 5 ng of template DNA, 2.5 U of *Taq* polymerase, 5 μ l of 10 \times buffer, 0.5 μ l of each primer (20 pM), 14 μ l of 25 mM MgCl₂, 0.75 μ l of 10 mM MnCl₂, 4 μ l of dNTPs (2.5 mM each), 4 μ l of dCTP (10 mM), and 4 μ l of dTTP (10 mM). The PCR was carried out for 13 cycles as follows: 94°C for 30 s, 50°C for 45 s, and 72°C for 3 min. An aliquot of the reaction mixture (0.5 μ l) was used as a template in a standard PCR amplification (3). The mutant Ty5 library was constructed by replacing the mutagenized *BspEI-PflMI* fragment with the corresponding fragment in a wild-type Ty5 element on pNK254 (27).

Mutations were made by PCR-based site-directed mutagenesis to define the TD. pNK254 was PCR amplified using the reverse primer and a mutagenic primer. The amplification product was digested with *EcoRI* and then inserted into the *EcoRI* site of pWW37, a Ty5 subclone containing an *HpaI-SacI* fragment. The *BspEI-PflMI* fragment of the recombinant plasmid was used to replace the corresponding fragment in pNK254. The mutagenic primers were as following: DVO754 for mutant pWW39 (5'-GGA-ATT-CAA-TCG-AAT-CTC-CTC-CAT-CGG-TGG-ATT-CAT-C), DVO755 for mutant pWW40 (5'-GGA-ATT-CAA-TCG-AAT-CTC-CTC-CAT-CGT-TGG-CTT-CAT-CGC-C), DVO756

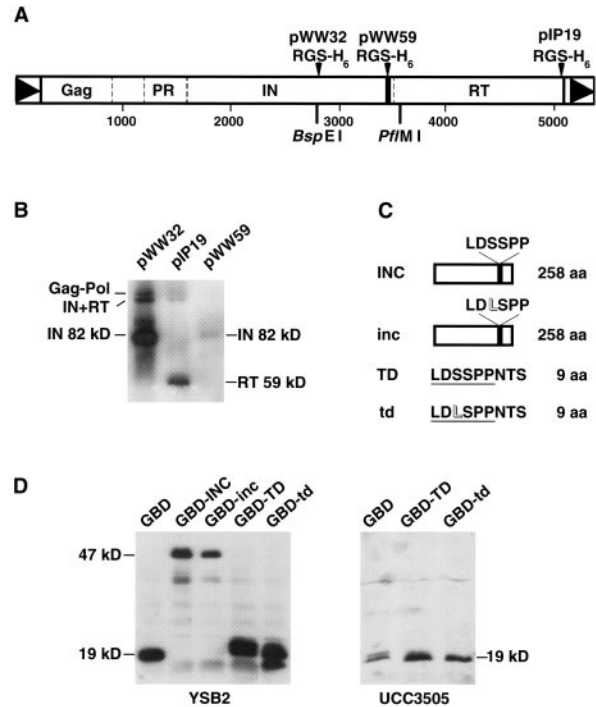


FIG. 1. The Ty5 TD is located at the C terminus of integrase. (A) Ty5 is 5,375 bp in length. It expresses a full-length protein of 182 kDa, which is processed by protease (PR) into Gag, IN, and RT (22). The cleavage sites, based on the mobilities of mature proteins by SDS-PAGE, are shown by dashed lines. The black bar marks the position of TD. The *BspEI* and *PflMI* sites define the region of IN used in the mutagenesis experiment. pIP19, pWW32, and pWW59 carry Ty5 elements that were modified by an RGS-H₆ tag. The tag replaced TD (pWW59) or was inserted either into the middle of IN (pWW32) or at the end of RT (pIP19) (22). (B) The modified Ty5 elements were expressed in yeast, and an anti-RGS-H₆ antibody was used to identify IN or RT on immunoblots. The partially processed and mature protein species are indicated. (C) The Ty5 IN fragments used throughout this paper are shown. INC and inc are 258 aa long (the small letters indicate the version with the S1094L mutation). TD and td represent the wild-type and mutant TD plus 3 flanking aa from Ty5. (D) Western blots demonstrating that the wild-type and mutant GBD fusion proteins are expressed at comparable levels in the test strains (YSB2 and UCC3505). GBD-INC and GBD-TD have molecular masses of approximately 47 and 19 kDa, respectively.

for mutant pWW41 (5'-GGA-ATT-CAA-TCG-AAT-CTC-CTC-CAT-CGT-TG G-ATT-CAT-CGG-CTC-CAA-ATA-C), DVO757 for mutant pWW42 (5'-GG A-ATT-CAA-TCG-AAT-CTC-CTC-CAT-CGT-TGG-ATT-CAT-CGC-CTC-C AG-CTA-CCT-CAT-TT), DVO631 for mutant pXW198 [5'-GGA-ATT-CAA-TCG-AA(T/G)-CTC-CTC-CA(T/G)-CGT-TGG-AT], DVO632 for mutants pXW199, pXW200, and pXW201 [5'-GGA-ATT-CAA-TCG-AAT-CT(C/G)-C T(C/G)-CAT-CGT-TG], and DVO634 for mutant pXW202 [5'-GGA-ATT-CG G-TCG-AAT-CTC-CTC-CAT-CGT-TGG-ATT-CAT-CGC-CTC-CAA-AT(A/G)-CC(T/G)-CAT-TTA-AC].

The Ty5 mutants rut3, rut15, rut31, rut38, rut41, and rut46 were constructed by replacing an *EcoRI-PflMI* fragment of the wild-type Ty5 element on pNK254 with the corresponding fragment from mutants ut3, ut15, ut31, ut38, ut41, and ut46. Because of the multiple *EcoRI* sites in Ty5, this was accomplished in two steps: first, an *EcoRI-EcoRI* fragment from the original mutant element was inserted into the *EcoRI* site of plasmid pWW37, which contains an *HpaI-SacI* fragment of Ty5; second, the *BspEI-PflMI* fragment of the resulting plasmid was used to replace the corresponding fragment of pNK254. For all mutant elements tested in this study, targeting was measured using our plasmid-based targeting assay described in detail in our previous study (14).

Tethered silencing. Gal4p DNA binding domain (GBD) fusion proteins were generated using pGBD plasmids, which contain the GAL4 DNA binding domain under the control of the *ADHI* promoter (24). The plasmid expressing GBD-INC (pXW140) contains amino acids (aa) 879 to 1136 of Ty5 inserted between the *Bam*HI and *Pst*I sites of pGBDU; the GBD-inc construct (pXW158) is identical, except that it carries the S1094L mutation (14). The plasmid expressing TD fused to GBD (GBD-TD) (pXW205) was generated by inserting into the *Eco*RI and *Bgl*II sites of pGBDU a short DNA fragment created from two complementary oligonucleotides, DVO690 (5'-AAT-TCT-TGG-ATT-CAT-CG C-CTC-CAA-ATA-CCT-CA) and DVO691 (5'-GAT-CTG-AGG-TAT-TTG-G AG-GCG-ATG-AAT-CCA-AG); the GBD-td construct (pXW213) is identical, except that it has the S1094L mutation. We generated versions of these plasmids (pWW48, GBD; pWW49, GBD-TD; pWW50, GBD-td) in which the *TRP1* marker gene of pGBD was replaced with *HIS3*. This was accomplished by replacing the *Eco*RV-*Xba*I fragment with a *HIS3*-containing *Nru*I-*Xba*I fragment. *LEU2*-based expression plasmids (pWW44, GBD; pWW45, GBD-TD; pWW46, GBD-td) were generated by swapping the Gal4p activation domain (GAD)-encoding *Sph*I fragment from pGAD (which has the *LEU2* marker [24]) with an *Sph*I fragment encoding the various GBD fusion proteins.

To test silencing of the *HMR* reporter gene, the above-described expression plasmids were transformed into strains with different *HMR*-E mutations (10). These strains include YSB1 (aeB but no UASg), YSB2 (aeB::3×UASg) and YSB35 (Aeb::3×UASg). *sir* derivatives of YSB2 and YSB35 include RS1072 and RS112 (*sir1::URA3*), RS1042 and RS1132 (*sir2::URA3*), RS1061 and RS1133 (*sir3::URA3*), and RS1067 (*sir4::URA3*) (1). The strain used to assess telomeric silencing was UCC3505 (kind gift of D. Gottschling) (16). Complementation of GBD-TD-induced loss of telomeric silencing was tested by introducing *SIR* genes (kind gift of J. Rine) on a 2- μ m plasmid (pRS424 [11]). The *SIR2* plasmid (pSZ270) carries a *Not*I-*Xho*I fragment from pRS315-*SIR2*, the *SIR3* plasmid (pSZ282) carries a *Bam*HI-*Sal*I fragment from pJR104, and the *SIR4* plasmid (pSZ269) carries a *Sac*II-*Cla*I fragment from pRS316-*SIR4*. To measure silencing, an overnight culture was grown to saturation for each strain and adjusted to an optical density at 600 nm of 1. Tenfold serial dilutions were made; 10 μ l of each dilution was spotted onto both the test plate and the control plate. The plates were incubated at 30°C for 2 days.

Two-hybrid assays. Two-hybrid assays were performed using yeast strain L40 (20), which has *HIS3* and *lacZ* reporter genes under the control of upstream LexA operators. The LexA-SIR4C construct was previously described (2) and includes the C-terminal region of Sir4p (aa 951 to 1358). Control strains expressed LexA from plasmid pBTM116 (4). GAD-INC fusions were constructed by inserting an *Xma*I-*Pst*I fragment from pXW140 into pGAD (24); the GAD-inc construct is identical, except that it carries the S1094L mutation. The control expressed GAD from pGAD. Strains with the relevant plasmids were inoculated into 2 ml of selective medium and shaken at 30°C for 24 h. Tenfold serial dilutions of the cultures were made, and 5 μ l of each dilution was plated onto synthetic complete medium (SC)-Trp-Leu or SC-Trp-Leu plus 5 mM 3-amino-1,2,4-triazole medium; plates were incubated at 30°C for 3 days.

Protein analyses. Immunoblot analyses of Ty5 and GBD proteins were conducted as previously described using antibodies specific to the RGS-H₆ tag (Qiagen) or GBD (Santa Cruz Biotechnology) (22). Two of the epitope-tagged Ty5 elements were described in that previous report (pWW32 and pIP19). The element with the TD replaced by the RGS-H₆ epitope (pWW59) was constructed by PCR mutagenesis (3). The Ty5 fragment within pWW37 (see above) was amplified using the universal primer and the reverse primers DVO1180 (5'-TG A-TGG-TGA-TGC-GAT-CCT-CTC-GAT-GGA-GGA-GAT-TCG-ATT-G) and DVO1181 (5'-CGC-ATC-ACC-ATC-ACC-ATC-ACA-ATA-CCT-CAT-TTA-A CG-CGG-C). The *Bsp*EI-*Pfl*MI fragment contained within the amplification product was used to replace the corresponding fragment in the wild-type Ty5 element carried on pSZ152 (54).

To measure in vitro interactions between the TD and Sir4p, we first constructed a plasmid expressing the C terminus of *SIR4* (aa 950 to 1358) by PCR amplifying pSZ269 (see above) with primers DVO1137 (5'-GAA-GGA-TCC-A GA-GGA-TCG-CAT-CAC-CAT-CAC-CAT-CAC-AGA-AGA-GTG-TCG-CA T-AGT-G) and DVO1085 (5'-TGA-TCT-CGA-GTC-AAT-ACG-GTT-TTA-T CT-CC). The amplification product was digested with *Bam*HI and *Xho*I and inserted into pCITE-2a(+) (Novagen) to generate pWW56. pWW56 DNA (0.5 μ g) was used in a 50- μ l coupled transcription-translation reaction mixture (Promega) containing 20 μ Ci of [³⁵S]methionine. GBD fusion proteins were immunoprecipitated as previously described (41) from 250-ml yeast cultures (optical density at 600 nm, 0.8 to 1.2) with either pXW205, pXW213, or pGBDU. Cells were harvested, washed with ice-cold water, and resuspended in 1.8 ml of lysis buffer B (50 mM HEPES [pH 7.5], 0.5 M NaCl, 10% glycerol, 0.1% IGPEL CA-630 [Sigma], 2 mM dithiothreitol, 2.5 mM benzamide, 0.5 mM phenyl-

methylsulfonyl fluoride, 2 μ g of leupeptin per ml, 2 μ g of bestatin per ml, and 2 μ g of pepstatin per ml). Cells were disrupted by the glass bead method (3), and the lysate was centrifuged at 12,000 rpm (JA-20 rotor; Beckman) at 4°C for 30 min. Levels of fusion protein in the supernatant were assessed by immunoblot analysis using anti-GBD antibodies (Santa Cruz Biotechnology). The GBD fusion proteins were immunoaffinity purified from 300 μ l of supernatant using 10 μ l of an anti-GBD agarose bead slurry (Santa Cruz Biotechnology). After a 2-h incubation, the beads were collected by centrifugation (500 \times g, 2 min), washed twice with 300 μ l of wash buffer (50 mM HEPES [pH 7.5], 150 mM NaCl, 5 mM Mg diacetate, protease inhibitors) and once with 1 \times PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM KH₂PO₄). The 50- μ l in vitro transcription-translation reaction mixture containing the labeled Sir4p was then added to the washed anti-GBD agarose beads, and the mixture was incubated at room temperature for 30 min. The beads were collected by centrifugation (500 \times g, 2 min) and washed three times with PBS. Ten microliters of 2 \times sodium dodecyl sulfate (SDS) sample buffer (3) was added to each tube; samples were heated (95°C for 10 min) and separated by SDS-polyacrylamide gel electrophoresis (PAGE). The gel was dried and exposed to X-ray film overnight.

RESULTS

Defining the Ty5 targeting domain. We previously found that a single amino acid substitution at position 1094 in the Ty5 polyprotein (S1094L) dramatically decreased targeting to the telomeres and silent mating loci (14). This indicated that Ty5 plays an active role in selecting targets and pointed to a possible targeting domain around S1094. To further define Ty5-encoded targeting determinants, a 758-bp *Bsp*EI-*Pfl*MI restriction fragment encompassing S1094 was mutagenized by PCR (Fig. 1A). The PCR products were used to replace the corresponding wild-type fragment, and sequencing of several recombinants revealed that each carried two to eight base pair changes. Over 2,100 mutagenized elements were screened for targeting defects using our plasmid-based targeting assay, which we have previously shown is an effective measure of chromosomal integration patterns (14). In this assay, targeting is quantified as the percentage of integration events that occur at a plasmid-borne *HMR* locus. We identified 11 elements that were impaired in targeting to various degrees (Table 1). DNA sequencing revealed multiple nucleotide changes in the *Bsp*EI-*Pfl*MI fragments of these elements. Whereas some mutations were silent, the number of amino acid substitutions ranged from two (mutants ut41 and ut38) to eight (mutant ut46) (Table 1).

We made three observations regarding the targeting mutants: (i) one mutant (ut5) had an S1094L substitution identical to the mutation described in our previous study; (ii) all of the remaining mutants had an amino acid substitution in the vicinity of S1094 (encompassing a span of 6 aa residues); and (iii) independent mutations in the same residue near S1094 had similar targeting defects (e.g., mutants ut23, ut31, and ut33). These observations suggested that the mutations near S1094 were primarily responsible for the loss of target specificity. To test this hypothesis, we constructed seven new mutants that carried only one amino acid substitution near S1094 (Table 1). The S1094L mutation was not included because it was previously characterized (14). Targeting assays indicated that in all seven cases, the single mutations conferred a targeting defect nearly identical to that of the original mutants. This was the case even for conservative substitutions; for example, the L1092V mutation dramatically decreased targeting. Based on these data, we concluded that Ty5 encodes a TD and that it may be limited to a short stretch of 6 aa (LDSSPP).

Because our mutagenesis recovered multiple substitutions in

TABLE 1. PCR mutagenesis defines the Ty5 TD

Element	Transposition efficiency (10^{-5}) ^a	Transposition fold decrease	% of targeted transposition ^b	Targeting fold decrease	No. of base changes ^c	Amino acid sequence near S1094 ^d
Wild type	13.3 ± 2.8	1.0	7.9	1.0	0	LDSSPP
Mutants identified in the screen						
ut3	3.43 ± 0.92	3.9	1.3	6.1	4	<u>SD</u> SSPP
ut29	0.97 ± 0.11	13.7	1.4	5.6	5	<u>VD</u> SSPP
ut5	1.93 ± 0.42	6.9	0.4	19.8	5	<u>LDL</u> SSPP
ut35	0.52 ± 0.08	25.6	0	NA ^e	7	<u>LD</u> <u>P</u> SSPP
ut41	1.49 ± 0.27	8.9	0.2	39.5	2	<u>LD</u> <u>P</u> SSPP
ut38	1.14 ± 0.22	11.7	1.1	7.2	2	<u>LD</u> <u>S</u> <u>L</u> PP
ut46	2.46 ± 0.68	5.4	0.5	15.8	8	<u>LD</u> SSPP
ut15	0.44 ± 0.19	30.2	0	NA	4	<u>LD</u> SS <u>P</u> L
ut23	1.42 ± 0.46	9.4	2.5	3.2	5	<u>LD</u> SS <u>P</u> L
ut31	1.55 ± 0.02	8.6	1.8	4.4	6	<u>LD</u> SS <u>P</u> Q
ut33	1.40 ± 0.06	9.5	2.1	3.8	4	<u>LD</u> SS <u>P</u> Q
Mutants with single amino acid substitutions in the vicinity of S1094						
rut3	2.40 ± 0.10	5.5	0.7	11.3	1	<u>SD</u> SSPP
pWW39	ND ^f	2.1	2.1	3.8	1	<u>VD</u> SSPP
rut41	2.12 ± 0.16	6.3	0.7	11.3	1	<u>LD</u> <u>P</u> SSPP
rut38	2.25 ± 0.66	5.9	1.1	7.2	1	<u>LD</u> <u>S</u> <u>L</u> PP
rut46	2.11 ± 0.98	6.3	0.6	13.2	1	<u>LD</u> SSPP
rut15	2.64 ± 0.23	5.0	3.0	2.6	1	<u>LD</u> SS <u>P</u> L
rut31	3.09 ± 0.89	4.3	2.0	4.0	1	<u>LD</u> SS <u>P</u> Q

^a Data compiled from three independent experiments.

^b Number of transposition events to the target plasmid divided by total number of transposition events. For the mutant whose value is 0%, transposition was too low to meaningfully determine the percentage of targeted transposition.

^c Base changes were identified by DNA sequencing of the *BspEI-PflMI* fragment from each of the mutants.

^d S1094 is in bold; missense mutations are underlined.

^e NA, not applicable.

^f ND, not determined.

the same amino acid, this suggested that the mutagenesis of the PCR fragment was saturated. To confirm the boundaries of the TD and to ensure that all critical amino acid residues near the LDSSPP motif had been identified, directed PCR mutagenesis was used to change residues near S1094 to alanine that were not identified as important for targeting. This included two residues (Asp1093 and Pro1096) that are located between other amino acids critical for targeting as well as residues upstream and downstream of the LDSSPP motif (Table 2). These substitutions had at most a modest effect on targeting (e.g., P1096, pWW41) or no effect at all (e.g., N1098, pWW42). It is interesting that each of the targeting mutants identified in

TABLE 2. Site-directed mutagenesis indicates that only 4 of the 6 aa in the Ty5 TD are required for integration specificity

Element	% of targeted transposition ^a	Amino acid sequence near the targeting domain ^b
Wild type	7.9	SPPSLDSSPPNTS
pWW40	7.0	SPPSLASSPPNTS
pWW41	5.9	SPPSLDSS<u>A</u>PPNTS
pXW198	8.6	<u>A</u>PPALDSSPPNTS
pXW199	8.0	<u>S</u>APSLDSSPPNTS
pXW201	8.2	<u>S</u>AASLDSSPPNTS
pXW200	8.6	<u>S</u>PASLDSSPPNTS
pXW202	7.9	SPPSLDSSPP<u>NAA</u>
pWW42	7.8	SPPSLDSSPP<u>A</u>TS

^a Data compiled from three independent experiments.

^b The TD is in bold; missense mutations are underlined.

this study had transposition frequencies from two- to sixfold lower than that of the wild type (Table 1). In contrast, the mutants without altered target specificity transposed at near wild-type levels (data not shown). This observation was also made in our original study, wherein the S1094L mutation caused a fourfold decrease in transposition (14). This indicates that TD mutations also affect transposition efficiency.

The TD is located in the integrase C terminus. Ty5 encodes a single open reading frame that is processed by a Ty5 gene-encoded protease into several proteins, including IN (80 kDa) and reverse transcriptase (RT; 59 kDa) (22). Extrapolating molecular weights from the Ty5 amino acid sequence, we predict that the protease cleavage site separating IN from RT is within the vicinity of the TD and therefore that the TD may reside either within the C terminus of IN or within the N terminus of RT. To distinguish between these possibilities, the TD was replaced with an epitope tag (RGS-H₆). Like the other TD mutants, the epitope-tagged element transposed at a level more than fivefold lower than that of the wild type (data not shown). Proteins were prepared from strains expressing the TD-tagged element (pWW59) as well as from control strains expressing elements with the same epitope at either the C terminus of RT (pIP19) or within the middle of the IN coding region (pWW32) (Fig. 1A). Immunoblot analysis indicated that the element with the tag located at the TD expressed a protein with the same mobility on SDS-PAGE gels as that of IN (Fig. 1B). Other tagged proteins revealed by immunoblotting represent various processing products or intermediates

(22). All lanes contained equivalent amounts of total protein, yet levels of the TD-tagged IN were severalfold lower than those of the other tagged IN, suggesting that mutations in the TD may affect protein stability. Nonetheless, we could conclude from this experiment that the TD resides within IN, consistent with its role in target site selection.

Tethering the TD to DNA nucleates silent chromatin. Our model for target specificity predicts that the TD interacts with silent chromatin to tether the integration apparatus to its target sites. If this is the case, then the converse may also be true: a TD tethered to DNA may recruit silencing factors and establish silent chromatin. To test this idea, we used an assay that evaluates a protein's ability to establish transcriptional silencing (10). The test protein was first fused to the GBD, and the fusion protein was expressed in a yeast strain with a weakened *HMR-E* silencer that contains Gal4p binding sites (UASg). The effectiveness of the tethered proteins in nucleating silent chromatin was measured by the transcriptional status of a reporter gene at *HMR* (e.g., *TRP1*). We constructed four Ty5-GBD fusion proteins (Fig. 1C), one of which has 258 aa of the Ty5 INC (GBD-INC) and another of which has only 9 Ty5 aa (GBD-TD), 6 of which constitute the TD. The remaining two fusion constructs differed only by the S1094L mutation (GBD-inc, GBD-td). Immunoblot analysis indicated that the wild-type and mutant forms of the fusion proteins were expressed equivalently in yeast (Fig. 1D).

The fusion constructs were introduced into yeast strains with two different *HMR-E* mutations: one lacking binding sites for ORC and Rap1p (aeB) and the other lacking binding sites for Rap1p and Abf1p (Aeb) (10). The ability of the fusion proteins to establish silencing was measured by spotting 10-fold serial dilutions of the various strains onto media lacking tryptophan. The GBD-INC fusion was found to repress transcription over 100-fold (Fig. 2). Surprisingly, even the 9-aa GBD-TD fusion protein silenced *TRP1* more than 10-fold. In both cases, the transcriptional silencing required the presence of the UASg, and in agreement with previous work (1, 10), the fusion proteins were more effective at silencers with a wild-type ORC binding site (Aeb). If the TD interacts with silent chromatin as predicted, then mutations that disrupt targeting should decrease its effectiveness in recruiting silencing factors. Consistent with this hypothesis, both GBD-inc and GBD-td were unable to establish transcriptional silencing (Fig. 2C). All of the above observations were confirmed in a strain with a *URA3* reporter gene at *HMR*, indicating that the silencing was not reporter gene dependent (data not shown).

Silent chromatin, by definition, requires the actions of Sir2p, Sir3p, and Sir4p. To determine whether the TD fusions establish silent chromatin or rather act in some other way to occlude the transcriptional machinery from the *TRP1* promoter, the TD fusions were introduced into various *sirΔ* strains (10) (Fig. 3). Sir1p was not required for TD-mediated silencing, consistent with its primary role in recruiting components of silent chromatin to the *HM* loci (48). However, the structural components of silent chromatin, Sir2p, Sir3p, and Sir4p, were all required, indicating that the Ty5 TD represses the reporter gene at *HMR* by establishing silent chromatin.

Overexpression of the TD disrupts telomeric silencing. Transcriptional silencing is very sensitive to the expression level of some components of silent chromatin. Overexpression

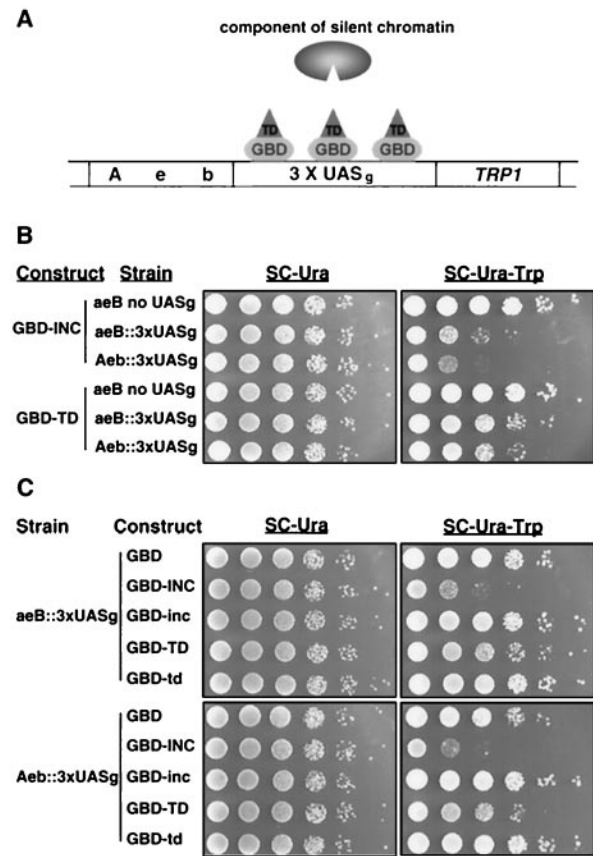


FIG. 2. The Ty5 TD nucleates silent chromatin. (A) A cartoon depicting the tethered silencing assay. Yeast strains were used with deletions in two of the protein binding sites in *HMR-E* (A, E, and B) and three copies of UASg, binding sites for Gal4p (10). In the example depicted, the E and B binding sites are deleted (e, b), resulting in derepression of transcription at *HMR*. Expression of a fusion protein between GBD and the Ty5 TD (GBD-TD) was tested for its ability to recruit components of silent chromatin and restore silencing. Silencing is measured by expression of the adjacent *TRP1* marker gene. (B) Silencing was established when GBD-INC and GBD-TD were tethered to the weakened *HMR* locus by the triple UASg. Serial, 10-fold dilutions of cells were plated onto control (SC-Ura) or test (SC-Ura-Trp) medium to measure silencing of the *TRP1* reporter gene at *HMR*. (C) A point mutation that abolishes targeting fails to restore silencing for both fusion proteins (GBD-inc and GBD-td) in both test strains.

of Sir4p, for example, disrupts telomeric silencing, presumably by titrating away other components of silent chromatin or by disrupting complex formation (23, 39) (Fig. 4B). We tested whether overexpression of GBD-TD could disrupt telomeric silencing by monitoring the expression of telomeric (and therefore normally silenced) *URA3* and *ADE2* genes (16). Overexpression of GBD-TD (Fig. 1D) resulted in the loss of telomeric silencing, as evidenced by the ability of cells to grow on media lacking uracil (Fig. 4A). *ADE2* expression was also evident by the white colony phenotype rather than the pinkish color characteristic of *ADE2* repression and adenine precursor accumulation. As observed in the tethering experiments, expression of the fusion with the S1094L mutation did not affect telomeric silencing.

To test whether GBD-TD disrupts telomeric silencing by titrating away components of silent chromatin, the *SIR* genes

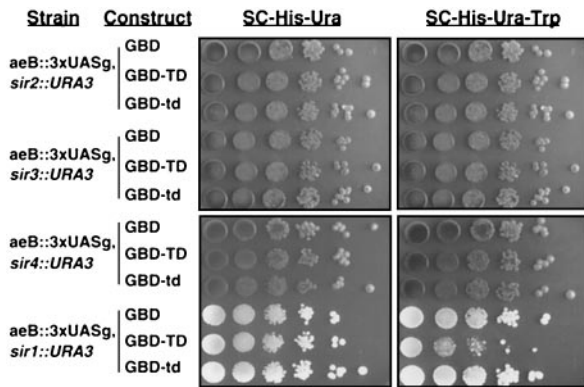


FIG. 3. Silencing conferred by the TD is Sir dependent. The assay system is described in the legend to Fig. 2. In strains with deletions of *SIR2*, *SIR3*, or *SIR4*, GBD-TD fails to establish silencing at a weakened *HMR* locus. Silencing, however, does not require the *SIR1* gene. The Ade⁻ phenotype of the *sir2Δ*, *sir3Δ*, and *sir4Δ* strains confers their dark color; the *sir1Δ* strain is Ade⁺.

were ectopically expressed by introducing them on high-copy-number 2μm plasmids (Fig. 4B). Overexpression of Sir4p by this means has previously been shown to disrupt telomeric silencing (39), and we made the same observation regardless of whether GBD-TD was coexpressed. Overexpression of Sir2p did not restore telomeric silencing to GBD-TD-expressing strains; however, it did increase telomeric silencing in GBD or GBD-td strains. In contrast, overexpression of Sir3p overcame the GBD-TD-dependent loss of telomeric silencing. This sug-

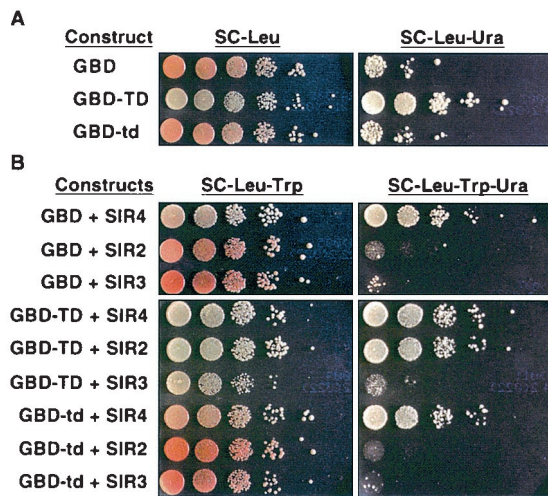


FIG. 4. Overexpression of the Ty5 TD disrupts telomeric silencing and loss of silencing is complemented by overexpression of Sir3p. Two reporter genes, *URA3* and *ADE2*, are located at telomeres VIII and VR, respectively (16). *URA3* expression was measured by growth of the yeast cells on selective media. *ADE2* expression is indicated by colony color; when *ADE2* is repressed, the colonies are red or pink. (A) Overexpression of GBD-TD disrupts telomeric silencing (i.e., causes *URA3* expression), in contrast to what occurs in its mutant form (GBD-td) and with GBD alone. (B) Overexpression of Sir2p and Sir3p strengthen silencing, and Sir4p breaks silencing, as previously reported (23, 39). This is demonstrated in the GBD control strains and GBD-td strains. When the TD and *SIR* genes are overexpressed in the same strains, Sir3p, but not Sir2p or Sir4p, restores silencing.

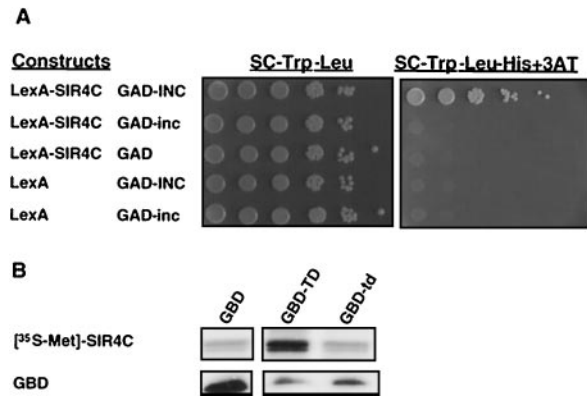


FIG. 5. The Ty5 INC interacts with the C terminus of Sir4p. (A) Yeast two-hybrid assays reveal an interaction between GAD-INC and LexA-SIR4C. Liquid cultures expressing the various LexA and GAD proteins were serially diluted 10-fold and spotted onto plates. A positive two-hybrid interaction was measured by transcriptional activation of the *HIS3* reporter, which allowed for growth on selective media (SC-Trp-Leu-His with 5 mM 3-amino-1,2,4-triazole [3AT]). Activation of the *HIS3* marker requires both SIR4C and the wild-type TD. (B) SIR4C interacts with GBD-TD in vitro. SIR4C was expressed and labeled with [³⁵S]methionine by coupled transcription and translation. GBD-TD and GBD-td were expressed in yeast and immunoaffinity purified with anti-GBD agarose beads. The top lanes indicate the amount of labeled SIR4C bound by the various GBD proteins. The bottom lanes are from an immunoblot performed with anti-GBD antibodies, and they indicate the levels of GBD proteins in the extract used for immunoaffinity purification.

gests that Sir3p is titrated away (either directly or indirectly) by interacting with the TD. Alternatively, excess Sir3p could dominantly restore silencing by bypassing a factor being titrated away by TD. In addition, a growth defect was observed on nonselective media when Sir3p and GBD-TD were both overexpressed (Fig. 4B).

Sir4p interacts with Ty5 IN. Several lines of evidence suggest that Sir3p and Sir4p are likely candidates for interacting with the Ty5 IN to mediate target specificity: (i) targeting is largely abolished in *sir3Δ* and *sir4Δ* strains (53); (ii) in *sir4-42* strains, Ty5 integration specificity changes with the chromosomal localization of Sir3p and Sir4p; and (iii) as described above, loss of telomeric silencing due to overexpression of the TD can be complemented by overexpression of Sir3p. To test whether Ty5 IN interacts with Sir3p or Sir4p, two-hybrid assays were conducted. An interaction was detected between the C terminus of Sir4p (SIR4C, aa 951 to 1358 expressed as a LexA fusion protein) and the Ty5 INC (expressed as a GAD fusion protein). This interaction strongly activated the *HIS3* reporter gene and enabled growth of yeast cells on selective media (Fig. 5A). Consistent with the role of the TD in silencing, this interaction required a wild-type TD: the S1094L mutation greatly weakened the two-hybrid interaction. These data indicate that INC binds to Sir4p (either directly or indirectly) and that the TD is required for this interaction.

To confirm the two-hybrid data, we tested whether the Sir4p C terminus and the Ty5 TD could interact in vitro. GBD, GBD-TD, and GBD-td were expressed in yeast and immunoaffinity purified using anti-GBD agarose beads. Beads with the bound GBD proteins were incubated with SIR4C that had been labeled with [³⁵S]methionine. The beads were washed,

and the proteins were eluted and separated by SDS-PAGE. Sir4p bound to GBD-TD. Lower, background levels of binding were observed for GBD and GBD-td (Fig. 5B). These *in vitro* data support the results obtained in the two-hybrid assays and collectively suggest that the biological activity of the TD is mediated by interactions with Sir4p.

DISCUSSION

In a simple model to explain LTR retroelement target specificity, the interaction between the preintegration complex and DNA-bound proteins tethers the integration machinery to target sites and results in integration site biases (7). Support for this model comes from the study of the yeast Ty retrotransposons. These elements have different target preferences: Ty1 and Ty3 prefer sites of RNAP III transcription (9, 13), and Ty5 prefers silent chromatin (56). In both cases, DNA-bound protein complexes are required for targeting. Human immunodeficiency virus (HIV) IN interacts with a human homolog of the transcription factor SNF5 in two-hybrid assays (25); however, there is no evidence that this interaction mediates target site choice. In this study, we define a Ty5-encoded TD (LXSSXP) and show that it interacts with Sir4p, a component of silent chromatin. To our knowledge, this interaction between a retroelement-encoded protein and a chromatin factor provides the first direct evidence for the targeting model.

Integrase C terminus and the TD. Retrotransposon and retroviral INs consist of three distinct domains (26): (i) an N-terminal region with a zinc-binding motif that is required for IN activity and likely binds cDNA, (ii) a catalytic domain that executes the integration reaction, and (iii) a C-terminal region which, for the retroviruses and some retrotransposons, is required for cDNA 3'-end processing. The retrotransposon integrase C termini are considerably larger than their retroviral counterparts; the Ty1 and Ty5 C termini constitute more than half of IN (e.g., HIV IN is 288 aa; Ty1 IN is 635 aa). Little is known about the function of the retrotransposon C-terminal extensions, with the exception that the very C terminus of Ty1 IN encodes a nuclear localization signal that is required for the preintegration complex to gain nuclear access (28, 42). For Ty1 and Ty5, the coding region of IN lies upstream of RT, and both are released from the polyprotein by proteolytic cleavage. In our earlier study of a Ty5 targeting mutant, we could not determine whether the targeting mutation was located in the C terminus of IN or the N terminus of RT (14). By replacing the TD with an epitope tag and comparing its electrophoretic mobility to those of other tagged forms of RT and IN, we demonstrate here that the TD resides within IN. This indicates that IN is responsible for target specificity and demonstrates a new function for the INC. For some members of the more distantly related Ty3/*gypsy* group retrotransposons (*Metaviridae*), the INC encodes a chromodomain, a motif implicated in targeting proteins to chromatin (37). The INC, therefore, may generally be used by retroelements for integration site selection.

Despite the large size of the INC, the Ty5 TD identified through our mutant screen spans only 6 aa residues. This short domain is biologically active: when as few as 9 aa encompassing the TD are expressed as part of a fusion protein (e.g., GBD-TD), they nucleate silent chromatin or disrupt telomeric si-

lencing. Fusion proteins expressing larger fragments of the INC are at least 10-fold more effective in nucleating silent chromatin. This suggests that other regions of the INC play a role in targeting. These regions may not have been identified by our mutant screen if they are required for transposition. Transposition defects were observed during characterization of several targeting mutants. For example, the transposition efficiency of mutant *ut15* was more than 30-fold lower than that of the wild type. This mutant had three missense mutations in the INC in addition to the mutation in the TD (Table 1). When the *ut15* TD was evaluated in isolation (see *rut15*, Table 1), the level of transposition was only fivefold lower than that of the wild type. Additionally, we observed that most of the mutagenized elements (~60%) were unable to transpose or showed greatly reduced levels of transposition (unpublished data). Some of these likely carry stop codons or frameshift mutations that prevent synthesis of the downstream RT; however, the frequency of nontransposing elements was too high based on the extent of mutagenesis (~0.5%). This implies that regions of the C terminus outside of the TD are important for transposition.

In addition to loss of target specificity, a second phenotype shared by all of the TD mutants is an overall decrease in transposition (two- to sixfold). For wild-type Ty5, more than 94% of Ty5 integration events occur within regions of silent chromatin whereas the remainder appear to be randomly distributed throughout the genome (55, 57). If the TD is not required for integration, then TD mutations should not alter transposition frequencies. If the TD is required only for targeted integration, then the transposition frequencies would be predicted to drop ~15-fold (94% divided by 6%). The intermediate effect of TD mutations on transposition suggests that the TD plays a role beyond tethering the integration apparatus to its preferred target sites. Possibilities include facilitating the integration reaction, facilitating reverse transcription (which is supported by recent data indicating that the INC is required to produce functional Ty1 RT *in vitro* [49]), or facilitating nuclear localization (the Ty1 nuclear localization signal is located in approximately the same region of the protein as the Ty5 TD [28, 42]). In strains expressing a modified Ty5 in which the TD was replaced by an epitope tag, we observed significantly lower levels of IN protein (Fig. 1). It is also possible that a wild-type TD is required for protein stability.

The TD and silent chromatin. A large number of proteins are involved in the assembly and maintenance of silent chromatin (reviewed in references 35 and 36). These include proteins such as Sir3p and Sir4p that carry out structural or scaffolding roles and proteins like Sir1p and Rap1p that recruit and nucleate the structural proteins at the silencers. Tethering either class of these proteins to weakened *HMR-E* silencers effectively establishes silencing (10, 38). Likewise, the INC is very effective in nucleating silencing and causes a 100- to 1,000-fold decrease in the expression of reporter genes at *HMR*. The TD does not simply recruit a protein complex that occludes the RNAP II machinery from the reporter gene. Rather, it establishes silent chromatin as defined by the requirement for Sir2p, Sir3p, and Sir4p. TD-mediated silencing does not need Sir1p, which primarily acts at *HM* loci and is not required for silencing at the telomeres or rDNA. At the *HM* loci, Sir1p interacts with the ORC (48), and it is interesting that the TD is more

effective when the ORC binding site is intact. This has also been observed with other nucleators of silent chromatin (1, 10). Although a number of *S. cerevisiae* proteins have motifs that match the TD consensus (LXSSXP), none are known components of silent chromatin.

We have previously shown that, in contrast to what occurs in *sir3Δ* and *sir4Δ* strains, targeting of Ty5 to the telomeres and *HM* loci is only partially impaired in *sir2Δ* strains and occurs at levels approximately 50% of those of the wild type (53). Furthermore, the loss of telomeric silencing caused by overexpressing the TD cannot be restored by overexpressing Sir2p. In light of recent findings that Sir2p is a histone deacetylase (21, 33, 47), this suggests that the Ty5 integration apparatus does not sense Sir2p-mediated acetylation patterns. The targeting defect in *sir2Δ* strains is likely a secondary consequence of perturbations in silent chromatin. In contrast, the loss of telomeric silencing caused by overexpressing the TD can be restored by additional *SIR3* expression. The Sir4p C terminus interacts directly with Sir3p, and it may be that overexpression of the TD disrupts this interaction, which, in turn, is stabilized by additional Sir3p. In that regard, it has been observed that loss of telomeric silencing caused by overexpression of the Sir4p C terminus can be complemented by overexpressing Sir3p (15). Finally, it is important to note that the combined expression of Sir3p and a wild-type TD causes a growth defect, indicating that, although the TD may not interact directly with Sir3p, Sir3p does modulate the biological activity of the TD.

The TD and Sir4p. Prior to the carrying out of our two-hybrid assays, no evidence distinguished the roles of Sir3p and Sir4p in Ty5 target specificity. The only notable difference was that Ty5 cDNA recombination increased more than 10-fold in *sir4Δ* strains and was only marginally affected in *sir3Δ* strains (53). In strains with the *sir4-42* allele, which expresses a C-terminally truncated form of Sir4p (aa 1 to 1237), Ty5 integrates preferentially into the rDNA (53). Here we demonstrate that the Ty5 TD interacts with the Sir4p C terminus (aa 951 to 1358), suggesting that the relevant region of interaction is located between aa 951 and 1237. The Sir4p C terminus interacts with many proteins, including Sir2p and Sir3p (41), Rap1p (6, 43), Sif2p (12), and Dis1p (52). Two-hybrid interactions may result if one or more of these proteins serve as a bridge between the TD and Sir4p. However, we have not observed two-hybrid interactions between IN and other components of silent chromatin (data not shown). Furthermore, the observed *in vitro* binding between Sir4p and the TD argues that these molecules interact directly. Nonetheless, because we used fusion proteins purified from yeast cells for these experiments, we cannot exclude the possibility that other factors copurified or modified the TD to yield productive interactions.

Concluding remarks. Previous studies demonstrated that DNA bound in silent chromatin is inaccessible to proteins like HO endonuclease, restriction enzymes, and transcription factors (34). However, the ability of Ty5 to integrate into silent chromatin suggests that this DNA is accessible to Ty5 IN. Our model suggesting that target specificity results from simply tethering the preintegration complex to chromatin may therefore require additional refinements. For example, the integration complex may induce changes in silent chromatin during integration, and the role of the TD in such processes is an important area of future research.

An increased understanding of targeting mechanisms may make it possible to manipulate retroelement target site choice. It may be possible to change the integration preference of retrotransposons by replacing TDs with peptide motifs that interact with specific chromosomal proteins. Such engineered retrotransposons may become useful tools for studying chromatin organization and may provide novel methods for genome manipulation. Retroviral vectors are widely used for DNA delivery in human gene therapy; however, uncontrolled, random integration into the host genome is one of their major drawbacks. It may now be possible to better control retroviral integration site choice to improve the efficacies of these vectors in gene delivery.

ACKNOWLEDGMENTS

W. Xie and X. Gai contributed equally to this work.

We thank Phil Irwin for his expert technical assistance. J. Rine and D. Gottschling graciously provided plasmids used in this study.

This work was supported by a grant from the American Cancer Society (RPG9510106MBC) to D.F.V. and by Hatch Act and State of Iowa funds.

REFERENCES

1. Andrusis, E. D., A. M. Neiman, D. C. Zappulla, and R. Sternglanz. 1998. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394**:592–595.
2. Ansari, A., and M. R. Gartenberg. 1997. The yeast silent information regulator Sir4p anchors and partitions plasmids. *Mol. Cell. Biol.* **17**:7061–7068.
3. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1987. *Current protocols in molecular biology*. Greene/Wiley Interscience, New York, N.Y.
4. Bartel, P. L., and S. Fields. 1995. Analyzing protein-protein interactions using two-hybrid system. *Methods Enzymol.* **254**:241–263.
5. Boeke, J. D., and S. E. Devine. 1998. Yeast retrotransposons: finding a nice quiet neighborhood. *Cell* **93**:1087–1089.
6. Buck, S. W., and D. Shore. 1995. Action of a RAP1 carboxy-terminal silencing domain reveals an underlying competition between *HMR* and telomeres in yeast. *Genes Dev.* **9**:370–384.
7. Bushman, F. 1995. Targeting retroviral integration. *Science* **267**:1443–1444.
8. Chalker, D. L., and S. B. Sandmeyer. 1993. Sites of RNA polymerase III transcription initiation and Ty3 integration at the U6 gene are positioned by the TATA box. *Proc. Natl. Acad. Sci. USA* **90**:4927–4931.
9. Chalker, D. L., and S. B. Sandmeyer. 1992. Ty3 integrates within the region of RNA polymerase III transcription initiation. *Genes Dev.* **6**:117–128.
10. Chien, C. T., S. Buck, R. Sternglanz, and D. Shore. 1993. Targeting of SIR1 protein establishes transcriptional silencing at *HM* loci and telomeres in yeast. *Cell* **75**:531–541.
11. Christianson, T. W., R. S. Sikorski, M. Dante, J. H. Shero, and P. Hieter. 1992. Multifunctional yeast high-copy-number shuttle vectors. *Gene* **110**:119–122.
12. Cockell, M., H. Renaud, P. Watt, and S. M. Gasser. 1998. Sif2p interacts with Sir4p amino-terminal domain and antagonizes telomeric silencing in yeast. *Curr. Biol.* **8**:787–790.
13. Devine, S. E., and J. D. Boeke. 1996. Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. *Genes Dev.* **10**:620–633.
14. Gai, X., and D. F. Voytas. 1998. A single amino acid change in the yeast retrotransposon Ty5 abolishes targeting to silent chromatin. *Mol. Cell* **1**:1051–1055.
15. Gotta, M., F. Palladino, and S. M. Gasser. 1998. Functional characterization of the N terminus of Sir3p. *Mol. Cell. Biol.* **18**:6110–6120.
16. Gottschling, D. E., O. M. Aparicio, B. L. Billington, and V. A. Zakian. 1990. Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* **63**:751–762.
17. Hecht, A., T. Laroche, S. Strahl-Bolsinger, S. M. Gasser, and M. Grunstein. 1995. Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell* **80**:583–592.
18. Henikoff, S. 1998. Conspiracy of silence among repeated transgenes. *Bioessays* **20**:532–535.
19. Heslop-Harrison, J. S., M. Murata, Y. Ogura, T. Schwarzacher, and F. Motoyoshi. 1999. Polymorphisms and genomic organization of repetitive DNA from centromeric regions of *Arabidopsis* chromosomes. *Plant Cell* **11**:31–42.
20. Hollenberg, S. M., R. Sternglanz, P. F. Cheng, and H. Weintraub. 1995.

- Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol. Cell. Biol.* **15**:3813–3822.
21. Imai, S., C. M. Armstrong, M. Kaerberlein, and L. Guarente. 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403**:795–800.
 22. Irwin, P. A., and D. F. Voytas. 2001. Expression and processing of proteins encoded by the *Saccharomyces* retrotransposon Ty5. *J. Virol.* **75**:1790–1797.
 23. Ivy, J. M., A. J. Klar, and J. B. Hicks. 1986. Cloning and characterization of four *SIR* genes of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **6**:688–702.
 24. James, P., J. Halladay, and E. A. Craig. 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**:1425–1436.
 25. Kalpana, G. V., S. Marmon, W. Wang, G. R. Crabtree, and S. P. Goff. 1994. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* **266**:2002–2006.
 26. Katz, R. A., and A. M. Skalka. 1994. The retroviral enzymes. *Annu. Rev. Biochem.* **63**:133–173.
 27. Ke, N., and D. F. Voytas. 1997. High frequency cDNA recombination of the *Saccharomyces* retrotransposon Ty5: the LTR mediates formation of tandem elements. *Genetics* **147**:545–556.
 28. Kenna, M. A., C. B. Brachmann, S. E. Devine, and J. D. Boeke. 1998. Invading the yeast nucleus: a nuclear localization signal at the C terminus of Ty1 integrase is required for transposition in vivo. *Mol. Cell. Biol.* **18**:1115–1124.
 29. Kennedy, B. K., N. R. Austriaco, Jr., J. Zhang, and L. Guarente. 1995. Mutation in the silencing gene *SIR4* can delay aging in *S. cerevisiae*. *Cell* **80**:485–496.
 30. Kennedy, B. K., M. Gotta, D. A. Sinclair, K. Mills, D. S. McNabb, M. Murthy, S. M. Pak, T. Laroche, S. M. Gasser, and L. Guarente. 1997. Redistribution of silencing proteins from telomeres to the nucleolus is associated with extension of life span in *S. cerevisiae*. *Cell* **89**:381–391.
 31. Kim, J. M., S. Vanguri, J. D. Boeke, A. Gabriel, and D. F. Voytas. 1998. Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res.* **8**:464–478.
 32. Kirchner, J., C. M. Connolly, and S. B. Sandmeyer. 1995. Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. *Science* **267**:1488–1491.
 33. Landry, J., A. Sutton, S. T. Tafrov, R. C. Heller, J. Stebbins, L. Pillus, and R. Sternglanz. 2000. The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc. Natl. Acad. Sci. USA* **97**:5807–5811.
 34. Laurenson, P., and J. Rine. 1992. Silencers, silencing, and heritable transcriptional states. *Microbiol. Rev.* **56**:543–560.
 35. Loo, S., and J. Rine. 1995. Silencing and heritable domains of gene expression. *Annu. Rev. Cell Dev. Biol.* **11**:519–548.
 36. Lustig, A. J. 1998. Mechanisms of silencing in *Saccharomyces cerevisiae*. *Curr. Opin. Genet. Dev.* **8**:233–239.
 37. Malik, H. S., and T. H. Eickbush. 1999. Modular evolution of the integrase domain in the Ty3/Gypsy class of LTR retrotransposons. *J. Virol.* **73**:5186–5190.
 38. Marcand, S., S. W. Buck, P. Moretti, E. Gilson, and D. Shore. 1996. Silencing of genes at nontelomeric sites in yeast is controlled by sequestration of silencing factors at telomeres by Rap1 protein. *Genes Dev.* **10**:1297–1309.
 39. Marshall, M., D. Mahoney, A. Rose, J. B. Hicks, and J. R. Broach. 1987. Functional domains of *SIR4*, a gene required for position effect regulation in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **7**:4441–4452.
 40. Moazed, D., and D. Johnson. 1996. A deubiquitinating enzyme interacts with SIR4 and regulates silencing in *S. cerevisiae*. *Cell* **86**:667–677.
 41. Moazed, D., A. Kistler, A. Axelrod, J. Rine, and A. D. Johnson. 1997. Silent information regulator protein complexes in *Saccharomyces cerevisiae*: a SIR2/SIR4 complex and evidence for a regulatory domain in SIR4 that inhibits its interaction with SIR3. *Proc. Natl. Acad. Sci. USA* **94**:2186–2191.
 42. Moore, S. P., L. A. Rinckel, and D. J. Garfinkel. 1998. A Ty1 integrase nuclear localization signal required for retrotransposition. *Mol. Cell. Biol.* **18**:1105–1114.
 43. Moretti, P., K. Freeman, L. Coodly, and D. Shore. 1994. Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes Dev.* **8**:2257–2269.
 44. Pimpinelli, S., M. Berloco, L. Fanti, P. Dimitri, S. Bonaccorsi, E. Marchetti, R. Caizzi, C. Caggese, and M. Gatti. 1995. Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. *Proc. Natl. Acad. Sci. USA* **92**:3804–3808.
 45. SanMiguel, P., A. Tikhonov, Y. K. Jin, N. Motchoulskaia, D. Zakharov, A. Melake-Berhan, P. S. Springer, K. J. Edwards, M. Lee, Z. Avramova, and J. L. Bennetzen. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**:765–768.
 46. Shafikhani, S., R. A. Siegel, E. Ferrari, and V. Schellenberger. 1997. Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization. *BioTechniques* **23**:304–310.
 47. Smith, J. S., C. B. Brachmann, I. Celic, M. A. Kenna, S. Muhammad, V. J. Starai, J. L. Avalos, J. C. Escalante-Semerena, C. Grubmeyer, C. Wolberger, and J. D. Boeke. 2000. A phylogenetically conserved NAD⁺-dependent protein deacetylase activity in the Sir2 protein family. *Proc. Natl. Acad. Sci. USA* **97**:6658–6663.
 48. Triolo, T., and R. Sternglanz. 1996. Role of interactions between the origin recognition complex and SIR1 in transcriptional silencing. *Nature* **381**:251–253.
 49. Wilhelm, M., M. Boutabout, and F. X. Wilhelm. 2000. Expression of an active form of recombinant Ty1 reverse transcriptase in *Escherichia coli*: a fusion protein containing the C-terminal region of the Ty1 integrase linked to the reverse transcriptase-RNase H domain exhibits polymerase and RNase H activities. *Biochem. J.* **2**:337–342.
 50. Yieh, L., G. Kassavetis, E. P. Geiduschek, and S. B. Sandmeyer. 2000. The brf and TATA-binding protein subunits of the RNA polymerase III transcription factor IIIB mediate position-specific integration of the *gypsy*-like element, Ty3. *J. Biol. Chem.* **275**:29800–29807.
 51. Zaccolo, M., D. M. Williams, D. M. Brown, and E. Gherardi. 1996. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255**:589–603.
 52. Zhang, Z., and A. R. Buchman. 1997. Identification of a member of a DNA-dependent ATPase family that causes interference with silencing. *Mol. Cell. Biol.* **17**:5461–5472.
 53. Zhu, Y., S. Zou, D. Wright, and D. F. Voytas. 1999. Tagging chromatin with retrotransposons: target specificity of the *Saccharomyces* Ty5 retrotransposon changes with the chromosomal localization of Sir3p and Sir4p. *Genes Dev.* **13**:2738–2749.
 54. Zou, S., N. Ke, J. M. Kim, and D. F. Voytas. 1996. The *Saccharomyces* retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. *Genes Dev.* **10**:634–645.
 55. Zou, S., J. M. Kim, and D. F. Voytas. 1996. The *Saccharomyces* retrotransposon Ty5 influences the organization of chromosome ends. *Nucleic Acids Res.* **24**:4825–4831.
 56. Zou, S., and D. F. Voytas. 1997. Silent chromatin determines target preference of the *Saccharomyces* retrotransposon Ty5. *Proc. Natl. Acad. Sci. USA* **94**:7412–7416.
 57. Zou, S., D. A. Wright, and D. F. Voytas. 1995. The *Saccharomyces* Ty5 retrotransposon family is associated with origins of DNA replication at the telomeres and the silent mating locus *HMR*. *Proc. Natl. Acad. Sci. USA* **92**:920–924.