Retrotransposon Target Site Selection by Imitation of a Cellular Protein[∇][†]

Troy L. Brady,[‡] Peter G. Fuerst,[§] Robert A. Dick, Clarice Schmidt, and Daniel F. Voytas*

*Department of Genetics, Development, and Cell Biology, 1035A Roy J. Carver Co-Laboratory,
Iowa State University, Ames, Iowa 50011*

Mobile elements rely on cellular processes to replicate, and therefore, mobile element proteins frequently interact with a variety of cellular factors. The integrase (IN) encoded by the retrotransposon Ty5 interacts with the heterochromatin protein Sir4, and this interaction determines Ty5's preference to integrate into heterochromatin. We explored the hypothesis that Ty5's targeting mechanism arose by mimicking an interaction between Sir4 and another cellular protein(s). Mutational analyses defined the requirements for the IN-Sir4 interaction, providing criteria to screen for cellular analogues. Esc1, a protein associated with the inner nuclear membrane, interacted with the same domain of Sir4 as IN, and 75% of mutations that disrupted IN-Sir4 interactions also abrogated Esc1-Sir4 interactions. A small motif critical for recognizing Sir4 was identified in Esc1. The functional equivalency of this motif and the Sir4-interacting domain of IN was demonstrated by swapping these motifs and showing that the chimeric IN and Esc1 proteins effectively target integration and partition DNA, respectively. We conclude that Ty5 targets integration by imitating the Esc1-Sir4 interaction and suggest molecular mimicry as a general mechanism that enables mobile elements to interface with cellular processes.

The impact that transposable elements have on genome organization and function is determined by the site of integration. For a number of transposable elements, target site choice is nonrandom (8, 10, 16, 43). Whereas some transposable elements recognize specific DNA sequences, integration sites of others are influenced by the epigenetic state of the target, including its transcriptional status. Chromatin effects on target site choice are particularly evident for the long terminal repeat retrotransposons and retroviruses (collectively referred to as retroelements), which copy a retroelement mRNA into cDNA by reverse transcription. The element-encoded integrase (IN) inserts this cDNA into the host genome, the site of which is often dictated by DNA-bound protein complexes.

The influence of chromatin and transcription complexes on target site choice has been most extensively studied for the retrotransposons of yeast. The Tf1 element of *Schizosaccharomyces pombe*, for example, integrates within a narrow window upstream of transcription start sites, consistent with a role for RNA polymerase II transcription complexes in target site choice (7, 48). In *Saccharomyces cerevisiae*, the Ty1 and Ty3 retrotransposons integrate preferentially upstream of genes transcribed by RNA polymerase III (Pol III), such as tRNA and 5S ribosomal genes. Ty1 requires Pol III transcription to integrate within its preferred target, a 700-bp window upstream of Pol III-transcribed genes. Both the TFIIIB subunit

Bdp1 and the chromatin remodeler Isw2 influence the pattern of Ty1 insertions upstream of target genes (6, 18). Ty3 integration site choice, on the other hand, is more precise, occurring within 1 to 3 bp of Pol III transcription start sites. This suggests that Ty3 recognizes a factor closely tied to Pol III transcription (12) and is consistent with the observation that the TFIIIB subunits Brf and TATA-binding protein are required for targeted integration (59).

In contrast to Ty1 and Ty3, the *S. cerevisiae* Ty5 retrotransposon integrates into regions of heterochromatin at the telomeres and silent mating-type loci. Ty5 selects integration sites through a direct interaction between Ty5 IN and the heterochromatin protein Sir4. Essential for Sir4 recognition is a 6-amino-acid motif of Ty5 IN, termed the targeting domain (TD) (58). The TD is phosphorylated, and substitutions in TD that prevent phosphorylation abrogate IN-Sir4 interactions and result in the random integration of Ty5 throughout the genome (17). Sir4 serves as a molecular scaffold at the nuclear periphery and interacts with many proteins, including the hypoacetylated tails of histone H3 (22). The TD recognizes the Sir4 C terminus (amino acids [aa] 951 to 1358), a region of the protein that encodes a coiled-coil domain with lamin-like heptad repeats (19). The recognition of Sir4 by IN is sufficient to mediate targeted integration: Ty5 integration occurs with high efficiency at DNA sites to which Sir4 is ectopically tethered (61). Furthermore, Ty5 target specificity can be changed by replacing TD with peptide motifs that recognize other DNA-bound protein partners (61).

Recognition of chromatin by integration complexes also appears to underlie retroviral target site choice. Both human immunodeficiency virus (HIV) and murine leukemia virus integrate preferentially at sites of RNA Pol II transcription (39, 44, 56), and HIV integration favors sites with transcription-associated chromatin modifications (55). For HIV, IN is the primary virally encoded determinant of integration site choice (32), and lens epithelium-derived growth factor (LEDGF)/p75, a chromatin-bound cotranscription factor, interacts with HIV

* Corresponding author. Mailing address: 1035A Roy J. Carver Co-Laboratory, Department of Genetics, Development, and Cell Biology, Iowa State University, Ames, IA 50011. Phone: (515) 294-1963. Fax: (515) 294-7155. E-mail: voytas@iastate.edu.
‡ Present address: Department of Microbiology, University of Pennsylvania School of Medicine, 402 Johnson Pavilion, 3610 Hamilton Walk, Philadelphia, PA 19104.
§ Present address: The Jackson Laboratory, Burgess Research Group, Bar Harbor, ME 04609.

IN and impacts target site preference. The absence or reduction of LEDGF levels in vivo significantly decreases HIV integration frequencies and target site bias (14, 35, 47, 51). Furthermore, HIV integration can be directed in vitro to sites of ectopically tethered LEDGF (15). Other chromatin-associated proteins also influence HIV integration. Barrier-to-autointegration factor (BAF) is an essential protein associated with chromatin structure and nuclear assembly (30, 45, 60) and prevents the autointegration of retroviral cDNA (31). Therefore, chromatin is emerging as a critical player in retroelement integration and target site choice.

In all cases examined to date, retrotransposons and retroviruses recognize aspects of chromatin or components of transcription complexes that serve critical cellular functions. This suggested to us that retroelements may mimic cellular factors that normally recognize DNA-bound protein complexes. We explored this hypothesis using Ty5 and specifically sought a cellular protein that interacts with Sir4 in a manner similar to that of Ty5 IN. Here, we demonstrate that Ty5 IN and Esc1, a protein associated with the nuclear periphery, are virtually equivalent in terms of their interactions with Sir4. We conclude that Ty5 targets integration to heterochromatin by mimicking the Esc1-Sir4 interaction.

MATERIALS AND METHODS

Yeast strains. YPH499 or isogenic strains with *esc1*, *ku70*, or *sir4* deletions were used in the tethered targeting and mitotic stability assays. Deletions were constructed by the one-step gene knockout method (5) using hygromycin B or G418 cassettes from pFA6a-hphNT1 and pFA6a-KANMX, respectively (28, 54). The yeast two-hybrid reporter strains L40 and PJ69 were used for measuring protein-protein interactions with LexA and Sir4 and GBD and Esc1 proteins, respectively (24, 27). Strains YSB1, YSB2, YSB35, and YSB41, used for monitoring the nucleation of heterochromatin, were described elsewhere previously (13). All cultures were grown at 30°C unless noted otherwise.

Plasmid constructs. All LexA-Sir4 constructs were made by PCR amplification of the relevant coding sequences (aa 950 to 1358) from a previously described LexA-Sir4 plasmid (3). PCR products were then digested with EcoRI/BglII and ligated into the EcoRI/BamHI sites of the LexA expression construct pBTM116 (40). A fusion between the Gal4 activating domain (GAD) and Sir4 (aa 950 to 1358) was made by moving the Sir4-containing EcoRI/PstI fragment from pYZ127 into the corresponding sites of pYZ277, a *TRP1*-marked version of pGAD-C1 (27). pPF204 (GAD-Esc1 aa 1361 to 1658) was isolated from the yeast two-hybrid screen. GAD-Esc1 (aa 1440 to 1473) was made by PCR amplification of pPF204 (GAD-Esc1 aa 1361 to 1658) and ligation of the PCR product into the EcoRI/BamHI sites of pGAD-C1. Mutant versions of the construct at aa 1440 to 1473 were made by primer extension of partially overlapping primers, followed by digestion and ligation into pGAD-C1. Other GAD-Esc1 constructs (aa 1443 to 1455 and aa 1448 to 1456) were made by annealing complementary primers and ligating them into the EcoRI/BamHI sites of pGAD-C1. GBD-Esc1 plasmids were generated by digesting GAD-Esc1 plasmids with EcoRI/EcoRV and ligating the resulting fragment into the EcoRI/MscI sites of pGBD-U (27). Chimeric Ty5 and Esc1 sequences were made using four-primer PCR to generate the desired coding sequence. For Ty5 constructs, PCR products were introduced into the BspEI/PflMI sites of Ty5 on pTB60, a pDR14 derivative with an Arg-Gly-Ser-His$_6$ epitope tag within IN (21, 25). For chimeric Esc1 constructs, PCR products were ligated into the ApaI sites of pDZ45, a centromeric plasmid containing the entire Esc1 reading frame that complements *esc1*Δ cells (2). For the tethered targeting assay, target plasmids with either zero or four LexA operators were created by swapping the *TRP1* marker from pYZ316 and pYZ317 with *LEU2* (61). For the reverse two-hybrid screens, a plasmid suitable for generating the LexA-Sir4 mutant library by recombination was made by substituting silent mutations into LexA-Sir4 aa 950 to 1358 at residues K971 and R1331 to introduce SacI and PpuMI sites, respectively. All plasmids constructed using PCR were verified by sequencing. Primer sequences are available upon request.

Analysis of protein function in yeast. Mutagenic PCR was used to create the mutant LexA-Sir4 library used for the reverse two-hybrid screens (11, 53). Briefly, Sir4-His$_6$ was PCR amplified in a 100-μl reaction mixture containing final concentrations of 7 mM MgCl$_2$, 0.5 mM MnCl$_2$, 0.2 mM dATP and dGTP, and 1 mM dCTP and dTTP. Mutagenized Sir4 fragments were cotransformed with a gapped pCS439 plasmid (SacI/PpuMI) using a yeast strain expressing GAD-IN (58). Transformants were stamped after 2 days of growth on selective medium lacking histidine to identify colonies in which Sir4 and IN no longer interacted. Candidates were retested, and those colonies in which Sir4 and IN were still unable to interact were subjected to colony PCR to determine whether they contained a Sir4 insert. The remaining pool of candidates was then screened by immunoblot analysis to detect Sir4 proteins that were unstable or truncated due to nonsense mutations. Plasmids were rescued from yeast by the glass bead method (5) and sequenced.

A previously described assay was used to measure the mitotic stability of unstable plasmids to which various LexA-Sir4 proteins were tethered (3). Cells were grown in 96-well plates until cultures reached an optical density at 600 nm of 3.0 to 4.0 (36 to 48 h). The effect of various chimeric Esc1 proteins on mitotic stability was tested in a similar manner in an *esc1*Δ strain. For these experiments, wild-type Esc1 or the Esc1 chimeras were encoded on a *LEU2* plasmid. The targeted integration of Ty5 to sites of tethered Sir4 was measured as described previously (9).

RESULTS

Defining the region of Sir4 that interacts with IN. We previously demonstrated that Ty5 IN interacts with the C terminus of Sir4 (residues 951 to 1358) and that N-terminal truncations beyond residue 971 impair interactions with IN (Fig. 1A) (61). To define the C-terminal boundary of the IN-interacting domain, yeast two-hybrid assays were performed with GAD-IN and various Sir4 C-terminal truncations fused to LexA. The coiled-coil domain of Sir4 was not required for the interaction, because the LexA fusion at aa 951 to 1251 still interacted with IN (Fig. 1A) even though it lacks the coiled-coil domain (residues 1271 to 1346) (41). Truncations past aa 1082 resulted in a low level of expression of the Sir4 fusion proteins (see Fig. S1A in the supplemental material), leading us to designate the minimal IN-interacting domain of Sir4 aa 971 to 1082.

Mutations in two Sir4 residues in the IN-interacting domain (W974 and R975) were previously shown to abrogate Sir4-IN interactions (61) (Fig. 1B). Four additional residues were identified by alanine-scanning mutagenesis of the region spanning aa 976 to 990 (Fig. 1B; see Fig. S1B in the supplemental material). The relatively large number of critical residues identified (6 out of 19 tested) suggested that the IN-interacting domain may be more extensive, and so a reverse two-hybrid screen was performed by randomly mutagenizing a DNA fragment encoding the Sir4 C terminus (aa 951 to 1358) and screening for variants that failed to interact with IN. Forty-one mutants with single or multiple amino acid substitutions were recovered, 90% of which contained at least one missense mutation in the minimal IN-interacting domain (see Table S1 and Fig. S1C in the supplemental material).

Six single point mutants and three double mutants were selected for further analyses. The protein expression levels of these mutants were comparable to levels of the wild-type protein (see Fig. S1D in the supplemental material). In Fig. 1B, the mutations in the six single and three double mutants are mapped onto an amino acid sequence alignment of Sir4 homologues from different yeast species. All of the single mutations reside within the minimal IN-interacting region. T957L and K1123R both lie outside this domain; however, each one was recovered in a double mutant, and we did not determine if the amino acid change responsible for the phenotype was within the IN interaction domain. All mutants were tested for
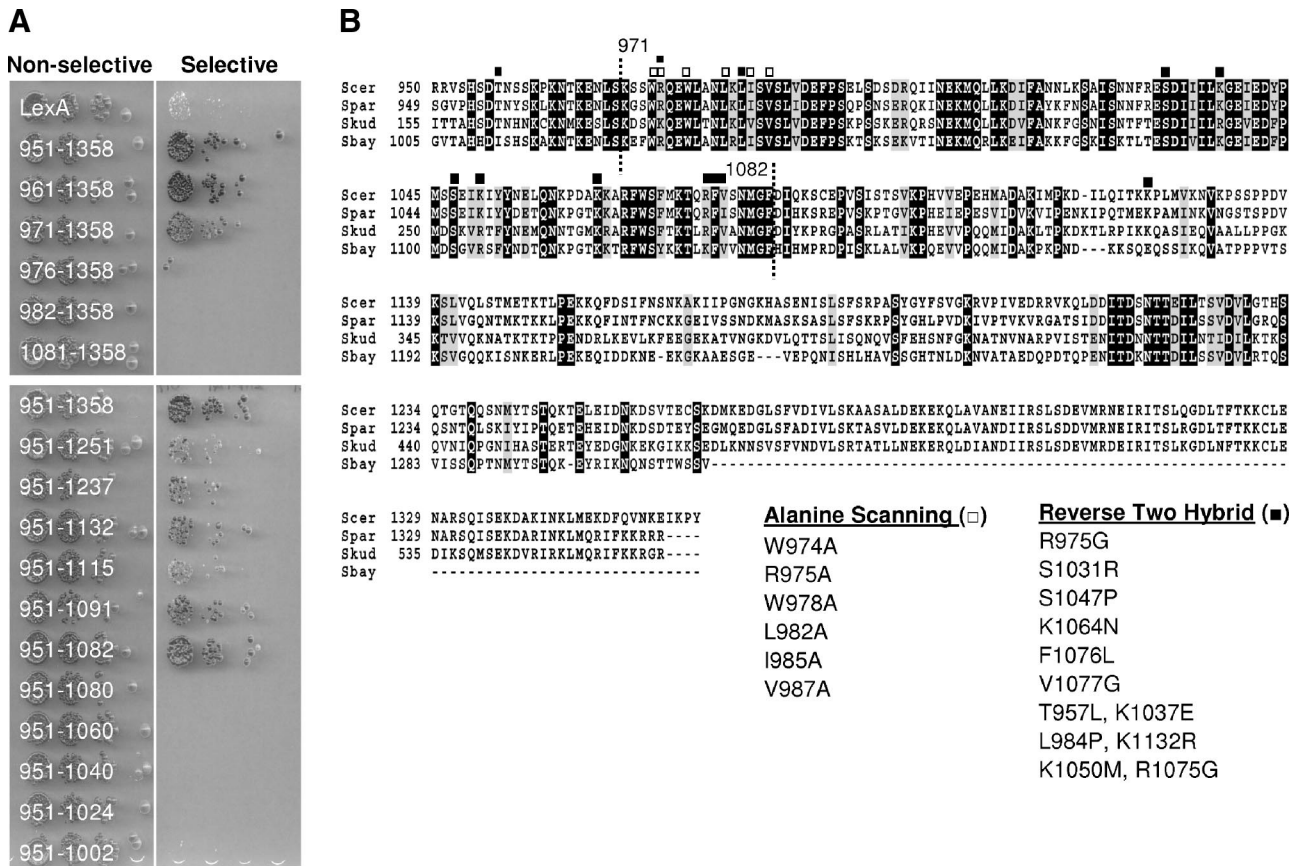
FIG. 1. A conserved region of Sir4 interacts with IN. (A) Delineation of the N- and C-terminal boundaries of the region of Sir4 that interacts with IN. Yeast two-hybrid assays were performed to measure interactions between GAD-IN and various Sir4 fragments fused to LexA. Cultures were serially diluted prior to spotting onto nonselective (synthetic complete minus leucine and tryptophan) and selective (synthetic complete minus leucine, tryptophan, and histidine plus 3 mM 3-amino 1,2,4-triazol) media. (B) The locations of Sir4 mutations are indicated on an alignment of Sir4 homologues from different *Saccharomyces* species. The area between the vertical dashed lines designates the boundaries of the IN-interacting domain of Sir4 (aa 971 to 1082). Amino acid substitutions are indicated by filled or open squares (■ and □), representing mutants isolated by the reverse two-hybrid screen or alanine scanning, respectively. Amino acid substitutions for each mutation are listed below the alignment. The GenBank accession numbers of the sequences used are AACG02000139, AACI02000029, AABY01000107, and NC_001136.

their abilities to interact with Sir4 in *sir2Δ*, *sir3Δ*, and *sir4Δ* strains with little or no observable effect, suggesting that the interaction is direct. Interestingly, all of the mutations are in residues that are conserved in at least three of the four Sir4 homologues, and 7 of the 11 residues in the single mutants are highly hydrophobic, implying that the IN interaction is mediated by a hydrophobic interface or that these residues are important for the proper folding of the protein. Based on the data from both mutagenesis approaches, we conclude that IN interacts with a discrete domain of Sir4. Furthermore, the amino acid sequence conservation suggests that this domain is likely to be important for Sir4 function.

Sir4 mutants affect targeted integration. Because targeting occurs as a consequence of the IN-Sir4 interaction, mutations in Sir4 that abrogate interactions with IN might affect Ty5 target specificity. To test this, we measured Ty5 integration specificity using a previously established targeting assay that measures integration in the vicinity of LexA operators to which the various mutant Sir4 proteins are tethered (61) (Fig. 2A). In this assay, a subset of Ty5 insertions occur near LexA operators on a target plasmid, and the ratio of target plasmids with

Ty5 to those without Ty5 provides a measure of targeting efficiency. Of the nine mutants tested in this assay, the majority showed a decrease in targeting efficiency of greater than threefold (Fig. 2B). Several Sir4 alleles still supported targeting, however, suggesting that endogenous Sir4 may affect targeting through the formation of complexes with the LexA-Sir4 fusions (see below) or that residual IN-Sir4 interactions that were undetectable by the two-hybrid assay were sufficient to target Ty5.

Ty5 insertions mediated by interactions between Sir4 and IN would be expected to cluster around the LexA operator, and so sites of Ty5 integration in the Sir4 mutants were determined by DNA sequencing and compared to the integration pattern observed with wild-type Sir4. Between 7 and 11 target plasmids with Ty5 insertions were analyzed for each of the nine Sir4 mutants (Fig. 2C). The majority of insertions mirrored the wild-type pattern, occurring within a window ±100 bp of the LexA operators. The S1047P allele, which supported targeting at background levels, displayed the only altered integration pattern, with 4 of 11 insertions occurring near the origin of replication on the target plasmid (data not shown). In a *sir4Δ*

A

B

▽ Denotes insertions with mutant Sir4
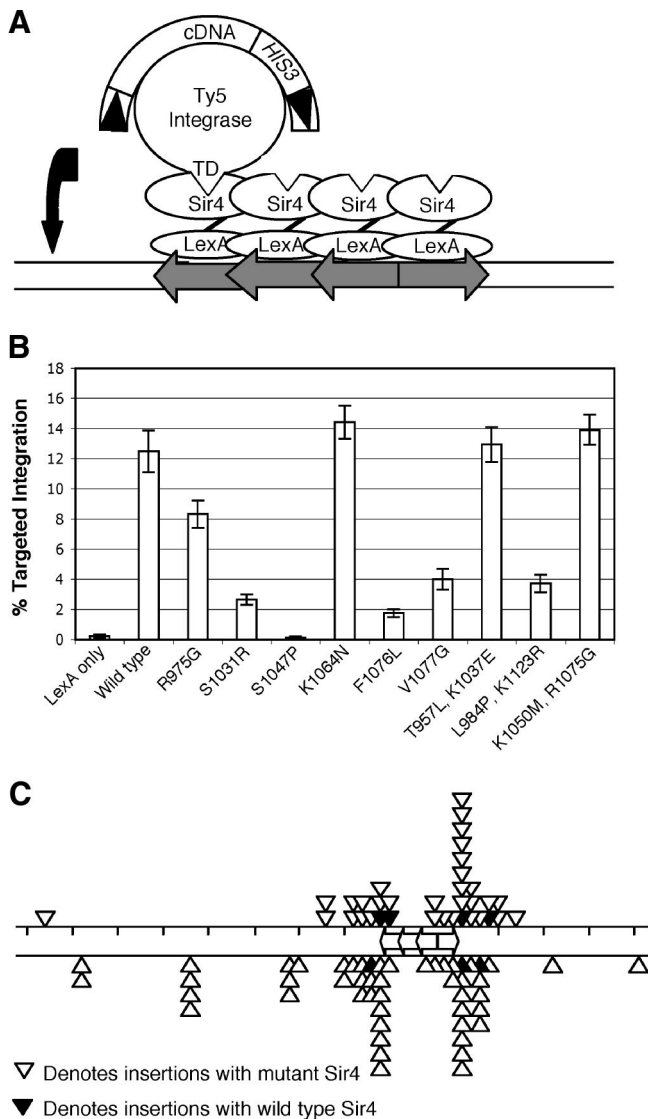
▼ Denotes insertions with wild type Sir4

FIG. 2. Targeted integration mediated by Sir4 mutants. (A) Tethered integration assay (61). The target plasmid carries LexA operators that bind LexA-Sir4, which in turn interact with the targeting domain of the Ty5 IN complexed to its cDNA. The target plasmid has selectable markers and replication origins that function in *Escherichia coli* (chloramphenicol resistance gene) (ColE1). After the induction of Ty5 transposition, target plasmids are recovered in *E. coli* by selecting either for chloramphenicol resistance (Chl^r) or for chloramphenicol resistance and histidine prototrophy (Chl^r His$^+$). The latter selection determines the number of target plasmids that carry Ty5 with its *HIS3* marker. The ratio of Chl^r His$^+$ to Chl^r indicates the efficiency with which Ty5 integrates into the plasmid. (B) Six of nine Sir4 mutants exhibited significant defects in targeted integration compared to the wild type. (C) Mutant Sir4 proteins tethered to LexA operators target Ty5 insertions in a manner similar to that of the wild type. LexA operators are indicated by arrows, and Ty5 insertions are shown in either the forward (▼) or reverse (▲) orientation.
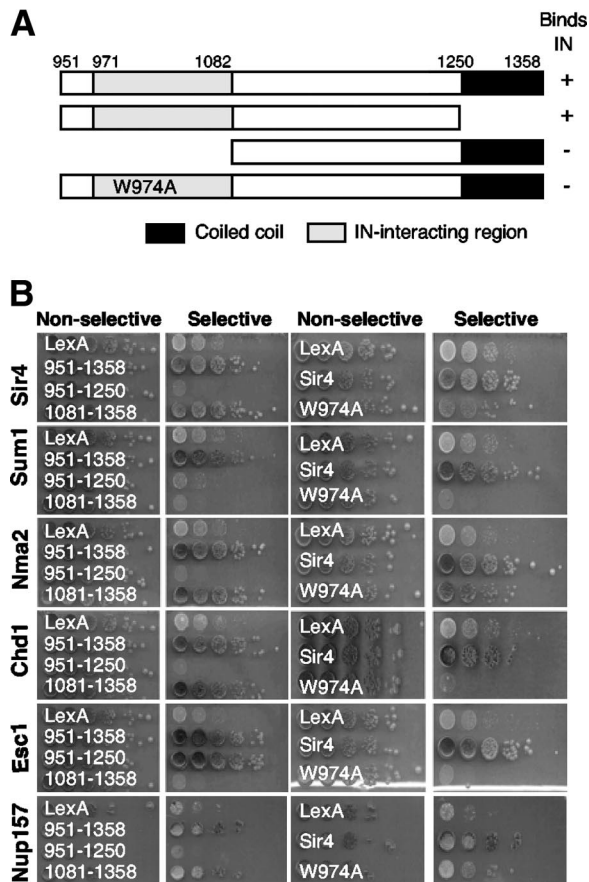

FIG. 3. Interaction profiles of Sir4-interacting proteins. (A) The various C-terminal fragments of Sir4 tested for two-hybrid interactions (B) are shown in graphic form with labeled interaction domains. For reference, the interaction of these constructs with IN is shown to the right. (B) Yeast two-hybrid assays measuring interactions of candidate proteins with the LexA-Sir4 constructs shown in panel A. Cultures were serially diluted and spotted onto medium as described in the legend to Fig. 1.

strain, this altered integration pattern was no longer observed (zero of seven insertions characterized) (data not shown). No significant changes in targeting efficiencies were observed for the other mutants in a *sir4Δ* strain, except for the F1076L mutant, which showed a sixfold decrease in targeting (data not

shown). These data indicate that endogenous Sir4 can contribute to both target specificity and integration efficiency in some instances where interactions between IN and the LexA-Sir4 fusion are attenuated.

Identifying a host analogue of IN. If the conserved, IN-interacting domain of Sir4 is critical for its function, then this domain may interact with other proteins to carry out its biological activity. To identify such proteins, a yeast two-hybrid screen was carried out using LexA-Sir4C (aa 951 to 1358) as bait. The screen identified six proteins: Sum1, Chd1, Nma2, and Nup157 (none of which were previously shown to bind Sir4), Esc1 (a known Sir4-interacting protein) (2), and Sir4 itself. Sum1 is a transcriptional repressor of middle-sporulation-specific genes and is involved in telomere maintenance (4, 57). Chd1 is a chromatin remodeler (50), and Nma2 (nicotinic acid mononucleotide adenylyltransferase) is involved in the NAD$^+$ salvage pathway (20). Nup157 is a member of the nucleoporin protein family (1), and Esc1 is important for the assembly of the nuclear pore complex (33) and tethers telomeres to the nuclear periphery (49). All of the proteins interacted with LexA-Sir4 in *sir2Δ*, *sir3Δ*, and *sir4Δ* strains
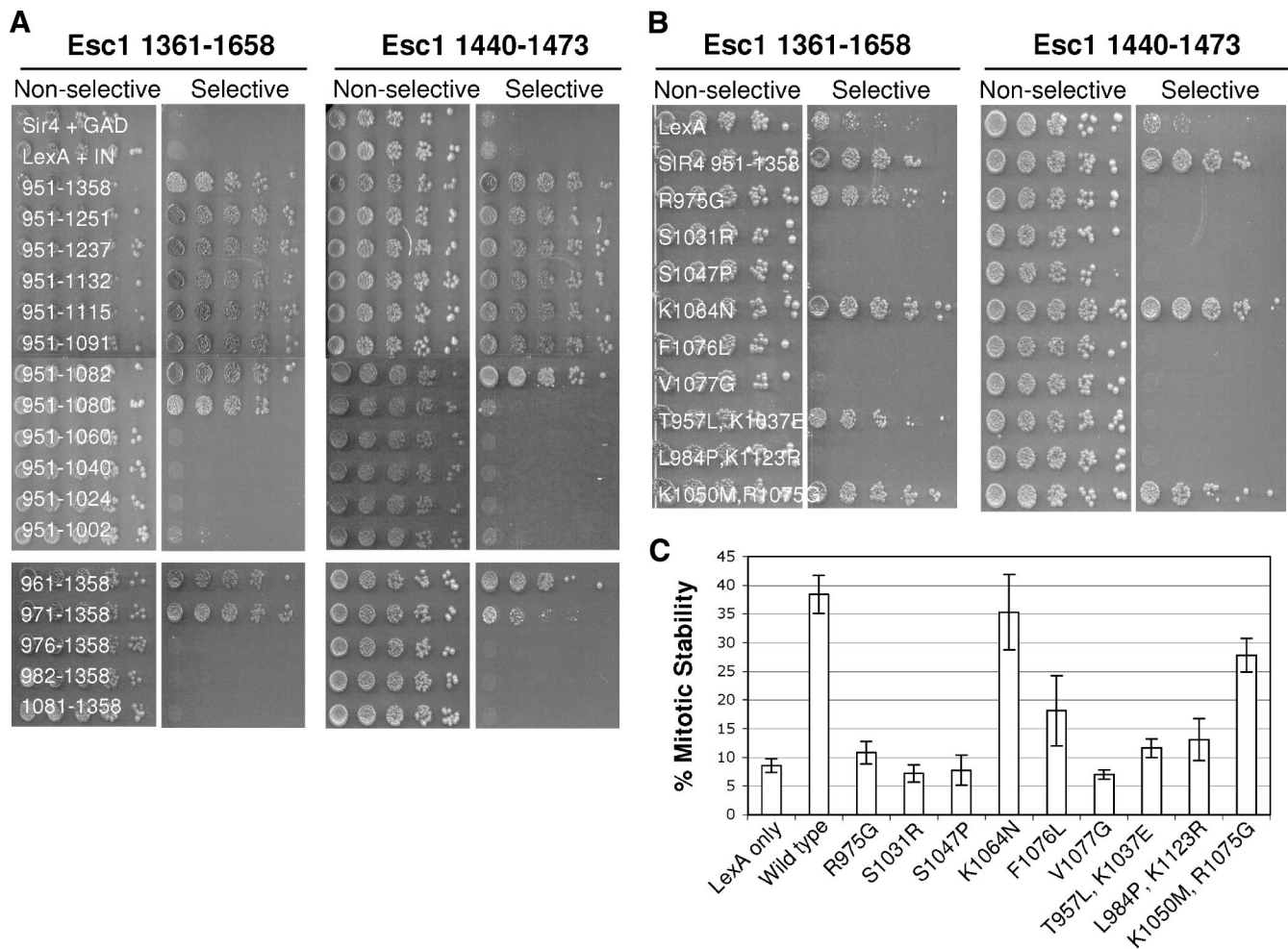
FIG. 4. Esc1 interacts with Sir4 in a manner similar to that of IN. (A) The domains of Sir4 that interact with Esc1 aa 1361 to 1658 and aa 1440 to 1473 span Sir4 residues 971 to 1080 and 961 to 1082, respectively. (B) The majority of Sir4 mutants that fail to interact with IN also fail to interact with either Esc1 fragment. Cultures in both panels were serially diluted and spotted as described in the legend to Fig. 1. (C) Sir4 mutants that fail to interact with Esc1 aa 1440 to 1473 also fail to reliably partition DNA. Mitotic stability refers to the number of cells retaining both the LexA-Sir4 and unstable reporter plasmids divided by those with a LexA-Sir4 plasmid only.

(data not shown), suggesting that the interaction between these proteins and Sir4 is direct.

To determine whether IN mimics one of the identified cellular proteins in its interaction with Sir4, we tested these proteins for interactions with selected Sir4 deletion constructs and the W974A mutant to assess whether their interactions resembled those of GAD-IN (Fig. 3A). In contrast to IN, Sum1 required the entire Sir4 C terminus for the interaction, and Nma2, Chd1, Nup157, and Sir4 required only the coiled-coil domain. As reported previously, Esc1 interacted with the Sir4 C terminus in the absence of the coiled-coil domain but, like IN, required the upstream N-terminal region (2). The W974A mutation abrogated interactions with Sum1, Chd1, and Esc1 but not interactions with Nma2, Nup157, or Sir4 itself (Fig. 3B). Based on these results, we concluded that Esc1 shares the most similarity with IN in terms of its Sir4 interactions.

Esc1 interacts with the same region of Sir4 as IN. To further characterize the similarity between the IN-Sir4 and Esc1-Sir4 interactions, Esc1-Sir4 two-hybrid assays were conducted with all of the Sir4 truncations and point mutants tested with IN.

Two fragments of Esc1 were used: a GAD-Esc1 C-terminal fragment (aa 1361 to 1658) isolated in our two-hybrid screen and a 34-amino-acid region of Esc1 (aa 1440 to 1473) reported previously to interact with Sir4 (2). Esc1 from aa 1361 to 1658 failed to interact with N-terminal truncations past aa 971 and C-terminal truncations preceding aa 1080. Esc1 from aa 1440 to 1473 had similar requirements, with N- and C-terminal boundaries of aa 961 and 1082, respectively (Fig. 4A).

Among the LexA-Sir4 mutants tested, 10 of 15 mutants failed to interact with the full-length Esc1 C terminus (Fig. 4B and data not shown). Of the six mutants that interacted, three (R975G, T957L/K1037E, and V987A) exhibited slow growth, indicating a weakened interaction. This effect was more evident on media containing higher concentrations of 3AT, an inhibitor of the *HIS3* reporter gene (data not shown). The smaller Esc1 fragment interacted only with K1064N and K1050M/R1075G mutants, both of which showed strong interactions with the larger Esc1 C terminus. These residues are therefore important for interactions with IN and are not critical for interactions with Esc1. Collectively, the data show that

A

```
                                    1440            1473
Scer 1401 ERSIEELNSEPEEAELYELEIEGPTETAASSKMNDDERQRGNIPSTDLPSDPPSDKEEVTDSYPYSNSENITAEKSAPTSPEVYEIFSDTPNEVP
Spar 1439 ERSIEDLHSEPEEDELFELERQVPTGIAVFGKSDDEERQRGTIPSTDLPSDPPSDREESTGSHTYSNSENASAGKGVPTSPEVYEIFSDTPNELP
Smik    2 DKSIDELHSEPEEGELFVLESKGPTQKETFCKSDVDGEQLG-IPSTDLPSDPPSDREGTAGSYANLSDNAATEMNVRTSPEVYEIFSDTPNEAP
Skud    1 EKSIEXHSEPEGEKLFALDMQTPTQTAAFAKSDDEEKQRGVIPSTDLPSDPPSDTEETTDSYTNLKMESKSPEVDAPISQEVYEILSDTPNEVL
Sbay 1527 EKSIDALHSEPEEEQFDFETQTPLETAASGESDGEEQIYGIIPSTDLPSDPPSDTEETADLYNNSNTESMEVETNSFASPEVYEISSDTPNEVP

Scer 1497 MEINDEIPATTLEKHDKTNVTSVLDDRSEHLSSHDVDNEPHDNSINIKVNEGEE-PEHQAVDIPVKVEVKEEQEMPSKSVLEEQKPSMELINDK
Spar 1535 MEINNELPNTSLGKDDKTTATSVLDDRLEDSSSNDK------NFINIKVNEGEEGSEHKAVDIPIEVDVKEEQEIPPNPVPEEQKLNEGLFNDK
Smik   97 IEIGDQSPNSVLGKGNATTMTPVLDGKSEDTVNHDVVNEASDNFINIKVTEAEE-PBAQAVDIPIEVDVKEEQEKMSSNSNLDEQKLNTEQSNDV
Skud   97 LETNDELSSAMLEKDNKGTKTPTLDDASEDSSYYNISHKSGRNSTSIKVDEGEG-VEFRAEDIPIEIEIEEEQEKMPSNSTLEGNKPKAE-SNNE
Sbay 1623 VEAGDKQLNAMSGKDSSMTTASILGVGSDDSPSHDVVNRTEEKSVDIKINEDGE-SESRAIDIPVEVYGKEEREEKPSDAKLEKEELKLD---DE

Scer 1591 SSPENNNDEETNREKDKTKAKKKSRKRNYNSRRRKRKITEGSSAASNTKRRRGHEPKSRGQNTHPSVDK
Spar 1624 GGLGSNNDEEINREEDKSKAKKKSRKRNYNSRRRKRKITEGSSAASNTKRRKRHEPRSRGKKTDSSVNK
Smik  191 EELGINSAEEINRGGDKVKAKKKSRKRNYNNRRRKRKITEDSPATSNSKRLKKHEAKERGKNTHRSANK
Skud  190 ESLGINDAEVIGPEEEESKTKKKSRKRNYNSRRRKRKITEDSSTNSRSKRSKGHGVKARKKNARSSGSR
Sbay 1714 EILDMKEAEKINLEESKSKPKKKSRKRNYNSRRRKRKITEDSGPGTRAKRSKGKDARPGGSK-------
```

B

```
                        *   *  **
Esc1    ...TDLPSDPPSD...
Ty5 IN  ...PSLDSSPPNT...
```

C

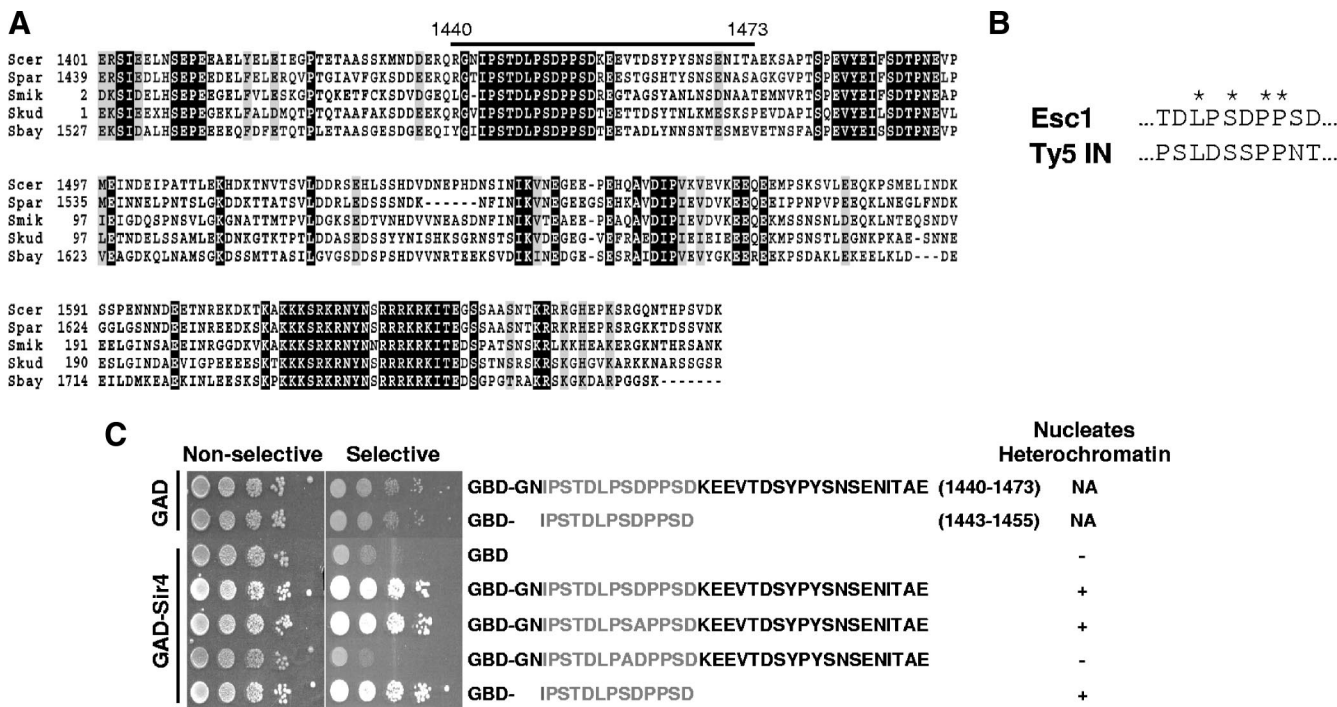|					Nucleates Heterochromatin
GAD	Non-selective	Selective	GBD-GNIPSTDLPSDPPSDKEEVTDSYPYSNSENITAE	(1440-1473)	NA
			GBD- IPSTDLPSDPPSD	(1443-1455)	NA
GAD-Sir4			GBD		-
			GBD-GNIPSTDLPSDPPSDKEEVTDSYPYSNSENITAE		+
			GBD-GNIPSTDLPSAPPSDKEEVTDSYPYSNSENITAE		+
			GBD-GNIPSTDLPADPPSDKEEVTDSYPYSNSENITAE		-
			GBD- IPSTDLPSDPPSD		+

FIG. 5. Esc1 encodes a conserved TD-like motif that interacts with Sir4. (A) Alignment of Esc1 homologues from various *Saccharomyces* species. Amino acid residues 1440 to 1473 of Esc1 are indicated above the alignment. Esc1 homologues were analyzed as described in the legend to Fig. 1. GenBank accession numbers of sequences used are AABY01000127, AABZ01000549, AACI02000321, AACG02000021, and NC_001145. (B) Amino acid sequence alignment of the Ty5 TD and the core of the conserved domain of Esc1. Asterisks indicate identical residues. (C) Yeast two-hybrid assays testing for interaction betweens GAD-Sir4 and various GBD-Esc1 fusion proteins. The Esc1 sequence in each GBD fusion is indicated to the right of the yeast strains. The text of the conserved amino acid sequence domain is gray. The ability of these constructs to silence an adjacent *TRP1* marker gene is indicated at the far right (see also Fig. S2B in the supplemental material). Cultures were serially diluted prior to spotting onto nonselective (synthetic complete minus tryptophan and uracil) and selective (synthetic complete minus tryptophan, uracil, and histidine plus 3 mM 3-amino 1,2,4-triazol) media. NA, not applicable.

Esc1 interacts with Sir4 in a fashion similar to that of IN and that the 34-amino-acid Esc1 fragment (aa 1440 to 1473) encodes a strong Sir4-interacting region that behaves like the larger Esc1 C terminus.

Sir4 mutants affect DNA partitioning by Esc1. The region of Sir4 that interacts with IN and Esc1 is involved in partitioning DNA to daughter cells during mitosis (3). This domain, referred to as the partitioning and anchoring domain (PAD), spans residues 950 to 1262. Partitioning of DNA by the PAD was demonstrated in assays in which a LexA-Sir4 PAD fusion was tethered to an otherwise unstable plasmid. The interaction of the Sir4 PAD with Esc1 is necessary for the plasmid to be partitioned to daughter cells, and partitioning is measured by the retention of a plasmid marker after cell division. While the full PAD is required for optimal partitioning, a smaller region (aa 950 to 1150) retains partial activity even in the absence of endogenous Sir4 (3). The minimal PAD roughly corresponds to the IN-interacting domain of Sir4.

To assess whether mutations that abrogate IN-Sir4 interactions affect Esc1 function in vivo, nine Sir4 mutants were tested for their ability to support DNA partitioning. The two mutants that interacted strongly with both Esc1 from aa 1361 to 1658 and aa 1440 to 1473 (K1064N and K1050M/R1075G) displayed the highest DNA-partitioning activity (compare Fig. 4B and C). The remaining seven mutants that showed little or no interaction with Esc1 from aa 1440 to 1473 had reduced levels of partitioning, which were significantly different from that of the wild type (Fig. 4C). The ability of these Sir4 mutants to engage in yeast two-hybrid interactions correlates well with their ability to efficiently partition plasmids. The data further suggest a role for aa 1440 to 1473 of Esc1 in DNA partitioning and support the hypothesis that IN imitates Esc1 in its interaction with Sir4.

The ability of the Sir4 mutants to interact with Esc1 in two-hybrid assays correlated well with their ability to target Ty5 integration (compare Fig. 2B and 4B). To test whether Esc1 or other associated proteins in the nuclear periphery affect Ty5 target specificity, targeted transposition assays were carried out in single- and double-deletion strains lacking Esc1, Sir4, and Ku70. We did not detect any effect on targeted integration in any of these deletion strains, indicating that these proteins either alone or in combination are not required for integration in the tethered targeting assay (see Table S2 in the supplemental material).

Similarities between the C termini of Ty5 IN and Esc1. In addition to having comparable Sir4 interaction profiles, the C termini of IN and Esc1 share amino acid sequence similarities. First, the Sir4-interacting domain of Esc1 (aa 1395 to 1551) is serine and proline rich (11.9% serine and 9.3% proline), a feature shared with the Ty5 IN C terminus (aa 934 to 1131, with 11.6% serine and 7.1% proline). This amino acid sequence composition is well above the average for proteins in

A

Ty5	AA sequence	% Targeted Integration

B

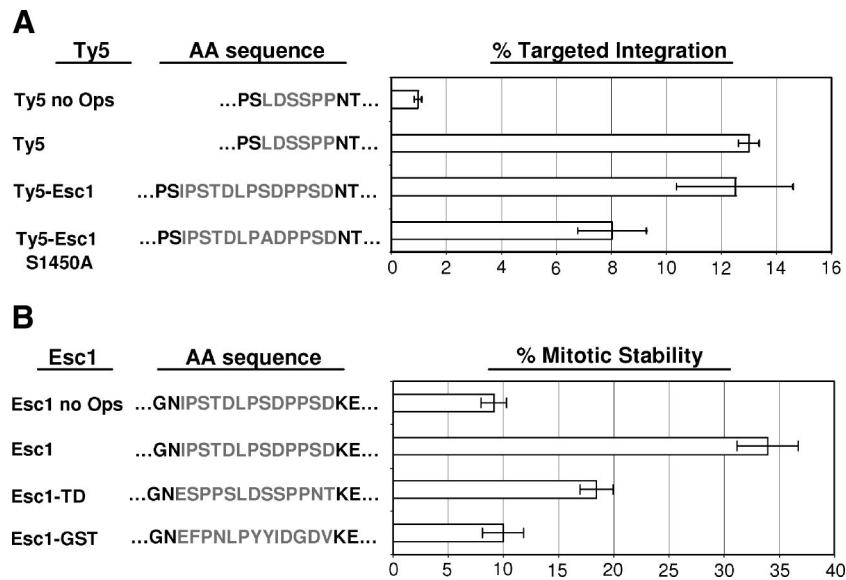| Esc1	AA sequence	% Mitotic Stability

FIG. 6. Targeting and DNA-partitioning activities of chimeric Ty5 elements and Esc1 proteins. For both panels, the amino acid sequences exchanged in each chimera are shown to the left of the chart. (A) Ty5 elements containing wild-type or mutant Esc1 sequences in place of the TD were tested for their abilities to target integration using the tethered targeting assay. The Ty5 chimeras transposed at wild-type levels (data not shown). (B) DNA partitioning efficiencies of chimeric Esc1 proteins that contain Ty5 TD or GST sequences. The TD partially restored partitioning, whereas the GST sequence did not.

the UniProtKB database (254,609 sequence entries), which average 6.8% serine and 4.8% proline. Second, an alignment of Esc1 homologues from five different yeast species revealed a stretch of 13 highly conserved residues within the region spanning aa 1440 to 1473, suggesting that this Sir4 interaction domain plays an important role in Esc1 function (Fig. 5A). Although there is little significant similarity in terms of amino acid sequence signatures between the Esc1 and Ty5 IN domains, we did notice that Esc1 encodes a sequence motif within the conserved block (aa 1448 to 1453, LPSDPP) that has three of the four residues in the Ty5 TD that are required for targeted integration (Fig. 5B).

To assess the functional relevance of the TD-like motif of Esc1, we performed two hybrid assays between GAD-Sir4 and a Gal4 DNA binding domain (GBD) fusion with the Esc1 fragment at aa 1440 to 1473 or two derivatives with alanine point mutations (S1450A and D1451A) (Fig. 5C). Of the proteins tested, only GBD-Esc1 (S1450A) failed to interact with Sir4. All proteins were expressed well, including the S1450A mutant (see Fig. S2A in the supplemental material). Comparable results were produced in a heterochromatin nucleation assay in which the GBD-Esc1 fusion proteins were assayed for their abilities to nucleate heterochromatin at a mutant *HMR* silencer (see Fig. S2B in the supplemental material). The Esc1 fragments, with the exception of the S1450A mutant, nucleated heterochromatin in a fashion similar to that of a GBD-TD fusion protein (see Fig. S2C in the supplemental material). In summary, the data indicate that a TD-like motif within the conserved Esc1 domain is capable of interacting with Sir4 and that a serine within the motif at position 1450 is critical for the interaction.

Functional equivalence of IN and Esc1 domains. Given the similar requirements for interactions of IN and Esc1 with Sir4 and the existence of small Sir4-interacting motifs in both pro-

teins, we hypothesized that these domains are functionally equivalent and therefore may functionally substitute for one another. To test this, Ty5's TD and adjacent amino acids were swapped with the conserved region of Esc1 (aa 1443 to 1455). An S1450A mutation was introduced into the Ty5-Esc1 chimera and tested along with the wild type for targeted integration using our tethered targeting assay. The chimera targeted as well as did wild-type Ty5, showing that the Esc1 motif can functionally substitute for TD (Fig. 6A). The S1450A mutation, which disrupted interactions with Sir4 in yeast two-hybrid assays and failed to nucleate heterochromatin, significantly reduced targeting when placed in the context of the Ty5 IN.

A reciprocal swap in which Esc1 at aa 1443 to 1455 was replaced with a TD-encoding sequence from Ty5 or an equivalently sized region of glutathione *S*-transferase (GST) predicted to have a structure similar to that of the Esc1 motif was also performed (data not shown). The Esc1 chimeras were introduced into an *esc1*Δ strain on single-copy plasmids and tested for their abilities to partition DNA as described above. The replacement of the conserved region of Esc1 with GST abolished all partitioning activity (Fig. 6B). The TD-containing Esc1 chimera showed an intermediate level of DNA partitioning that was well above levels for the negative control. The ability of these domains to replace one another demonstrates their functional equivalency and supports the hypothesis that Ty5 targets integration by mimicking the interaction between Esc1 and Sir4.

DISCUSSION

Different retroelements use different strategies to populate their host genomes. Retrotransposons often integrate into "safe havens," benign regions of the genome that can withstand insertions without deleterious consequences to host fit-

ness, thereby ensuring that both the retrotransposon and host persist (8, 10, 16, 43). The retroviruses, on the other hand, often integrate preferentially into regions of the genome conducive to transcription, presumably so that progeny viruses can be produced to amplify the infection. How do retroelements evolve their targeting strategies? Here, we present evidence that one mechanism is for IN to mimic a chromatin factor already associated with the desired target sites. For Ty5, targeting integration to heterochromatin is accomplished by mimicking the interaction between Esc1 and the heterochromatin protein Sir4.

Esc1 and IN share functionally equivalent Sir4 interaction domains. Although Esc1 and IN act in very different biological processes, both proteins have short motifs in their C termini that mediate the interaction with Sir4. Importantly, the motifs are functionally equivalent: Esc1 chimeras with the Ty5 TD partition DNA during mitosis, and Ty5 chimeras with the Esc1 motif target integration to sites of tethered Sir4. The amino acid sequence signature of critical residues in the Ty5 targeting domain are strikingly similar to the core of the Esc1 motif (LDSSPP for Ty5 [underlined residues are important for function] and LPSDPP for Esc1 [underlined residues are shared with the TD]). The only difference is the second serine in the TD (S1095), which is an aspartate in Esc1 (D1451). We have recently shown that S1095 of the TD is phosphorylated and that phosphorylation is required for both interactions with Sir4 and targeted integration (17). Furthermore, negatively charged amino acids can functionally replace S1095. This likely explains the ability of the Esc1 motif to serve as a targeting determinant when it is substituted for the TD. The two motifs, however, are not completely equivalent: a D1451A mutation in Esc1, for example, does not affect interactions with Sir4, whereas mutations at the analogous position of TD (S1095) do (17). This difference may reflect the adaptation of this Sir4 binding motif for optimal function in distinct biological processes. Interestingly, Esc1 has been reported to be a phosphoprotein (23), and we found by mass spectrometry that the 34-aa fragment of Esc1 that interacts with Sir4 is multiply phosphorylated (data not shown). The phosphorylation status of Esc1 underscores the similarities between the small Sir4-interacting regions of these two proteins. Whether Esc1 phosphorylation plays a role in its affinity for Sir4 is currently under investigation. It will also be interesting to know if this Esc1 domain has any role in telomere localization and/or nuclear pore function (33, 49).

Although the Ty5 TD and the Esc1 motif are essential for interactions with Sir4 in vivo, they are not the only Sir4-interacting regions of these proteins. Both motifs reside in extended serine/proline-rich domains, a feature shared by some protein interaction domains regulated by phosphorylation (29, 42). Larger fragments of IN show a significantly stronger interaction with Sir4 in yeast two-hybrid assays than does TD alone (58). Similarly, regions of Esc1 other than those described here can nucleate heterochromatin, presumably through their interaction with Sir4 (2). Although other regions of Esc1 and IN appear to bind Sir4, both the Esc1 motif described here and the TD are critical for productive Sir4 interactions. These motifs may serve as communication modules for a larger protein scaffold that directs their respective Sir4 associations.

How did Ty5 acquire its TD? Esc1 and Ty5 IN may have independently evolved a common motif for interacting with

Sir4. Alternatively, and what seems to be a more likely scenario, Ty5 acquired the TD by transducing the motif from Esc1 or a related protein, similar to the mechanism by which oncoviruses acquire oncogenes. Some retrotransposons have chromodomains in their IN C termini (38), which, in cellular genes, recognize specific histone modifications and therefore are logical targeting determinants. It may be that the acquisition of chromatin interaction modules has been commonly employed by retroelements to evolve new target site specificities.

Esc1-Sir4 interaction and Ty5 target specificity. The strategy of tethering integration complexes through mimicry raises interesting questions about the dynamics of the integration process. Because both IN and Esc1 interact with the same region of Sir4, it is possible that these two proteins compete for binding in vivo. However, we saw no transposition or targeting phenotype in *esc1Δ* strains. We also saw no effect in mitotic stability assays when Ty5 was overexpressed or when IN was overexpressed as a GAD or GBD fusion protein (our unpublished observations). Finally, no effect on transposition was observed upon GBD-Esc1 overexpression (our unpublished observations). Previous colocalization studies of Sir4 and an overexpressed GFP-Esc1 fusion protein showed that the two proteins only partially colocalize (2). Although this observation may not reflect the association of these proteins at endogenous expression levels, it is likely that free Sir4 with which both IN and Esc1 can interact is available.

Although Sir4 appears to be the main targeting determinant, several lines of evidence suggest that other factors play a role. For example, the level of targeted integration in *sir4Δ* strains is significantly higher than expected for random integration, suggesting the involvement of targeting cofactors (62). Sir4 interacts with numerous proteins, and we have not exhaustively tested other Sir4-interacting proteins for transposition phenotypes, including the novel Sir4 interactors identified in this study. Among the Sir4 interactors that we did analyze, namely, Esc1, Ku70, and Sir4 itself, we did not observe an effect on Ty5 transposition frequencies or target specificity as measured by our tethered targeting assay (see Table S2 in the supplemental material) (data not shown). However, there is a hint that Esc1 or the Esc1-Sir4 interaction contributes to target specificity: many of the Sir4 mutants that failed to interact with IN retained the ability to target Ty5 integration at wild-type levels, and all of these mutants also interacted with the C terminus of Esc1. A subtle effect on targeting in *esc1Δ* strains may not be revealed by our tethered targeting assays but may be evident by surveying integration patterns of chromosomal insertions.

Also of interest is how integration complexes access DNA during the integration process. One mechanism for Ty5 may be to link integration to certain times in the cell cycle, perhaps when heterochromatin is dismantled or newly established. At the telomeres, heterochromatin formation begins with Rap1 binding to the telomeric repeats, followed by the recruitment of Sir4 (36). At this point, Ty5 could gain access to DNA targets as Esc1 begins to tether the telomeres to the nuclear periphery (Fig. 7). Regulation by phosphorylation could help coordinate the timing of such events, raising the possibility that the IN-Sir4 and Esc1-Sir4 interactions are controlled by the same regulatory pathways. Work to determine the timing of Ty5 integration during the cell cycle, to test for a role for phosphorylation in the Esc1-Sir4 interaction, and to identify
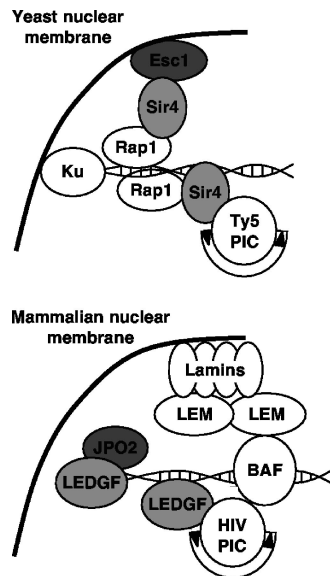
FIG. 7. Models for tethered integration of Ty5 and HIV. Chromatin-associated factors act to target integration through their interactions with IN. PIC indicates the preintegration complex, and LEM represents LEM domain-containing proteins (Lap2, emerin, and Man1), which bind BAF. Shaded light gray indicates Sir4 and LEDGF, which interact with the IN of Ty5 and that of HIV, respectively. Shaded dark gray are Esc1 and JPO2, cellular proteins that interact with Sir4 and LEDGF in a manner similar to those of the Ty5 and HIV IN.

kinases involved in regulating the interactions of these proteins is currently under way.

Retroviral integration and mimicry of host factors. Recent findings demonstrate that HIV also targets integration through a protein-protein interaction between HIV IN and LEDGF/p75, a general transcription factor that is constitutively chromatin bound (52). The depletion of LEDGF blocks HIV integration (35), indicating the importance of chromatin localization in the infection process. Emerin, a protein localized at the nuclear periphery, has also been reported to be required for HIV infectivity (26). Although the role of emerin remains controversial (46), clear parallels can be drawn between Ty5 and HIV integration in terms of the impact of chromatin-bound proteins, some of which localize to the nuclear periphery (Fig. 7).

Another similarity between Ty5 and HIV is the recent finding that LEDGF interacts with the Myc-binding host factor JPO2 (37). This LEDGF-JPO2 association stabilizes JPO2 protein levels, and LEDGF tethers JPO2 to chromatin in a manner analogous to the way it tethers IN (34, 52) (Fig. 7). Mutations that affect the IN binding surface of LEDGF disrupt interactions with JPO2, suggesting that an HIV IN-JPO2 relationship is similar to that of Ty5 IN and Esc1. Thus, mimicry of host factor interactions may be a general phenomenon utilized by retroelements to tether integration complexes to target sites. A further understanding of the role of chromatin in target site selection will likely provide insight into how these elements impact genome organization and may offer new therapeutic opportunities for treating retroviral diseases.

REFERENCES

1. **Aitchison, J. D., M. P. Rout, M. Marelli, G. Blobel, and R. W. Wozniak.** 1995. Two novel related yeast nucleoporins Nup170p and Nup157p: complementation with the vertebrate homologue Nup155p and functional interactions with the yeast nuclear pore-membrane protein Pom152p. J. Cell Biol. **131:** 1133–1148.
2. **Andrulis, E. D., D. C. Zappulla, A. Ansari, S. Perrod, C. V. Laiosa, M. R. Gartenberg, and R. Sternglanz.** 2002. Esc1, a nuclear periphery protein required for Sir4-based plasmid anchoring and partitioning. Mol. Cell. Biol. **22:**8292–8301.
3. **Ansari, A., and M. R. Gartenberg.** 1997. The yeast silent information regulator Sir4p anchors and partitions plasmids. Mol. Cell. Biol. **17:**7061–7068.
4. **Askree, S. H., T. Yehuda, S. Smolikov, R. Gurevich, J. Hawk, C. Coker, A. Krauskopf, M. Kupiec, and M. J. McEachern.** 2004. A genome-wide screen for *Saccharomyces cerevisiae* deletion mutants that affect telomere length. Proc. Natl. Acad. Sci. USA **101:**8658–8663.
5. **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, NY.
6. **Bachman, N., M. E. Gelbart, T. Tsukiyama, and J. D. Boeke.** 2005. TFIIIB subunit Bdp1p is required for periodic integration of the Ty1 retrotransposon and targeting of Isw2p to *S. cerevisiae* tDNAs. Genes Dev. **19:**955–964.
7. **Behrens, R., J. Hayles, and P. Nurse.** 2000. Fission yeast retrotransposon Tf1 integration is targeted to 5′ ends of open reading frames. Nucleic Acids Res. **28:**4709–4716.
8. **Boeke, J. D., and S. E. Devine.** 1998. Yeast retrotransposons: finding a nice quiet neighborhood. Cell **93:**1087–1089.
9. **Brady, T. L., C. L. Schmidt, and D. F. Voytas.** 2007. Targeting integration of the *Saccharomyces* Ty5 retrotransposon, p. 153–164. *In* G. D. Davis and K. J. Kayser (ed.), Chromosomal mutagenesis, methods in molecular biology, vol. 435. Humana Press, Totowa, NJ.
10. **Bushman, F. D.** 2003. Targeting survival: integration site selection by retroviruses and LTR-retrotransposons. Cell **115:**135–138.
11. **Cadwell, R. C., and G. F. Joyce.** 1994. Mutagenic PCR. PCR Methods Appl. **3:**S136–S140.
12. **Chalker, D. L., and S. B. Sandmeyer.** 1992. Ty3 integrates within the region of RNA polymerase III transcription initiation. Genes Dev. **6:**117–128.
13. **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.
14. **Ciuffi, A., and F. D. Bushman.** 2006. Retroviral DNA integration: HIV and the role of LEDGF/p75. Trends Genet. **22:**388–395.
15. **Ciuffi, A., T. L. Diamond, Y. Hwang, H. M. Marshall, and F. D. Bushman.** 2006. Modulating target site selection during human immunodeficiency virus DNA integration in vitro with an engineered tethering factor. Hum. Gene Ther. **17:**960–967.
16. **Craig, N. L.** 1997. Target site selection in transposition. Annu. Rev. Biochem. **66:**437–474.
17. **Dai, J., W. Xie, T. L. Brady, J. Gao, and D. F. Voytas.** 2007. Phosphorylation regulates integration of the yeast Ty5 retrotransposon into heterochromatin. Mol. Cell **27:**289–299.
18. **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.
19. **Diffley, J. F., and B. Stillman.** 1989. Transcriptional silencing and lamins. Nature **342:**24.
20. **Emanuelli, M., F. Carnevali, M. Lorenzi, N. Raffaelli, A. Amici, S. Ruggieri, and G. Magni.** 1999. Identification and characterization of YLR328W, the *Saccharomyces cerevisiae* structural gene encoding NMN adenylyltransferase: expression and characterization of the recombinant enzyme. FEBS Lett. **455:**13–17.
21. **Gao, X., D. J. Rowley, X. Gai, and D. F. Voytas.** 2002. Ty5 *gag* mutations increase retrotransposition and suggest a role for hydrogen bonding in the function of the nucleocapsid zinc finger. J. Virol. **76:**3240–3247.
22. **Gasser, S. M., and M. M. Cockell.** 2001. The molecular biology of the SIR proteins. Gene **279:**1–16.
23. **Gruhler, A., J. V. Olsen, S. Mohammed, P. Mortensen, N. J. Faergeman, M. Mann, and O. N. Jensen.** 2005. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol. Cell. Proteomics **4:**310–327.
24. **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.

25. **Irwin, P. A., and D. F. Voytas.** 2001. Expression and processing of proteins encoded by the *Saccharomyces* retrotransposon Ty5. J. Virol. **75:**1790–1797.
26. **Jacque, J. M., and M. Stevenson.** 2006. The inner-nuclear-envelope protein emerin regulates HIV-1 infectivity. Nature **441:**641–645.
27. **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.
28. **Janke, C., M. M. Magiera, N. Rathfelder, C. Taxis, S. Reber, H. Maekawa, A. Moreno-Borchart, G. Doenges, E. Schwob, E. Schiebel, and M. Knop.** 2004. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast **21:**947–962.
29. **Kay, B. K., M. P. Williamson, and M. Sudol.** 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. FASEB J. **14:**231–241.
30. **Lee, K. K., T. Haraguchi, R. S. Lee, T. Koujin, Y. Hiraoka, and K. L. Wilson.** 2001. Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. J. Cell Sci. **114:**4567–4573.
31. **Lee, M. S., and R. Craigie.** 1998. A previously unidentified host protein protects retroviral DNA from autointegration. Proc. Natl. Acad. Sci. USA **95:**1528–1533.
32. **Lewinski, M. K., M. Yamashita, M. Emerman, A. Ciuffi, H. Marshall, G. Crawford, F. Collins, P. Shinn, J. Leipzig, S. Hannenhalli, C. C. Berry, J. R. Ecker, and F. D. Bushman.** 2006. Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS Pathog. **2:**e60.
33. **Lewis, A., R. Felberbaum, and M. Hochstrasser.** 2007. A nuclear envelope protein linking nuclear pore basket assembly, SUMO protease regulation, and mRNA surveillance. J. Cell Biol. **178:**813–827.
34. **Llano, M., S. Delgado, M. Vanegas, and E. M. Poeschla.** 2004. Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. J. Biol. Chem. **279:**55570–55577.
35. **Llano, M., D. T. Saenz, A. Meehan, P. Wongthida, M. Peretz, W. H. Walker, W. Teo, and E. M. Poeschla.** 2006. An essential role for LEDGF/p75 in HIV integration. Science **314:**461–464.
36. **Luo, K., M. A. Vega-Palas, and M. Grunstein.** 2002. Rap1-Sir4 binding independent of other Sir, yKu, or histone interactions initiates the assembly of telomeric heterochromatin in yeast. Genes Dev. **16:**1528–1539.
37. **Maertens, G. N., P. Cherepanov, and A. Engelman.** 2006. Transcriptional co-activator p75 binds and tethers the Myc-interacting protein JPO2 to chromatin. J. Cell Sci. **119:**2563–2571.
38. **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.
39. **Mitchell, R. S., B. F. Beitzel, A. R. Schroder, P. Shinn, H. Chen, C. C. Berry, J. R. Ecker, and F. D. Bushman.** 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. **2:**E234.
40. **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.
41. **Murphy, G. A., E. J. Spedale, S. T. Powell, L. Pillus, S. C. Schultz, and L. Chen.** 2003. The Sir4 C-terminal coiled coil is required for telomeric and mating type silencing in *Saccharomyces cerevisiae*. J. Mol. Biol. **334:**769–780.
42. **Pelech, S. L., and J. S. Sanghera.** 1992. Mitogen-activated protein kinases: versatile transducers for cell signaling. Trends Biochem. Sci. **17:**233–238.
43. **Sandmeyer, S.** 2003. Integration by design. Proc. Natl. Acad. Sci. USA **100:**5586–5588.
44. **Schroder, A. R., P. Shinn, H. Chen, C. Berry, J. R. Ecker, and F. Bushman.** 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell **110:**521–529.
45. **Segura-Totten, M., A. K. Kowalski, R. Craigie, and K. L. Wilson.** 2002. Barrier-to-autointegration factor: major roles in chromatin decondensation and nuclear assembly. J. Cell Biol. **158:**475–485.
46. **Shun, M. C., J. E. Daigle, N. Vandegraaff, and A. Engelman.** 2007. Wild-type levels of human immunodeficiency virus type 1 infectivity in the absence of cellular emerin protein. J. Virol. **81:**166–172.
47. **Shun, M. C., N. K. Raghavendra, N. Vandegraaff, J. E. Daigle, S. Hughes, P. Kellam, P. Cherepanov, and A. Engelman.** 2007. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. Genes Dev. **21:**1767–1778.
48. **Singleton, T. L., and H. L. Levin.** 2002. A long terminal repeat retrotransposon of fission yeast has strong preferences for specific sites of insertion. Eukaryot. Cell **1:**44–55.
49. **Taddei, A., and S. M. Gasser.** 2004. Multiple pathways for telomere tethering: functional implications of subnuclear position for heterochromatin formation. Biochim. Biophys. Acta **1677:**120–128.
50. **Tran, H. G., D. J. Steger, V. R. Iyer, and A. D. Johnson.** 2000. The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. EMBO J. **19:**2323–2331.
51. **Vandekerckhove, L., F. Christ, B. Van Maele, J. De Rijck, R. Gijsbers, C. Van den Haute, M. Witvrouw, and Z. Debyser.** 2006. Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. J. Virol. **80:**1886–1896.
52. **Vanegas, M., M. Llano, S. Delgado, D. Thompson, M. Peretz, and E. Poeschla.** 2005. Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. J. Cell Sci. **118:**1733–1743.
53. **Vidal, M., R. K. Brachmann, A. Fattaey, E. Harlow, and J. D. Boeke.** 1996. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. Proc. Natl. Acad. Sci. USA **93:**10315–10320.
54. **Wach, A., A. Brachat, R. Pohlmann, and P. Philippsen.** 1994. New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. Yeast **10:**1793–1808.
55. **Wang, G. P., J. Leipzig, C. C. Berry, and F. D. Bushman.** 2007. HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. Genome Res. **17:**1186–1194.
56. **Wu, X., Y. Li, B. Crise, and S. M. Burgess.** 2003. Transcription start regions in the human genome are favored targets for MLV integration. Science **300:**1749–1751.
57. **Xie, J., M. Pierce, V. Gailus-Durner, M. Wagner, E. Winter, and A. K. Vershon.** 1999. Sum1 and Hst1 repress middle sporulation-specific gene expression during mitosis in *Saccharomyces cerevisiae*. EMBO J. **18:**6448–6454.
58. **Xie, W., X. Gai, Y. Zhu, D. C. Zappulla, R. Sternglanz, and D. F. Voytas.** 2001. Targeting of the yeast Ty5 retrotransposon to silent chromatin is mediated by interactions between integrase and Sir4p. Mol. Cell. Biol. **21:**6606–6614.
59. **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.
60. **Zheng, R., R. Ghirlando, M. S. Lee, K. Mizuuchi, M. Krause, and R. Craigie.** 2000. Barrier-to-autointegration factor (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex. Proc. Natl. Acad. Sci. USA **97:**8997–9002.
61. **Zhu, Y., J. Dai, P. G. Fuerst, and D. F. Voytas.** 2003. Controlling integration specificity of a yeast retrotransposon. Proc. Natl. Acad. Sci. USA **100:**5891–5895.
62. **Zhu, Y., S. Zou, D. A. 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.