

Invading the Yeast Nucleus: a Nuclear Localization Signal at the C Terminus of Ty1 Integrase Is Required for Transposition In Vivo

MARGARET A. KENNA, CARRIE BAKER BRACHMANN,
SCOTT E. DEVINE, AND JEF D. BOEKE*

*Department of Molecular Biology and Genetics, Johns Hopkins University
School of Medicine, Baltimore, Maryland 21205*

Received 5 September 1997/Returned for modification 22 October 1997/Accepted 7 November 1997

Retrotransposon Ty1 faces a formidable cell barrier during transposition—the yeast nuclear membrane which remains intact throughout the cell cycle. We investigated the mechanism by which transposition intermediates are transported from the cytoplasm (the presumed site of Ty1 DNA synthesis) to the nucleus, where they are integrated into the genome. Ty1 integrase has a nuclear localization signal (NLS) at its C terminus. Both full-length integrase and a C-terminal fragment localize to the nucleus. C-terminal deletion mutants in Ty1 integrase were used to map the putative NLS to the last 74 amino acid residues of integrase. Mutations in basic segments within this region decreased retrotransposition at least 50-fold in vivo. Furthermore, these mutant integrase proteins failed to localize to the nucleus. Production of virus-like particles, reverse transcriptase activity, and complete in vitro Ty1 integration resembled wild-type levels, consistent with failure of the mutant integrases to enter the nucleus.

Yeast retrotransposon Ty1 is similar in many aspects to retroviruses (7). Ty1 must generate more copies of itself for insertion into the genome of the host cell. Ty1 mRNA is transcribed and processed in the nucleus and then transported to the cytoplasm, where it is translated into Gag and Gag-Pol proteins. These are assembled into virus-like particles (VLPs) resembling retroviral core particles (34, 50). Reverse transcription of Ty1 RNA into Ty1 cDNA occurs in the VLPs, which also contain integrase (IN) (30). Unlike retroviruses, in which the viral core becomes enveloped and leaves the producer cell to infect a new cell and integrate into the target genome, Ty1 cDNA reenters the producer cell nucleus to integrate at a new target site. How extrachromosomal Ty1 cDNA enters the nucleus where it can access its target site DNA remains unresolved.

The question of how Ty1 enters the nucleus is especially important because in all fungi, including the Ty1 host *Saccharomyces cerevisiae*, the nuclear envelope (NE) does not break down during mitosis (43). This significant intracellular barrier prevents Ty1 DNA from accessing the genome via mitotic NE dissociation, a cellular process apparently exploited by many retroviruses to gain nuclear access (55). However, this opportunity exists only in cells that are undergoing growth and division. Lentiviruses like human immunodeficiency virus type 1 (HIV-1) are able to infect nondividing cells (15, 17), which, like growing yeast cells, do not offer a mitotic window of opportunity to the nucleus. Thus, an understanding of the mechanism used by Ty1 to enter the nucleus should shed light on the analogous process in lentiviruses.

Several possible mechanisms for nuclear entry can be envisioned. VLPs might enter the nucleus intact. However, VLPs are estimated to be 600 Å in diameter and thus seem too large to be transported through the nuclear pore complex (NPC)

(13, 19, 34, 50). Maximal pore size estimates range from 90 Å (for proteins that enter the nucleus passively) to 260 Å (for NLS-bearing proteins of >~40 kDa) (23, 29). Alternatively, VLPs might enter the nucleus by means of a conformational change allowing passage through pores. However, electron micrographs show that most if not all VLPs are cytoplasmic (34, 50). Another formal possibility is that VLPs dock near the pore and inject their contents into the nucleus (7) much like what has been proposed for adenovirus (21). A more likely possibility is that a subassembly of the particle (similar to retroviral preintegration complexes) is transported into the nucleus by host cell machinery.

IN is a Ty1 protein that must enter the nucleus since its intranuclear function of inserting a Ty1 cDNA copy into the chromosomal DNA is absolutely required for retrotransposition. In this study, we demonstrated that Ty1 IN indeed has a bipartite nuclear localization signal (NLS) that is essential for retrotransposition. This NLS differs from other bipartite signals in its long spacer region. An analysis of related protein sequences found in the complete set of yeast protein sequences suggests there may be other NLSs of this type.

MATERIALS AND METHODS

Strains and media. The plasmids used are listed in Table 1. DNA cloning was performed by use of standard techniques described by Sambrook et al. (56). Selection against Ura⁺ strains was accomplished by culture on solid synthetic complete (SC) medium containing 1 mg of 5-fluoro-orotic acid (5-Foa) per ml (6).

The following yeast strains were used: W303a (*MATa ade2-1 ura3-1 his3-11 trp1-1 leu2-3 can1-100*), W303 α (*MAT α ade2-1 ura3-1 his3-11 trp1-1 leu2-3 can1-100*), YH50 (*MATa his3 Δ 200 ura3-167 leu2 Δ 1 trp1 Δ 1 spt3- Δ 202*), and YM603 (*MATa ura3-52 his3 Δ 200 ade2-101 hys2-801 met⁻ reg1-501*). W303a and - α were obtained from L. Davis. YM603 was obtained from M. Johnston. YH50 was constructed by H. Xu in this laboratory.

Indirect immunofluorescence. The anti-IN monoclonal antibody (MAb) 8B11 (30) and polyclonal anti-E1 (raised against an avian coronavirus glycoprotein [53]) antibodies were used for immunofluorescence experiments essentially as described by Adams and Pringle (1). To visualize nuclear DNA, cells were stained with 0.5 μ g of DAPI (4',6-diamidino-2-phenylindole) per ml. Treated cells were mounted on glass slides in a 1-mg/ml concentration of *p*-phenylene-diamine in 90% glycerol. The cells were examined by using fluorescence microscopy with a Zeiss Axioscope fitted with UV and fluorescein isothiocyanate

* Corresponding author. Mailing address: Johns Hopkins University School of Medicine, Dept. of Molecular Biology and Genetics, 725 N. Wolfe St., Baltimore, MD 21205. Phone: (410) 955-2481. Fax: (410) 614-2987. E-mail: jef.boeke@gmail.bs.jhu.edu.

TABLE 1. Plasmids used in this study

Plasmid	Markers	Comments	Source
pGEU414	<i>CEN TRP1</i>	<i>GAL1 E1</i>	28
pGEU415	<i>CEN LEU2</i>	<i>GAL1 E1</i>	28
pCAR224	<i>CEN LEU2</i>	C-terminal IN in p415GEU1	This study
pCAR227	<i>CEN LEU2</i>	Full-length IN in p415GEU1	This study
pCAR607	<i>CEN LEU2</i>	N-terminal IN in p415GEU1	This study
pMAK60	<i>CEN TRP1</i>	Full-length IN in p414GEU1	This study
pMAK61	<i>CEN TRP1</i>	IN Δ <i>BglII-ClaI</i>	This study
pMAK62	<i>CEN TRP1</i>	IN Δ <i>EcoRI</i>	This study
pMAK63	<i>CEN TRP1</i>	IN Δ <i>Clal</i>	This study
pMAK64	<i>CEN TRP1</i>	IN Δ <i>Clal</i> + <i>EcoRI</i>	This study
pMAK65	<i>CEN TRP1</i>	IN Δ <i>BglII-ClaI</i> / Δ <i>EcoRI</i>	This study
pMAK66	<i>CEN TRP1</i>	C-terminal IN in pGEU414	This study
pMAK67	<i>CEN TRP1</i>	N-terminal IN in pGEU414	This study
pX3	<i>CEN URA3 TRP1</i>	Ty1- <i>TRP1</i> , <i>URA3</i> on plasmid	58
pMAK69	<i>CEN URA3 TRP1</i>	IN B1 in pX3	This study
pMAK70	<i>CEN URA3 TRP1</i>	IN B2 in pX3	This study
pGM315	<i>CEN URA3 TRP1</i>	Linker insertion mutation in IN	51a
90-1553	<i>CEN URA3 TRP1</i>	DD-YD change at YXDD domain in Ty1 RT	A. Gabriel (unpublished data)
pSD546	2 μ m <i>HIS3 bla</i>	Target plasmid for in vitro IN assay	26

(FITC) optics. Images were collected by a cooled charge-coupled device camera (Photometrix) by using IPLab Spectrum software from Signal Analytics (Vienna, Va.) and imported into Adobe Photoshop 3.0 for figure presentation. The following antibodies were used at the dilutions indicated in parentheses: 1 mg of protein G-Sepharose-purified MA6 8B11 (1:2,500 to 1:5,000) per ml, polyclonal R1-F serum (1:2,000), and polyclonal E1 serum (1:500). A diploid yeast strain made by mating W303a and W303 α was used for all immunofluorescence experiments.

Construction of GAL-IN subclones: E1-tagged IN constructs. Three plasmids encoding E1-IN (full length), E1-N-term IN, and E1-C-term IN (N-term and C-term IN are N- and C-terminal fragments of IN and are defined relative to the internal *BglII* site within the coding region of IN) were made. The E1 epitope tag is derived from the coronavirus glycoprotein (53). The IN regions were amplified by PCR and cloned into expression vector p415GEU1 (28). The fragment encoding full-length IN was amplified by using primer JB581 (nucleotides [nt] 2041 to 2063 with respect to the Ty1-H3 sequence [4]) incorporating a *XhoI* site 5' to nt 2041 (see oligonucleotide list below) and primer JB584 (nt 4056 to 4077) incorporating a stop codon and a *PstI* site 3' to nt 4077. The fragment encoding N-term IN was amplified by using primers JB581 and JB582 (nt 3286 to 3306) incorporating a stop codon and a *PstI* site 3' to nt 3306. The fragment encoding C-term IN was amplified by using primer JB583 (nt 3301 to 3318) incorporating a *XhoI* site 5' to nt 3301 and primer JB584. The PCR products were digested with *XhoI* and *PstI*, gel purified, and ligated to p415GEU1 digested with *XhoI* and *PstI*, creating pCAR153 (E1-IN), pCAR150 (E1-N-term IN), and pCAR151 (E1-C-term IN). Because these initial constructs did not faithfully reproduce the native C terminus of IN protein (amino acid sequence RIHLLA), we corrected the constructs encoding full-length and C-term IN as follows. The *Sali-SmaI* (*SmaI* is a polylinker site downstream of the *PstI* site used in cloning) fragment of IN in pCAR153 was replaced with the *Sali-PvuII* fragment of IN from pJEF724 (4, 5), creating pCAR226. pCAR224 was constructed in a similar fashion, with the *BglII-SmaI* fragment of pCAR151 being replaced with the *BglII-PvuII* fragment of IN from pJEF724. The *BssHII* fragments of pCAR224 and pCAR226 were then subcloned into pRS414 (*CEN TRP1*), transferring the *GAL-E1-IN* cassettes and generating pMAK67 and pMAK60 (see Table 1). IN Δ *BglII-ClaI* was created by cutting pMAK60 with *BglII* and *Clal*, filling in with Klenow fragment, and ligating the blunt ends, creating pMAK61. This created an in-frame fusion with the rest of the C terminus of IN. IN Δ *EcoRI* was created by digesting pMAK60 with *EcoRI* and religating, creating pMAK62. IN is truncated by a fortuitous stop codon within 3 amino acids of the *EcoRI* fusion point. IN Δ *Clal* was created by doubly digesting pMAK60 with *Clal* and *EcoRI*. The ends were filled in, and the blunt ends were religated, generating pMAK63. This blunt ligation creates a new *EcoRI* site. In this construct, IN is truncated by a fortuitous stop codon 3 amino acids after the *Clal/EcoRI* fusion point. IN Δ *Clal+EcoRI* was created by subcloning the 200-bp *EcoRI* fragment into the single *EcoRI* site that was generated by filling in and fusing *Clal* to *EcoRI*, generating pMAK64. IN Δ *BglII-ClaI/EcoRI* was created by digesting the IN Δ *BglII-ClaI* with *EcoRI* and religating, generating pMAK65.

Oligonucleotides. The oligonucleotides used in the present study and their sequences were as follows: JB581, CGCGCTCGAGGAATGTCCATACAAGT GAAAGTAC; JB582, GCGCTCGAGTCAAGATCTCTTCTTTGTGGAAAG; JB583, CGCGCTCGAGATCTAGCAACCCCA; JB584, GCGCTCGAGT CACAGTTGATTGACTTCTTTGTGG; JB1391, GCGGATCCAGCAGCGTTG ATAGTAGTATAGGCATT; JB1392, GCGGATCCGCTGCTAATGAACT

GAAATTAAGGTA; JB1393, GCGGATCCAGCAGCTTCTAACTACGCA TATTCTT; JB1394, GCGGATCCGCTGCTCACCTGATTGCAGGGGGGA TC; JB1395, GCGGATCCGCTGCTCACCTGATTGCAGCTGTAA; JB1396, CAATATCCAGTACGGGTG; JB1397, CAGATGGTAAGCCCTCC; JB1498, AGTTTGGGTGGTATTGGTGACTCTAATGCCTATACTACTACTTCTAGA AGATAATGAAACTGAA; JB1499, AGTTTGGGTGGTATTGGTGACTCT AATGCCTATACTACTATATCGCAGCAGGATCCGAGCATTAGAAGATA ATGAAACTGAA; JB1500, AATTAACCCTCACTAAA; JB1523, AATATG CGTAGTTTAGAACCCTCCGGCAGCAGGATCCGAGCACACCTGATTG CAGGGGGGATC; JB1524, GATCCCCCTGCAATCAGGTGTGCTGCGG ATCCTGCTGCCGAGGTTCTAACTACGCATATT; JB1618, AATGCCT ATACTACTACAACGGTAAGAAGAGATCTAAGGCTGCTGGATCCGC TGCTAATG; JB1619, CATTAGCAGCGGATCCAGCAGCCTTAGATCTCT TCTACCGTTGATAGTAGTATAGGCATT.

Site-directed mutagenesis. Mutations (block substitutions or additions) in IN were made in one of two ways by utilizing fusion PCR strategies (38). (i) In the first strategy, the IN mutants in Ty1 were made in the parent vector pX3 (58) (see Fig. 4). Mutant DNA fragments of Ty1 IN were generated by PCR and directly ligated into pCRII (Invitrogen) and sequenced. Amino acid residues of interest (Fig. 1, underlined residues) were replaced by the sequence AAGSAA, which includes a diagnostic *BamHI* site (51). The first PCR fragment for the IN basic region 1 (B1) mutant was synthesized by using pX3 as the template and oligonucleotides JB1391 and JB1396 as primers. The second PCR fragment was synthesized by using pX3 as the template and oligonucleotides JB1392 and JB771 as primers. pCRII containing the 5' segment was restricted with *KpnI* and *BamHI*. pCRII containing the 3' segment was restricted with *BamHI-AflII*. Both insert fragments were ligated to pX3 previously digested with *KpnI* and *AflII*. The final construct was also completely sequenced in the regions derived from PCRs. The IN basic region 2 (B2) mutant was similarly assembled by using primer pairs JB1393-JB1396 and JB1395-JB771. (ii) To create block substitutions or insertion mutations, another PCR strategy was also employed (57). 3' and 5' primers homologous to the insertion site, each containing the desired insertion or substitution sequences, were designed. These were then used in separate amplification reactions, paired with the appropriate primer from either an upstream or downstream site. The two PCR products of IN thus generated overlap at the ends containing the mutation site; they were then used as the template to amplify the entire mutated construct with the two outside primers. The resulting fragment was cleaved with unique restriction sites and subcloned into the IN expression plasmid of choice.

Protein analysis. For fractionation studies, yeast strains were grown to mid-log phase and cells were turned into spheroplasts as described by Kalinich and Douglas (41). The spheroplasts were lysed by resuspension in 10 ml of ice-cold solution A (20 mM HEPES-HCl [pH 7.5], 5 mM MgCl₂, protease inhibitors) per g of yeast cell pellet (41), followed by vortexing and incubation on ice for 5 min. Lysates were centrifuged at 9,000 rpm in an SA600 rotor for 45 min at 4°C. The supernatant (S1), representing soluble cellular components, was collected. The pellet was resuspended in 10 ml of ice-cold solution B (20 mM HEPES-HCl [pH 7.4], 5 mM MgCl₂, 1 M NaCl, protease inhibitors) per g of yeast cell pellet, incubated on ice for 10 min, and centrifuged at 9,000 rpm for 45 min. This supernatant (S2) was collected. Eight hundred microliters of S1 fractions and 1,000 μ l of S2 fractions were precipitated with trichloroacetic acid (TCA) for

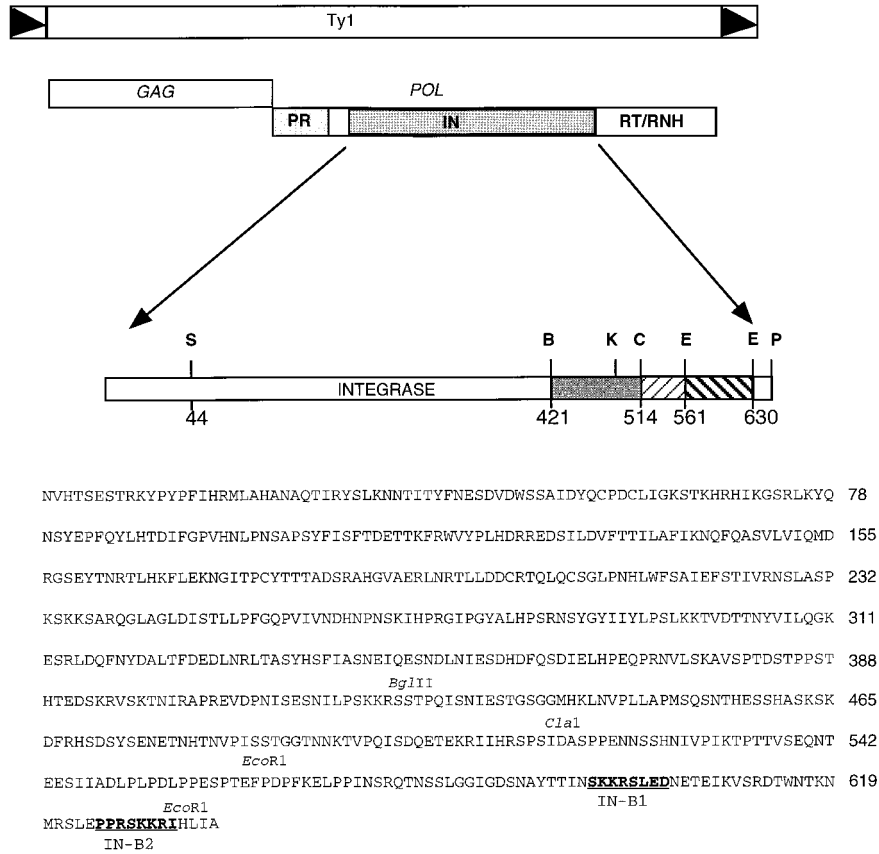


FIG. 1. Ty1 and IN structure. The structure of the Ty1 retrotransposon is diagrammed. Black triangles indicate the long-terminal-repeat regions. The GAG and POL ORFs and the element-encoded proteins derived from POL are indicated. The complete sequence of Ty1 IN is shown. Restriction site abbreviations: S, *Sal*I; B, *Bgl*II; C, *Cla*I; E, *Eco*RI; K, *Kpn*I; P, *Pvu*II. The amino acid number is indicated beneath each restriction site. Key restriction sites are indicated above the corresponding amino acid sequence. IN NLS mutations, B1 and B2, made in the GAL-Ty1 plasmid are in boldface type. The numbers to the right of the sequence represent the last amino acid residue on that line of sequence.

protein gel analysis and resuspended in Laemmli buffer (45). Proteins were separated electrophoretically on sodium dodecyl sulfate (SDS)-8% polyacrylamide gels as described below. Rabbit antihexokinase antibody was a gift from Rob Jensen's lab. Mouse antinucleoporin (MAB 414) was a gift from Laura Davis.

Whole-cell yeast extracts were made by harvesting 5 ml of cells (*A*₆₀₀: 0.8 to 1.0), washing the cells once with H₂O, and resuspending them in 200 µl of H₂O. This suspension was frozen in liquid N₂ and stored at -80°C. Glass beads and 40 µl of 100% TCA were added, and the yeast cells were broken in a Mini Bead Beater 8 (Biospec Products) as directed by the manufacturer. One milliliter of 5% TCA was added, and samples were spun in a microcentrifuge for 20 min at 4°C. The supernatant was removed, and the pellet was resuspended in H₂O and spun and washed as described for the previous step. Proteins were extracted by resuspending the pellet in sample buffer (6% SDS, 0.5 M Tris base) at 45 to 50°C for 10 min two times, pooling the supernatants. One-third volume of a solution of 0.25 M dithiothreitol, 50% glycerol, and 0.2% bromophenol blue was added prior to loading the gel.

Proteins were separated electrophoretically on SDS-8% or -10% polyacrylamide gels and transferred to nitrocellulose by use of a Bio-Rad semidry transfer apparatus. The following antibodies were used at the dilutions indicated in parentheses: 1 mg of protein G-Sepharose-purified MAB 8B11 (1:5,000) per ml, polyclonal R1-F serum (1:2,000), polyclonal E1 serum (1:500). Filters were incubated with the appropriate secondary antibodies and then incubated with an enhanced chemiluminescence reagent (Amersham) and exposed to film.

Transposition assays. The transposition assay method is similar to that previously described (9). pX3 and its derivatives were transformed into YH50 (46). This strain is an *spt3* mutant; hence, transposition events arising from endogenous Ty1 copies are eliminated (8). Colonies from each strain were patched onto SC-Ura-Trp (SC medium lacking uracil and tryptophan) and incubated at 30°C for 2 days. The cells from the plates were replica plated onto SC-Ura-Trp galactose and incubated for 3 days at 22°C for induction of transposition. The cells from the plates were replica plated onto yeast extract-peptone-dextrose (YPD) and incubated at 30°C for 24 h to allow for plasmid loss. After replica plating the cells onto YPD once more, the cells from the final YPD plates were

replica plated onto SC+5-Foa-Trp and grown at 30°C. Cells that lost the *URA3*-based plasmid and contain transpositions grow on SC+5-Foa-Trp. To quantify results, patches from the second YPD replica were scraped into sterile H₂O, and appropriate dilutions were plated for single colonies onto YPD plates and replica plated onto SC-Ura and SC-Trp plates. Transposition frequencies were determined by dividing the number of Trp⁺ Ura⁻ derivatives by the total number of Ura⁻ derivatives, and then converting the value to a percentage of wild-type frequency, normalized to 100%.

VLP preparations, RT assays, and in vitro IN assays. VLPs were isolated from a *regI* mutant strain (YM603) which allows galactose induction in the presence of glucose (39). A single colony was grown in 10 ml of SC-Ura medium supplemented with 0.8 mM adenine and 2% glucose for 18 to 24 hours at 30°C. This culture was used to inoculate 200 ml of the same medium and was incubated at 30°C for another 24 h. For induction of VLPs, 500 ml of yeast nitrogen base medium without ammonium sulfate and amino acids (Difco) supplemented with 20 g of Casamino Acids per liter, 5 g of ammonium sulfate per liter, 0.8 mM adenine, 2% glucose, and 2% galactose in a 2-liter flask was inoculated with enough cells from the 200-ml culture to give a starting *A*₆₀₀ of 0.25 to 0.5. This culture was incubated at 22°C in a floor shaker with vigorous shaking (350 rpm) for 36 h. The final *A*₆₀₀ was between 4.0 and 6.0. The cells were harvested, and VLPs were prepared as described previously (10, 30). Reverse transcriptase (RT) activity was assayed as described previously (30, 34), and in vitro integration activity was performed as reported before (25, 26).

RESULTS

Ectopically expressed IN localizes to the nucleus. The question of exactly how Ty1 cDNA crosses the yeast NE (which remains intact throughout the cell cycle) and then inserts itself at its target sites within the yeast genome is intriguing. Other than intact VLPs (30), preintegration complexes have not yet been defined for Ty1, but these should consist minimally of a

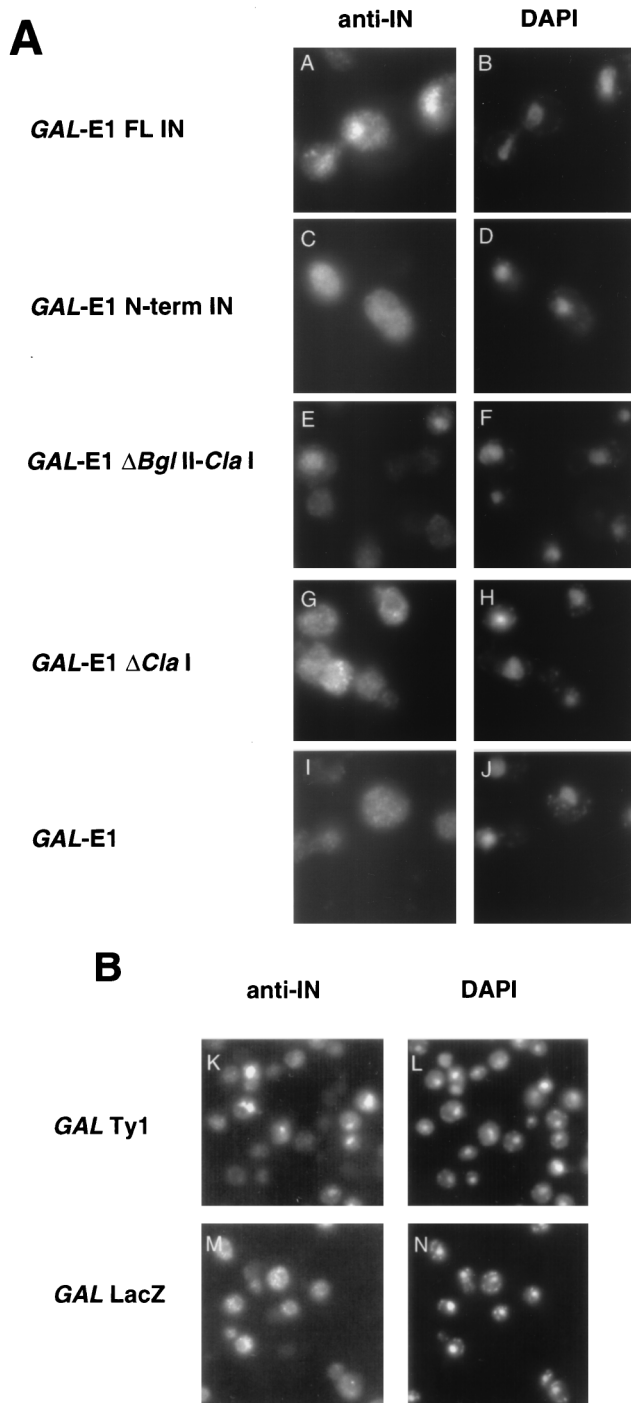


FIG. 2. Ty1 IN localizes to the nucleus. Anti-IN antibody was used on cells induced for the production of E1-IN from the *GALI* promoter (A) or on cells induced for Ty1 transposition (B, panels K and L [panels M and N show *GAL-LacZ* negative controls]). FITC-stained panels are on the left; DAPI-stained panels are on the right. Constructs tested for IN localization are indicated to the left of the micrographs. FL, full length.

cDNA-IN complex, since IN must cross the NE to direct integration of new Ty1 copies. This led us to examine whether Ty1 IN was nuclear localized.

Full-length Ty1 IN was fused at its N terminus to a *GALI*-regulated E1 epitope (see Materials and Methods). Indirect

immunofluorescence using an anti-IN MAb, 8B11 (30), shows that IN is nuclear localized (Fig. 2A, panels A and B). The IN-coding region was further subcloned to identify which region conferred nuclear localization. An N-terminal fusion consisting of residues 1 to 422 and a C-terminal fusion containing residues 423 to 636 were constructed and analyzed by immunofluorescence. The N-terminal fusion failed to localize to the nucleus, as assayed by MAb 8B11 staining (Fig. 2A, panels C and D). This finding suggests that an NLS resides at the C-terminal third of IN. To analyze the C-terminal construct, we used a polyclonal antibody against the E1 epitope (28, 53) because the MAb 8B11 epitope is not present on this fragment. The C-terminal IN fusion localized to the nucleus (Fig. 3A). Further mapping was carried out by using a series of deletion constructs. Figure 2 shows immunofluorescence of a key subset of these constructs. Since the deletion of C-terminal residues between *Bgl*II and *Cla*I (Fig. 1) had no effect on nuclear localization but the deletion of residues after the *Cla*I site did have an effect (Fig. 2A, panels G and H), we conclude that an NLS resides in the last 115 amino acids of IN.

Ty1 IN from cells undergoing transposition localizes to the nucleus. Pol proteins are expressed as translational read-through products from the *GAG* open reading frame (ORF) (3, 20, 49). The Ty1 protease (PR) processes the Gag/Pol polyprotein (190 to 200 kDa) to mature Ty1 proteins (p90 IN and p65 RT/RNase H [35]). We examined the localization of the IN expressed in this *GAG-POL* context. Cells containing a plasmid-borne Ty1 element under control of the *GALI* promoter were induced for transposition. IN localized to the nucleus of yeast cells, as demonstrated by MAb 8B11 immunofluorescence (Fig. 2B, panels K and L). We observed nuclear localization in about 40% of the cells and no signal in about 60% of the cells. The lack of signal in 60% of the cells is directly correlated with the high rate of plasmid loss of this *GAL-Ty1* plasmid in galactose medium (even when cells are grown selectively; data not shown). Figure 2B, panels M and N, show a negative control of a *GAL-lacZ* plasmid. Several additional controls were used, including a comparison of uninduced and induced cells, an empty 2 μ m vector, or secondary antibody alone (data not shown). In each case, these negative controls failed to show nuclear staining, indicating that the observed nuclear staining was IN specific.

C-terminal deletions of IN identify a candidate NLS(s). To more precisely map the regions of the C terminus representing the putative NLS, a series of C-terminal deletions were made in the *GAL-E1-IN* construct, and plasmids encoding these truncated proteins were transformed into yeast. The transformants were induced by growth on galactose and analyzed by immunoblotting and immunofluorescence. Figure 3B shows that all of the truncated forms of IN are expressed. Immunofluorescence with MAb 8B11 shows that IN is nuclear localized in all constructs that maintain the last 74 residues (encoded within the *Eco*RI-*Eco*RI fragment) (Fig. 3A). This fragment, when added back to the Δ *Cla*I construct, restores nuclear localization (Fig. 3A; IN Δ *Cla*I+*Eco*RI). Therefore, an NLS resides within the last 74 residues.

To biochemically confirm the immunofluorescence data, we fractionated protein extracts from several of the transformants into soluble and nuclear components. It has previously been shown that nuclear proteins such as nucleoporins fractionate with the nuclear pellet and can then be solubilized with 1 M NaCl (24). The subcellular localization of the IN fusions was crudely analyzed by separating soluble (S1) and particulate fractions by low-speed centrifugation of spheroplast lysates. The particulate fraction was then extracted with 1 M NaCl (S2) (see Materials and Methods). Immunoblot analysis of these

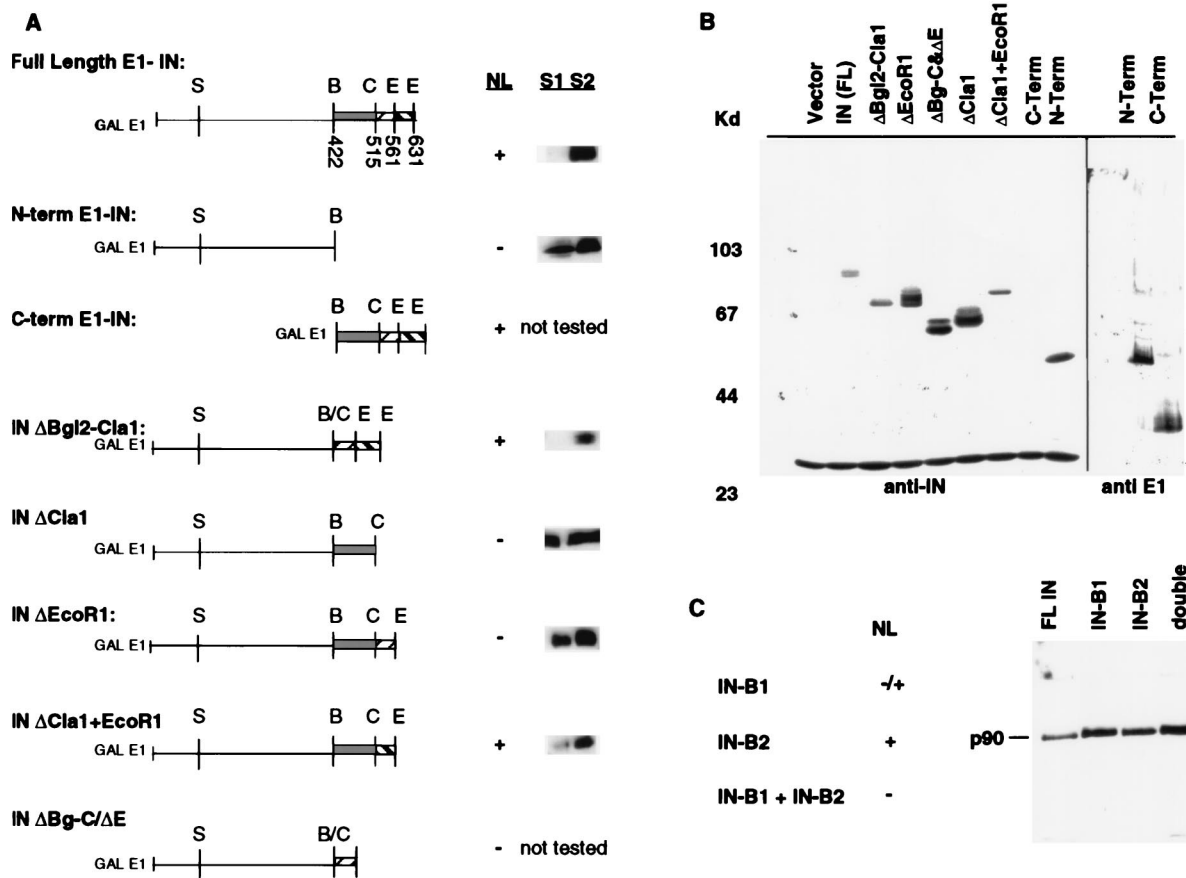


FIG. 3. Analysis of E1-IN fusion proteins. (A) C-terminal deletion constructs map the NLS region within the last 74 amino acids of IN. Constructs tested are diagrammed, with restriction sites indicated; abbreviations for these are defined in the legend to Fig. 1. Numbers shown below restriction sites in full-length E1-IN represent amino acid residues. Nuclear localization (NL) is indicated as present (+) or absent (-); these results are from the indirect immunofluorescence experiments with anti-IN antibody described in the legend to Fig. 2, except for that for the C-terminal construct that necessitated the use of anti-E1 (see text). Cell fractionation experiments (described in Materials and Methods) were performed on strains bearing the E1-IN truncation constructs. Bands corresponding to the indicated E1-IN protein are shown to the right of the NL data. (B) Immunoblot analysis of the truncated IN proteins. Whole-cell extracts were electrophoresed on an SDS-8% polyacrylamide gel. Proteins were detected with either MAb 8B11 (1:5,000) or anti-E1 (1:500). Molecular mass markers are shown at the left in kilodaltons. MAb 8B11 also recognizes a non-Ty1 band of ~30 kDa. Since the C-terminal truncation does not contain the epitope for MAb 8B11, this protein extract, along with the N-terminal one, was probed with anti-E1. (C) Similar immunoblot analysis of the substitution mutant *GAL*-E1-IN constructs B1 (NSKKRS→AAGSAA) and B2 (RSKKRI→AAGSAA). NL is indicated as present (+), partial (-/+), or absent (-).

two fractions showed that deletions removing the NLS region result in the release of a significant portion of the E1-IN construct into the S1 fraction (Fig. 3A, N-term E1-IN, INΔ*Cla*I, and INΔ*Eco*RI). Constructs that retain this region (full-length IN, INΔ*Bgl*II-*Cla*I, and INΔ*Cla*I+*Eco*RI) fractionate with the nuclear-enriched fraction and are released into the S2 supernatant only by treatment with 1 M NaCl. Although this crude fractionation is not sufficient to unambiguously determine subcellular localization, these data fully support the conclusions from the immunofluorescence experiments. The same immunoblot was reprobed with an antibody against a soluble protein, hexokinase, as a marker for the cytosolic fraction and with an antibody against a nucleoporin, MAb 414, as a marker for the nuclear fraction. These results (data not shown) indicate that the S1 and S2 fractions represent cytoplasmic and nuclear compartments, respectively.

Site-directed mutagenesis. Canonical NLSs are most typically regions of consecutive basic amino acids. There are two basic segments in IN that somewhat resemble previously identified NLS motifs. We tested whether either basic region 1 (B1) or 2 (B2) (Fig. 1) acted as an NLS by replacing these basic

regions with the neutral residues AAGSAA. Similar mutations were made in both the *GAL*-Ty1 and the *GAL*-E1-IN vectors.

For cells expressing the *GAL*-E1-IN construct, we found that the IN B1 mutation caused a clear but partial loss of nuclear signal, with some signal instead accumulating in the cytoplasm, as compared to cells expressing wild-type IN (Fig. 3C). The IN B2 mutation apparently had no detectable effect by itself on the nuclear localization of IN, since staining of the mutant with MAb 8B11 looked similar to staining of the wild type. However, when both mutations were introduced into the same full-length E1-IN construct, nuclear localization of IN was completely abolished. Thus, site-directed alterations to constructs expressing the E1-IN fusion protein reveal that these two regions are each important for nuclear localization, reminiscent of the bipartite nuclear targeting sequence of nucleoplasmin (27).

Nuclear localization and retrotransposition defects. The advantage of the *GAL*-E1-IN system lies in the ease with which the immunofluorescence analysis can be performed, presumably because more IN protein is produced. However, the more relevant question is what happens to IN produced within the

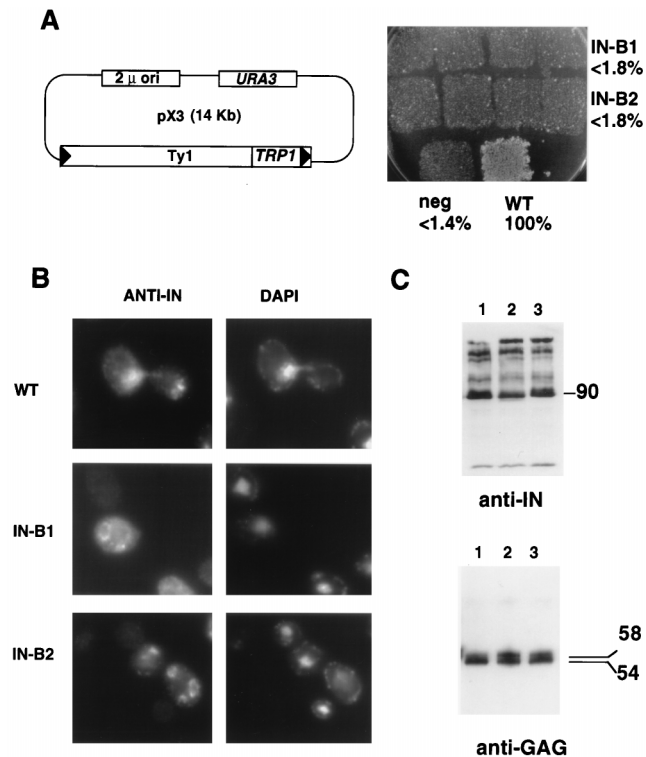


FIG. 4. Two sites required for nuclear localization and transposition. (A) (Left) The pX3 *GAL-Ty1-TRP1* plasmid (58) was used as the parent for making the IN B1 and B2 mutants (residues underlined in amino acid sequence of Fig. 1 were replaced with AAGSAA). Cells containing these pX3 derivatives were assayed for Ty1-*TRP1* transposition. (Right) Transposition results for pX3 (wild-type IN [WT]) and mutants B1, B2, and RT (the negative control [neg]) are shown. Percentages of wild-type transposition frequency are indicated. (B) Indirect immunofluorescence. Anti-IN (MAb 8B11) panels are shown on the left (stained with FITC); DAPI staining is shown on the right. (C) (Top) Immunoblot with anti-IN (MAb 8B11) of whole-cell extracts prepared from wild-type (lane 1), IN B1 (lane 2), and IN B2 (lane 3) strains. The apparent molecular mass of fully processed IN is indicated on the right in kilodaltons. (Bottom) Immunoblot with R1-F (anti-Gag antibody). The apparent molecular masses of unprocessed and processed Gag proteins are indicated on the right in kilodaltons. Lanes are as described for the top panel.

context of a native Ty1 element. In this natural context, *in vivo* function, transposition, and other aspects of the Ty1 life cycle can be assessed along with effects on nuclear localization.

When strains carrying IN B1 or B2 mutations in the *GAL-Ty1-TRP1* plasmid were assayed by immunofluorescence, we detected no MAb 8B11 staining in the nucleus of cells of either mutant strain (Fig. 4B). In these mutants, we observed a punctate or patchy cytoplasmic signal. In some instances, we observed ring-like structures (also seen in cells with the wild-type plasmid but to a lesser extent); however, these ring structures are clearly not surrounding the nucleus, as judged by the location of the DAPI staining regions, and therefore do not represent a nuclear-rim staining pattern. Perhaps this staining reflects mutant IN trapped in the cytoplasm or in VLPs, unable to enter the nucleus. When strains with these plasmids were assayed for transposition, both the IN B1 and B2 mutants were severely defective in transposition, as compared to that of the wild type (Fig. 4A). Their transposition defect is as severe as that of a previously constructed RT mutant with a missense mutation in the active site region (FVDD→FVYD) (31a), and the level of transposition is at least 50-fold lower than wild-type Ty1-*TRP1* transposition. Taken together, these data suggest that these two basic segments are key components of an un-

usually long bipartite NLS. The spacer region between the KKR motifs is 30 residues long. For example nucleoplasmin, the model for this class of NLSs, has a spacer of only 10 residues (54).

Mutant VLPs have normal PR, RT, and strand transfer activities. To ensure that these mutations were specific for a defect in nuclear transport of IN, we analyzed other retrotransposon functions. We first examined the IN protein by immunoblotting to evaluate expression level and whether it is properly released from the Gag-Pol precursor by the Ty1 PR. This is important because the IN B2 mutation lies very close to the IN/RT junction (Fig. 1). Figures 4C (anti-IN panel) and 5D show that IN is made at near wild-type levels and is processed properly. Additionally, similar levels of cDNA for the mutant and the wild-type INs were recovered (data not shown). We also probed with an anti-Gag antibody (R1-F [12]) and observed Gag proteins of the expected mobilities (p58 and p54), indicating that proper processing to mature Gag occurred in the mutant VLPs (Fig. 4C). VLPs were isolated by running 20 to 70% sucrose density gradients, and the VLP profiles were similar for the wild type and mutants (Fig. 5A). RT activity cofractionated with the VLP peak, and levels of encapsidated IN protein were similar in all of the samples (Fig. 5B). Most importantly, *in vitro* IN activity in the NLS mutants was nearly identical to that in the wild type (Fig. 5B). Previous experiments demonstrated that the C terminus is dispensable for the single-end joining function of IN (11). However, the assay used here closely mimics the *in vivo* integration reaction. The reaction consists of mixing an artificial transposon substrate with Ty1 termini, a target plasmid, and the VLPs to be tested (Fig. 5C). Formation of a recombinant target plasmid bearing the artificial transposon requires concerted joining of both transposition ends to target DNA, with appropriate formation of a 5-bp target site duplication (25, 26). The B1 and B2 mutant INs are catalytically indistinguishable from wild-type IN in this *in vitro* reaction (Fig. 5B). VLPs isolated from these mutants are fully functional transposition intermediates. Thus, an essential function of the basic C-terminal region of IN is to transport the IN protein and perhaps a larger preintegration complex into the nucleus.

A new family of NLS sequences? Since the Ty1 IN NLS appeared to consist of two repeats of the sequence SKKR separated by about 30 amino acid residues, we used this information to search the complete yeast genome sequence for possible additional members of this class of NLS. By using the search parameters (SKKR)_{N20-40}(SKKR) with the program CoSMoS (pre-release version made available by D. Bassett and P. Hieter; the final version will be deposited on the XREF database, <http://www.ncbi.nlm.nih.gov/XREFdb/>) to search the complete set of yeast ORFs, we identified only IN sequences from Ty1 and Ty2 elements. By relaxing the stringency to (KKR)_{N20-40}(KKR), we identified 16 additional single copy sequences (Table 2). Of these, 10 ORFs were sequences about which little is known, and six represented proteins of known or suspected function. Remarkably, five of the six known proteins are nuclear localized or very likely to be so. In several cases, the putative NLS is near the C terminus as in Ty1 IN.

DISCUSSION

Retrotransposition of Ty1 absolutely depends on IN. IN is responsible for the strand-joining reaction that unites Ty1 cDNA with host cell target DNA. The studies presented here indicate that IN must be transported to the nucleus for retrotransposition to occur. We have identified a bipartite NLS at

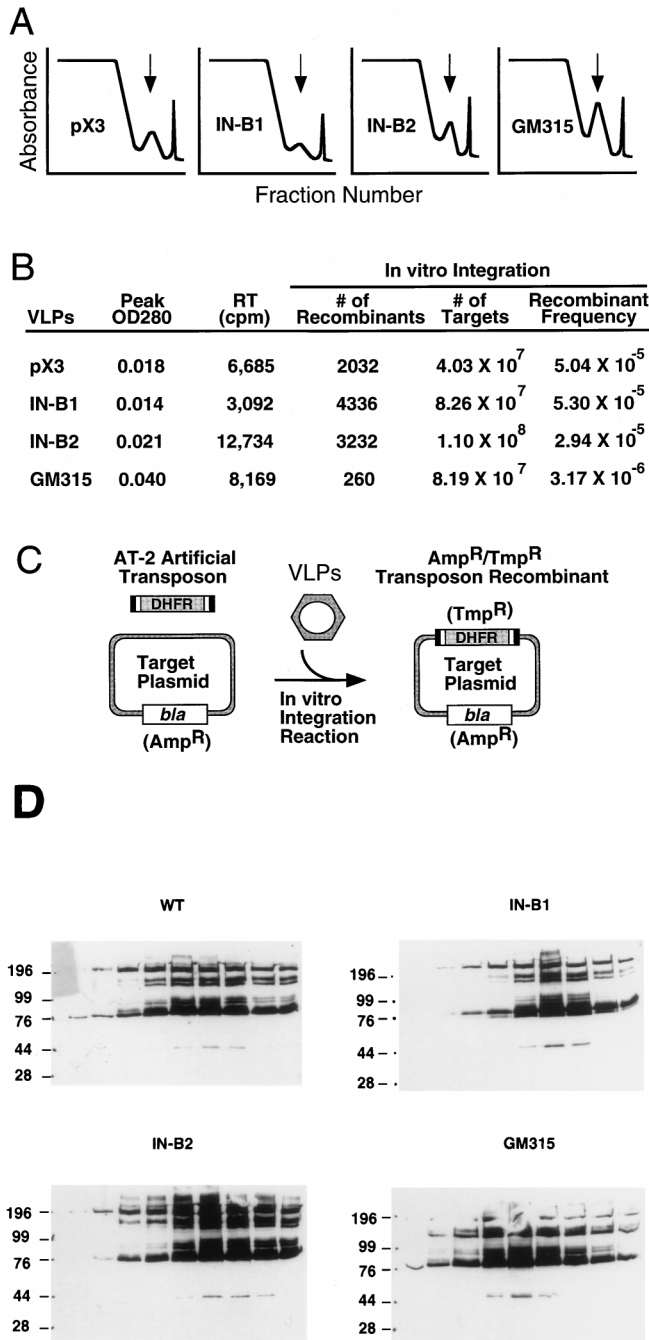


FIG. 5. Analysis of VLPs from wild-type and NLS mutant Ty1 strains. (A) Sucrose gradient fractionation profiles showing VLP peaks. The top of the gradient (20% sucrose) is presented at the left side of the x axis; the bottom of the gradient (70% sucrose) is at the right. The y axis represents absorbance at 254 nm. It is important to note that VLPs could be isolated from wild-type and NLS mutant Ty1 strains (small peaks under vertical arrows). The first peak at the top of the gradient is bulk protein, and the last peak at the right is the extreme bottom of the gradient. (B) Biochemical analysis of VLPs. The peak fraction from each VLP preparation was examined for RT and integration activities in vitro. All of the preparations showed normal RT activity. The activities of positive control avian myeloblastosis virus RT and the negative control, no enzyme (data not shown), were 93,500 and 465 cpm, respectively. VLPs isolated from the wild-type strain (pX3), or the NLS mutant strains (IN B1 and B2) supported normal levels of integration, whereas the negative control IN mutant strain (GM315) was 16-fold less active. OD280, optical density at 280 nm. (C) In vitro integration assay (25). The artificial transposon AT-2 serves as a substrate for Ty1 IN and is integrated into the target plasmid in vitro when VLPs are added. AT-2 carries the *dhfr* gene which confers trimethoprim resistance and

the C terminus of IN which is required for both its nuclear localization and its in vivo retrotransposition function.

An essential NLS in Ty1 IN. Several lines of evidence reveal the Ty1 IN NLS, including immunofluorescence studies on *GAL*-Ty1-produced IN and ectopically expressed tagged IN (*GAL*-E1-IN). Furthermore, independent studies at this laboratory (50a) and in the Garfinkel laboratory (51b) examined Ty1 IN fusions to green fluorescent protein (GFP). In all cases, IN was unambiguously nuclear localized.

Active transport of nuclear proteins depends on the presence of short stretches of basic amino acids that associate with the cytosolic receptor karyopherin (36). Although many proteins contain such sequences, no single consensus has emerged for NLSs. They have, however, been placed into the following two classes: (i) those with a single contiguous stretch of predominately arginine or lysine residues (e.g., simian virus 40 T antigen), and (ii) those having a bipartite structure (e.g., *Xenopus* nucleoplamin) (54). Naturally occurring bipartite NLSs consist of two short stretches of basic residues separated by a spacer of about 10 nonconserved residues. Experiments with artificial bipartite NLSs suggest that they tolerate deletion, insertion, and amino acid changes (48, 54). Our results show that Ty1 IN represents a member of a new class of bipartite NLSs with a much longer spacer element; a similar element was recently reported for HIV-1 IN (32).

We have shown that site-specific block substitution mutations in basic regions of the C terminus of Ty1 IN identified functional NLS sequences. Two different systems were analyzed, *GAL*-E1-IN and *GAL*-Ty1, and two regions important for nuclear localization were revealed in both contexts. In the ectopically expressed IN (*GAL*-E1-IN), it is interesting to note that IN regions B1 and B2 must both be mutated to block nuclear import, suggesting that in this context each basic region is important but that neither is essential for simple nuclear import. However, defects in transposition and localization are evident with either B1 or B2 mutated in the context of the full transposon (*GAL*-Ty1), suggesting that the NLS functions as a single bipartite signal, relying on both basic regions. As shown in the accompanying paper (51b), Moore et al. performed mutagenesis studies similar to ours and identified the same two basic regions near the Ty1 IN C terminus as essential for NLS function.

The difference in behavior of the B1 and B2 mutations in the two different experimental systems used to analyze IN nuclear localization has several possible explanations. One possible explanation stems from the fact that subtly different mutations were made in our study in the *GAL*-E1 and *GAL*-Ty1 plasmids. However, this explanation seems unlikely given that Moore et al. (51b) have shown that a completely different set of specific NLS mutations in the same regions behave similarly; complete deletion of B2 has no deleterious effect in the *GAL*-GFP construct; a localization defect is observed only when this deletion is combined with a substitution mutation(s) in B1. More likely, the relatively subtle defects caused by mutating only B1 or B2 have relatively little effect when IN is expressed

thus serves as a selectable marker when the transposition reaction mixture is transformed into *Escherichia coli*. The target plasmid contains the *bla* gene conferring ampicillin resistance, while transposon recombinants carry both markers. The number of recombinants (Amp^r Tmp^r colonies) divided by the total number of targets recovered (Amp^r colonies) provides the recombinant frequency. (D) Immunoblots of nine peak fractions from each gradient probed with MA8B11, showing similar yields of encapsidated IN protein in the four preparations. Molecular masses are indicated on the left in kilodaltons. WT, wild type; GM315, negative control IN mutant strain.

TABLE 2. Putative bipartite NLSs in other yeast ORFs

ORF name (gene name)	(Putative) NLS sequence ^a	Distance between KKR motifs (aa) ^b	Distance from C terminus (aa) ^c	Protein length (aa) ^d	Function ^e	Localization
Ty1 and Ty2 IN	KKRSL EDNEETEIKVSRD TWNTK NMRSL EPSP KKR	29	5	635	Ty1 integration	Nuclear
YIL126w (<i>STH1</i>)	KKRVY DDGLTEEOFL EA VE..... DDNMS LEDAI KKR	27	274	1,359	Helicase-related; <i>SNF2</i> homolog	Nuclear
YBR142w (<i>MAK5</i>)	KKRGH NSIIIGHEK..... TNAL ETL KKKKR	20	2	773	DEAD box helicase; RNA splicing	Nuclear ^f
YOL109c (<i>JNO4</i>)	KKRR RSKKINKLTDG..... QIRIN HVS KKR	23	94	151	Transcription factor	Nuclear ^f
YKL114c (<i>APN1</i>)	KKRAG TRRKKATAEP..... SDN ILL SQWTKR	23	5	368	AP endonuclease; DNA repair	Nuclear ^f
YKL129c (<i>MYO3</i>)	KKRS AKIKKATPDANKK EVG ISDLTL LLSK ... ISDE SIN ENL KKR	37	1,212	1,272	Myosin nonessential; synthetic lethal with <i>myo5</i>	Cytoskeleton? ^f
YKL021c (<i>MAK11</i>)	KKRDA ETADIGDQSEV ES FT EE EL KKI MFGE KKL NKKR	33	13	468	Essential; 60S ribosome subunit biogenesis	Nucleolar? ^f
YML034w	KKRDP SDDDWSESN S EN K IND NK HL N LL S SD SE IE OD Y OKAKR	39	438	657	Unknown	Unknown
YML093w	KKRV IDDEDK EV DT LP GW GE WAG AG S K PK NKKR	29	93	900	Unknown	Unknown
YKL082c	KKRKL ES EQE Q DE IA SD SD ME DI SD LE EN S KKR	31	186	435	Unknown	Unknown
L8300.12	KKRR RGNGHLSR..... KE KK R W ER AD KKR	20	840	899	Unknown	Unknown
YOR004w	KKRKL GP K AP FL SV KK KK V NS P SD EV KD KE DT SK E KKR	35	21	296	Unknown	Unknown
YGL136c	KKRQ AVFT N VHK FK PD A SR DE SK ET Y I GL KKR	29	11	320	Unknown	Unknown
YGL133w	KKRKP TE V ND S EN S SE ED KK..... GON VT S ET S KKR	29	912	1,264	Unknown	Unknown
YDL094c	KKR NGSM HM Y T AS V FL LF SN AY SD..... FS R V NE FL KKR	34	49	169	Unknown	Unknown
YGL141w	KKRKL SS LR FR KL E IL LE LP FF IP FE ERV ... DL F Y MF IA L DK KKR	36	380	910	Unknown	Unknown

^a The complete yeast genome sequence was probed by using the pattern matching program CoSMoS (2a) with the query (KKR)_{N₂₀₋₄₀}(KKR). The results are summarized in this table. KKR motifs are in boldface type, and acidic residues in the spacer region are underlined. The motifs are arbitrarily split 10 amino acid residues upstream of the downstream KKR motif. The proteins derived from known genes are listed first, with the anonymous ORFs listed below.

^b aa, amino acids.

^c Distance of putative NLS from predicted C terminus of protein. aa, amino acids.

^d Total length of protein (predicted). aa, amino acids.

^e Functional data obtained from Saccharomyces Genome Database (URL: <http://genome-www.stanford.edu/Saccharomyces/>).

^f Localization inferred from function.

ectopically at high levels in the form of an E1-IN fusion protein (or IN-GFP) and much more of an effect in the context of the intact *GAL*-Ty1, from which less IN protein is synthesized. Moreover, in the latter case, the Ty1 IN may be burdened by association with Ty1 DNA (see below) and other preintegration complex components, whereas the E1-IN fusion protein is unlikely to be so encumbered. An interesting observation, noted by Moore et al., is that an acidic region (amino acids LED) immediately adjacent to the basic residue in IN B1 may contribute to this region's ability to act as an NLS. We substituted this region in the IN B1 mutation in the *GAL*-Ty1 vector but not in the *GAL*-E1-IN vector. In fact, acidic amino acids, as well as other nonbasic components, have been shown to be important for NLS activity (22, 48). Thus, our understanding of the rules governing NLS behavior is incomplete. Mutations that inactivate a signal can sometimes be rescued by a secondary mutation at a remote site, and neutral and acidic residues can play critical roles when they flank a short cluster of basic residues (48).

The role of the Ty1 IN NLS may differ from that of conventional NLSs, which need only facilitate the import of a single protein. The IN NLS may facilitate the import of ~6,000 bp of cDNA (and possibly additional preintegration complex proteins). This might explain why in the *GAL*-E1-IN constructs, mutations in just one of the basic clusters results in proper nuclear localization, but in the *GAL*-Ty1-*TRP1* plasmid, a mutation in just one cluster is sufficient to prevent IN from localizing to the nucleus. Direct evidence for the ability of large nucleic acids to transit the NE comes from the influenza virus system. Influenza virus RNA is imported via an RNA-associated protein(s) containing an NLS and requires host cell transport factors (karyopherin, Ran, and p10) in an in vitro assay (52).

The Ty1 IN has a long spacer region and is flanked by identical SKKR motifs. Does this represent a special class of NLS or will a generic NLS function equally well? Placing the NLS from the yeast histone H2B gene directly upstream of the inactivated IN B1 mutant in Ty1 failed to restore its transposition (data not shown), suggesting that the specific sequence of the Ty1 IN NLS (or its immediate context) may be important. While it is tempting to speculate about the exact nature of these signals, the main purpose for using the *GAL*-E1-IN system was to facilitate indirect immunofluorescence assays. The more relevant information comes from the situation in which IN is produced from an intact Ty1 element, and we reached similar conclusions by analyzing both systems; both basic clusters are important for the correct nuclear localization of IN.

Relationship to retroviral mechanisms. Ty1 and retroviruses are often compared because of the similarity between their life cycles and their mechanism of integrating cDNA into the respective host genomes. One major difference is that Ty1 is not infectious and does not exit the host cell. But the mechanism by which Ty1 enters the nucleus of yeast in which the NE does not break down may represent an important functional similarity to lentiviruses such as HIV-1.

HIV-1 can infect both dividing and nondividing cells (47). During the course of infecting nondividing cells, for example, terminally differentiated macrophages, the HIV-1 preintegration complex is actively transported across the intact interphase NE (15). The matrix (MA) protein (17 kDa) contains an NLS capable of targeting reporter proteins to the nuclei. Viruses with mutations in the MA NLS and the Vpr gene can be propagated only in dividing cells, not in cell cycle-arrested cells (14). Vpr was subsequently shown to contain an NLS that is sufficient for HIV-1 import in a strain containing a defective MA NLS (37). MA associates with karyopherin α (Rch1) (33).

Thus, it appears that MA and Vpr proteins contain redundant NLSs that are sufficient but not individually necessary for HIV-1 nuclear transport. There are conflicting reports regarding the possible presence of an NLS in HIV-1 IN (33, 40, 44). Gallay et al. (33) have suggested that at least one other NLS-bearing protein plays a role in import of HIV-1 and that HIV-1 IN has sequences that resemble karyophilic motifs. Recent evidence suggests that HIV-1 IN does have a C-terminal, bipartite NLS, mutations in which block nuclear import and association with karyopherin. However, viruses mutated in this region appear to have a generalized replication defect, even in dividing cells (32). Thus, the relevance of the HIV-1 IN NLS to HIV-1 replication is not clear at the present time.

The avian sarcoma virus (ASV) retrovirus IN contains a potent NLS which also somewhat resembles the Ty1 NLS in that it has lysine residues at the ends of a ~20-amino-acid-residue signal (44). However, because ASV replicates only in dividing cells, it is not clear whether the NLS is actually required for viral replication because ASV preintegration complexes can gain nuclear access during mitosis. Perhaps the NLS in ASV IN allows it to transit the NE during nonmitotic parts of the cell cycle or allows replication in some select subset of cell types. In contrast to the retroviral cases, the importance of the IN NLS is very clear in the Ty1 system; the yeast NE never breaks down and IN is the only element-encoded protein whose function is required in the nucleus. Furthermore, IN is the only Ty1 protein known to contain an NLS. The reports of NLSs in retroviral INs hint at the possibility of a conserved import mechanism.

The finding that IN has an NLS that specifies its nuclear localization predicts that it interacts with the host cell protein import machinery and the NPC. In an initial survey to identify genes encoding host factors important for IN uptake, we measured transposition rates in several mutants affected in NPC components and cytoplasmic nuclear transport factors (data not shown). As expected, several mutants did indeed show reductions in transposition rates, but since these mutants were also defective in nuclear import or RNA export, it is possible that the effect on transposition is indirect. At least one mutant, *nup170Δ*, showed a significant reduction in transposition. Other studies have shown that *nup170Δ* has no effect on gross import or RNA export (2, 42), suggesting that this effect on transposition may be rather specific to IN uptake. Nevertheless, these data support the idea that the IN B1 and B2 sequences serve as NLSs that interact with import machinery rather than simple nuclear retention signals.

Cell biology of the Ty1 integration machinery. Ty1 VLPs apparently serve as the center for organizing the complex steps of reverse transcription, a process assumed to occur in the cytoplasm, where most if not all VLPs are found (34, 50). Supporting this, a Gag-LacZ fusion protein was shown to localize to ~50 cytoplasmic dots, consistent with the number of VLPs expected per cell (18). Furthermore, VLPs contain the components necessary for an in vitro integration reaction (30). Thus far, it has not been possible to isolate a complex capable of efficiently mediating Ty1 integration other than the VLP itself. However, since IN is expected to be inside the VLP, together with the Ty1 cDNA, and IN contains an NLS, some specific event must expose the IN NLS to the nuclear import machinery. The actual identity and composition of the species of preintegration complex that transits the NE are not clear; at a minimum, the complex must include cDNA and IN, and at a maximum, it could include the entire VLP. The latter possibility seems unlikely for several reasons: (i) the IN NLS is probably inaccessible in intact VLPs, (ii) VLPs are very large, and (iii) in HIV-1, there is good evidence for a preintegration

complex intermediate containing a subset of the core virion proteins and viral cDNA (16, 31). Two possible processes that seem likely are the following: (i) docking of Ty1 VLPs at a nuclear pore, coupled with transfer of the preintegration complex into the nucleus, requiring the NLS, and (ii) a specific VLP disassembly process that releases a preintegration complex competent for subsequent nuclear uptake. A third possibility is that IN and the DNA are transported to the nucleus independently, but this seems unlikely based on retroviral precedents. The discovery and characterization of the IN NLS provide a useful new tool to further explore mechanistic aspects of preintegration complex uptake into yeast nuclei as well as the further steps involved in defining and acquiring targets in chromosomal DNA.

ACKNOWLEDGMENTS

We thank Y. Eby for assistance with DNA sequencing, L. Davis for generously supplying strains and antibodies, and S. Moore and D. Garfinkel for communicating data prior to publication. We thank A. Davis and R. Jensen for supplying the antihexokinase antibody. We are especially grateful to D. Bassett and P. Hieter for assistance with the prerelease CoSMoS program and members of the Boeke lab for their helpful comments, especially J. Smith and E. Hoff. M.A.K. thanks R. Skibbens for suggestions and support.

This work was supported in part by NIH grants 5-T-32-CA09139-22 to M.A.K. and PO1-CA16519 to J.D.B.

REFERENCES

- Adams, A. E., and J. R. Pringle. 1991. Staining of actin with fluorochrome-conjugated phalloidin. *Methods Enzymol.* **194**:729-731.
- 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.
- Bassett, D. Personal communication.
- Belcourt, M. F., and P. J. Farabaugh. 1990. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell* **62**:339-352.
- Boeke, J. D., D. J. Eichinger, D. Castrillon, and G. R. Fink. 1988. The *Saccharomyces cerevisiae* genome contains functional and nonfunctional copies of transposon Ty1. *Mol. Cell. Biol.* **8**:1432-1442.
- Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink. 1985. Ty elements transduce through an RNA intermediate. *Cell* **40**:491-500.
- Boeke, J. D., F. Lacroute, and G. R. Fink. 1984. A positive selection for mutants lacking orotidine 5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.* **197**:345-346.
- Boeke, J. D., and S. B. Sandmeyer. 1991. Yeast transposable elements, p. 193-261. In J. R. Broach, E. Jones, and J. Pringle (ed.), *The molecular and cellular biology of the yeast Saccharomyces: genome dynamics, protein synthesis, and energetics*, vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Boeke, J. D., C. A. Styles, and G. R. Fink. 1986. *Saccharomyces cerevisiae* *SPT3* gene is required for transposition and transpositional recombination of chromosomal Ty elements. *Mol. Cell. Biol.* **6**:3575-3581.
- Boeke, J. D., H. Xu, and G. R. Fink. 1988. A general method for the chromosomal amplification of genes in yeast. *Science* **239**:280-282.
- Braiterman, L. T., and J. D. Boeke. 1994. In vitro integration of retrotransposon Ty1: a direct physical assay. *Mol. Cell. Biol.* **14**:5719-5730.
- Braiterman, L. T., and J. D. Boeke. 1994. Ty1 in vitro integration: effects of mutations in *cis* and in *trans*. *Mol. Cell. Biol.* **14**:5731-5740.
- Braiterman, L. T., G. M. Monokian, D. J. Eichinger, S. L. Merbs, A. Gabriel, and J. D. Boeke. 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: phenotypic analysis. *Gene* **139**:19-26.
- Brookman, J. L., A. J. Stott, P. J. Cheeseman, N. R. Burns, S. E. Adams, A. J. Kingsman, and K. Gull. 1995. An immunological analysis of Ty1 virus-like particle structure. *Virology* **207**:59-67.
- Bukrinsky, M. I., S. Haggerty, M. P. Dempsey, N. Sharova, A. Adzhubel, L. Spitz, P. Lewis, D. Goldfarb, M. Emerman, and M. Stevenson. 1993. A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* **365**:666-669.
- Bukrinsky, M. I., N. Sharova, M. P. Dempsey, T. L. Stanwick, A. G. Bukrinskaya, S. Haggerty, and M. Stevenson. 1992. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc. Natl. Acad. Sci. USA* **89**:6580-6584.
- Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, W. G.

- Tarpley, and M. Stevenson. 1993. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. USA* **90**:6125–6129.
17. Bukrinsky, M. I., T. L. Stanwick, M. P. Dempsey, and M. Stevenson. 1991. Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection. *Science* **254**:423–427.
 18. Burns, N., B. Grimwade, P. B. Ross-Macdonald, E. Y. Choi, K. Finberg, G. S. Roeder, and M. Snyder. 1994. Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. *Genes Dev.* **8**:1087–1105.
 19. Burns, N. R., H. R. Saibil, N. S. White, J. F. Pardon, P. A. Timmins, S. M. H. Richardson, B. M. Richards, S. E. Adams, S. M. Kingsman, and A. J. Kingsman. 1992. Symmetry, flexibility and permeability in the structure of yeast retrotransposon virus-like particles. *EMBO J.* **11**:1155–1164.
 20. Clare, J. J., M. Belcourt, and P. J. Farabaugh. 1988. Efficient translational frameshifting occurs within a conserved sequence of the overlap between the two genes of a yeast Ty1 transposon. *Proc. Natl. Acad. Sci. USA* **85**:6816–6820.
 21. Dales, S., and Y. Chardonnet. 1973. Early events in the interaction of adenoviruses with HeLa cells. IV. Association with microtubules and the nuclear pore complex during vectorial movement of the inoculum. *Virology* **56**:465–483.
 22. Dang, C. V., and W. M. F. Lee. 1988. Identification of the human *c-myc* protein nuclear translocation signal. *Mol. Cell. Biol.* **8**:4048–4054.
 23. Davis, L. I. 1995. The nuclear pore complex. *Annu. Rev. Biochem.* **64**:865–896. (Review.)
 24. Davis, L. I., and G. R. Fink. 1990. The NUP1 gene encodes an essential component of the yeast nuclear pore complex. *Cell* **61**:965–978.
 25. Devine, S. E., and J. D. Boeke. 1994. Efficient integration of artificial transposons into plasmid targets *in vitro*: a useful tool for DNA mapping, sequencing and genetic analysis. *Nucleic Acids Res.* **18**:3765–3772.
 26. 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.
 27. Dingwall, C., J. Robbins, S. M. Dilworth, B. Roberts, and W. D. Richardson. 1988. The nucleoplasmic nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J. Cell Biol.* **107**:841–849.
 28. Doheny, K. F., P. K. Sorger, A. A. Hyman, S. Tugendreich, F. Spencer, and P. Hieter. 1993. CTF13 is an essential component of the *Saccharomyces cerevisiae* kinetochore. *Cell* **73**:761–774.
 29. Dworetzky, S. I., and C. M. Feldherr. 1988. Translocation of RNA-coated gold particles through the nuclear pores of oocytes. *J. Cell Biol.* **106**:575–584.
 30. Eichinger, D. J., and J. D. Boeke. 1988. The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. *Cell* **54**:955–966.
 31. Farnet, C. M., and W. A. Haseltine. 1991. Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J. Virol.* **65**:1910–1915.
 - 31a. Gabriel, A., and J. D. Boeke. Unpublished data.
 32. Gallay, P., T. Hope, D. Chin, and D. Trono. 1997. HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc. Natl. Acad. Sci. USA* **94**:9825–9830.
 33. Gallay, P., V. Stitt, C. Mundy, M. Oettinger, and D. Trono. 1996. Role of the karyopherin pathway in human immunodeficiency virus type 1 nuclear import. *J. Virol.* **70**:1027–1032.
 34. Garfinkel, D. J., J. D. Boeke, and G. R. Fink. 1985. Ty element transposition: reverse transcriptase and virus-like particles. *Cell* **42**:507–517.
 35. Garfinkel, D. J., A.-M. Hedge, S. D. Youngren, and T. D. Copeland. 1991. Proteolytic processing of *pol-TYB* proteins from the yeast retrotransposon Ty1. *J. Virol.* **65**:4573–4581.
 36. Gorlich, D., and I. W. Mattaj. 1996. Nucleocytoplasmic transport. *Science* **271**:1513–1518.
 37. Heinzinger, N. K., M. I. Bukrinsky, S. A. Haggerty, A. M. Ragland, V. Kewalramani, M.-A. Lee, H. E. Gendelman, L. Ratner, M. Stevenson, and M. Emerman. 1994. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc. Natl. Acad. Sci. USA* **91**:7311–7315.
 38. Higuchi, R., B. Krummel, and R. Saiki. 1988. A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* **16**:7351–7367.
 39. Hovland, P., J. Flick, M. Johnston, and R. A. Sclafani. 1989. Galactose as a gratuitous inducer of GAL gene expression in yeasts growing on glucose. *Gene* **83**:57–64.
 40. Jones, K. S., J. Kulkosky, and A. M. Skalka. 1991. Analyses of HIV integration components, p. 21–26. *In A. Kumar* (ed.), *Advances in molecular biology and targeted treatment for AIDS*. Plenum Press, New York, N.Y.
 41. Kalinich, J. F., and M. G. Douglas. 1989. *In vitro* translocation through the yeast nuclear envelope. Signal-dependent transport requires ATP and calcium. *J. Biol. Chem.* **264**:17979–17989.
 42. Kenna, M. A., J. G. Petranka, J. L. Reilly, and L. I. Davis. 1996. Yeast N1e3p/Nup170p is required for normal stoichiometry of FG nucleoporins within the nuclear pore complex. *Mol. Cell. Biol.* **16**:2025–2036.
 43. Kubai, D. F. 1975. The evolution of the mitotic spindle. *Int. Rev. Cytol.* **43**:167–227.
 44. Kukolj, G., K. S. Jones, and A. M. Skalka. 1997. Subcellular localization of avian sarcoma virus and human immunodeficiency virus type 1 integrases. *J. Virol.* **71**:843–847.
 45. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685.
 46. Lauerermann, V., K. Nam, J. Trambley, and J. D. Boeke. 1995. Plus-strand strong-stop DNA in retrotransposon Ty1. *J. Virol.* **69**:7845–7850.
 47. Lewis, P. F., and M. Emerman. 1994. Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.* **68**:510–516.
 48. Makkerh, J. P. S., C. Dingwall, and R. A. Laskey. 1996. Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr. Biol.* **6**:1025–1027.
 49. Mellor, J., A. M. Fulton, M. J. Dobson, W. Wilson, S. M. Kingsman, and A. J. Kingsman. 1985. A retrovirus-like strategy for the expression of a fusion protein encoded by yeast transposon Ty1. *Nature* **313**:243–246.
 50. Mellor, J., M. H. Malim, K. Gull, M. F. Tuite, S. McCreedy, T. Dibbayawan, S. M. Kingsman, and A. J. Kingsman. 1985. Reverse transcriptase activity and Ty RNA are associated with virus-like particles in yeast. *Nature* **318**:583–586.
 - 50a. Merkulov, G., and J. D. Boeke. Submitted for publication.
 51. Merkulov, G. V., K. M. Swiderek, C. B. Brachmann, and J. D. Boeke. 1996. A critical proteolytic cleavage site near the C terminus of the yeast retrotransposon Ty1 Gag protein. *J. Virol.* **70**:5548–5556.
 - 51a. Monokian, G. M., L. T. Braiterman, and J. D. Boeke. 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: mutations, transposition, and dominance. *Gene* **139**:9–18.
 - 51b. 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.
 52. O'Neill, R. E., R. Jaskunas, G. Blobel, P. Palese, and J. Moroiianu. 1995. Nuclear import of influenza virus RNA can be mediated by viral nucleoprotein and transport factors required for protein import. *J. Biol. Chem.* **270**:22701–22704.
 53. Pluta, A. F., N. Saitoh, I. Goldberg, and W. C. Earnshaw. 1992. Identification of a subdomain of CENP-B that is necessary and sufficient for localization to the human centromere. *J. Cell Biol.* **116**:1081–1093.
 54. Robbins, J., S. M. Dilworth, R. A. Laskey, and C. Dingwall. 1991. Two interdependent basic domains in nucleoplasmic nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**:615–623.
 55. Roe, T., T. C. Reynolds, G. Yu, and P. O. Brown. 1993. Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* **12**:2099–2108.
 56. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed., vol. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 57. Vallette, F., E. Merge, A. Reiss, and M. Adesnik. 1989. Construction of mutant and chimeric genes using the polymerase chain reaction. *Nucleic Acids Res.* **17**:723–733.
 58. Xu, H., and J. D. Boeke. 1987. High frequency deletion between homologous sequences during retrotransposition of Ty elements in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **84**:8553–8557.