The EMBO Journal vol.12 no.12 pp.4885-4895, 1993

Novel gene expression mechanism in a fission yeast retroelement: Tf1 proteins are derived from a single primary translation product

Henry L.Levin[1,3], Daniel C.Weaver[2] and Jef D.Boeke

Department of Molecular Biology and Genetics, The Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205 and [1]National Institutes of Child Health and Human Development, Bethesda, MD 20892, USA
[2]Present address: Program in Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80302, USA
[3]Corresponding author

Communicated by G.Simchen

In sharp contrast to the single ORF of the *Schizosaccharomyces pombe* retrotransposon Tf1, retroviruses and most retrotransposons employ two different ORFs to separately encode the Gag and Pol proteins. The different ORFs are thought to allow for overexpression of the Gag protein relative to Pol protein presumed necessary for the assembly of functional retrovirus particles and virus-like particles (VLPs). The results of *in vivo* experiments designed to detect the transposition of Tf1 show that Tf1 is indeed active and can insert itself into the host genome via a true retrotransposition process. Thus, a paradox emerged between the lack of any obvious means of overexpressing Tf1 Gag protein and the demonstrated functionality of the element. Epitope tagging experiments described here confirm that the Tf1 large ORF is intact and that there is no translational or transcriptional mechanism used to overexpress the Tf1 Gag protein. In addition, we used sucrose gradients and antisera specific for Tf1 capsid (CA) and integrase (IN) to show that the Tf1 proteins do assemble into uniform populations of macromolecular particles that also co-sediment with Tf1 reverse transcription products. This evidence suggests that Tf1 proteins form VLPs without using the previously described mechanisms that retroviruses and retrotransposons require to overexpress Gag proteins.

Key words: nmt1 promoter/retrotransposition/retrotransposon/*Schizosaccharomyces pombe*/Tf1

Introduction

Both retroviruses and retrotransposons assemble particle structures that encapsulate protease (PR), reverse transcriptase (RT) and integrase (IN). The structural shell of these particles is composed of capsid (CA) protein which is encoded by the *gag* gene (Varmus and Brown, 1989; Boeke and Sandmeyer, 1991). Retrovirus and retrotransposon particles possess a common architectural solution to providing the appropriate conditions for the complex series of steps comprising reverse transcription of RNA and integration of the cDNA into the host genome. A molar excess of Gag protein molecules is assembled around a small number of enzymes that are responsible for protein processing, reverse transcription and integration (Varmus and Brown, 1989). The function of these particles depends on assembly conditions, which include a molar excess of Gag molecules (Felsenstein and Goff, 1988; Belcourt and Farabaugh, 1990; Weaver *et al.*, 1990; Xu and Boeke, 1990; Park and Morrow, 1991; Dinman and Wickner, 1992). Interestingly, retroviruses and most retrotransposons use similar mechanisms to provide a molar excess of Gag. Ty1 and Ty2 (Belcourt and Farabaugh, 1990; Xu and Boeke, 1990), Ty3 (Hansen and Sandmeyer, 1990), *gypsy* (Marlor *et al.*, 1986), RSV (Jacks and Varmus, 1985), MMTV (Hizi *et al.*, 1987) and HIV (Jacks *et al.*, 1988) are examples of elements that use a reading frameshift at the end of Gag to produce excess Gag protein. A rare exception is MuLV that has no frameshift between *gag* and *pol*, and uses a nonsense suppression mechanism to overexpress Gag (Jamjoom *et al.*, 1977). An exceptional case is *copia*, the *Drosophila* retrotransposon with a single *gag/pol* ORF that uses a transcriptional splicing mechanism instead of a translational strategy to overexpress Gag. *copia* depends on the splicing of its mRNA to remove Pol sequences from the transcripts (Miller *et al.*, 1989; Brierley and Flavell, 1990; Yoshioka *et al.*, 1990).

In sharp contrast to the multiple ORFs required for translational methods of modulating Gag and Pol protein expression, the retrotransposon Tf1 of *Schizosaccharomyces pombe* has one single ORF encoding both Pol and putative Gag proteins (Levin *et al.*, 1990). High levels of transposition activity of this element have been observed (Levin and Boeke, 1992) (and results described here) despite the apparent lack of any reading frame separation that could allow for overexpression of a capsid protein. This seeming paradox led us to ask whether Tf1 employs a novel mechanism of retrotransposition that uses a different expression strategy. The related elements SURL (Springer *et al.*, 1991) and Tf2 (Weaver *et al.*, 1993) also have a single ORF, suggesting that this subfamily of *gypsy* family elements may share a common gene expression strategy.

Tf1 was recently isolated from a wild-type strain of *S.pombe* that contains >20 Tf1 copies in its genome (Levin *et al.*, 1990). The complete sequence of Tf1 reveals a single large ORF with amino acid sequences that are similar to the sequences of the PR, RT and IN domains of retroviruses and retrotransposons. Upstream of the PR sequence there is coding sequence sufficient to encode a protein of ~260 amino acids. We show here that this protein is expressed as a 27 kDa protein; this expression depends on the Tf1 PR activity. We refer to this protein as CA because of its similarity in molecular weight and genomic position to the CA of retroviruses and retrotransposons. Tf1 also possesses two long terminal repeats (LTRs) of 358 nucleotides which flank the body of the element and serve to initiate as well as terminate the transcription of the full-length, 4.4 kb mRNA (Levin *et al.*, 1990). The results of *in vivo* experiments designed to detect the transposition of Tf1-*neo*

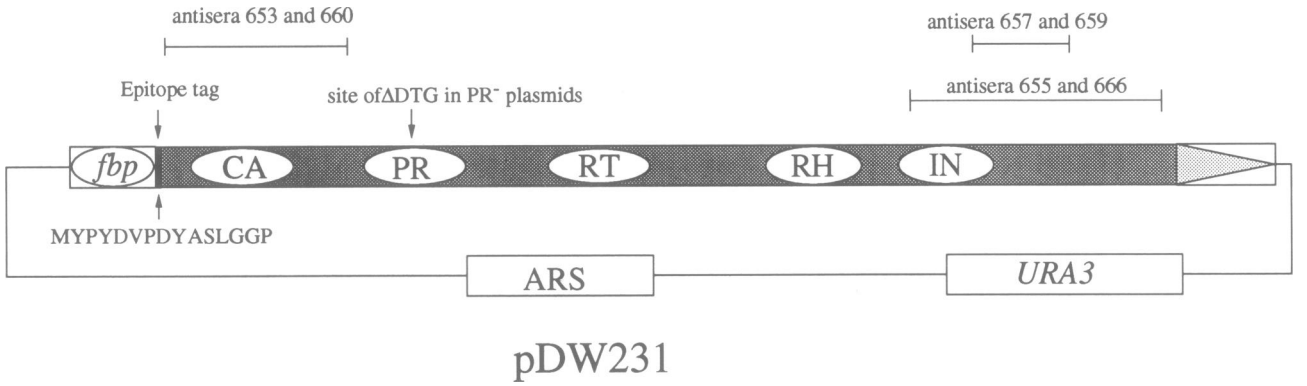

Fig. 1. The structure of pDW231, a Tf1 expression plasmid with an N-terminal epitope tag. *fbp* indicates the position of the *S.pombe* promoter and its mRNA leader from *fbp*. The position of the epitope tag is indicated by a thick line and the sequence is shown. The dark shaded rectangle represents the single full-length ORF of Tf1. The positions of the protein domains CA, PR, RT, RH and IN are marked. The portions of the Tf1 ORF that were used to raise antisera are labeled with the number of the corresponding rabbits. The light-shaded triangle represents the Tf1 3′ LTR and the boxes depict the *S.cerevisiae URA3* gene and *S.pombe ars1*.

fused to a strong *S.pombe* promoter show that Tf1 is indeed active and can insert itself into the host genome via a true retrotransposition process (Levin and Boeke, 1992). Taken together, the sequence and the transposition data represent a paradox between the lack of any obvious means of overexpressing CA and the demonstrated functionality of the element. Either Tf1 uses a unique mechanism to overexpress CA, or the transposition process is novel and does not require the typical ratio of Gag and Gag/Pol proteins.

To resolve this paradox, we first designed an experiment to confirm the presence of a single Tf1 ORF. An epitope tag was fused to the N terminus of the Tf1 ORF within an *S.pombe* expression plasmid. Epitope-specific antibodies were used to probe immunoblots of protein from the plasmid-bearing strain. The results described here demonstrate that Tf1 coding sequence is contained within a single ORF and that there is no translational or transcriptional mechanism to overexpress Gag. Thus Tf1 expression is unique among the retroviruses and retrotransposons whose gene expression has been studied. We also performed sucrose gradient analysis on *S.pombe* extracts to determine whether Tf1 proteins do assemble in virus-like particles (VLPs) that contain Tf1-specific nucleic acid. Analysis of these gradients with antisera raised against two Tf1 proteins demonstrated that they do assemble into a uniform population of particles that co-sediment with a product of Tf1 reverse transcription.

Results

Tf1 RNA is translated into a single primary product

In order to identify the primary translation products encoded by the Tf1 ORF, we fused a hemagglutinin epitope tag to the N terminus of the Tf1-107 ORF (Kolodziej and Young, 1991). pDW231, the construct used to express the epitope tag−Tf1 fusion protein, contains the promoter and a 272 bp leader from the *S.pombe* fructose bisphosphatase (*fbp*) gene fused to the epitope-tagged Tf1 ORF (Figure 1). The presence of the 15 amino acid tag enabled us to detect the proteins by immunoblot analysis, using a monoclonal antibody that interacted with the hemagglutinin epitope. The size of the detected species was expected to correspond to the length of the primary translation products. Were translational termination to occur after CA, a 28 kDa species should be detected on the immunoblot. If the full-length 1340 amino

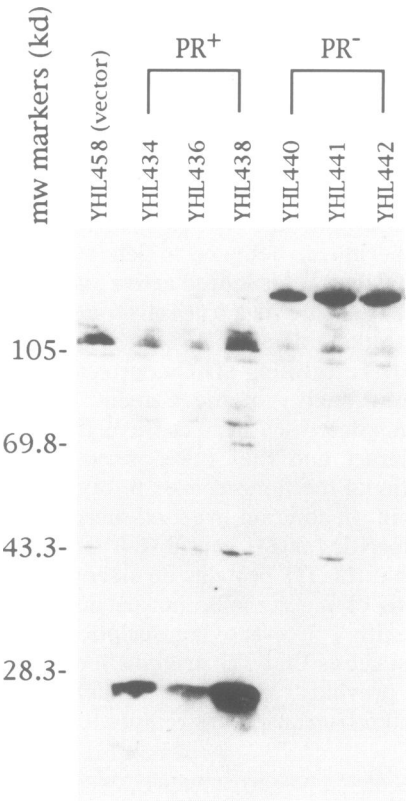

Fig. 2. Immunoblot of epitope-tagged Tf1 proteins. The antibody 12CA5 was used to probe lanes of an SDS−polyacrylamide gel that contained protein expressed from Tf1 plasmids bearing an epitope tag at the N terminus of the full-length ORF. PR^+ and PR^- indicate which strains carried a mutation in the Tf1 PR. YHL440, YHL441 and YHL442 are independent transformants that contained the mutant PR plasmid while YHL434, YHL436 and YHL438 are wild-type. As discussed in the plasmid construction section, each of the duplicate transformants were also produced from separate ligations of independently synthesized PCR products. YHL458 contains the *fbp* promoter plasmid that lacks both the epitope and the Tf1 sequence. The bands migrating between the 28 kDa and full-length proteins are not associated with Tf1 because they can be observed in YHL458, the vector only strain which lacks the epitope tag altogether.

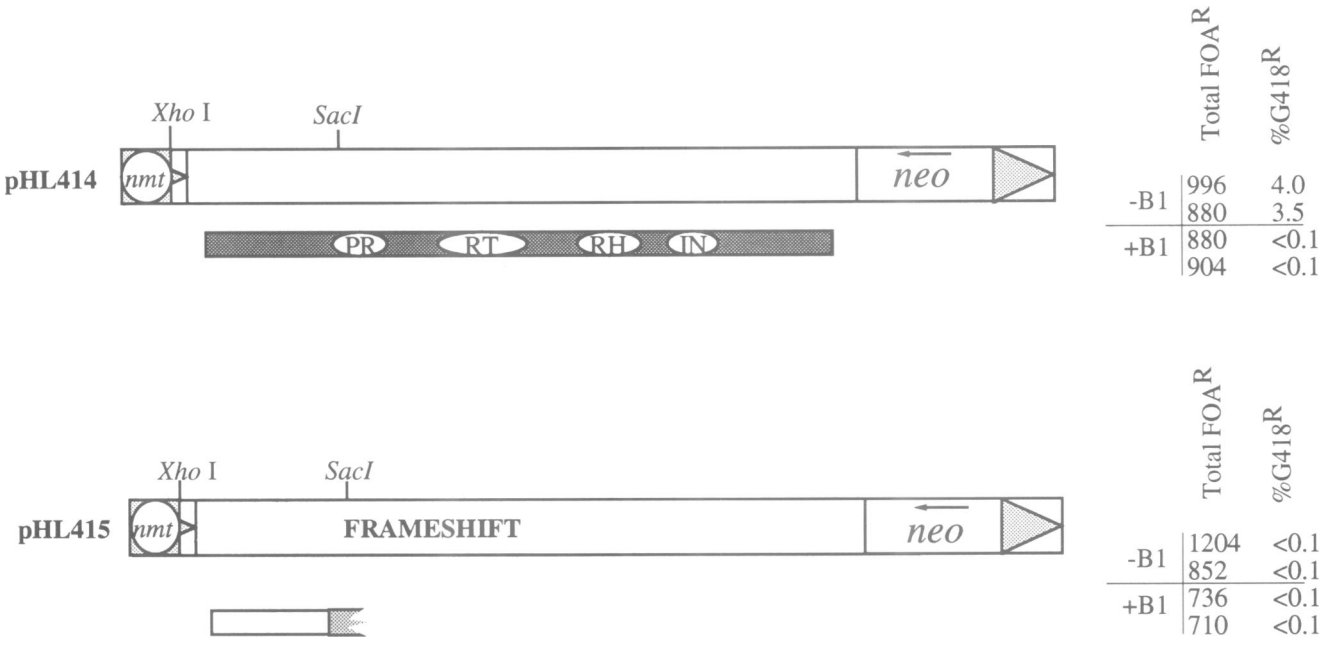

Fig. 3. Transposition assay of pHL414 and pHL415. The structure of Tf1 in the *nmt1* expression plasmids pHL414 and pHL415 is shown. The circle labeled *nmt* represents the *nmt1* promoter and the shaded triangles indicate the position of the Tf1 LTRs. The ORFs are shown as long rectangles that contain ovals indicating the protein domains, PR, RT and IN. The broken rectangle in pHL415 represents the mutation that disrupts the ORF. *neo* shows the position of the bacterial *neo* gene and the arrow above it indicates the direction of *neo* transcription. The data shown in the right margin are the results from the transposition assay of two independent transformants of pHL414 and pHL415 grown in the presence and absence of vitamin B1 (thiamine).

acid polypeptide is the only primary translation product, the blot should show a single 140 kDa band. To avoid the complexities introduced by the activity of the transposon-encoded PR, which cleaves the primary translation products into smaller products, we used a PR mutant. The highly conserved PR active site residues DTG (Hutchison *et al.*, 1991) within Tf1 were deleted in order to block processing events expected to occur between the Tf1 proteins. Should Tf1 transcripts be translated into an unexpectedly short product due to frameshifts, stop signals, mRNA editing or splicing events, less than full-length proteins would be observed on the protein blots. Otherwise, only a single full-length translation product should be observed. Strain YHL232 was transformed with plasmids that contained the *fbp* promoter and hemagglutinin tag fused to the Tf1 ORF with (pDW231) and without (pDW232) the PR mutation. We grew these strains under conditions that induced *fbp* transcription and made protein extracts that were loaded onto an SDS−polyacrylamide gel and transferred to the blot shown in Figure 2. A single protein species of 140 kDa was detected by the hemagglutinin antibody in the PR mutants, whereas the wild-type strains all produced a 28 kDa species corresponding in size to the putative CA protein encoded by the Tf1 ORF upstream of the PR domain. The molecular weight of this protein corresponds well to the weights of CA proteins from retroviruses such as Moloney (30 kDa) and RSV (27 kDa) (Hatfield *et al.*, 1992) as well as *gypsy* family retrotransposons Ty3 (26 kDa) (Hansen *et al.*, 1992) and TED (37 kDa) (Lerch and Friesen, 1992). These demonstrate that the epitope-tagged Tf1 transcripts are translated into a single, full-length primary translation product. The apparent absence of splicing, editing or frameshifting mechanisms to overproduce the N-terminal domain of the Tf1 ORF indicate that Tf1 is unique among the LTR-containing retrotransposons and retroviruses that have been examined for protein expression. Given that Tf1 lacks these typical mechanisms of CA protein overexpression, we determined whether Tf1 proteins assemble into VLPs. In order to address the issue of assembly biochemically, it was necessary to improve transcriptional efficiency of our vectors as well as to generate antibodies that recognize native Tf1 proteins.

The nmt1 promoter fusion gives rise to regulated transposition

To optimize the expression of Tf1 proteins, we fused the inducible *S.pombe nmt1* promoter to Tf1 at the first base of the transposon transcript (Figure 3). The *nmt1* fusion plasmid was designed to initiate transcription at the same position as endogenous Tf1 elements. The production of *nmt1* transcripts can be reduced 80-fold by growing the cells in 10 μM thiamine (Basi *et al.*, 1993). To compare the levels of maximum expression of the *nmt1* plasmid with the *fbp* fusion, we measured the levels of *in vivo* transposition produced by the *nmt1* plasmids and compared them with those produced by plasmids bearing the *fbp* promoter. Figure 3 shows the structure of Tf1−*neo* in these plasmids. Transposition is measured by growing cells with the transposon plasmid in the absence of thiamine to induce transposition and spreading these cells onto FOA plates to select for colonies of cells that lack the plasmid (see Levin and Boeke, 1992). These colonies can then be replica plated onto G418 plates to determine which colonies acquired genomic copies of Tf1−*neo*. To demonstrate that these events are not due to simple homologous recombination, we inserted a frameshift mutation into the Tf1 PR to assess what fraction of the colonies became resistant to G418 due to transposition. As seen in Figure 3, YHL1089 and YHL1090, the independent transformants of the plasmid with the intact

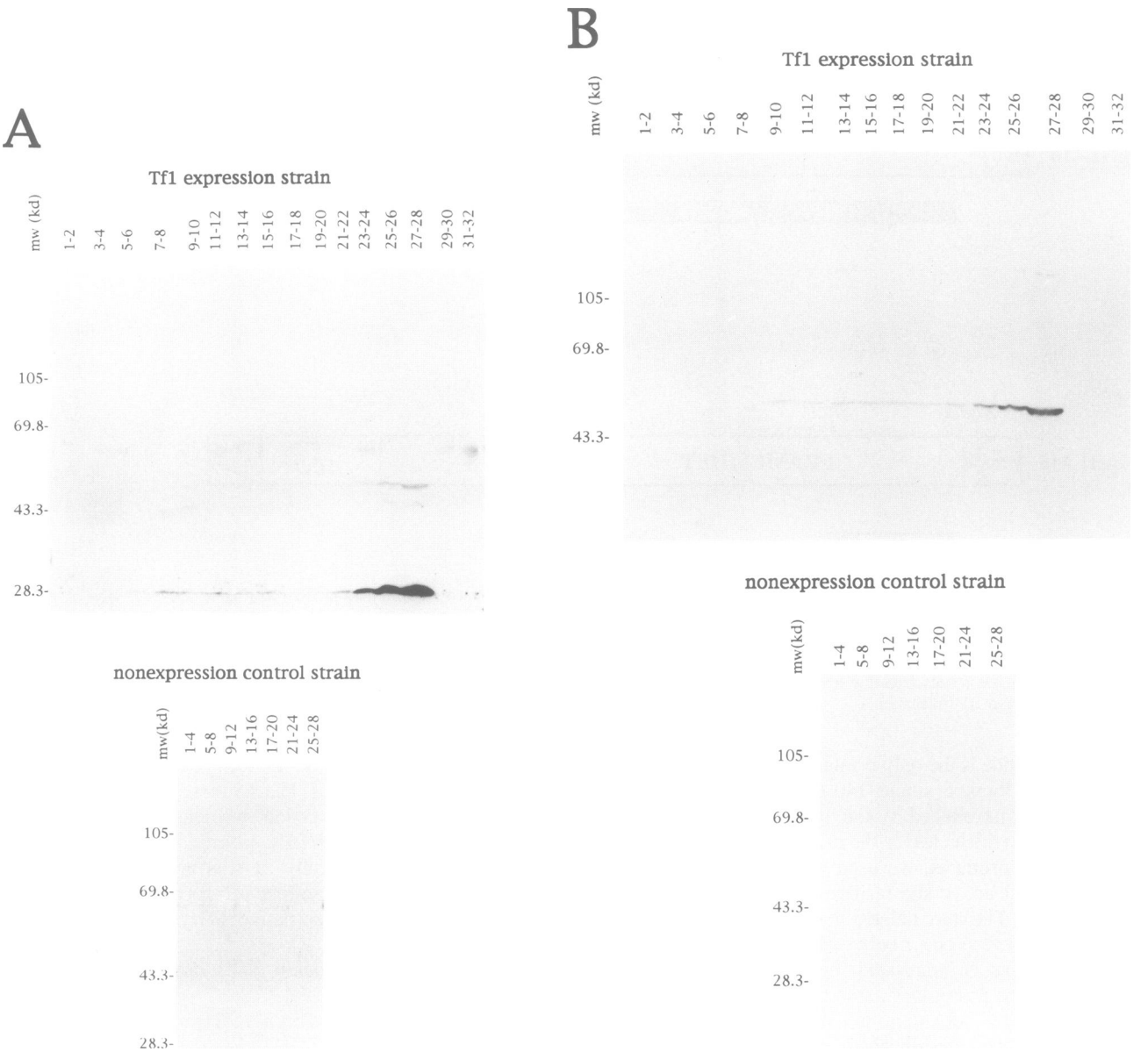

Fig. 4. (**A**) Sucrose step gradient analysis. The top panel is an immunoblot of a 7.5% SDS−polyacrylamide gel with fractions from a sucrose step gradient that contained extract from a Tf1 expression strain (YHL1051). The antibody probe was anti-Gag serum CA653. The fractions were pooled in pairs and TCA precipitated before being loaded onto the gel. The bottom panel was also probed with the anti-Gag serum and is an immunoblot of pooled fractions from a control strain (YHL1032) that has a vector without Tf1 sequence. (**B**) These are immunoblots of the same fractions in panel A probed with an anti-IN antiserum IN666. The top panel is derived from the Tf1 expression strain and the bottom panel is derived from the vector control strain.

Tf1 ORF, pHL414, gave rise to G418 resistant colonies at a frequency of 3.5−4.0% when grown in the absence of thiamine supplement. When these identical strains were grown in the presence of 10 μM thiamine, <0.1% transposition was observed since no G418 resistant colonies grew. The strains with the PR frameshift, YHL1091 and YHL1092, also produced no G418 resistant colonies regardless of whether they were grown in 10 μM thiamine. The 4.0% transposition from the *nmt1* promoter represents a 20-fold increase over that produced by the *fbp* promoter system when similar conditions are used (Levin and Boeke, 1992). Therefore, the *nmt1* promoter plasmids were chosen for the Tf1 biochemical analyses.

Evidence for VLPs from sucrose step gradients

The fact that Tf1 encodes a CA protein of the appropriate size and genomic position suggests that it has a VLP intermediate. To determine whether Tf1 proteins assemble into VLPs, we used the antibodies anti-CA653 and anti-IN666 to characterize the behavior of the Tf1 CA and IN proteins in sucrose step gradients. Conditions have been developed for isolating large macromolecular species such as Ty1-VLPs in a sucrose step gradient (Eichinger and Boeke, 1988, 1990). Figure 4A contains protein blots of 1.2 ml gradient fractions of extracts from strains that had either the Tf1 expression plasmid or the vector, pFL20. Both panels were probed with antiserum specific for Tf1 Gag

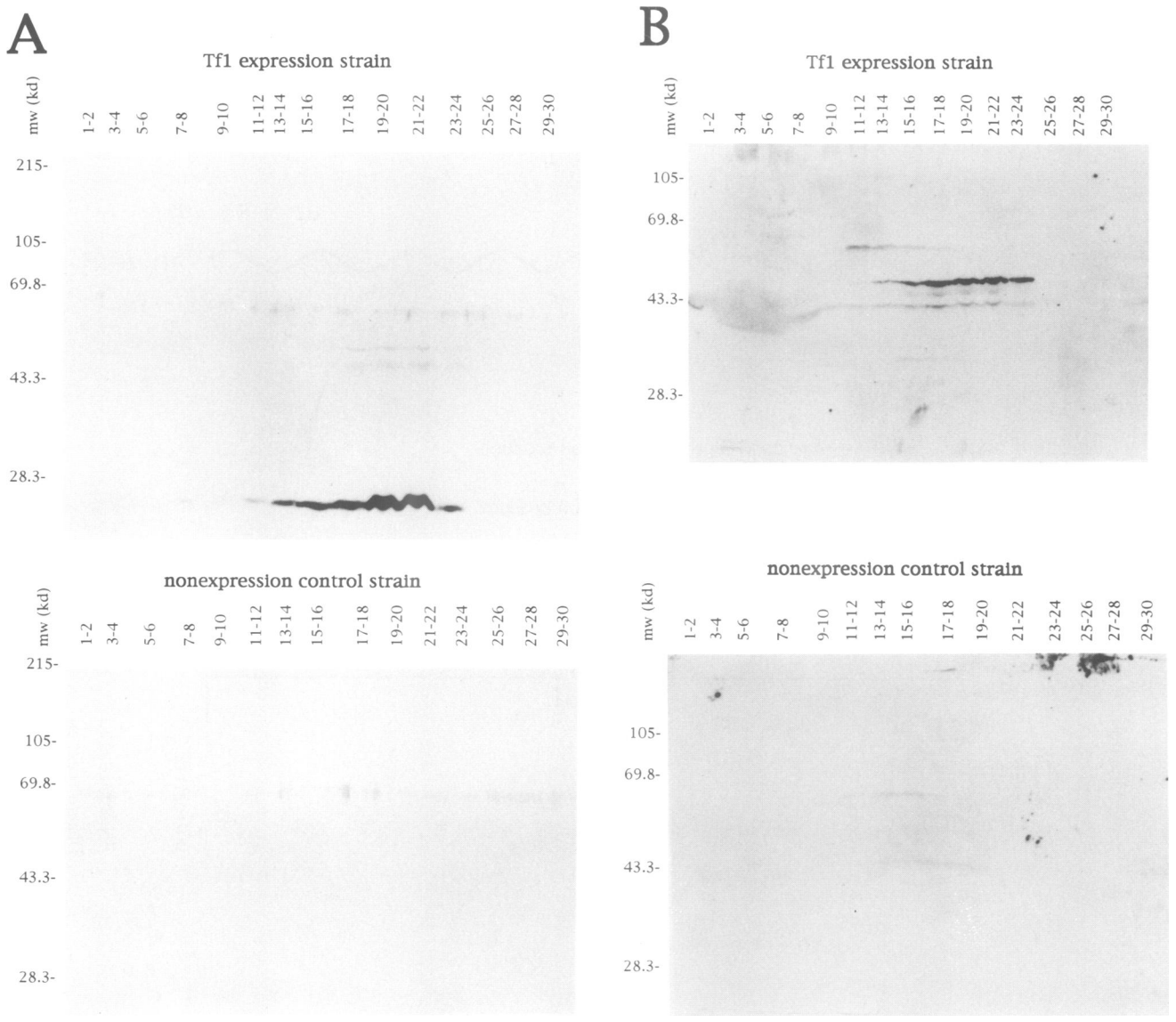

Fig. 5. (**A**) Linear sucrose gradient analysis. These are immunoblots of pooled sucrose gradient fractions from a Tf1 expression strain (top) and a vector control strain (bottom) probed with anti-Gag antiserum. (**B**) These are the same fractions shown in panel A probed with anti-IN antiserum IN666. The top panel contains fractions from a Tf1 expression strain whereas the bottom panel contains fractions from a vector control strain.

(anti-CA653). The top panel has fractions from the Tf1 expression strain and shows the 27 kDa CA protein in fractions 23–28. The bottom blot contains protein from the control strain and does not show the 27 kDa band. Figure 4B has an identical set of filters as Figure 4A except that they were probed with antiserum specific for the IN protein (anti-IN666). In this case, the expression strain (top) showed the 56 kDa IN in fractions 23–28, whereas the vector only control produced no Tf1 protein signal. Thus, the majority of the 56 kDa IN and 27 kDa CA proteins were found in fractions 23–28 which correspond to the 30–70% sucrose interface. The presence of both proteins within the same fractions at the top of the 70% step demonstrates that Tf1 proteins assembled into large particle species. Although Tf1 proteins co-sedimented to the 70% sucrose step, we could not rule out the possibility that the Tf1 proteins had formed large, random aggregates and not a uniform population of VLPs.

Tf1 proteins: behavior in linear sucrose gradients

To determine whether the fast sedimenting Tf1 proteins sediment with uniform distributions, we loaded extracts from the Tf1 expression and control strains onto 30 ml 20–70% linear sucrose gradients that were spun for 24 h (Braiterman *et al.*, 1993). This type of gradient can help determine whether Tf1 proteins are assembled into random aggregates or VLPs. Figure 5A contains the protein blots of 1.2 ml sucrose gradient fractions from the Tf1 expression strain (top) and the vector control strain (bottom), probed with anti-Gag antiserum. Whereas a small amount of the 27 kDa CA protein from the expression strain was found within fractions 1–6, the majority was concentrated in a second peak, fractions 15–24. The control strain did not produce the 27 kDa band. Figure 4B contains the identical set of filters as shown in Figure 4A except they were probed with anti-IN antisera. The 56 kDa IN band formed a peak within fractions 15–24 of the expression strain (top) and was not

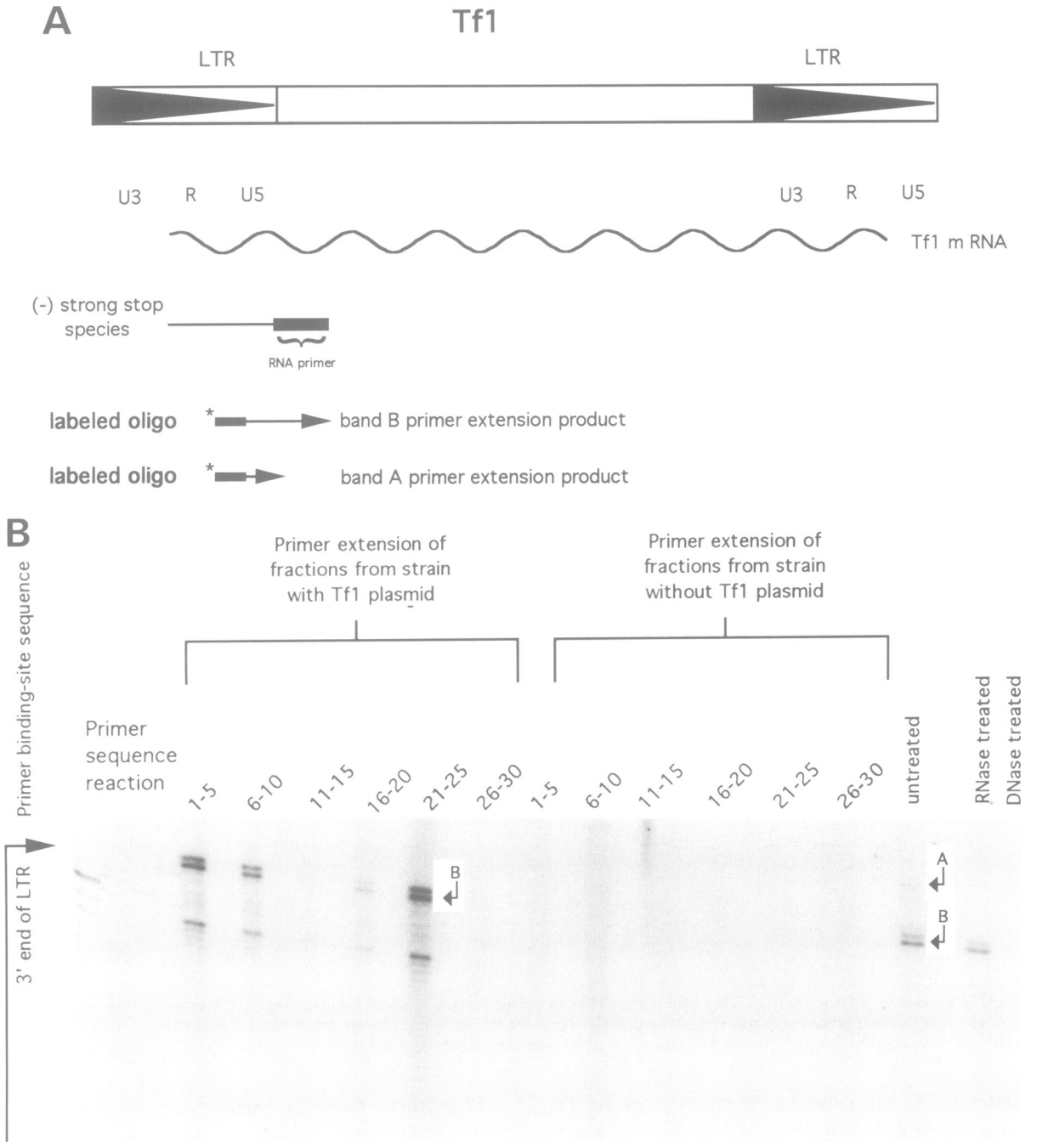

Fig. 6. (**A**) Cycle primer extension used to detect Tf1 reverse transcripts. Besides RT and Gag proteins, retrovirus and retrotransposon particles also contain various species of RNA and DNA. The (−) strong-stop species has a 5′ end composed of a small RNA primer. The (−) strong-stop is the first RT product made during synthesis of the double-stranded DNA genome of the element. In an effort to detect Tf1 nucleic acids in the fractions suspected of containing particles, we used multiple cycles of primer extension from a labeled oligo that hybridizes to bp 265−281 of the LTR (JB341). The presence of (−) strong-stop molecules in these fractions would cause a specific length primer extension product to be observed that should not be produced from any molecules that are not the products of reverse transcription. Full-length reverse transcripts could also give the same length of product, although these might be expected to lack an RNA primer. The primer extension species representing bands A and B of panel B are labeled here. (**B**) Primer extension products. Primer extension products produced from nucleic acid extracted and RNased from either the expression strain (YHL1051) or the vector control strain (YHL1032) are shown. The sequence ladder was produced using the same 'kinased' oligo so that the migration of the extension products is the same as the bands in the sequence lanes. The bands (doublet) marked B were seen in most of the lanes from the expression strain. The lane marked untreated was not pretreated with RNase and contained a band marked A that was nine bases longer than the A species. Apparently, Taq polymerase can copy at least a small number of ribonucleotide template residues under the conditions used. The DNase treated lane contained nucleic acid pretreated with DNase before the extension reaction.

observed in the fractions from the control strain (bottom). Thus, both Tf1 proteins tested co-sedimented within this 20−70% linear sucrose gradient.

Reverse transcription products are associated with CA and IN proteins in linear sucrose gradients

Because retroviral and retrotransposon particles contain

nucleic acid as well as the CA, RT and IN proteins, we examined whether any of the 20–70% sucrose gradient fractions were associated with Tf1 nucleic acid. We initially examined sucrose fractions for Tf1 RNA by blot analysis but were unable to distinguish between Tf1 transcripts associated with ribosomes and transcripts packaged in VLPs (not shown). Instead we examined these fractions for products of reverse transcription. The primary distinguishing property between the products of reverse transcription and genomic or plasmid Tf1 DNA sequences is that the reverse transcripts have discrete ends. Many retroviral and retrotransposon particles contain, as a predominant DNA species, the minus strong-stop replication intermediate as diagrammed in Figure 6A. To assay gradient fractions for minus strong-stop and related DNA species such as full-length reverse transcripts, we used a modified primer extension assay that included 50 extension cycles of a ^{32}P-labeled oligo. Since the positions of minus strong-stop priming in other retrotransposons and retroviruses has always been found adjacent to the 5′ LTR, the size of the primer extension products can be predicted and are shown in Figure 6A. The 30 fractions from each 20–70% sucrose gradient were combined into six pools of five fractions each. The pools were concentrated, treated with proteinase K, phenol extracted, ethanol precipitated and treated with RNase. Pooled material from expression and control strains was primer extended (Figure 6B). Two peaks of extension products were observed in the expression strain gradient with oligo JB341 while none were seen in the control gradient. The extension signals were produced in the lanes of pooled fractions 1–5, 6–10 and 21–25, suggesting that some Tf1 nucleic acid species were associated with Tf1 proteins in a fast sedimenting form, but others were slow sedimenting. The top two bands (doublet) marked B are composed of DNA that ends either at the last base of the LTR or the downstream adjacent base. The larger band is probably due to the propensity of Taq polymerase to add an untemplated base onto the end of an extension product (Mead *et al.*, 1991). When the material from fractions 21–25 were pretreated with DNase the signal disappeared, demonstrating that the Tf1 nucleic acid template is composed of DNA. Material treated with neither RNase nor DNase included an additional species (band A) nine nucleotides longer than band B, suggesting an LTR template covalently attached to at least nine nucleotides of RNA.

To confirm that oligo JB341 was hybridizing to the appropriate sequence within the Tf1 LTR, we used an additional oligo, JB340, designed to bind upstream of the JB341 site. The extension products of the two oligos should differ in size by 54 nucleotides because the 5′ end of JB340 is 54 bases upstream of the 5′ end of JB341. Because the difference in size of the extension products made by JB340 versus JB341 was ~50 bases, we concluded that all the oligos hybridized specifically to their predicted positions within the Tf1 sequence (data not shown).

Discussion

The translation products of Tf1

The presence of a single ORF in Tf1 suggests that a unique regulatory mechanism may be required to express the appropriate ratios of Gag to Pol proteins. Therefore, we determined the nature of Tf1 primary translation products because the great majority of retroviruses and retrotransposons contain translational stop signals between the sequence encoding the Gag and Pol proteins (Hatfield *et al.*, 1992). The position and efficiency of these stop signals serve to restrict the synthesis of Pol proteins, thus allowing for the expression of a 10- to 20-fold molar excess of CA protein (Hatfield *et al.*, 1992). The overexpression of CA protein may very well be uniformly essential for particle assembly since the only example of a retrovirus or an LTR retrotransposon that is known to produce particles and lacks a translational stop signal is *copia*, a *Drosophila* element that nonetheless overexpresses CA via an mRNA splicing event (Yoshioka *et al.*, 1990).

The hemagglutinin epitope fused to the N terminus of the Tf1 ORF enabled us to observe directly the synthesis of Tf1 primary translation products from elements that either have or lack functional PR. The absence of any less than full-length Tf1 product confirmed the lack of internal stop codons or frameshifts. In addition, this result also indicated that there is no transcriptional splicing, editing or translational mechanism that functions to overexpress CA. If CA is available in molar excess, there must be a novel mechanism for this that has not been observed to occur during the assembly of other retrotransposons or retroviruses. One possible mechanism for this would be ordered processing of Gag-Pol proteins, with release of CA followed by selective degradation of Pol proteins.

Expression and detection of Tf1 proteins

The incorporation of the *nmt1* promoter into the Tf1 vectors greatly improved the expression of the Tf1 proteins and allowed us to characterize their assembly. One result of the increased level of expression was the 20-fold higher levels of transposition (4%) that the *nmt1* promoter achieved when compared with the *fbp* promoter. We have now optimized growth conditions so that *nmt1*–Tf1 transposition frequencies are 20% (E.Sweeney and H.Levin, unpublished). In addition to the increased expression, the *nmt1* promoter also results in the first regulated transposition system of *S.pombe*, with transposition frequencies that were below the level of detection (<0.1%) when cells were grown in 10 μM thiamine. This regulated system will facilitate the study of retrotransposons in *S.pombe* as well as their application in the design of genomic mapping and tagging techniques.

The characterization of Tf1 protein association required the use of antibodies that recognized specific functional proteins and not just the hemagglutinin epitope. Antisera produced from *trpE* protein fusions expressed in bacteria were specific for Tf1 CA and IN individually. Not only did these antisera allow us to study Tf1 protein assembly but they also enabled us to identify the sizes of the processed CA and IN proteins expressed by a functional Tf1 element.

The assembly of the Tf1 proteins and nucleic acid

The formation of VLPs by retrotransposons has now been observed for *copia* (Shiba and Saigo, 1983), Ty1 (Garfinkel *et al.*, 1985), Ty3 (Hansen *et al.*, 1992) and TED (Lerch and Friesen, 1992). The similarity in structure of VLPs to retrovirus particles is great and includes the general size and shape as well as the presence of full-length transcripts, tRNA primers, replication intermediates and double-stranded reverse transcripts (Boeke and Sandmeyer, 1991; Gabriel and Boeke, 1993). Whether or not Tf1 proteins assemble

into VLPs is a significant question because of the apparent lack of a biosynthetic mechanism for producing a molar excess of CA.

The antisera raised against the Tf1 proteins used in conjunction with the *nmt1* expression plasmids allowed us to detect readily Tf1 proteins. The co-fractionation of the Tf1 IN and CA in sucrose step gradients demonstrated that the proteins do assemble into large macromolecular complexes and provided the first evidence that IN and CA associate with each other. Since the overwhelming majority of the IN and CA sedimented to the 30–70% interface, the efficiency of this assembly process appears to be high. When the same conditions of centrifugation are used on extracts containing either Ty1 or Ty3 proteins, the particles behave as the Tf1 proteins did, sedimenting at the 30–70% interface (Eichinger and Boeke, 1988; Hansen *et al.*, 1992). However, because step gradients tend to concentrate protein complexes of heterogeneous sizes at the interfaces, this experiment alone does not rule out the possibility that Tf1 proteins were forming random aggregates instead of particles. To address this issue we used linear sucrose gradients to determine the uniformity of the Tf1 complexes. The result that CA and IN co-sediment in fractions 15–24 with similar concentration profiles provided further evidence that these proteins are co-assembled. In the absence of an antiserum that interacts specifically with RT, and because we have failed to detect RT activity, we have no direct evidence that RT is co-fractionating with the IN and CA proteins. Given the overall similarity of Tf1 to other retrotransposons, we think it likely that RT is among the proteins assembled into particles.

Furthermore, the 50-cycle primer extension reactions clearly detected the presence of Tf1 reverse transcripts in the same sucrose gradient fractions that contained CA and IN. The co-fractionation of CA, IN and the Tf1 DNA in fractions 21–25 of the YHL1051 extract and not in fractions from YHL1032 (the strain without the Tf1 expression plasmid) provides stronger evidence for the presence of Tf1 VLPs. The nine nucleotides on the 3′ end of the Tf1 DNA that are degraded by RNase indicate that Tf1 reverse transcription begins with an RNA primer, confirming the identity of the Tf1 template as true retrotransposon reverse transcripts. The primer for reverse transcription is unlikely to be a conventional full-length tRNA because it does not end in CCA. Searches of databases have not identified likely candidates for the primer (E.Hoff, H.Levin and J.Boeke, unpublished data). We suspect that the Tf1 nucleic acid in fractions 1–10 must represent reverse transcripts that were once associated with the VLPs but have since become dissociated.

Models for the assembly of protein particles

The conclusion that Tf1 can form particles is surprising since the assembly of retroviruses and retrotransposons requires separate Gag and Pol ORFs to produce a molar excess of CA protein. In contrast, we have shown that Tf1 Gag and Pol proteins are expressed within the same primary translation product and as a result, in equal molar ratios. Two distinctly different types of particles could exist that would be consistent with our observations. One possibility is that the Tf1 particles do contain an excess of CA protein with respect to the Pol proteins. Perhaps the extra polymerase proteins transiently associate with the particles, but ultimately are excluded by steric limitations. This would require that a significant

Table I. Yeast strains used

Strain	Genotype	Source	Plasmid description[a]
YHL232	*h⁻, ura4-D18*	Grimm *et al.* (1988)	
YHL434	YHL222/pHL229	this paper	Epitope tagged Tf1 with PR^+
YHL436	YHL232/pHL230	this paper	Epitope tagged Tf1 with PR^+ (from independent PCR reaction)
YHL438	YHL232/pHL231	this paper	Epitope tagged Tf1 with PR^+ (from independent PCR reaction)
YHL440	YHL232/pHL232	this paper	Epitope tagged Tf1 with PR^- (from independent PCR reaction)
YHL441	YHL232/pHL232	this paper	Epitope tagged Tf1 with PR^- (independent transformant of YHL440)
YHL442	YHL232/pHL233	this paper	Epitope tagged Tf1 with PR^- (from independent PCR reaction)
YHL458	YHL232/pCYH20	this paper	Vector control has *fbp* promoter but lacks Tf1 sequence
YHL1032	*ura4-294, leu1-32*/pFL20	this paper	Vector control
YHL1051	*ura4-294, leu1-32*/pHL411	this paper	*nmt1* promoter fused to Tf1
YHL1089	*ura4-294, leu1-32*/pHL414	this paper	*nmt1* promoter fused to Tf1–*neo*
YHL1090	*ura4-294, leu1-32*/pHL414	this paper	Independent transformant of YHL1089
YHL1091	*ura4-294, leu1-32*/pHL415	this paper	*nmt1* promoter fused to Tf1–*neo* that has frameshift in PR
YHL1092	*ura4-294, leu1-32*/pHL415	this paper	Independent transformant of YHL1091

[a]All plasmids are described here except pCYH20 (Hoffman and Winston, 1989) and pFL20 (Losson and Lacroute, 1983).

Table II. Oligonucleotides used

Name	Sequence (5′–3′)	Use
JB68	CCAAGGTCTGGTAGGAAG	Construction of pWD223
JB106	CTATAGAGCTCTCAGACAAGGTATGGAGC	Construction of pWD226
JB107	CATATCTCGAGCATATGAAAAACTCATCACAG	Construction of pWD223 and pWD226
JB108	TATGTACCCATACGATGTTCCAGATTACGCTAGCTTGGGTGGTCC	Epitope tag insert
JB109	TAGGACCACCCAAGCTAGCGTAATCTGGAACATCGTATGGGTACA	Epitope tag insert
JB278	GATGGATCCCGCCATAAAAGACAG	*nmt1* promoter fusion
JB279	TTATTCACTCGAGGATATGCCAGGA	*nmt1* promoter fusion
JB340	GTTATGAGCTATATAAA	Primer extension
JB341	ATACTCAGTTGTTACTC	Primer extension

amount of proteolytic processing precedes assembly. In retroviruses, processing takes place during or after assembly. A second possibility is that Tf1 particles contain equimolar amounts of Pol proteins. These unique particles would presumably be much bulkier with less available space within the core. The observation that RSV and *copia* particles package equimolar amounts of CA and PR provides a precedent for molar excesses of a Pol-encoded protein. Whether Tf1 VLPs contain equal molar amounts of CA and Pol proteins or Pol proteins are specifically degraded, studies of the VLP assembly mechanism of this unique retrotransposon are likely to be quite interesting.

Materials and methods

Media

The *S.pombe* minimal liquid and plate media were composed of EMM (Bio 101) prepared as recommended by the manufacturer. Where indicated, amino acids and pyrimidines were supplemented individually at a concentration of 250 μg/ml. Ten micromolar vitamin B1 (thiamine) was added to EMM plates when indicated. FOA (PRC Inc., Gainesville, FL) plates were made by adding 1 g/l FOA to EMM supplemented with 250 μg/ml uracil. YE-G418 plates contained 500 μg/ml Geneticin (Gibco) and were prepared as described by Levin and Boeke (1992).

Strains and plasmid constructions

The yeast strains used in this paper are listed in Table I. The oligonucleotides used are listed in Table II. LE392 is an *Escherichia coli* strain with the genotype *supE*44 *supF*58 *hsdR*514 *galK*2 *galT*22 *metB*1 *trpR*55 *lacY*1 (Borck *et al.*, 1976). pHL411 contains the *nmt1* promoter (Maundrell, 1990) fused to Tf1 at the first base of the Tf1 transcript in a vector derived from pFL20 (Losson and Lacroute, 1983). pHL411 was generated by ligating three fragments together: the *Nhe*I−*Bgl*II fragment of pCYH20 containing the plasmid backbone, *ars1*, the stabilization fragment and the *Saccharomyces cerevisiae URA3* gene (Hoffman and Winston, 1989), a *Bam*HI−*Xho*I PCR fragment of the *nmt1* promoter, and the *Xho*I−*Nhe*I fragment from pHL404 that included an intact Tf1 with a *Bgl*II site previously introduced downstream of the ORF. We used oligos JB278 and JB279 to produce an *nmt1* fragment by PCR that included bases −1162 to +1 with base +1 falling within the *Xho*I recognition sequence. The *Xho*I sites were positioned within the *nmt1* and Tf1 fragments such that the first base of *nmt1* transcript was fused to the first base of the Tf1 transcript in order to produce a mRNA indistinguishable from natural Tf1 mRNA. pHL412 is identical to pHL411 except it received the *Xho*I−*Nhe*I fragment from pHL405 which contains a frameshift mutation in the PR domain as described earlier (Levin and Boeke, 1992). pHL414 and pHL415 were generated by the insertion of a 1 kb *neo*-containing *Bam*HI fragment (Joyce and Grindley, 1984) into the *Bgl*II sites of pHL411 and pHL412 respectively. pHL431 was constructed to be identical to the Tf1−*neo* plasmid pHL414, except that a frameshift mutation was made in the beginning of IN. The IN frameshift was first made in pHL338 at base 3288 by filling-in a *Bsp*HI site in pHL286 which is identical to pHL297 except for the orientation of the *neo* fragment insert (Levin and Boeke, 1992). Thus pHL431 was produced in a three-piece ligation reaction that included the *Bam*HI−*Xho*I vector fragment of pHL414, the *Xho*I−*Apa*I piece of Tf1 from pHL338, and the *Apa*I−*Bam*HI fragment of pHL414 that contained the *neo* marker in the 3′ end of Tf1.

pHL333, pHL334 and pHL336 were pATH vector (Korner *et al.*, 1991) fusions to various restriction fragments. pHL333 was generated by inserting an *Eco*RI fragment, bp 563−1312, into the *Eco*RI site of pATH1. To generate pHL334, pATH3 was cut with *Eco*RI and the 3′ ends were filled-in with the Klenow fragment of DNA polymerase I to allow for the insertion of bp 3711−4099. Similarly, pHL336 was produced by inserting bp 3553−4495 into *Bam*HI cut pATH1 that had the 3′ ends filled-in.

To construct plasmids with the 15 amino acid hemagglutinin epitope tag an *Nde*I restriction site (CATATG) was introduced into the Tf1 ORF at the start codon using a PCR-based procedure. Simultaneously, a mutation was introduced into the PR domain in one of the constructs. This mutation results in the precise deletion of the amino acid sequence DTG, representing the essential catalytic triad of retroviral PRs (Hutchison *et al.*, 1991). This was carried out by PCR amplification using oligos JB107 and JB68 generating a wild-type *Xho*I−*Sac*I fragment, or JB107 and JB106, generating a PR mutant *Xho*I−*Sac*I fragment. The restriction fragments were subcloned into pBSII digested with *Xho*I and *Sac*I, generating plasmids pDW217 (wild-type) and pDW220 (PR mutant). In a second step, double-stranded oligos JB108 and JB109 were used to insert sequences encoding an 15 amino acid hemagglutinin epitope tag into the *Nde*I sites of the above plasmids generating pDW223 (wild-type) and pDW226 (PR mutant). The final expression plasmids pDW231 (wild-type) and pDW232 (PR mutant) with the *fbp1* promoter driving, epitope-tagged Tf1 transcripts were generated in three-piece ligations that included the *Xho*I−*Bam*HI vector fragment of pCYH21 (Hoffman and Winston, 1989), the *Xho*I−*Sac*I fragments encoding the 5′ end of Tf1 from pDW223 (wild-type) and pDW226 (PR mutant) and the *Sac*I−*Bam*HI 3′ fragment of Tf1.

Protein extractions and VLP preparations

The protein preparations used to produce the material for the epitope-tagged blot were based on the protocol of Xu and Boeke (1990) with the following modifications. The yeast strains were grown in 10 ml cultures of YNB with 3% glycerol (Hoffman and Winston, 1989) to saturation. The cultures were washed twice in and resuspended in 0.4 ml of buffer (15 mM KCl, 10 mM HEPES−KOH, pH 7.8 and 5 mM EDTA), 5 mM DTT and 2 mM PMSF and then lysed in 13 × 100 mm glass tubes with glass beads by vortexing at top speed for 5 min at 4°C. One hundred microliters of buffer were added to each sample and the vortexed supernatant was removed and mixed with an equal volume of 2 × protein sample buffer (20% w/v glycerol, 10% β-mercaptoethanol, 0.125 M Tris−Cl, pH 6.8 and 5% SDS) before they were boiled for 5 min. The tubes were spun for 10 min in a microfuge and the supernatant was loaded onto a SDS−polyacrylamide gel. The total protein concentrations were determined so that equal amounts of protein could be loaded onto each lane of the SDS−polyacrylamide gels.

The preparations of large-scale yeast extracts and the subsequent analysis on sucrose gradients were based on previously published protocols (Garfinkel *et al.*, 1985; Eichinger and Boeke, 1988, 1990). Five hundred millilitres of EMM + 250 μg/ml leucine, 2% glucose medium (Rose *et al.*, 1990) was inoculated at an OD_{600} of 0.1 and grown at 30°C for 2 days to reach stationary phase. The cells were harvested, washed and broken as previously described (Garfinkel *et al.*, 1985; Eichinger and Boeke, 1988, 1990) except that higher (>90%) efficiency of *S.pombe* cell breakage required that the cell suspension be vortexed with the glass beads in a 25 mm glass test tube for ~20 min. Beads loaded to just above the meniscus greatly increased the efficiency of breakage and recovery of VLPs. Two and a half millilitres of harvested cells were resuspended in 5 ml of buffer B−EDTA (15 mM KCl, 10 mM HEPES−KOH, pH 7.8, 5 mM EDTA) containing 3 mM dithiothreitol and 2 mM PMSF. Five millilitres of supernatant recovered from a 10 000 spin (SS34 rotor) of the cell extract were loaded onto either a sucrose step gradient (20 ml of 20%, 5 ml of 30%, 5 ml of 70%; percent solutions are w/v) or a 20−70% linear gradient of sucrose in buffer B−EDTA. The step gradients were spun for 150 min at 25 000 r.p.m. (SW28 rotor) and the linear gradients were spun for 24 h at 25 000 r.p.m. in a Beckman SW28 rotor. Fractions (1.2 ml) were collected and 100 μl from each was precipitated in 10% TCA. The pellets were washed in cold acetone, resuspended in sample buffer, and loaded onto SDS−polyacrylamide gels for immunoblot analysis.

Immunoblots

The protein preparations used to produce the material for the epitope-tagged blot were loaded onto a 10% SDS−polyacrylamide gel with equal amounts of total protein loaded in each lane. Standard electrotransfer techniques were used (Towbin *et al.*, 1979) with nitrocellulose as the membrane. The detection method used was the ECL system as described by the manufacturers (Amersham) except that the secondary antibody, HRP-conjugated donkey anti-rabbit Ig, was used at a dilution of 10 000-fold. The primary monoclonal antibody used for each filter was 6 μg of purified 12CA5 kindly provided by Dr Jeff Corden. The above protocol was also used for blots of sucrose gradient fractions except that the primary antibodies were the polyclonal sera described under antibody preparation.

Preparation of antisera

pHL333 is the pATH1 vector with *trpE* fused to a 748 bp Tf1 restriction fragment that encodes most of the Gag protein. The plasmid pHL334 was made with pATH3 such that *trpE* was fused to a 388 bp restriction fragment that encodes the N-terminal half of IN. The 942 bp fragment inserted into pATH3 to make pHL336 contained greater than two-thirds of the IN sequences. The position of each of the three Tf1 restriction fragments used to produce antigens is shown in Figure 1. Plasmids pHL333, pHL334 and pHL336 were transformed into *E.coli* strain LE392. The resulting strains were grown at 30°C in 100 ml cultures of M9, 1% casamino acids and 100 μg/ml ampicillin until they reached OD_{600} of 0.2 (~9 h of growth) when indoleacrylic acid was added to a final concentration of 20 μg/ml. After an additional 4 h of growth, the cells were pelleted, washed in 10 mM Tris−Cl (pH 8.0) and resuspended in 20 ml of 50 mM Tris−Cl (pH 8.0)

containing 5 mm EDTA and 3 mg/ml lysozyme. After a 2 h incubation on ice, 1.4 ml of 5 M NaCl and 10% NP-40 was added and followed by a 30 min incubation on ice. The cell lysates were sonicated for 30 s to lower viscosity and spun for 10 min in an Eppendorf minifuge. The pellets were resuspended in 500 μl of loading buffer (20% glycerol, 10% β-mercaptoethanol, 0.125 M Tris−Cl, pH 6.8, 5% SDS) and boiled. The supernatants of an additional 5 min spin were loaded onto a 1.5 mm-thick 10% SDS−polyacrylamide preparative gel. The gel was stained by first washing in water for 1 min and then staining in 0.3 M $CuSO_4$ for 5 min (Lee *et al.*, 1987). This incubation was followed by a 2 min wash in water and the gel was stored in water. The TrpE protein fusions were cut from the gel, the acrylamide pieces were extruded through syringe needles and the resulting material was injected into rabbits by Hasleton Research (Denver, PA). Each antigen was injected into two rabbits, each of which had been pretested as negative for any cross-reaction with *S.pombe* proteins. Fusion protein encoded by pHL333 was injected into rabbits 653 and 660, protein from pHL334 was injected into rabbits 657 and 659, and protein from pHL336 was injected into rabbits 655 and 666. Three sets of injections of all three fusions were spaced by 3 weeks and a fourth injection of only the IN fusions was made. The bleeds were made every 3 weeks starting 1 month after the first injections. The expunge bleeds were done 1 month after the last set of injections.

Antibody characterization

The antisera from five individual bleeds of these six rabbits were used to probe separate membrane strips of Immobilon P (Millipore) transferred with crude protein extracted from YLH1051, a strain that contained an *nmt1* expression plasmid of Tf1. The strips probed with antisera anti-CA653 and anti-CA660 consistently produced a band of ~27 kDa whereas anti-IN657, anti-IN659, anti-IN655 and anti-IN666 all detected a species of 56 kDa. The molecular weight value of 27 kDa for the CA protein corresponds well with the value of 28 kDa obtained for the slightly larger epitope-tagged CA from the experiment described above. The two antisera chosen for all the following experiments were anti-CA653 and anti-IN666. Preimmune antisera failed to detect either of these proteins.

To confirm the identity of the 27 and 56 kDa proteins detected by these antibodies, two plasmids were constructed, pHL415 and pHL431, which were identical to the *nmt* expressed Tf1-*neo* plasmid (pHL414) except they contained frameshift mutations in the beginning of PR and IN, respectively. The 56 kDa protein was only produced in the intact ORF strains while none of the frameshift strains produced a signal. The absence of a band in the IN frameshift strains indicates that the 56 kDa species is the IN protein (data not shown). An identical filter was subsequently probed with anti-CA660. The 27 kDa protein was produced in all lanes except for the two from the strain with the PR frameshift which had a slightly larger species of ~35 kDa. The disappearance of the 27 kDa band in the PR frameshift strain coupled with the appearance of a new slightly larger species demonstrated that the 27 kDa protein was CA because the mutation was in the beginning of PR causing defective processing of the Tf1 proteins which would result in a CA protein fused to a fragment of PR.

Transposition assay

Strains bearing plasmids to be tested for transposition frequency were first streaked out for single colonies on EMM + leucine, +thiamine plates. Individual colonies were picked and used to inoculate patches of ~1 cm^2 on EMM + leucine, −thiamine plates. After 2 days growth at 30°C, the patches were used to inoculate 10 ml liquid cultures of EMM + leucine, −thiamine, at a cell density of 0.1 OD_{600}. These cultures were grown for another 2 days at 30°C before plating a 1000-fold dilution on FOA (Boeke *et al.*, 1984) plates that contained EMM with 250 μg/ml leucine and uracil. The FOA resistant colonies were replica plated onto YE-G418 plates to determine the percentage of colonies that had suffered a transposition event.

Cycle primer extension assay for Tf1 DNA

The pools of sucrose gradient fractions were concentrated by first diluting 3-fold in buffer B−EDTA and spinning for 1 h at 35 000 r.p.m. in a SW41 rotor. The pellets were resuspended in 200 μl of buffer B−EDTA overnight. Five microlitres of 1 mg/ml proteinase K were added to 150 μl of the resuspended pellet and the samples were incubated for 30 min at room temperature. Two phenol extractions, a phenol−chloroform extraction and an ethanol precipitation, were performed on these samples (Eichinger and Boeke, 1988). The ethanol pellets were resuspended in 100 μl; 10 μl were used for each primer extension reaction.

Three hundred nanograms of oligos JB340 and JB341 were labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. Each primer extension reaction contained 100 ng of kinased oligo and 10 μl of the nucleic acid sample as well as all the specified components recommended by the manufacturer for a Taq polymerase (Perkin Elmer−Cetus) PCR reaction except that only one oligo was included. The total reaction volume was 50 μl. Fifty cycles of annealing and extensions were used to maximize production of the extension products.

To determine whether the extension templates were sensitive to RNase digestion, the nucleic acid samples were pretreated with RNase A at a concentration of 20 μg/ml. The DNase I (Promega RQ1 RNase-free DNase) digestions were done in nick translation buffer (50 mM Tris−Cl, pH 7.2, 10 mM $MgSO_4$, 100 μg DTT) at an enzyme concentration of 0.1 unit/10 μl reaction for 1 h at 37°C. After DNase digestion the samples were phenol extracted and ethanol precipitated before being loaded onto the denaturing sequencing gel.

Acknowledgements

We gratefully acknowledge the technical support of Erin Sweeney. We thank V.Lauermann for suggesting the use of cycle primer extension to detect reverse transcription intermediates. We thank Dan Voytas for comments on the manuscript. Supported in part by an ACS postdoctoral fellowship and PHS training grant CA09139 from the NIH (H.L.L.); a Howard Hughes Undergraduate Summer Research Fellowship (D.C.W.); ACS Faculty Research Award FRA-366 and PHS grant CA16519 from NIH (J.D.B.).

References

Basi,G., Schmid,E. and Maundrell,K. (1993) *Gene*, in press.
Belcourt,M.F. and Farabaugh,P.J. (1990) *Cell*, **62**, 339−352.
Boeke,J.D. and Sandmeyer,S.B. (1991) In Broach,J., Pringle,J. and Jones,E. (eds), *The Molecular and Cellular Biology of the Yeast Saccharomyces*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 193−261.
Boeke,J.D., Lacroute,F. and Fink,G.R. (1984) *Mol. Gen. Genet.*, **197**, 345−346.
Borck,K., Beggs,J., Brammar,W., Hopkins,A. and Murray,N. (1976) *Mol. Gen. Genet.*, **146**, 199.
Braiterman,L., Monokian,G.M., Eichinger,D.J., Merbs,S.L., Gabriel,A. and Boeke,J.D. (1992) *Gene*, in press.
Brierley,C. and Flavell,A.J. (1990) *Nucleic Acids Res.*, **18**, 2947−2951.
Dinman,J.D. and Wickner,R.B. (1992) *J. Virol.*, **66**, 3669−3676.
Eichinger,D.J. and Boeke,J.D. (1988) *Cell*, **54**, 955−966.
Eichinger,D.J. and Boeke,J.D. (1990) *Genes Dev.*, **4**, 324−330.
Felsenstein,K.M. and Goff,S.P. (1988) *J. Virol.*, **62**, 2179−2182.
Gabriel,A. and Boeke,J.D. (1993) In Skalka,H. and Goff,S. (eds), *Reverse Transcriptase*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Garfinkel,J.D., Boeke,J.D. and Fink,G.R. (1985) *Cell*, **42**, 507−517.
Grimm,C., Kohli,J., Murray,J. and Maundrell,K. (1988) *Mol. Gen. Genet.*, **215**, 81−86.
Hansen,L.J. and Sandmeyer,S.B. (1990) *J. Virol.*, **64**, 2599−2607.
Hansen,L.J., Chalker,D.L., Orlinsky,K.J. and Sandmeyer,S.B. (1992) *J. Virol.*, **66**, 1414−1424.
Hatfield,D., Levin,J., Rein,A. and Oroszlan,S. (1992) *Adv. Virus Res.*, **41**, 193−239.
Hizi,A., Henderson,L.E., Copeland,T.D., Sowder,R.C., Hixson,C.V. and Oroszlan,S. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 7041−7045.
Hoffman,C.S. and Winston,F. (1989) *Gene*, **84**, 473−479.
Hutchison,C.A., Swanstrom,R. and Loeb,D.D. (1991) *Methods Enzymol.*, **202**, 356−390.
Jacks,T. and Varmus,H.E. (1985) *Science*, **230**, 1237−1242.
Jacks,T., Power,M.D., Masiarz,F.R., Luciw,P.A., Barr,P.J. and Varmus,H.E. (1988) *Nature*, **331**, 280−282.
Jamjoom,G.A., Naso,R.B. and Arlinghaus,R.B. (1977) *Virology*, **78**, 11−34.
Joyce,C.M. and Grindley,N. (1984) *J. Bacteriol.*, **158**, 636−643.
Koerner,T.J., Hill,J.E., Myers,A.M. and Tzagoloff,A. (1991) *Methods Enzymol.*, **194**, 477−490.
Kolodziej,P. and Young,R. (1991) *Methods Enzymol.*, **194**, 508−519.
Lee,C., Levin,A. and Branton,D. (1987) *Anal. Biochem.*, **166**, 308−312.
Lerch,R.A. and Friesen,P.D. (1992) *J. Virol.*, **66**, 1590−1601.
Levin,H.L. and Boeke,J.D. (1992) *EMBO J.*, **11**, 1145−1153.
Levin,H.L., Weaver,D.C. and Boeke,J.D. (1990) *Mol. Cell. Biol.*, **10**, 6791−6798.
Losson,R. and Lacroute,F. (1983) *Cell*, **32**, 371−377.
Marlor,R.L., Parkhurst,S.M. and Corces,V.G. (1986) *Mol. Cell. Biol.*, **6**, 1129−1134.
Maundrell,K. (1990) *J. Biol. Chem.*, **265**, 10857−10864.

Mead,D., Pey,N., Herrnstadt,C., Marcil,R. and Smith,L. (1991) *Bio/Technology*, **9**, 657–663.
Miller,K., Rosenbaum,J., Zbrzezna,V. and Pogo,A.O. (1989) *Nucleic Acids Res.*, **17**, 2134.
Park,J. and Morrow,C.D. (1991) *J. Virol.*, **65**, 5111–5117.
Rose,M.D., Winston,F. and Hieter,P. (1990) *Methods in Yeast Genetics: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Shiba,T. and Saigo,K. (1983) *Nature*, **302**, 119–124.
Springer,M.S., Davidson,E.H. and Britten,R.J. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8401–8404.
Towbin,H., Staehelin,T. and Gordon,J. (1979) *Proc. Natl Acad. Sci. USA*, **76**, 4350–4354.
Varmus,H. and Brown,P. (1989) In Berg,D. and Howe,M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC.
Weaver,D.C., Shpakovski,G.V., Caputo,E., Levin,H.L. and Boeke,J.D. (1993) *Gene*, in press.
Weaver,T.A., Talbot,K.J. and Panganiban,A.T. (1990) *J. Virol.*, **64**, 2642–2652.
Xu,H. and Boeke,J.D. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 8360–8364.
Yoshioka,K., Honma,H., Zushi,M., Knodo,S., Togashi,S., Miyaek,T. and Shiba,T. (1990) *EMBO J.*, **9**, 535–541.

Received on June 9, 1993; revised on August 23, 1993