Volume 14 Number 17 1986 Nucleic Acids Research

Expression strategies of the yeast retrotransposon Ty: a short sequence directs ribosomal frameshifting

Wilma Wilson, Michael H.Malim, Jane Mellor, Alan J.Kingsman* and Susan M.Kingsman

Department of Biochemistry, South Parks Road, Oxford OX1 3QU, UK

Received 8 July 1986; Accepted 11 August 1986

ABSTRACT

The Ty element of yeast is a member of a class of eukaryotic transposons which bear a striking resemblance to retroviral proviruses in their structure and expression strategies (1,2). A direct comparison can be drawn between the production of a fusion protein encoded by Ty, resulting from a frameshift event which fuses two out-of-phase open reading frames TYA and TYB, and the production of $Pr180^{gag-pol}$ in a retrovirus such as Rous Sarcoma Virus (RSV) (3,4). We present data which shows, definitively, that RNA splicing is not responsible for the frameshift in Ty. By in vitro mutation of a class I element, Tyl-15, we demonstrate that 31 nucleotides contained within the region where the TYA and TYB open reading frames overlap direct the frameshift. Within this short sequence there is a region of homology with a class II element which we show is also able to frameshift.

INTRODUCTION

Most Ty elements are about 5.9 kb long with long terminal repeats (LTRs) of about 340 base pairs (bp) called delta sequences. There are about 35 copies of the element in a haploid yeast genome and these fall into two classes, I and II, that show two large regions of non-homology while maintaining the same general organization (3, fig. 1). The major transcript from Ty elements is a 5.7 kb species that starts in the left or 5' delta and ends in the right or 3' delta such that the RNA has a 50 nucleotide terminal redundancy (4,5). The gross organization of the element is therefore similar to that of a retroviral provirus and to some other eukaryotic transposons such as the copia-like sequences of Drosophila (1,2). This similarity to retroviral proviruses has been extended recently by the observation that Ty transposes via an RNA intermediate (6) and that Ty proteins are assembled into virus-like particles, Ty-VLPs, that contain Ty RNA and reverse transcriptase (7,8).

In addition to these structural and functional similarities an unusual mode of gene expression that, in principle, is shared

by retroviruses has also been described (9,10). In a class I element such as Ty1-15, the transcriptional unit is divided into two open reading frames, TYA and TYB (9,10,11, fig. 1). TYA starts at nucleotide 299 and ends at 1,619 (10,11). TYB starts with an ACA (threonine) codon at nucleotide 1,581 and ends near the 3' delta at 5,676 (12,13). TYA and TYB therefore overlap by 38 bp and TYB is in the +1 reading frame with respect to TYA. TYA is expressed by the simple translation of the 5.7 kb transcript to produce a 50 kilodalton (kd) protein designated p1 (14). TYB, on the other hand, is expressed as a TYA:TYB fusion protein, p3, of about 190 kd (9,10,15). The production of p3 is achieved via a frameshifting event that leads to translational avoidance of the TYA termination codon and a shift into the TYB reading phase. This expression strategy is directly analogous to the production of $Pr180^{gag-pol}$ in a retrovirus such as RSV (2,16).

The mechanism of this frameshift is not known for Ty or retroviruses. It has recently been shown that RSV does not use RNA splicing to fuse gag and pol (17) and preliminary data for Ty is consistent with this (9,10). In this paper we address some simple questions that bear on the mechanism of this frameshift event. We show, in a carefully controlled S1 analysis, that RNA splicing is not involved, that a short sequence can mediate the shift and that this sequence has a region of homology with a class II Ty element that we show here, can also frameshift.

MATERIALS AND METHODS

Bacterial strains, yeast strains and media.

E.coli strain AKEC28 (C600 *thrC leuB6 thyA trpC1117 hsdRk hsdMk*) was used for plasmid manipulation and preparation. *Saccharomyces cerevisiae* strain MD40-4c (*ura2 trp1 leu2-3 leu2-112 his3-15*). *E.coli* media were prepared according to Miller (18). Yeast media were prepared according to Hawthorne and Mortimer (19)

DNA isolation

Plasmid DNA was isolated as described by Chinault and Carbon (20). Mini preparations of plasmid DNA were according to Holmes and Quigley (21).

Enzymes and fragment purification

Enzymes were purchased from BRL and used according to the manufacturers' instructions. DNA fragments were purified from agarose gels by the method of Tabak and Flavell (22).

DNA sequencing

The dideoxy-nucleotide chain termination method of Sanger et al was used (23).

Yeast transformation

The method of Hinnen et al was used (24).

Yeast extracts and protein blotting

Protein extracts of whole yeast cells were prepared as described previously (25). SDS-12.5% polyacrylamide gels (30:0.8 acrylamide/bisacrylamide) were run according to Laemmli (26). Protein blotting was carried out as described previously (10).

RNA isolation and blotting

Total RNA was isolated as described previously (14), transferred to nitrocellulose filters and hybridized (27). Probe fragments were labelled by nick translation (28).

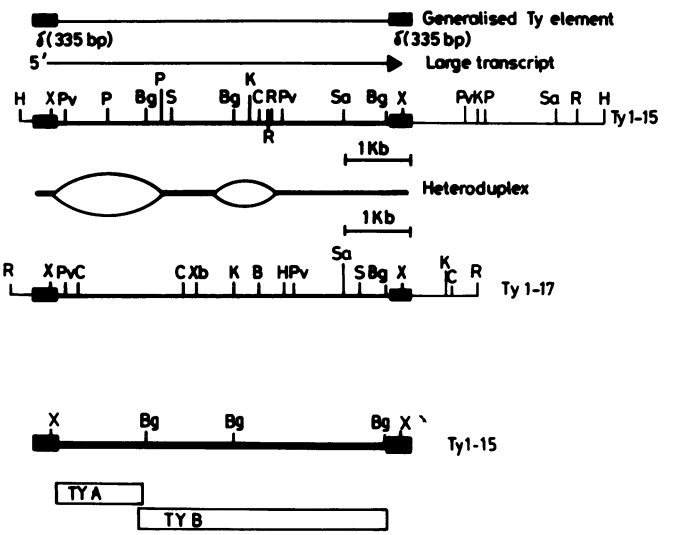
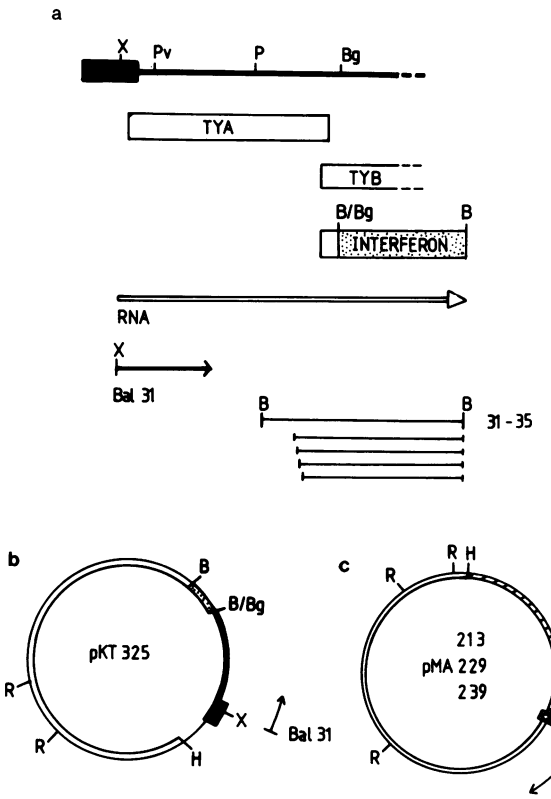


Figure 1. Restriction maps and gross sequence relationships of Ty1-15 and Ty1-17. The top diagram shows the general organization of a Ty element and its major transcript. A 9.6kb HindIII fragment containing Ty1-15 and a 7.2kb EcoRI fragment containing Ty1-17 are shown with a schematic diagram of the heteroduplex formed between the two elements (3). The lower diagram shows the organization of Ty1-15, showing the two open reading frames TYA and TYB. Solid boxes, delta sequences; Thick lines, internal region; Thin lines, Ty flanking sequences. H = HindIII; X = XhoI; Pv = PvuII; Bg = BglII; S = SalI; K = KpnI; C = ClaI; R = EcoRI; Sa = SacI; Xb = XbaI; B = BamHI.



d

```
                                                                  31
GAA CTG TTC TTA GAT ATC CAT GCT ATT TAT GAA GAA CAA CAG GGA

   1050
TCG AGA AAC AGT AAA CCT AAT TAC AGG AGA AAT CCG AGT GAT GAG

            1100
AAG AAT GAT TCT CGC AGC TAT ACG AAT ACA ACC AAA CCC AAA GTT

                    1150
ATA GCT CGG AAT CCT CAA AAA ACA AAT AAT TCG AAA TCG AAA ACA

                           1200
GCC AGG GCT CAC AAT GTA TCC ACA TCT AAT AAC TCT CCC AGC ACG

                                    1250
GAC AAC GAT TCC ATC AGT AAA TCA ACT ACT GAA CCG ATT CAA TTG

32        33                        34                35
                                          1300
AAC AAT AAG CAC GAC CTT CAT CTT AGG CCA GAA ACT TAC TGA.
 aca ata agc acg acc ttc atc tta ggc cag aaa ctt act gaa tct aca
```

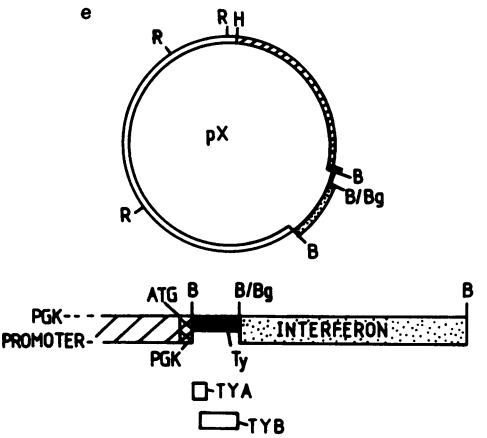


Figure 2a. Schematic representation of the deletion analysis strategy. Solid box, 5' delta sequences; Thick line, internal sequences. The two open reading frames TYA and TYB are shown. Fusion of an interferon coding sequence at the first BglII site of Tyl-15 in phase with TYB is illustrated, along with the resulting 2.3kb RNA species. The BamHI fragments, designated fragments 31-35, produced as a result of Bal 31 exonuclease digestion from the XhoI site are shown. **2b.** Map of plasmid pKT325. Open region E.coli/yeast shuttle vector sequences containing the yeast LEU2 gene and the 2 micron plasmid origin of replication (40); Thick line, Tyl-15 sequences; Thin line, Tyl-15 flanking sequences; Stippled box, alpha 2-interferon coding sequence; Arrow marks the direction of Bal 31 digestion. H = HindIII; Pv = PvuII; X = XhoI; Bg = BglII; S = SalI; B = BamHI; R = EcoRI; B/Bg = BamHI/BglII junction. **2c.** Map of the generalized structure of the PGK vectors pMA229, pMA239 and pMA213 (33). Their end-points are at nucleotide 14, 37 and 102 respectively (numbered from A, nucleotide 1, of the ATG at the start of the PGK coding sequence). Hatched region, PGK promoter; Criss-cross box, part of the PGK coding sequence; Arrow marks the direction of transcription. **2d.** Nucleotide sequence of the 3' end of TYA (11), showing the deletion end-points, numbered 31-35. The open reading frame corresponding to TYA is represented in triplets. The sequence starts at nucleotide 1,000 (numbered from A, nucleotide 1, of the ATG at the start of TYA) and ends at the TYA termination codon. The beginning of the overlapping TYB open reading frame is shown in lower case. **2e.** Map of the generalized structure of plasmids pMA229-31, pMA229-32, pMA213-33, pMA229-34 and pMA239-35 designated plasmid pX. These plasmids arise from the insertion of BamHI fragments 31-35, comprising a portion of TYA, the TYA:TYB overlap and the interferon cDNA, into the BamHI site of the appropriate PGK vector. The bottom diagram is a schematic representation of the region of fusion of PGK, Ty and interferon sequences in these constructions. Hatched region, PGK promoter; Criss-cross box, part of the PGK coding sequence; Thick line, Tyl-15 sequences; Stippled box, alpha 2-interferon coding sequence; Open boxes, part of the two open reading frames TYA and TYB.

S1 nuclease protection analysis

The RNA/DNA hybridization was performed in the presence of 80% formamide, 0.4M NaCl, 50mM PIPES pH6.4 and 1mM EDTA. The mix was heated at 82°C for 13 minutes, then hybridized at 52°C for 3 hours. The S1 treatment was carried out in the presence of 300mM NaCl, 30mM NaOAc pH4.6, 5mM $ZnSO_4$ and 20ug / ml salmon sperm DNA and the S1 enzyme concentration used was 500U/ml. After incubation at 30°C for 60 minutes, the reaction was stopped by the addition of EDTA. NaOH was added to a concentration of 100mM, and the mix was heated to 70°C for 5 minutes. After phenol extraction and ethanol precipitation, the DNA fragments protected from S1 nuclease digestion were separated on an alkaline agarose gel (29), transferred to nitrocellulose (30) and hybridized to nick translated interferon fragment 1.

Frameshift analysis

We have previously described an assay system for frameshifting in Ty expression (10), Briefly, this relies upon fusing an interferon alpha 2 cDNA in frame with TYB and detecting interferon-containing fusion proteins in a Western blot probed with interferon-specific monoclonal antibody. Such a protein can only be produced via a frameshift. In order to produce the appropriate reading frame in different TYB fusion constructions we use two different interferon cDNAs designated fragment 1 and fragment 2 that differ only in the relationship of their reading phases relative to their 5' Bam HI sites (31).

Plasmid constructions

(a) Defining the minimal requirement for TYA:TYB frameshifting in class I elements. We performed a deletion analysis of the TYA region as outlined in figure 2a. Plasmid pKT40b (32), which is a yeast/E.coli shuttle vector carrying a 9.6kb HindIII fragment containing Ty1-15 (fig. 1) was cleaved with BamHI and BglII and interferon fragment 1 was inserted in frame with TYB to produce plasmid pKT325 (fig. 2b). This was digested with Bal 31 exonuclease from the unique Xho I site to produce a collection of molecules with increasing deletions of TYA. The deleted molecules were ligated in the presence of BamHI linkers (GGATCCATGGG) and the BamHI fragments from 5 such plasmids, designated fragments 31 to 35, were inserted into the BamHI site of one of three PGK translational fusion vectors , pMA213, pMA229 or pMA239 (33, fig. 2c). The choice of vector depended on the end-point of the Bal31 generated deletion with respect to the reading frame of TYA (fig. 2d). The resulting plasmids, designated pMA229-31, pMA229-32,

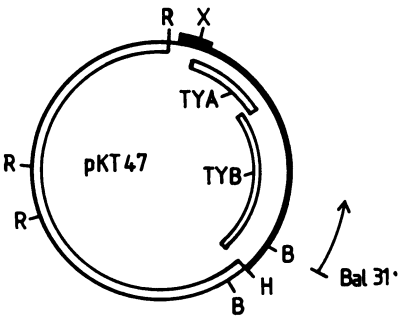


Figure 3. Map of plasmid pKT47 (11). Open region, E.coli/yeast shuttle vector sequences, containing the yeast LEU2 gene and the 2 micron plasmid origin of replication. Solid box, delta sequence; Thick line, Tyl-17 sequences; Thin line, Tyl-17 flanking sequences; Arrow marks the direction of Bal 31 digestion. The positions of TYA and TYB are shown. X = XhoI; R = EcoRI; B = BamHI; Bg = BglII; H = HindIII.

pMA213-33, pMA229-34 and pMA239-35 have the initiating ATG plus a few codons of PGK fused in frame to the residual portion of TYA, the TYA:TYB overlap region and an interferon cDNA in frame with TYB (fig. 2e). These plasmids were used to transform yeast strain MD40-4c to leucine independence.

(b) Analysis of frameshifting in the class II element, Tyl-17. To insert an interferon cDNA near the start of TYB in Tyl-17 a set of fragments was generated by Bal 31 deletion as there was no convenient restriction site. Plasmid pKT47 (11, fig. 3) which contains an EcoRI:HindIII fragment corresponding to the left 'half' of Tyl-17 (fig. 1) was digested with Bal 31 exonuclease from the unique Bam HI site and the deleted molecules were ligated in the presence of BglII linkers (CAAAAGATCTTTTG) and then screened for appropriate individual deletions. Two such molecules, plasmids pKT337 and pKT338 whose end-points are 79 and 75 nucleotides beyond the end of TYA (34), were used in this analysis. Interferon cDNA fragments were inserted into TYB in all three reading frames by fusing interferon fragment 2 into pKT337 and pKT338 and interferon fragment 1 into pKT338 to produce plasmids pKT337-2, pKT338-2 and pKT338-1. All three plasmids were used to transform yeast strain MD40-4c to leucine independence.

RESULTS

The experimental system

Our assay system for the Ty frameshift is based on fusing an

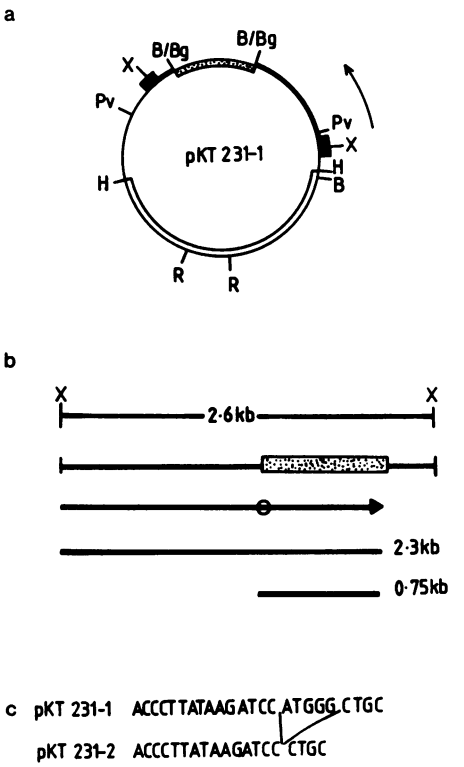


Figure 4a. Map of plasmid pKT231-1 (10). Open region, E.coli/yeast shuttle vector sequences, containing the yeast LEU2 gene and the 2 micron plasmid origin of replication. Solid boxes, delta sequences; Thick line, Tyl-15 DNA; Stippled box, interferon coding sequence; Arrow marks the direction of transcription. **4b.** Schematic representation of the 2.6kb XhoI fragment of pKT231-1. The full length transcript is shown and the region of mismatch between RNA from yeast transformant containing plasmid pKT231-2 and pKT231-1 DNA is circled. The protected fragments (2.3kb and 0.75kb) obtained in the S1 analysis are shown. **4c.** Comparison of the nucleotide sequence of plasmids pKT231-1 and pKT231-2 at the Ty:interferon junction highlighting the 5 bp deletion in pKT231-2 (10).

interferon cDNA in frame with TYB and asking if we can detect, in a Western blot, interferon-containing fusion proteins. Such proteins can only be produced via the frameshift. A yeast plasmid construction that achieves this is pKT231-1 (10, fig. 4a). This plasmid directs the synthesis of a 2.3 kb Ty:interferon fusion transcript which starts in the 5' Ty delta, runs through TYA, into TYB, through the interferon coding sequence and

terminates in the fortuitous transcription terminator in the 3' untranslated region of the cDNA. This transcript is translated via the frameshift to produce a 70 kd Ty:interferon fusion protein. (10, fig. 6b).

The TYA:TYB frameshift is not a result of RNA splicing

Previous reports which claimed that an RNA splicing event did not mediate the Ty frameshift did not provide a positive control for the detection of an intron of a few nucleotides (9,10) and therefore this possibilty could not be excluded. The possible presence of an intron in the hybrid transcript from pKT231-1 was tested in an S1 protection experiment using the 2.6 kb XhoI fragment from pKT231-1 as a probe (fig. 4b). S1 protected products were detected by Southern blotting using an interferon specific probe to avoid 'noise' from partially homologous RNAs encoded by the endogenous Ty elements. To control for detection of a very small intron, RNA from a yeast transformant containing a second plasmid, pKT231-2 (10), was used. pKT231-2 is the same as pKT231-1 except it has a 5 bp deletion at the Ty:interferon junction (fig. 4c). The RNA from this plasmid will therefore mimic a pKT231-1 transcript that has undergone the removal of a 5 nucleotide intron. Both the pKT231-1 and pKT231-2 transcripts are about 2.3 kb in length. This will be the size of the protected fragment that has homology with the interferon probe if there is no splicing. If, however, a splicing event is involved a smaller fragment will be produced that is slightly larger than the interferon cDNA as we would expect the putative intron to have its 3' end within the TYA:TYB overlap region. The pKT231-2 RNA will produce a 'mock' spliced product the size of the interferon component (0.75 kb). The data in figure 5 show that while the 5 bp mismatch between the pKT231-2 RNA and the 2.6 kb XhoI fragment from pKT231-1 can be readily detected there is no evidence for an intron in the pKT231-1 RNA. We can conclude from this that an RNA processing step is not involved in the fusion of TYA and TYB and therefore some other mechanism must be invoked.

Minimal sequence requirements for the frameshift

If RNA processing is not involved in the frameshift event it seems likely that the TYA:TYB fusion occurs during translation. In the most general terms this would require recognition of a sequence or structure within the RNA that peturbs normal translational fidelity. To identify such a sequence we have carried out a Bal 31 deletion analysis of the TYA region to

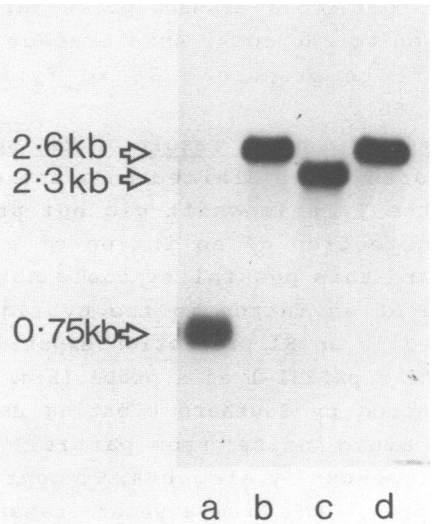


Figure 5. Southern blot of products from S1 nuclease protection analysis. Lanes a and b, pKT231-1 XhoI fragment hybridized to RNA from yeast transformant containing plasmid pKT231-2 : (a), treated with S1 nuclease; (b), in the absence of S1. Lanes c and d, pKT231-1 XhoI fragment hybridized to RNA from yeast transformant containing plasmid pKT231-1 : (c), treated with S1 nuclease; (d), in the absence of S1. All lanes were probed with interferon fragment 1.

define the requirements for the production of interferon via the frameshift (**Materials and Methods**, figs 2a and d). These deleted derivatives all contained the interferon cDNA in frame with TYB in the same configuration as in pKT231-1 and from each a cassette comprising a portion of TYA, the TYA:TYB overlap region and the interferon cDNA could be removed on a BamHI fragment. These BamHI cassettes, designated fragments 31-35, were then inserted into the expression site of one of three yeast PGK-based expression vectors, pMA213, pMA229 or pMA239, (fig. 2c) so that the ATG initiation codon and a few PGK codons were provided by the expression vector. The vector was chosen so that the small amount of PGK coding sequence in the vector and the remains of the TYA coding sequence after deletion were in phase (fig. 2d). The resulting plasmids, pMA229-31, pMA229-32, pMA213-33, pMA229-34 and pMA239-35 (fig. 2e) were used to transform yeast strain MD40-4c to leucine independence, producing transformants T229-31, T229-32, T213-33, T229-34 and T239-35. Extracts of these transformants were assayed for interferon by Western blotting.

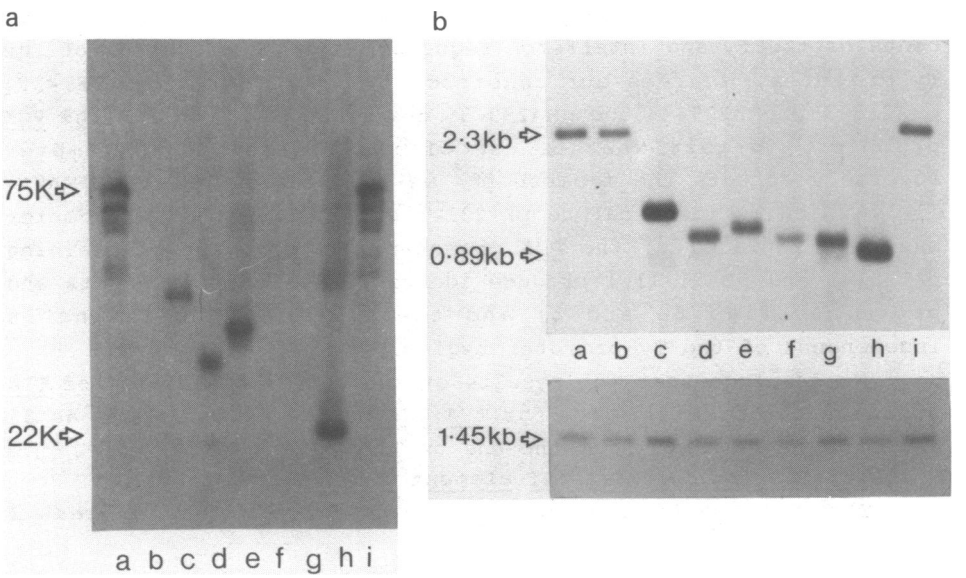


Figure 6a. Western blot analysis of interferon fusion proteins in extracts of yeast transformants T231-1 (lane a), T231-2 (b), T229-31 (c), T229-32 (d), T213-33 (e), T229-34 (f), T239-35 (g), T230-1 (h), T91-11-1 (i). All sizes are in kilodaltons. T231-1 is an MD40-4c tranformant containing plasmid pKT231-1 in which the frameshift event occurs , whereas T231-2 is an MD40-4c transformant containing plasmid pKT231-2 in which the frameshift event does not take place (10). T230-1 is an MD40-4c transformant containing plasmid pMA230-1 (31) which encodes a PGK/interferon fusion protein of M_r 22K. This was included as a positive control. T91-11-1 is an MD40-4c transformant containing plasmid pMA91-11-1 (see text). **6b.** Northern blot analysis of transcripts from yeast transformants T231-1 (lane a), T231-2 (b), T229-31 (c), T229-32 (d), T213-33 (e), T229-34 (f), T239-34 (g), T230-1 (h) and T91-11-1 (i) probed with interferon fragment 1 (upper panel). A 2.3kb transcript is obtained from T231-1, T231-2 (fig. 5) and T91-11-1. A 0.89kb transcript is obtained from T230-1 (31). Lanes a-i were re-probed with a 2.95kb HindIII PGK fragment from pMA91 (41, bottom panel). The 1.45kb transcript from the chromosomal PGK gene is present in all the RNA preparations and serves as a loading control.

Data in figure 6a shows that transformants T229-31, T229-32 and T213-33 produced interferon fusion proteins whereas T229-34 and T239-35 and did not. These differences are not due to differences in RNA levels as shown by the Northern blot of RNA from the same transformants (fig. 6b). These results show that a specific sequence in the TYA:TYB overlap region is necessary for frameshifting to occur. Our initial observation of the frameshift

event arose from the analysis of plasmid pKT231-1. This molecule contains the Ty and interferon sequences under the control of the Ty promoter, whereas our constructions pMA229-31, pMA229-32, pMA213-33, pMA229-34 and pMA239-35 described here contain the PGK promoter. To test whether use of the heterologous promoter adversely affects the frameshift, we have tested plasmid pMA91-11-1, which is identical to pKT231-1 except that the Ty promoter has been replaced by the PGK promoter. Transformants containing pKT231-1 and pMA91-11-1 produce identical fusion transcripts and proteins (figs 6a and b) and therefore frameshifting is independent of the Ty promoter region.

We conclude that the results of the deletion analysis of TYA localize the requirements for TYA:TYB fusion to as little as 31 nucleotides upstream from the end of TYA.

Frameshifting in the class II element Ty1-17

The overlap of TYA and TYB is 38 nucleotides in class I

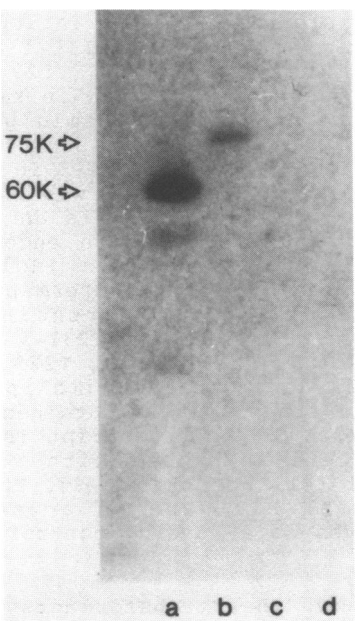


Figure 7. Western blot analysis of interferon fusion proteins in extracts of yeast transformants T361-1 (lane a), T337-2 (b), pKT338-1 (c) and T338-2 (d). All sizes are in kioldaltons. T361-1 is an MD40-4c transformant containing plasmid pMA361-1 (31) which encodes a PGK/interferon fusion protein of M_r 60K. An extract from this transformant was included as a positive control for the Western blot.

elements and 44 nucleotides in class II elements (11) and these two classes share little homology in this region. Furthermore TYB in class I elements starts with an ACA (threonine) codon whilst in class II it starts with a GCG (alanine) codon (11). As these regions showed such differences, we decided to determine whether a class II element, Ty1-17, fuses TYA and TYB via a frameshift. A similar strategy to that used for Ty1-15 was adopted and plasmids analogous to pKT231-1 were constructed. However, because Ty1-17 does not have a convenient BglII site near the start of TYB appropriate fragments were created by Bal 31 exonuclease digestion (see **Materials and Methods**). By inserting the appropriate interferon fragments into the deleted molecules, the interferon coding sequence was inserted into TYB in all three reading frames. Plasmids pKT338-2, pKT337-2 and pKT338-1 have interferon in phase with reading phase 1 (i.e. TYA), reading phase 2 (i.e. TYB) and reading phase 3 respectively. These plasmids were introduced into yeast strain MD40-4c to produce transformants T338-2, T337-2 and T338-1. Protein extracts of these transformants were analysed by Western blotting and the data in figure 7 shows that interferon is expressed only in T337-2 where interferon is in phase with TYB (reading frame 2). The protein produced has a relative molecular mass of about 75K. About 55K of Ty-encoded protein must be derived from TYA as there is no significant coding capacity in reading frames 2 or 3 to account for this increase (11). The remaining 20K comes from TYB and the fused interferon which must be brought into phase with TYA via a frameshifting event. This demonstrates that TYB in Ty1-17, as in Ty1-15 (9,10), is expressed as a TYA:TYB fusion protein.

DISCUSSION

We have sought to analyse some aspects of the frameshift event in Ty which leads to the TYA:TYB fusion. One obvious means by which this event could occur is RNA splicing, a mechanism which had been assumed for many years (35) to account for the corresponding gag:pol fusion in retroviruses. Our analysis presented here and the lack of appropriate yeast consensus RNA splicing signals (36,37) clearly demonstrate that RNA splicing does not mediate the TYA:TYB fusion. This conclusion is consistent with that drawn from recent studies on the gag:pol fusion in RSV (17) which showed that the same proteins as are produced *in vivo* are produced when an SP6-driven (38) RSV

transcript is translated in vitro. Furthermore, it is known that Moloney murine leukemia virus synthesises its gag:pol fusion protein as a result of occasional translational supression of the amber stop codon separating the gag and pol genes (39). Therefore RNA splicing is not involved in either of the retroviral gag:pol fusion systems analysed to date or in the fusion of TYA:TYB in yeast.

Since RNA processing does not account for the Ty frameshift it seems probable that the TYA:TYB fusion occurs via a ribosomal mechanism. This mechanism is unclear, but the analysis presented here defines the sequences required for the fusion. We have shown that the frameshift event still occurs when most of TYA is deleted and most of TYB is replaced with interferon. This observation suggests that the primary and secondary structure of the TYA:TYB RNA is not important and that the TYA gene product, p1, is not involved. Although the unlikely possibility of involvement of trans-acting p1 from endogenous Ty elements cannot be excluded. Again similar results have been obtained with RSV (17).

Our data shows that the information necessary to direct the ribosomal frameshift must reside in the 31 nucleotides at the end of the TYA:TYB overlap. In addition, we have shown that a class II Ty element is also able to fuse its TYA and TYB genes via a frameshift event. Interestingly, although these elements share little homology over the overlap region there is an 11 nucleotide block that is completely conserved between Ty1-15 and Ty1-17 and which is located within the 31 nucleotides that are required for

```
          1240                          1260                    gcg atg
Ty1-17    ATT AAT GAA TCA ACC GTT TCA TCA CAA TAC TTA AGC GAT

Ty1-15    ATC AGT AAA TCA ACT ACT GAA CCG ATT CAA TTG AAC AAT
                                                        aca ata

           1280                          1300
           aca acg aac tta gtc tta ggc cag caa cag aaa gaa tct aag
Ty1-17     GAC AAC GAA CTT AGT CTT AGC CCA GCA ACA GAA AGA ATC TAA.
                         * *** *** *** *
TY1-15     AAG CAC GAC CTT CAT CTT AGG CCA GAA ACT TAC TGA.
           agc acg acc ttc atc tta ggc cag aaa ctt act gaa tct aca
```

Figure 8. Nucleotide sequence comparison between the TYA:TYB overlap region of Ty1-15 and Ty1-17 (11). Upper sequence, Ty1-17; Lower sequence, Ty1-15. Both sequences start at nucleotide 1,237 (numbered from A, nucleotide 1, of the ATG at the start of TYA) and end at their TYA termination codons. The open reading frame corresponding to TYA is represented in triplets. The beginning of the overlapping TYB open reading frames is shown in lower case. The 11 nucleotide block of homology is marked by a row of '*'.

frameshifting (fig. 8). The significance of this homology block is not yet clear, but it is worth noting that in our construction pMA229-34 where 5 of these 11 conserved nucleotides have been deleted the frameshift event does not take place.

ACKNOWLEDGEMENTS

We would like to thank Frank Caddick for excellent artwork and photography. M.H.M. is supported by an SERC Research Studentship. J.M. is supported by an SERC/Celltech cooperative grant. Major support comes from an MRC project grant.

* To whom correspondence should be addressed.

REFERENCES

1. Finnegan,D.J. (1985) Int. Rev. Cytol. **93**, 281-326.
2. Varmus,H.E. (1983) in Shapiro, J.A. (Ed) Mobile Genetic Elements. Academic Press, N.Y., pp329-361.
3. Kingsman,A.J., Gimlich,R.L., Clarke,L., Chinault, A.C., and Carbon,J.A. (1981) J. Mol. Biol. **145**, 619-632.
4. Roeder,G.S. and Fink,G.R. (1983) in Shapiro,J.A. (Ed) Mobile Genetic Elements. Academic Press, N.Y., pp299-328.
5. Elder,R.T., Loh,E.Y. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA **80**, 2432-2436.
6. Boeke,J.D., Garfinkel,D.J., Styles,C.A. and Fink,G.R. (1985) Cell **40**, 491-500.
7. Garfinkel,D.J., Boeke,J.D. and Fink, G.R. (1985) Cell **42**, 507-517.
8. Mellor,J., Malim,M.H., Gull,K., Tuite,M.F., McCready,S. Dibbayawan,T., Kingsman,S.M. and Kingsman,A.J. (1985) Nature **318**, 583-586.
9. Clare,J. and Farabaugh,P. (1985) Proc. Natl. Acad. Sci. USA **82**, 2829-2833.
10. Mellor,J., Fulton,A.M., Dobson,M.J., Wilson,W., Kingsman, S.M., and Kingsman,A.J. (1985) Nature **313**, 243-246.
11. Fulton,A.M., Mellor,J., Dobson,M.J., Chester,J., Warmington,J.R., Indge,K.J., Oliver,S.G., de la Paz,P., Wilson,W., Kingsman,A.J. and Kingsman,S.M. (1985) Nucl. Acids Res. **13**, 4097-4112.
12. Hauber,J., Nelbock-Hockstetter,P. and Feldmann,H. (1985) Nucl. Acids Res. **13**, 2745-2758.
13. Fulton et al, unpublished data.
14. Dobson,M.J., Mellor,J., Fulton,A.M., Roberts,N.A., Bowen,B.A., Kingsman,S.M. and Kingsman,A.J. (1984) EMBO J. **3**, 1115-1119.
15. Mellor,J., Fulton,A.M., Dobson,M.J., Roberts,N.A., Wilson,W. Kingsman,A.J. and Kingsman,S.M. (1985) Nucl. Acids Res. **13**, 6249-6263.
16. Schwartz,D.E., Tizard,R. and Gilbert,W. (1983) Cell **32**, 853-869.
17. Jacks,T. and Varmus,H.E. (1985) Science **230**, 1237-1242.
18. Miller,J.H. (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, N.Y.

19. Hawthorne,D.C. and Mortimer,R.K. (1960) Genetics **45**, 1085-1110.
20. Chinault,A.C. and Carbon,J.A. (1979) Gene 5, 111-126.
21. Holmes,D.S. and Quigley,M. (1981) Analyt. Biochem. **114**, 193-197.
22. Tabak,H.F. and Flavell,R.A. (1978) Nucl. Acids Res. **5**, 2321-2332.
23. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467.
24. Hinnen,A., Hicks,J.B. and Fink,G.R. (1978) Proc. Natl. Acad. Sci. USA **75**, 1929-1933.
25. Mellor,J., Dobson,M.J., Roberts,N.A., Tuite,M.F., Emtage,J.S., White,S., Lowe,P.A., Patel,T., Kingsman,A.J. and Kingsman,S.M. (1983) Gene **24**, 1-14.
26. Laemmli,U.K. (1970) Nature **227**, 680-685.
27. Thomas,P.S. (1980) Proc. Natl. Sci. USA **77**, 5201-5204.
28. Rigby,P.W.J., Dieckman,M., Rhodes,C. and Berg,P. (1977) J.Mol. Biol. **113**, 327-251.
29. Maniatis,T., Fritsch,E.F. and Sambrook,T. (1982) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, N.Y., ppl71-172.
30. Southern,E.M. (1975) J. Mol. Biol. **98**, 503-517.
31. Tuite,M.F., Dobson,M.J., Roberts,N.A., King,R.M., Burke,D.C., Kingsman,S.M. and Kingsman,A.J. (1982) EMBO J. **1**, 603-608.
32. Bowen,B.A., Fulton,A.M., Tuite,M.F. Kingsman,S.M. and Kingsman,A.J. (1984) Nucl.Acids Res. **12**, 1627-1640.
33. Kingsman,S.M., Dobson,M.J., Tuite,M.F., Mellor,J., Roberts,N.A., and Kingsman,A.J. (1983) in Korhola,M. and Vaisanen,E. (Eds) Gene Expression in Yeast. Proceedings of the Alko Yeast Symposium Helsinki 1983. Foundation for Biotechnical and Industrial Fermentation Research **1**, 95-114.
34. Warmington,J.R., Waring,R.B., Newlon,C.S., Indge,K.J., and Oliver,S.G. (1985) Nucl. Acids Res. **13**, 6679-6693.
35. Weiss,R., Teich,H., Varmus,H. and Coffin,J. (Eds.) (1982) Molecular Biology of Tumor Viruses : RNA Tumor Viruses. Cold Spring Harbor Laboratory, N.Y., chap. 6.
36. Teem,J.L. and Rosbasch,M. (1983) Proc. Natl. Acad. Sci. USA **80**, 4403-4407.
37. Langford,C.J., Klinz,F.-J., Donath,C. and Gallwitz,D. (1984) Cell **36**, 645-653.
38. Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) Nucl. Acids Res. **12**, 7035-7056.
39. Yoshinaka,Y., Katoh,I., Copeland,T.D. and Orozlan,S. (1985) Proc. Natl. Acad. Sci. **82**, 1618.
40. Beggs,J.D. (1978) Nature **275**, 104-109.
41. Mellor,J., Dobson,M.J., Roberts,N.A., Kingsman,A.J. and Kingsman,S.M. (1985) Gene **33**, 215-226.