Volume 15 Number 15 1987 Nucleic Acids Research

A transcriptional activator is located in the coding region of the yeast *PGK* gene

J.Mellor, M.J.Dobson[1], A.J.Kingsman and S.M.Kingsman

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU and [1]Department of Botany, School of Biological Sciences, University of Nottingham, University Park, Nottingham N67 2RD, UK

Received May 5, 1987; Revised and Accepted July 2, 1987

ABSTRACT

Expression of heterologous genes from the PGK promoter on high copy number plasmids in yeast is relatively poor compared to the intact PGK gene because of low steady-state RNA levels. In this paper we show that low levels of heterologous RNA are not due to instability of mRNA but result from inefficient transcription due to a defect in RNA synthesis. A comparison of RNA levels from homologous and heterologous transcription units allowed the identification of a positive activator for transcription within the PGK coding region which is required for efficient expression of the PGK gene. Deletion of this region, the "downstream activator sequence", causes a six to ten fold reduction in transcriptional efficiency from the PGK 5' non-coding region.

INTRODUCTION

The yeast PGK gene encodes phosphoglycerate kinase, one of the efficiently expressed glycolytic enzymes. The PGK protein constitutes about 1% of total cell protein and its mRNA is correspondingly abundant (1). Expressing the PGK gene on a high copy number plasmid produces 50-80% of the total cell protein as PGK and there is a similar increase in the steady-state levels of the RNA (2). Increasing the number of copies of the gene therefore results in a corresponding increase in the amount of PGK RNA and protein. It is generally considered that the DNA sequences which are required for efficient transcription in eukaryotes lie upstream from the initiating ATG (3,4) and the upstream sequences from the PGK gene have been used to construct vectors for the expression of heterologous proteins. PGK based vectors constitute one of the most efficient expression systems for use in S. cerevisiae (2,5,6,7,8). However, the yields of heterologous products , typically around 1 to 2% of the total cell protein, are well below the theoretical maximum yield of 50 to 80% of the total cell protein, which is the yield of PGK

protein which is observed when the intact gene is placed in a high copy number plasmid (2). This can be explained in part by the relative instability of some heterologous products in yeast but the main reason is that the steady-state RNA levels transcribed from the heterologous template under the control of the PGK promoter are low (2). Low RNA levels are observed with a number of heterologous genes, both of yeast and mammalian origin (Mellor, Kingsman and Kingsman, unpublished data). A comparison of the steady-state RNA levels transcribed from the PGK promoter directing the expression of a number of foreign genes with the levels of PGK RNA from the intact gene shows that in general, heterologous RNA levels are five to ten fold lower than those from the intact PGK gene (2,6).

The low levels of heterologous RNA could be due either to instability of the mRNA or to inefficient transcription. In this paper we have distinguished between these two possibilities by comparing RNA stability and transcription initiation rates for homologous and heterologous transcription units. We show that it is a defect in the synthesis of heterologous mRNA which is responsible for the low steady-state mRNA levels and identify a positive activator for transcription within the PGK coding region.

MATERIALS AND METHODS

Bacterial and yeast strains and media: Plasmids were propagated in E.coli strain AKEC28 (C600, thrC, leuB6, thyA, trpC1117, hsdR, hsdM). Saccharomyces cerevisiae strain MD40/4c (x, ura2, trp1, leu2-3, leu2-112, his3-11, his3-15). E.coli cultures were grown in Luria broth (9) with ampicillin (Sigma) at 25ug/ml when necessary. Yeast cultures were grown at 30°C in YEPD (1% yeast extract, 2% peptone and 2% glucose) or in defined minimal medium (10) with 1% glucose as the carbon source. All cultures were harvested at 6 x 10^6 cells/ml for the preparation of RNA and DNA. DNA manipulations and plasmid constructions: Restriction enzymes, T_4 DNA ligase, Klenow fragment and nuclease Bal31 were purchased from BRL. Radioactively labelled isotopes ([α^{32}P]dTTP, [α^{32}P]UTP, [^{3}H] uracil and ^{32}P) were purchased from Amersham International. DNA manipulations and plasmid constructions followed standard procedures (11). Prior to ligation all fragments were isolated from agarose gels. Bal31 digestions followed standard conditions (12). The reactions products were filled in with Klenow and BamHI linkers (CCGGATCCGG, Celltech

Ltd.) were added. All end points were sequenced using the dideoxy method (13). Plasmids that have been described previously are pMA27 (2), pMA1 (6), pMA230-1 (2) pMA301-1 (2), pMA91-11 (7). Yeast transformants are designated T followed by the plasmid number; for instance MD40/4c transformed with pMA27 is designated T27.

Hybridisation probes: The following DNA fragments were isolated from agarose gels prior to nick translation using reagents supplied by BRL. A 2.95Kb (PGK) HindIII fragment was isolated from pMA1 (6), a 0.9Kb (IFN) BamHI fragment from pAT153-1 (6), a 1.25Kb (Ty) PvuII/BglII fragment from pMA91-11 (7), a 1.8Kb fragment (rDNA with homology to rRNA) and a 2.2Kb fragment (rDNA with homology to spacer DNA) by EcoRI digestion of pYRIG12 (14).

DNA blotting and filter binding assays: For preparation of nitrocellulose filters (see figures 2 and 4), pMA1 was restricted with HindIII, pAT153-1 with BamHI and pYRIG12 with EcoRI, and 1µg of the digested DNA was electrophoretically separated on multiple tracks of a 1% agarose gel. The gel was blotted onto 0.1µ nitrocellulose using standard procedures (15) and the filter was cut into individual sections prior to hybridisation with radiolabelled RNA. For filter binding assays, 0.1µg of isolated specific DNA fragments from pMA1, pAT153-1 and pMA91-11 were transfered onto small pieces of 0.1µ nitrocellulose (16).

Copy number estimation: Total yeast DNA was isolated from yeast using the method of Cryer et al. (17), restricted with EcoRI, electrophoretically separated on 1% agarose gels, stained with ethidium bromide and blotted onto 0.45µ nitocellulose (12,15). The filters were hybridised to isolated restriction fragments as described in the results and the copy numbers estimated by cutting out regions of the filter and scintillation counting the radioactivity associated with the plasmid and ribosomal bands.

RNA preparation: RNA was isolated from yeast and Northern blotted using the method decribed in Kim et al (12). Quantitation of the counts associated with each specific transcript was done by cutting regions from the nitrocellulose filters and scintillation counting (18).

RNA labelling experiments: (a). RNA turnover experiments: For preparation of uniformly ^{32}P labelled RNA, cultures of yeast were harvested at 2 x 10^4 cells/ml, washed twice in warm low phosphate minimal medium (prepared by adding 1ml of 1M $MgSO_4$ and 1ml of aqueous ammonia to 100ml of 0.67% yeast nitrogen base (Difco), precipitating the phosphate by stirring very gently at room

temperature for 30 mins, filtering through Whatman no. 1 paper, adusting to pH5.8 with HCl, autoclaving and adding sterile glucose to 1%), then resuspended in warm low phosphate medium with 50μCi/ml carrier free phosphate and allowed to grow for at least one generation time (3 hours). The cells were then filtered out of the medium using a Nalgene unit (0.45μ nitrocellulose), washed quickly with two lots of warm low phosphate medium containing 7mM phosphate, resuspended in this medium and grown shaking at 30°C for a further hour. Samples were taken at various intervals, frozen at -70°C, then RNA was prepared using standard procedures. The specific activity of this RNA was routinely 1 x 10^6cpm/μg. 2μg of total RNA were heat denatured at 70°C, added to hybridisation buffer (5X SSC, 50% formamide, 20mM phosphate pH6.5, 1X Denhardts, 0.1% SDS, 250μg/ml sonicated denatured salmon sperm DNA, and either 100μg/ml or 500μg/ml yeast tRNA (BRL)), and hybridised to Southern type filters, which had been prehybridised for 16 hours in the same mixture, at 42°C for at least 48 hours. Standard washing conditions (0.1X SSC, 0.1%SDS, 20mM phosphate pH6.5 at 42°C) were used. Addition of yeast tRNA to these hybridisations reduced the intensity of the signal to the 1.8Kb rDNA, as these preparations of tRNA are contaminated with significant amounts of yeast rDNA. All experiments were done in triplicate.

(b).Filter binding assays: Yeast cultures at 4 x 10^6 cells per ml were washed twice in warm synthetic complete medium minus uracil and grown for three hours in this medium supplemented with 1mCi/20 mls of tritiated uracil. The conditions for chasing out the label and preparation of RNA are as described above. 2μg of total yeast RNA were hybridised to nitrocellulose filters as described by Montgomary et al. (16). The counts associated with each filter were assessed by scintillation counting. All experiments were done in triplicate.

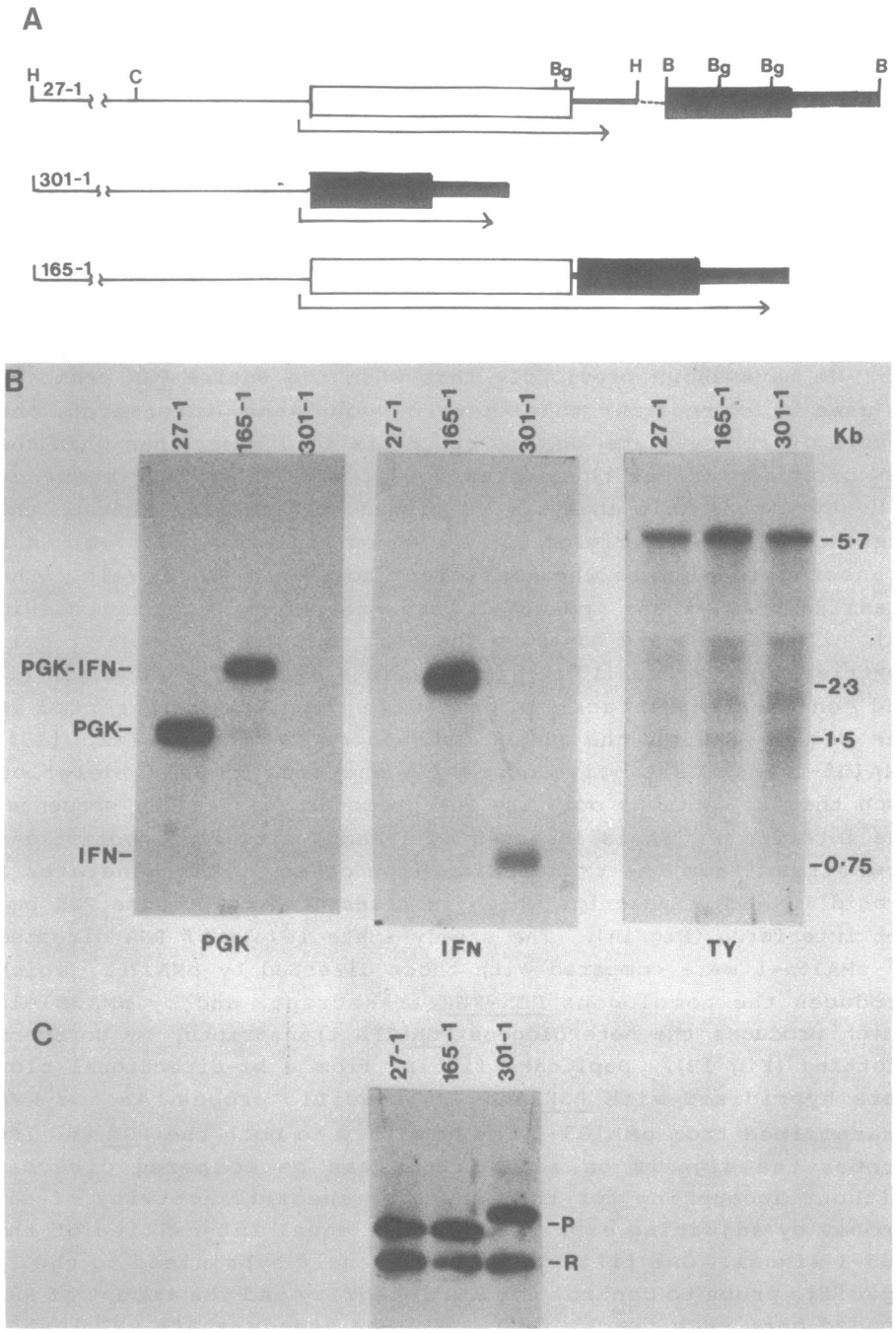
(c).Pulse-labelling RNA in whole yeast: 25ml of cells at 6 x 10^6/ml were harvested at 4°C, transferred to an eppendorf tube and washed twice, using a 30 second spin in a microfuge, in 1ml of TMN (10mM Tris/HCl pH7.4, 100mM NaCl, 5mM $MgCl_2$). The cells were resuspended in 0.95ml of ice cold water and 0.05ml of 10% N-Lauroyl sarcosine (Sodium salt, Sigma L-5125), then left on ice for 15 minutes. The cells were spun out in a microfuge (30 second spin) and resuspended in 120μl of reaction mix (50mM Tris/HCl pH7.9; 100mM KCl; 5mM $MgCl_2$; 1mM $MnCl_2$; 2mM DTT; 0.5mM ATP; 0.25mM GTP; 0.25mM CTP; 10mM phosphocreatine; 1.2μg/ul

creatine phosphokinase; 1uCi/ul $\alpha[^{32}P]$ UTP) and incubated at 25°C for 90 seconds. The cells were spun out of the reaction at 4°C, washed with cold TMN containing 50µM UTP, frozen at -70°C and RNA prepared using the standard procedure. The RNA was hybridised to DNA filters as described in the above section. The specific activity of the RNA was routinely about 2 to $8x10^4$ cpm/µg. All experiments were done in duplicate.

RESULTS

Steady-state mRNA levels of heterologous and homologous transcription units.

We have shown previously that when the entire PGK promoter is used to express our model heterologous gene, interferon, the levels of steady-state RNA are about six fold lower than when the PGK promoter drives the expression of the PGK coding sequence (2). To extend this analysis we wished to determine whether the presence of the interferon (IFN) cDNA on the plasmid or within a transcriptional unit was sufficient to affect RNA levels. The plasmids pMA27-1 and pMA165-1 were therefore constructed (Fig 1A). All plasmids are based on the high copy number yeast-E. coli shuttle vector pMA3a (19). pMA27-1 is a derivative of pMA27 (2) and contains the intact PGK gene with the IFN cDNA inserted at the BamHI site in the pBR322 sequences on the plasmid (19). pMA165-1 is a derivative of pMA27 generated by BAL31 deletion such that it contains only the PGK promoter and coding sequence. The interferon cDNA is inserted at a BamHI site four nucleotides downstream from the translation stop codon. This generates a hybrid PGK-IFN transcript which is translated to produce PGK but not interferon (Fig 1A). The steady-state levels of RNA directed by pMA165-1 were compared with those directed by pMA27-1, which produces the homologous PGK-PGK transcript, and by pMA301-1, which produces the heterologous PGK-IFN transcripts, by Northern blotting (Fig 1B). Duplicate filters from a bi-directional blot were hybridised with PGK and IFN specific probes. As the RNA transcribed from pMA165-1 has homology to both the PGK and IFN probes the signals on each filter can be compared directly without accounting for the length or specific activity of the probes by adjusting exposures to give equal intensities of the 165-1 signal. One filter was subsequently hybridised to the Ty specific probe to control for RNA integrity and the amount of RNA loaded onto each track of the gel. The steady-state RNA levels from T27-1 (see Materials and Methods) are identical to those


A
H
27-1
C
Bg
H
B
Bg
Bg
B
301-1
165-1
B
27-1
165-1
301-1
27-1
165-1
301-1
27-1
165-1
301-1
Kb
PGK-IFN-
PGK-
IFN-
-5·7
-2·3
-1·5
-0·75
PGK
IFN
TY
C
27-1
165-1
301-1
-P
-R

previously determined for T27, that is about 50 fold higher than those for the chromosomal PGK. This indicates that the presence of IFN cDNA on the plasmid does not affect RNA levels or plasmid copy number (Figs 1B and 1C). As expected, there is about a six fold difference in the levels of RNA from T27-1 and T301-1 (2). However, T165-1 RNA levels are high, about 40 fold higher than the chromosomal PGK, and only slightly lower than those from T27-1. The long 3' untranslated region of the hybrid RNA may account for the slight drop in the steady-state RNA levels. Nevertheless, this result suggests that the presence of IFN sequences within a PGK directed transcript does not significantly affect mRNA production.

Heterologous transcripts are stable in yeast.

In designing the approach to demonstrate the rate of degradation of the IFN and PGK transcripts in yeast it was decided that the rate of degradation of the IFN transcript should be compared initially with that of the chromosomal PGK transcript in the same transformant, T301-1. Subsequently, these rates were compared with the rate of degradation in T165-1, which produces high RNA levels. Cultures of T301-1 and T165-1 were labelled for one generation, about 2.75 hours, with ^{32}P, then the label was chased out with 7mM inorganic phosphate and samples of RNA were prepared at the times indicated. The RNA was hybridised to identical nitrocellulose filters containing DNA digests of plasmids pMA1, PAT153 and pYRG12 to produce PGK, IFN and rDNA specific fragments. Figure 2 shows two exposures of the same set of filters. On each filter the ribosomal RNA signal was used to control for the amount of RNA and the degree of hybridisation, as

Figure 1. **A.** Hybrid transcriptional units 1.5Kb PGK 5' noncoding region (thin line); PGK coding region (open box); PGK non-coding region (thick line); Interferon coding region (filled thick box); Interferon 3' non-coding region (filled medium box); pBR322 sequences (broken line). The transcript expected from each hybrid transcription unit is shown by the thin arrowed line below the gene. H=HindIII, B=BamHI, Bg=BglII, C=ClaI. The transcription units were inserted as HindIII/BamHI fragments into the high copy number yeast-*E. coli* shuttle vector pMA3a (19). **B.** Northern analysis. Each panel is the same filter hybridised with PGK (left), IFN (middle) and Ty (right) probes. The position of the PGK-IFN hybrid transcript (2.3 Kb), the normal PGK transcript (1.5 Kb), the 0.7 Kb IFN transcript and the 5.7 Kb Ty transcription are shown. **C.** Southern analysis. Total yeast DNA restricted with EcoRI and hybridised to the 2.2Kb rDNA (R) and pBR322 DNA (P).

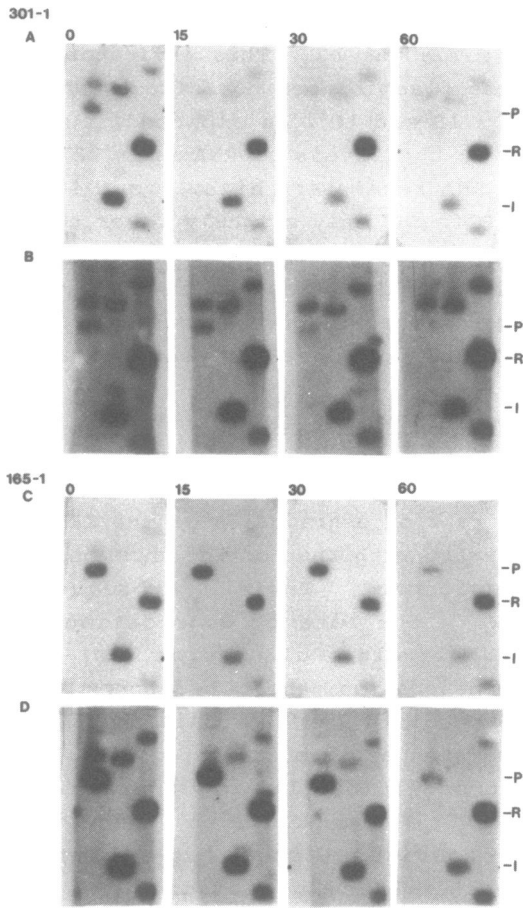


Figure 2. Rate of degradation of IFN, chromosomal PGK and PGK-IFN hybrid transcripts. Panels A and B show two exposures (16 hours and 72 hours) of the same set of filters hybridised to RNA from T301-1 and panels C and D show two exposures (8 hours and 48 hours) of the same set of filters hybridised to RNA from T165-1. 0, 15, 30 and 60 are the time in minutes from the start of the chase. The position of the PGK (P), IFN (I), and rDNA (R) specific signals are shown.

these rRNAs are degraded slowly in yeast. Therefore the PGK and IFN RNAs can both be compared with the rRNA throughout the experiment. This experiment is not designed to measure accurately the half-life of the RNAs but rather to compare their rates of degradation. The shorter exposure (Fig 2A) for T301-1 shows the IFN RNA turning over and in the longer exposure (Fig 2B) the

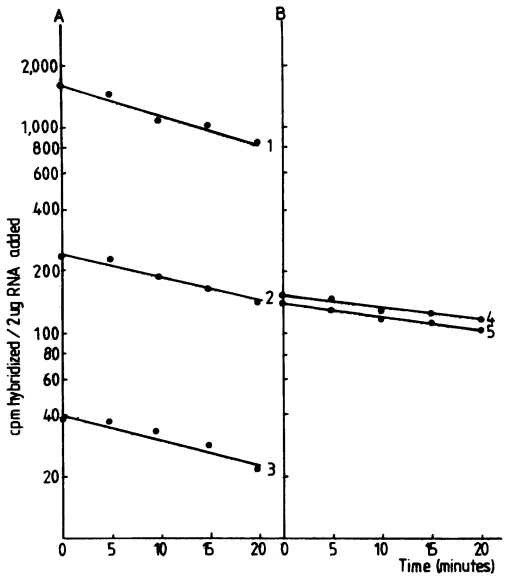


Figure 3. Filter binding assay. Semi-logarthmic plot of cpm hybridised against time after the removal of [^{3}H] uracil from medium. **A.** Line 1. PGK RNA from T27; 2. IFN RNA from T3Ø1-1; 3. PGK RNA from T3Ø1-1; **B.** 4. Ty RNA from T3Ø1-1; 5. Ty RNA from T27.

chromosomal PGK RNA, which is less abundant than the IFN RNA, can be seen. It is clear that both the IFN and PGK RNAs have similar rates of degradation in the same transformant. The bottom two panels (Fig 2C and D) show two exposures of autoradiographs for T165-1 RNA turnover. The rate of degradation is similar to both the IFN and chromosomal PGK RNA in the top panels. This result was confirmed in a filter binding assay (16). In this experiment, cultures of T3Ø1-1 and T27 were labelled to steady-state with (^{3}H)-uridine, instead of ^{32}P, and the label was chased out for the times indicated on the Figure. The overexpressed PGK RNA in T27 was observed to be degraded in a similar way to the chromosomal PGK RNA and the IFN RNA in T3Ø1-1 (Figure 3A). Figure 3B shows a control for this experiment. Here, the RNA from the three transformants was hybridised to a filter containing DNA from the yeast Ty element (7). This RNA is turning over at a similar rate in all three transformants.

These experiments indicate that the IFN transcripts are as stable as the authentic highly expressed PGK RNA, the chromosomal

PGK RNA, and the highly expressed PGK-IFN RNA from T165-1. The approximate values for the half life of the PGK RNA, about 30 minutes, agree with those determined independently (P.Piper, per. comm.) If the rates of degradation of these RNAs are the same but they achieve different steady state levels, then the rates of transcription of the two types of RNA must be different. The defect in heterologous RNA production is in the rate of transcription.

To test the hypothesis that a defect in the rate of initiation of IFN RNA synthesis from the PGK promoter results in low steady-state RNA levels, a run-on transcription or pulse labelling assay, originally developed by J. Warner (Albert Einstein College,N.Y.), was used. Two transformants T301-1, which produces low RNA levels and T27-1, with high RNA levels were chosen for this experiment. In order to look at the distribution of the radioactivity in the pulse-labelled total RNA population and to compare it with total RNA labelled to steady-state with inorganic phosphate, 1µg aliquots of RNA were separated by electrophoresis through formaldehyde/agarose and then blotted onto nitrocellulose. Autoradiography of the nitrocellulose showed label distributed throughout the pulse-labelled RNA population while in the steady-state labelled RNA the ribosomal 25S and 18S RNAs are the most abundant species (Fig 4a). The copy number and steady-state RNA levels were also determined in cells from the same culture and were as described previously (results not shown). The labelled RNA was then hybridised to nitrocellulose filters similar to those described in the RNA turnover experiments (Fig 2). Again, the amount of hybridisation of the 1.8Kb ribosomal RNA to the ribosomal DNA acts as an internal control so that RNA from different transformants can be directly compared. The RNA transcribed from the pBR322 and LEU2/2µ regions of the plasmids can also be detected and used as a second internal control. Figure 4b shows a comparison of representative filters hybridised to RNA pulse labelled for 90 seconds and RNA labelled to steady state with ^{32}P. Irrespective of the duration of labelling the signal from the PGK RNA on the T27 filter is about 6 fold more intense than the IFN RNA signal on the T301-1 filter. The IFN signal on the T301-1 filter is about seven fold more abundant than the chromosomal PGK signal in the same transformant. These differences reflect the steady-state difference in these RNAs previously observed by Northern blotting. Therefore, the rate

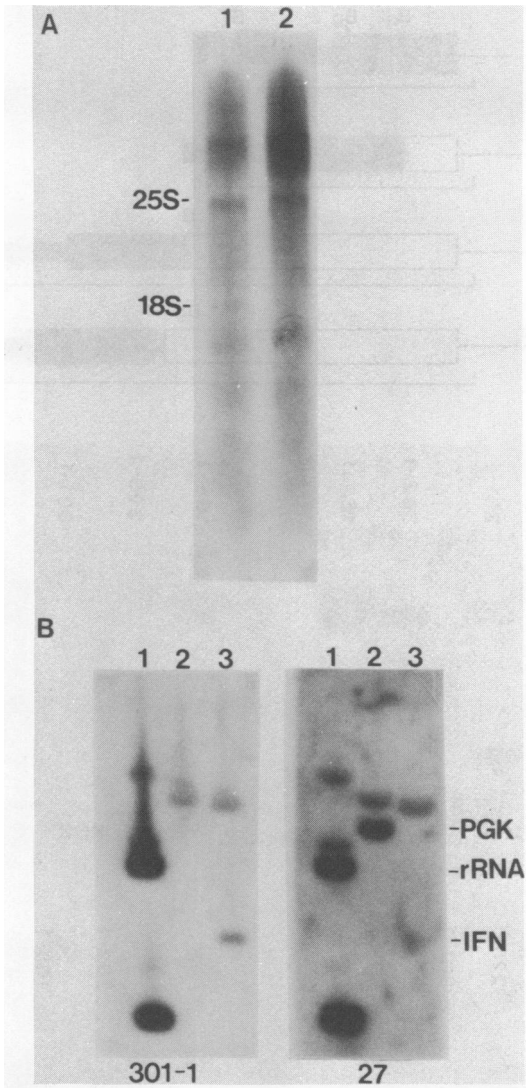


Figure 4. Pulse-labelling of yeast RNA. **A.** Autoradiograph of 1 μg of total RNA pulse labelled with ^{32}P for 90 seconds on a 1% agarose, 6% formaldehyde gel from T301-1 (track 1) and T27 (track 2). The position of migration of the 25S and 18S RNA's is shown. **B.** The panels show autoradiographs of nitrocellulose filters after hybridisation to labelled RNAs; left, RNA from 301-1, right RNA from T27. The positions at which specific transcripts hybridize to the PGK, IFN and rDNA DNA fragments are marked. Faint bands towards the top of the filters is RNA homologous to the LEU2, 2μ and pBR322 DNA. Track 1, pYIRG12 digested with EcoRI; 2, pMA1 digested with HindIII; 3, pMA153-1 digested with BamHI.

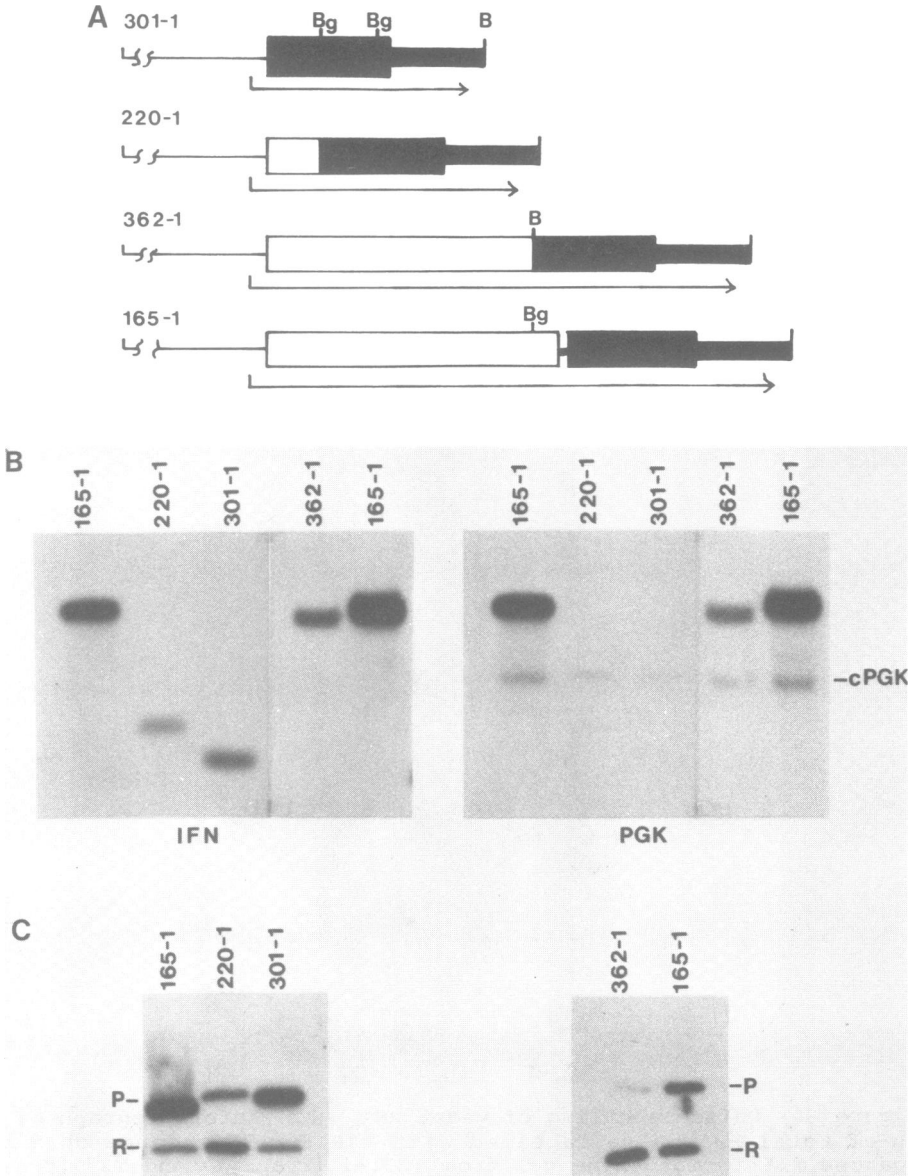


Figure 5. Localization of the Downstream Activator Sequence. **A.** Hybrid transcription units (see Figure 1A for symbols).
B. Autoradiographs of bidirectional Northern blots hybridised to IFN (left) and PGK (right) specific probes. **C.** Copy number determination. Autoradiographs of Southern blots of total yeast DNA hybridised to IFN (P)and 2.2Kb rDNA (R) specific probes.

synthesis of the PGK RNA in T27 is different to the rate of synthesis of IFN RNA in T301-1. As the time of the 90 second pulse-label in the "run-on" assay is short compared to the observed rate of degradation of the RNAs (30 minutes), these data show that IFN RNA production in T301-1 is limited at the point of rate of synthesis of RNA.

Evidence for an internal transcriptional activator region in the PGK gene.

A comparison of T165-1 and T301-1 shows that the former produces high levels of RNA while the latter does not. The only difference in these constructions is the presence of all 412 amino acids of the PGK coding region on the plasmid in T165-1 which expressed high levels of RNA. This suggests that the sequences in the coding region of the PGK gene are required for efficient expression. As we have previously shown, the plasmid pMA230-1 which contains 11 amino acids of the PGK N-terminus also shows low levels of expression (2). This locates the important signal between amino acid 11 and 412. To locate the cis-acting sequences more precisely the IFN gene was fused in frame with the PGK gene such that 79 and 310 amino acids of the PGK protein would preceed the IFN protein (pMA220-1 and pMA362-1, Fig 5A). The copy number (Fig 5C) and steady-state RNA levels (Fig 5B) of transformants T301-1, T220-1 and T362-1 were compared to T165-1. A bidirectional Northern blot analysis of the steady-state RNA from T301-1, T220-1 and T362-1 showed low RNA levels whereas in T165-1 the RNA levels were high (Fig 5B). However, the copy number of plasmids pMA220-1 and pMA362-1 were dramatically reduced compared to pMA301-1 and pMA165-1 (Fig 5C). The regions of the nitrocellulose filters from both the northern and Southern blots were cut out and counted so that an accurate quantitation of RNA levels and plasmid copy number could be made. The copy number of plasmids in T301-1 and T165-1 were normal, about 100 copies per cell, but the copy number of plasmids in T220-1 and T362-1 were repeatedly measured at about 5 copies per cell. This represents a twenty fold drop in template number, however, as the RNA is twice as abundant as the chromosomal PGK RNA, the steady-state RNA per plasmid copy must be high and equivalent to the level in T165-1. The reason for the drop in copy number during production of large PGK-IFN fusion proteins is not known.

It may be significant that these transformants showed a reduced growth rate (Mellor, unpublished data) suggesting that truncated PGK proteins may be toxic to the cells. There may, therefore,

be a selective advantage to cells which have a reduced gene dosage. When fully quantitated adjustments for copy numbers are made, it is clear that the fusion which contains 79 amino acids of PGK in pMA22Ø-1 is transcribed efficiently.

As our previous data showed that a fusion with only eleven amino acids (pMA23Ø-1) is poorly transcribed, this places the cis-acting sequences required for high effiency transcription from the PGK "promoter" between 37 bp and 236 bp downstream from the initiating ATG. We call this region the Downstream Activator Sequence (DAS).

DISCUSSION

In this paper we have investigated further the reason why heterologous gene expression from the PGK promoter in yeast is less efficient than homologous expression. We have shown that the low levels of heterologous RNA do not result from instability of foreign mRNAs in yeast but results from inefficient transcription from the PGK promoter when it is linked to a heterologous coding region. We have identified a region in the PGK coding sequence which is required for efficient transcription from the PGK promoter. Deletion of this region, which we call the downstream activator sequence (DAS) results in a five to ten fold reduction in the rate of synthesis of PGK RNA in hybrid transcriptional units or of any heterologous RNA fused into the PGK 5' and 3' non-coding regions in PGK based high copy number expression vectors. Thus, heterologous gene expression is limited because of low RNA levels.

The nature of the DAS and its relationship to other sequences responsible for the efficiency of the PGK "promoter" is at present uncertain. Previous deletion analysis has located an upstream activator sequence (UAS) in the PGK promoter (18). However, after deletion of the UAS and all sequences to within 91 bp of the ATG there is still some residual transcription (18). This implies that other sequences in this "promoterless" gene contribute to expression. One potential region is a CT rich sequence, known as the the CT..CAAG, which is conserved around the RNA initiation site (RIS) of a number of highly expressed yeast genes (19). The other possibilty is that the downstream activator sequence is a component of efficient expression. Thus both the upstream and downstream sequences may interact with the RIS to potentiate high level expression.

A precedent already exists for intragenically located transcriptional control regions in polII transcribed genes. The immunoglobulin heavy chain genes contain B-lymphocyte-specific enhancer sequences located within the second intron of the rearranged, activated genes (20, 2, 22). The genes of the human globin family have sequences located intragenically, in the second intervening sequence for α and β genes, which regulate the expression of the genes (23, 24, 25). In addition, the information sufficient to specify regulation of the chicken thymidine kinase gene during differentiation is entirely intragenic (26). Further evidence for internally located cis-acting sequences comes from studies using hybrid gene fusions expressed in transgenic mice. Both the human HPRT and the mouse growth hormone genes appear to contain intragenic sequences that determine tissue specific expression (27, 28). The intragenic location of these enhancing and regulating sequences may represent ecomony in gene organization. However, it is possible that the differential use of upstream and downstream cis-acting controlling elements may allow subtle changes in the pattern of gene expression. Transcriptional enhancers have been located in the protein coding regions of three oncogenic viral genomes. Early expression of the E1A proteins in adenovirus 5 is dependent on four regions with enhancer activity, three located in the 5' non-coding region and the fourth located about 400bp downstream from the transcription initiation site. These enhancers seem to be required only during early E1A expression (29). The hepatitis B and the bovine papilloma viruses also contain enhancers located in potential coding regions. Both these sequences were identified by their ability to activate a heterologous promoter in an orientation and position independent manner (30, 31, 32). Our data with PGK suggest that the coding region has evolved to fulfil requirements for both protein structure and function and to ensure efficient transcription. It is unlikely that this has been to achieve economy of expression as suggested for viruses. It is possible however, that genes such as PGK which play an essential housekeeping role in the cell have additional features to guarantee their expresssion.

ACKNOWLEDGEMENTS

This work was funded by a SERC-Celltech cooperative grant (SMK and AJK).

REFERENCES

1. Holland, J.P. and Holland, M.J. (1978) Biochem. **17**, 4900-4907.
2. Mellor, J., Dobson, M.J., Roberts, N.A., Kingsman, A.J. and Kingman, S.M. (1985) Gene **33**, 215-226.
3. Serfling, E., Jasin, M. and Schaffner, W. (1985) Trends in Genetics **1**, 224-230.
4. Guarente, L. (1984) Cell **36**, 799-800.
5. 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.
6. 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.
7. Dobson, M.J., Mellor, J., Fulton, A.M., Roberts, N.A., Kingsman, S.M. and Kingsman, A.J. (1984) EMBO J. **3**, 1115-1121.
8. Kingsman, S.M., Kingsman, A.J., Dobson, M.J., Mellor, J. and Roberts, N.A. (1985) In Biotechnology and Genetic Engineering Reviews **3**, 337-416.
9. Miller, J.H. (1972) Experiments in Molecular Genetics, published by Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
10. Hawthorne, D.C. and Mortimer, R.K. (1960) Genetics **45**, 1085-1110.
11. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press, N.Y.
12. Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Molec. Cell Biol. **6**, 4251-4358.
13. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463-5467.
14. Petes, T.D., Hereford, L.M. and Konstantin, K.G. (1978) J. Bacteriol. **134**, 295-305.
15. Southern, E.M. (1975) J. Mol. Biol. **98**, 503-517.
16. Montgomery, O.L., Leung, D.W., Smith, M., Shalit, P., Faye, G. and Hall, B.D. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 541-545.
17. Cryer, D.R., Eccelshall, R. and Marmur, J. (1975) Methods Cell. Biol. **12**, 39-44.
18. Ogden, J.E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Molec. Cell Biol. **6**, 4335-4343.
19. Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J., Kingsman, S.M., Perkins, R.E., Conroy, S.C., Dunbar, B. and Fothergill, L.A. (1982) Nucleic Acids Res. **10**, 2625-2637.
20. Sakonjo, S., Bogenhagen, D.F. and Brown, D.D. (1980) Cell **19**, 27-35.
21. Banerji, J., Olson, L. and Schaffner, W. (1983) Cell **33**, 729-740.
22. Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983) Cell **33**, 717-728.
23. Queen, C. and Baltimore, D. (1983) Cell **33**, 741-748.
24. Wright, S., Rosenthal, A., Flavell, R. and Grosfeld, F. (1984) Cell **38**, 265-273.
25. Charney, P., Treisman, R., Mellon, P., Chad, M., Axel, R. and Maniatis, T. (1984) Cell **38** 251-263.

26. Kosche, K.A., Dobkin, C. and Bank, A. (1985) Nucleic Acids Res. **13**, 7781-7793.
27. Merrill, G.F., Hauschka, S.D. and McKnight, S.L. (1984) Mol. Cell. Biol. **4**, 1777-1784.
28. Stout, J.T., Chen, H.Y., Brennand, J., Caskey, C.T. and Brinster, R.L. (1985) Nature **317**, 250-252.
29. Swanson, L.W., Simmons, D.M., Arriza, J., Hammer, R., Brinster, R., Rosenfeld, M.G. and Evans R.M. (1985) Nature **317**, 363-366.
30. Osbourne, T.F., Arvidson, D.N., Tyau, E.S., Dunsworth-Browne, M. and Berk, A.J. (1984) Mol. Cell. Biol. **4**, 1293-1305.
31. Lusky, M., Berg, L., Weiher, H. and Botchen, M. (1983) Mol. Cell. Biol. **3**, 1108-1122.
32. Shaul, Y., Rutter, W.J. and Laub, O. (1985) EMBO J. **4**, 427-430.
33. Tognoni, A., Cattaneo, R., Serfling, E. and Schaffner, W. (1985) Nucleic. Acids Res. **13**, 7457-7472.