Rapid Selection Using G418 of High Copy Number Transformants of *Pichia pastoris* for High-level Foreign Gene Expression

Carol A. Scorer, Jeffrey J. Clare, William R. McCombie[1,2], Michael A. Romanos* and Koti Sreekrishna[1,3]

Department of Cell Biology, Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK. [1]Phillips Research Center, Bartlesville, OK 74004. [2]Present address: Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11727. [3]Present address: Marion Merrell Dow Inc., 2110 E. Galbraith Rd., Cincinatti, OH 45215. *Corresponding author.

***Pichia pastoris* is a methylotrophic yeast increasingly important in the production of therapeutic proteins. Expression vectors are based on the methanol-inducible *AOX1* promoter and are integrated into the host chromosome. In most cases high copy number integration has been shown to be important for high-level expression. Since this occurs at low frequency during transformation, we previously used DNA dot blot screens to identify suitable clones. In this paper we report the use of vectors containing the *Tn903 kan*^{r} gene conferring G418-resistance. Initial experiments demonstrated that copy number showed a tight correlation with drug-resistance. Using a G418 growth inhibition screen, we readily isolated a series of transformants, containing progressively increasing numbers (1 to 12) of a vector expressing HIV-1 ENV, which we used to examine the relationship between copy number and foreign mRNA levels. Northern blot analysis indicated that ENV mRNA levels from a single-copy clone were nearly as high as *AOX1* mRNA, and increased progressively with increasing copy number so as to greatly exceed *AOX1* mRNA. We have also developed protocols for the *selection*, using G418, of high copy number transformants following spheroplast transformation or electroporation. We anticipate that these protocols will simplify the use of *Pichia* as a biotechnological tool.**

Received 17 August 1993; accepted 1 November 1993.

The yeast *Saccharomyces cerevisiae* has been used extensively as a host for the expression of foreign genes, combining advantages of microbial and eukaryotic organisms (for review see ref.1). However, frequent low yields and difficulties in secreting some proteins have led to the investigation of alternatives, such as the expression system based on *Pichia pastoris*. *P. pastoris* is a methylotrophic yeast initially selected for single-cell protein production because of its ability to grow to very high cell density in simple defined media. An expression system was developed based on the powerful methanol-inducible *AOX1* promoter, using vectors that can integrate at the *HIS4* or *AOX1* loci in chromosomal DNA[2]. Since a single copy of the *AOX1* gene is induced to yield alcohol oxidase up to 30% of cell protein it was believed that single-copy integration of the vector would be sufficient for high-level expression. Double homologous recombination to replace *AOX1* (transplacement) was the favoured method for generating stable single-copy transformants. However, single-copy integrants give disappointingly low yields in a number of examples[1]. In contrast, integration of multiple copies of the expression vector usually results in greatly increased yields of intracellular proteins (frequently > 10% of cell protein) and secreted proteins[3–5]. At the high cell densities (130 g/l dry mass) that can be achieved in fermentors these values translate to very high volumetric yields (eg. > 12 g/l for tetanus toxin fragment C[4]).

Multi-copy integration is expected to arise during integration by single homologous recombination, where it results in tandem integrated copies of the vector. However, we have also found multi-copy transformants to occur during transplacement in *P. pastoris*; transformants containing over 20 tandem repeats of the transplacing fragment have been isolated[4]. The mechanism of multi-copy transplacement probably involves an initial transplacement followed by repeated integration of circles formed by *in vivo* ligation[4]. Because the frequency of either type of multi-copy integrant can be low and unpredictable, we have used a DNA dot blot technique to screen for rare 'jackpot' transformants among several hundred grown in microtitre wells[6]. Though this method is rapid, it would be preferable to avoid the manipulations involved and directly *select* for multi-copy integrants. Here we investigate the dose-dependence of G418-resistance in *P. pastoris* using the *Tn903 kan*^{r} gene, and describe the use of a set of *P. pastoris* vectors that allow rapid G418-selection of rare high copy number transformants for high-level expression.

Results and Discussion

***P. pastoris* polylinker vectors with G418 selection.** Figure 1 illustrates a set of polylinker vectors for foreign protein expression and secretion in *P. pastoris*. pPIC3 is designed for intracellular expression, with foreign genes ligated at their 5′ ends to the unique BamHI (or NcoI site containing the initiating ATG). The secretion vector pPIC9 contains the *S. cerevisiae* α-factor leader with an engineered XhoI site adjacent to the DNA encoding the lys-arg cleavage site of the preproprotein; foreign genes may be ligated using a synthetic oligonucleotide to span the junction[5]. The vectors pPIC3K and pPIC9K are identical but contain in addition the *Tn903 kan*^{r} gene, so that transformants can be selected both by histidine prototrophy and G418-resistance. Like other *P. pastoris* vectors, these may be digested with SacI to promote integration 5′ to the chromosomal *AOXI* gene, or with BglII for transplacement[1].

Copy number dependence of G418 resistance. Our initial aim was to establish whether there was a relationship between the copy number of the integrated vector and the level of resistance to G418. Therefore we transformed the host strain, GS115, with SacI-digested vector (in this case an HIV-1 ENV vector, pPIC3K-ENV, see Experimental Protocol) and tested randomly-selected transformants for growth on G418. (GS115 itself was found to be sensitive to G418 (0.25 mg/ml) in YPD agar, though cells streaked at high density showed greater resistance).

Ninety-six transformants were tested for growth on YPD agar plates containing increasing concentrations of G418 (0.25 to 4.0 mg/ml). All the transformants grew on 0.25 mg/ml G418, but a decreasing number grew on the higher drug concentrations. Only a minority grew at the highest concentrations (eg. 15

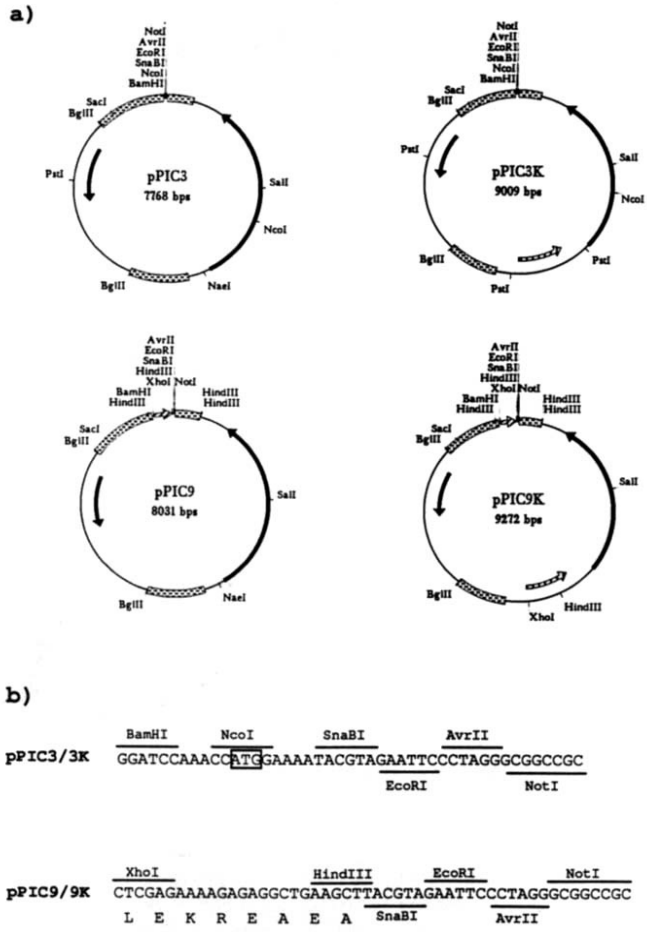

FIGURE 1. (a) Maps of *P. pastoris* polylinker expression vectors. Bacterial genes are indicted by arrows inside, yeast genes by arrows on the circle. The following genes are shown: ampicillin-resistance (black arrow), kanamycin-resistance (vertically-striped arrow), *P. pastoris AOX1* promoter, terminator and 3′ sequences (dashed boxes), HIS4 (hatched arrow), *S. cerevisiae* α-factor leader (diagonally-striped arrow). (b) Sequences of polylinkers in pPIC3 and pPIC3K, and pPIC9 and pPIC9K.

TABLE 1. Vector copy numbers for a selection of pPIC3K-ENV transformants of GS115 having different G418 minimum inhibitory concentrations.

Transformant	G418 concentration (mg/ml)[a]	Copy number
G2	≤0.25	1[b]
G12	≤0.25	0
A6	0.50	1[b]
C6	0.50	1
A1	0.75	2
A2	0.75	2
D1	1.00	2
D5	1.00	3
E8	1.50	4
G4	1.50	3
E5	1.75	5
F8	1.75	5
A4	2.00	7
E12	2.00	5
A8	>4.00	9
F11	>4.00	12
G9	>4.00	9
G10	>4.00	8
A12	>4.00	7

[a]The minimum G418 concentration required to completely inhibit growth.
[b]These transformants were shown by Southern blotting analysis to have, unexpectedly, expression cassettes integrated 3′ to *AOX1*.

at 1 mg/ml, 10 at 2 mg/ml), and these had the highest number of integrated copies, as visualized using the DNA dot blot technique (Fig. 2).

A number of the transformants were then selected for quantitative DNA dot blot analysis in order to determine the precise copy number. The data revealed an excellent correlation between minimum inhibitory G418 concentration and copy number, such that it was possible to predict the approximate copy number of a transformant from its level of resistance, up to a maximum of 7 copies (Table 1). This assay could therefore provide a rapid method for screening for high copy number transformants, or for generating a copy number series (1 to ≥7). Such a series could be useful for studies on gene dosage or in optimizing copy number for foreign protein secretion[7].

Selection for multi-copy transformants from spheroplast transformations. The G418 screening results suggested that it should be possible to directly *select* for high-copy transformants. The most efficient reported transformation method for *P. pastoris* is the spheroplast method[8], which also yields high copy number integrants while transformation of intact cells in the presence of lithium apparently does not (K.S., unpublished). However, primary selection of transformed spheroplasts with G418 is impractical due to the conflicting requirements to (a) regenerate the spheroplasts in agar and (b) allow a period of time for expression of drug-resistance before addition of G418. Therefore we transformed GS115 with pPIC3K-based plasmids cut with SacI or BglII, and isolated primary His+ transformants which were then pooled and subjected to *secondary* G418-selection on YPD agar plates containing different concentrations of G418 (0 to 2 mg/ml). Preliminary experiments had established that the level of resistance for any single transformant showed a dependence on plating density above 10^5 cells per plate; above this density even single-copy transformants showed resistance to high concentrations of G418. Therefore cells were always plated at a maximum of 10^5 per plate; this still allows selection from a large pool of transformants.

In this selection procedure, only a proportion of transformants were able to grow at 0.25 mg/ml G418. This number decreased with increasing G418 concentration. Starting with 1000 SacI primary transformants, 26 G418-selected colonies grew at 1.5 mg/ml, and 7 at 2 mg/ml in secondary selection; with 450 BglII transformants, 6 colonies grew at 1.5 mg/ml and none at 2 mg/ml. The vector copy numbers of the most resistant colonies were quantitated and found to be in the range 3 to 8 copies for SacI integrants and 7 to 18 copies for BglII integrants. Thus it appears that this method will allow the selection of high copy number transformants. However, in contrast to the growth inhibition screen (Table 1), no inferences can be made about the exact copy number from the G418 concentration used for selection.

We have observed that the relative frequency of multi-copy integration appears to be higher using BglII-digested rather than SacI-digested DNA. This may be because transformation with BglII-digested DNA occurs at lower frequency, indicating that perhaps double homologous recombination is a rate-limiting event, more likely to occur in the fraction of competent cells that have taken up several copies of the DNA.

Selection for multi-copy transformants following electroporation. Since spheroplast transformation is laborious and impractical for primary G418 selection, we examined the use of electroporation for transformation of *P. pastoris* and subsequent isolation of multi-copy integrants. Transformation frequencies were routinely 20-fold higher with SacI-cut than BglII-cut DNA, and 2 to 4-fold higher using the *aox1* strain KM71 rather than GS115. Therefore KM71 was electroporated with SacI-digested pPIC3K-derived plasmids (see Experimental Protocol) and plated on minimal agar plates for His+ selection, or on YPD

agar containing G418 (0.5 or 2 mg/ml) for primary drug selection, after an overnight expression period in YPD broth. Electroporation resulted in relatively high transformation frequencies (eg. 1000 His^+ colonies from 2.5 g of SacI-cut plasmid).

In the secondary selection experiment, pooled His^+ SacI transformants were plated at a density of 10^5 cells per plate on YPD agar containing 0 to 2 mg/ml G418, and transformants growing at the highest concentration picked for further analysis. In a typical experiment, the numbers of resistant colonies were: 100 at 1 mg/ml, 32 at 1.5 mg/ml, and 5 at 2.0 mg/ml. With primary G418-selection, colonies appeared on 0.5 mg/ml G418 after 2 days, and on 2 mg/ml after 3 days; all transformants selected in this way were also His^+.

Copy numbers of the most resistant transformants derived from both selection protocols were determined by quantitative DNA dot blot. With primary selection on G418, copy numbers were low (1 to 3 copies). This is probably due to the fact that plating cells directly from a transformation results in a much higher plating density than is optimal for G418 selection. However, secondary selection on G418 gave higher copy numbers (in the range 5 to 9 copies). Thus, this method appears to be almost as efficient as the spheroplast technique in yielding multi-copy SacI integrants. We have not tested BglII integrants due to the lower frequency of transformation. The unusually large amount of DNA that we used in these electroporations may be required for the generation of these transformants, since previously multi-copy transformants had not been detected from electroporation (K.S., unpublished).

Thus, although spheroplast transformation probably results in a higher frequency of multi-copy transformants, the electroporation method coupled with G418 selection greatly facilitates their isolation.

Correlation of mRNA level and copy number. The single-copy *P. pastoris AOXI* gene can be induced by methanol to yield alcohol oxidase at levels as high as 30% of cell protein. Therefore it was initially believed that the *AOXI* promoter was powerful enough so that single-copy vectors would give maximal foreign gene expression. However, subsequently it was found that multiple copies were required for high-level foreign gene expression. We were interested in finding out whether foreign mRNA levels increased with increasing copy number, thereby confirming that the copy number effect on protein yield is general rather than protein-specific. We investigated the relative concentrations of ENV mRNA in pPIC3K-ENV transformants with different vector copy numbers. The G418 copy number screen was particularly useful in rapidly identifying a series of transformants with increasing copy number from 1 to 12. In Figure 3A, ENV mRNA may be compared to *P. pastoris GAPDH* mRNA on the same Northern blot as an internal control, showing a gradual increase in foreign gene mRNA with increasing copy number (confirmed by quantitation of the bands, data not shown).

Since these transformants contained intact *AOXI* we were able to directly compare levels of the ENV mRNA to *AOXI* mRNA using a common hybridization probe from the *AOXI* 5′ and 3′ untranslated regions (Fig.3B). The results indicate that, at single-copy, ENV mRNA concentration is 2 to 3-fold lower than that of *AOXI*, while at 3 copies it exceeds it several-fold. Similar results are obtained for tetanus toxin fragment C mRNA (data not shown). In contrast to alcohol oxidase which reaches 5–30% of cell protein from single copy *AOXI*, ENV protein only reaches 2.5% even at 12 copies of the vector (data not shown). These results show that foreign gene expression is generally limited by vector copy number, and suggest that levels of alcohol oxidase in *P. pastoris* are not fully determined by *AOXI* promoter strength, but may be due to another factor, such as unusual protein stability. The high relative level of foreign mRNA to *AOXI* mRNA is in marked contrast to the situation in *S. cerevisiae*, where, at least in the case of the *PGK* promoter, inefficient foreign gene expression has been ascribed to greatly reduced steady-state levels of the foreign mRNA compared to *PGK* mRNA[9]. This may be explained by the presence of promoter elements within the *PGK* coding sequence[10].

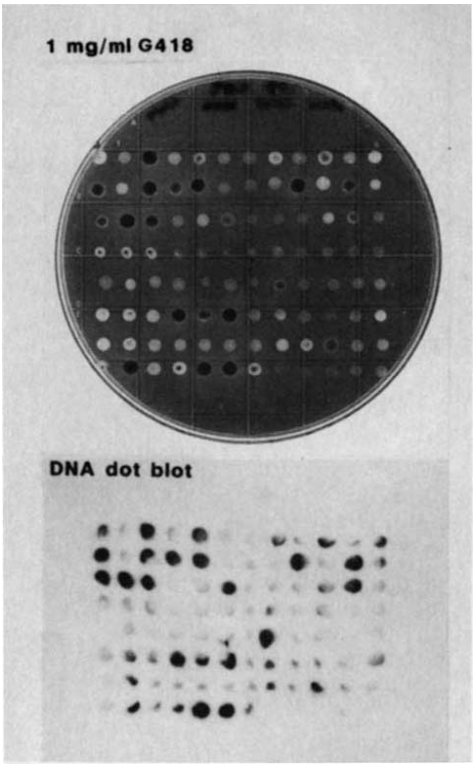

FIGURE 2. Growth of 96 randomly-selected transformants (GS115, SacI-cut pPIC3K-ENV) on YPD agar containing 1 mg/ml G418. The minority of transformants that grew in this screen (seen as darker spots due to illumination from below) showed the highest vector copy number in a DNA dot blot of lysed whole cells (bottom panel).

Our conclusion from the mRNA analysis is that the dependence of expression levels on vector copy number should be a general phenomenon, independent of the foreign gene, provided that other factors, eg. product toxicity, do not predominate. Western blot analysis (Fig. 4) indicates that ENV protein accumulation increased progressively up to 12 copies. A similar progressive increase in yield was seen with fragment C[4], and high copy number has been shown to improve yields in a number of examples[1].

Experimental Protocol

Construction of polylinker vectors. The sequence of the linker in the secretion vector pPIC9 is shown in Figure 1; pPIC9 was constructed by inserting an oligonucleotide between the XhoI and NotI sites of pPIC9-EGF[5]. Similarly, the intracellular expression vector pPIC3 (Fig. 1) was constructed by inserting an oligonucleotide between the BamHI and NotI sites of pPIC9-EGF. A HincII fragment containing the *Tn903 kan*^{r} gene from pUC4K was cloned into the unique NaeI site, resulting in pPIC9K and pPIC3K, respectively. The vector pPIC3K-ENV was constructed by inserting a partially-synthetic DNA encoding mature HIV-1 gp120 into pPIC3K[11].

Transformation of *P. pastoris*. The host strains used were GS115 (*his4*) or the derived *his4 aox1* strain KM71 (ref. 8). Transformation frequencies are 2 to 4-fold higher for KM71 by all methods used. Spheroplast transformations were carried out using 5 to 20 g SacI- or BglII-digested DNA[8]. Up to 3 μg of DNA was used for electroporation, which was carried out essentially by the method of Becker and Guarente[12], with cells pulsed in 0.2 cm sterile electroporation cuvettes at 1.5kV, 25μF, 400Ω, using a BioRad Gene Pulser with Pulse Controller. Immediately after pulsing, 1ml of cold 1M sorbitol was added to the cuvette. Cells were then either plated directly onto YNBD agar (His^+ selection) or selected directly with G418 (see below).

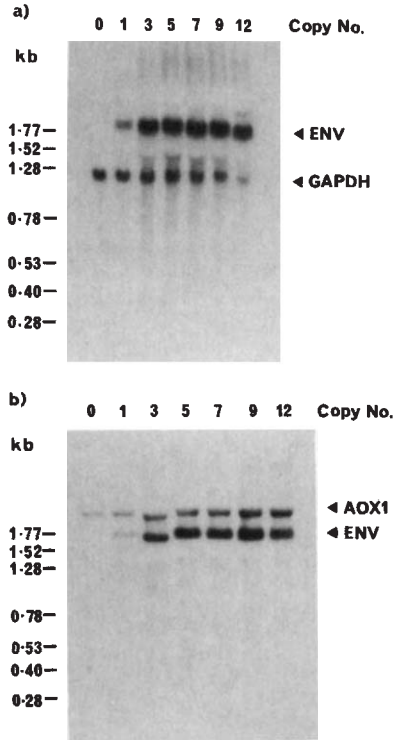

FIGURE 3. Northern blot analysis of mRNA from induced GS115 containing increasing copy numbers (0, 1, 3, 5, 7, 9, 12) of pPIC3K-ENV (a) probed with ENV DNA and *P. pastoris* GAPDH DNA to show ENV mRNA levels relative to *GAPDH* as an internal standard, (b) probed with DNA corresponding to the 5′ and 3′ untranslated regions of *AOX1*, in order to compare ENV to *AOX1* mRNA levels.

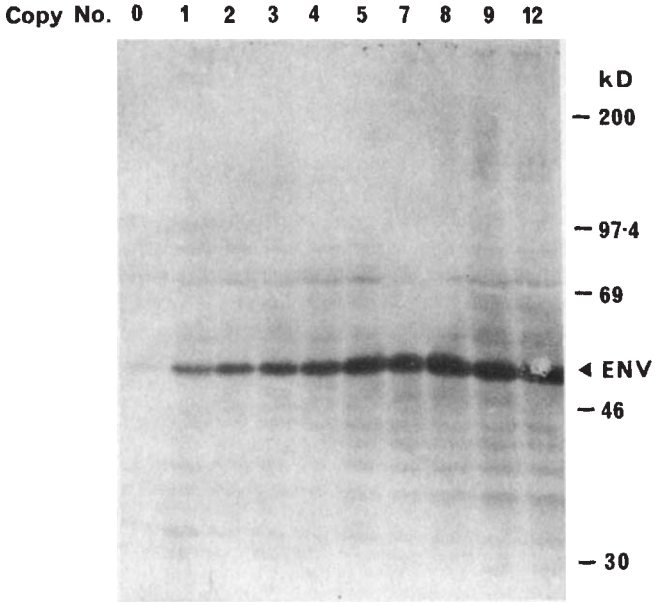

FIGURE 4. Western blot analysis of ENV protein from induced GS115 containing increasing copy numbers (0, 1, 2, 3, 4, 5, 6, 7, 9, 12) of pPIC3K-ENV.

Copy number screen using G418. His$^+$ colonies from a transformation were picked into a 96 well microtiter plate and grown statically overnight in YPD (200 μl per well) at 30°C. In order to ensure uniform growth, each culture was mixed by pipetting. Using a multi-channel pipette, 5 μl was sub-cultured into 195 μl fresh YPD in a microtitre plate, and grown overnight again. After mixing again, 1 μl of each culture was then spotted onto large (14 cm) petri dishes containing YPD agar with different amounts of G418 (0 to 4 mg/ml). After 2 days, growth on the different G418 concentrations was assessed and the minimum inhibitory concentration of G418, ie. lowest concentration which prevents growth, determined for each transformant.

Primary selection with G418 following electroporation. Electroporated cells were allowed to recover for 60 min in 0.5 ml 1M sorbitol after pulsing. YPD (0.5 ml) was then added and the cells were left at room temperature overnight in order to express G418-resistance. Cells were pelleted by centrifugation and resuspended in 200 μl distilled water, after which 100 μl was plated on 0.5 mg/ml G418 and 100 μl on 2 mg/ml G418. Resistant colonies appeared after 2 to 5 days and were checked for the His$^+$ phenotype since it is possible for spontaneous G418-resistant colonies to appear.

Secondary selection for multi-copy transformants using G418. After initial selection for His$^+$, transformants were pooled and plated on YPD agar containing G418. With spheroplast transformations, the top layer of soft agarose containing regenerated His$^+$ transformants was removed using a sterile spatula, vortexed vigorously in sterile distilled water, and the agarose allowed to settle. For electroporation, the transformants were suspended in sterile water using a spreader. Cell density was determined using a haemocytometer and cells were plated at 10^5 cells per standard 8.5 cm petri dish on YPD agar containing G418 (0 to 2 mg/ml). A range of concentrations of G418 was used (eg. 0.25, 0.5, 1 and 2 mg/ml), since the numbers of resistant colonies drop steeply above 0.5 to 1 mg/ml G418. Resistant transformants appeared from 2 to 5 days.

DNA and RNA analysis. Chromosomal DNA was analyzed by quantitative DNA dot blot in order to determine the copy number of the foreign gene[4]. Total RNA was prepared from transformed cells induced with 1% methanol for 8 hours and analyzed by Northern blot as described previously[13].

Protein analysis. Cultures were induced for 24 hours in shake flasks, and cell lysates prepared as described previously[4]. A 7.5% SDS-polyacrylamide gel was used to separate proteins (50 μg per lane). Proteins were then transferred to nitrocellulose and ENV protein was visualized using a sheep polyclonal antiserum to CHO cell-derived gp120 (ADP401) and a second anti-sheep antibody conjugated to alkaline phosphatase.

Acknowledgments

C.A.S. was supported during this work by a grant from the Medical Research Council AIDS Directed Programme.

References

1. Romanos, M. A., Scorer, C. A. and Clare, J. J. 1992. Foreign gene expression in yeast: a review. Yeast **8**:423–488.
2. Cregg, J. M., Tschopp, J. F., Stillman, C., Siegel, R., Akong, M., Craig, W. S., Buckholz, R. G., Madden, K. R., Kellaris, P. A., Davies, G. R., Smiley, B. L., Cruze, J., Torregrossa, R., Velicelebi, G. and Thill, G. P. 1987. High-level expression and efficient assembly of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris*. Bio/Technology **5**:479–485.
3. Sreekrishna, K., Nelles, L., Potenz, R., Cruze, J., Mazzaferro, P., Fish, W., Motohiro, F., Holden, K., Phelps, D., Wood, P. and Parker, K. 1989. High-level expression, purification, and characterisation of recombinant human tumour necrosis factor synthesised in the methylotrophic yeast *Pichia pastoris*. Biochemistry **28**:4117–4125.
4. Clare, J. J., Rayment, F. B., Ballantine, S. P., Sreekrishna, K. and Romanos, M. A. 1991. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. Bio/Technology **9**:455–460.
5. Clare, J. J., Romanos, M. A., Rayment, F. B., Rowedder, J. E., Smith, M. A., Payne, M. M., Sreekrishna, K. and Henwood, C. A. 1991. Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. Gene **105**:205–212.
6. Romanos, M. A., Clare, J. J., Beesley, K. M., Rayment, F. B., Ballantine, S. P., Makoff, A. J., Dougan, G., Fairweather, N. F. and Charles, I. G. 1991. Recombinant *Bordetella pertussis* pertactin (P69) from the yeast *Pichia pastoris*: high-level production and immunological properties. Vaccine **9**:901–906.
7. Thill, G. P., Davis, G. R., Stillman, C., Holtz, G., Brierley, R., Engel, M., Buckholtz, R., Kinney, J., Provow, S., Vedvick, T., and Seigel, R. S. 1990. Positive and negative effects of multi-copy integrated expression vectors on protein expression in *Pichia pastoris*, p. 477–490. *In:* Heslot, H., Davies, J., Florent, J., Bobichon, L., Durand, G. and Penasse, L. (Eds.). Proceedings of the 6th International Symposium on Genetics of Microorganisms, Vol II. Societe Francaise de Microbiologie, Paris.
8. Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. 1985. *Pichia pastoris* as a host system for transformations. Mol. Cell Biol. **5**:3376–3385.
9. Chen, C. Y., Oppermann, H. and Hitzeman, R. A. 1984. Homologous versus heterologous gene expression in the yeast, *Saccharomyces cerevisiae*. Nucl. Acids Res. **12**:8951–8970.
10. Mellor, J., Dobson, M., Kingsman, A. J. and Kingsman, S. M. 1987. A transcriptional activator is located in the coding region of the yeast *PGK* gene. Nucl. Acids Res. **15**:6243–6259.
11. Scorer, C. A., Buckholz, R. G., Clare, J. J. and Romanos, M. A. 1993. The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast Pichia pastoris. Gene. *In press.*
12. Becker, D. M. and Guarente, L. 1991. High-efficiency transformation of yeast by electroporation. Methods in Enzymol. **194**:182–187.
13. Romanos, M. A., Makoff, A. J., Fairweather, N. F., Beesley, K. M., Slater, D. E., Rayment, F. B., Payne, M. M. and Clare, J. J. 1991. Expression of tetanus toxin fragment C in yeast: gene synthesis is required to eliminate fortuitous polyadenylation sites in AT-rich DNA. Nucl. Acids Res. **19**:1461–1467.