Published in final edited form as:

Yeast. 2009 June ; 26(6): 349–353. doi:10.1002/yea.1667.

Sequence of the yeast protein expression plasmid pEG(KT)

B. Daniel Pierce and Beverly Wendland*

Department of Biology, Johns Hopkins University, 3400 N. Charles Street, Mudd Hall, Baltimore, MD 21218, USA

Abstract

The plasmid pEG(KT) is a widely used plasmid for expressing high levels of GST fusion proteins in the yeast *Saccharomyces cerevisiae*. Unfortunately, a complete sequence file has been lacking, thus complicating efforts to design cloning projects or to modify the plasmid for other uses (e.g. exchanging selection markers, epitope tags or protease cleavage sites to remove the epitope tag). Here, the complete sequence of the pEG(KT) plasmid is reported, thus facilitating its use. Additionally, its use as a vector backbone for high-level expression of a TAP-tagged protein is shown.

Keywords

pEG(KT); pEG(KG); *leu2-d*; protein expression plasmid; Pan1; yeast

Introduction

A variety of systems can be used to express and purify large amounts of proteins for biochemical and structural experiments, including bacteria, yeast and insect cells (Verma *et al.*, 1998). Often, the desired protein is initially generated as a chimeric protein, encoded as a translational fusion to an epitope that facilitates purification of the protein (Arnau *et al.*, 2006). Finally, it is generally useful to control the expression of the protein using an inducer, such as arabinose, IPTG or galactose (Stagoj *et al.*, 2006; Terpe, 2006). Using these various strategies, proteins of many types are expressed and purified for a variety of purposes.

The yeast plasmid pEG(KT) was first reported in 1993 as a tool to express proteins such as Ras2 that require eukaryotic post-translational modification for full functionality (Mitchell *et al.*, 1993). Many labs have since used this plasmid and its derivatives to produce large amounts of proteins for biochemical assays (e.g. Seaman *et al.*, 1998; Goode *et al.*, 1999; Duncan *et al.*, 2001; Zhu *et al.*, 2001). One shortcoming in further application of this very useful plasmid is the lack of a complete sequence file. We have sequenced the pEG(KT) plasmid to overcome this problem.

*Correspondence to: Beverly Wendland, Department of Biology, Johns Hopkins University, 3400 N. Charles Street, Mudd Hall, Baltimore, MD 21218, USA. bwendland@jhu.edu.

Materials and methods

Plasmid sequencing

pEG(KT) plasmid was purified using Qiagen Miniprep spin columns. Oligos were obtained from Operon (Huntsville, AL) and resuspended at 100 μ M in water. The plasmid DNA was mixed with the appropriate oligos (Table 1) and sent for commercial sequencing, using fluorescent di-deoxy Sanger sequencing. Reads were typically 600–800 bp. Sequences were used in BLAST searches to compare to known vector backbones, and chromatograms of the sequencing runs were inspected to resolve points of conflict. Gene Construction Kit (Textco BioSoftware, West Lebanon, NH, USA) was used to merge all of the sequences into one circular plasmid file.

Yeast expression of proteins from pEG(KT)-based plasmid

BJ2168 protease-deficient yeast cells with the genotype *MAT a leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2* (Aris and Blobel, 1989; Jones, 1991) were transformed with plasmids using the LiAc method (Schiestl and Gietz, 1989). For large scale purification, plasmid-containing cells should be grown to an absorbance of at least 2.0 at 600 nm (A_{600}) in synthetic medium supplemented with 8 mg/ml tryptophan and 2% w/v raffinose at 26 °C. Cultures were induced overnight (12–16 h) by the addition of 10 g/l yeast extract, 20 g/l Bactotryptone and 100 ml/l 20% w/v galactose at 26 °C. Harvested cells are then washed in water, resuspended in a small volume of water ($\sim 0.2 \times$ pellet volume) and poured in a thin stream into liquid nitrogen for quick freezing. The frozen pellets should be stored at -80 °C. Cell pellets are broken into a powder under liquid nitrogen in a metal Waring blender and the powder reconstituted in a $2 \times$ buffer to generate lysates for protein purification. For the samples used in Figure 2, a smaller-scale expression protocol was followed, which yields results that are qualitatively similar to the large-scale induction. For the small-scale expression approach, plasmid-containing protease-deficient yeast cells were grown to an absorbance of at least 1.0 at 600 nm (A_{600}) in synthetic medium supplemented with 8 mg/ml tryptophan and 2% raffinose at 30 °C, and induced for 4 h in 2% galactose at 30 °C. Whole-cell extracts were then prepared, using precipitation by 10% ice-cold trichloroacetic acid, washing twice in ice-cold acetone, drying, and then breaking the cells open by glass bead lysis in protein sample buffer.

Protein analysis

Cell extracts were heated to 70 °C for 5 min and then separated by SDS–PAGE on 7.5% acrylamide gels. The gels were then either stained with Gel-Code Blue (Pierce, Rockford, IL, USA) or transferred to nitrocellulose by electrophoresis, blocked with 5% non-fat dried milk in Tris buffer, and incubated with HRP-conjugated donkey anti-rat antibodies to recognize the IgG-binding component of the TAP tag. After washing, the nitrocellulose blot was developed using chemiluminescence reagents (Pierce, Rockford, IL, USA) and the image captured using an Alpha Innotech Fluorochem (San Leandro, CA, USA).

Results

Using 14 oligonucleotides (Table 1), the sequence of the pEG(KT) plasmid was obtained. All restriction mapping that is based on the sequence obtained matched perfectly to the bands observed when the actual plasmid was digested and analysed by agarose gel electrophoresis. A map of the pEG(KT) plasmid is shown in Figure 1, with unique restriction enzyme sites in bold and restriction enzymes with two sites in plain font. The sequence of pEG(KT) is available as a nucleotide file through GenBank (Accession No. FJ526990). In addition, we have generated a sequence file of pEG(KG) using the polylinker sequence reported in Mitchell (1993) (Accession No. FJ526991).

In an effort to express high levels of the endocytic scaffold protein Pan1 (Wendland and Emr, 1998), we used a series of restriction enzyme digests and ligations to make modifications to a pEG(KT)-based plasmid that expresses an N-terminally truncated GST–Pan1 protein (Duncan *et al.*, 2001; Miliaras *et al.*, 2004); the end result was a plasmid encoding full-length Pan1 with a C-terminal TAP tag in the pEG(KT) backbone. These modifications were done prior to the sequencing of the pEG(KT) plasmid, and relied primarily upon restriction mapping, along with what could be gleaned from the literature regarding the fragments used to construct the original pEG(KT) plasmid. To generate the Pan1–TAP-expressing [pEG(KT)] plasmid, the GST tag was excised while retaining the intact galactose promoter, a 5′ fragment that included the natural *PAN1* start codon and 5′ coding sequence was introduced, and a 3′ fragment encoding the C-terminal part of Pan1 fused to a C-terminal TAP tag was introduced (Toshima *et al.*, 2005). The latter fragment was excised from a more conventional high-copy yeast expression plasmid encoding galactose-regulated full-length Pan1 with a C-terminal TAP tag in the pRS426 backbone (Christianson *et al.*, 1992; Toshima *et al.*, 2005).

The Pan1–TAP [pEG(KT)] and Pan1–TAP [pRS426] plasmids were transformed into the protease-deficient yeast strain BJ2168, using the *URA3* selection marker of each plasmid. When cells were grown under inducing conditions in the presence of galactose as the carbon source, the Pan1–TAP protein was produced, while in the presence of the repressing carbon source dextrose, no or low levels of Pan1–TAP protein was translated (Figure 2). Higher levels of Pan1–TAP were synthesized from the pEG(KT)-based plasmid, as is evident from the Coomassie blue band visible near the 200 kDa marker in Figure 2A, and the signal from the Western blot detecting the TAP tag in Figure 2B. When the additional selection of the *leu2-d* allele, which has a truncated *LEU2* promoter (Baldari *et al.*, 1987), was applied by growing the cells in the absence of leucine, the highest levels of Pan1–TAP protein were produced (Figure 2). Under the conditions used for large-scale expression (see Materials and methods), we estimate that ~50–100 mg/l Pan1–TAP protein is expressed, based on Coomassie staining of whole cell extracts.

Now that the complete sequence of the pEG(KT) is available, future manipulations of the plasmid will be much easier to accomplish, including using such convenient approaches as cloning by homologous recombination (Ma *et al.*, 1987).

Discussion

The utility of the pEG(KT) plasmid is extremely high, due to the independent ways in which expression levels can be controlled. First, it is a so-called 2 μ plasmid, which imparts a high copy number in yeast cells (Rose and Broach, 1990). Second, its galactose promoter controls the expression of the desired protein product, which also contributes to high level of expression when the cells are grown in galactose-containing medium (Johnston and Davis, 1984). Third, it has the *leu2-d* allele of *LEU2*, which due to its very short promoter allows for a further selection of cells that have extremely high numbers of plasmids by growing the cells in medium lacking leucine (Erhart and Hollenberg, 1983).

The utility of the pEG(KT) plasmid is also limited, since it has *URA3* as the primary selection marker, GST as the epitope tag, the epitope tag located at the N-terminus of the protein, and the thrombin cleavage site for removing the epitope tag. For some circumstances, it might be desirable to use a selection marker other than *URA3*, an epitope tag other than GST, a placement of the epitope at the protein's C-terminus, or a different protease cleavage site, such as a TEV cleavage site (Parks *et al.*, 1994). Thus, one might want to retain the galactose-regulated expression and the *leu2-d* additional high-level expression selection, but alter other features of the plasmid.

Now that the sequence of pEG(KT) is available, many more investigators may find that protein expression in yeast is a convenient and powerful way to express high levels of proteins that might be toxic in other cells. For example, the central and C-terminal parts of the Pan1 protein are toxic in bacteria (Sachs and Deardorff, 1992; and our unpublished observations), but these same regions can be expressed in yeast cells (Duncan *et al.*, 2001; Miliaras *et al.*, 2004; Toshima *et al.*, 2005).

Acknowledgments

We thank Jiro Tosima for the Pan1-TAP (pRS426) vector and the TAP-tagged protein purification protocol. B.W. acknowledges funding from the NIH, Grant No. R01 GM60979.

References

- Aris JP, Blobel G. Yeast nuclear envelope proteins cross react with an antibody against mammalian pore complex proteins. *J Cell Biol.* 1989; 108:2059–2067. [PubMed: 2661560]
- Arnau J, Lauritzen C, Petersen GE, Pedersen J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr Purif.* 2006; 48:1–13. [PubMed: 16427311]
- Baldari C, Murray JA, Ghiara P, et al. A novel leader peptide which allows efficient secretion of a fragment of human interleukin 1 β in *Saccharomyces cerevisiae*. *EMBO J.* 1987; 6:229–234. [PubMed: 3034576]
- Christianson TW, Sikorski RS, Dante M, et al. Multifunctional yeast high-copy-number shuttle vectors. *Gene.* 1992; 110:119–122. [PubMed: 1544568]
- Duncan MC, Cope M, Goode BL, et al. Yeast Eps15-like endocytic protein, Pan1p, activates the Arp2/3 complex. *Nat Cell Biol.* 2001; 3:687–690. [PubMed: 11433303]
- Erhart E, Hollenberg CP. The presence of a defective *LEU2* gene on 2 μ DNA recombinant plasmids of *Saccharomyces cerevisiae* is responsible for curing and high copy number. *J Bacteriol.* 1983; 156:625–635. [PubMed: 6313610]

- Goode BL, Wong JJ, Butty AC, et al. Coronin promotes the rapid assembly and cross-linking of actin filaments and may link the actin and microtubule cytoskeletons in yeast. *J Cell Biol.* 1999; 144:83–98. [PubMed: 9885246]
- Johnston M, Davis RW. Sequences that regulate the divergent GAL1–GAL10 promoter in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 1984; 4:1440–1448. [PubMed: 6092912]
- Jones EW. Tackling the protease problem in *Saccharomyces cerevisiae*. *Methods Enzymol.* 1991; 194:428–453. [PubMed: 2005802]
- Ma H, Kunes S, Schatz PJ, Botstein D. Plasmid construction by homologous recombination in yeast. *Gene.* 1987; 58:201–216. [PubMed: 2828185]
- Miliaras NB, Park JH, Wendland B. The function of the endocytic scaffold protein Pan1p depends on multiple domains. *Traffic.* 2004; 5:963–978. [PubMed: 15522098]
- Mitchell DA, Marshall TK, Deschenes RJ. Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. *Yeast.* 1993; 9:715–722. [PubMed: 8368005]
- Parks TD, Leuther KK, Howard ED, et al. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal Biochem.* 1994; 216:413–417. [PubMed: 8179197]
- Rose AB, Broach JR. Propagation and expression of cloned genes in yeast: 2 μ m circle-based vectors. *Methods Enzymol.* 1990; 185:234–279. [PubMed: 2199781]
- Sachs AB, Deardorff JA. Translation initiation requires the PAB-dependent polyA ribonuclease in yeast. *Cell.* 1992; 70:961–973. [PubMed: 1339314]
- Schiestl RH, Gietz RD. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet.* 1989; 16:339–346. [PubMed: 2692852]
- Seaman MN, McCaffery JM, Emr SD. A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast. *J Cell Biol.* 1998; 142:665–681. [PubMed: 9700157]
- Stagoj MN, Comino A, Komel R. A novel GAL recombinant yeast strain for enhanced protein production. *Biomol Eng.* 2006; 23:195–199. [PubMed: 16707274]
- Terpe K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol.* 2006; 72:211–222. [PubMed: 16791589]
- Toshima J, Toshima JY, Martin AC, Drubin DG. Phospho-regulation of Arp2/3-dependent actin assembly during receptor-mediated endocytosis. *Nat Cell Biol.* 2005; 7:246–U244. [PubMed: 15711538]
- Verma R, Boleti E, George AJ. Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems. *J Immunol Methods.* 1998; 216:165–181. [PubMed: 9760222]
- Wendland B, Emr SD. Pan1p, yeast eps15, functions as a multivalent adaptor that coordinates protein-protein interactions essential for endocytosis. *J Cell Biol.* 1998; 141:71–84. [PubMed: 9531549]
- Zhu H, Bilgin M, Bangham R, et al. Global analysis of protein activities using proteome chips. *Science.* 2001; 293:2101–2105. [PubMed: 11474067]

Yeast. Author manuscript; available in PMC 2015 November 04.

Figure 2.

High levels of Pan1–TAP are produced when growing cells in the absence of uracil and leucine. Cells grown in raffinose to stationary phase were induced with 2% galactose for 4 h and extracts from 0.2 OD₆₀₀ were resolved in each lane by 7.5% SDS–PAGE. (A) Whole cell extracts were stained with Coomassie blue to show relative protein levels. Bio-Rad molecular weight markers were used (bands from the top down correspond to: 200, 116.25, 97.4, 66.2 and 45 kDa). The Pan1–TAP expression in [pEG(KT)] vs. (pRS.426) plasmids is compared; in both cases Pan1–TAP is under the control of the galactose promoter and is expressed in protease-deficient yeast. The additional absence of leucine from the medium (–URA, –LEU) selected for yeast cells with very high plasmid copy number, which leads to enhanced expression of the Pan1–TAP protein relative to cells grown with –URA selection. (B) An immunoblot of a gel similar to that shown in (A) was probed with HRP-conjugated donkey IgG, which binds to the TAP tag, and visualized by chemiluminescence

Table 1

Oligonucleotides used for sequencing (listed 5' to 3')

Oligo No.	Sequence
548	GGGCTGGCAAGCCACGTTTGGTG
749	CGGTGTGGTGGGCCAGG
750	CCTGGGCCCACCACACCG
753	GGCCGCTTTCATGGCCCTAC
1474	GTAAACGCGGGAAGTGGAGTCAGGC
1475	CGAGGACGCACGGAGGAGAGTC
1476	GGTCGTTCGGCTGCGGCGAGCGG
1477	GTGATTTAGGTGGTTCCAACAGTACCACCG
1478	CCATGAGTGATAAACTGCGGCC
1479	GGCCGCAGTGTATCACTCATGG
1480	GTCAGAGGTTTTACCGTCATCACC
1481	AGGTGGGTGGGTCTTAAGTAGG
1711	CCTGTAGCGGCGCATTAAAGCGC
1712	GCGGCATCAGAGCAGATTGTACTG