Appl Biochem Biotechnol (2007) 142:105–124
DOI 10.1007/s12010-007-0003-x

Expression of Recombinant Proteins in *Pichia Pastoris*

Pingzuo Li · Anukanth Anumanthan · Xiu-Gong Gao · Kuppusamy Ilangovan · Vincent V. Suzara · Nejat Düzgüneş · V. Renugopalakrishnan

Received: 14 April 2006 / Revised: 16 May 2006 / Accepted: 23 May 2006 /
Published online: 25 April 2007

Abstract *Pichia pastoris* has been used extensively and successfully to express recombinant proteins. In this review, we summarize the elements required for expressing heterologous proteins, and discuss various factors in applying this system for protein expression. These elements include vectors, host strains, heterologous gene integration into the genome, secretion factors, and the glycosylation profile. In particular, we discuss and evaluate the recent progress

P. Li
Shanghai Research Center of Biotechnology, Chinese Academy of Sciences, Shanghai 200233, People's Republic of China

P. Li · V. Renugopalakrishnan (✉)
Children's Hospital, Harvard Medical School, 300 Longwood Avenue, Boston, MA 02115, USA
e-mail: bionanotechnology@charter.net

A. Anumanthan
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA

X.-G. Gao
Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, 37 Convent Drive, 37/1B23, Bethesda, MD 20892, USA

K. Ilangovan
Centro de Investigacion en Calidad Ambiental, Tecnologico de Monterrey, Campus Estado de México, Monterrey, Mexico

V. V. Suzara · N. Düzgüneş
Department of Microbiology, Arthur A. Dugoni School of Dentistry, University of the Pacific, 2155 Webster Street, San Francisco, CA 94115, USA

V. Renugopalakrishnan
Bionanotechnology Group, Florida International University, Miami, FL 33174, USA

V. Renugopalakrishnan
Biophotovoltaic Group, Division of Bioengineering and Department of Mechanical Engineering, NUS Nanotechnology Initiative (NUSNNI), National University of Singapore, Singapore -117576, Singapore

in optimizing the fermentation process to improve the yield and stability of expressed proteins. Optimization can be achieved by controlling the medium composition, pH, temperature, and dissolved oxygen, as well as by methanol induction and feed mode.

Keywords *Pichia pastoris* · Protein expression · Methanol induction · Dissolved oxygen · Gene integration · Alcohol oxidase promoter · AOX1

Introduction

Recombinant protein production in the yeast strain *Pichia pastoris* has several advantages over other eukaryotic and prokaryotic expression systems: (1) rapid growth rate, coupled with ease of high cell-density fermentation; (2) high levels of productivity in an almost protein-free medium; (3) elimination of endotoxin and bacteriophage contamination; (4) ease of genetic manipulation of well-characterized yeast expression vectors; (5) absence of known human pathogenicity in the spectrum of lytic viruses that prey on *P. pastoris*; (6) diverse posttranslational modifications that include polypeptide folding, glycosylation, methylation, acylation, proteolytic adjustment, and targeting to subcellular compartments; and (7) the ability to engineer secreted proteins that can be purified from growth medium without harvesting the yeast cells themselves.

Most *P. pastoris* expression systems use the methanol-induced alcohol oxidase (AOX1) promoter [1]. Upon induction by methanol, the fraction of total soluble protein that is composed of alcohol oxidase can typically rise to 30% [2], indicating the power of this promoter element. AOX1 has been characterized and incorporated into a series of commercially available *P. pastoris* expression vectors, which require the following elements: (1) 5′-AOX1 (the alcohol oxidase promoter upstream of the gene of interest); (2) SIG (a secretion signal sequence); (3) MCS (a multiple cloning site with many, preferably unique, endonuclease sites); (4) TT (a transcription termination site); (5) HIS4 (a marker for selection by hydroxyhistidinase); (6) Amp^r (for selection with ampicillin); and (7) ColB1 (a replication element for plasmid propagation in *E. coli*) [3]. Candidate SIGs that have been used for successfully expressing recombinant proteins include PHO_5 (acid phosphatase), SUC2 (invertase), the 128-kDa pGKL killer protein, the leader sequence of the *Pichia acaciae* killer toxin, a phytohemagglutinin signal sequence (PHA-E) from *Phaseolus vulgaris*, and alpha-MF (a yeast mating factor). In general, the stability of secreted proteins recovered from a *P. pastoris* fermentation system can be improved upon by the addition of amino acid- or peptone-rich supplements, and by proper pH management of culture media. *P. pastoris* grows in a broad acidic pH range of 3.3 to 7.0. This property of the yeast is useful when adjustment of pH is necessary for minimizing the degradation of proteins secreted into the culture medium.

Protein purification from *P. pastoris* is also straightforward. The secreted and soluble proteins can be directly recovered by clarification of the *P. pastoris* culture media by centrifugation. Samples can be concentrated and purified by subjecting the supernatant to ultrafiltration, precipitation, and/or adsorption/elution chromatography. The yield of secreted protein can be increased dramatically during fermentation by utilizing multistage processes that sequentially scale up yeast from small “starter cultures” in a nonfermenting fashion, increase biomass by feed-batch fermentation and, finally, induce the gene of interest [4]. Typically, a three-stage system of this type for *P. pastoris* would be composed of the following:

Stage 1 Small batch culture of engineered yeast from archive in a nonfermentable carbon source such as glycerol; usable static volume 500 ml–1 l.

Stage 2 Increasing volume feed-batch culture where glycerol is added at growth-limiting amounts to optimize biomass production; 1–10 l in scale-up.

Stage 3 Actual induction of the gene of interest driven by the AOX1 promoter by addition of methanol to the fermentation culture; 10 l up to industrial production volumes, optimizing cell growth rates and metabolic rates *vs* secreted protein production [5].

Expression Vectors and Host Strains

Expression of a foreign gene in *P. pastoris* requires three basic steps: (a) the insertion of the gene into an expression vector; (b) integration of the expression vector into the *P. pastoris* genome; and (c) selection of potentially expressing strains for the foreign gene [6]. A variety of *P. pastoris* expression vectors and host strains are available. More detailed information on vectors and strains can be found in Higgins and Cregg [5]. In addition, the DNA sequences of many vectors are available at the Invitrogen website (www.invitrogen.com).

Expression Vectors

General Features of a Typical P. pastoris Expression Vector

One of the major features shared by all *P. pastoris* expression vectors is an expression cassette, which is composed of a promoter sequence (most often the *AOX1* promoter), a transcriptional termination sequence derived from *AOX1* that directs efficient 3′ processing and polyadenylation of the mRNAs, and, between them, single or multiple cloning sites for insertion of the foreign gene [7]. The design of the expression cassette ensures that the resulting transcription product (cap-*AOX1* 5′ UTR-ORF-3′ UTR-polyA) is a mature mRNA structure that is familiar to the yeast cellular machinery and does not contain cryptic sequences that may affect message stability or translational efficiency [1].

The insertion of a foreign coding sequence into the expression vector is usually carried out in *Escherichia coli*; therefore, all expression vectors of *P. pastoris* have been designed as *E. coli*/*P. pastoris* shuttle vectors. Besides the expression cassette, they also contain an origin of replication for plasmid maintenance in bacteria, along with selectable markers for transformation of the vector in both organisms. For secretion of foreign proteins, some expression vectors contain sequences encoding a secretion signal that are in frame with the foreign gene. These include the secretion signals of *P. pastoris* acid phosphatase (*PHO1*) and *S. cerevisiae* α-mating factor (α-MF). Figure 1 is the map of pPICZα vectors including the general features.

Some vectors also contain a fragment originating from *AOX1* 3′ sequences that can be used along with the *AOX1* promoter sequences to carry out *AOX1* gene replacement, as will be described in the "Integration Strategies" section.

Alternative Promoters for Expression in P. Pastoris

In circumstances where the *AOX1* promoter is not suitable, other promoters like *GAP*, *FLD1*, *PEX8*, and *YPT7* can be used. The *GAP* promoter is derived from the *P. pastoris* glyceraldehyde-3-phosphate dehydrogenase (GAP) gene [8]. The advantage of using this promoter is that methanol is not required for induction, nor is it necessary to shift cultures from one carbon source to another, making strain growth and protein expression more

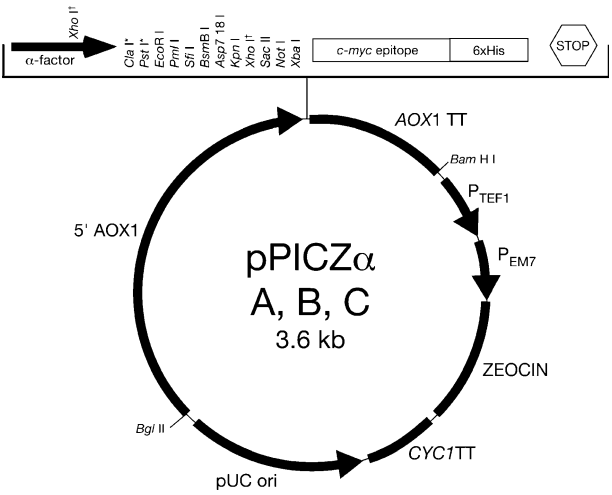

Fig. 1 Map of pPICZα. 5′ AOX1 promoter region: bases 1–941. 5′ AOX1 priming site: bases 855–875. α-factor signal sequence: bases 941–1207. α-factor priming site: bases 1144–1164. Multiple cloning site: bases 1208–1276. c-myc epitope: bases 1275–1304. Polyhistidine (6×His) tag: bases 1320–1337. 3′ AOX1 priming site: bases 1423–1443. AOX1 transcription termination region: bases 1341–1682. TEF1 promoter: bases 1683–2903. EM7 promoter: bases 2905–2162. Sh ble ORF: bases 2163–2537. CYC1 transcription termination region: bases 2538–2855. PUC origin: bases 2866–3539. * PstI is in Version B only. ClaI is in Version C only. + The two XhoI sites in the vector allow the user to clone their gene in frame with the Kex2 cleavage site, resulting in expression of their native gene without additional amino acids at the N-terminus. (Adapted from www.invitrogen.com)

convenient and straightforward. The *GAP* promoter is also convenient for expressing labeled proteins for NMR studies if culture is restricted to single carbon source. Because the *GAP* promoter is expressed constitutively, it is not a good choice for the production of proteins that may be toxic to the yeast.

The *FLD1* promoter is derived from the *P. pastoris FLD1* gene [9]. It is induced by either methanol as a sole carbon source (with ammonium sulfate as a nitrogen source), or methylamine as a sole nitrogen source (with glucose as a carbon source), but it is repressed in medium with glucose and ammonium sulfate. Thus, the *FLD1* promoter offers the flexibility of inducing high levels of protein expression with methylamine, which is an inexpensive nitrogen source.

The *PEX8* and *YPT1* promoters are useful when moderate expression levels are desirable [10]. The *PEX8* gene encodes a peroxisomal matrix protein that is essential for peroxisome biogenesis [11]. The level of expression from the *PEX8* promoter in methanol is significantly lower than that from the *FLD1* and *AOX1* promoters. The *YPT1* gene encodes a GTPase involved in secretion. The *YPT1* promoter provides a low but constitutive level of expression in media containing glucose, methanol, or mannitol as carbon sources [12].

Selectable Markers

Although classical and molecular genetic techniques are generally well developed for *P. pastoris*, few selectable marker genes have been described for the molecular genetic manipulation of the yeast. Existing markers are limited to the biosynthetic pathway genes *HIS4* from either *P. pastoris* or *Saccharomyces cerevisiae*: *ARG4* from *S. cerevisiae*; the

bacterial kanamycin-resistance gene (*kan*^{R}), which confers resistance to high levels of G418 [13]; and the *Sh ble* gene from *Streptoalloteichus hindustanus*, which confers resistance to the bleomycin-related drug zeocin [14, 15, 16]. The stable expression of human type III collagen illustrates the need for multiple selectable markers in *P. pastoris* [6].

Recently, a new set of biosynthetic markers has been isolated and characterized from *P. pastoris*: the *ADE1* (PR-amidoimidazole succinocarboxamide synthase), *ARG4* (argininosuccinate lyase), and *URA3* (orotidine 5′-phosphate decarboxylase) genes [17]. Each of these selectable markers has been incorporated into expression vectors.

Host Strains

All *P. pastoris* expression strains are derivatives of NRRL-Y 11430 (Northern Regional Research Laboratories, Peoria, IL). Most have a mutation in the histidinol dehydrogenase gene (*HIS4*) to allow for selection of expression vectors containing *HIS4* upon transformation [14]. Other biosynthetic gene/auxotrophic mutant host marker combinations are also available, but are used less frequently (cf. "Selectable Markers"). Before transformation, all these strains grow on complex media, but require supplementation with histidine (or other appropriate nutrient) for growth on minimal media.

Methanol Utilization Phenotype

Three types of *P. pastoris* host strains are available that vary with regard to their ability to utilize methanol. Most strains grow on methanol at the wild-type rate (Mut$^+$, methanol utilization plus phenotype). Two other types of host strains (Mut^s and Mut$^-$, methanol utilization slow and minus phenotype, respectively), however, have deletions in one or both *AOX* genes. Strains with deleted *AOX* genes sometimes are better producers of foreign proteins than wild-type strains [18, 19]. In addition, these strains require much less methanol to induce expression, which is an advantage in large-scale fermentation where large quantities of methanol are considered a significant fire hazard.

The most commonly used expression host of *P. pastoris* is GS115 (*his4*), which is wild type with regard to the *AOX1* and *AOX2* genes, and grows on methanol at the wild-type rate (Mut$^+$) [14]. KM71 (*his4 arg4 aox1*Δ: *SARG4*) is a strain in which the chromosomal *AOX1* gene is largely deleted and replaced with the *S. cerevisiae ARG4* gene [20]. As a result, this strain must rely on the much weaker *AOX2* gene for AOX and grows on methanol at a slow rate (Mut^s). The third host, MC100-3 (*his4 arg4 aox1*Δ: *SARG4 aox2*Δ: *Phis4*), is deleted of both *AOX* genes and is unable to grow on methanol (Mut$^-$) [15]. All of these strains, even the Mut$^-$ strain, retain the ability to induce expression at high levels from the *AOX1* promoter [21].

Protease-deficient Host Strains

Some secreted foreign proteins are unstable in the *P. pastoris* culture medium because of rapid degradation by proteases. Vacuolar proteases appear to be a significant factor in protein degradation, particularly in fermentor cultures, owing to the high-cell-density environment, in combination with the lysis of a small percentage of cells. Several protease-deficient strains, SMD1163 (*his4 pep4 prb1*), SMD1165 (*his4 prb1*), and SMD1168 (*his4 pep4*), have been shown to be effective in reducing degradation of some foreign proteins [10, 22]. The *pep4* gene encodes proteinase A, a vacuolar aspartyl protease required for the

activation of other vacuolar proteases, such as carboxypeptidase Y and proteinase B. Proteinase B has about half the activity before processing and activation by proteinase A. The *prb1* gene encodes proteinase B. Therefore, proteinase A and carboxypeptidase Y activities are substantially decreased or eliminated in *pep4* mutants, and proteinase B activity is decreased substantially. In the *prb1* mutant, only proteinase B activity is eliminated, whereas in *pep4 prb1* double mutants all these protease activities are reduced substantially or eliminated.

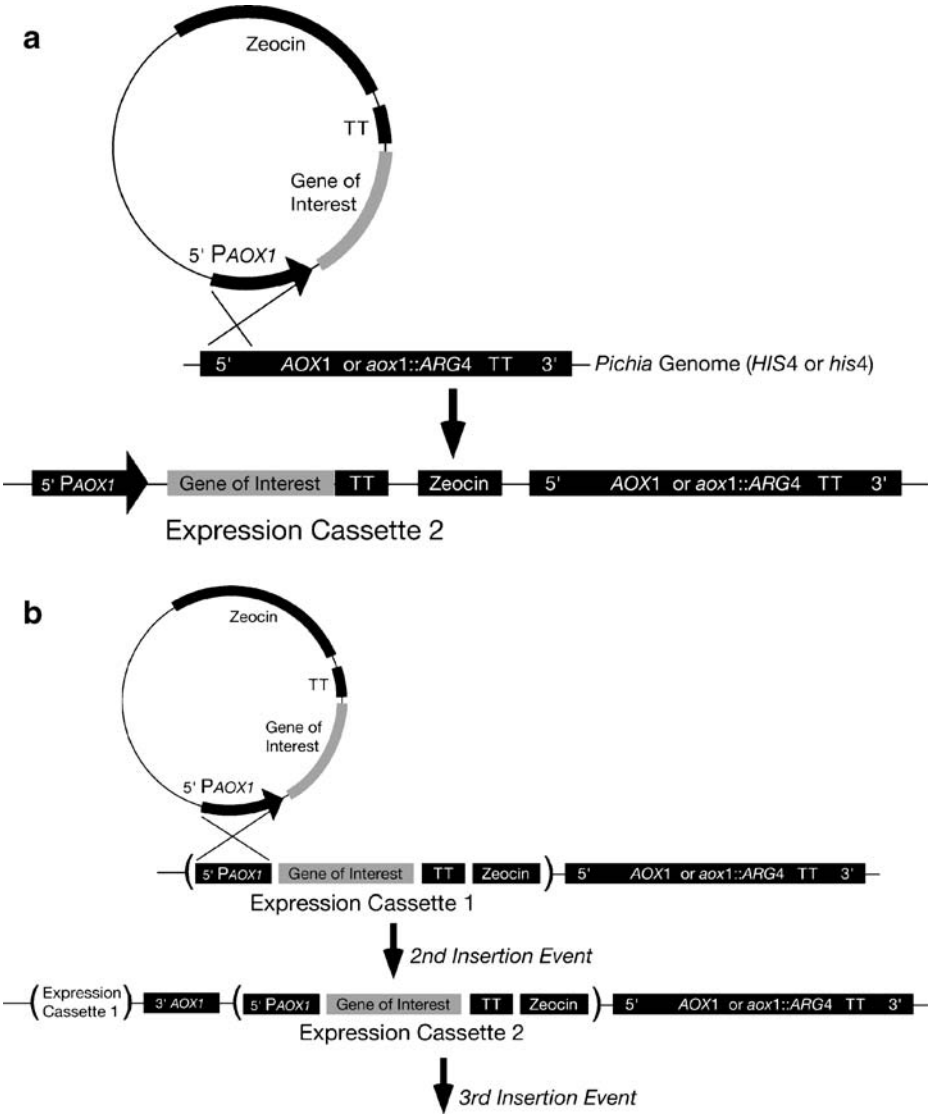

Fig. 2 Map of integration of a heterologous gene into the genome of *Pichia pastoris.* (A) Single copy integration; (B) Multiple copy integration (Adapted from www.invitrogen.com)

Integration, Transformation, and Selection

Integration of Expression Vectors into the *P. pastoris* Genome

Because no stable episomal vectors have been developed for *P. pastoris*, expression vectors are usually integrated into the *P. pastoris* genome (Fig. 2) to obtain stable expression strains, via either of the two methods described below. The integrants thus generated are stable in the absence of selective pressure, even when present as multiple copies.

Integration Strategies

As in *S. cerevisiae*, linear vector DNAs can generate stable transformants of *P. pastoris* via homologous recombination between sequences shared by the vector and host genome [5, 17]. The first, and simplest, way of integration is to digest the vector at a unique site with a restriction enzyme within either the marker gene (*HIS4*) or the *AOX1* promoter sequences, and then to transform the linearized vector into the appropriate auxotrophic mutant (*his4* mutant). The free DNA termini stimulate the vector to recombine homologously at the cut locus via a single crossover event with high frequency (50–80% of His^+ transformants). The remaining transformants have undergone gene conversion events in which only the marker gene from the vector has integrated into the mutant host locus without other vector sequences. Alternatively, certain *P. pastoris* expression vectors can be digested in such a way that the resulting DNA fragment containing the expression cassette and marker gene (*HIS4*) is flanked by 5′ and 3′ AOX1 terminal sequences that stimulate gene replacement events at *AOX1*, leading to *AOX1* gene deletion and replacement by the expression cassette and marker gene. The resulting replacement strains, which constitute approximately 10–20% of transformants, are prototrophic for histidine (His^+). In addition, because of the disruption of the *AOX1* gene, these strains rely on the transcriptionally weaker *AOX2* gene for growth on methanol [2] and, as a result, these strains have a Mut^s phenotype. These gene replacement strains are identified easily among transformed colonies by replica-plating in the presence of methanol and selecting those colonies with a reduced ability to grow on methanol. The graphic integration for pPICZ vectors inserted with the objective gene is shown in Fig. 2.

With either single-crossover or gene replacement integration strategies and selection for His^+ transformants, a significant percentage of transformants will not contain the expression vector. Thus, transformant colonies should be further confirmed for the presence of the foreign gene by Southern blot or PCR analysis, for foreign message RNA expression by Northern blot, and for protein expression by immunoblot or functional assay for the foreign protein.

Generating Multicopy Strains

Strains that contain multiple integrated copies of an expression cassette often produce larger amounts of foreign protein than do single-copy strains [13]. Therefore, after confirming that a single-copy *P. pastoris* strain produces significant amounts of the correct-sized, biologically active protein, it is advisable to construct and examine protein expression by "multicopy strains."

There are three different approaches that can be used reliably to generate multicopy expression strains of *P. pastoris*. The first approach involves constructing a vector with multiple head-to-tail copies of an expression cassette [10]. The second approach entails the

use of an expression vector that contains both the *P. pastoris HIS4* and the bacterial kanamycin resistance gene kan^R, which also confers resistance to the related eukaryotic antibiotic G418 [23]. The third approach involves the use of a vector carrying the bacterial *Sh ble* gene, which confers resistance to the antibiotic zeocin [16]. A detailed description of these methods can be found in Higgins and Cregg [5].

Multicopy expression strains of all three types have proven to be stable under the selective pressure of production in fermentor cultures [16, 23].

Transformation

Methods employed for DNA-mediated transformation of *P. pastoris* are similar to those for *S. cerevisiae* and lead to similar results [15]. *P. pastoris* can be transformed by spheroplast fusion, or by whole-cell methods such as electroporation, or DNA coprecipitation with lithium chloride, calcium chloride, or polyethylene glycol [14, 24–26. Electroporation is a simple and fast method for transforming *P. pastoris*. Although it gives a lower frequency of multicopy transformants than the spheroplast technique, when combined with G418 selection it is ideal for the rapid isolation of multicopy transformants for routine laboratory use [23]. In practice, the electroporation/G418 selection method works best using *AOX1* single-crossover integration in the strain KM71, because two- to fourfold higher transformation frequencies (*e.g.*, 1,000–2,000 colonies/μg DNA) can be achieved with this strain.

The transformation frequencies by electroporation achieved for gene replacement integration are typically 20-fold lower compared to those obtained using the single-crossover method. Therefore, the spheroplast transformation method is recommended for gene replacement integration, especially where very high copy number transformations are required.

Strategies for Efficient Selection

As discussed in "Selectable Markers," a variety of selectable markers has been characterized and incorporated into the expression vectors of *P. pastoris*. These markers can be divided into two subgroups: (a) biosynthetic markers, including *HIS4*, *ARG4*, *ADE1*, *URA3*; and (b) drug-resistance markers, including the kan^R gene, and the *Sh ble* gene that confer resistance to G418 and zeocin, respectively. Combinations of these different markers are employed for efficient selection of true transformants.

The first strategy adopted for generating recombinant strains was single-copy transplacement of the foreign gene at the *AOX1* site, as this type of transformant is the most stable, and in some early studies yielded reasonable levels of product. However, as we described in "Generating Multicopy Strains," numerous examples have now accumulated where multicopy transformants have been used to increase yields dramatically. An important aspect of efficient selection of *P. pastoris* transformants is the selection of multicopy transformants and, thereafter, the determination of the exact copy number. Methods for obtaining multicopy transformants have been discussed above. In addition, the method of rapid, semiquantitative DNA dot blot of whole-cell lysates can be used for screening the resulting transformants. Unlike selection with G418, this method can identify the very high number "jackpot" clones [27].

To determine the absolute copy number of the vector, total genomic DNA from transformed strains is isolated first, and Southern blot analysis is then used to determine the chromosomal structure of the integrated vector DNA (*i.e.* the site of integration, the copy

number of *AOX1* or *HIS4* genes, and whether *AOX1* gene replacement has occurred). A typical protocol can be found in Tuite et al. [28]. In some cases, quantitative dot-blot analysis gives more accurate results [13]. In cases where a good assay is available for the protein expressed, selection can be carried out directly by initial high throughput expression screening of transformants, without any prior knowledge of copy number [29].

Glycosylation

Glycosylation is one of the critical posttranslation processing events in the synthesis of proteins. The role of glycosylation of proteins in protein folding, oligomer assembly, structural stability, specific signal transduction, recognition and secretion processes, and in the clearance of glycoproteins has been well documented [30, 31]. Yeast and most higher eukaryotes utilize an evolutionarily conserved *N*-linked oligosaccharide biosynthetic pathway that involves the formation of a Glc3Man9GlcNAc2-PP-dolichol lipid-linked precursor, the glycan portion of which is transferred cotranslationally in the endoplasmic reticulum (ER) to suitable Asn residues on nascent polypeptides [32]. Subsequently, glycohydrolases in the ER remove the three glucoses and (with the exception of *S. pombe*) a single, specific mannose residue. Processing sugar transferases in the Golgi lead to the formation of core-sized structures ($Hex_{(<15)}GlcNac_{(2)}$) and cores with an extended poly-α1, 6-Man backbone.

Glycosylation Characteristics for *Pichia pastoris*

For *P. Pastoris*, the oligosaccharide chains attached to proteins are shorter and more authentic than in *S. cerevisiae* [33]. The average chain length of glycoproteins expressed by *P. pastoris* is only 8–14 mannose residues, whereas that by *S. cerevisiae* is 40∼150 residues. Analysis of the carbohydrates present on recombinant enzymes indicated the predominant presence of *N*-linked, high-mannose structures rather than complex carbohydrates. In addition, the oligosaccharides secreted by *S. cerevisiae* have terminal α-1,3-linked mannose residues, which will increase the antigenic activity. Oligosaccharides secreted by *P. pastoris*, however, do not have the terminal α-1,3-linkages, and the site of glycosylation is Asn-X-Ser/Thr, the same as that in mammalian cells [34].

Thermostability of the proteins expressed in *P. pastoris* is intriguing. In some cases, introducing glycosylation can increase thermostability, but in most cases, glycosylation does not have an effect on thermostability or may decrease it. Alkalophilic *Bacillus* alpha-amylase (ABA) was produced in *P. pastoris* with a yield of 50 mg/l of culture supernatant. The recombinant protein, rABA, was glycosylated at seven of the nine sites for potential *N*-glycosylation, as identified by automated peptide sequencing and MALDI-TOF MS of tryptic fragments. Extensive *N*-glycosylation, however, reduces thermal stability [35]. Penicillin G amidase from *Providencia rettgeri* is a heterodimer of 92 kDa. The glycosylated rPAC(Pr) produced in *P. pastoris* resulted in improved thermostability compared to its nonglycosylated counterparts and other bacterial penicillin G amidases. The enzyme activities of the rPACPr produced in *P. pastoris* and PACEc after 20 min of incubation at 50°C remained high, whereas more than 90% of the rPACPr protein produced in *S. cerevisiae* became inactive after 10 min [36].

Glycosylation can have effects on proteins other than thermostability, but the effect depends on individual proteins. Table 1 shows some glycosylated proteins recently expressed in *P. pastoris*.

Table 1 Glycosylated proteins expressed in *P. pastoris*.

Proteins	Oligosaccharide	Glycosylation Type	Bio-activity	Reference
α-galactosidase A		*N*-linked	+	[40]
Alkalophilic *Bacillus* α-amylase	$(Man)_{8\text{-}18}$ GlcNAc	*N*-linked	–	[35]
α-1,6 glucan-6-glucanohydrolase	$Man_{(7\text{-}14)}$ $GlcNAc_{(2)}$	*N*-linked	+	[80]
Penicillin G amidase	/	*N*-linked	++	[36]
Carbohydrate-binding modules	$(Man)_{1\text{-}4}GlcNAc$	*O*-linked	+	[81]
Mannose 6-phosphate receptors	$Man_{(8\text{-}12)}$ $GlcNAc_{(2)}$	*N*-linked	+	[82]
Antithrombin	$Man_{(9\text{-}12)}$ $GlcNAc_{(2)}$	*N*-linked	+	[83]
Chicken cystatin	/	*N*-lined	++	[84]
Copper-dependent Fe(II) oxidase	/	*O*-linked, *N*-linked	+	[85]
α-lactolbumin	/	*N*-linked	++	[86]
Candida antarctica lipase B	/	*N*-linked	+	[87]
Surface antigen 1	$Man_{(1\text{-}5)}GlcNAc$	*O*-linked	–	[88]
Allergen	$Hexoses_{(1\text{-}3)}$	*N*-linked	+	[89]
Cutinase	/	*N*-linked	++	[90]
Phytase	/	*N*-linked	+	[91]
Cellobiohydrolase	/	*N*-linked	+	[92]
Aspartic protease	$Man_{(6\text{-}17)}$ $GlcNAc_{(2)}$	*N*-linked	+	[93]
Single-chain Fv		*N*-linked	++	[94]
Carcinoembryonic antigen	$Man_{(9)}GlcNAc_{(2)}$	*N*-linked	+	[95]
Gelatinase B	$Man_{(8\text{-}15)}$ $GlcNAc_{(2)}$	*N*-linked, *O*-linked	+	[96]
Porcine dollicle-stimulating hormone	High-mannose	*N*-linked	+	[97]
α-N-Acetylgalactosaminidase	$Man_{(9\text{-}14)}GlcNAc$	*N*-linked	+	[98]
Placental alkaline phosphatase	/	*N*-linked	+	[99]
Neuraminidase	$Man_{(30\text{-}40)}$ GlcNAc	*N*-linked	+	[100]
Bovine opsin	/	*N*- linked	+	[101]
Angiotensin I-converting Enzyme	/	*N*-linked	+	[102]

+: Bioactivity of the glycosylated protein is the same as that of the native protein

++: Bioactivity of the glycosylated protein is higher than that of the native protein

–: Bioactivity of the glycosylated protein is lower than that of the native protein

/: Result was not reported

Improvement of Glycosylation Profiling in the *P. pastoris* System

Although many glycosylated proteins have been expressed successfully in *P. pastoris*, therapeutic glycoprotein production in this system is hampered by the differences in the protein-linked carbohydrate biosynthesis between this yeast and the target organisms, such as humans. One method to improve this shortage is *N*-glycan engineering. A significant step toward the generation of human-compatible *N*-glycans in this organism is the

conversion of the yeast-type high-mannose glycans to mammalian-type high-mannose and/or complex glycans. Callewaert et al. [37] have coexpressed an endoplasmic reticulum-targeted *Trichoderma reesei* 1,2-alpha-D-mannosidase with two glycoproteins: influenza virus hemagglutinin and *Trypanosoma cruzi trans*-sialidase. The results indicated that the *N*-glycans of the two purified proteins showed a >85% decrease in the number of α-1,2-linked mannose residues, and the human-type high-mannose oligosaccharide $Man_{(5)}$-$GlcNAc_{(2)}$ was the major *N*-glycan of the glyco-engineered *trans*-sialidase.

Another method is using combinatorial genetic libraries to humanize *N*-linked glycosylation [38]. The secretory pathway of *P. pastoris* is genetically reengineered to perform sequential glycosylation reactions that mimic early processing of *N*-glycans in humans and other higher mammals. After eliminating nonhuman glycosylation by deleting the initiating alpha-1,6-mannosyltransferase gene from *P. pastoris*, several combinatorial genetic libraries are constructed to localize active alpha-1,2-mannosidase and human beta-1,2-*N*-acetylglucosaminyltransferase I (GnTI) in the secretory pathway. Recombinant expression of a human reporter protein in these engineered strains leads to the formation of a glycoprotein with GlcNAc-(Man)(5)-(GlcNAc)(2) as the primary *N*-glycan. This strategy has opened the door for engineering yeast to perform complex human-like glycosylation.

Fermentation Process Optimization

Fermentation is essential for secreted proteins because yields correlate largely with the cell density. In some cases, switching from shake-flask expression to fermentation can cause a dramatic increase in yield, with reports of expression levels in fermentation being 10-fold higher than in shake flasks. For *P. pastoris*, two-phase fermentation is generally applied. In the first phase (I) the cells are grown until glycerol is depleted. In the second phase (II) gene expression begins by feeding methanol to the fermenter. Up to now, many heterologous proteins have been expressed in *P. pastoris* because the level expression is equivalent to that of *E. coli* and significantly higher than that of *S. cerevisiae* [1]. Particularly important is the organism's ability to secrete proteins: the secreted product can comprise more than 80% of the total protein in the culture medium [19]. Expression and secretion of these heterologous proteins, however, not only depend on gene dosage, but also on other factors, such as signal sequence recognition and processing, proteolysis, fermentation, and glycosylation.

With regard to fermentation, there are several factors affecting production yield, including culture medium composition, strain type, and non-nutritional factors, such as culture pH, agitation rate, dissolved oxygen, methanol induction, and fermentation strategy. In this section, means to control or improve the fermentation process are presented. It should be pointed out that the general protocols for *P. pastoris* fermentation are provided by Invitrogen, but many studies indicate that recombinant protein production in *P. pastoris* should be optimized according individual processes following established principles.

Medium Composition

Like other yeasts, growth of *P. pastoris* needs sources of carbon and nitrogen. The most common carbon sources are glucose and glycerol, and nitrogen sources are peptone, yeast extract, and yeast nitrogen. For seed culture, MGY and YPD are the usual media, whereas

for scale-up fermentor fermentation, two basic media formulations, basal salts and FM22, have shown good results. Medium composition is thought to influence heterologous protein expression in yeast by affecting cell growth and viability [39, 40]. Yeast extract, casamino acids, or EDTA appear to enhance protein accumulation by *P. pastoris* [41]. Supplementation of the induction medium with 0.4 M L-arginine, 5 mM EDTA, or 2% casamino acids in the BMMY induction medium increased scFv (single chain antibody variable region fragments) production approximately three to fivefold, reaching 25 mg/l of functional scFV [42]. A report on expressing recombinant human bile salt-stimulated lipase (rhBSSL) in *P. pastoris* [43] indicated that in the presence of sorbitol and skimmed milk in the media, together with other optimized conditions, 0.8–1 g of rhBSSL was secreted in 1 l. Salts were also necessary for the production of heterologous proteins in *Pichia*. The presence of at least 200 μM copper was needed for optimal laccase activity in the *Pichia* culture [44]. Adding 100 μM ferric ion to the medium resulted in a significant improvement in the expression of rPLF (recombinant porcine lactoferrin) in *P. pastoris*: expression levels were approximately 12 mg/l, much higer than in *Saccharomyces cerevisiae* [45]. The addition of Triton X-114 during rPIN-a (recombinant puroindoline-a) fermentation increased the production yield of the protein by 10-fold to 13 mg/l and inverted the ratio between secreted and membane-bound rPIN-a [46]. The addition of 0.01% (*v*/*v*) Triton X-100 to a feeding medium reduced partially the proteolysis of a urokinase-type plasminogen activator and increased the secretion level [47].

Temperature

Proteins may be susceptible to misfolding for a variety of reasons, including the formation of intermolecular disulfide bonds and exposure of their hydrophobic surfaces. Intermolecular disulfide bonds are preferentially formed at higher temperatures when the protein is expressed in *E. coli* [48]. Sulfhydryl group shuffling may result in protein aggregation. Higher temperatures may also lead to exposure of more hydrophobic surfaces during peptide folding and favor hydrophobic interactions, and thus may predispose proteins to aggregation [49]. The misfolded and aggregated proteins are more susceptible to intracellular proteolytic degradation [49]. Low temperatures have been shown to improve the solubility of heterologous proteins in *E. coli* [50, 51]. Among the reasons for this observation is the reduced rate of protein synthesis at lower growth temperatures, which may in turn allow more time for the nascent peptide chains to fold properly. An added benefit of lowering the temperature is to reduce the proteolytic degradation of the recombinant protein in the culture medium. Lower culture temperatures have been used frequently to produce proteins in *E. coli* to obtain soluble recombinant proteins that are recalcitrant to expression at 37°C [52].

There are several reports on expressing heterologous proteins in *P. pastoris* at low temperatures (Table 2). Incubation temperatures of 30, 27, 25, and 23°C have been examined in attempts to minimize extracellular proteolysis [3, 53]. *P. pastoris* could be propagated at temperatures as low as 15°C, leading to reduced protease levels and greatly enhanced periods of scFv production [42]. Upon induction, the cell number continued to increase at approximately the same rate at 30 or 15°C. At 15°C, however, intact scFv could be recovered from the culture medium up to 96 h after induction, compared to 24–48 h at 30°C, with peak recovery at 48–72 h, corresponding to the period of maximal number of viable cells. In summary, low-temperature expression may be applied to increase the yields of aggregation-prone and/or unstable gene products in *P. pastoris*, although the fermentation period is longer than at 30°C.

Table 2 Expression of heterologous proteins in *P. pastoris* at low temperature.

Protein	Temperature (°C)	Yield	Reference
scFv	15	25 mg/l	[43]
CBM-CABL	22	1.5 g/l	[55]
Rh-midkine	20	0.36 g/l	[56]
Laccase	20	11,500 U/l	[103]
rhBSSL	20	0.8–1 g/l	[43]
hAFP	23	>5 mg/l	[3]
Galactose oxidase	25	0.5 g/l	[104]

pH-controlled Fermentation

Optimum pH is critical for cell growth, protein formation, and protein stability. Therefore, pH-controlled fermentation is often chosen in *P. pastoris* expression systems [54]. The kinetics of proteolytic reactions, in the presence or absence of cells, were shown to be influenced by pH. Jahic et al. [55] reported that decreasing the pH from 5.0 to 4.0 in bioreactor cultures resulted in an increase in the fraction of full-length product from 40 to 90% during the expression of CBM (cellulose-binding module)-CALB (cellulose 6A and lipase B) in *P. pastoris*. Maintaining the culture at pH 3 or lower in the methanol induction phase has been reported to be effective in protecting product proteolysis in an IGF-1 (insulin-like growth factor) expression system [10]. Approximately 70% of purified recombinant human midkine secreted from *P. pastoris* was truncated when the induction pH was kept at 3. When the methanol induction was kept at pH 3 instead of pH 5, however, fermentation substantially changed the pattern of proteolysis and was highly effective in the purification of authentic rh-midkine [56]. A small multifunctional cytokine, growth-blocking peptide (GBP), from the armyworm *Pseudaletia separata* larvae, was expressed as a soluble and active recombinant peptide in *P. pastoris* when the pH in the fermenter was kept at 3.0 [57]. The medium was collected within 48 h post-methanol shift to minimize exposure of the target peptide to proteases, and up to 50 mg GBP was recovered per liter of yeast culture supernatant. The laccase from *Trametes versicolor* was expressed in *P. pastoris*, and when alanine was added to the medium the pH could be kept at 3, resulting in higher levels of laccase activity compared to cultures grown in the absence of alanine [44]. From the above studies, it appears that *P. pastoris* fermentation should be kept at lower pH, such as 3.0.

Nevertheless, the optimal pH is best determined by running a series of fermentations at different pH. When recombinant porcine lactoferrin was expressed in *P. pastoris*, increasing the initial pH of the culture medium from 6.0 to 7.0 resulted in significant improvements, and the expression level could reach 12 mg/l [45]. The expression of the EGFP-human mu-opioid receptor fusion protein in *P. pastoris* was optimized and monitored using both fluorescence and ligand-binding experiments. A set of parameters, including gene copy number, strain type, temperature, pH, and methanol inducer levels, was studied for its effect on the production of the recombinant protein. The maximum level was reached at a lower temperature and a higher pH than normally used [58]. Therefore, controlling the medium pH during the fermentation process is necessary, the optimum pH depending on individual protein properties, especially stability.

Dissolved Oxygen Control and Methanol Induction

Dissolved oxygen is one of the most important factors for *P. pastoris* cell growth and heterologous protein expression during the fermentation process. Dissolved oxygen can be

controlled via the agitation rate, and the air or O_2 flow rate (Fig. 3). In practice, during the glycerol batch phase, which usually lasts 20–24 h, the dissolved oxygen is controlled by first changing the agitation rate up to 800 rpm for a 4-l fermentor, and 450 rpm for a 60-l fermentor, and then adjusting manually the air or O_2 flow into the fermentor [59]. When glycerol is consumed completely, the dissolved oxygen value will rise rapidly. For most *P. pastoris* fermentations, the dissolved oxygen should be kept consistently at 30~35%, but different proteins need different optimal levels that can only be determined by preliminary experiments.

A methanol induction strategy is critical for expressing successfully heterologous proteins in *P. pastoris*, and the methanol feeding is related closely to the dissolved oxygen level. Therefore, here we consider both of these parameters as part of the induction phase strategies. Considering that different phenotypic strains of *P. pastoris* possess different properties (see Table 3), we suggest that protocols from Invitrogen (Fermentation Guidelines, Invitrogen, San Diego, CA), or other sources [59] for fermentation with each of the mutant phenotype strains be followed. A common feature for the three-phenotype strains is that the methanol level in the fermentor must be monitored carefully either directly by gas chromatography, or indirectly by the dissolved oxygen.

To keep the methanol concentration within optimal limits, three different methanol feed-batch strategies can be introduced: (A) The methanol-feeding rate is controlled according to the concentration in the culture media as determined by gas chromatography. With this strategy it is not easy to control the feeding rate online, because a sample needs to be taken from the fermentor for analysis; thus, it takes at least 30 min to obtain the methanol concentration. (B) The methanol-feeding rate can be controlled by the dissolved oxygen

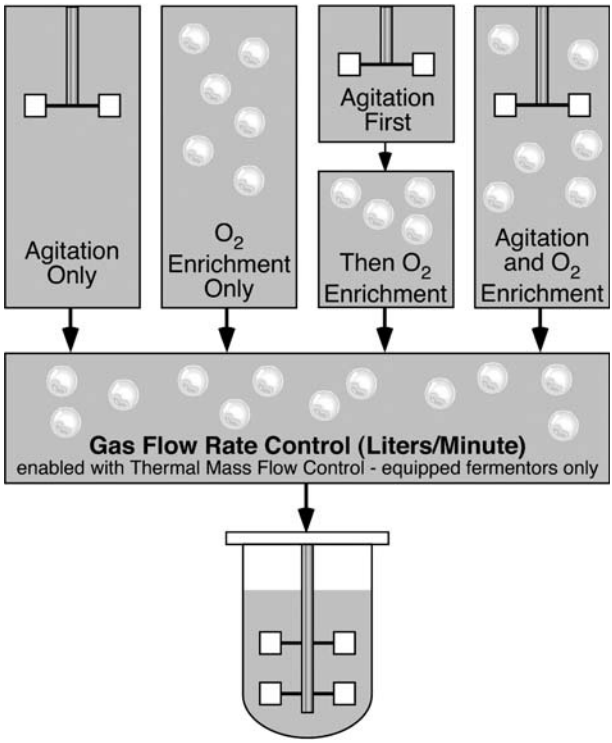

Fig. 3 Dissolved oxygen control options in fermentation mode. In fermentors equipped with a thermal mass flow meter (TMM), DO concentrations may be readily optimized through control of agitation and/or O_2 enrichment. In fermentors equipped with a thermal mass flow controller (TMC), DO control options may be extended by cascading any of these modes either sequentially or simultaneously with air flow. Further optimization may be achieved with the use of optional Bio–Command software to control additional operating parameters such as feeding strategy, temperature, or pH (adapted from www.nbsc.com, with permission from the New Brunswick Scientific Company, Inc.)

Table 3 Properties for different phenotype strains of *P. pastoris*.

Strain type	AOX gene	Residual methanol (%)	Specific growth rate (h^{-1})	Methanol feeding rate
Mut^+	AOX_1^+, AOX_2^+	<0.5	0.14	+++
Mut^s	AOX_1^-, AOX_2^+	0.2–0.8	0.04	++
Mut^-	AOX_1^-, AOX_2^-	0.5	0.0	–

AOX_1^+, AOX_2^+: Strain has AOX_1 and AOX_2 genes

AOX_1^-, AOX_2^-: Strain does not have AOX_1 and AOX_2 genes

+++ = higher feeding rate, ++ = high feeding rate, – = no methanol feeding

value. When the dissolved oxygen increases rapidly, the methanol-feeding rate should be increased automatically, and *vice versa*. In this strategy, the methanol concentration is not given precisely via the dissolved oxygen; thus, the methanol concentration probably exceeds the limiting value when the dissolved oxygen is increasing. This method is in general an experimental strategy. (C) The third strategy for controlling the methanol-feeding rate is according to the specific growth rate during the induction phase. If the growth kinetic model of *P. pastoris* is well set up and described, it is an effective way to keep the methanol concentration within an optimal limit. Minning et al. [60] adjusted the methanol-feeding rate according to the dissolved oxygen, and obtained a 2.5-fold higher productivity in the fermentation of lipases by engineered *P. pastoris*. To improve the expression level, glycerol and methanol were fed simultaneously, followed by a single methanol feed, and resulted in the highest productivity (12,888 U $l^{-1}h^{-1}$), which is 13.6-fold higher than the dissolved oxygen-based strategy. The highest yield of sea raven type II antifreeze protein (30 mg/l) was obtained via a methanol–glycerol mixed feed-batch strategy during the fermentation of engineered *P. pastoris* [61]. The mixed feeding strategy based on growth kinetic studies was also applied to express the heavy-chain fragment C of botulinum neurotoxin serotype C (BoNT/C(Hc)). The results of studies on the relationship between the growth rate and the glycerol/methanol consumption rate have shown that the optimal glyerol/methanol ratio is around 2 for obtaining the highest BoNT/C(Hc) protein, about 3 mg/g wet cells [62].

Several laboratories have investigated the feed strategy based on growth kinetics. High-cell-density cultivation of *P. pastoris* exhibited oscillatory metabolic behavior when fed methanol under closed-loop operations, using a dissolved-oxygen-based bioreactor feed controller (DO-stat). A simple mathematical model of the closed-loop DO-stat was developed, describing the biological process and the components of the standard proportional-integral feedback controller. Inputs into the process model included metabolic pathway information, oxidative metabolism stoichiometry, and substrate uptake kinetics. After analysis, the authors concluded that, when the rate of oxygen transfer approaches the rate of oxygen utilization, the potential for controller destabilization is greatest [63, 64].

Continuous fermentation with an optimal dilution rate or specific growth rate is another strategy to improve expression levels of heterologous proteins in *P. pastoris* (Table 3). When a continuous culture was carried out with a low dilution rate (D=0.025 h^{-1}), the production yield of recombinant puroindoline-a (rPIN-a) was increased by 10-fold, and 80% of the rPIN-a was soluble [46]. A $Mut^{(S)}$ *P. pastoris* strain that had been genetically modified to produce and secrete sea raven antifreeze protein was used as a model system to demonstrate the implementation of a rational, model-based approach to improve process productivity. A set of glycerol/methanol mixed-feed, continuous stirred-tank reactor (CSTR) experiments was performed at the 5-l scale to characterize the relationship

between the specific growth rate and the cell yield on methanol, the specific methanol consumption rate, the specific recombinant protein formation rate, and the productivity, based on secreted protein levels. Two exponential feed-batch fermentations were conducted according to the predicted feeding strategy at specific growth rates of 0.03 h^{-1} and 0.07 h^{-1}. The overall volumetric productivity of both runs was approximately 2.2 mg $l^{-1}h^{-1}$, representing a 10-fold increase in the productivity compared with a heuristic feeding strategy [65]. Recombinant ovine interferon-tau (r-oIFN-tau) production by *P. pastoris* was conducted with different specific growth rates. The r-oIFN-tau concentration in the culture began to decline, despite continued cell growth after 50±6 h of induction, which was associated with an increase in proteolytic activity in the fermentation broth. The best results were obtained when the specific growth rate was stepped down from 0.025 to 0.02 h^{-1} at 38 h of induction, whereby the active production period was prolonged up to 70 h after induction, and the broth protease activity was correspondingly reduced. The corresponding maximum protein yield was 392 mg/l after 70 h of fermentation. The production profile of r-oIFN-tau was found to be significantly different from other secreted and intracellular recombinant protein processes, which is an indication that recombinant protein production in *P. pastoris* needs to be optimized as individual processes following established principles [66]. Similarly, controlling the growth rate at 0.02 h^{-1}, total endostatin production reached 400 mg in 3 l of initial fermentation volume [67].

A more effective strategy for keeping the methanol within a suitable concentration during the induction phase may be designed by the yeast metabolism responding to the methanol consumption. In this strategy, a program correlating growth kinetics with methanol concentration is set up; when the methanol concentration matches a set point the control system automatically feeds the additives needed for the optimized production of the expressed protein. A good control is only achieved via a robust feedback control system. A typical design of a methanol feed control system based on a growth model was described in detail by Zhang et al., [64] (Fig. 4).

Prospects and Summary

The practice for improving protein thermostability via protein engineering and other expression systems can also implement in *P. pastoris* with ease and efficiency. The thermal stability of a given protein can be improved by site-directed mutagenesis at the gene level, producing mutants with higher T_m and longer shelf lives [68, 69]. This may be performed either by a directed fashion producing mutations based on rational site-directed mutagenesis selection algorithm (RSDMSA) [70], or by a randomized method to screen and select the objective thermal protein from the mutant library. Additionally, introduction of posttranslational modification sites, such as glycosylation, phosphorylation, and other schemes, may be applied to improve the thermal stability. For example, the protein sequence Arg-Arg-X-Ser-Thr-Tyr (RRXSTY) seems be frequently phosphorylated [71, 72], so we can introduce such sequence into the objective protein through the mutation methods.

The purification of membrane proteins from cellular membranes is tedious. When membrane proteins are expressed in *P. pastoris*, however, they are secreted into the medium in membrane vesicles, rendering purification more straightforward. Several studies have indicated that the methylotrophic yeast contains an inducible import pathway for peroxisomal matrix proteins with an N-terminal targeting signal [73, 74, 11, 25]. The peroxisomal matrix proteins are localized in a small vesicular compartment [75], indicating

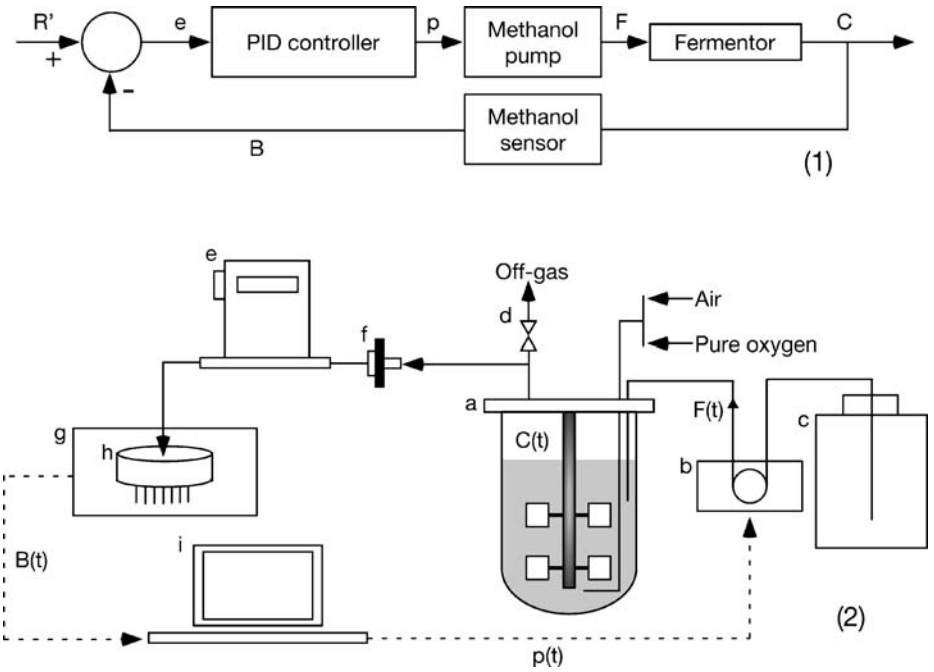

Fig. 4 A closed-loop methanol control system. (1) Schematic diagram: (R′) setpoint; (B) measured value; (e)=R′ − B; (p) PID controller output; (F) methanol pump feed rate; (C) methanol concentration in fermentor. (2) Experimental setup: (→) mass flow direction; (...>) control signal flow direction. (a) Bioflo 3000 5-l fermentor; (b) methanol feed pump; (c) methanol reservoir; (d) valve; (e) MFC 1104 flow rate controller for off-gas; (f) self-locking filter; (g) MC-168 methanol monitor and controller; (h) Figaro TGS822 alcohol sensor; (i) PID controller built in NBS AFS-BioCommand multiprocess management hard/software. (Adapted with permission from Dr. Wenhui Zhang, Departments of Chemical Engineering and Food Science and Technology, University of Nebraska-Lincoln, and United States Army Medical Research Institute of Infectious Diseases (USAMRIID)), Fort Detrick, MD

that the heterologous protein was synthesized in the cytosol and subsequently sorted to the peroxisomal matrix. Several exogenous membrane proteins have been expressed successfully in the *Pichia* system, including 5-HT5A serotonin receptor [76], human μ-opioid receptor [77], CD40 ligand [78], and phytochrome [79].

Pichia pastoris has many advantages as a eukaryotic expression system for recombinant proteins, including multiple copy selection, strong promoter activity, and facilitation of secretion. This system is particularly suitable for the production of proteins that form inclusion bodies in *E. coli*, and whose expression levels are very low in mammalian cell lines. *P. pastoris* has been used for the production of vaccines, coagulation inhibitors, fibrinolytic compounds, allergens, antibodies, protease inhibitors, hormones, cytokines, receptors, and ligands. Proteins for structural studies, such as single or dual ^{13}C and ^{15}N-labeled proteins, have also been produced in this system, which may facilitate the analysis of membrane proteins. Thus, *P. pastoris* has become the preferred option among the various yeast expression systems. A commercial kit based on *P. pastoris* is now available and subject to further improvement. Fermentation process optimization is likely to make this system more competitive and reproducible in producing relevant compounds on both laboratory and industrial scales.

Acknowledgments We thank Dr. Geoff Lin-Cereghino (Department of Biology, University of the Pacific) for helpful comments on the manuscript, and Mr. Neal Johnson (School of Dentistry, University of the Pacific) for preparing the figures. The research described here was supported by US AFOSR, NSF, the Wallace Coulter Foundation, and Harvard Medical School.

References

1. Romanos, M. A., Scorer, C. A., & Clare, J. J. (1992). *Yeast, 8*, 423–488.
2. Cregg, J. M., Vedvick, T. S., & Raschke, W. C. (1993). *Biotechnology (NY), 11*, 905–910.
3. Li, P. Z., Gao, X.-G., Arellano, R. O., & Renugopalakrishnan, V. (2001). *Protein Expression and Purification, 22*, 369–380.
4. Cereghino, G. P., Cereghino, J. L., Ilgen, C., & Cregg, J. M. (2002). *Current Opinion in Biotechnology, 13*, 329–332.
5. Higgins, D. R., & Cregg, J. M. (1998). *Methods in Molecular Biology, 103*, 1–15.
6. Lin-Cereghino, G. P., Sunga, A. J., Lin Cereghino, J., & Cregg, J. M. (2001). *Genetic Engineering (NY), 23*, 157–169.
7. Koutz, P. J., Davis, G. R., Stillman, C., Barringer, K., Cregg, J., & Thill, G. (1989). *Yeast, 5*, 167–177.
8. Waterham, H. R., Digan, M. E., Koutz, P. J., Lair, S. V., & Cregg, J. M. (1997). *Gene, 186*, 37–44.
9. Shen, S., Sulter, G., Jeffries, T. W., & Cregg, J. M. (1998). *Gene, 216*, 93–102.
10. Brierley, R. A. (1998). *Methods in Molecular Biology, 103*, 149–177.
11. Liu, H., Tan, X., Russell, T. A., Veenhuis, M., & Cregg, J. M. (1995). *Journal of Biological Chemistry, 270*, 10940–10951.
12. Sears, I. B., O'Connor, J., Rossanese, O. W., & Glick, B. S. (1998). *Yeast, 14*, 783–790.
13. Clare, J. J., Romanos, M. A., Rayment, F. B., Rowedder, J. E., Smith, M. A., Payne, M. M., et al. (1991). *Gene, 105*, 205–212.
14. Cregg, J. M., Barringer, K. J., Hessler, A. Y., & Madden, K. R. (1985). *Molecular and Cellular Biology, 5*, 3376–3385.
15. Cregg, J. M., & Madden, K. R. (1989). *Molecular and General Genetics, 219*, 320–323.
16. Higgins, D. R., Busser, K., Comiskey, J., Whittier, P. S., Purcell, T. J., & Hoeffler, J. P. (1998). *Methods in Molecular Biology, 103*, 41–53.
17. Cereghino, J. L., & Cregg, J. M. (2000). *FEMS Microbiology Reviews, 24*, 45–66.
18. Veenhuis, M., van Dijken, J. P., & Harder, W. (1983). *Advances in Microbial Physiology, 24*, 1–82.
19. Tschopp, J. F., Sverlow, G., Kosson, R., Craig, W., & Grinna, L. (1987). *Biotechnology (NY), 5*, 1305–1308.
20. Cregg, J. M., & Madden, K. R. (1987). In. G. G. Stewart, I. Russell, R. D. Klein, & R. R. Hiebsch (Eds.), *Biological Research on Yeasts*, Vol. 2. Boca Raton, FL.: CRC Press.
21. Chiruvolu, V., Cregg, J. M., & Meagher, M. M. (1997). *Enzyme and Microbial Technology, 21*, 277–283.
22. White, C. E., Hunter, M. J., Meininger, D. P., White, L. R., & Komives, E. A. (1995). *Protein Engineering, 8*, 1177–1187.
23. Scorer, C. A., Clare, J. J., McCombie, W. R., Romanos, M. A., & Sreekrishna, K. (1994). *Biotechnology (NY), 12*, 181–184.
24. Liu, H., Tan, X., Veenhuis, M., McCollum, D., & Cregg, J. M. (1992). *Bacteriology, 174*, 4943–4951.
25. Waterham, H. R., de Vries, Y., Russell, K. A., Xie, W., Veenhuis, M., & Cregg, J. M. (1996). *Molecular and Cellular Biology, 16*, 2527–2536.
26. Cregg, J. M., & Russell, K. A. (1998). *Methods in Molecular Biology, 103*, 27–39.
27. Sreekrishna, K., Potenz, R. B., Cruze, J. A., McCombie, W. R., Parker, K. A., Nelles, L., et al. (1988). *Journal of Basic Microbiology, 28*, 265–278.
28. Tuite, M. F., Clare, J. J., & Romanos, M. A. (1999). In S. J. Higgins & B. D. Hames (Eds.), Protein expression, a practical approach. Oxford University Press.
29. Laroche, Y., Storme, V., De Meutter, J., Messens, J., & Lauwereys, M. (1994). *Biotechnology (NY), 12*, 1119–1124.
30. Wright, A., & Morrison, S. L. (1997). *Trends in Biotechnology, 15*, 26–32.
31. Jenkins, N., Parekh, R. B., & James, D. C. (1996). *Nature Biotechnology, 14*, 975–981.
32. Gemmill, T. R., & Trimble, R. B. (1999). *Biochimica et Biophysica Acta, 1426*, 227–237.
33. Grinna, L., & Tschopp, J. F. (1989). *Yeast, 5*, 107–115.
34. Trimble, R. B., Atkinson, P. H., Tschopp, J. F., Townsend, R. R., & Maley, F. (1991). *Journal of Biological Chemistry, 266*, 22807–22817.

35. Tull, D., Gottschalk, T. E., Svendsen, I., Kramhoft, B., Phillipson, B. A., Bisgard-Frantzen, H., et al. (2001). *Protein Expression and Purification, 21*, 13–23.
36. Sevo, M., Degrassi, G., Skoko, N., Venturi, V., & Ljubijankic, G. (2002). *FEM Yeast Research, 4*, 271–277.
37. Callewaert, N., Laroy, W., Cadirgi, H., Geysens, S., Saelens, X., MinJou, W., et al. (2001). *FEBS Letters, 503*, 173–178.
38. Choi, B. K., Bobrowicz, P., Davidson, R. C., Hamilton, S. R., Kung, D. H., Li, H., et al. (2003). *Proceedings of the National Academy of Sciences of the United States of America, 100*, 5022–5027.
39. Kang, H. A., Choi, E. S., Hong, W. K., Kim, J. Y., Ko, S. M., Sohn, J. H., et al. (2000). *Applied Microbiology and Biotechnology, 53*, 575–582.
40. Chen, D. C., Wang, B. D., Chou, P. Y., & Kuo, T. T. (2000). *Yeast, 16*, 207–217.
41. Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J. T., Smith, P. L., et al. (1997). *Gene, 190*, 55–62.
42. Shi, X., Karkut, T., Chamankhah, M., Alting-Mees, M., Hemmingsen, S. M., & Hegedus, D. (2003). *Protein Expression and Purification, 28*, 321–330.
43. Murasugi, A., Asami, Y., & Mera-Kikuchi, Y. (2001). *Protein Expression and Purification, 23*, 282–288.
44. O'Callaghan, J., O'Brien, M. M., McClean, K., & Dobson, A. D. (2002). *Journal of Industrial Microbiology and Biotechnology, 29*, 55–59.
45. Wang, S. H., Yang, T. S., Lin, S. M., Tsai, M. S., Wu, S. C., & Mao, S. J. (2002). *Protein Expression and Purification, 25*, 41–49.
46. Issaly, N., Solsona, O., Joudrier, P., Gautier, M. F., Moulin, G., & Boze, H. J. (2001). *Applied Microbiology, 90*, 397–406.
47. Tsujikawa, M., Okabayashi, K., Morita, M., & Tanabe, T. (1996). *Yeast, 12*, 541–553.
48. Schein, C. H., & Noteborn, M. N. H. (1988). *Biotechnology (NY), 6*, 291–294.
49. Lee, S. C., Choi, Y. C., & Yu, M. H. (1990). *European Journal of Biochemistry, 187*, 417–424.
50. Broeze, R. J., Solomon, C. J., & Pope, D. H. (1978). *Journal of Bacteriology, 34*, 861–874.
51. Han, K. G., Lee, S. S., & Kang, C. (1999). *Protein Expression and Purification, 16*, 103–108.
52. Makrides, S. C. (1996). *Microbiological Reviews, 60*, 512–538.
53. Curvers, S., Brixius, P., Klauser, T., Thommes, J., Weuster-Botz, D., Takors, R., et al. (2001). *Biotechnology Progress, 17*, 495–502.
54. Soden, D. M., O'Callaghan, J., & Dobson, A. D. (2002). *Microbiology, 148*, 4003–4014.
55. Jahic, M., Gustavsson, M., Jansen, A. K., Martinelle, M., & Enfors, S. O. (2003). *Journal of Biotechnology, 102*, 45–53.
56. Murasugi, A., & Tohma-Aiba, Y. (2003). *Protein Expression and Purification, 27*, 244–252.
57. Koganesawa, N., Aizawa, T., Shimojo, H., Miura, K., Ohnishi, A., Demura, M., et al. (2002). *Protein Expression and Purification, 25*, 416–425.
58. Sarramegna, V., Demange, P., Milon, A., & Talmont, F. (2002). *Protein Expression and Purification, 24*, 212–220.
59. Stratton, J., Chiruvolu, V., & Meagher, M. (1998). *Methods in Molecular Biology, 103*, 107–120.
60. Minning, S., Serrano, A., Ferrer, P., Sola, C., Schmid, R. D., & Valero, F. (2001). *Journal of Biotechnology, 86*, 59–70.
61. Loewen, M. C., Liu, X., Davies, P. L., & Daugulis, A. J. (1997). *Journal of Microbiology and Biotechnology, 48*, 480–486.
62. Zhang, W., Hywood Potter, K. J., Plantz, B. A., Schlegel, V. L., Smith, L. A., & Meagher, M. M. (2003). *Journal of Industrial Microbiology & Biotechnology, 30*, 210–215.
63. Chung, J. D. (2000). *Biotechnology & Bioengineering, 68*, 298–307.
64. Zhang, W., Smith, L. A., Plantz, B. A., Schlegel, V. L., & Meagher, M. M. (2002). *Biotechnology Progress, 18*, 1392–1399.
65. d'Anjou, M. C., & Daugulis, A. J. (2001). *Biotechnology & Bioengineering, 72*, 1–11.
66. Sinha, J., Plantz, B. A., Zhang, W., Gouthro, M., Schlegel, V., Liu, C. P., et al. (2003). *Biotechnology Progress, 19*, 794–802.
67. Trinh, L. B., Phue, J. N., & Shiloach, J. (2003). *Biotechnology & Bioengineering, 82*, 438–444.
68. Li, Z., Xiong, F., Lin, Q., d´Anjou, M., Daugulis, A., Yang, D. S. C., et al. (2001). *Protein Expression and Purification, 21*, 438–445.
69. Oobatake, M., Yamasaki, T., Simmer, J. P., & Renugopalakrishnan, V. (2006). *Proteins, 62*, 461–469.
70. Renugopalakrishnan, V., Strzelczyk, A., Li, P. Z., Mokhnatyuk, A. A., Gursahani, S. H., Nagaraju, M., et al. (2003). *International Journal of Quantum Chemistry, 95*, 625–631.
71. Pearson, R., & Kemp, B. (1991). *Methods in Enzymology, 200*, 62–81.
72. Takagi, T., Suzuki, M., Baba, T., Minegishi, K., & Sasaki, S. (1984). *Biochemical and Biophysical Research Communications, 121*, 592–597.

73. Waterham, H. R., Titorenko, V. I., Haima, P., Cregg, J. M., Harder, W., & Veenhuis, M. (1994). *Journal of Cell Biology, 127*, 737–749.
74. Faber, K. N., Haima, P., Gietl, C., Harder, W., Ab, G., & Veenhuis, M. (1994). *Proceedings of the National Academy of Sciences of the United States of America, 91*, 12985–12989.
75. De Hoop, M. J., & Geert, A. B. (1992). *Biochemical Journal, 286*, 657–669.
76. Weiss, H. M., Haase, W., Michel, H., & Reilander, H. (1995). *FEBS Letters, 377*, 451–456.
77. Talmont, F., Sidobre, S., Demange, P., Milon, A., & Emorine, L. J. (1996). *FEBS Letters, 394*, 268–272.
78. McGrew, J. T., Leiske, D., Dell, B., Klinke, R., Krasts, D., Wee, S. F., et al. (1997). *Gene, 187*, 193–200.
79. Wu, S.-H., & Lagarias, J. C. (1996). *Proceedings of the National Academy of Sciences of the United States of America, 93*, 8989–8994.
80 Betancourt, L. H., Garcia, R., Gonzales, J., Montesino, R., Quintero, O., Takao, T., et al. (2001). *FEMS Yeast Research, 1,* 151–160.
81. Boraston, A. B., Sandercock, L. E., Warren, R. A., Kilburn, D. G. (2003). *Journal of Molecular Microbiology and Biotechnology, 5,* 29-36.
82. Reddy, S. T., Dahms, N. M. (2002). *Protein Expression and Purification*, 26, 290-300.
83. Hirose, M., Kameyama, S., Ohi, H. (2002). *Yeast, 19,* 1191-1202.
84. Jiang, S. T., Chen, G. H., Tang, S. J., Chen, C. S. (2002). *Journal of Agricultural and Food Chemistry, 50,* 5313-5317.
85. Watmann, T., Stephan, U. W., Bube, I., Boer, E., Melzer, M., Manteuffel, R., et al. (2002). *Yeast, 19,* 849-862.
86. Saito, A., Usui, M., Song, Y., Azakami, H., Kato, A. (2002). *Journal of Biochemistry (Tokyo)*, *132*, 77-82.
87. Rotticci-Mulder, J. C., Gustavsson, M., Holmquist, M., Hult, K., Martinelle, M. (2001). *Protein Expression and Purification, 21,* 386-392.
88. Letourneur, O., Gervasi, G., Gaia, S., Pages, J., Watelet, B., Jolivet, M. (2001). *Biotechnology and Applied Biochemistry, 33,* 35-45.
89. Best, E. A., Stedman, K. E., Bozic, C. M., Hunter, S. W., Vailes, L., Chapman, M. D. (2000). *Protein Expression and Purification, 20,* 462-471.
90. Sagt, C. M., Kleizen, B., Verwaal, R., de Jong, M. D., Muller, W. H., Smits, A. (2000). *Applied and Environmental Microbiology, 66,* 4940-4944.
91. Rodriguez, E., Wood, Z. A., Karplus, P. A., Lei, X. G. (2000). *Archives of Biochemistry and Biophysics, 382,* 105-112.
92. Boer, H., Teeri, T. T., Koivula, A. (2000). *Biotechnology and Bioengineering*, *69*, 486-494.
93. Montesino, R., Nimtz, M., Quintero, O., Garcia, R., Falcon, V., Cremata, J. A. (1999). *Glycobiology, 9,* 1037-1043.
94. Wang, M., Lee, L. S., Nepomich, A., Yang, J. D., Conover, C., Whitlow, M. et al. (1998). *Protein Engineering, 11,* 1277-1283.
95. You, Y. H., Hefta, L. J., Yazaki, P. J., Wu, A. M., Shively, J. E. (1998). *Anticancer Research, 18,* 3193-3201.
96. Van den Steen, P., Rudd, P. M., Proost, P., Martens, E., Paemen, L., Kuster, B. et al. (1998). *Biochimica et Biophysica Acta, 1425*, 587-598.
97. Richard, F, Robert, P, Remy, JJ, Martinat, N, Bidart, JM, Salesse, R., et al. (1998). *Biochemical and Biophysical Research Communications, 245*, 847-852.
98. Zhu, A., Wang, Z. K., Beavis, R. (1998). *Archives of Biochemistry and Biophysics, 352,* 1-8.
99. Heimo, H., Palmu, K., Suominen, I. (1998). *Protein Expression and Purification, 12*, 85-92.
100. Martinet, W., Saelens, X., Deroo, T., Neirynck, S., Contreras, R., Min Jou, W., et al. (1997). *European Journal of Biochemistry, 247*, 332-338.
101. Abdulaev, N. G., Popp, M. P., Smith, W. C., Ridge, K. D. (1997). *Protein Expression and Purification, 10,* 61-69.
102. Williams, T. A., Michaud, A., Houard, X., Chauvet, M. T., Soubrier, F., Corvol, P. (1996). *Biochemical Journal, 318*, 125-131.
103. Hong, F., Meinander, N. Q., Jonsson, L. J. (2002). *Biotechnology Bioengineering, 79*, 438-449.
104. Whittaker, M. M., Whittaker, J. W. (2000). *Protein Expression and Purification, 20*, 105-111.