

Review

Heterologous protein production using the *Pichia pastoris* expression system

Sue Macauley-Patrick,* Mariana L. Fazenda, Brian McNeil and Linda M. Harvey

Strathclyde Fermentation Centre, Department of Bioscience, University of Strathclyde, 204 George Street, Glasgow G1 1XW, UK

*Correspondence to:
Sue Macauley-Patrick,
Strathclyde Fermentation Centre,
Department of Bioscience,
University of Strathclyde, 204
George Street, Glasgow G1
1XW, UK.
E-mail:
s.e.macauley@strath.ac.uk

Abstract

The *Pichia pastoris* expression system is being used successfully for the production of various recombinant heterologous proteins. Recent developments with respect to the *Pichia* expression system have had an impact on not only the expression levels that can be achieved, but also the bioactivity of various heterologous proteins. We review here some of these recent developments, as well as strategies for reducing proteolytic degradation of the expressed recombinant protein at cultivation, cellular and protein levels. The problems associated with post-translational modifications performed on recombinant proteins by *P. pastoris* are discussed, including the effects on bioactivity and function of these proteins, and some engineering strategies for minimizing unwanted glycosylations. We pay particular attention to the importance of optimizing the physicochemical environment for efficient and maximal recombinant protein production in bioreactors and the role of process control in optimizing protein production is reviewed. Finally, future aspects of the use of the *P. pastoris* expression system are discussed with regard to the production of complex membrane proteins, such as G protein-coupled receptors, and the industrial and clinical importance of these proteins. Copyright © 2005 John Wiley & Sons, Ltd.

Received: 13 September 2004
Accepted: 30 November 2004

Keywords: foreign protein production; heterologous expression; methylotrophic; *Pichia pastoris*; post-translational modification; process optimization; yeast

Contents

Introduction	249
The <i>Pichia</i> expression systems	251
Selectable markers	253
Controlling proteolysis	253
Gene dosage	256
Post-translational modification of secreted proteins	256
Production of heterologous proteins in <i>P. pastoris</i>	259
Future aspects	263
Conclusions	265
References	266

Introduction

General

Over the last few decades, geneticists have learned how to manipulate DNA to identify, move and place genes into a variety of organisms that are quite different from the source organism. A major use

for many of these recombinant organisms is to produce proteins. Since many proteins are of immense commercial value, numerous studies have focused on finding ways to produce them efficiently and in a functional form.

The production of a functional protein is intimately related to the cellular machinery of the organism producing the protein. The yeast *Pichia pastoris* is a useful system for the expression of milligram-to-gram quantities of proteins for both basic laboratory research and industrial manufacture. The fermentation can be readily scaled up to meet greater demands, and parameters influencing protein productivity and activity, such as pH, aeration and carbon source feed rate, can be controlled [46]. Compared with mammalian cells, *Pichia* does not require a complex growth medium or culture conditions, is genetically relatively easy to manipulate, and has a eukaryotic protein synthesis pathway. Because of these characteristics, discussed in more detail in the following sections, some proteins, such as G protein-coupled receptors, that cannot be expressed efficiently in bacteria, *Saccharomyces cerevisiae* or the insect cell/baculovirus system, have been successfully produced in functionally active form in *P. pastoris* [20,21,33]. This methylotrophic yeast is particularly suited to foreign protein expression for a number of reasons, including ease of genetic manipulation, e.g. gene targeting, high-frequency DNA transformation, cloning by functional complementation, high levels of protein expression at the intra- or extracellular level, and the ability to perform higher eukaryotic protein modifications, such as glycosylation, disulphide bond formation and proteolytic processing [29]. *Pichia* can be grown to very high cell densities using minimal media [113,117] and integrated vectors help genetic stability of the recombinant elements, even in continuous and large-scale fermentation processes [90]. Simple purification of secreted recombinant proteins is possible due to the relatively low levels of native secreted proteins [28]. Therefore, the powerful genetic techniques available, together with its economy of use, make *P. pastoris* a system of choice for heterologous protein expression.

Background

The Phillips Petroleum Company was the first to develop media and protocols for growing *P. pastoris* on methanol in continuous culture at high cell densities (>130 g/dry cell weight [18]). During the 1970s, *P. pastoris* was evaluated as a potential source of single-cell protein due to the ability of this yeast to utilize methanol as sole carbon source. Unfortunately, the oil crisis of the 1970s caused a dramatic increase in the cost of methane (the source of the methanol). Simultaneously, the price of soybeans, the major alternative source of animal feed, fell. Therefore, the economics of single cell protein (SCP) production from methanol became highly unfavourable. The ICI equivalent product, Pruteen (*Methylophilus methylotrophus*), underwent a very similar pattern of development and economic non-viability in the same time period [11,31].

However, in the following decade, Phillips Petroleum, together with the Salk Institute Biotechnology/Industrial Associates Inc. (SIBIA, La Jolla, CA, USA), studied *P. pastoris* as a system for heterologous protein expression. The gene and promoter for alcohol oxidase were isolated by SIBIA, who also generated vectors, strains and corresponding protocols for the molecular manipulation of *P. pastoris*. What began more than 20 years ago as a program to convert abundant methanol to a protein source for animal feed has developed into what are today two important biological tools: a model eukaryote used in cell biology research, and a recombinant protein production system. *Pichia* has gained widespread attention as an expression system because of its ability to express high levels of heterologous proteins. As a result, recombinant vector construction, methods for transformation, selectable marker generation, and fermentation methods have been developed to exploit the productive potential of this system [91].

Research Corporation Technologies (Tucson, AZ, USA) are the current holders of the patent for the *P. pastoris* expression system, which they have held since 1993, and the *P. pastoris* expression system is available in kit form from Invitrogen Corporation (Carlsbad, CA, USA).

Methanol metabolism in *Pichia*

Methylotrophic yeasts are able to utilize methanol as the sole carbon and energy source. All such strains identified to date belong to only four genera: *Hansenula*, *Pichia*, *Candida* and *Torulopsis* [36]. They all share a specific methanol utilization pathway involving several unique enzymes. The initial reactions take place in specialized microbodies, the peroxisomes, followed by subsequent metabolic steps in the cytoplasm. Peroxisomes play an indispensable role during growth, as they harbour the three key enzymes for methanol metabolism, viz. alcohol oxidase, catalase and dihydroxyacetone synthase. The subsequent reactions of methanol assimilation and dissimilation are localized in the cytosol. The methanol metabolism is fully described in Cereghino and Cregg [18], Jahic [53] and Gellissen [38].

Methanol utilization phenotypes

There are three phenotypes of *P. pastoris* host strains with regard to methanol utilization. The Mut⁺, or methanol utilization plus phenotype, grow on methanol at the wild-type rate and require high feeding rates of methanol in large-scale fermentations [18]. The Mut^s, or methanol utilization slow phenotype, have a disruption in the *AOX1* gene. Since the cells must then rely on the weaker *AOX2* for methanol metabolism, a slower growing and slower methanol utilization strain is produced. Mut^s strains have been found to be advantageous for production of hepatitis B surface antigen [26]. The Mut⁻, or methanol utilization minus phenotype, are unable to grow on methanol, since these strains have both *AOX* genes deleted. One of the advantages of this phenotype is that low growth rates may be desirable for production of certain recombinant products [26,27]. Currently, the majority of researchers use the Mut⁺ phenotype [47,102,129], although some researchers are also using the Mut^s phenotype [1,71,83,85,126].

The *Pichia* Expression Systems

The expression of any foreign gene in *P. pastoris* comprises three principal steps: (a) insertion of the gene into an expression vector; (b) introduction of the expression vector into the *P. pastoris* host; and (c) examination of potential strains for the expression of the foreign gene. More information about the expression vectors available for this system can be found in Cereghino and Cregg [18], Cregg *et al.* [28], Koutz *et al.* [61] and Sreekrishna *et al.* [103]. More recent developments in the use of *Pichia* expression vectors are discussed below.

The majority of heterologous protein production in *P. pastoris* is based on the fact that enzymes required for the metabolism of methanol are only present when cells are grown on methanol [35]. This has been the most successful system reported for this organism [19]. The *AOX* promoters have therefore been the most widely utilized; however, other promoter options are available for the production of foreign proteins in *Pichia*, and these are discussed comprehensively by Cereghino and Cregg [18]. The advantages and disadvantages of using the *AOX1* promoter are represented in Table 1.

The *AOX1* promoter has been the most widely reported and utilized of all the available promoters for *P. pastoris* [19]. One of the reasons that the constitutive *GAP* (glyceraldehyde 3-phosphate dehydrogenase) promoter has not been widely used is the belief that constitutive production of foreign proteins in *P. pastoris* may have cytotoxic effects [73]. However, recent studies have found not only that cytotoxic effects are not necessarily observed, but also that production levels of a recombinant exo-levanase (LsdB) using the *GAP* promoter were similar to those using the *AOX1* promoter [73]. The levanase activity reportedly achieved using each of these promoters was 21.1 U/ml under the control of the *AOX1* promoter and 26.6 U/ml with the *GAP* promoter. However, the biomass levels and fermentation time for each of these processes differed remarkably: 96 h and 115.2 g/l (dry cell weight) for LsdB produced under the control of the *AOX1* promoter, and 39 h and 59.7 g/l (dry cell weight) for LsdB produced using the *GAP* promoter. These results appear to favour the *GAP* promoter for production

Table 1. Advantages and disadvantages of using the *AOX1* promoter system

Advantages	Disadvantages
Transcription of the foreign protein is tightly regulated and controlled by a repression/derepression mechanism	Monitoring methanol during a process is often difficult, due to the unreliability of on-line probes and the complications of measuring off-line
High levels of foreign proteins can be expressed, even if they are toxic to the cell	Methanol is a fire hazard; therefore, storing the large quantities required for these processes is undesirable
The repression of transcription by the initial carbon source ensures that good cell growth is obtained before the gene product is overexpressed	Methanol is mainly derived from petrochemical sources, which may be unsuitable for use in the production of certain food products and additives
Induction of transcription is easily achieved by the addition of methanol	Two carbon sources are required, with a switching over from one to the other at a precise time point

of *LsdB*; however, no direct comparison can be made without knowing the production levels of this recombinant protein relative to the biomass produced.

Combining the *GAP* and *AOX1* promoters in a strain expressing human granulocyte-macrophage colony-stimulating factor (hGM-CSF) resulted in a two-fold increase in production of the recombinant protein, i.e. 90 mg/l produced using the *GAP* promoter and 180 mg/l produced using *GAP/AOX1* combined promoters [123]. By using this combined promoter for sequential expression of a constitutively produced enzyme required for post-translational modification, an inducible recombinant protein can be modified to a specified form within the fermentation process itself [17,122].

The *YPT1* (a small GTPase involved in secretion [96]) and *PEX8* (a peroxisomal matrix protein [65]) promoters have not been widely used, probably because of their low expression levels, since the goal of the majority of researchers using the *P. pastoris* expression system is to obtain maximum amounts of expressed foreign proteins [18]. The remainder of this section will discuss the more recent developments regarding *P. pastoris* promoters.

AOX2 promoter

The *AOX2* gene also produces alcohol oxidase, although this gene yields 10–20 times less AOX activity than the *AOX1* gene [27]. However, there have been some successful reports of increased expression levels using the *AOX2* promoter or with a truncated version of the promoter [75]. Production levels of a recombinant human serum albumin using the *AOX2* promoter were enhanced by the addition of a small amount of oleic acid (0.01%), resulting in 80 mg/l of the protein as opposed to 40 mg/l in the absence of oleic acid [59]. This demonstrates that the *AOX2* promoter can be efficient if the physicochemical environment has been optimized, and that addition of anti-foam agents, such as oleic acid, may improve transcriptional regulation of *AOX2*. Generally, expression levels of foreign proteins under the control of the *AOX1* promoter have been much higher than those reported for the *AOX2* promoter. For this reason, the majority of researchers using the methanol induction system opt for the *AOX1* promoter [19].

FLD1 promoter

Formaldehyde dehydrogenase (*FLD*) is a key enzyme involved in the methanol metabolic pathway in *Pichia* [113]. It is also involved in protection of the cell from toxic effects caused by formaldehyde during methylamine metabolism, which can be used as the sole nitrogen source for the cells [97].

The *FLD1* gene is inducible by both methanol and methylamine. The advantage of this is that when using methylamine for induction of foreign protein expression, glucose or glycerol can be used as the carbon source instead of methanol, or methanol itself can be used as the sole carbon source and also for induction [89]. This, coupled with the *AOX1*-comparable tight regulation and transcriptional efficiency, makes the *FLD1* promoter an attractive alternative for methanol-independent expression of recombinant

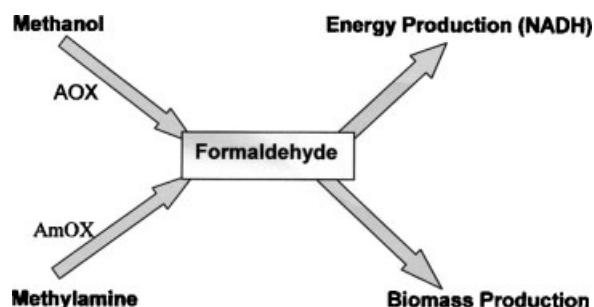


Figure 1. The intermediate formaldehyde is common to both the methanol and methylamine pathways in the yeast *P. pastoris*. Formaldehyde is produced by alcohol oxidase (AOX) in the methanol pathway, with the co-production of hydrogen peroxide (H_2O_2); and also by amine oxidase (AmOX) in the methylamine pathway, with the co-production of H_2O_2 and NH_3 .

proteins in *Pichia* [97]. Figure 1 shows the role of the common intermediate, formaldehyde, in both the methanol and methylamine pathways.

Resina *et al.* [89] found that induction levels of the *FLD1* promoter by methanol were dependent on the nitrogen source used, with co-induction using methanol and methylamine resulting in higher expression levels (19 U/ml) of a *Rhizopus oryzae* lipase than induction of the *FLD1* promoter with methanol and ammonium (12 U/ml).

ICL1 promoter

Recently, the *ICL1* (isocitrate lyase) gene of *P. pastoris* was investigated as an alternative promoter to the commonly used *AOX1* and *GAP* promoters [72]. A plasmid (pPICLDEX) containing a single copy of the gene encoding dextranase from *Penicillium minioluteum* was transformed into *P. pastoris*. This gene was under the control of the *ICL1* promoter. The authors reported that high levels of dextranase were produced in cultures containing low levels of glucose or with ethanol as the sole carbon source; however, more dextranase was produced in the absence of glucose than after induction with ethanol. No actual figures were offered on the exact levels of expression obtained using *ICL1* as an alternative promoter and further work will need to be carried out to determine its usefulness for heterologous protein production in *Pichia*.

Selectable markers

All *P. pastoris* expression strains are derived from NRRL-Y11430 (Northern Regional Research Laboratories, IL, USA), the wild-type strain. One or more auxotrophic mutants are often present in these strains, allowing for selection during transformation of strains containing the appropriate selectable marker gene. A number of selectable marker genes are known for the molecular genetic manipulation of *P. pastoris*, as represented in Table 2, and some of which are described in Cereghino and Cregg [18].

The genetic manipulation of *P. pastoris* for the production of various heterologous proteins is simplified by the use of a wide range of selectable markers and promoters. Choosing the correct markers and promoters is essential for obtaining a high productivity from this system and, as has been described here, is different for every heterologous protein. This demonstrates that a great deal of optimization at the molecular level may be required, as well as optimization of the protein production process.

Controlling proteolysis

The genetic background of a *Pichia* host strain can influence the level of transcription, translation efficiency, the secretory pathway, protein quality, plasmid stability and plasmid copy number [34]. This

Table 2. Selectable marker genes for use with the *Pichia pastoris* expression system

Gene	Origin	Function	Reference
<i>HIS4</i> (histidinol dehydrogenase gene)	<i>P. pastoris</i> or <i>S. cerevisiae</i>	Biosynthetic pathway gene	
<i>ARG4</i> (argininosuccinate lyase gene)	<i>S. cerevisiae</i>	Biosynthetic pathway gene	
<i>Zeo^R</i> (zeocin resistance gene)	<i>Streptoalloteichus hindustanus</i>	Confers resistance to the bleomycin-related drug	
Blasticidin S deaminase gene	<i>Aspergillus tereris</i>	Confers resistance to the drug blasticidin	19
<i>ADE1</i> –PR-amidoimidazolesuccinocarboximide synthase	<i>P. pastoris</i>	Biosynthetic pathway gene	19
<i>URA3</i> –orotidine 5'-phosphate decarboxylase	<i>P. pastoris</i>	Biosynthetic pathway gene	19
<i>Sor^R</i> –acetyl-CoA carboxylase	<i>Sorangium cellulosum</i>	Confers resistance to the macrocyclic polyketide soraphen A	115

is best illustrated by the use of protease-deficient strains to improve the quality and yields of various heterologous proteins [111].

In general, the secreted recombinant proteins can potentially be proteolytically degraded in the culture medium by extracellular proteases, cell-bound proteases [57] and/or by intracellular proteases from lysed cells. *P. pastoris* extracellular proteases are not well documented, however, and this yeast reportedly secretes only low levels of endogenous proteins [18,53].

Several problems due to proteolysis can be foreseen in the production of recombinant proteins: (a) reduction of product yield when the product is degraded; (b) loss of biological activity when the product is truncated; and (c) contamination of the product by degradation intermediates in downstream processing because of their similar physicochemical and/or affinity characteristics [53].

Several strategies can be employed to control proteolysis in *P. pastoris*. These are based on modification at the cultivation, cell, and recombinant protein level.

Cultivation-level strategies

Cultivation techniques can influence proteolysis of recombinant proteins. *P. pastoris* is capable of growing across a relatively broad pH range (3.0–7.0). This range does not affect the growth significantly, which allows considerable freedom in adjusting the pH to one that is not optimal for a problem protease. Different pH values were found to be optimal from the point of view of a recombinant protein's stability: pH 6.0 was optimal in production of recombinant mouse epidermal factor and human serum albumin [25,59] and pH 3.0 was optimal in production of insulin-like growth factor-I and cytokine growth-blocking peptide (50 mg/l) [13,60].

The product stability is further enhanced by addition of amino acid-rich supplements (e.g. peptone, casamino acids) to the culture medium, possibly by acting as alternative and competing substrates for one or more problem proteases, and these supplements can also repress protease induction caused by nitrogen limitation [25,99,103,118].

Lower cultivation temperature can also influence yields of recombinant protein, possibly due to poor stability of the recombinant protein at higher temperatures, release of more proteases from dead cells, and folding problems at higher temperatures [48]. Li and co-workers [64] have shown that lowering the process temperature (from 30 °C to 23 °C) increased the yield of herring antifreeze proteins from 5.3 mg/l to 18.0 mg/l, and also increased cell viability. Hong *et al.* [48] achieved higher laccase activity by decreasing the cultivation temperature (from 30 °C to 20 °C) and by reducing methanol concentration (from 1.0% to 0.5%). Jahic *et al.* [54] obtained much higher concentration of a fusion protein by applying a temperature-limited fed-batch (TLFB) technique compared with the traditional methanol fed-batch technique. In TLFB, the common methanol limitation is replaced by temperature limitation, in order to avoid oxygen limitation at high cell density. A lower cell death rate was obtained in the TLFB process, which correlated with a lower protease activity in the culture supernatant, due to lower temperature and higher AOX activity.

The specific growth rate has also been used to reduce proteolysis. This was controlled by the addition of excess methanol (2–10 g/l). Proteolytic degradation of hirudin was greatly reduced when the specific growth rate was kept below the maximum, and was particularly obvious when the specific growth rate was maintained at 0.02–0.047 h⁻¹. This corresponded with a methanol concentration of 3.09 g/l, i.e. methanol was kept at growth rate-limiting quantities [133].

Addition of specific protease inhibitors to the culture medium may also be an option. Shi *et al.* [98] identified three types of proteases present in *P. pastoris* culture expressing a single-chain antibody (scFv) targeted against *Mamestra configurata* serpins. These were aspartic, cysteine and serine-type proteases. Total protease activity was reduced by 53% when a serine protease inhibitor was added to the culture medium and by 30% when an aspartic protease inhibitor was used. However, on an industrial scale, use of specific protease inhibitors could prove to be cost-prohibitive.

The combination of one or all of these cultivation-level strategies would prove an effective means for obtaining intact recombinant proteins from *P. pastoris* with minimum proteolytic degradation. The work described in this section demonstrates the necessity for process optimization at all levels, not only to achieve maximum growth and protein production, but also to obtain maximal amounts of the desired protein in an intact and bioactive form.

Cell-level strategies

The use of protease-deficient strains, such as SMD1163 (*his4 pep4 prb1*), SMD1165 (*his4 pep4*), and SMD1168 (*his4 pep4*), has been found to enhance the yield and the quality of various heterologous proteins [18]. These strains have a disruption in the genes encoding proteinase A (*PEP4*) and/or proteinase B (*PRB1*) [104]. Proteinase A is a vacuolar aspartyl protease necessary for the activation of vacuolar proteases, such as carboxypeptidase Y and proteinase B. Proteinase B has about half the activity of the processed enzyme before being activated by proteinase A. Therefore, *pep4* mutants eliminate the activity of proteinase A and carboxypeptidase Y, and partially reduce proteinase B activity. The *prb1* mutants eliminated activity of proteinase B, whereas *pep4 prb1* double mutants showed a significant reduction or elimination of all three of these protease activities [54].

In fermenter cultures of *P. pastoris*, a relatively high concentration of these vacuolar proteases is present in the culture broth, resulting from the combination of high cell densities and lysis of a small percentage of cells [18].

These protease-deficient strains, combined with other strategies to reduce proteolysis, have been invaluable in the production of insulin-like growth factor-I [13], ghilanten [10] and laccase [56].

Cereghino and Cregg [18] reported that protease-deficient strains are not as vigorous as wild-type strains with respect to *PEP4*, that they have lower viability, exhibit lower specific growth rates and are more difficult to transform, although no evidence was given to support this view. Brankamp *et al.* [10] observed higher secretion levels of a recombinant ghilanten and higher cell growth rates using a protease-deficient strain (SMD1168) than those obtained using a 'protease-normal' strain. However, since the protease-normal strain had a slow methanol utilization phenotype (Mut^s), it can not be concluded that the protease-deficient strain performed any better with regard to recombinant protein and biomass production levels than if the comparison had been made with a methanol plus phenotype (Mut⁺) strain. Comparison of protease-deficient and wild-type strains for the production of herring antifreeze protein demonstrated that, while the wild-type strain began product secretion a day earlier than the protease-deficient strain, a higher yield of intact product was obtained after purification using the protease-deficient strain (12.8% as opposed to 4% from the wild-type) [64]. This was probably due to the lack of proteolytic degradation of the product using the protease-deficient strain. It is difficult to ascertain the usefulness of protease-deficient strains for the production of foreign proteins from the results summarized above, and they should perhaps only be used, as Cereghino and Cregg [18] suggested, when all other methods for reducing proteolysis have been exhausted.

Protein-level strategies

If a linker between the domains of a fusion protein contains an amino acid sequence recognized by native proteases, it could be particularly sensitive to degradation. Accordingly, the amino acid sequence can be deleted if it is not essential for the function of the protein. Gustavsson and co-workers [41] designed stable linker peptides for a cellulose-binding domain lipase fusion protein in order to decrease proteolysis. The fusion product was produced at levels of around 10 mg l and the activity of the cellulose binding domain was found to be lower when the linker peptide chain was shorter than when the linker was longer. Paradoxically, those fusion products with shorter linker lengths were less susceptible to protease activity, and so a trade-off between high product activity and high protease activity on the one hand, and lower product activity coupled with lower protease activity on the other, would have to be made if this particular approach were to be considered.

Gene dosage

Gene copy number has been identified as a 'rate-limiting' step in the production of recombinant proteins from *P. pastoris* [25]. Increasing the number of copies of the expression cassette generally has the effect of increasing the amount of protein expressed [25,90,112].

Pichia cells expressing hepatitis B surface antigen (HbsAg) under the control of the constitutive GAP promoter were identified as undergoing spontaneous multiple gene insertions at a single locus. Multicopy clones containing up to four copies of the *HbsAg* gene were found to have a four-fold higher yield of HbsAg than those containing a single copy of the gene. No limitation at the transcriptional level was identified [112]. Thus, for this construct the relationship between copy number and product level was very simple.

However, increased copy number might reasonably be expected to exert a knock-on effect on transcription and translation, both of which may become rate-limiting due to a lack of resources, such as precursors and energy [47], as has been previously reported for recombinant protein production in *E. coli* [62,93]. However, in *Pichia* it has been proposed that it is more likely that any limitations are due to post-translational events, such as folding within the endoplasmic reticulum (ER), membrane translocation and signal sequence processing [47].

Sunga and Cregg [107] also demonstrated that increasing the copy number of a β -galactosidase (β -gal) gene (*lacZ*), under the control of the FLD1 promoter, increased the relative activity of the enzyme proportionately. They reported a 17-fold increase in activity of β -gal when 22 copies of the *lacZ* gene were present, relative to the activity observed when a single copy of the gene was present.

In contrast to the work described above, where various researchers found a direct correlation between gene copy number and subsequent yield and activity of the gene product, Hohenblum *et al.* [47] did not note any increase in the expression of a recombinant human trypsinogen when the gene dosage was increased from one to three copies under the control of the GAP promoter. However, an increase in the expression level of this protein was observed using the AOX1 promoter when the gene copy number was increased to two, but levels of expression fell upon further copy number increases. A number of researchers have had significant success with increasing gene dosage utilizing both the AOX1 and GAP promoters [25,70,112], therefore, before considering increasing the gene copy number as an optimization strategy for recombinant protein production from *P. pastoris*, the identity of the promoter must be considered in advance.

Post-translational modification of secreted proteins

P. pastoris, unlike bacterial expression systems, has the ability to perform many of the post-translational modifications usually performed in higher eukaryotes, e.g. correct folding, disulphide bond formation, O-

and N-linked glycosylation and processing of signal sequences. Folding and disulphide bond formation have been identified in some cases as the 'rate-limiting' step in the production of foreign proteins from *Pichia* [47], i.e. the ability of the organism to process, fold and secrete the recombinant products determines the productivity of the *Pichia* expression system. It is therefore difficult to optimize production of recombinant proteins unless these rate-limiting factors can be identified. Some of the post-translational modification options are discussed below.

Secretion signals and disulphide bond formation

P. pastoris can produce expressed foreign proteins either intracellularly or extracellularly. Extracellular production of foreign proteins is most desirable in order to avoid the usual first steps of purification, e.g. cell lysis to release the cellular contents, followed by clarification to remove cell debris. *Pichia* also secrete very low levels of native proteins, thus making it easier to recover the foreign secreted protein from the fermentation fluid by simple removal of whole cells by filtration or centrifugation. Moreover, secretion signals can be attached to the protein of interest, causing it to be exported out of the cell. Foreign genes may be cloned in *P. pastoris* vectors to align them in the correct reading frame with either the native secretion signal for the protein of interest, the *S. cerevisiae* α -factor prepro-peptide, or the *P. pastoris* acid phosphatase (*PHO1*) signal [18]. The *S. cerevisiae* α -factor prepro-signal is the most widely used and successful secretion signal, being in some cases better than the leader sequence of the native heterologous protein. However, variability in the number of N-terminal amino acids is commonly reported with heterologous proteins secreted using the α -factor prepro-leader [14].

In *P. pastoris* cells producing maize Mir1 cysteine proteinase, the protein was not detected in the culture medium or in the soluble fraction of the cell lysate. It was detected in the membrane fraction, where the protein was retained after secretion was initiated. The researchers concluded that this occurred due to the presence of the yeast *PHO1* signal sequence at the protein's modified N-terminal [84].

In other cases where the standard α -factor or *PHO1* secretion signals have not worked, new signal sequences had to be used, such as the leader sequence of the *Pichia acaciae* killer toxin and the phytohaemagglutinin signal sequence (*PHA-E*) from *Phaseolus vulgaris* [87]. Unfortunately, trial and error experiments are often required to find the optimum secretion signal for a specific protein [21]. In proteins such as the hepatitis B surface antigen, which has multiple transmembrane domains, and is likely to be sequestered in the membrane, the use of the α -factor secretion signal did not significantly improve recovery [112].

The *P. pastoris* expression system has been successfully used to produce proteins that are highly disulphide-bonded [119]. Prokaryotic systems have been generally unsuccessful in achieving this, due to the reducing environment of the cytoplasm, resulting in the necessity to refold disulphide-bonded proteins from inclusion bodies, or to secrete the proteins into the periplasmic space. Among those proteins with high disulphide-bonding that have been produced in *P. pastoris* are a fragment of thrombomodulin containing two epidermal growth factor-like domains, and coagulation protease (Factor XII) and some of its amino-terminal domains. The presence of disulphide bonds may have an effect on the binding activity of certain proteins, such as juvenile hormone binding protein, and thus the positions of these were elucidated using mass spectrometry (Cys¹⁰–Cys¹⁷ and Cys¹⁵¹–Cys¹⁹⁵) [32].

O- and N-linked glycosylation

Glycosylation is one of the most common post-translational modifications performed by *P. pastoris*. It is also one of the more complex [34]. It is thought that since many mammalian native proteins are glycosylated, it must be necessary to have the correct glycosylation patterns on recombinant proteins to ensure their biological activity. In addition, the glycosylated gene products generally have much shorter glycosyl chains than those expressed in *S. cerevisiae*, thus making *P. pastoris* a much more attractive host for the expression of recombinant proteins [12].

Pichia, like other yeasts and fungi, add O-oligosaccharides to the hydroxyl groups of serine and threonine of secreted proteins. These are composed of mannose residues only, whereas higher eukaryotes,

such as mammals, have a more varied sugar composition in these oligosaccharides. It is possible that *Pichia* will glycosylate heterologous proteins, even when those proteins are not normally glycosylated by the native host; and even when the protein is glycosylated in the native host, *Pichia* may not glycosylate it on the same serine and threonine residues [18].

The extent of *O*-glycosylation by *P. pastoris* has been studied in a glucoamylase catalytic domain from *Aspergillus awamori* [44]. The molecular weight of the secreted protein was 20 kDa heavier than the native protein. About 10 kDa of this weight could be attributed to *N*-glycosylation, meaning that the rest could be attributed to *O*-linked glycosides, probably consisting of 20–30 mannose residues. It was concluded from this study that the extent of glycosylation of proteins by *P. pastoris* was substantially less than that by *S. cerevisiae*, but that the extent of glycosylation of different proteins secreted by *P. pastoris* differed greatly. In the production of recombinant human antithrombin III from *P. pastoris*, it was noted that *O*-glycosylation occurred near the reactive site and resulted in the recombinant protein having half the inhibitory activity against thrombin when compared with the native antithrombin III [75]. The occurrence of glycosylation in this case may have actually changed the function of the recombinant protein, where the native enzyme is normally responsible for inactivating thrombin, the recombinant may act as a substrate for thrombin.

N-linked glycosylation, in eukaryotes, begins on the cytoplasmic side of the endoplasmic reticulum (ER) with the generation of a branched heptasaccharide intermediate. This is translocated across the ER, initiating the second phase of the synthesis process [45]. There are three classes of *N*-linked glycans that are synthesized by the co-translational addition of Glc₃Man₉GlcNAc₂ to the polypeptide chain by oligosaccharyl transferase; these are composed of Man_{5–6}GlcNAc₂ (high mannose), a mixture of several different sugars (complex) or a combination of both (hybrid), and are represented in Figure 2 [39]. Addition and reduction reactions as the carbohydrate structure passes through the ER and Golgi apparatus results in the final glycan structure, which will incorporate one of the three classes of *N*-linked carbohydrates described above. However, in fungi and yeasts, such as *Pichia*, the outer oligosaccharide chain of secreted proteins is mostly unaltered and consists of Man_{8–9}GlcNAc₂ [76]. Despite the differences observed in the amount and complexity of *N*-linked glycosylation, there is evidence for a single consensus recognition sequence for initiation of glycosylation, i.e. the sequon Asn–Xaa–Ser/Thr, which has been found to be necessary for *N*-linked glycosylation, but is not always sufficient [22]. The *PNO1* (phosphomannosylation of *N*-linked oligosaccharides) gene was cloned and was found to

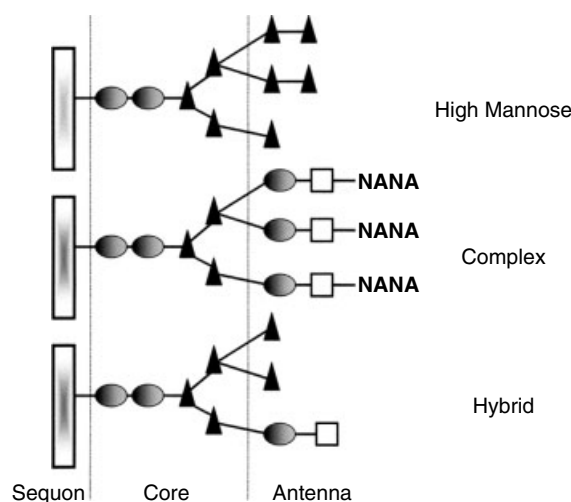


Figure 2. Protein glycosylation patterns. Oligosaccharide chains consist of a sequon, which is attached to the core by an Asn residue. The core is attached to the antenna, which can be high-mannose, complex or hybrid. GlcNAc, ●; mannose, ▲; galactose, □

promote *N*-linked glycosylation only to the core oligosaccharides and not to the outer sugar chain, unlike the *MNN4* gene of *S. cerevisiae* [74]. This would explain why proteins expressed in *P. pastoris* have shorter glycosylation chain lengths than those expressed in *S. cerevisiae*.

Humanizing the glycosylation patterns of foreign proteins secreted by *P. pastoris* is desirable, since the extent and positioning of the glycosides may affect the activity of the protein. This can be achieved through specific engineering of the strains [21]. In a *P. pastoris* strain producing human midkine, a heparin-binding protein, three of the threonine residues involved in *O*-glycosylation were replaced with three alanine residues by site-specific mutagenesis. The resultant human midkine recombinant protein contained no mannose residues and promoted proliferation of CHO cells as well as the native protein, despite the alteration to the amino acid sequence [2].

Pichia-derived glycosylated proteins have the potential to trigger inappropriate immune responses if used as pharmaceuticals. The immunogenicity of *Pichia*-derived proteins is an issue that has attracted interest in the literature with regard to humanizing the glycosylation patterns [2,24]. One group of researchers disrupted the *PNO1* gene of *P. pastoris*, resulting in a reduction of *N*-linked glycosylation of a recombinant human antithrombin. They reported that the glycosylation had fallen from around 20% in the wild-type strain to less than 1% in the engineered strain [74], showing that suppression of *N*-linked glycosylation is possible.

However, these methods cannot guarantee that the products' glycosylation pattern would be sufficiently humanized in order to utilize these proteins as therapeutic pharmaceuticals in humans. One group of researchers has shown that it is possible to engineer the glycosylation pathway in *P. pastoris* in order to obtain recombinant human proteins that have uniform, complex *N*-glycosylation. This was achieved by removing the host's endogenous glycosylation pathways and localizing five key eukaryotic enzymes to obtain an *in vivo*, synthetic glycosylation pathway, which produced a complex human *N*-glycan [5,42]. This important research into humanizing glycoproteins produced in *P. pastoris* has the potential to elucidate the structure–function relationships of *N*-glycans and thus accelerate their use as human therapeutics.

This ability to engineer *Pichia* cells is possible due to the increased knowledge available about the way in which these complex carbohydrate structures are added to the recombinant proteins, and to how these structures are related to the native proteins. This provides the opportunity to optimize the host-cell background in order to produce proteins with the desired carbohydrate structures. However, it is not possible to generalize about the optimal glycosylation patterns, since each recombinant protein must be assessed, first in the context of the system in which it is expressed, and second in terms of the purpose for which it will be used.

Production of heterologous proteins in *P. pastoris*

The *Pichia* expression system has been widely used to produce a variety of different heterologous proteins, many of which are listed by Cereghino and Cregg [18]. Since 2002 there has been a great increase in foreign protein production in *Pichia*, which is briefly summarized in Table 3. Cell growth is particularly important for secreted protein production in bioreactors, since the concentration of the product in the extracellular medium is roughly proportional to the concentration of cells in the culture in many instances. The yields of proteins expressed intracellularly in bioreactors are also high, due to the efficiency of the *AOX1* promoter [18]. Production of large amounts of heterologous proteins in shake-flask culture is difficult, due to the limitations of volume, oxygen transfer, substrate addition and an inability to monitor these factors efficiently. Methods of monitoring methanol in shake flasks are available, such as organic solvent vapour detectors [40], and the oxygen transfer of these systems can be increased by the use of baffles [114]. On-line measurement of oxygen and carbon dioxide transfer rates in shake flasks has also been demonstrated using a respiration activity monitoring system (RAMOS) [100]. However, the use of bioreactors is preferable, since all of these parameters can be monitored and controlled simultaneously, allowing more efficient production of the desired heterologous protein. The use of bioreactors for the

Table 3. Summary of a range of foreign proteins produced using the *Pichia* expression system since 2002

Protein expressed	Function	Expression level	Reference
Bacteria			
<i>Clostridium botulinum</i> serotype F [BoNTF(Hc)] heavy chain fragment C	Antigen	205 mg/kg of cells	55
<i>Escherichia coli</i> AppA	Use in the animal feed industry to release inorganic phosphate	117 ± 1 U/mL	105
<i>Escherichia coli</i> phytase	Use in animal feed industry to improve phosphorus utilization	6.4 g/L	23
<i>Gluconoacetobacter diazotrophicus</i> exo-levanase	Fructose-releasing potential for high-fructose syrup production	0.46 g/L	73
<i>Thermus aquaticus</i> YT-I aqualysin I	Heat-stable subtilisin-type serine protease	1 g/L	81
Fungi			
<i>Aspergillus oryzae</i> tannase	Hydrolyses the ester and depside bonds of gallotannins and gallic acid esters	72 mg/L	132
<i>Candida antarctica</i> CBM–CALB fusion protein	Lypase	1.3 g/L	54
<i>Candida parapsilosis</i> lipase/acyltransferase	Catalyses alcoholysis of esters	5.8 g/L	15
<i>Rhizopus oryzae</i> lipase	Catalyses breakdown of triglycerides	19 U/mL	89
<i>Trametes versicolor</i> cellobiose dehydrogenase (CDH)	Oxidation of reducing-end groups of cellobiose	225 U/mL	106
<i>Trametes versicolor</i> laccase (lccI)	Phenoloxidase	23.9 U/mL	79
<i>Trigonopsis variabilis</i> D-amino acid oxidase (DAAO)	Flavoprotein, oxidoreductase	1.46 U/mg	129
Invertebrates			
<i>Boophilus microplus</i> BM95 antigen	Protects cattle from <i>B. microplus</i> infestation	550.0 mg/L	7
<i>Pandalus borealis</i> (shrimp) cathepsin L	Cysteine proteinase	60.0 mg/L	1
Vertebrates (non-human)			
<i>Bothrops atrox</i> moojeni venom batroxobin	Cleaves fibrinogen α -chain	7 mg/L	128
Bovine enterokinase light chain (EK _L)	Duodenal serine proteinase	350.0 mg/L	85
Equine lactoferrin (ELF)	Iron-binding protein, family of transferrins	40.0 mg/L	83
Hen egg H5-lysozyme	Bactericidal effects	2.1 mg/L	66
Hen egg white ovotransferrin	Iron-binding protein, family of transferrins	57.0 mg/L	63
Mouse endostatin	Tumour growth suppressor	133.33 mg/L	110
Ovine interferon- τ (r- α IFN- τ)	Antiviral and antiproliferative activity	391.7 mg/L	101
Porcine lactoferrin (PLF)	Iron-binding glycoprotein, family of transferrins	12.0 mg/L	116
Humans			
μ -Opioid receptor (HuMOR)	G-protein coupled receptor (GPCR)	4.0 pmol/mg	94
μ -Opioid receptor fused to green fluorescent protein	G-protein coupled receptor (GPCR)	1250 pmol/l	86
Angiostatin	Used in long-term therapy in suppression of metastases	108.0 mg/L	124
Anti-HBs Fab fragment	Prevention and treatment of hepatitis B virus	50.0 mg/L	77
DNA topoisomerase I (hTopo I)	Role in DNA replication, transcription, recombination, chromosome assembly	11–58 mg/L	126
Dopamine D2S receptor	G-protein	3–13 pmol/mg	33
Granulocyte-macrophage colony-stimulating factor (hGM-CSF)	Regulation of activity of mature, differentiated granulocytes and macrophages	180.0 mg/L	123
Serine protease inhibitor SERpin CI-INhibitor (CI-Inh)	Plasma glycoprotein, inhibitor of inflammatory reactions	180.0 mg/L	6
Serum albumin	Binding and transport, colloid osmotic pressure	13.4 mg/g	16
Serum amyloid P (SAP)	G-protein, inhibition of DNA auto-antibody formation	12.0 mg/L	8
Transforming growth factor- β type II receptor extracellular domain (SolRII)	TGF- β antagonist, regulates cell proliferation and differentiation	20.0 mg/L	95
Plants			
Amaryllidaceae snowdrop agglutinin	Binding of α -D-mannose groups	80.0 mg/L	3
<i>Arabidopsis thaliana</i> core α 1,3-fucosyltransferase	Biosynthesis of N- and O-linked oligosaccharides	1.0 U/L	4
Gsp (Panax ginseng C, medicinal peptide)	Potential use as drug against diabetes	800.0 mg/L	125
Maize cysteine proteinase (MirI)	Seed germination, senescence	Insoluble/inactive	84
Rice α -amylase (Amy 1A/3D)	Hydrolysis of starch	340.0 mg/L	63
Rice β -D-fructofuranosidase (Os β fruct3)	Hydrolysis of sucrose	0.35 U/L	37
<i>Solanum tuberosum</i> (potato apyrase)	Phosphohydrolase, platelet-aggregation inhibition	1.0 mg/L	78
<i>Thaumatococcus daniellii</i> thaumatin II	Sweet-tasting protein	25 mg/L	69
Viruses			
Measles virus nucleoprotein (MeN)	Responsible for viral genome packaging	29% of total protein	102

production of recombinant proteins from *P. pastoris* is reviewed by Cereghino *et al.* [21]; in this section the importance of medium composition and control in these processes is discussed.

Nutrient considerations

The majority of publications in the field of recombinant protein production from *P. pastoris* use complex media for both growth and induction of protein expression. The media formulations are published by Invitrogen Corporation (Carlsbad, CA, USA), and are recommended for use with the strains they provide. While these media are excellent for providing the correct chemical environment required for *Pichia* cell growth, they have some limitations when considering larger-scale and industrial fermentations. One of the common limitations is the initial concentration of substrates (glucose or glycerol) used. While it is important to maintain the initial substrate at low non-inhibitory concentrations, it is important to remember that the production of foreign proteins in *P. pastoris* is growth-related, and it is therefore essential to optimize biomass production in order to obtain large quantities of the desired proteins. This can be achieved by increasing the initial substrate concentration, or by employing a feeding regime whereby the biomass can be increased dramatically before induction of protein production is initiated. The use of yeast extracts and peptones in complex media means that batch-to-batch variations would occur. It is therefore desirable to eliminate any complex components from the medium, making it easier to standardize the production process and to validate the medium and process itself. The Invitrogen formulations are, however, useful in the initial studies, where the laboratory researcher can follow the instructions and get reproducible results, thus making it easier to screen strains and compare expression levels of the desired protein. However, to produce large quantities of the heterologous protein, either for characterization or structural studies, or for manufacture, requires the use of a defined medium, as well as a basic understanding of process development in order to manipulate the physicochemical environment to maximize growth and protein production.

A common formulation for a defined (basal) medium for *P. pastoris* is listed in Table 4 with a formulation for trace salts solution in Table 5 [16]. Boze and co-workers [9] found that supplementation of a basal medium with seven vitamins and two trace elements increased production of a recombinant porcine follicle-stimulating hormone from 93 mg/l (basal) to 187 mg/l (supplemented), showing that vitamin and trace element requirements have an important effect on cell growth and recombinant protein production in *P. pastoris*.

The nitrogen source is usually provided by the addition of ammonium hydroxide, which also has the effect of controlling the pH to the desired level. Increased concentrations of ammonium in the medium can prolong the lag phase and thus inhibit cell growth, especially at concentrations of 0.6 M and above [127]. Low ammonium ion concentrations have been shown to encourage degradation of a recombinant hirudin, while higher concentrations ($> 0.4 \text{ M NH}_4^+$) inhibited hirudin production [127]. However, other research has shown that ammonium ion concentrations above 0.1 M can have inhibitory effects on cell growth and protein production [124]. From these differing reports, it can be concluded that the ammonium ion requirements and effects on protein production should be thoroughly investigated

Table 4. Defined medium composition for *P. pastoris* fermentations [16]

Constituent	Quantity (g/L)
Glycerol	40.0
CaSO ₄	0.93
K ₂ SO ₄	18.2
MgSO ₄ ·7H ₂ O	14.9
KOH	4.13
K ₂ HPO ₄	7.0
+22.7 ml 85% H ₃ PO ₄	+12 ml trace solution

Table 5. Trace salts solution [16]*

Constituent	Quantity (g/L)
CuSO ₄ .5H ₂ O	6.0
NaI	0.08
MnSO ₄ .H ₂ O	3.0
Na ₂ MoO ₄ .2H ₂ O	0.2
H ₃ BO ₃	0.02
CoCl ₂	0.5
ZnCl ₂	20.0
FeSO ₄ .7H ₂ O	65.0
Biotin	0.2

* +5 ml conc. H₂SO₄.

as part of the optimization process for each individual case. Addition of adequate ammonium ions to the growth medium was responsible for increasing the yield of a recombinant angiostatin by around 2.8-fold to 108 mg/l, as opposed to 39 mg/l in the absence of ammonium [124].

Other nitrogen sources affect recombinant protein production in *P. pastoris*. The use of yeast extract and casamino acids increased protein secretion and accumulation [104], and L-arginine and EDTA increased accumulation of a single-chain antibody (scFv) three- to five-fold, although no expression levels were mentioned specifically [98].

The use of ammonium ions as a nitrogen source during recombinant protein production in *P. pastoris* is probably the most commonly used method of overcoming nitrogen-limitation problems. However, it is clear from the examples described here that careful consideration of the nitrogen source, as well as the concentration ranges utilized, is advisable when optimizing the production of proteins in this way.

The carbon source used may also affect the amount of recombinant protein retained within the cells, even when secretion signals have been used [47]. Hohenblum *et al.* [47] constructed a clone containing the GAP promoter to investigate the production of a recombinant trypsinogen in *P. pastoris*. Their research indicated that higher amounts of intracellular trypsinogen were observed when cells containing the clone were grown on glucose than in cells grown on methanol. When these results were compared with those obtained for a clone containing the AOX1 promoter and expressing trypsinogen, the authors concluded that retention of heterologous proteins intracellularly was carbon source-dependent and not promoter-dependent.

Non-expression-repressing carbon sources, such as alanine, sorbitol and mannitol, have also been investigated for use with Mut⁻ *P. pastoris* strains [52]. Each of these carbon sources was shown to increase the production of a recombinant β -galactosidase as compared to cells grown with glycerol or glucose, as well as reduce the amount of methanol required for induction of protein expression. This could be beneficial in the industrial production of recombinant proteins, meaning that smaller quantities of methanol would be required, thus reducing the hazards of storage of large amounts of methanol, as well as reducing costs.

Process control

Protein levels in fermenter culture are typically much higher than in shake-flask cultures, although this may be due to poor optimization of methanol-feeding strategies in shake-flask cultures [40]. For this reason, scale-up from shake flask to bioreactor requires much more reoptimization than scale-up from small to large bioreactors [21]. The medium components used in *P. pastoris* processes (glycerol, methanol, biotin, salts, trace elements) are economical and well defined and, as such, are almost ideal for large-scale production of heterologous proteins in bioreactors. Fermentative carbon sources, such as glucose, tend to be avoided in favour of the non-fermentative carbon sources, such as glycerol. This is because the by-product ethanol has been found to repress the AOX promoter, even at levels of around 10–50 mg/l

[51]. Monitoring methanol in a *Pichia* process is extremely important, since high levels of methanol (above 5 g/l) can be toxic to the cells [130] and low levels of methanol may not be enough to initiate transcription [18].

Methods for monitoring methanol can be problematic. For example, using a biochemistry analyser to determine methanol concentrations by means of enzymatic reactions of alcohol oxidase, or using gas chromatography, involves processing of the sample before presentation to the analyser, which is usually located some distance from the bioreactor. These methods take time and can be expensive, as well as running the risk of allowing the methanol to evaporate before the concentration can be determined. So, an accurate picture of what is happening within the bioreactor at any given time during the process cannot be obtained. On-line analysis of methanol concentrations in shake flasks has been demonstrated [40,134] using sensors that detect methanol vapour; however, sequential injection analysis (SIA) has been used for on-line methanol determination in a small bioreactor with 1 l working volume [108]. The SIA analyser worked by removing a sample from the fermenter automatically, treating it and then subjecting the sample to enzymatic analysis. In this way, four analyses/h could be performed, and the feed rate of methanol could be adjusted accordingly; however, an in-line dilution step was required to keep the methanol concentration within the range of the analyser (up to 2 g/l).

Other physical parameters within the bioreactor have an effect on the expression levels of foreign proteins in *P. pastoris*. The yield of a recombinant galactose oxidase from *Dactylium dendroides*, expressed in *P. pastoris*, was sensitive to culture conditions such as process pH [121]. Galactose oxidase is unstable at pH values lower than pH 6.0. Yield of this protein was therefore improved by increasing the process pH to 6.0 during induction with methanol. Process temperature was also found to be an important factor in yield and activity of this protein. Around four times as much galactose oxidase was produced when the process temperature was decreased from 30°C to 25°C during methanol induction, giving a final concentration of 500 mg/l.

Monitoring of methanol, as well as glycerol and heterologous protein production, is necessary to obtain the maximum amount of product possible. This was achieved using Fourier transform mid-infrared spectroscopy (FT-MIR) [30], and information about all of the medium components, including biomass, was generated simultaneously. Mid-infrared spectroscopy is particularly suited to *Pichia* processes, since low concentrations of heterologous protein can be detected, typically less than 1 g/l, and, since methanol levels within these processes must be strictly controlled to induce and maintain protein production, this near-real-time analysis technique improves control and optimization of any *Pichia* process.

Additionally, in a minimal medium typically containing no complex organic N sources such as peptides, it is potentially possible to derive information about the folding characteristics of the secreted protein using FT-MIR. Thus, the potential relevance of such spectroscopic techniques to protein optimization from *Pichia* is considerable.

Given the influence optimization of the physicochemical environment in the bioreactor has upon levels of protein expression, it is vitally important when attempting to make a desired protein using *Pichia* to consider the whole process (genetic alteration and physiological manipulation) as an integral. Thus, the genetic construction of the strain should correctly be looked upon as conferring a potential for protein production upon a strain that is only realizable with the appropriate handling at the physiological level, via close regulation of the physicochemical environment.

In essence, this requirement for an integral approach poses a challenge for the biotechnological community, best met by collaboration between groups having complimentary skills in the two above areas.

Future aspects

The *Pichia* expression system has been widely used for the production of soluble proteins, which are easily secreted into the liquid phase of the bioprocess, (many of these are detailed in Table 3) and are easily recovered from the liquid phase in an active form using conventional downstream processes. However, the current challenge in using this expression system is the production and recovery of the largely

hydrophobic membrane proteins. Many membrane proteins have specific lipid and sterol requirements, and these lipids can be involved in the folding and stability of the protein as well as in their function [82]. The lipid content of various host cells is presented by Opekarová & Tanner [82], who suggest that one of the obstacles for producing mammalian membrane proteins in *P. pastoris* could be a shortage of cholesterol. *Pichia* cells might replace this and other plant and animal sterols with ergosterol, the main fungal sterol; however, this could affect the functionality of any membrane proteins produced using this system. Membrane proteins have therefore been difficult to overexpress using *P. pastoris* and to obtain in an active form, making it extremely difficult to perform any structural or functional studies [67].

Ruf and co-workers [92] recently described a process for the production of a human oxidosqualene cyclase, a monotopic membrane protein which catalyses the formation of lanosterol in mammals, in *P. pastoris*. As with the majority of membrane proteins, it had been difficult to obtain large enough quantities of the protein in order to perform biochemical and structural studies, the goal of this research being to identify inhibitors of oxidosqualene cyclase and thus reduce cholesterol production. The authors discovered that the protein of interest had not been altered post-translationally and had no glycosylation, making it almost identical to the native human protein. Additionally, it had been produced in reasonable quantities (3 mg/g cell mass), enough for the purpose of structural studies. Other attempts at producing mammalian membrane proteins [109] in *P. pastoris* have not had the same success. The rat serotonin transporter, rSERT, was found to be misfolded and glycosylated when produced in a protease-deficient *P. pastoris* strain, as well as having a very low binding activity.

Non-mammalian membrane proteins have also been expressed successfully in *P. pastoris*, e.g. the SARS-CoV M protein was produced at a relatively high concentration of 6 mg/L and was suspected to have undergone glycosylation, which did not affect the protein's use in the detection of antibodies, but may not have been identical to the native virus protein [43]; and the aquaporin PM28A from spinach leaf plasma membranes, which was produced at 25 mg/L and was shown to be active and correctly folded despite the addition of extra amino acids [58].

These examples demonstrate that the *P. pastoris* expression system can be used for the production of membrane proteins from a wide variety of organisms, to varying degrees of success. The challenge in using this expression system for this purpose is to select the most appropriate strain, selection marker, promoter, etc., as this can have an enormous effect on the success of producing full-length, correctly folded and correctly glycosylated membrane products [29].

G-protein-coupled receptors (GPCRs) are a superfamily of membrane proteins which contain seven transmembrane segments. They have the ability to transduce a variety of external signals (e.g. hormones, neurotransmitters, etc.) to the interior of the cell [88]. G-protein-mediated signalling and cell signalling through thromboxane A₂ receptors are reviewed by Offermans [80] and Huang *et al.* [49]. Since the effects of many pharmaceutical substances are mediated through GPCRs, these proteins have attracted a great deal of interest from the pharmaceutical industry [94] and could prove to be useful in the treatment of human diseases [120]. However, GPCRs, as with other membrane proteins, have very low natural levels of expression, resulting in little opportunity for structural and functional studies of these proteins, as well as ensuring that the cost of such studies remains prohibitive.

Various vector and host cell expression systems have been explored for the overexpression of GPCRs, including adenovirus/CHO cells for the production of a human and mouse galanin-receptor like GalRL [50] and human κ - and μ -opioid receptors [131]. However, production levels of these proteins were not quoted, making it difficult to ascertain how efficient these systems were for producing significant quantities of these membrane proteins. The Semliki Forest virus (SFV) has been used with a variety of mammalian host cells for the overexpression of a large number of GPCRs. These have been summarized by Lundström [68] in an informative table detailing production levels for some of these GPCRs in the range approximately 2–10 mg/L. The baculovirus/insect cell system has also been used to overexpress GPCRs with some success, e.g. Perret *et al.* [86] produced 1250 pmol/L of enhanced green fluorescent protein-amino-tagged human μ -opioid receptor. Bacterial expression systems such as *E. coli* are generally not used for the overexpression of mammalian GPCRs, first because the integration of these proteins into the bacterial cell membrane is often toxic to the cell, and second because *E. coli* is generally incapable of

the complicated post-translational modifications and processing required for the correct folding of such proteins [68]. For this reason, and because other expression systems utilizing mammalian or insect cells tend to be lengthy and expensive, research has turned to the *Pichia* expression system in the hope that this situation can be remedied.

The human μ -opioid receptor (HuMOR) has been overexpressed in *P. pastoris* [94]. Through classical ligand binding studies, it was revealed that 1 pmol/mg membrane protein was produced. However, the use of a green fluorescent reporter protein indicated that 16 pmol of the protein remained in the membrane fraction and that 100 pmol/mg of protein was present in the whole cells. These studies indicate that, while overall production of this GPCR increased, the ligand-binding studies threw some doubt over the functionality of the overexpressed protein. This may have been due to the specific lipid requirements of this GPCR [82], but this aspect was not considered by the authors. de Jong *et al.* [33] produced a recombinant human dopamine D2S receptor in *P. pastoris* and obtained variable results. Expression levels of this protein were in the range 3–13 pmol/mg total protein, and the D2S receptor appeared to be fully functional.

Recovery of GPCRs from a *Pichia* process is more complicated than the recovery of soluble proteins. This is mainly due to the insolubility and hydrophobicity of these proteins and because GPCRs tend to be found in the membrane fractions or in the whole cells, even when secretion signals have been used [33]. Whereas, with soluble proteins, recovery may involve just a few steps, such as centrifugation and one or two chromatographic steps, with membrane proteins/additional processing is necessary. This would normally involve a complicated membrane preparation step using glass beads rather than the harsher French press or sonication techniques to obtain intact protein molecules, followed by a solubilization step with non-ionic detergents, before the purification steps suggested above can be attempted.

Despite the variable results achieved with the production of GPCRs in *Pichia*, such as unreproducible production levels and low secretion levels of functional protein, and the relative difficulty of recovering GPCRs from a *Pichia* process, it is anticipated that scale-up of these processes to fermenter vessels will significantly improve the production levels and thus make their use for pharmaceutical research and applications more financially viable.

Conclusions

The use of *P. pastoris* for the production of heterologous proteins is highly effective. The wide range of promoters available, as well as selectable markers, secretion signals, methods for coping with proteases and a better understanding of glycosylation patterns, have given researchers diverse means to achieve the production of foreign proteins. This diversity has enabled the production of even complex membrane proteins, such as the G-protein-coupled receptors. Although the *Pichia* expression systems available are efficient and easy to use with well-defined process protocols, some degree of process optimization is required to achieve maximum production of the heterologous protein, as well as maximum activity of the protein. In fact, yield and activity are often dependent upon the physical parameters of the culture vessel, e.g. pH, temperature and O₂ availability, and they are dependent on the residual concentrations of methanol. These factors can all be closely monitored to ensure the exact conditions required.

Over the last few years, the numbers of publications detailing the use of this expression system for research purposes have increased rapidly. This has been reflected in the increase in use of this system for the industrial production of heterologous proteins, replacing more conventional methods, such as production in *E. coli*, *S. cerevisiae*, and extraction of native proteins from, for example, plasma. A common theme amongst these publications is the lack of detail about the actual production processes. The production levels for each protein are often omitted, or not stated in comparable units, e.g. many researchers publish the specific activity of their overexpressed protein, or give the production levels in bioluminescence units, without relating these figures to the productivity of their particular process. This makes it extremely difficult to draw comparisons between production of a particular protein using different expression systems, as well as between the work of different research groups.

Is *P. pastoris* a better expression system than *E. coli*, *Saccharomyces*, baculovirus/insect cells, SFV or adenovirus/mammalian host cells for the production of heterologous proteins? For soluble, secretable proteins, undoubtedly, it is the easiest and most reliable expression system with the potential to produce grams of the desired protein. For insoluble membrane proteins, it is difficult to draw a conclusion when there is so little information available about the efficacy of this system, and little understanding of how *Pichia*-derived membrane proteins differ from native proteins and exactly what these differences mean in terms of the use of the recombinant product. If the difficulties of producing membrane proteins such as the GPCRs can be addressed, i.e. low expression levels, lipid and sterol requirements, recovery and bioactivity, the *Pichia* expression system has the potential to produce membrane proteins as successfully as it does soluble proteins.

References

1. Aoki H, Ahsan MN, Watabe S. 2003. Heterologous expression in *Pichia pastoris* and single-step purification of a cysteine proteinase from northern shrimp. 2003. *Prot Exp Pur* **31**(2): 213–221.
2. Asami Y, Nagano H, Ikematsu S, Murasugi A. 2000. An approach to the removal of yeast specific *O*-linked oligomannoses from human midkine expressed in *Pichia pastoris* using site-specific mutagenesis. *J Biochem* **128**: 823–826.
3. Baumgartner P, Harper K, Raemaekers RJM, et al. 2003. Large-scale production and purification of recombinant *Galanthus nivalis* agglutinin (GNA) expressed in the methylotrophic yeast *Pichia pastoris*. *Biotechnol Lett* **25**: 1281–1285.
4. Bencurova M, Rendic D, Fabini G, et al. 2003. Expression of eukaryotic glycosyltransferases in the yeast *Pichia pastoris*. *Biochimie* **85**: 413–422.
5. Bobrowicz P, Davidson RC, Li H, et al. 2004. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. *Glycobiology* **14**(9): 757–766.
6. Bos IGA, de Bruin EC, Karuntu YA, et al. 2003. Recombinant human C1-inhibitor produced in *Pichia pastoris* has the same inhibitory capacity as plasma C1-inhibitor. *Biochim Biophys Acta* **368**64: 1–9.
7. Boue O, Farnos O, Gonzalez A, et al. 2004. Production and biochemical characterisation of the recombinant *Boophilus microplus* BM95 antigen from *Pichia pastoris*. *Exp Appl Acarol* **32**(1–2): 119–128.
8. Boysen S, Fogh-Schultz B, Andersen I, et al. 2004. recombinant human serum amyloid P component from *Pichia pastoris*: production and characterisation. *Prot Exp Pur* **35**(2): 284–292.
9. Boze H, Laborde C, Chemardin P, et al. 2001. High-level secretory production of recombinant porcine follicle-stimulating hormone by *Pichia pastoris*. *Process Biochem* **36**: 907–913.
10. Brankamp RG, Sreekrishna K, Smith PL, Blankenship DT, Cardin AD. 1995. Expression of a synthetic gene encoding the anticoagulant-antimetastatic protein ghilanten by the methylotropic yeast *Pichia pastoris*. *Prot Exp Pur* **6**: 813–820.
11. Braude R, Hosking ZD, Mitchell KG, Plonka S, Sambrook IE. 1977. Pruteen, a new source of protein for growing pigs. I. Metabolic experiment: utilization of nitrogen. *Livestock Prod Sci* **4**: 79–89.
12. Brethauer RK, Castellino FJ. 1999. Glycosylation of *Pichia pastoris*-derived proteins. *Biotechnol Appl Biochem* **30**: 193–200.
13. Brierley RA, Davis RG, Holtz CG. 1994. Production of insulin-like growth factor-I in methylotrophic yeast cells. US Patent No. 5,324,639.
14. Brocca S, Schmidt-Dannert C, Lotti M, Alberghina L, Schmid RD. 1998. Design, total synthesis, and functional overexpression of the *Candida rugosa* *lip1* gene coding for a major industrial lipase. *Protein Sci* **7**: 1415–1422.
15. Brunel L, Neugnot V, Landucci L, et al. 2004. High-level expression of *Candida parasilosis* lipase/acyltransferase in *Pichia pastoris*. *J Biotechnol* **111**: 41–50.
16. Bushell ME, Rowe M, Avignone-Rossa CA, Wardell JN. 2003. Cyclic fed-batch culture for production of human serum albumin in *Pichia pastoris*. *Biotech Bioeng* **82**(6): 678–683.
17. Callewaert N, Laroy W, Cadirgi H, et al. 2001. Use of HDEL-tagged *Trichoderma reesei* mannosyl oligosaccharide 1,2- α -D-mannosidase for *N*-glycan engineering in *Pichia pastoris*. *FEBS Lett* **503**: 173–178.
18. Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* **24**: 45–66.
19. Cereghino GPL, Cereghino JL, Sunga AJ, et al. 2001a. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*. *Gene* **263**(1–2): 159–169.
20. Cereghino GPL, Sunga AJ, Cereghino JL, Cregg JM. 2001b. Expression of foreign genes in the yeast *Pichia pastoris*. In *Genetic Engineering: Principles and Methods* Setlow JK (ed.). Kluwer Academic/Plenum: London; 157–169.
21. Cereghino GPL, Cereghino JL, Ilgen C, Cregg J. 2002. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr Opin Biotechnol* **13**(4): 329–332.
22. Charlwood J, Bryant D, Skehel JM, Camilleri P. 2001. Analysis of *N*-linked oligosaccharides: progress towards the characterisation of glycoprotein-linked carbohydrates. *Biomol Eng* **18**: 229–240.

23. Chen C-C, Wu P-H, Huang C-T, Cheng K-J. 2004. A *Pichia pastoris* fermentation strategy for enhancing the heterologous expression of a *Escherichia coli* phytase. *Enz Microb Technol* **35**(4): 315–320.
24. Choi B-K, Bobrowicz P, Davidson RC, *et al.* 2003. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. *Appl Biol Sci* **100**(9): 5022–5027.
25. Clare JJ, Rayment FB, Ballantyne SP, Sreerashtra K, Romanos MA. 1991. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *BioTechnology* **9**: 455–460.
26. Cregg JM, Tschopp JF, Stillman C, *et al.* 1987. High-level expression and efficient assembly of hepatitis B surface antigen in *Pichia pastoris*. *BioTechnology* **5**: 479–485.
27. Cregg JM, Madden KR. 1988. Development of the methylotrophic yeast, *Pichia pastoris*, as a host system for the production of foreign proteins. *Dev Ind Microbiol* **29**: 23–42.
28. Cregg JM, Vedvick TS, Raschke WC. 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. *BioTechnology* **11**: 905–910.
29. Cregg JM, Cereghino L, Shi J, Higgins DR. 2000. Recombinant protein expression in *Pichia pastoris*. *Mol Biotech* **16**: 23–52.
30. Crowley J, McCarthy B, Nunn NS, Harvey LM, McNeil B. 2000. Monitoring a recombinant *Pichia pastoris* fed batch process using Fourier transform mid-infrared spectroscopy (FT-MIRS). *Biotechnol Lett* **22**: 1907–1912.
31. Davidson JF. 1995. The origin of insights in chemical engineering: planned and unplanned research. *Chem Eng Sci* **50**(23): 3661–3684.
32. Debski J, Wyslouch-Ciesznka A, Dadlez M, *et al.* 2004. Positions of disulfide bonds and N-glycosylation site in juvenile hormone binding protein. *Arch Biochem Biophys* **421**: 260–266.
33. de Jong LAA, Grunewald S, Franke JP, Uges, DRA, Bischoff R. 2004. Purification and characterisation of the recombinant human dopamine D2S receptor from *Pichia pastoris*. *Prot Exp Pur* **33**: 176–184.
34. Eckart MA, Bussineau CM. 1996. Quality and authenticity of heterologous proteins synthesized in yeast. *Curr Opin Biotechnol* **7**: 525–530.
35. Egli T, van Dijken JP, Veenhuis M, Harder W, Fiechter A. 1980. Methanol metabolism in yeasts: regulation of the synthesis of catabolic enzymes. *Arch Microbiol* **124**: 115–121.
36. Faber KN, Harder W, Ab G, Veenhuis M. 1995. Review: methylotrophic yeasts as factories for the production of foreign proteins. *Yeast* **11**: 1331–1344.
37. Fu R-H, Wang A-Y, Wang Y-C, Sung H-Y. 2003. A cDNA encoding vacuolar type β -D-fructofuranosidase (*Os β fruct3*) of rice and its expression in *Pichia pastoris*. *Biotechnol Lett* **25**: 1525–1530.
38. Gellissen G. 2000. Heterologous protein production in methylotrophic yeasts. *Appl Microbiol Biotechnol* **54**: 741–750.
39. Goochee CF, Gramer MJ, Andersen DC, Bahr JB, Rasmussen JR. 1991. The oligosaccharides of glycoproteins — bioprocess factors affecting oligosaccharide structure and their effect on glycoprotein properties. *Bio Technology* **9**(12): 1347–1355.
40. Guarna MM, Lesnicki GJ, Tan BM, *et al.* 1997. On-line monitoring and control of methanol concentration in shake-flask cultures of *Pichia pastoris*. *Biotechnol Bioeng* **56**(3): 279–286.
41. Gustavsson M, Lehtio J, Denman S, *et al.* 2001. Stable linker peptides for a cellulose-binding domain–lipase fusion protein expressed in *Pichia pastoris*. *Protein Eng* **14**(9): 711–715.
42. Hamilton SR, Bobrowicz P, Bobrowicz B, *et al.* 2003. Production of complex human glycoproteins in yeast. *Science* **301**(5637): 1244–1246.
43. Han X, Bartlam M, Jin Y-H, *et al.* 2004. The expression of SARS–CoV *M* gene in *P. pastoris* and the diagnostic utility of the expression product. *J Virol Methods* **122**: 105–111.
44. Heimo H, Palmu K, Suominen I. 1997. Expression in *Pichia pastoris* and purification of *Aspergillus awamori* glucoamylase catalytic domain. *Prot Exp Pur* **10**: 70–79.
45. Helenius J, Aebi M. 2002. Transmembrane movement of dolichol-linked carbohydrates during N-glycoprotein biosynthesis in the endoplasmic reticulum. *Cell Dev Biol* **13**: 171–178.
46. Higgins DR, Cregg JM (eds). 1998. *Methods in Molecular Biology: Pichia Protocols*. Humana: Totowa, NJ.
47. Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D. 2004. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol Bioeng* **85**(4): 367–375.
48. Hong F, Meinander NQ, Jonsson LJ. 2002. Fermentation strategies for improved heterologous expression of laccase in *Pichia pastoris*. *Biotechnol Bioeng* **79**(4): 438–449.
49. Huang J-S, Ramamurthy SK, Lin X, Le Breton GC. 2004. Cell signalling through thromboxane A₂ receptors. *Cell Signal* **16**: 521–533.
50. Ignatov A, Hermans-Borgmeyer I, Schaller HC. 2004. Cloning and characterization of a novel G-protein-coupled receptor with homology to galanin receptors. *Neuropharmacology* **46**: 1114–1120.
51. Inan M, Meagher MM. 2001a. The effect of ethanol and acetate on protein expression in *Pichia pastoris*. *J Biosci Bioeng* **92**(4): 337–341.
52. Inan M, Meagher MM. 2001b. Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of *Pichia pastoris*. *J Biosci Bioeng* **92**(6): 585–589.
53. Jahic M. 2003. Process Techniques for Production of Recombinant Proteins with *Pichia pastoris*. PhD Thesis, Department of Biotechnology, Royal Institute of Technology, S-10691 Stockholm, Sweden; ISBN 91-7283-522-2.
54. Jahic M, Gustavsson M, Jansen A-K, Martinelle M, Enfors S-O. 2003. Analysis and control of proteolysis of a fusion protein in *Pichia pastoris* fed-batch processes. *Bio Technology* **102**: 42–53.

55. Johnson SK, Zhang W, Smith LA, et al. 2003. Scale-up of the fermentation and purification of the recombinant heavy chain fragment C of botulinum neurotoxin serotype F, expressed in *Pichia pastoris*. *Prot Exp Pur* **32**: 1–9.
56. Jönsson LJ, Saloheimo M, Penttilä M. 1997. Laccase from the white-rot fungus *Trametes versicolor*: cDNA cloning of *lcc1* and expression in *Pichia pastoris*. *Curr Genet* **32**: 425–430.
57. Kang HA, Choi ES, Hong WK, et al. 2000. Proteolytic stability of recombinant human serum albumin secreted in the yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **53**: 575–582.
58. Karlsson M, Fotiadis D, Sjövall S, et al. 2003. Reconstitution of water channel function of an aquaporin overexpressed and purified from *Pichia pastoris*. *FEBS Lett* **537**: 68–72.
59. Kobayashi K, Kuwae S, Ohya T, et al. 2000. Addition of oleic acid increases expression of recombinant human serum albumin by the AOX2 promoter in *Pichia pastoris*. *J Biosci Bioeng* **89**(5): 479–484.
60. Koganesawa N, Aizawa T, Shimojo H, et al. 2002. Expression and purification of a small cytokine growth-blocking peptide from armyworm *Pseudaletia separata* by an optimized fermentation method using the methylotrophic yeast *Pichia pastoris*. *Prot Exp Pur* **25**(3): 416–425.
61. Koutz P, Davis GR, Stillman C, et al. 1989. Structural comparison of the *Pichia pastoris* alcohol oxidase genes. *Yeast* **5**: 167–177.
62. Kramer W, Elmecker G, Weik R, Mattanovich D, Bayer K. 1996. Kinetic studies for the optimisation of recombinant protein formation. *Ann NY Acad Sci* **782**: 323–333.
63. Lee CY, Nakano A, Shioimi N, Lee EK, Katoh S. 2003. Effects of substrate feed rates on heterologous protein expression by *Pichia pastoris* in DO-stat fed-batch fermentation. *Enz Microb Technol* **33**: 358–365.
64. Li Z, Xiong F, Lin Q, et al. 2001. Low temperature increases the yield of biologically active herring antifreeze protein in *Pichia pastoris*. *Prot Exp Pur* **21**(3): 438–445.
65. Liu H, Tan X, Russell KA, Veenhuis M, Cregg JM. 1995. *PER3*, a gene required for peroxisome biogenesis in *Pichia pastoris*, encodes a peroxisomal membrane protein involved in protein import. *J Biol Chem* **270**(18): 10 940.
66. Liu S-T, Saito A, Azakami H, Kato A. 2003. Expression, purification, and characterisation of an unstable lysozyme mutant in *Pichia pastoris*. *Prot Exp Pur* **27**: 304–312.
67. Loll PJ. 2003. Membrane protein structural biology: the high throughput challenge. *J Struct Biol* **142**: 144–153.
68. Lundstrom K. 2003. Semliki Forest virus vectors for rapid and high-level expression of integral membrane proteins. *Biochim Biophys Acta* **1610**: 90–96.
69. Masuda T, Tamaki S, Kaneko R, et al. 2003. Cloning, expression, and characterisation of recombinant sweet-protein thaumatin II using the methylotrophic yeast *Pichia pastoris*. *Biotechnol Bioeng* **85**(7): 761–769.
70. McGrew JT, Leiske D, Dell B, et al. 1997. Expression of trimeric CD40 ligand in *Pichia pastoris*: use of a rapid method to detect high-level expressing transformants. *Gene* **187**: 193–200.
71. McKinney J, Knappskog PM, Pereira J, et al. 2004. Expression and purification of human tryptophan hydroxylase from *Escherichia coli* and *Pichia pastoris*. *Protein Expr Purif* **33**: 185–194.
72. Menendez J, Valdes I, Cabrera N. 2003. The *ICLI* gene of *Pichia pastoris*, transcriptional regulation and use of its promoter. *Yeast* **20**: 1097–1108.
73. Menendez J, Hernandez L, Banguela A, Pais J. 2004. Functional production and secretion of the *Gluconoacetobacter diazotrophicus* fructose-releasing exo-levanase (LsdB) in *Pichia pastoris*. *Enz Microb Technol* **34**(5): 446–452.
74. Miura M, Hirose M, Miwa T, Kuwae S, Ohi H. 2004. Cloning and characterisation in *Pichia pastoris* of *PNO1* gene required for phosphomannosylation of N-linked oligosaccharides. *Gene* **324**: 129–137.
75. Mochizuki S, Hamato N, Hirose M, et al. 2001. Expression and characterisation of recombinant human antithrombin III in *Pichia pastoris*. *Prot Exp Pur* **23**: 55–65.
76. Montesino R, Garcia R, Quintero O, Cremata JA. 1998. Variation in N-linked oligosaccharide structures on heterologous proteins secreted by the methylotrophic yeast *Pichia pastoris*. *Prot Exp Pur* **14**: 197–207.
77. Ning D, Junjian X, Xunzhang W, et al. 2003. Expression, purification, and characterisation of humanised anti-HBs Fab fragment. *J Biochem* **134**: 813–817.
78. Nourizad N, Ehn M, Gharizadeh B, Hober S, Nyren P. 2003. Methylotrophic yeast *Pichia pastoris* as a host for production of ATP-diphosphohydrolase (apyrase) from potato tubers (*Solanum tuberosum*). *Prot Exp Pur* **27**: 229–237.
79. O'Callaghan J, O'Brien MM, McClean K, Dobson ADW. 2002. Optimisation of the expression of a *Trametes versicolor* laccase gene in *Pichia pastoris*. *J Ind Microbiol Biotechnol* **29**: 55–59.
80. Offermanns S. 2003. G-proteins as transducers in transmembrane signalling. *Prog Biophys Mol Biol* **83**: 101–130.
81. Oledzka G, Dabrowski S, Kur J. 2003. High-level expression, secretion, and purification of the thermostable aqualysin I from *Thermus aquaticus* YT-1 in *Pichia pastoris*. *Prot Exp Pur* **29**: 223–229.
82. Opekarová M, Tanner W. 2003. Specific lipid requirements of membrane proteins — a putative bottleneck in heterologous expression. *Biochim Biophys Acta* **1610**: 11–22.
83. Paramasivam M, Saravanan K, Uma K, et al. 2002. Expression, purification and characterization of equine lactoferrin in *Pichia pastoris*. *Prot Exp Pur* **26**: 28–34.
84. Pechan T, Ma PWK, Luthe DS. 2004. Heterologous expression of maize (*Zea mays* L) Mir1 cysteine proteinase in eukaryotic and prokaryotic expression systems. *Prot Exp Pur* **34**(1): 134–141.
85. Peng L, Zhong X, Ou J, et al. 2004. High-level secretory production of recombinant bovine enterokinase light chain by *Pichia pastoris*. *J Biotechnol* **108**: 185–192.

86. Perret BG, Wagner R, Lecat S, *et al.* 2003. Expression of EGFP-amino-tagged human mu opioid receptor in *Drosophila* Schneider 2 cells: a potential expression system for large-scale production of G-protein coupled receptors. *Prot Exp Pur* **31**: 123–132.
87. Raemaekers RJM, de Muro L, Gatehouse JA, Fordham-Skelton AP. 1999. Functional phytohaemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris*: correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. *Eur J Biochem* **265**: 394–403.
88. Reilander H, Weiss HM. 1998. Production of G-protein-coupled receptors in yeast. *Curr Opin Biotechnol* **9**: 510–517.
89. Resina D, Serrano A, Valero F, Ferrer P. 2004. Expression of a *Rhizopus oryzae* lipase in *Pichia pastoris* under control of the nitrogen source-regulated formaldehyde dehydrogenase promoter. *J Biotechnol* **109**(1–2): 103–113.
90. Romanos M. 1995. Advances in the use of *Pichia pastoris* for high-level gene expression. *Curr Opin Biotechnol* **6**: 527–533.
91. Rosenfeld SA, Nadeau D, Tirado J, *et al.* 1999. Production and purification of recombinant hirudin expressed in the methylotrophic yeast *Pichia pastoris*. *Methods Enzymol* **306**: 154–169.
92. Ruf A, Müller F, D'Arcy B, *et al.* 2004. The monotopic membrane protein human oxidosqualene cyclase is active as monomer. *Biochem Biophys Res Commun* **315**: 247–254.
93. Sanden AM, Prytz I, Tubulekas I, *et al.* 2003. Limiting factors in *Escherichia coli* fed-batch production of recombinant proteins. *Biotechnol Bioeng* **81**: 159–166.
94. Sarramegna V, Talmont F, de Roch MS, Milon A, Demange P. 2002. Green fluorescent protein as a reporter of human μ -opioid receptor overexpression and localization in the methylotrophic yeast *Pichia pastoris*. *J Biotechnol* **99**: 23–39.
95. Scharstuhl A, Glansbeek H, Vitters EL, van der Kraan PM, van den Berg WB. 2003. Large scale production of the extracellular domain of the transforming growth factor- β type II receptor using the *Pichia pastoris* expression system. *J Chromatogr B* **786**: 271–277.
96. Segev N, Mulholland J, Bostein D. 1988. The yeast GTP-binding YPT1 protein and a mammalian counterpart are associated with the secretion machinery. *Cell* **52**: 915–924.
97. Shen S, Sulter G, Jeffries TW, Cregg JM. 1998. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **216**: 93–102.
98. Shi X, Karkut T, Chamankhah M, *et al.* 2003. Optimal conditions for the expression of a single-chain antibody (scFv) gene in *Pichia pastoris*. *Prot Exp Pur* **28**: 321–330.
99. Siegel RS, Buckholz RG, Thill GP, Wondrack LM. 1990. Production of epidermal growth factor in methylotrophic yeasts. *International Patent Application*, Publ. No. WO 90/10 697.
100. Silberbach M, Maier B, Zimmermann M, Buchs J. 2003. Glucose oxidation by *Gluconobacter oxydans*: characterization in shaking flasks, scale-up and optimization of the pH profile. *Appl Microbiol Biotechnol* **62**: 92–98.
101. Sinha J, Plantz BA, Zhang W, *et al.* 2003. Improved production of recombinant ovine interferon- τ by Mut⁺ strain of *Pichia pastoris* using an optimised methanol feed profile. *Biotechnol Prog* **19**: 794–802.
102. Slibinskas R, Samuel D, Gedvilaite A, Staniulis J, Sasnauskas K. 2004. Synthesis of the measles virus nucleoprotein in yeast *Pichia pastoris* and *Saccharomyces cerevisiae*. *J Biotechnol* **107**: 115–124.
103. Sreekrishna K, Prevatt WD, Thill JP, *et al.* 1993. Production of *Bacillus* entomotoxins in methylotrophic yeast. *European Patent Application*, Publ. No. EP0586 892 A1.
104. Sreekrishna K, Brankamp RG, Kropp KE, *et al.* 1997. Strategies for optimal synthesis of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene* **190**: 55–62.
105. Stahl CH, Wilson DB, Lei XG. 2003. Comparison of extracellular *Escherichia coli* AppA phytases expressed in *Streptomyces lividans* and *Pichia pastoris*. *Biotechnol Lett* **25**: 827–831.
106. Stapleton PC, O'Brien MM, O'Callaghan J, Dobson ADW. 2004. Molecular cloning of the cellobiose dehydrogenase gene from *Trametes versicolor* and expression in *Pichia pastoris*. *Enz Microb Technol* **34**: 55–63.
107. Sunga AJ, Cregg JM. 2004. The *Pichia pastoris* formaldehyde dehydrogenase gene (*PNO1*) as a marker for selection of multicopy expression strains of *P. pastoris*. *Gene* **330**: 39–47.
108. Surribas A, Cos O, Montesinos JL, Valero F. 2003. On-line monitoring of the methanol concentration in *Pichia pastoris* cultures producing an heterologous lipase by sequential injection analysis. *Biotechnol Lett* **25**: 1795–1800.
109. Tate CG, Haase J, Baker C, *et al.* 2003. Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim Biophys Acta* **1610**: 141–153.
110. Trinh LB, Phue JN, Shiloach J. 2003. Effect of methanol feeding strategies on production and yield of recombinant mouse endostatin from *Pichia pastoris*. *Biotechnol Bioeng* **82**(4): 438–444.
111. Van den Hazel HB, Kielland-Brandt MC, Winther JR. 1996. Review: biosynthesis and function of yeast vacuolar proteases. *Yeast* **12**: 1–16.
112. Vassileva A, Chugh DA, Swaminathan S, Khanna N. 2001. Expression of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris* using the GAP promoter. *J Biotechnol* **88**: 21–35.
113. Veenhuis M, van Dijken JP, Harder W. 1983. The significance of peroxisomes in the metabolism of one-carbon compounds in yeast. *Adv Microb Physiol* **24**: 1–82.
114. Villatte F, Hussein AS, Bachmann TT, Schmid RD. 2001. Expression level of heterologous proteins in *Pichia pastoris* is influenced by flask design. *Appl Microbiol Biotechnol* **55**: 463–465.
115. Wan H, Sjölander M, Schairer HU, Leclercq A. 2004. A new dominant selection marker for transformation of *Pichia pastoris* to soraphen A resistance. *J Microbiol Methods* **57**: 33–39.

116. Wang S-H, Yang T-S, Lin S-M, et al. 2002. Expression, characterisation and purification of recombinant porcine lactoferrin in *Pichia pastoris*. *Prot Exp Pur* **25**: 41–49.
117. Wegner G. 1990. Emerging applications of the methylotrophic yeasts. *FEMS Microbiol Rev* **7**: 279–283.
118. Werten MWT, van den Bosch TJ, Wind RD, Mooibroek H, de Wolf FA. 1999. High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **15**(11): 1087–1096.
119. White CE, Kempf NM, Komives EA. 1994. Expression of highly disulfide-bonded proteins in *Pichia pastoris*. *Structure* **2**(11): 1003–1005.
120. White JF, Trinh LB, Shiloach J, Grishammer R. 2004. Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett* **564**: 289–293.
121. Whittaker MM, Whittaker JW. 2000. Expression of recombinant galactose oxidase by *Pichia pastoris*. *Prot Exp Pur* **20**: 105–111.
122. Wu JM, Chieng LL, Hsu TA, Lee CK. 2003a. Sequential expression of recombinant proteins and their separate recovery from *Pichia pastoris* cultivation. *Biochem Eng J* **3715**: 1–8.
123. Wu JM, Lin JC, Chieng LL, Lee CK, Hsu TA. 2003b. Combined use of *GAP* and *AOX1* promoter to enhance the expression of human granulocyte-macrophage colony-stimulating factor in *Pichia pastoris*. *Enz Microb Technol* **33**(4): 453–459.
124. Xie J, Zhang L, Ye Q, et al. 2003. Angiostatin production in cultivation of *Pichia pastoris* fed with mixed carbon sources. *Biotechnol Letts* **25**: 173–177.
125. Yan Y, Chen J, Li J. 2003. Overexpression of a small medicinal peptide from ginseng in the yeast *Pichia pastoris*. *Prot Exp Pur* **29**: 161–166.
126. Yang G, Zhou H, Lu Y, Lin Y, Zhou S. 2004a. Comparing expression of different forms of human DNA topoisomerase I in *Pichia pastoris*. *Enz Microb Technol* **34**(2): 139–146.
127. Yang Z, Zhou X, Zhang Y. 2004b. Improvement of recombinant hirudin production by controlling NH_4^+ concentration in *Pichia pastoris* fermentation. *Biotechnol Letts* **26**: 1013–1017.
128. You W-K, Choi W-S, Koh Y-S, et al. 2004. Functional characterisation of recombinant batroxobin, a snake venom thrombin-like enzyme, expressed in *Pichia pastoris*. *FEBS Lett* **571**: 67–73.
129. Yu J, Li D-Y, Zhang Y-J, et al. 2002. High expression of *Trigonopsis variabilis* D-amino acid oxidase in *Pichia pastoris*. *J Mol Cat B Enz* **18**(4–6): 291–297.
130. Zhang W, Bevins MA, Plantz BA, Smith LA, Meagher MM. 2000. Modelling *Pichia pastoris* growth on methanol and optimising the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. *Biotechnol Bioeng* **70**: 1–8.
131. Zhen Z, Bradel-Tretheway BG, Dewhurst S, Bidlack JM. 2004. Transient overexpression of κ and μ opioid receptors using recombinant adenovirus vectors. *J Neurosci Methods* **136**(2): 133–139.
132. Zhong X, Peng L, Zheng S, et al. 2004. Secretion, purification and characterisation of a recombinant *Aspergillus oryzae* tannase in *Pichia pastoris*. *Prot Exp Pur* **36**: 165–169.
133. Zhou X-S, Zhang Y-X. 2002. Decrease of proteolytic degradation of recombinant hirudin produced by *Pichia pastoris* by controlling the specific growth rate. *Biotechnol Lett* **24**(17): 1449–1453.
134. Zhou X-S, Lu J, Fan WM, Zhang Y-X. 2002. Development of a responsive methanol sensor and its application in *Pichia pastoris* fermentation. *Biotechnol Lett* **24**: 643–646.