Secretory expression and characterization of insulin in *Pichia pastoris*

Thomas Kjeldsen[1], Annette Frost Pettersson and Morten Hach

Department of Insulin Research and Protein Chemistry, Novo Nordisk A/S, Novo Alle 6B S.90, DK-2880 Bagsvaerd, Denmark

The yeasts *Pichia pastoris* and *Saccharomyces cerevisiae* have similar overall features regarding the secretory expression of insulin. The *S. cerevisiae* mating factor α (α-factor) prepro-leader facilitated the secretion of an insulin precursor, but not proinsulin expressed in *P. pastoris*. Synthetic prepro-leaders developed for the secretory expression of the insulin precursor in *S. cerevisiae* also facilitated the secretion of the insulin precursor expressed in *P. pastoris*. In contrast with *S. cerevisiae*, only insulin precursor and no unprocessed hyperglycosylated α-factor pro-leader/insulin precursor fusion protein was secreted from *P. pastoris*. A spacer peptide in the fusion protein increased the fermentation yield of the insulin precursor in *P. pastoris*. A synthetic prepro-leader, but not an α-factor prepro-leader lacking N-glycosylation sites, facilitated the secretion of the insulin precursor in *P. pastoris*. *P. pastoris* has a capacity for secretory expression of the insulin precursor that is equal to or better than that of *S. cerevisiae*. Peptide mapping and MS indicated a structure of the insulin precursor expressed in *P. pastoris* identical with that of human insulin.

Introduction

Yeast is well suited to the expression of heterologous proteins of pharmaceutical importance. Yeast's secretory pathway exhibits much of the structure and function of the mammalian secretory system and has the capacity to fold, to process proteolytically, to glycosylate and to secrete proteins. The yeast *Saccharomyces cerevisiae* is one of the best characterized eukaryotic expression systems. The entire *S. cerevisiae* genome has been sequenced and more than 1000 genes have been characterized [1,2]. An example of the importance of yeast within biotechnology is the production of insulin for treatment of diabetes mellitus; approximately half of the world's need for insulin is produced by *S. cerevisiae*. An increasing number of alternative yeast species have become important expression organisms for the production of recombinant heterologous proteins. Among these are methylotrophic yeast species like *Pichia pastoris*, *Hansenula polymorpha*, *Pichia methanolica* and the budding yeast *Kluyveromyces lactis* [3–8]. In particular, the ascomycetous yeast *P. pastoris* has become successful in the production of high levels of a broad range of heterologous proteins [4,5,9,10]. High-level expression, high density of cells, easy scaling up, strong and tightly regulated promoters as well as limited or absent hyperglycosylation, have all been implied as the major advantages of *P. pastoris*. In certain aspects the secretory expression of heterologous proteins in both *S. cerevisiae* and *P. pastoris* can be different from that in other eukaryotic expression systems, and association with secretion pro-leaders might be necessary to confer secretion competence on heterologous proteins. Relatively few prepro-leaders are available to facilitate the export of heterologous proteins from yeast. The *S. cerevisiae* mating factor α (α-factor) prepro-leader is the classical yeast prepro-leader; it has been used for the secretory expression of numerous heterologous proteins in both *S. cerevisiae* and *P. pastoris* [4,6,11,12]. The *S. cerevisiae* α-factor prepro-leader consists of a 19-residue signal (pre) sequence followed by a 66-residue pro-sequence containing three consensus N-linked glycosylation sites and a dibasic Kex2 endoprotease (an endopeptidase processing *S. cerevisiae* α-factor precursor) processing site [13].

The peptide hormone insulin is essential for glucose homeostasis, and recombinant insulin is important in the treatment of diabetes mellitus. Insulin is normally synthesized as preproinsulin in the pancreatic β-cells and is processed to the two-chain 51-residue globular peptide hormone in the secretory pathway. The 21-residue A-chain features two α-helices (A2–A8 and A13–A19) joined by a loop (A9–A11) and one intrachain disulphide bridge (A6 to A11). The A-chain is joined to the 30-residue B-chain by two disulphide

Abbreviations used: α-factor, mating factor α; IP, human insulin precursor sequence comprising the first 29 residues of the B chain and the 21 residues of the A chain of human insulin connected by the peptide sequence Ala-Ala-Lys; Kex2 endoprotease, endopeptidase processing *S. cerevisiae* α-factor precursor; MALDI–MS, matrix-assisted laser desorption ionization MS; PNGase F, peptide N-glycosidase F; RP-HPLC, reverse-phase HPLC.

[1] To whom correspondence should be addressed.

© 1999 Portland Press Ltd

bridges (A7 to B7 and A20 to B19). Proinsulin is not readily expressed in *S. cerevisiae*, but the expression of a smaller insulin precursor (IP), comprising the first 29 residues of the insulin B-chain and the 21 residues of the insulin A-chain connected by the peptide sequence Ala-Ala-Lys, as a fusion protein with a suitable prepro-leader, promotes efficient secretion [14–17]. Proinsulin and the IP were expressed in *P. pastoris* by using different prepro-leaders to confer secretory competence. The secreted IPs were characterized by reverse-phase HPLC (RP-HPLC) and MS. The structure of purified IPs was evaluated by peptide mapping combined with on-line RP-HPLC–MS.

Materials and methods

Materials

Components for microbial media and skimmed milk were obtained from Difco (Detroit, MI, U.S.A.) or Merck (Damstadt, Germany). DNA restriction enzymes, T4 ligase and peptide N-glycosidase F (PNGase F) were from New England Biolabs. Precast SDS/PAGE gels, buffers and electrophoresis apparatus were obtained from Novex (San Diego, CA, U.S.A.); dNTPs and Taq polymerase were from Boehringer Mannheim (Mannheim, Germany). PCR was performed in a Perkin Elmer thermal cycler 480. PVDF membrane was from Millipore (Bedford, MA, U.S.A.). Monoclonal antibody F19 against insulin was a gift from Dr. Per Nobert Jørgensen (Novo Nordisk) and anti-mouse antiserum conjugated to horseradish peroxidase was from Dako (Glostrup, Denmark). The enhanced chemiluminescence (ECL®) detection system was from Amersham (Little Chalfont, Bucks., U.K.). The 1 mm × 250 mm micro-HPLC column packed with Vydac 5 mm particles was obtained from LC Packings (Amsterdam, The Netherlands). Other chemicals were from Sigma Chemical Co. (St. Louis, MO, U.S.A.) or Merck.

Stains, media and fermentation

The *P. pastoris* expression system was purchased from InVitrogen (Carlsbad, CA, U.S.A.) and evaluated as recommended by the supplier. The *P. pastoris* strain GS115 (his4) was selected for expression of the IP [18]. The *S. cerevisiae* strain MT663 (*MATa/MATα pep*4-3/*pep*4-3 *HIS*4/*his*4 *tpi*::*LEU*2/*tpi*::*LEU*2 Cir$^+$) and the DNA encoding proinsulin were a gift from Dr. Mogens Trier Hansen (Novo Nordisk). For comparison, both *P. pastoris* and *S. cerevisiae* were grown in medium based on non-buffered peptone and yeast extract. Media were: MGY medium, 1% (w/v) yeast extract/2% (w/v) peptone/1.34% (w/v) yeast nitrogen base/400 μg/l biotin/1% (v/v) glycerol; MMY medium, as MGY medium except that 1% (v/v) methanol was used instead of glycerol. MD agar [1.34% (w/v) yeast nitrogen base/400 μg/l biotin/1% (w/v) glucose] was used for screening of HIS$^+$ transformants. *P. pastoris* was grown for 12 h in 5 ml of MGY medium at 30 °C, after which the media was changed to MMY medium and fermentation continued for 96 h; 1% (v/v) methanol was added every 24 h. Fermentation of *S. cerevisiae* was performed in YPD medium [1% (w/v) yeast extract/2% (w/v) peptone/2% (w/v) glucose] at 30 °C for 72 h. *P. pastoris and S. cerevisiae* were grown under identical physical conditions.

Construction of expression vectors

Secretory expression of the IP in *P. pastoris* was performed by association with a suitable prepro-leader to facilitate secretion, as described for expression in *S. cerevisiae* [14–17]. DNA fragments encoding a fusion protein with the following configuration was constructed: signal peptide–pro-leader–KR–spacer–IP, where KR is a potential dibasic endoprotease-processing site. The DNA fragments encoding the different prepro-leader/IP fusion proteins were subcloned into the *P. pastoris* expression vector pPIC3.5 by using standard methods [19]. The prepro-leaders (shown in Table 1) have, with the exception of TA39, previously been described in connection with their secretory expression of the IP in *S. cerevisiae* [17,20]. DNA encoding the prepro-leader TA39 was generated by PCR with appropriate oligonucleotides as primers and cloned by using standard methods. The DNA sequences were verified by using standard methods. To optimize the processing of the fusion protein in *P. pastoris* by an endopeptidase similar to the *S. cerevisiae* Kex2 endoprotease, DNA encoding a spacer peptide was inserted between the DNA encoding the prepro-leader and the IP in most of the fusion proteins (Table 1) [16,17]. Expression of the IP in *S. cerevisiae* was performed with the *S. cerevisiae* strain MT663 and the 2μ-based yeast expression vector CPOT as described previously [14–17]. All IP constructs expressed in *P. pastoris* and *S. cerevisiae* lacked residue 30 in the insulin B chain. After purification of the IP, the Thr^{B30} could be added to Lys^{B29} by trypsin-mediated transpeptidation to generate human insulin [14]. Thus the IP was secreted as a single-chain N-terminally extended proinsulin-like peptide with a short synthetic peptide, Ala-Ala-Lys, connecting Lys^{B29} and Gly^{A1}.

Transformation and screening of yeasts for secretory expression of IP

The *P. pastoris* strain GS115 (*his*4) was transformed with the various expression plasmids digested with the restriction endonuclease *Sac*I. Approx. 2 μg of linearized plasmid DNA was used to transform *P. pastoris* cells that had been turned

© 1999 Portland Press Ltd

Table 1 Secretory expression of IP by prepro-leaders in *P. pastoris* and *S. cerevisiae*

The IP fermentation yield (N=4) was determined by RP-HPLC of the culture supernatant relative to the fermentation of the *S. cerevisiae* strain yAK708 (α-factor prepro-leader/EEAEAEAEPK/IP). *P. pastoris* and *S. cerevisiae* cultures both reached an optical density at 600 nm of approximately 22–23, indicating that the two yeast species achieve a similar biomass under the fermentation condition used. Culture supernatant from *P. pastoris* strains yAK1022, yAK1023, and yAK1024 contained IP's with two slightly different retention times on RP-HPLC. The fermentation yield is given as total of insulin related peaks. The amino acid sequences of the α-factor pro-leader, the LA19 pro-leader [17] and the consensus pro-leader[b] (yAK928) [25] are given in full. The amino acid sequences of additional prepro-leaders are only indicated at the position where they differ from either the α-factor pro-leader or the LA19 pro-leader. The α-factor prepro-leader was modified at amino acids positions 63 and 64 in the C-terminus from LD to MA as previously described [16]. Consensus N-linked glycosylation sites as well as mutated non-functional glycosylation sites are underlined.

P. pastoris strain	*S. cerevisiae* strain	Pro-leader	Pro-leader amino acid sequence	Spacer	*P. pastoris* yield	*S. cerevisiae* yield
yAK1002	yAK944	α	APVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTNNGLL FINTTIASIAAKEEGVSMAKR	–	54%	28%
yAK1022	yAK708	α	---NTT-------------------------------NST----- ---NTT-------------KR	EEAEAEAEPK	167%	100%
yAK1023	yAK904	TA71	---QTT-------------------------------QST----- ---QTT-------------KR	EEAEAEAEPK	<1%	15%
yAK1024	yAK721	LA19	QPIDDTESNTTSVNLMADDTESRFATNTTLALDVVNLISMAKR	EEAEAEAEPK	54%	119%
yAK1031	yAK817	TA39	--------NTT---------------NTT--GGLDVVNLISMAKR	EEGEPK	67%	156%
yAK1026	yAK855	TA57	--------QTT-----------A---QTNSGGLDVVGLISMAKR	EEGEPK	93%	153%
yAK1027	yAK928	Pro-leader[b]	QPVISTTVGSAAEGSLDKR	EEAEAEAEPK	<1%	12%

into spheroplasts. Individual clones of each construct were grown in liquid culture; the culture supernatant was tested for IP production and the clone secreting the greatest amount of IP was re-isolated and used for further studies. Quantification of the IP yield in the culture supernatants was performed by RP-HPLC analysis with human insulin as an external standard [22]. Similarly, *S. cerevisiae* clones of each construct were tested for IP production and the clone secreting the greatest IP amount was re-isolated and used for further studies.

SDS/PAGE analysis and Western blot analysis

Proteins were separated by SDS/PAGE [Tricine 10% (w/v) gel with Mes/SDS buffer] with NuPAGE gels (Novex, San Diego, CA, U.S.A.) in accordance with the manufacturers' instructions. Deglycosylation by PNGase F was performed as described previously [16]. Cell-free supernatant (15 μl) was loaded in each well. Protein was revealed by silver staining. For Western blotting, cellular and secreted proteins were separated by SDS/PAGE. Cells were isolated from 100 μl of culture and the cells were lysed by adding 100 μl of glycosidase buffer [0.05 M Hepes (pH 8.0)/0.1% SDS/0.5% n-octyl-β-D-glucopyranoside/0.1% 2-mercaptoethanol], 1.3 μl of PMSF (100 mM in DMSO) and acid-treated glass beads (425–600 μm; Sigma) to approx. 2 mm below the surface. The samples were mixed for 20 min at 4 °C, incubated for 3 min at 100 °C and then centrifuged for 10 min at 20000 **g** and 4 °C. Proteins were separated by SDS/PAGE and then transferred to PVDF membrane; additional protein binding was blocked by incubation with 5% (w/v) skimmed milk for 18 h at 4 °C. Immunoreactive protein was detected by the monoclonal antibody F19 (Novo Nordisk) against insulin. Peroxidase-conjugated antibody was used as secondary antibody (Dako) and detected by enhanced chemiluminescence. Cell-free supernatant (15 μl), or material derived from cells isolated from 15 μl of culture, was loaded in each well.

MS of IPs secreted by *P. pastoris* and *S. cerevisiae*

The molecular masses of IPs secreted by *P. pastoris* and *S. cerevisiae* were determined by MS by using the culture supernatant directly. Matrix-assisted laser desorption ionization (MALDI)–MS was performed with a Voyager-DE mass spectrometer (PerSeptive Biosystems, Framingham, MA, U.S.A.). Cell-free culture supernatants were mixed 1:10 with sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid) before being subjected to MALDI–MS.

Purification and characterization of IPs expressed in *P. pastoris* by peptide mapping combined with on-line RP-HPLC–MS

IPs were purified from the fermentation supernatants of the *P. pastoris* clones yAK1022, yAK1024, yAK1026 and yAK1002 by RP-HPLC as described previously [21]. Fractions containing the isolated IPs were dried in vacuum. The dried samples were reconstituted in 40 μl of 0.4 M Tris/HCl, pH 7.5. Subsequently, 4 μg, corresponding to 2 units, of *Staphylococcus aureus* V8 protease dissolved in 4 μl of water was added to each sample and the digestion was left for 6 h at 37 °C. The reaction (digestion) was terminated by the addition of 10 μl of 2 M acetic acid. On-line RP-HPLC–MS analysis of the digested IP samples was performed on a

© 1999 Portland Press Ltd

1 mm $\times$ 250 mm RP-HPLC column packed with Vydac 5 μm C_{18} particles (LC Packings) as described previously [21].

Results and discussion

Secretory expression of the IP in *P. pastoris* by using different prepro-leaders

Expression of an α-factor-leader/proinsulin fusion protein in *P. pastoris* did not result in a secretion of proinsulin. Replacement of the C-peptide with a short connecting peptide, Ala-Ala-Lys, and association with the α-factor prepro-leader (yAK1002) resulted in the efficient secretory expression of the IP in *P. pastoris* (Table 1 and Figure 1, lanes 1 and 2). Presumably a *P. pastoris* protease analogue of the *S. cerevisiae* Kex2 endoprotease was able to cleave the fusion protein. Similar results have been published on the secretory expression of insulin in *S. cerevisiae* [14–16]. Mutation of the α-factor prepro-leader consensus N-linked glycosylation sites impaired the secretion of the IP, indicating the importance of α-factor prepro-leader N-linked carbohydrates in the secretion of the IP in *P. pastoris* (Table 1 and Figure 1, lane 5). The N-linked glycosylation of the α-factor prepro-leader has previously been shown to be important for the ability to facilitate the efficient transport of both the α-factor and the IP through the secretory pathway of *S. cerevisiae* [21,23,24]. These results indicate that *P. pastoris* and *S. cerevisiae* have similar overall characteristics in their expression and secretion of insulin. There were, however, a number of differences between *P. pastoris* and *S. cerevisiae*

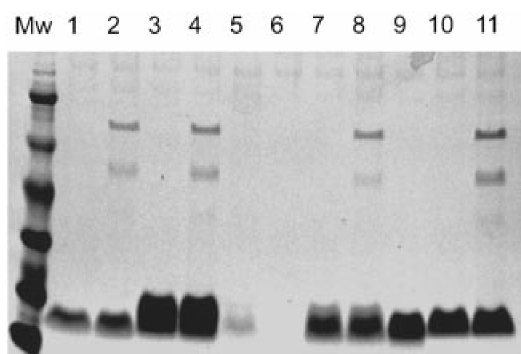

Figure 1 SDS/PAGE of culture supernatant from *P. pastoris* strains expressing the IP by different prepro-leaders

Culture supernatants from yeast strains carrying various expression plasmids were separated by Tricine/SDS/PAGE to evaluate the efficiency of processing and secretion of prepro-leader/IP fusion protein. Culture supernatant from *P. pastoris* strain GS115 carrying various expression plasmids was analysed without (lanes 1, 3, 5, 6, 7, 9 and 10) or with (lanes 2, 4, 8 and 11) treatment with PNGase F. PNGase F has a molecular mass of 36 kDa. Lanes 1 and 2, yAK1002; lanes 3 and 4, yAK1022; lane 5, yAK1023; lane 6, yAK1027; lanes 7 and 8, yAK1024; lane 9, yAK1026, lanes 10 and 11, yAK1031. Lane Mw, rainbow molecular marker (Amersham) with molecular masses of 46, 30, 21.5, 14.3, 6.5 and 3.4 kDa.

regarding the secretory expression of the IP. An α-factor prepro-leader/IP fusion protein expressed in *S. cerevisiae* was processed only partly by the Kex2 endoprotease, and both IP and hyperglycosylated fusion protein were secreted to the culture supernatant [16]. The identical fusion protein expressed in *P. pastoris* was completely processed and no hyperglycosylated fusion protein was secreted; only IP was detected in the culture supernatant (Figure 1, lanes 1 and 2). In *S. cerevisiae*, the processing of the fusion protein by Kex2 endoprotease could be optimized and the fermentation yield increased by the insertion of a spacer peptide (yAK708) between the prepro-leader and the IP [16,17]. The insertion of a spacer peptide between the α-factor prepro-leader and the IP (yAK1022) also increased the fermentation yield in *P. pastoris* (Table 1 and Figure 1, lanes 3 and 4). Whether the spacer peptide optimized the endoprotease processing of the fusion protein or secretion directly in *P. pastoris* is unclear. RP-HPLC of culture supernatant from *P. pastoris* strains yAK1022, yAK1023 and yAK1024 showed two slightly different retention times, indicating post-translational modification of the IP N-terminally extended with the amino acid sequence EEAEAEAEPK. SDS/PAGE of culture supernatant of *P. pastoris* strain yAK1022 showed a broad band, further indicating that some modification of the secreted N-terminally extended IP might have occurred (Figure 1).

Synthetic prepro-leaders developed for secretory expression in *S. cerevisiae* also facilitated the secretion of the IP in *P. pastoris* (Table 1 and Figure 1, lanes 7–11) [20]. The synthetic leaders resulted in a higher secretion of the IP in *S. cerevisiae* than in *P. pastoris* under the growth conditions used (Table 1). This is in contrast with the α-factor prepro-leader, which resulted in a higher secretion of the IP secretion in *P. pastoris* (Table 1). This difference might reflect the fact that the synthetic prepro-leaders were developed to facilitate the secretion of the IP expressed in *S. cerevisiae* and some optimization might be necessary to achieve the same secretory competence in *P. pastoris*. In contrast with the α-factor prepro-leader, synthetic prepro-leader lacking N-linked glycosylation sites (TA57) efficiently facilitated the secretion of the IP expressed in *P. pastoris*, as has also been reported for *S. cerevisiae* (Table 1 and Figure 1, lane 11) [20]. A 19-residue synthetic prepro-leader without consensus N-linked glycosylation sites facilitated the secretion of epidermal growth factor in *S. cerevisiae* and Tick anticoagulant peptide in *P. pastoris* [25,26]. This prepro-leader poorly facilitated the secretion of IP in both *P. pastoris* and, as described previously, in *S. cerevisiae* (Table 1 and Figure 1, lane 6) [20].

Western blot analysis showed that very little or no IP was retained intracellularly in *P. pastoris* cells (Figure 2B). No fusion protein was present intracellularly or extracellularly, indicating the efficient proteolytic cleavage in *P. pastoris* of fusion proteins based on both the α-factor prepro-leader

© 1999 Portland Press Ltd

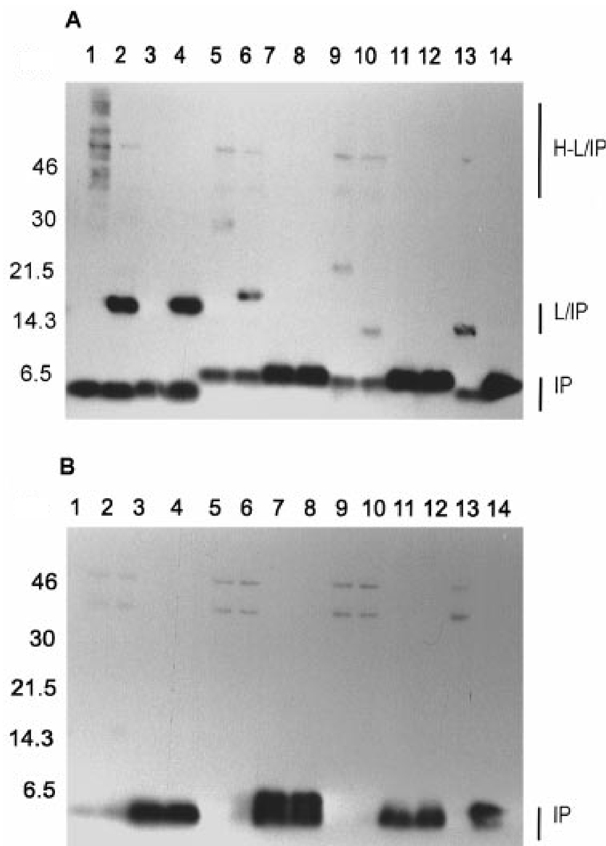

Figure 2 Western blot analysis of cells and culture supernatant from *S. cerevisiae* (A) and *P. pastoris* (B) expressing the IP

S. cerevisiae (**A**) and *P. pastoris* (**B**) cells were analysed for IP immunoreactive material with (lanes 2, 6 and 10) or without (lanes 1, 5, 9 and 13) treatment by PNGase F. *P. pastoris* and *S. cerevisiae* culture supernatants were also analysed for IP imumunoreactive material with (lanes 4, 8 and 12) or without (lanes 3, 7, 11 and 14) treatment by PNGase F. (**A**) Lanes 1–4, yAK944; lanes 5–8, yAK708; lanes 9–12, yAK721; lanes 13 and 14, yAK855. (**B**) Lanes 1–4, yAK1002; lanes 5–8, yAK1022; lanes 9–12, yAK1024; lanes 13 and 14, yAK1026. The positions of molecular mass markers are shown (in kDa) at the left. Abbreviations: H-L/IP, hyperglycosylated pro-leader/IP; L/IP, pro-peptide/IP.

and the synthetic prepro-leaders (Figure 2B). Western blot analysis of culture supernatant of *P. pastoris* strain yAK1022 showed a broad band, indicating that some modification of the secreted N-terminally extended IP might have occurred (Figure 2B, lanes 7–8). Western blot analysis of *S. cerevisiae* cells expressing the IP showed intracellular retention of the IP (Figure 2A). In addition, hyperglycosylated prepro-leader/IP fusion protein, not cleaved by the Kex2 endoprotease, was retained intracellularly (Figure 2A). Insertion of the spacer peptide into the α-factor prepro-leader/IP fusion protein (yAK708) increased the secretion and decreased the intracellular retention of the IP (Figure 2A, lanes 1–8). The lack of intracellular retention of the IP might indicate that *P. pastoris* has a secretory pathway with a higher capacity than that of *S. cerevisiae* for folding and transporting recombinant proteins. In contrast with the *S. cerevisiae* MT663 strain, the *P. pastoris* GS115 strain had intact vacuole proteolytical degradation. The lack of intracellular IP in *P. pastoris* cells is therefore not necessarily a result of efficient secretion but might reflect a more efficient vacuole degradation in the *P. pastoris* strain GS115.

Molecular mass of IP expressed in *P. pastoris* and *S. cerevisiae*

MALDI–MS of cell-free culture supernatant from *P. pastoris* and *S. cerevisiae* expressing the IP was performed to evaluate the potential post-translational modification of secreted IP. MALDI–MS of cell-free culture supernatant from the *P. pastoris* strain yAK1002 showed a single peak with a molecular mass equivalent to the expected theoretical molecular mass (Table 2). MALDI–MS of cell-free culture supernatant from *S. cerevisiae* cells expressing various IPs showed one major peak with a molecular mass equivalent to the expected theoretical molecular mass (Table 2). Thus neither the *P. pastoris* strain GS115 nor the *S. cerevisiae* strain MT663 degraded the IP proteolytically under the fermentation conditions used here. Expression of the α-factor prepro-leader/spacer/IP fusion protein in *P. pastoris* (yAK1022) increased IP fermentation yield but also resulted in the secretion of IPs with different molecular masses (Table 2). The lowest molecular mass found in the culture supernatant was equivalent to the IP lacking the spacer/N-terminal extension, indicating a cleavage after the lysine in the spacer and before insulin Phe^{B1}. The highest molecular mass found in the culture supernatant was equivalent to the unprocessed N-terminal extended IP, indicating that only part of the secreted IP was cleaved proteolytically. In addition, IPs with molecular masses of 6583.64 and 6382.73 Da were present in the culture supernatant, indicating an additional processing, presumably internally in the spacer peptide (Table 2). The intermediate molecular masses might represent the enzymic removal of EEA and EEAEA (single-letter amino-acid code) from the N-terminal extended IP. A similar pattern of molecular masses was also found when the IP was expressed in *P. pastoris* in the context of the synthetic prepro-leaders and the same spacer/N-terminal extension (EEAEAEAEPK) (Table 2). The results indicate that only the spacer peptide/N-terminal extension and not the IP was cleaved in *P. pastoris*. Thus proteolytical processing is of no importance in the quantity and quality of the expressed product, because the N-terminal extension is removed enzymically when the IP is converted to human insulin. The removal of certain types of N-terminal extension from the IP expressed in *S. cerevisiae* by the yeast aspartyl protease 3 has been described previously [16]. A *P. pastoris* analogue of *S. cerevisiae* yeast aspartyl protease 3 might remove the N-terminal extension from the IP. Several

© 1999 Portland Press Ltd

Table 2 MALDI–MS of IPs expressed by *P. pastoris* and *S. cerevisiae*

The molecular masses of IPs secreted by *P. pastoris* and *S. cerevisiae* were determined by MS with the culture supernatant directly. Cell-free culture supernatants were mixed 1:10 with sinapinic acid before being subjected to MALDI–MS.

P. pastoris/*S. cerevisiae* strain	Theoretical molecular mass of expressed IP (Da)	Experimentally determined mass of IP expressed in *P. pastoris* (Da)	Experimentally determined mass of IP expressed in *S. cerevisiae* (Da)
yAK1002/yAK944	5958.90	5956.29	5957.22
yAK1022/yAK708	7043.02	7041.13	7040.03
		6583.64	
		6382.73	
		5957.55	
yAK1024/yAK721	7043.02	7041.14	7040.4
		6583.74	
		6383.24	
		5957.95	
yAK1026/yAK855	6628.60	6627.12	6627.1
yAK1031/yAK817	6628.60	6627.17	6625.58

Table 3 Characterization of purified IPs expressed in *P. pastoris* by peptide mapping and on-line RP-HPLC-MS analysis

Numbers in column 1 refer to positions in the insulin molecule and list the peptides obtained by digestion with *Staph. aureus* V8 of the IPs with the correct disulphide bridges. Abbreviation: n.d., not determined.

Fragments	Theoretically determined molecular mass (Da)	Experimentally determined molecular mass (Da)
Generated by enzymic processing of (B1–B29–AAK–A1–A21) IP (yAK1002)		
B14–B29/A18–A21	1377.56	1377.14
B22–B29–AAK–A1–A4	1683.97	1684.22
B1–B13/A5–A17	2969.39	2969.18
Generated by enzymic processing of [E(EA)$_3$EPK–B1–B29–AAK–A1–A21] IP (yAK1022)		
EEAEAE	658.56	n.d.
B14–B29/A18–A21	1377.56	1377.32
B22–B29–AAK–A1–A4	1683.97	1683.86
AEPK–B1–B13/A5–A17	3394.88	3394.14
Generated by enzymic processing of [E(EA)$_3$EPK–B1–B29–AAK–A1–A21] IP (yAK1024)		
EEAEAE	658.56	n.d.
B14–B29/A18–A21	1377.56	1377.31
B22–B29–AAK–A1–A4	1683.97	1683.95
AEPK–B1–B13/A5–A17	3394.88	3394.45
Generated by enzymic processing of (E$_2$GEPK–B1–B29–AAK–A1–A21) IP (yAK1026)		
EE	276.24	n.d.
B14–B29/A18–A21	1377.56	1377.38
B22–B29–AAK–A1–A4	1683.97	1683.97
GEPK–B1–B13/A5–A17	3080.85	3380.56

approaches are possible to limit the proteolytic processing of the N-terminal extension in *P. pastoris*: changing the pH in the medium, using a protease-deficient *P. pastoris* strain or modifying the spacer peptide. The spacer peptide EEGEPK has previously been shown to function well in *S. cerevisiae* [20]. This spacer/N-terminal extension also increased the fermentation yield of the IP expressed in *P. pastoris* (yAK1026 and yAK1031) and was resistant to the observed proteolytic processing of the IP (Tables 1 and 2). Electrospray MS analysis of selected purified IPs expressed in *P. pastoris* (Table 3) confirmed the results obtained by MALDI–MS with the culture supernatants directly (results not shown).

Characterization of IP expressed in *P. pastoris*

Selected IPs were characterized by peptide mapping and on-line RP-HPLC–MS to evaluate the structure of the IP expressed in *P. pastoris*. Digestion of IPs, with correct localization of the three disulphide bridges, with *Staph. aureus* V8 protease will result in three major peptides (fragments I, II and III) and one or two minor peptides depending on the specific IP. The three disulphide bridges in the IP will be localized in fragment I and III (shown in Table 3). The experimentally determined molecular masses of the peptides obtained from digestion with *Staph. aureus* V8 of the secreted, purified and endoprotease-processed IPs derived from the yAK1002, yAK1022, yAK1024 and yAK1026 constructs are shown in Table 3. The molecular masses of fragments I–III derived from *Staph. aureus* V8 digested IPs were all similar to the theoretical molecular masses (Table 3). Fragment III has two disulphide bridges and the position of these cannot be definitely determined by this method. However, fragment III has a retention time on RP-HPLC identical to that of the corresponding fragment derived from human insulin (results not shown), supporting the notion that disulphide bridge formation in the IPs is identical to that in human insulin (A6 to A11 and A7 to B7). The similarity in the experimentally determined molecular masses and the theoretical molecular masses illustrates that both the α-factor prepro-leader and the synthetic prepro-leaders facilitate the secretion of IPs with correctly formed disulphide bridges expressed in *P. pastoris*. Similar results on the localization of the disulphide bridges of the IPs expressed in *S. cerevisiae* has been published previously [20,21].

P. pastoris has a capacity for expression and secretion of IP that is equal to or better than the capacity observed with *S. cerevisiae*. The *P. pastoris* expression system has the

© 1999 Portland Press Ltd

possibility of isolating high-yield clones with multiple copies of the gene. It is possible that large-scale screening will lead to the isolation of *P. pastoris* clones with a higher IP expression level than those reported here. A direct comparison between the two yeast species regarding secretory expression of the IP is obviously complex owing to the number of variables between the two expression systems. The *S. cerevisiae* expression system utilizes a constitutive promoter, whereas the *P. pastoris AOX1* promoter is tightly regulated by methanol. Methanol was, however, present during most of the fermentation of *P. pastoris*. The *S. cerevisiae* expression system used here is based on episomal vectors, whereas the *P. pastoris* expression system is based on integrating vectors. Presumably the copy number of the gene encoding the prepro-leader/IP fusion protein is different in the described *P. pastoris* strains and in *S. cerevisiae* strains. The fermentation yield of secreted heterologous proteins expressed in yeast is a result of not only gene transcription and mRNA level, but also the overall capacity of, as well as individual bottlenecks in, the secretory pathway. The prepro-leaders' ability to interact with the secretory pathway or to function as intramolecular chaperones also influence secretion and fermentation yield. The prepro-leaders differ in their ability to facilitate the secretion of the IP expressed in *P. pastoris* and in *S. cerevisiae*. Therefore both prepro-leader and yeast species should be considered when expressing heterologous proteins in yeast. The ability of the synthetic leaders to facilitate secretion of the IP in *P. pastoris* might be improved by increasing the number of amino acids, especially of negatively charged and hydrophilic amino acids, and by the insertion of additional consensus N-linked glycosylation sites. The pro-leader TA57, which lacks potential N-linked glycosylation sites, facilitates the secretion of the IP both in *P. pastoris* and in *S. cerevisiae*, indicating that carbohydrate is not necessary for secretion of the IP in yeast. *P. pastoris* and *S. cerevisiae* have similar overall features in their secretory expression of insulin; the IP fermentation yields obtained with various prepro-leaders and spacer peptides were similar under the growth conditions used. The expression of recombinant proteins in yeast generally uses the Kex2 endoprotease or an analogous protease to cleave a fusion protein proteolytically. One advantage of *P. pastoris* seems to be a Kex2 endoprotease analogue that is significantly more efficient than the *S. cerevisiae* Kex2 endoprotease in processing the pro-leader/IP fusion protein. The yeast species *P. pastoris* and *S. cerevisiae* are clearly both suited to the production of recombinant proteins of pharmaceutical importance, here exemplified by the secretory expression of the IP. An important aspect of any cellular production system is its easy and efficient adaptation to large-scale fermentation conditions used for the production of recombinant proteins. Clearly, the potential of *P. pastoris* for easy up-scaling and high cell density is a attractive feature in the production of recombinant proteins.

Acknowledgments

We thank Lene Drube, Kate Müggler, Tine Christensen, Lone Albrechtsen, Betina Greis and Vibeke Petersen for technical assistance, and Hanne Kronstrøm and Connie Eriksen for assistance with MALDI–MS analysis.

References

1 Heinisch, J. J. and Hollenberg, C. P. (1993) in Biotechology, 2nd edn. (Rehm, H.-J. and Reed, G., eds.), pp. 469–514, Verlag VCH, Weinheim

2 Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M. et al. (1996) Science **274**, 546–567

3 Weydemann, U., Keup, P., Piontek, M., Strasser, A. W., Schweden, J., Gellissen, G. and Janowicz, Z. A. (1995) Appl. Microbiol. Biotechnol. **44**, 377–385

4 Romanos, M. (1995) Curr. Opin. Biotechnol. **6**, 527–533

5 Faber, K. N., Harder, W., Ab, G. and Veenhuis, M. (1995) Yeast **11**, 1331–1344

6 Hollenberg, C. P. and Gellissen, G. (1997) Curr. Opin. Biotechnol. **8**, 554–560

7 Raymond, C. K., Bukowski, T., Holderman, S. D., Ching, A. F. T., Vanja, E. and Stamm, M. R. (1998) Yeast **14**, 11–23

8 Gellissen, G. and Hollenberg, C. P. (1997) Gene **190**, 87–97

9 Cregg, J. M., Vedvick, T. S. and Raschke, W. C. (1993) Bio/Technology **11**, 905–910

10 Scorer, C. A., Clare, J. J., McCombie, W. R., Romanos, M. A. and Sreekrishna, K. (1994) Bio/Technology **12**, 181–184

11 Romanos, M. A., Scorer, C. A. and Clare, J. J. (1992) Yeast **8**, 423–488

12 Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J.-T., Smith, P. L., Wierschke, J. D., Subramaniam, A. and Birkenberger, L. A. (1997) Gene **190**, 55–62

13 Kurjan, J. and Herskowitz, I. (1982) Cell **30**, 933–943

14 Markussen, J., Damgaard, U., Diers, I., Fiil, N., Hansen, M. T., Larsen, P., Norris, F., Norris, K., Schou, O., Snel, L. et al. (1987) in Peptides 1986 (Theodoropoulos, D., ed.), pp. 189–194, de Gruyter, Berlin

15 Thim, L., Hansen, M. T., Norris, K., Hoegh, I., Boel, E., Forstrom, J., Ammerer, G. and Fiil, N. P. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 6766–6770

16 Kjeldsen, T., Brandt, J., Andersen, A. S., Egel-Mitani, M., Hach, M., Pettersson, A. F. and Vad, K. (1996) Gene **170**, 107–112

17 Kjeldsen, T., Pettersson, A. F., Hach, M., Diers, I., Havelund, S., Hansen, P. H. and Andersen, A. S. (1997) Protein Express. Purif. **9**, 331–336

© 1999 Portland Press Ltd

18 Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. (1985) Mol. Cell Biol. **5**, 3376–3385
19 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20 Kjeldsen, T., Hach, M., Balschmidt, P., Havelund, S., Pettersson, A. F. and Markussen, J. (1999) Protein Express. Purif., in the press
21 Kjeldsen, T., Andersen, A. S., Hach, M., Diers, I., Nikolajsen, J. and Markussen, J. (1998) Biotechnol. Appl. Biochem. **27**, 109–115
22 Snel, L. and Damgaard, U. (1988) Horm. Metabol. Res. **20**, 476–488
23 Caplan, S., Green, R., Rocco, J. and Kurjan, J. (1991) J. Bacteriol. **173**, 627–635
24 Julius, D., Schekman, R. and Thorner, J. (1984) Cell **36**, 309–318
25 Clements, J. M., Catlin, G. H., Price, M. J. and Edwards, M. (1991) Gene **106**, 267–272
26 Laroche, Y., Storme, V., De Meutter, J., Messens, J. and Lauwereys, M. (1994) Bio/Technology **12**, 1119–1124

Received 28 October 1998; accepted 6 November 1998

© 1999 Portland Press Ltd