HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

New applications of heterologous gene expression in *Hansenula polymorpha* for protein and metabolite production

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Eva Eilert
aus Velbert

Düsseldorf, März 2013

aus den Instituten für Molekulare Mykologie und Mikrobiologie der Heinrich-Heine Universität Düsseldorf.

Die praktischen Arbeiten wurden in den Laboren der ARTES Biotechnology GmbH durchgeführt. Teile dieser Arbeit wurden durch das Bundesministerium für Bildung und Forschung (BMBF) unterstützt.

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf.

Referent:	Prof. Dr. J. F. Ernst & Prof. Dr. C. P. Hollenberg
Korreferent:	Prof. Dr. M. Bott
Externer Betreuer:	Dr. M. Suckow

Tag der mündlichen Prüfung: 7. Mai 2013

Fantasie haben heißt nicht, sich etwas ausdenken;
es heißt sich aus den Dingen etwas machen.

Thomas Mann in *Meerfahrt mit Don Quijote*

Abstract

The methylotrophic yeast *Hansenula polymorpha* has been established as a highly efficient expression system. Strong promoters derived from the methanol metabolism pathway and mitotically stable high-copy integration have made this organism very effective for heterologous protein expression.

To improve the processing of secretory proteins targeted by the MFα prepro sequence, variations in the amino acid composition near the processing site of Kex2p were investigated (Manuscript I). The insertion of Ala-Pro between MFα prepro sequence and protein of interest leads to an improved processing of the signal sequence in case of interleukin-6 (IL-6) and interferon-α 2a (IFNα-2a). Modeling of the transition sequences confirmed a better exposition of the Kex2p-cleavage site in dependency of Ala-Pro insertions. Analysis of the secreted proteins revealed activity of an Ala-Pro specific endoprotease. Partial removal of the N-terminal Ala-Pro dipeptide was also detected during IFNα-2a processing. The putative protease responsible for the processing of Ala-Pro dipeptides was identified by gene deletion. Significant similarities with *Saccharomyces cerevisiae* Ste13p were identified.

To generally improve productivity in the *H. polymorpha* system new fusion protein strategies were investigated (Manuscript II). The recombinantly strongly expressed cytosolic ferritin heavy chain (FTH1) was analyzed with regard to its potential as fusion partner for difficult-to-express target proteins. A fragment of the parathyroid hormone (PTH) was chosen to explore whether fusion to FTH1 enhances PTH stability. We could show that the expression level of this fusion protein was at least of the magnitude of that of FTH1 without fusion partner. The solubility of the fusion protein was increased in comparison to FTH1 alone and no degradation of PTH occurred when fused C-terminally to FTH1.

Ectoine and 5-hydroxyectoine belong to the family of compatible solutes which are known to contribute to the adaptation of the cell to osmotic stress by mediation of a constant turgor. *H. polymorpha* was engineered to express the enzymes of the *H. elongata* 5-hydroxyectoine biosynthesis pathway (Manuscript III). Expression of the enzymes EctB, EctA, EctC and EctD in *H. polymorpha* resulted in production of 5-hydroxyectoine in gram per liter scale with almost full conversion of ectoine to 5-hydroxyectoine. We showed that 5-hydroxyectoine synthesis was influenced by feeding of methanol during cultivation. Furthermore, 5-hydroxyectoine synthesis was downregulated by the addition of methionine, lysine or threonine indicating the existence of a feedback regulated aspartate kinase in *H. polymorpha*.

Zusammenfassung

Die methylotrophe Hefe *Hansenula polymorpha* hat sich als effizientes System zur heterologen Genexpression etabliert. Besondere Merkmale dieses Systems sind die starken Promotoren des Methanolstoffwechsels und die Fähigkeit zur mitotisch stabilen Integration einer Expressionskassette in einer hohen Kopienzahl.

Am Beispiel der MFα prepro Sequenz sollte die Prozessierung sekretorisch exprimierter Proteine verbessert werden (Manuskript I). Hierfür wurden Veränderungen in der Nähe der Kex2p Erkennungssequenz durchgeführt. Durch das Einfügen von Ala-Pro zwischen MFα prepro Sequenz und Zielprotein konnte die Prozessierung im Fall von Interleukin-6 (IL-6) und Interferon-α2a (IFNα-2a) verbessert werden. Eine Modellierung dieses Sequenzübergangs zeigte eine verbesserte Präsentation der Kex2p-Schnittstelle. Nach Analyse des sekretierten IL-6 wurde außerdem die Aktivität einer Ala-Pro spezifischen Endoprotease festgestellt. Auch im Falle von IFNα-2a konnte eine partielle Abspaltung des N-terminalen Ala-Pro-Dipeptids festgestellt werden. Ein Protease, welche mutmaßlich an der Prozessierung dieser Ala-Pro-Dipeptide beteiligt ist, wurde per Gendeletion identifiziert. Diese Protease wies signifikante Ähnlichkeiten mit der *S. cerevisiae* Ste13p auf.

Zur allgemeinen Verbesserung der Ausbeute bei der Produktion heterologer Proteine wurden neue Fusionsprotein-Strategien untersucht (Manuskript II). Hierfür wurde das rekombinant cytosolisch stark exprimierte Protein Ferritin (schwere Kette, FTH1) im Hinblick auf das Potenzial als Fusionspartner für schwierig zu exprimieren Zielproteine analysiert. Ein Fragment des Parathormons (PTH) wurde gewählt, um zu untersuchen, ob die Fusion beider Proteine die Stabilität von PTH erhöht. Wir konnten zeigen, dass die Expression dieses Fusionsproteins mindestens in der Größenordnung des FTH1 ohne Fusionspartner war. Die Löslichkeit des Fusionsproteins wurde gegenüber FTH1 allein erhöht und es trat kein Abbau des PTH in Fusion mit FTH1 auf.

Ectoin und 5-Hydroxyectoin gehören zur Familie der kompatiblen Solute, die hauptsächlich zur Anpassung der Zelle bei osmotischem Stress durch Anpassung des Turgors beitragen. Ein *H. polymorpha*-Stamm wurde entwickelt, welcher die Enzyme der 5-Hydroxyectoin-Biosynthese aus *H. elongata* exprimiert (Manuskript III). Die Expression der Enzyme EctB, EctA, EctC und EctD in *H. polymorpha* resultierte in der Produktion von 5-Hydroxyectoin im Gramm pro Liter Bereich. Es fand außerdem eine nahezu vollständige Umwandlung von Ectoin in 5-Hydroxyectoin statt. Wir konnten zeigen, dass die 5-Hydroxyectoin-Synthese durch Zugabe von Methanol während der Kultivierung beeinflusst wurde. Desweiteren wurde die 5-Hydroxyectoin-Synthese durch die Zugabe von Methionin, Lysin oder Threonin herunterreguliert, was andeutet, dass in *H. polymorpha* eine Feedback-regulierte Aspartatkinase existiert.

Contents

List of Abbreviations

aa ..... amino acid(s)
ADP ..... adenosine diphosphate
ATP ..... adenosine triphosphate
BiP ..... binding protein
Catp ..... catalase
dcw ..... dry cell weight
Dhasp ..... dihydroxyacetone synthase
e.g. ..... abbreviation Latin 'exempli gratia', for example
EctA ..... L-2,4-diaminobutyrate acetyltransferase
EctB ..... L-2,4-diaminobutyrate transaminase
EctC ..... ectoine synthase
EctD ..... ectoine hydroxylase
EDTA ..... ethylenediaminetetraacetic acid
Ek ..... enterokinase cleavage site
ER ..... endoplasmic reticulum
Fghp ..... S-formylglutathione hydrolase
Fldp ..... formaldehyde dehydrogenase
Fmdp ..... formate dehydrogenase
FTH1 ..... human ferritin heavy chain 1
g ..... gravity or gram
G3 ..... three glycine spacer
GPI ..... glycophosphatidylinositol
GST ..... glutathione S-transferase
GTP ..... guanosine triphosphate
h ..... hour(s)
HARS1 ..... *Hansenula* autonomous replication sequence
IFNα-2a ..... interferon-α 2a

IL-6 ..... interleukin-6
kbp ..... kilo base pairs
MeOH ..... methanol
MES ..... 2-(N-morpholino)ethanesulfonic acid
MFα ..... mating factor α
min ..... minute(s)
Moxp ..... methanol oxidase
NADH ..... nicotinamide adenine dinucleotide
ORF ..... open reading frame
PCR ..... polymerase chain reaction
PTH ..... human parathyroid hormone, N-terminal fragment (amino acids 1-34)
PTS ..... peroxisomal targeting sequence
Seb1p ..... Sec61 complex subunit beta
SRP ..... signal recognition particle
SUMO ..... small ubiquitin-like modifier protein
SYN6 ..... synthetic minimal medium
YNB ..... yeast nitrogen base

1 General Introduction

1.1 The expression system *Hansenula polymorpha*

Heterologous gene expression is often performed using prokaryotic expression systems for which *Escherichia coli* is the most prominent example. Besides prokaryotic systems there are several microbial eukaryotic systems available which offer solutions that cannot be handled using a prokaryotic system. Production of proteins of pharmaceutical interest frequently requires posttranslational modifications such as glycosylation, acylation, phosphorylation or proper formation of hydrogen bonds which are typical for eukaryotes (van Dijk et al., 2000). Besides *Saccharomyces cerevisiae* other yeasts such as *Kluyveromyces lactis*, *Schizosaccharomyces pombe*, *Yarrowia lipolytica*, *Pichia pastoris* and *Hansenula polymorpha* have been established as efficient expression systems (Gellissen and Hollenberg, 1997; Müller et al., 1998). During the last two decades the methylotrophic yeast *H. polymorpha* has been successfully applied for industrial applications. Due to strong promoters derived from the methanol (MeOH) metabolism pathway and the mitotically stable high-copy integration, effective protein expression could be shown for numerous applications (for example Janowicz et al., 1991; Mayer et al., 1999; Weydemann et al., 1995) including whole-cell biocatalysis (Gellissen et al., 1996).

1.1.1 Promoters derived from the methanol utilization pathway

MeOH utilization was only known in bacteria until Ogata et al. (1969) were the first to describe a yeast capable to utilize MeOH as a sole carbon source. Since then, many other yeasts were identified sharing this ability, amongst them *P. pastoris*, *Candida boidinii*, *Pichia methanolica* and *H. polymorpha*. MeOH utilization was studied extensively and the key enzymes were identified. The most essential steps of the MeOH metabolism are summarized in Fig. 1.1.
MeOH utilization takes place in two compartments of the cell: peroxisomes and cytosol. When MeOH enters the cell it is oxidized to formaldehyde and hydrogen peroxide by methanol oxidase (Moxp) in the peroxisomes. The hydrogen peroxide is subsequently detoxified by catalase (Catp) which is also localized in the peroxisome. Alongside assimilation and dissimilation of formaldehyde occurs. In the dissimilation pathway formaldehyde is converted to CO_2 by the enzymes formaldehyde dehydrogenase (Fldp), S-formylglutathione hydrolase (Fghp) and formate dehydrogenase (Fmdp). These latter steps all take place in the cytosol. During the assimilation pathway formaldehyde is converted to dihydroxyacetone and glyceraldehyde 3-phosphate by the enzyme dihydroxyacetone synthase (Dhasp) in the

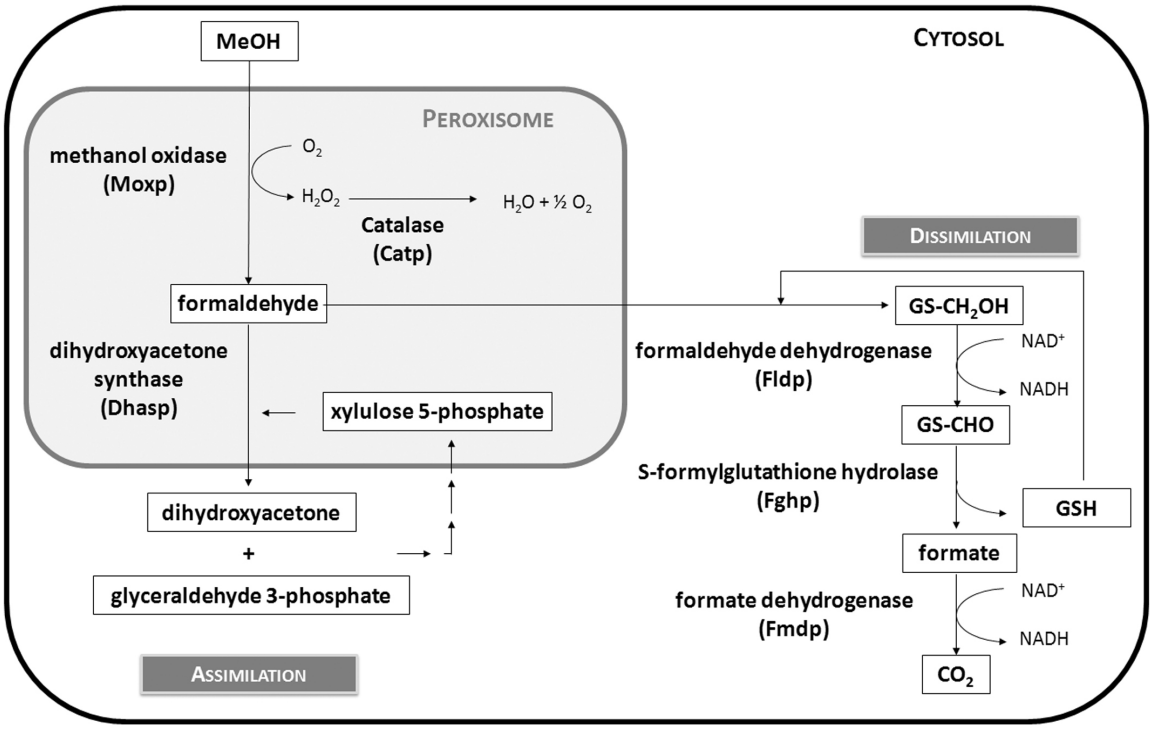


Figure 1.1: **Schematic illustration of the methanol utilization pathway in *H. polymorpha* with regard to the major participating enzymes**

peroxisome. Dihydroxyacetone and glyceraldehyde 3-phosphate subsequently take part in the xylulose 5-phosphate regeneration (for a recent review of the MeOH utilization see Yurimoto et al., 2011). Moxp, Fmdp and Dhasp are considered the key enzymes of MeOH utilization since the are responsible for utilization of MeOH and the conversion products formaldehyde and formate. In presence of MeOH the expression of these three enzymes is highly upregulated at the transcript level (Moxp 17.3-fold up, Fmdp 347-fold up, Dhasp 19-fold up)(van Zutphen et al., 2010; Gödecke et al., 1994). Several strong promoters responsible for this upregulation are employed for expression of heterologous genes in methylotrophic yeasts. The *H. polymorpha* promoters derived from the *MOX* (Ledeboer et al., 1985) and *FMD* genes (Hollenberg and Janowicz, 1995) have been characterized in detail. In other yeasts the use of *FMD* promoter is rather rare and has been solely described for *C. boidinii* (Sakai et al., 1997).

Regulation of these promoters is dependent on the carbon source used during cultivation. Glucose and ethanol are both repressing carbon sources. Contrarily both promoters are strongly inducible with MeOH (Eggeling and Sahm, 1978) and, unlike in *P. pastoris*, in *H. polymorpha* are not repressed using carbon sources like glycerol, ribose, xylose or sorbitol (Eggeling and Sahm, 1978). Using these carbon sources results in derepression which is also the case during cultivation under glucose starvation (Egli et al., 1980). During derepression genes under the control of the *MOX* promoter or *FMD* promoter are strongly transcribed, however the transcription level is lower in comparison to MeOH induction (Eggeling and Sahm, 1980). MeOH induction can be combined with cultivation under derepressing conditions, resulting in higher cell densities with high protein expression yields.

1.1.2 Genomic integration of heterologous genes

Apart from the choice of the promoter, the expression of a heterologous gene can be influenced by several other factors. After integration of a gene of interest the position in the genome, the copy number and rearrangements which might occur during the integration process can have a major influence on expression yield (Agaphonov et al., 1995).
For strain generation, plasmids harboring the *HARS1* sequence (*Hansenula* autonomous replication sequence) are used (Roggenkamp et al., 1986). After transformation the plasmid is maintained episomally only for a limited time and integrates into the genome spontaneously in high frequency (Janowicz et al., 1991). The integration occurs in multi-copy tandem repeats with head-to-tail arrangement (Roggenkamp et al., 1986). Supertransformation of a strain with the same expression cassette that has been introduced earlier can result in up to 100 copies per haploid genome (Guengerich et al., 2004). The resulting strains are of high mitotic stability which have been proven exemplary for several strains (Guengerich et al., 2004). This kind of stability is required for a consistent production processes (Böer et al., 2007).

1.1.3 Protein production and biotransformation

In general, heterologous gene expression pursues two goals:

1. Generation of purified proteins such as pharmaceutical proteins or enzymes.
2. Production of a biocatalyst which usually uses the whole cell.

For both goals either intracellular or secretory targeting of the protein(s) of interest is optional. In most cases secretion of the protein of interest is favorable for the production of purified enzymes. *H. polymorpha* does not secrete significant amounts of homologous proteins, thus purification of the desired protein from the culture supernatant only requires limited downstream processing steps. Secretion of proteins can be limited by several factors which will be addressed in the following sections. Recognition of a signal peptide may be inefficient or processing may be incomplete due to limitations within the secretory pathway. Even after successful secretion of the protein of interest degradation due to proteolysis or instability may occur. In some cases glycosylation of a protein is undesired and intracellular protein production may be an alternative.
Apart from protein production, recombinant host strains have been used for biotransformation processes. Purified enzymes can be used as biocatalysts and can be produced as described above. For metabolic engineering or the application of whole cells, proteins typically need to be expressed within the cell or immobilized to the outer cell wall. In the following targeting to extracellular and intracellular compartments as applied in *H. polymorpha* is described in more detail.

1.2 Processing of secretory proteins

1.2.1 Secretion in eukaryotic systems

In eukaryotic systems different two modes of translocation are found: co-translational and post-translational translocation (for reviews see Rapoport, 2007; Park and Rapoport, 2012).

During co-translational translocation, the strongly hydrophobic signal peptide mediates binding of the nascent polypeptide and the ribosome to the signal recognition particle (SRP). During the attachment to the SRP translation is interrupted. Another domain of the SRP binds to SRP-receptor which is an ER-membrane protein. The ribosome-protein-complex interacts with the translocon-pore and the nascent polypeptide is inserted into the channel. Translation is continued as the polypeptide is transported through the translocon-pore. The signal peptide is removed within the lumen of the ER by the signal peptidase and the polypeptide is entirely translated while passing the translocon-pore. Folding of the protein is induced by binding to the protein binding protein (BiP) on ER-translocation. During post-translational translocation, a moderately hydrophobic signal peptide mediates binding of the fully translated polypeptide to the translocon-pore by a so far unknown mechanism. After insertion of the polypeptide all bound chaperones are released from the polypeptide. Subsequently the translocation into the lumen of the ER occurs, which is conducted by the binding of BiP. The attachment of BiP prevents any backwards movement of the peptide in direction of the cytosol. Upon on movement towards the ER lumen further BiP molecules bind to the peptide. Finally the BiP molecules are released under exchange of ADP for ATP. Again the signal peptide is removed.

After post-translational modification such as glycosylation the folded protein is incorporated in vesicles and then transported to the Golgi. Further processing of the protein occurs within the Golgi. Besides additional glycosylation, disulphide bond formation, glycosyl-phosphatidyl-inositol addition also processing of signal sequences may continue in this compartment (for review see Idiris et al., 2010). During the last step of the secretory pathway the mature protein is transported to the cell surface in Golgi vesicles and is then released to the extracellular space or other cell compartments (for review see Nakano and Luini, 2010).

1.2.2 Application of signal sequences for secretion of heterologous proteins in *H. polymorpha*

In general, a signal sequence of a homologous protein which is efficiently secreted is favorable for protein secretion. *H. polymorpha* wild-type strains do not secrete significant amounts of homologous proteins. The signal sequence for secretion from homologous repressible acid phosphatase (Pho1p) was identified, but has not yet been tested (van Dijk et al., 2000).

Even though the secretory pathway is highly conserved among eukaryotes, not all heterologous signal sequences are functional in *H. polymorpha*. Different signal sequences were tested for heterologous

protein secretion in *H. polymorpha*. Agaphonov et al. (1995) evaluated the killer toxin signal sequence derived from *K. lactis*, the glucoamylase signal sequence derived from *Aspergillus niger* and an optimized signal sequence derived from human lysozyme. Except for the glucoamylase signal sequence the other sequences resulted in correctly secreted protein. The MFα prepro sequence derived from *S. cerevisiae* remains the most commonly used signal sequence.

1.2.3 Secretory targeting using the MFα prepro sequence derived from *S. cerevisiae*

The *S. cerevisiae* α-mating pheromone precursor is processed during passing of the ER and Golgi and four mature α-factor molecules per precursor are then released from the cell (Tanaka et al., 1977; Kurjan and Herskowitz, 1982). The processing of the α-factor mating pheromone precursor is shown in Figure 1.2 A.

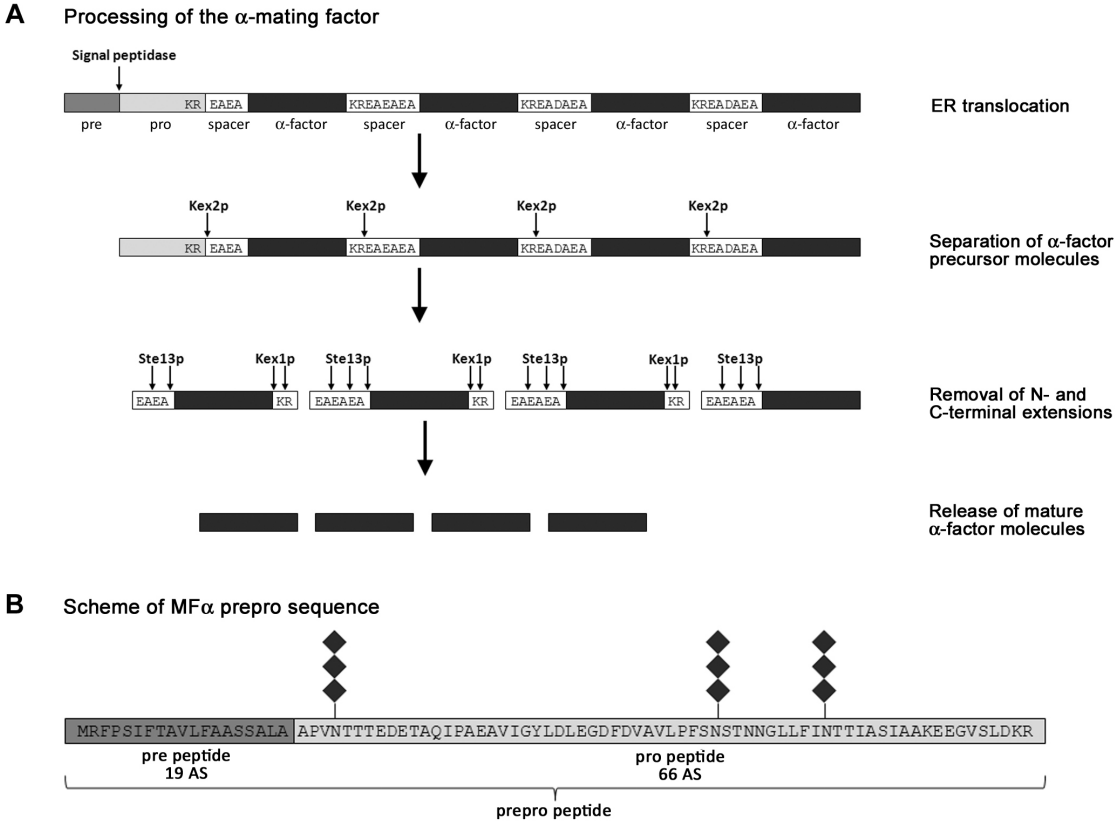


Figure 1.2: **Scheme of the α-factor mating pheromone derived from *S. cerevisiae*.** (A) Processing of the α-factor mating pheromone precursor during passage of the secretory pathway. (B) MFα prepro sequence, glycosylation sites are indicated at the conserved Asn sites.

The N-terminus consists of a pre- and a pro-sequence and comprises 85 amino acids (compare Figure 1.2 B). The first 19 amino acids are the pre-sequence. This region is rather hydrophobic. It is followed by the 66 amino acids comprising pro-sequence. This region is rather hydrophilic (Fuller et al., 1988).

Translocation of the polypeptide occurs posttranslational (Meyer, 1988). The pre-sequence is removed upon ER-translocation by the signal peptidase. Three *N*-linked core oligosaccharides are added to the pro-sequence during ER-passage (Fuller et al., 1988). Later on hypermannosylation occurs during passage of the late Golgi (Julius et al., 1984a). The C-terminus of the MFα prepro sequence and the spacers are carrying a dibasic Lys-Arg motif which is recognized by the Kex2 protease. The Kex2p also mediates removal of the pro-sequence (Buckholz, 1993; Julius et al., 1984b). Removal of the pre-sequence may occur upon ER-translocation (Waters et al., 1988). During α-mating pheromone processing the remaining C-terminal Lys-Arg residues are removed by the Kex1 protease (Dmochowska et al., 1987). Further modifications of the N-terminus are conducted by the Ste13p protease which removes the redundant dipeptides (Brenner and Fuller, 1992).The MFα prepro sequence has been widely applied for the expression of heterologous proteins in *S. cerevisiae* (Brake et al., 1984) and in various other yeasts (*P. pastoris* (Clare et al., 1991),*K. lactis* (Chen et al., 1992), *Zygosaccharomyces bailii* (Porro et al., 2005), *Ogataea minuta* (Akeboshi et al., 2007) and *H. polymorpha* (Weydemann et al., 1995)). High yields and an authentic N-terminus of the protein of interest resulting from the processing of the MFα prepro sequence are major advantages of the application of this prepro sequence.

1.2.4 Modifications for improvement of secretion

For secretion of a heterologous protein different aspects need to be considered. Besides the reasonably simple synthesis of the polypeptide many co-factors are needed for co- or post-translational translocation to the ER. Furthermore proper protein folding and the subsequent quality control inside the ER, glycosylation in the ER and Golgi, sorting and further processing are required. Hence genetic modification of production strains with regard to elimination of bottle-necks in the secretory pathway may enhance productivity (Idiris et al., 2010).
Besides the *S. cerevisiae* wildtype sequence, different variants of the MFα prepro sequence have been used for protein secretion. Ghosalkar et al. (2008) applied an altered variant of the MFα prepro sequence with an amino acid exchange at position 83 (aspartate exchanged for glutamate, `D83E`) resulting in increased secretion of correctly processed interferon-α 2b. During overexpression of homologous and heterologous proteins, processing of the MFα prepro sequence sometimes is not complete. To overcome this problem an additional Kex2p construct can be retained in the ER via fusion to the `KDEL` retention signal. This retention can help to process and subsequently to secrete heterologous proteins which are otherwise stuck in the ER and can only be processed by Kex2p *in vitro* (Chaudhuri et al., 1992). Incomplete of the Kex2p processing may result in a mixture of correctly as well as unprocessed forms of the protein (Seeboth and Heim, 1991). In addition Kex2p can be expressed in a soluble form by truncation of 200 amino acids at the C-terminus. The protein is efficiently secreted and retains full activity (Brenner and Fuller, 1992; Seeboth and Heim, 1991; Rolf, 2011). This soluble Kex2p can be used for *in vitro* processing of mis- or unprocessed isoforms of secreted proteins.
Often Ste13p activity is not sufficient for complete processing of the Glu-Ala-spacer and extended forms are released from the cells (Brake et al., 1984). Limitation of processing by Ste13p was also

shown by Brenner and Fuller (1992) during Kex2p overexpression. In this case remaining Leu-Pro and Val-Pro dipeptides indicated incomplete processing of the N-terminus. To overcome this bottleneck the Glu-Ala-spacer was often discarded and the protein of interest was directly fused to the C-terminus of the Lys-Arg-motif. It was shown that the removal of the Glu-Ala-spacer has no effect on the secretion efficiency (Zsebo et al., 1986).

1.3 Targeting to intracellular compartments

In general, targeting of an heterogeneously expressed protein to different cell compartments is possible. Signal sequences for mitochondrial, peroxisomal and vacuolar targeting as well as for ER-retention have been identified and tested. In addition proteins can be targeted to the cell wall and thus be accessible from the outside.

1.3.1 Intracellular protein production

Peroxisome

In *H. polymorpha* peroxisomal targeting has been studied most extensively. Three trafficking variants have been described for peroxisomal proteins: PTS type I, PTS type II and Non-PTS proteins. Dhasp was identified as a PTS type I protein in *H. polymorpha* (Hansen et al., 1992). It harbors a Asn-Lys-Leu (`NKL`) sequence at the C-terminus (peroxisomal targeting sequence I, PTS1) which shows high similarity to the commonly found PTS1 sequence Ser-Lys-Leu (`SKL`) first identified for firefly luciferase (Gould et al., 1989). In addition to this consensus sequence the residues of 9 adjacent amino acids are involved in interaction of the cargo protein with the soluble protein HpPex5p. This complex targets towards the peroxisomal membrane and associates with resident membrane proteins. Subsequently pore formation is mediated by HpPex14p and HpPex5p resulting in the translocation of the cargo protein (Rucktäschel et al., 2011). Another example for PTS type I proteins is the Catp (Williams et al., 2012). It harbors a Ser-Lys-Ile (`SKI`) sequence which also shows similarity to the more commonly found `SKL` sequence.
Moxp was identified as a PTS type II protein (Hansen et al., 1992). It harbors a Ala-Arg-Phe (`ARF`) sequence at the C-terminus (peroxisomal targeting sequence II, PTS2). PTS2 is recognized by the soluble protein Pex7p. Import of the cargo protein into the peroxisome requires HpPex20p which is acting as a co-receptor (Otzen et al., 2005; Rucktäschel et al., 2011). The subsequent import mechanism has not yet been fully identified, however HpPex14p may be involved in pore formation (de Vries et al., 2007). For functionality, the targeting signals PTS1 and PTS2 need to be positioned at the extreme C-terminus (Hansen et al., 1992; Faber et al., 1993; Gould et al., 1989).
Amine oxidase is an example for non-PTS proteins. Non-PTS proteins are defined to neither contain PTS1 nor PTS2 (Gunkel et al., 2004). In case of amine oxidase the mechanism for translocation to the peroxisome has not yet been identified. There are several possibilities how PTS-independent

targeting may occur. One mechanism has been described for Moxp, which indeed is a PTS type I protein, but can also be imported as a non-PTS protein. It was shown that Moxp interacts with the N-terminal region of HpPex5p (Gunkel et al., 2004). However this binding occurs independent from PTS1 and the PTS-recognition site of HpPex5p which is located in the C-terminal domain.
Only few approaches have been published on fusion proteins targeted to an intracellular compartment. Amine oxidase was studied as an intracellular carrier protein for targeting to the peroxisome. Faber et al. (1996) fused the gene encoding for IGF-II to the C-terminus of the amine oxidase. The fusion protein was successfully expressed, the amount detected exceeded 20% of the total cellular protein. Furthermore IGF-II was found to be more stable when accumulated in the peroxisome.

Vacuole

Vacuolar targeting was studied by characterization of the vacuolar protein carboxypeptidase Y (Bellu et al., 1999). The carboxypeptidase Y harbors a N-terminal targeting sequence (Gln-His-Pro-Leu, `QHPL`). The `QXPX` motif seems to be conserved for several yeasts since it can also be found in *P. pastoris* and *S. cerevisiae* carboxypeptidase Y homologues. Yet, this import mechanism has not been used for industrial applications.

Mitochondria

The import of mitochondrial proteins to different compartments has been analyzed in detail but not been used for protein production. Sorting to mitochondria was observed for various mitochondrial alcohol dehydrogenases in yeast. Suwannarangsee et al. (2012) characterized the *H. polymorpha* alcohol dehydrogenase 3. The targeting signal was predicted to be located in the N-terminal region. The first 28 amino acids are removed upon translocation. A recent review on mitochondrial protein import mechanism was provided by Schmidt et al. (2010).

Golgi and ER

The sequences Lys-Asp-Glu-Leu (`KDEL`) (Munro and Pelham, 1987) and His-Asp-Glu-Leu (`HDEL`) (Pelham et al., 1988) mediate the retention of soluble proteins in the ER. A `HDEL` signal sequence was detected at the extreme C-terminus of HpBiPp (van der Heide et al., 2002). Agaphonov et al. (2005) showed that protein-O-mannosyltransferase (HpPmt1p) has a `KDEL` signal at the extreme C-terminus. Since HpPMT1p is an integral membrane protein `KDEL` also mediates retention of membrane proteins. Ste13p and Kex2p are two well characterized proteins which are located in the Golgi. Ste13p is retained by Phe-Xaa-Phe-Xaa-Asp (`FXFXD`) in the cytoplasmic domain (Nothwehr et al., 1993). A second signal (TLS2) in the cytoplasmic domain delays the transport to the *trans*-Golgi Network and the prevacuolar compartment (Bryant and Stevens, 1997). A similar mechanism applies for retention of Kex2p, again two retention signals are located in the cytoplasmic domain (TLS1 and TLS2) (Brickner and Fuller, 1997).

Surface display techniques

Peptides and proteins can be targeted to the cell-surface by fusion to an anchoring motif which is derived from a homologous membrane protein (for review see Lee et al., 2003). The protein of interest is transported to the cell-surface were it is linked to the membrane via the anchoring motif. Positioning of the anchoring motif at N- or C-terminus may significantly influence activity and accessibility of the protein of interest. The application of glycophosphatidylinositol(GPI)-anchored proteins (or parts thereof) to display heterologous proteins is the most popular technique for surface display in yeast and was first described by Schreuder et al. (1993). There are several *H. polymorpha*-derived anchor proteins which find application in this yeast or have been evaluated using *P. pastoris* strains. The *H. polymorpha*-derived proteins HpCwp1p and HpSed1p are both GPI-anchored cell wall proteins which were successfully used for cell-surface display (Kim et al., 2002, 2007; Lin et al., 2012).

1.3.2 Fusion proteins

Also the fusion of a protein of interest to another functional protein has been conducted. The targeting signal of the carrier protein may be sufficient to achieve intra- or extracellular targeting for the fusion protein. The stability of a protein is influenced by several factors such as pH and media components, degradation by proteases and polypeptide composition and also may be enhanced when fused to the carrier protein. Varshavsky (1997) could show that a protein's *in vivo* half-life is strongly influenced by the N-terminal residues. A low level expression of the protein of interest may also be improved when fused to a highly expressed carrier protein. Furthermore, solubility may be improved (La Vallie and Riggs, 1994). However the authentic N- or C-terminus the protein of interest needs to be recovered from the fusion protein. Removal of the carrier protein can be conducted with a site-specific protease (Terpe, 2003).

1.4 Biotransformation

Starting in the 1970s, organic chemists found biotransformation processes to be beneficial for certain problems which arise with the synthesis of organic compounds (Yamada and Shimizu, 1988). Application of enzymes for the conversion of a precursor to fine chemicals, building blocks or pharmaceuticals has become an attractive technique and alternative to organic synthesis. Due to high specificities of enzyme's substrate-specific regio- and stereo-selective reactions can be performed (Straathof et al., 2002). It is even possible to replace a multi-step chemical synthesis with just one biotransformation process (for review see Wohlgemuth, 2010). Highest activity of the enzymes is most commonly achieved by maintaining physiological temperature and pressure as well as neutral pH (Yamada and Shimizu, 1988). In contrast to organic chemistry, the use of organic solvents is limited and only required in rare cases.

To further reduce costs, which arise during the production and purification of the desired enzymes, the application of whole cells as biocatalysts has been extensively evaluated. The conversion of a substrate to a product using microorganisms as a catalyst is consequently called microbial transformation. The synthesis of a product from carbon and nitrogen sources using growing microorganisms is called fermentation.

1.4.1 Microbial transformation

For microbial transformation either cell extracts or whole cells can be applied (Fig. 1.3). In case of whole-cell biocatalysis proper permeability for both substrate and product are essential. Intracellular expression of an enzyme for a biocatalytic application may result in transport limitations of substrate and product. Vitality of the cells is negligible after the enzymes have been properly produced, hence, the application of toxic substrates may be possible. Though some transformation processes require co-factor regeneration, it was shown for *S. cerevisiae* that this process may be independent from cell growth (Chin-Joe et al., 2001). To overcome bottle-necks concerning membrane permeability cells can be permeabilized or they can be engineered to display the enzyme(s) of interest on the cell-surface.

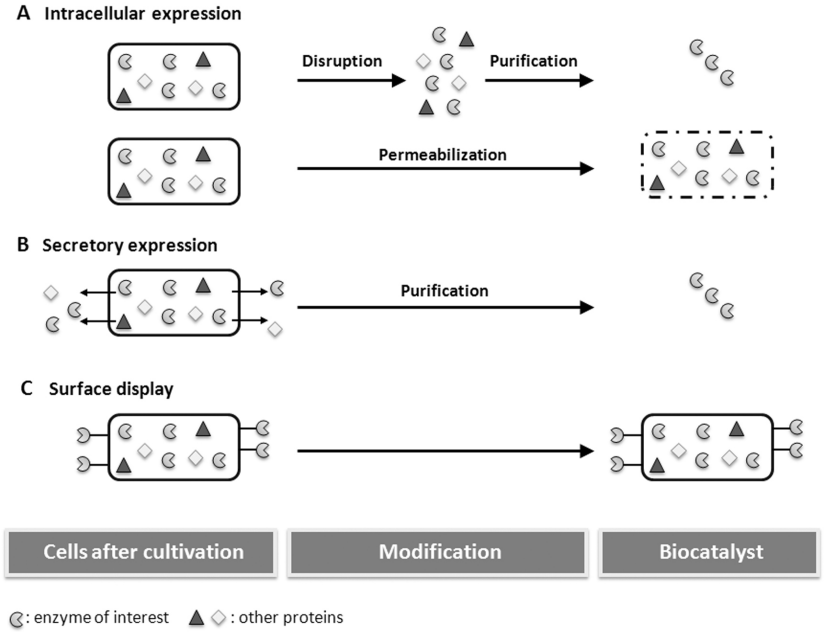


Figure 1.3: **Generation of biocatalysts on basis or intracellular, secretory of cell wall-targeted protein expression.** Cells are generated through cultivation of strains expressing the desired enzyme(s). Subsequently modification or purification of the enzyme(s) or cells is performed if needed. The resulting biocatalysts then are either purified enzymes or whole cells. Furthermore, immobilization of either biocatalyst my be performed subsequently.

Permeabilization

Permeability of the cell wall may be enhanced via application of various permeabilization techniques (for review see Chen, 2007). Treatment of *P. pastoris* cells with benzalkonium chloride or isopropanol

resulted efficient permeabilization and thus bioconversion (Gough et al., 2001; Liu et al., 2000). A more recent example is the generation of a D-amino acid oxidase *P. pastoris* expression strain for which the conversion during whole-cell biocatalysis was optimized via permeabilization using isopropanol (Abad et al., 2010). *H. polymorpha* was also proven to be a suitable host for whole-cell biocatalysis coexpressing spinach glycolate oxidase and *S. cerevisiae*-derived catalase T (Gellissen et al., 1996). After generation of biomass the cells were permeabilized with isobutyric acid. Subsequently the cells were used for the conversion of glycolate to glyoxylic acid catalyzed by the glycolate oxidase while the peroxide formed during this conversion was detoxified by the catalase.

1.4.2 Fermentation - Pathway engineering for metabolite production

The *de novo* production of biologicals from carbon and nitrogen sources is the most beneficial approach since no cost-intensive intermediate has to be provided for conversion. Engineering of a metabolic pathway aims at the enhanced production of an already existing metabolite or production of a new compound. Four main aspects are addressed by metabolic engineering (Kern et al., 2007):

- enhancement of productivity of a homologous producer organism
- reduction of side-products which either decrease the overall yield or complicate purification processes
- expansion of substrate range for more economically processes
- introduction of a new biosynthesis pathway

Corynebacterium glutamicum is the most prominent example of improvement of already existing metabolic pathways. Wildtype strains have been extensively used for production of the amino acid L-glutamate and have been further modified to improve the yield. In addition the organism has been exploited for the industrial production of various other amino acids (L-lysine, L-valine, L-isoleucine, L-aspartate and L-alanine among others (for review see Hermann, 2003)).

In contrast to *C. glutamicum* other microorganisms, amongst them various yeasts, were engineered to express enzymes of a complete heterologous pathway to synthesize a new product. Nguyen and Nevoigt (2009) were able to have *S. cerevisiae* produce dihydroxyacetone from glycerol by expression of glycerol dehydrogenase from *H. polymorpha*. Glycerol is synthesized by *S. cerevisiae* from glucose under osmotic stress conditions. The synthesized dihydroxyacetone was protected from further conversions by the deletion of two dihydroxyacetone kinase genes. In a different study, *H. polymorpha* was engineered to produce penicillin from the precursors L-cystein, L-valine and α-aminoadipic acid. The three enzymes involved in this conversion where overexpressed and penicillin was efficiently produced and secreted into the medium (Gidijala et al., 2009).

Insufficient co-factor supply can be overcome by regeneration systems. In *P. pastoris* NADH regeneration was enhanced for whole-cell biocatalysis or bioconversion (Schroer et al., 2010). NADH

regeneration was coupled to the MeOH utilization where two molecules of NADH are formed during dissimilation of one molecule MeOH.

1.4.3 Combining two techniques: microbial transformation and fermentation

For efficient generation of the desired product microbial transformation and fermentation are often combined. In this case one or more substrates or co-factors are provided by supplementation of the culture medium, whereas the other part of substrates is provided by the cell. The conversion of pyruvate and benzaldehyde to L-phenylacetylcarbinol was achieved and optimized using *H. polymorpha* in a combination of microbial transformation and fermentation (Popa et al., 2007). L-phenylacetylcarbinol is a key intermediate for L-ephedrine synthesis. Pyruvate is an internal compound and is generated from glucose or another substrate whereas benzaldehyde is added to the culture broth.

1.4.4 Ectoine and 5-hydroxyectoine biosynthesis as an example of metabolic engineering

Ectoine and 5-hydroxyectoine belong to the family of compatible solutes which are known to stabilize the cell's essential functions by maintaining the stability of proteins under stress conditions like high salinity, heat or aridity. Hence, compatible solutes are synthesized under elevated saline conditions and then accumulate in the cell (Lippert and Galinski, 1992). The synthesis of ectoine takes place in three steps starting with L-aspartate-semialdehyde which is an intermediate of the biosynthesis pathway of lysine, methionine and isoleucine (Peters et al., 1990) (Fig. 1.4). L-aspartate-semialdehyde is converted by EctB to L-2,4-diaminobutyrate via transamination. Next the Nγ-acetyl-2,4-diaminobutyrate is formed acetylation catalyzed by EctA. In the last step ectoine is generated via condensation by EctC. An additional hydroxylation step catalyzed by EctD converts ectoine into 5-hydroxyectoine (Bursy et al., 2007).

Ectoine was first discovered in *Ectothiorhodospira halochloris* (Galinski et al., 1985), later more halophilic or halotolerant organisms with the ability to produce ectoine, or in addition 5-hydroxyectoine, were identified (Inbar and Lapidot, 1988). Among moderate halophilic bacteria, *Halomonas elongata* is considered to be an important model organism and has been studied extensively (García-Estepa et al., 2006). The synthesis of ectoine and 5-hydroxyectoine in *H. elongata* only occurs under a variety of stress conditions, e.g. salt (Wohlfarth et al., 1990).

Due to its chiral carbons ectoine and 5-hydroxyectoine can only be produced via biological approaches (Schiraldi et al., 2006). Lentzen and Schwarz (2006a) developed the so far most effective 'permanent-milking' process for ectoine production which finds application in industrial production. By using the genuine producer *H. elongata* yields of up to 20 gram of ectoine/L are achieved. 5-hydroxyectoine production is also possible by modification of the 'bacterial milking' process (Lentzen and Schwarz, 2006b). However, Seip et al. (2011) stated that even after optimization of the production process by increased salinity and elevated temperature, conversion of ectoine to 5-hydroxyectoine did not

exceed 50%. Just recently (Rodriguez-Moya et al., 2013) published a study where *Chromohalobacter salexigens* was engineered to achieve a conversion of ectoine to 5-hydroxyectoine of about 90%. In addition, *Marinococcus* M52 was developed for 5-hydroxyectoine production, however, due to several limitations this process did not yet find industrial application (Frings et al., 1995; Schiraldi et al., 2006).

E. coli has been utilized for the expression of the *ectABC* cluster derived from different species such as *Bacillus halodurans* (Anbu Rajan et al., 2008) or *C. salexigens* (Schubert et al., 2007), leading to yields of up to 6 gram of ectoine/L. Another efficient ectoine process reported using *E. coli* is 'bacterial milking' introduced by Sauer and Galinski (1998). The 5-hydroxyectoine biosynthesis was optimized recently using the homologous producer *Pseudomonas stutzeri* and recombinant *E. coli* expressing the *P. stutzeri ectABCD-ask* cluster (Seip et al., 2011). These approaches resulted in almost complete conversion of ectoine to 5-hydroxyectoine.

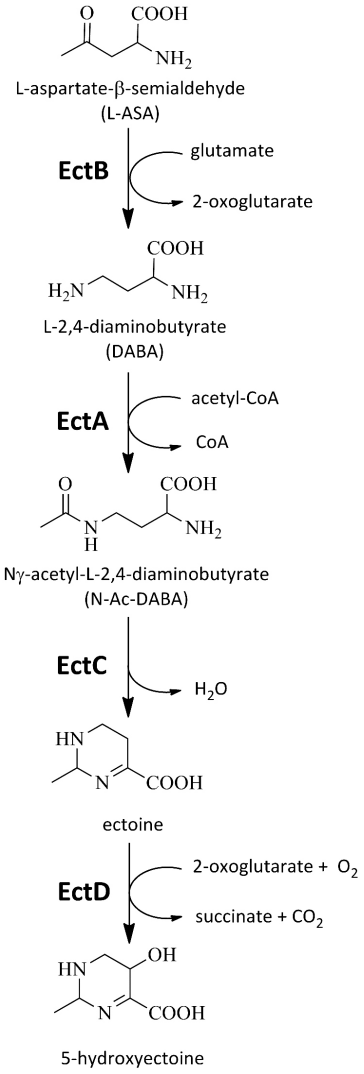


Figure 1.4: **Overview of the 5-hydroxyectoine biosynthetic pathway in *H. elongata*.** Conversion of L-aspartate-β-semialdehyde, a metabolite of the methionine, lysine and isoleucine biosynthesis, to the compatible solute 5-hydroxyectoine is catalyzed by the four enzymes EctB (L-2,4-diaminobutyrate transaminase), EctA (L-2,4-diaminobutyrate acetyltransferase), EctC (ectoine synthase) and EctD (ectoine hydroxylase) in the cytosol.

2 Aim of this study

The aim of this PhD thesis was to investigate new approaches on heterologous gene expression in *H. polymorpha*. The main goal was to improve extra- and intracellular protein expression and exploit new applications thereof. The study was subdivided into three parts:

Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence (Manuscript I)

To achieve secretory targeting of a protein of interest the MFα prepro sequence is most frequently used. However misprocessing of the signal sequence may occur and can cause cutbacks in the overall yield. Previous experiments showed that during expression of MFα prepro interferon-α 2a (IFNα-2a) partly misprocessing of the signal sequence occurred. Wendel (2009) could show that processing was improved by variation of the MFα prepro sequence via an exchange of one amino acid within the sequence. During the expression of interleukin-6 (IL-6) a similar misprocessing pattern was detected which could be eliminated by variation of the connection of the signal sequence to the protein. Improvement of the processing of IFNα-2a and IL-6 was target of optimization. To influence the processing behavior in *H. polymorpha* the accessibility of the Kex2p cleavage site was enhanced by introduction of additional amino acids between signal sequence and target protein. By modeling of the connection of the MFα prepro sequence to the protein of interest further insight into the processing behavior was obtained. The model was compared to experimental results. Furthermore a putative protease involved in the removal of Ala-Pro dipeptides was identified through generation of deletion mutants.

The use of highly expressed FTH1 as carrier protein for cytosolic targeting in *Hansenula polymorpha* (Manuscript II)

Human ferritin heavy chain (FTH1) was found to be highly expressed intracellularly in *H. polymorpha*. To evaluate FTH1 as a carrier protein for intracellular targeted fusion proteins, IFNα-2a was fused to FTH1 and expressed to provide a proof-of-concept. A human parathyroid hormone fragment (PTH) comprising of 34 amino acids has already been expressed and secreted in *H. polymorpha*, however it

was found to be highly unstable in the culture supernatant. Different approaches regarding protease deficient strains and medium optimization were of limited success (unpublished data). Therefore the FTH1 fusion was chosen to study the intracellular expression of PTH. The fusion protein was analyzed with regard to fermentation yield, stability and solubility.

Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha* (Manuscript III)

The experience of using *H. polymorpha* for pathway engineering and bioconversion is limited. First application of whole-cell biocatalysis using *H. polymorpha* was performed for the conversion of glycolate to glyoxylic acid (Gellissen et al., 1996). To further investigate the potential of *H. polymorpha* in this respect, the sequential four-step conversion of L-aspartate-semialdehyde to 5-hydroxyectoine as catalyzed by the four *H. elongata*-derived enzymes EctB, EctA, EctC and EctD was introduced. A stable whole cell biocatalyst which catalyzes the transformation of monomeric substrates to 5-hydroxyectoine was developed. The yields of ectoine and hydroxyectoine were analyzed and optimized.

3 Manuscript I: Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence

Author's contribution

Journal: Journal of Biotechnology
Impact Factor: 3.045
1^{st} author
Author's contribution: 85%

- Experimental design
- Cloning and strain generation
- Expression and protein analysis (except *in vitro* Kex2p treatment)
- Modeling of tertiary structures and analysis of cleavage site accessibility (together with Andreas Heumaier)
- Phylogenetic analysis
- Authoring manuscript

G Model
BIOTEC-6174; No. of Pages 7
ARTICLE IN PRESS

Journal of Biotechnology xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence

Eva Eilert[a], Theresa Rolf[a], Andreas Heumaier[b], Cornelis P. Hollenberg[a,c], Michael Piontek[a], Manfred Suckow[a,*]

[a] *ARTES Biotechnology GmbH, Elisabeth-Selbert-Str. 9, 40764 Langenfeld (Rheinland), Germany*
[b] *ZBH Zentrum für Bioinformatik, Universität Hamburg, Bundesstr. 43, 20146 Hamburg, Germany*
[c] *Institut für Mikrobiologie, Heinrich-Heine-Universität, Universitätsstr. 1, 40225 Düsseldorf, Germany*

ARTICLE INFO

Article history:
Received 15 May 2012
Received in revised form 27 August 2012
Accepted 31 August 2012
Available online xxx

Keywords:
Hansenula polymorpha
Interleukin-6
Interferon-α 2a
Alpha mating factor
Secretion
Processing
Signal sequence
Kex2p

ABSTRACT

The literature as well as databases are ambiguous about the exact start of human interleukin-6 (IL-6) – three possibilities for the initiation of the mature protein are described. These three variants of IL-6, different in the exact initiation of the mature protein (A28, P29, or V30), were expressed in *Hansenula polymorpha* using the *Saccharomyces cerevisiae* MFα prepro sequence instead of the homologous pre sequence. All three IL-6 variants were secreted but the processing by the Kex2 protease showed significant differences. V30-IL-6 showed correctly processed material but also a molecule species of higher molecular weight indicating incomplete processing of the MFα pro peptide. P29-IL-6 did not yield any correctly processed IL-6, instead only the unprocessed pro form was found in the culture supernatant. Only A28-IL-6 led to 100% correctly processed material. N-terminal sequencing of this material revealed a start at V30 – obviously the first two amino acids (Ala28-Pro29) have been removed by a so far unknown protease. Thus expression of both A28-IL-6 and V30-IL-6 as MFα prepro fusion proteins resulted in the very same mature V30-IL-6, however, the ratio of correctly processed molecules was significantly higher in the case of A28-IL-6.

The expression of an MFα prepro-interferon α-2a (IFNα-2a) fusion protein in *H. polymorpha* leads to about 50% correctly processed molecules and 50% misprocessed forms which contain part of the pro peptide at the N-termini. The insertion of A28 and P29 of IL-6 between the pro peptide and the start of the mature IFNα-2a led to correct processing and elimination of all high molecular weight isoforms observed in earlier experiments

1. Introduction

The MFα prepro sequence derived from *Saccharomyces cerevisiae* has been very successfully used for secretory expression of a wide variety of heterologous target proteins in various yeasts. It is derived from the α-mating factor precursor and is essential for the correct processing and secretion of pheromone. During translocation the pre sequence (residues 1–19) is cleaved off by the signal peptidase. During the passage through the ER and Golgi the pro sequence is glycosylated at three N-glycosylation sites near its C-terminus. In the late Golgi the Kex2 protease cleaves off the pro sequence (66 amino acids) and releases 4 α-factor precursor units (Julius et al., 1984). The C-terminal KR-residues are removed by the Kex1 carboxypeptidase (Wagner and Wolf, 1987) followed by the removal of the N-terminal Glu-Ala-Glu-Ala spacer by the action of the Ste13 dipeptidyl aminopeptidase A (Julius et al., 1983). With this last step the processing is completed, and mature α-factor becomes secreted. The MFα prepro sequence has been successfully used in heterologous gene expression amongst others in the following yeast derived expression systems: *S. cerevisiae* (Brake et al., 1984), *Pichia pastoris* (Clare et al., 1991), *Kluyveromyces lactis* (Chen et al., 1992), *Zygosaccharomyces bailii* (Porro et al., 2005), *Ogataea minuta* (Akeboshi et al., 2007) and *Hansenula polymorpha* (Weydemann et al., 1995). The finding that these yeasts are able to correctly process the MFα prepro sequence via the dibasic motif supports that Kex2p homologs exist in all these organisms.

Interleukin-6 (IL-6) is a multifunctional human protein which belongs to the group of cytokines. It has a pre sequence which in the natural host is cleaved off during ER uptake of the protein. However the exact transition between signal sequence and sequence of the mature IL-6 is unclear (Fig. 1A). Parekh et al. (1992) described the

Abbreviations: IL-6, interleukin-6; IFNα-2a, interferon-α 2a; IFNα-2b, interferon-α 2b; MFα, mating factor α.

* Corresponding author. Tel.: +49 0 2173 2758715; fax: +49 0 2173 2758777.
E-mail addresses: e.eilert@artes-biotechnology.com (E. Eilert), theresa.rolf@unitybox.de (T. Rolf), aheumaier@zbh.uni-hamburg.de (A. Heumaier), Cornelis.Hollenberg@uni-duesseldorf.de (C.P. Hollenberg), m.piontek@artes-biotechnology.com (M. Piontek), m.suckow@artes-biotechnology.com (M. Suckow).

0168-1656/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbiotec.2012.08.024

G Model
BIOTEC-6174; No. of Pages 7
ARTICLE IN PRESS
2 E. Eilert et al. / Journal of Biotechnology xxx (2012) xxx–xxx

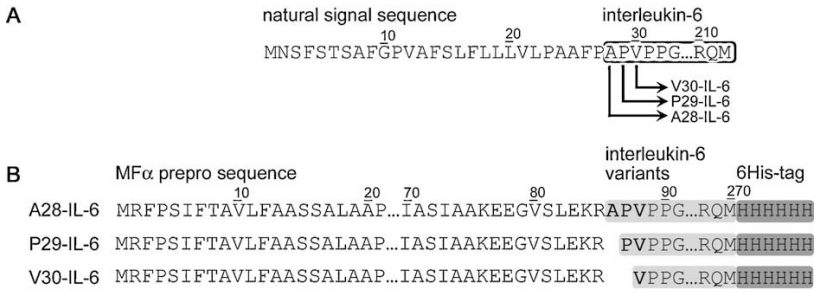


Fig. 1. (A) Transition from signal peptide to mature protein in the case of interleukin-6. The arrows indicate the three possible initiation amino acids described in the literature and the resulting nomenclature. (B) Due to the three potential N-terminal ends, three IL-6 variants were fused to the MFα prepro sequence for expression in *H. polymorpha*.

N-terminal sequence of human IL-6 to commence with the amino acids A28-P29-V30, but additionally they found two other sequence initiations in which one (A28) or two amino acids (A28-P29) were removed. Since the exact start also in the databases is ambiguous, three possibilities for the initiation of the mature protein need to be considered, which are referred to as V30-, P29-, or A28-IL-6 (mature IL-6 starting at the indicated amino acid residue, respectively).

IL-6 has been expressed in various yeast systems using the MFα prepro sequence. Guisez et al. (1991) expressed P29- and A28-IL-6 in *S. cerevisiae*. They observed different processing for each type. The P29-IL-6 showed incorrect processing of the pro peptide, whereas in the case of A28-IL-6 the first two amino acids A28 and P29 were removed. Steinborn et al. (2006) expressed V30-IL-6 in *S. cerevisiae* and observed N-terminal truncation of eight amino acids, however in *Arxula adeninivorans* they observed correct processing and the full length mature protein was secreted.

The methylotrophic yeast *H. polymorpha* has been established as a highly efficient expression system (Janowicz et al., 1991) amongst others due to the strong promoters derived from the methanol metabolism pathway and the possibility of mitotically stable high-copy integration. Degelmann et al. (2002) expressed the V30-IL-6 in *H. polymorpha*. Secretion was observed but no further sequence analysis was performed. Steinborn et al. (2006) also expressed the V30-IL-6 in a *H. polymorpha* system. These authors found an eight-amino-acid truncation of the N-terminus of the secreted molecules.

In this study we investigated the influence of three different start amino acids (A28, P29, and V30) on expression and processing of mature IL-6 using in all three cases the MFα prepro sequence as a signal and transport sequence. Furthermore, we applied the positive processing results from A28-IL-6 for the secretion of interferon α-2a.

2. Materials and methods

2.1. Strains and culture conditions

For construction and propagation of plasmid DNA *Escherichia coli* NEB10β (araD139 Δ(ara-leu)7697 fhuA lacX74 galK (ϕ80 Δ(lacZ)M15) mcrA galU recA1 endA1 nupG rpsL (Str^R) Δ(mrr-hsdRMS-mcrBC); New England Biolabs, Frankfurt a. M., Germany) was used. Cultivation took place in Luria Bertani media, when needed supplemented with 50 μg ampicillin per ml. For protein expression *H. polymorpha* strain RB11 deficient in orotidine-5′-phosphate decarboxylase (*ura3*) (Suckow and Gellissen, 2002) or KLA8-1 deficient in orotidine-5′-phosphate decarboxylase and β-isopropylmalate dehydrogenase (*ura3 leu2*) (Latchev et al., 2002) were used. Cultivation took place either in yeast extract-peptone-dextrose (YPD) medium or yeast nitrogen base (YNB) medium. All cultivations were performed at 37 °C.

2.2. Vector construction

To generate the expression plasmids pFPMT-MFaE-A28-IL-6, pFPMT-MFaE-P29-IL-6, and pFPMT-MFa-V30-IL-6, a synthetic open reading frame for each variant based on Universal Protein Resource Knowledgebase (UniProt ID: P05231) were ligated with the 6.8 kb *Hind*III/*Aat*II fragment of pFPMT-MFaE-IFNα-2a (Wendel, 2009). This plasmid contains all features of the vector pFPMT121 (Degelmann et al., 2002) and in addition the MFα prepro sequence with an amino acid exchange at position (Asp to Glu at position 83, D83E). A His-tag was added C-terminally to the IL-6 protein sequence. The cloned IL-6 sequences with additional flanking recognition sites (5′ *Hind*III and 3′*Aat*II) were synthesized by GeneArt (Regensburg, Germany).

To generate the expression plasmid pFPMT-MFaE-AP-IFNα-2a a PCR fragment was amplified with oligonucletide primers AP IFN hin (5′-AGGGGTAAGCTT GGAAAAGAGAGCTCCATGTGATTTGCCACAGACACACTCCCTGGGC) and AP IFN her (5′-GCGAGGGGGATCCTTATTACTCCTTCGATC) using plasmid pFPMT-MFaE-IFNalpha-2a as a template. The PCR product has been digested with *Hind*III/*Bam*HI resulting in a 0.523 kb fragment. This fragment has been ligated with the 7.217 kb *Hind*III/*Bam*HI fragment of pFPMT-MFaE-IFNalpha-2a. *H. polymorpha* was transformed by electroporation as described in Faber et al. (1994). The following strain generation was performed according to Guengerich et al. (2004). The passaging procedure lasted for four passages.

2.3. Protein expression and preparation of protein material

For heterologous protein expression recombinant strains were subjected to derepression/MeOH induction. Cells were cultivated in 3 ml scale in YNB/2% glycerol for 48 h and then for additional 24 h in YNB/1% methanol (37 °C). After cultivation the cells were harvested by centrifugation, and the supernatant was collected and stored at −20 °C.

2.4. Protein analysis

For molecular weight determination the samples were analyzed by SDS-PAGE according to Laemmli (1970). For the electrophoretic separation 4–12% Bis–Tris acrylamide XT gels (BioRad, Hercules, CA, USA) were used. The gels were stained either with the Coomassie stain SimplyBlue (Invitrogen, Karlsruhe, Germany) or the protein patterns were transferred to nitrocellulose-membranes. The protein bands were then visualized by immunoblotting. For immunological detection a His-tag-specific antibody (Micromol, Karlsruhe, Germany) or an IFNα-specific antibody (US Biologicals, Massachussetts, MA, USA) was applied. As secondary antibody a goat-anti-mouse-AP-conjugate (BioRad) was applied

Please cite this article in press as: Eilert, E., et al., Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence. J. Biotechnol. (2012), http://dx.doi.org/10.1016/j.jbiotec.2012.08.024

G Model
BIOTEC-6174; No. of Pages 7
ARTICLE IN PRESS
E. Eilert et al. / Journal of Biotechnology xxx (2012) xxx–xxx

respectively. Development was performed by using NBT/BCIP (Roche, Mannheim, Germany).

2.5. Determination of N-terminal amino acid sequence

The protein samples were separated on acrylamide gels just as described above and blotted on a PVDF membrane (Invitrogen). The membrane was stained with 0.1% Coomassie Brilliant Blue G 250 and the protein bands of interest were isolated. The N-terminal amino acid sequence has been determined by TOPLAB (Martinsried, Germany).

2.6. HPLC detection of IFNα-2a

For HPLC analysis samples were prepared by filtration with a cutoff at 0.22 μm. HPLC analyses were carried out as described in the European Pharmacopoeia 5.0 (Council of Europe, 2005).

2.7. Modeling of IL-6- and IFNα-2a-tertiary structures

Native local protein structure properties were obtained by artificial neural networks methods using *NetSurfP* version 1.1 (Petersen et al., 2009) for protein surface accessibility and secondary structure prediction. Automated 3D-model building using *profile–profile matching* (Ohlson et al., 2004) and tertiary structure modeling was carried out using *Protein Homology/analogY Recognition Engine* V 2.0 *(Phyre2) remote homology modeling server* described in Kelley and Sternberg (2009). All resulting structures were aligned using MatchMaker in the program *UCSF Chimera* version 1.6.1 (Pettersen et al., 2004) with the Needleman–Wunsch algorithm (BLOSUM62) (Needleman and Wunsch, 1970) to obtain a structural comparison. Furthermore this structural alignment was visualized using *UCSF Chimera*.

2.8. Homology search

For the detection of remote protein homologs *Domain Enhanced Lookup Time Accelerated BLAST* (*DELTA-BLAST*) was used as described in Boratyn et al. (2012).

3. Results

3.1. Evaluation of peptide initiation by construction of three variants of IL-6 with different N-terminal sequences

According to the literature three different N-terminal ends of the mature IL-6 protein are known as summarized in Fig. 1A. We aimed to determine the influence of these different IL-6 initiations on the processing of secretory MFα prepro-IL-6 fusion proteins. For each variant a synthetic ORF was designed and inserted into the *H. polymorpha* expression vector pFPMT121. In all cases a C-terminal his tag was added for further analyses of the expressed protein (Fig. 1B). In the three resulting expression plasmids pFPMT-MFaE-V30-IL-6-H6, pFPMT-MFaE-P29-IL-6-H6, and pFPMT-MFaE-A28-IL-6-H6 the IL-6 transcription is controlled by the strong carbon source regulated *FMD* promoter. Corresponding IL-6 producer strains were generated on the basis of the *H. polymorpha* host strain RB11 (*ura3*). Selected strains resulting from this procedure were subjected to small scale methanol induction, and the culture supernatants were analyzed for the presence of secreted IL-6 products by SDS-PAGE/Western blot using a His-tag-specific primary antibody. The strains with the highest productivity were chosen for further analyses (V30-IL-6 #11/1, P29-IL-6 #2/2 and A28-IL-6 #3/1; data not shown).

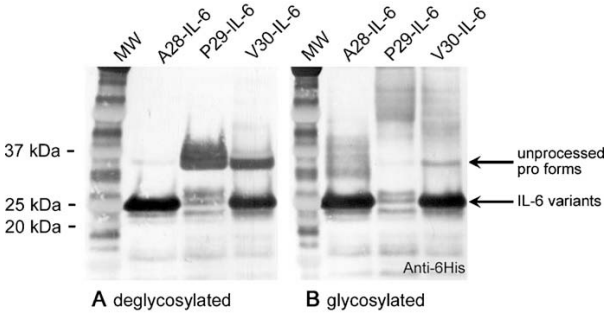


Fig. 2. SDS-PAGE of the supernatant of IL-6 producing strains and a vector control strain. The indicated strains were subjected to methanol induction at the 3 ml scale. The culture supernatant was harvested and analyzed by SDS-PAGE/Western blot. Secreted IL-6 products were visualized by immunodetection via a His-tag specific antibody. MW: molecular weight marker.

3.2. The MFα prepro sequence processing performance is influenced by the exact initiation of the IL-6 target protein

All three variants were efficiently secreted (Fig. 2). However, in Western blots the processing of the MFα pro peptide by the Kex2 protease appears to be different for each variant. The mature V30-IL-6 has a calculated molecular weight of 21 kDa. After deglycosylation, the V30-IL-6 culture supernatant showed correctly processed material with an apparent molecular weight of 25 kDa and in addition a band with a higher molecular weight with an apparent molecular weight of 35 kDa (Fig. 2A) probably due to unprocessed MFα pro peptide. The shift between calculated and apparent molecular weight is in agreement with the findings of Guisez et al. (1991). In contrast to V30-IL-6, the P29-IL-6 found in the culture supernatant appears to consist exclusively of the unprocessed pro form indicating that Kex2p was not able to access the Lys-Arg site (Fig. 2). The fact that we could only detect highly glycosylated material is due to the presence of three N-glycosylation sites within the MFα pro peptide. Only in case of the A28-IL-6 construct, the MFα pro peptide seems to be quantitatively removed by the Kex2 protease – there was only one band detected corresponding to the expected molecular weight of 25 kDa (Fig. 2).

3.3. Expression of both MFα prepro-V30-IL-6 and -A28-IL-6 result in the same mature V30-IL-6

N-terminal sequencing revealed identical protein initiation in case of the 21 kDa material of the V30-IL-6 and A28-IL-6 samples. This result indicates that in the case of V30-IL-6, Kex2p has cleaved behind the dibasic motif Lys-Arg, while in the case of A28-IL-6 an additional cleavage behind the Ala-Pro dipeptide might have taken place. Guisez et al. (1991) found this processing pattern also in the case of MFα prepro-A28-IL-6 expression in *S. cerevisiae*. They showed that the Ste13 protease is responsible for the cleavage behind the Ala-Pro residues (Fig. 3).

3.4. Variation of the transition between MFα prepro sequence and interferon-α 2a (IFNα-2a)

Partial misprocessing as observed with the MFα prepro-V30-IL-6 has also been described for other MFα prepro fusion proteins, for example MFα prepro-IFNα-2a in *H. polymorpha* (Wendel, 2009) and MFα prepro-IFNα-2b in *P. pastoris* (Salunkhe et al., 2010). The culture supernatant of the respective strains typically displayed both correctly Kex2p-processed material as well as a slower migrating form due to -11-misprocessing resembling the situation with the MFα prepro-V30-IL-6 shown above. The overall analogy between

Please cite this article in press as: Eilert, E., et al., Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence. J. Biotechnol. (2012), http://dx.doi.org/10.1016/j.jbiotec.2012.08.024

G Model
BIOTEC-6174; No. of Pages 7
ARTICLE IN PRESS
4 E. Eilert et al. / Journal of Biotechnology xxx (2012) xxx–xxx

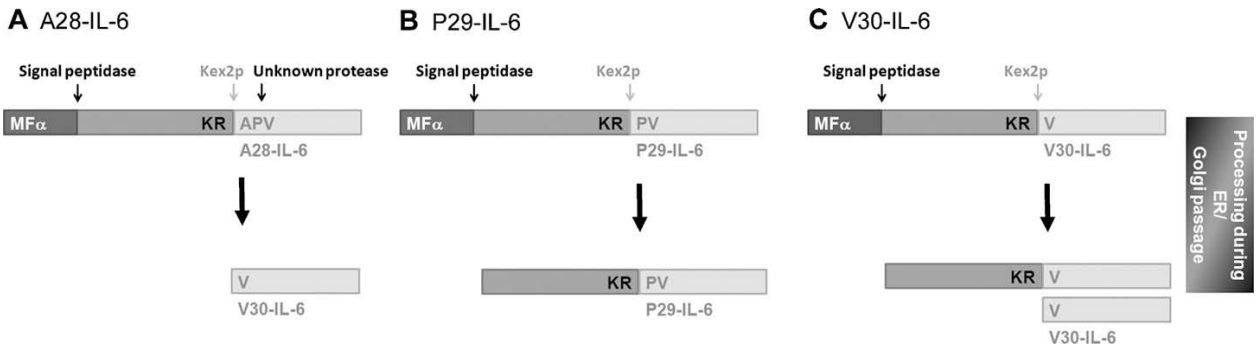


Fig. 3. Proposed processing of A28-, P29-, and V30-IL-6 in *H. polymorpha* based on Western blot analyses and N-terminal sequencing results. Processing of the IL-6 variants in *H. polymorpha* is Ala-Pro dependent. In *S. cerevisiae* the 'unknown protease' was identified as Ste13p (Guisez et al., 1991).

the processing of MFα prepro-IL-6- and -IFNα-2a fusion proteins makes IFNα-2a a suitable test protein to study whether the IL-6 effect of quantitative processing behind Ala-Pro described above can be transferred towards another protein.

To test this a pFPMT121 derivative encoding the MFα prepro-Ala-Pro-IFNα-2a fusion protein was constructed. Respective producer strains were generated on the basis of the *H. polymorpha* host strain RB11 (*ura3*) and KLA8-1 (*ura3 leu2*) as described above. IFNα-2a was expressed at very high levels in several strains. The strains with the highest productivity were chosen for further analyses (RB11/AP-IFNα-2a #1/1 and KLA8-1/AP-IFNα-2a #4/1).

3.5. Processing of the MFα pro sequence in the context of IFNα-2a is improved by insertion of Ala-Pro

We compared the processing of MFα prepro-IFNα-2a and MFα prepro-Ala-Pro-IFNα-2a in two different *H. polymorpha* strain backgrounds. Expression of MFα prepro-Ala-Pro-IFNα-2a in the RB11 strain resulted in an increase of 21 kDa material comigrating with the mature IFNα-2a as compared with MFα prepro-IFNα-2a (Fig. 4: RB11). HPLC analyses further confirmed this observation. The overall yield was increased and misprocessing was reduced (Fig. 5). However, in the case of AP-IFNα-2a the 21 kDa band contained two differently processed forms of IFNα-2a: mature IFNα-2a and most probably AP-IFNα-2a (Fig. 5B). We assume that the majority of the product is efficiently cleaved behind the Kex2p cleavage site resulting in AP-IFNα-2a, and that a subgroup of this material becomes additionally cleaved behind Ala-Pro to release mature IFNα-2a. Importantly, the MFα prepro sequence was processed much more efficiently as the MFα prepro-IFNα-2a devoid of an

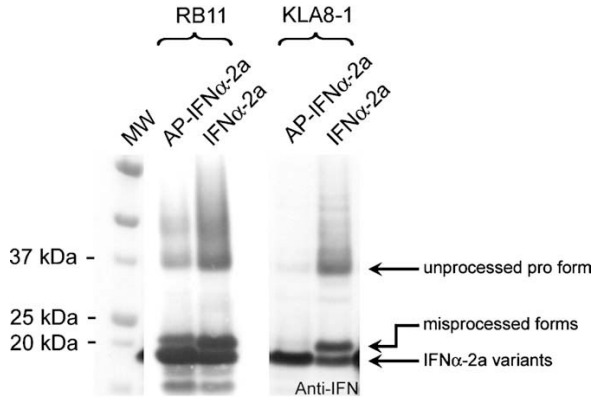


Fig. 4. SDS-PAGE of the culture supernatant of IFNα-2a producing strains in two different strain backgrounds: RB11 and KLA8-1. The indicated strains were subjected to methanol induction at the 3 ml scale. The supernatant was harvested and analyzed by SDS-PAGE/Western blot. Secreted IFNα-2a products were visualized by immunodetection via an IFN-specific antibody. MW: molecular weight marker. For further explanations see text.

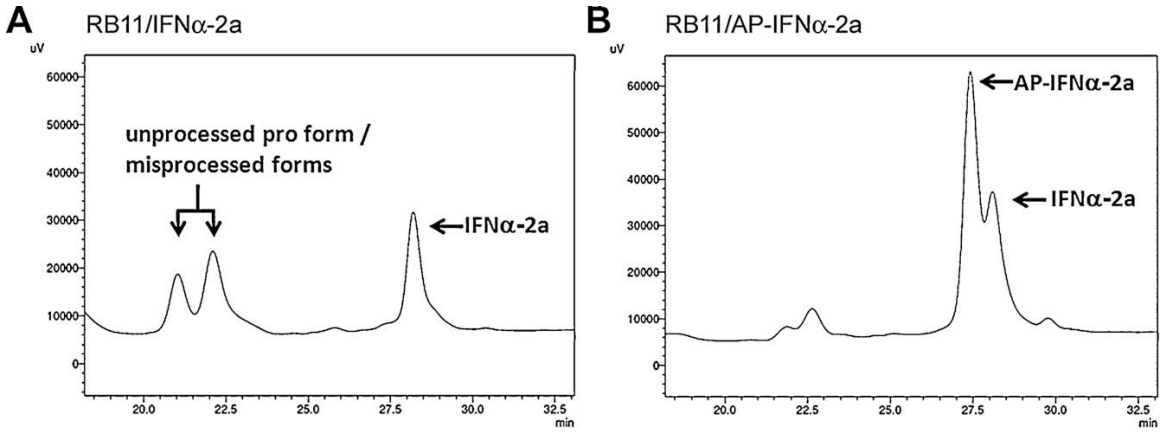


Fig. 5. HPLC analysis was performed according to the Pharmacopoeia method. In the case of AP-IFNα-2a less misprocessed material was detected. The main peak seems to correspond to AP-IFNα-2a in which the two additional amino acids Ala-Pro were not removed. Mature IFNα-2a was also detected but to a lesser degree.

Please cite this article in press as: Eilert, E., et al., Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence. J. Biotechnol. (2012), http://dx.doi.org/10.1016/j.jbiotec.2012.08.024

G Model
BIOTEC-6174; No. of Pages 7
ARTICLE IN PRESS
E. Eilert et al. / Journal of Biotechnology xxx (2012) xxx–xxx

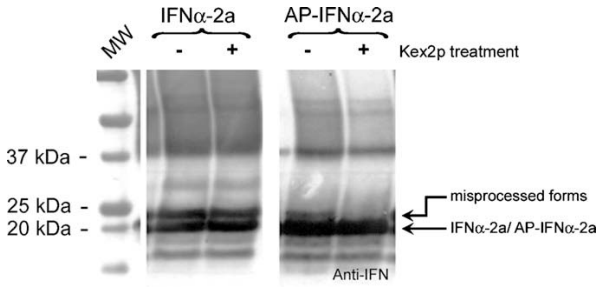


Fig. 6. Effect of Kex2p treatment on misprocessed MFα prepro sequence was analyzed by SDS-PAGE. Culture supernatants of IFNα-2a producing strains were treated in vitro with soluble Kex2p and analyzed by SDS-PAGE/Western blot. IFNα-2a products were visualized by immunodetection via an IFN-specific antibody. MW: molecular weight marker.

Ala-Pro insertion, especially, the amount of high molecular isoforms was significantly decreased.

We repeated the experiment in a second strain background (KLA8-1). The Ala-Pro dependent effect was even more pronounced in this strain background (Fig. 4: KLA8-1) – the Ala-Pro construct leads to almost no misprocessed IFNα-2a molecules.

3.6. Processing of the MFα pro-AP sequence is further improved by in vitro Kex2 treatment

Since the expression of AP-IFNα-2a in the KLA8-1 strain background resulted in nearly complete processing of the MFα prepro sequence at the Kex2p cleavage site we wondered whether AP-IFNα-2a material derived from RB11 strain background containing misprocessed IFN forms can be processed in vitro by soluble Kex2p. The treatment with a *H. polymorpha*-produced soluble fraction containing the *S. cerevisiae* Kex2p resulted in complete elimination of the misprocessed forms of AP-IFNα-2a (Fig. 6). In contrast the various mis- or unprocessed molecules provided by IFNα-2a without Ala-Pro insertion did not at all respond to in vitro treatment with soluble Kex2p. We therefore assume that Ala-Pro leads to a better exposition of the Lys-Arg motif and therefore an improved accessibility for Kex2p.

3.7. Modeling of tertiary structures reveals Ala-Pro dependent accessibility of Kex2-cleavage site

To analyze how the tertiary structure of the MFα prepro sequence transition to the protein of interest is influencing the overall accessibility of the Lys-Arg motif *profile–profile matching* models of the tertiary structure were generated using *Phyre2* remote homology modeling server. The structures were calculated with an overall sequence identity of more than 66% modeled at >90% confidence. The resulting structures were aligned with a cut-off of 2.0 Å RSMD using *UCSF Chimera*. All amino acids relevant in the transition between MFα and protein (Lys84-Arg85-Ala86-Pro87/-Pro86, respectively) were calculated with a confidence score >80%.

Structural alignment of MFα-A28-, -P29- and -V30-IL-6 showed exposition of the Lys-Arg motif for all three variants (Fig. 7B). However structure analysis indicated a different strain motif for MFα-P29-IL-6 resulting in a different accessibility of the Kex2p cleavage site. To visualize changes in surface accessibility ΔRSA was calculated based on *NetSurfP* results (data not shown). An altered surface volume can lead to possible inaccessibility of adjacent regions. A comparison of these results confirmed significant shifts in surface accessibility at Pro86 and Val87 for P30-IL-6 which is indicated by the ΔRSA (Fig. 7C). There are no severe shifts found in case of V30- and A28-IL-6. Due to this tertiary structure motif changes a drift in the surface volume is indicated which can lead to possible inaccessibility of the Kex2p cleavage site for P30-IL-6. These findings are in agreement with the experimental data described earlier.

In the case of IFNα-2a structural alignment showed no exposition of the Lys-Arg motif. The Kex2p cleavage site is buried inside the protein and badly accessible for the Kex2p (Fig. 7E), which is again consistent with the experimental data. In contrast, in the case of AP-IFNα-2a the Lys-Arg motif is exposed and shows a tendency of good accessibility for the Kex2p (Fig. 7E).

4. Discussion

Specific proteases are involved in the processing of precursors to generate the final product. Within these precursors Ala and Pro can be commonly found at even-numbered positions. A stepwise removal of such dipeptides can be found in a variety of organisms such as yeasts, insects, humans but also – rarely – in prokaryotes (Chiravuri et al., 2000; Kreil, 1990; Vanhoof et al., 1995). Additional functions of conserved proline motifs were reviewed by Vanhoof et al. (1995). Especially Xaa-Ala and Xaa-Pro motifs seem to be advantageous if positioned close to a cleavage site since they form exposed hinge regions. The processing of those stretches as well as adjacent domains may therefore be generally improved (Erni et al., 1989).

In this study we found that V30-IL-6 was correctly processed in *H. polymorpha* which is in contrast to the findings of Steinborn et al. (2006). They reported quantitative misprocessing with a N-terminal truncation of 8 amino acids of a comparable MFα-V30-IL-6 construct in a different *H. polymorpha* strain background (HP102). We could also observe protein material of lower molecular weight which might correspond to the truncated material described by Steinborn et al. (2006), however, we found the majority of the expressed protein to be correctly processed. In our system in case of A28-IL-6 the first two IL-6 residues (Ala-Pro) were removed resulting in the same mature V30-IL-6 as in the case of the expression of MFα prepro-V30-IL-6. Thus in this case the Ala28-Pro29 dipeptide has been quantitatively removed. The observations of Steinborn et al. (2006) concerning V30-IL-6 expression could be explained by the use of a different promoter and different cultivation conditions. The author's expression cassette included the constitutive *TPS1* promoter for IL-6 expression while in our system the carbon source regulated *FMD* promoter was used. Differences in the consistencies of a target protein in dependence of the promoter used for expression have been observed previously (see for example Gellissen et al., 2002). Other possible explanations for the observed differences between our data and the data shown in Steinborn et al. (2006) may be the use of different media (YPG vs. HMM), different culturing temperatures (37 °C vs. 30 °C), and/or the use of different carbon sources (glycerol vs. glucose). A further significant difference between the two studies is the host strain used for V30-IL-6 expression (RB11 vs. HP102). Although RB11 and HP102 derive from the same parental strain *H. polymorpha* CBS4732 it is conceivable that the genotypes of these strains differ in more genes than only *URA3* and *LEU2*.

A partial removal of Ala-Pro was also found during AP-IFNα-2a expression when these two amino acids were inserted between the Kex2p cleavage site and the mature IFNα-2a part. In this heterologous context the Ala-Pro removal was not as efficient as in the case of the IL-6 which naturally contains the Ala-Pro dipeptide between signal peptide and mature form.

S. cerevisiae Ste13 protease removes N-terminal Xaa-Ala dipeptides (Julius et al., 1983). Guisez et al. (1991) found Ste13p responsible for removal of the Ala-Pro dipeptide from the N-terminus of heterologously expressed IL-6 in *S. cerevisiae*. Thus

Please cite this article in press as: Eilert, E., et al., Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence. J. Biotechnol. (2012), http://dx.doi.org/10.1016/j.jbiotec.2012.08.024

G Model
BIOTEC-6174; No. of Pages 7
ARTICLE IN PRESS
6 E. Eilert et al. / Journal of Biotechnology xxx (2012) xxx–xxx

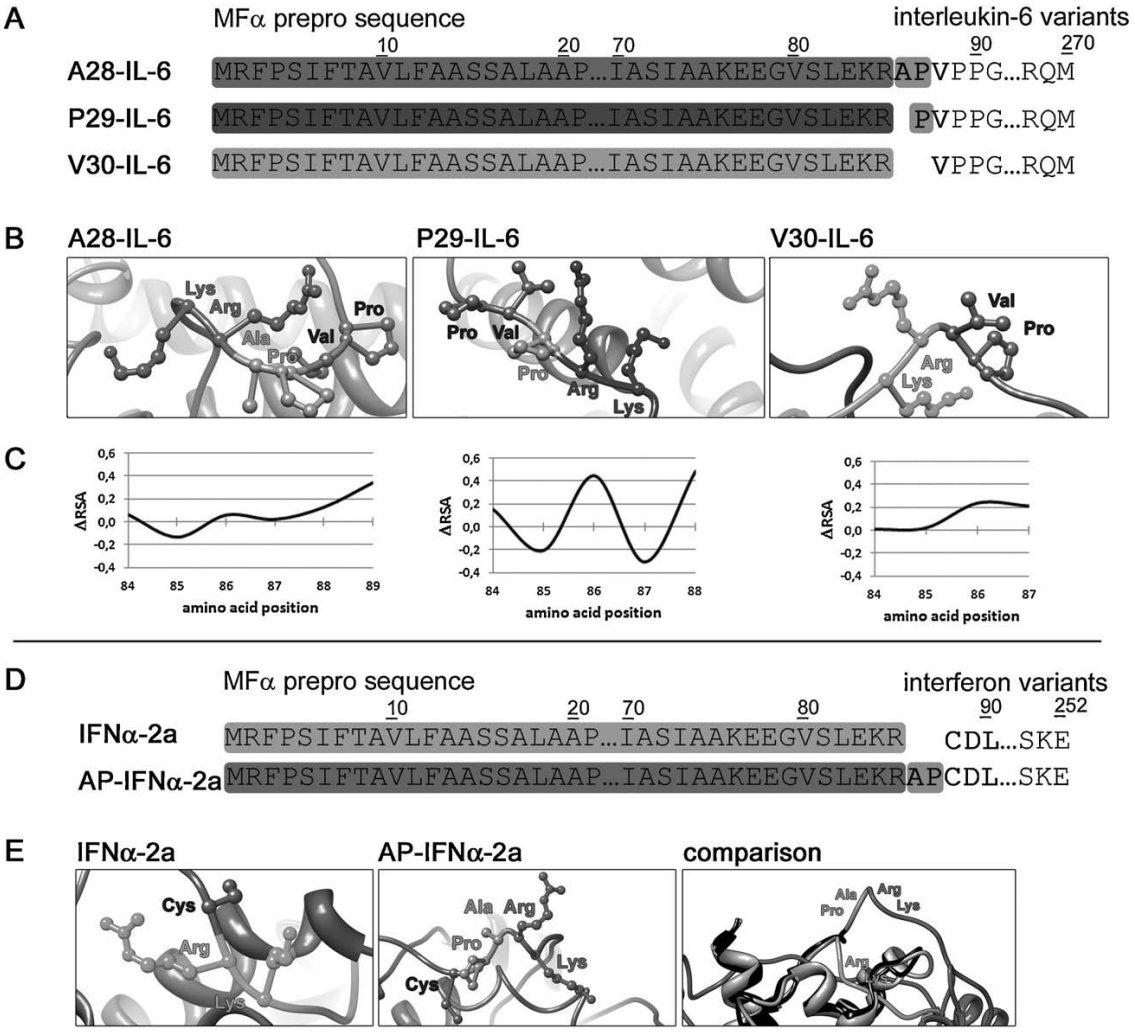


Fig. 7. Structural comparison of MFα prepro sequence transitions and exposition of the Kex2p cleavage site. The transitions between MFα prepro and protein of interest in dependence of Ala and Pro residues were visualized using *UCSF Chimera*. Side chains of relevant amino acids are displayed in ball and stick representation. The protein backbone is shown in a ribbon representation. In A and B (D and E, respectively) the same color code is used. (A) Protein sequences of A28-IL-6, P29-IL-6 and V30-IL-6 fused to the MFα prepro sequence. (B) Structural alignment of the three MFα-IL-6 variants. (C) Determination of structural changes in dependency of drifts in surface volume. ΔRSA was calculated and plotted as a function of the amino acid position in the MFα-IL-6 transitions starting with Lys84 for A28-, P29- and V30-IL-6. (D) Protein sequences of IFNα-2a and AP-IFNα-2a fused to the MFα prepro sequence. (E) Influence of Ala-Pro residues on Lys-Arg accessibility in the two MFα-IFNα-2a variants.

the Ste13 protease is not limited to the cleavage of the Xaa-Ala from the Glu-Ala-Glu-Ala spacer of the natural α-mating factor, but is also able to cleave Xaa-Pro in other contexts. Homology search of different yeast using *DELTA-BLAST* resulted in maximum identities ranging from 32% to 50%. Even for yeasts closely related to *S. cerevisiae* no exact homologs have been identified yet. For these less conserved homologs similar as well as different specificities have been observed. In *Saccharomyces kluyveri* a Ste13p homolog was described to remove both Xaa-Ala and Xaa-Pro residues during α-factor maturation. In *S. kluyveri* the Lys-Arg motif of the MFα prepro sequence is followed by a Glu-Ala-Asp-Ala-Glu-Pro spacer which is sequentially removed also by a Ste13p homolog (Egel-Mitani and Hansen, 1987). Matoba and Ogrydziak (1989) detected dipeptidyl aminopeptidase activity with Xaa-Ala and Xaa-Pro specificity also in *Yarrowia lipolytica*. In contrast, for *P. pastoris* a Ste13-homolog with a different specificity has been described (Prabha et al., 2009). The *P. pastoris* Ste13p homolog performs removal of N-terminal Gly-His dipeptides. Kreil (1990) suspected that the processing of Xaa-Pro residues may also be caused by non-Ste13 proteases. Considering these observations made in other yeasts it may be possible that there may be a Ste13p homolog responsible for the removal of Ala-Pro residues in the cases of IL-6 and IFNα-2a in *H. polymorpha*. With *DELTA-BLAST* homology search a putative Ste13p-homologs was identified for *H. polymorpha* DL-1 (max. identity 34%) which so far has not been further characterized.

Liu et al. (2001) found N-terminal extension of interferons when expressed in *P. pastoris*. They analyzed Interferon α-1 and found 9- and 11-amino acid extensions. Salunkhe et al. (2010) could also show that a -11 isoform occurs during IFNα-2b expression in *P. pastoris*. The authors succeeded to improve the processing of this isoform by addition of at least 3% DMSO to the culture medium, however the total yield was decreased with increasing DMSO concentration. Purkarthofer et al. (2012) detected a 14-amino acid extension during IFNα-2a expression in the same system. They

Please cite this article in press as: Eilert, E., et al., Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence. J. Biotechnol. (2012), http://dx.doi.org/10.1016/j.jbiotec.2012.08.024

G Model
BIOTEC-6174; No. of Pages 7
ARTICLE IN PRESS
E. Eilert et al. / Journal of Biotechnology xxx (2012) xxx–xxx

improved processing of IFNα-2a by co-expression of heterologous protein disulfide isomerase (PDI) and/or Kex2p. In *H. polymorpha* the co-expression of full length IFNα-2a with *S. cerevisae* wild-type Kex2p resulted also in an improved processing (Wendel, 2009), however, in contrast co-expression with a soluble Kex2p fragment did not show an effect (data not shown).

We were able to improve the processing of IFNα-2a significantly by variation of the N-terminal end without reducing the overall yield. Also the choice of the strain background had a significant impact. The yields obtained with MFα prepro-Ala-Pro were higher than those with the unmodified MFα prepro sequence. However, the removal of the Ala-Pro is still incomplete. Our findings may be useful for the expression of technical enzymes, for which higher yields due to improved processing of the MFα prepro sequence are a key factor and the exact protein initiation is not relevant for the protein activity.

The increased cleavage of the MFα prepro sequence by Kex2p when Lys-Arg was followed by Ala-Pro motif is in agreement with modeling data which indicate a Ala-Pro dependent improved accessibility of the cleavage site. The Lys-Arg motif is more exposed when Pro is located at an even-numbered position relative to the cleavage site.

Acknowledgments

Excellent technical assistance was provided by Manuela Bußmann and Nina Wingbermühle.

The research work was supported by grants from the German National Ministry of Education and Research (BMBF) within the ETB EuroTransBio program.

Molecular graphics and analyses were performed with the *UCSF Chimera* package. *Chimera* is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health (National Center for Research Resources grant 2P41RR001081, National Institute of General Medical Sciences grant 9P41GM103311).

References

Akeboshi, H., Chiba, Y., Kasahara, Y., Takashiba, M., Takaoka, Y., Ohsawa, M., Tajima, Y., Kawashima, I., Tsuji, D., Itoh, K., Sakuraba, H., Jigami, Y., 2007. Production of recombinant beta-hexosaminidase A, a potential enzyme for replacement therapy for Tay-Sachs and Sandhoff diseases, in the methylotrophic yeast *Ogataea minuta*. Applied and Environmental Microbiology 73 (15), 4805–4812.

Boratyn, G.M., Schaffer, A.A., Agarwala, R., Altschul, S.F., Lipman, D.J., Madden, T.L., 2012. Domain enhanced lookup time accelerated BLAST. Biology Direct 7 (1), 12.

Brake, A.J., Merryweather, J.P., Coit, D.G., Heberlein, U.A., Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P., Barr, P.J., 1984. Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences of the United States of America 81 (15), 4642–4646.

Chen, X., Gao, B., Shi, W., Li, Y., 1992. Expression and secretion of human interferon alpha A in yeast *Kluyveromyces lactis*. Yi Chuan Xue Bao 19 (3), 284–288.

Chiravuri, M., Agarraberes, F., Mathieu, S.L., Lee, H., Huber, B.T., 2000. Vesicular localization and characterization of a novel post-proline-cleaving aminodipeptidase, quiescent cell proline dipeptidase. Journal of Immunology 165 (10), 5695–5702.

Clare, J.J., Romanos, M.A., Rayment, F.B., Rowedder, J.E., Smith, M.A., Payne, M.M., Sreekrishna, K., Henwood, C.A., 1991. Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. Gene 105 (2), 205–212.

Council of Europe, 2005. European Pharmacopoeia: Interferon Alfa-2 Concentrated Solution, 5th ed. 2, Strasbourg.

Degelmann, A., Müller, F., Sieber, H., Jenzelewski, V., Suckow, M., Strasser, A., Gellissen, G., 2002. Strain and process development for the production of human cytokines in *Hansenula polymorpha*. FEMS Yeast Research 2 (3), 349–361.

Egel-Mitani, M., Hansen, M.T., 1987. Nucleotide sequence of the gene encoding the *Saccharomyces kluyveri* alpha mating pheromone. Nucleic Acids Research 15 (15), 6303.

Erni, B., Zanolari, B., Graff, P., Kocher, H.P., 1989. Mannose permease of *Escherichia coli*. Domain structure and function of the phosphorylating subunit. Journal of Biological Chemistry 264 (31), 18733–18741.

Faber, K.N., Haima, P., Harder, W., Veenhuis, M., AB, G., 1994. Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*. Current Genetics 25 (4), 305–310.

Gellissen, G., Müller, F., Sieber, H., Tieke, A., Jenzelewski, V., Degelmann, A., 2002. Production of cytokines in *Hansenula polymorpha*. In: Gellissen, G. (Ed.), *Hansenula polymorpha*: Biology and Applications. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Guengerich, L., Kang, H., Gehle, B., Gellissen, G., Suckow, M., 2004. A platform for heterologous gene expression based on the methylotrophic yeast *Hansenula polymorpha*. In: Kueck, U. (Ed.), Genetics and Biotechnology. Springer Verlag, Berlin, Heidelberg.

Guisez, Y., Tison, B., Vandekerckhove, J., Demolder, J., Bauw, G., Haegeman, G., Fiers, W., Contreras, R., 1991. Production and purification of recombinant human interleukin-6 secreted by the yeast *Saccharomyces cerevisiae*. European Journal of Biochemistry 198 (1), 217–222.

Janowicz, Z., Melber, K., Merckelbach, A., Jacobs, E., 1991. Simultaneous expression of the S and L surface antigens of hepatitis B, and formation of mixed particles in the methylotrophic yeast, *Hansenula polymorpha*. Yeast, 431–443.

Julius, D., Blair, L., Brake, A., Sprague, G., Thorner, J., 1983. Yeast alpha factor is processed from a larger precursor polypeptide: the essential role of a membrane-bound dipeptidyl aminopeptidase. Cell 32 (3), 839–852.

Julius, D., Schekman, R., Thorner, J., 1984. Glycosylation and processing of prepro-alpha-factor through the yeast secretory pathway. Cell 36 (2), 309–318.

Kelley, L.A., Sternberg, M.J.E., 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nature Protocols 4 (3), 363–371.

Kreil, G., 1990. Processing of precursors by dipeptidylaminopeptidases: a case of molecular ticketing. Trends in Biochemical Sciences 15 (1), 23–26.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (5259), 680–685.

Latchev, K.L., Semenova, V.D., Tolstorukov, I.I., van der Klei, I., Veenhuis, M., 2002. Isolation and properties of genetically defined strains of the methylotrophic yeast *Hansenula polymorpha* CBS4732. Archives of Microbiology 177, 150–185.

Liu, P.T., Ta, T.V., Villarete, L.H., 2001. High-yield expression and purification of human interferon alpha-1 in *Pichia pastoris*. Protein Expression and Purification 22 (3), 381–387.

Matoba, S., Ogrydziak, D.M., 1989. A novel location for dipeptidyl aminopeptidase processing sites in the alkaline extracellular protease of *Yarrowia lipolytica*. Journal of Biological Chemistry 264 (11), 6037–6043.

Needleman, S.B., Wunsch, C.D., 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology 48 (3), 443–453.

Ohlson, T., Wallner, B., Elofsson, A., 2004. Profile–profile methods provide improved fold-recognition: a study of different profile–profile alignment methods. Proteins 57 (1), 188–197.

Parekh, R.B., Dwek, R.A., Rademacher, T.W., Opdenakker, G., van Damme, J., 1992. Glycosylation of interleukin-6 purified from normal human blood mononuclear cells. European Journal of Biochemistry 203 (1-2), 135–141.

Petersen, B., Petersen, T.N., Andersen, P., Nielsen, M., Lundegaard, C., 2009. A generic method for assignment of reliability scores applied to solvent accessibility predictions. BMC Structural Biology 9, 51.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera – a visualization system for exploratory research and analysis. Journal of Computational Chemistry 25 (13), 1605–1612.

Porro, D., Sauer, M., Branduardi, P., Mattanovich, D., 2005. Recombinant protein production in yeasts. Molecular Biotechnology 31 (3), 245–259.

Prabha, L., Govindappa, N., Adhikary, L., Melarkode, R., Sastry, K., 2009. Identification of the dipeptidyl aminopeptidase responsible for N-terminal clipping of recombinant Exendin-4 precursor expressed in *Pichia pastoris*. Protein Expression and Purification 64 (2), 155–161.

Purkarthofer, T., Dib, I., Gruber, U., Imp, E., Plank, H., Bona, R., Weis, R., 2012. Production and processing of interferon in pichia – a cut above. Poster Presentation at Pichia – Conference 2012, 28 February 2012, Alpbach.

Salunkhe, S., Soorapaneni, S., Prasad, K., Raiker, V., Padmanabhan, S., 2010. Strategies to maximize expression of rightly processed human interferon alpha2b in *Pichia pastoris*. Protein Expression and Purification 71 (2), 139–146.

Steinborn, G., Böer, E., Scholz, A., Tag, K., Kunze, G., Gellissen, G., 2006. Application of a wide-range yeast vector (CoMed) system to recombinant protein production in dimorphic *Arxula adeninivorans*, methylotrophic *Hansenula polymorpha* and other yeasts. Microbial Cell Factories 5, 33.

Suckow, M., Gellissen, G., 2002. The expression platform based on *H. polymorpha* strain RB11 and its derivatives – history, status and perspectives. In: Gellissen, G. (Ed.), *Hansenula polymorpha*: Biology and Applications. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Vanhoof, G., Goossens, F., Meester, I., de Hendriks, D., Scharpé, S., 1995. Proline motifs in peptides and their biological processing. FASEB Journal 9 (9), 736–744.

Wagner, J.-C., Wolf, D.H., 1987. Hormone (pheromone) processing enzymes in yeast. The carboxy-terminal processing enzyme of the mating pheromone α-factor, carboxypeptidase yscα, is absent in α-factor maturation-defective *kex1* mutant cells. FEBS Letters 221, 423–426.

Wendel, F., 2009. Untersuchungen zur Expression von Interferon alpha in der methylotrophen Hefe *Hansenula polymorpha*. Diploma thesis. Fachhochschule Aachen Standort Jülich.

Weydemann, U., Keup, P., Piontek, M., Strasser, A.W., Schweden, J., Gellissen, G., Janowicz, Z.A., 1995. High-level secretion of hirudin by *Hansenula polymorpha* – authentic processing of three different preprohirudins. Applied Microbiology and Biotechnology 44 (3-4), 377–385.

Please cite this article in press as: Eilert, E., et al., Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence. J. Biotechnol. (2012), http://dx.doi.org/10.1016/j.jbiotec.2012.08.024

3.2 Supplementary Data

3.2.1 Abstract

The MFα prepro sequence is the secretion signal which most commonly finds application for secretion of heterologous proteins in various yeast systems. Processing at the Kex2p cleavage is dependent on the amino acid sequence adjacent to the cleavage site of the signal sequence. Introduction of an Ala-Pro dipeptide adjacent to the Kex2p cleavage site highly improved IFNα-2a processing probably by improved accessibility of the cleavage site. However, only a partial removal of the Ala-Pro dipeptide occurred.

We identified a protease with the ability to remove Ala-Pro dipeptides in *H. polymorpha*. Deletion of the corresponding gene of the putative protease HpS9-15130 resulted in an altered processing of AP-IFNα-2a in comparison to the wildtype host strain. The N-terminal dipeptide of AP-IFNα-2a was not removed at all. The putative protease HpS9-15130 shows significant homology to *S. cerevisiae* Ste13p as well as other Ste13p homologs.

3.2.2 Introduction

Julius et al. (1983) observed that a particular deletion mutant of *S. cerevisiae* resulted in a sterile phenotype. The lack of the membrane-bound heat-stable dipeptidyl aminopeptidase Ste13p (DPAP A) was identified to be responsible for inability of proper mating. The lack of Ste13p resulted in unprocessed isoforms of the α-mating factor which showed decreased biological activity. Hence, the removal of the Glu-Ala spacer is necessary for activation of the α-factor molecules (Brake et al., 1984). The expression of Ste13p is induced when cells are treated with α-factor (pheromone-mediated expression) (Achstetter, 1989). Ste13p is located in the late Golgi where it is retained by a Phe-Xaa-Phe-Xaa-Asp motif within the cytoplasmic domain (Nothwehr et al., 1993; Fuller et al., 1988). This protease shows strong homology to the human DPPIV which is also commonly found bound to the membrane (Cunningham and O'Connor, 1997). The catalytic site of DPPIV is located at the C-terminal region. The mating deficiency can be partly restored by overexpression of a gene encoding Dap2p (DPAP B) (Julius et al., 1983). Both proteases share similar features regarding primary structure and topology (Jones, 1991).

Although the main task of Ste13p is processing of the α-mating factor, it has shown ability to remove other dipeptides besides Glu-Ala and Asp-Ala. Brenner and Fuller (1992) found that during maturation of Kex2p Leu-Pro and Val-Pro dipeptides are removed. It was suggested that Ste13p is responsible for this processing. Additionally Guisez et al. (1991) could detect Ste13p-mediated removal of Ala-Pro dipeptides. Even though Ste13p is not specific for the removal of Xaa-Pro dipeptides it is considered to be a proline specific N-terminal exopeptidase (Walter et al., 1980).

Putative Ste13p homologs have also been found in other yeasts. For *Yarrowia lipolytica* (Matoba and Ogrydziak, 1989), *C. albicans* (Bautista-Muñoz et al., 2005) and *P. pastoris* (Prabha et al., 2009) *S. cerevisiae* Ste13p homologs showed variation. However the cleavage specificity was also reported to be diversified. The *Y. lipolytica* homolog removes 10 consecutive dipeptides with the Xaa-Pro or Xaa-Ala motif during processing of the extracellular alkaline protease. The *P. pastoris* homolog was shown to remove a Gly-His dipeptide at the N-terminus of exendin-4. The specificity of *C. albicans* has not yet been determined.

Since we observed that Ala-Pro dipeptides were removed in *H. polymorpha*, we suspected that this species would also harbor a putative Ste13p. A gene encoding a protease responsible for Ala-Pro removal during IFNα-2a expression was identified by deletion mutation. Subsequently this putative *H. polymorpha* Ste13p homolog was compared to other Ste13p homologs derived from other yeasts.

3.2.3 Materials and methods

Additional experiments were performed as described below, all other experiments were performed as described in Manuscript I.

Strains and culture conditions

For deletion of putative proteases *H. polymorpha* KLA8-1/AP-IFNα-2a #4/1 deficient in β-isopropylmalate dehydrogenase (*leu2*) was used.

Vector construction

To generate the expression plasmids pDS15-16551-lx(Ls), pDS9-15130-lx(Ls) and pDS9-102261-lx(Ls) up- and downstream fragments were amplified via PCR using the primers for synthetic open reading frames for the enzymes listed in Tab. 3.1 resulting in amplicons of approximately 1.1 kbp length. The PCR products were digested with *Pme*I/*Pac*I (upstream fragments) or *Sbf*I/*Not*I (downstream fragments). These fragments have been ligated with the 3.165 kbp *Not*I/*Pme*I fragment of p-lx(Ls) and the 1.645 kbp *Pac*I/*Sbf*I fragment of p-lx(Ls). The three different plasmids contain the up- and downstream fragment for the target locus as well as a selection marker (*LEU2syn*) and an origin of replication for propagation in *E. coli*. The deletion cassette was excised via digestion with *Not*I/*Pme*I. *H. polymorpha* was transformed by electroporation as described in Faber et al. (1994).

Table 3.1: **Primers for amplification of up- and downstream fragments of three proteases with putative Xaa-Pro specificity.** Search for peptidases with Xaa-Pro specificity was performed using the *H. polymorpha* NCYC 495 *leu1.1* genome database. PCR amplification results in fragments of approximately 1.1 kbp length for either up- or downstream fragment. Unique restriction sites for cloning were introduced by PCR amplification due to the primer design.

Name	Primer sequence
16551_up_A	5'ATCGATGTTTAAACAGTAGTTGTACATGATCTCCTCAAGTCCTGTTTC
16551_up_B	5'CTGCCCTTAATTAACCATGCTTCTCGGCCGATCGACCCCAGTG
16551_down_A	5'GCGCGTCCTGCAGGTCGAGACCCGAATTGAGCCCCACGCGCGG
16551_down_B	5'CGGCTGGCGGCCGCGGTCCCTGTACAGCGGCACATGGTGGAAAAG
15130_up_A	5'TGTTGGGTTTAAACATTGATTTCATGCCGCAAGGACAAGACAAACTGG
15130_up_B	5'GCTTTCTTAATTAATCGGATGAATTAGAACAGAATGTACAAATTAGATCAG
15130_down_A	5'CGACCCCCTGCAGGTTTCGGGGCTGATATCAGTTTGATCAATAAAGTAC
15130_down_B	5'GTGAACGCGGCCGCTTATATCTGTGAAGAAAGCTGATGCAAACATCAGC
102261_up_A	5'ATCTAAGTTTAAACTCGAGGCTTTGATCAGCTAACTATACTAACTCAG
102261_up_B	5'CTTTTCTTAATTAAGTTGTAAATTTAACCTAAGAGGAAATATTTCTATTTAAG
102261_down_A	5'GTCATGCCTGCAGGATTTGGGTGCTGTTTTTATACTCAAACAAGGTCCC
102261_down_B	5'GGTGCGGCGGCCGCCGCACGTTCGATCAGACGCTTCAAATGGTATTC

Bioinformatic analysis

Sequence data for *H. polymorpha* were obtained from *The Genome Portal of the Department of Energy Joint Genome Institute* via http://genome.jgi-psf.org/Hanpo2/Hanpo2.home.html (*H. polymorpha* strain NCYC 495 *leu1.1*). Conserved domains were identified using *Conserved Domain Database (CDD)* (Marchler-Bauer et al., 2011). Prediction of subcellular localization was searched with *PSORT II* (Nakai and Horton, 1999). For the detection of remote protein homologs *Domain Enhanced Lookup Time Accelerated BLAST* (*DELTA-BLAST*) was used as described in (Boratyn et al., 2012). For internal searches in the *H. polymorpha* strain NCYC 495 *leu1.1* genome database the *protein-protein BLAST* algorithm was used (Altschul et al., 1997). Sequences were aligned and a phylogenetic tree was constructed using *MEGA5 Molecular Evolutionary Genetics Analysis* (Tamura et al., 2011).

Identification of a putative α-mating pheromone precursor

To locate the α-mating pheromone precursor the *H. polymorpha* NCYC 495 *leu1.1* Genome Database was filtered according to the following regular expression matchers in Fig. 3.1 in the file Hanpo2_all_proteins_20100927.aa.fasta using *BioRuby* (Goto et al., 2010). The search algorithm was designed by Andreas Heumaier. All matches issued were checked manually.

```
def search
      @db.each_entry do |entry|
        if hit =  entry.seq.match(/[Q,A][P][V,A].*/)
          if rhit =
hit.to_s.match(/[K,R]R[N,E,D,S,V,A][A,P,S].........*[K,R]R[N,E,D,S,V,A][A,P,S]/)
            puts entry.definition
            puts rhit
          end
        end
      end
    end
```

Figure 3.1: **Search algorithm for the identification of α-mating pheromone precursors in databases with FASTA format.**

3.2.4 Results

Modeling of tertiary structures reveals Ala-Pro dependent accessibility of Kex2p-cleavage site

In addition to the modeling of MFα-IFNα-2a and -AP-IFNα-2a in Manuscript I the surface accessibility RSA was determined using *NetSurfP*. Based on these data ΔRSA was calculated in order to visualize changes in surface accessibility. As described earlier significant shifts in RSA may indicate possible inaccessibility of adjacent regions and therefore inaccessibility of adjacent cleavage sites like the Kex2p cleavage site. Structural alignment of MFα-IFNα-2a and -AP-IFNα-2a showed a strong exposition of the Lys-Arg motif for MFα-AP-IFNα-2a (Fig. 3.2B) but not for MFα-IFNα-2a. In the latter case the Kex2p cleavage site was buried inside the protein and thus badly accessible. Additional calculation of ΔRSA for AP-IFNα-2a showed no significant changes between Lys84, Arg85 and Ala87 (Fig. 3.2C). Thus the Lys-Arg motif shows a tendency to good accessibility for the Kex2p. In contrast, calculation of ΔRSA for IFNα-2a showed distinct changes for Lys84, Arg85 and Cys86. These changes in surface accessibility may provide additional evidence for only partial recognition and cleavage of the Kex2p recognition site.

Identification of proteases with Xaa-Pro specificity

Evaluation of the *H. polymorpha* NCYC 495 *leu1.1* genome revealed three proteases with predicted Xaa-Pro specificity. The protease with protein ID 16551 (HpS15-16551p) harbors a conserved domain of the Peptidase S15 family. The proteases with protein ID 15130 (HpS9-15130p) and protein ID 102261 (HpS9-102261p) both belong to the Peptidase S9 family. To study whether one of these proteases is responsible for the removal of Ala-Pro dipeptides in *H. polymorpha* all three candidate genes were disrupted separately in strain KLA8-1/AP-IFNα-2a #4/1. For each protease a flanking up- and downstream fragment of an approximate size of 1.1 kbp was amplified via PCR. These fragments were used to generate the plasmids pDS15-16551-lx(Ls), pDS9-15130-lx(Ls) and pDS9-102261-lx(Ls). Transformation of *H. polymorpha* was performed with the linearized deletion cassette

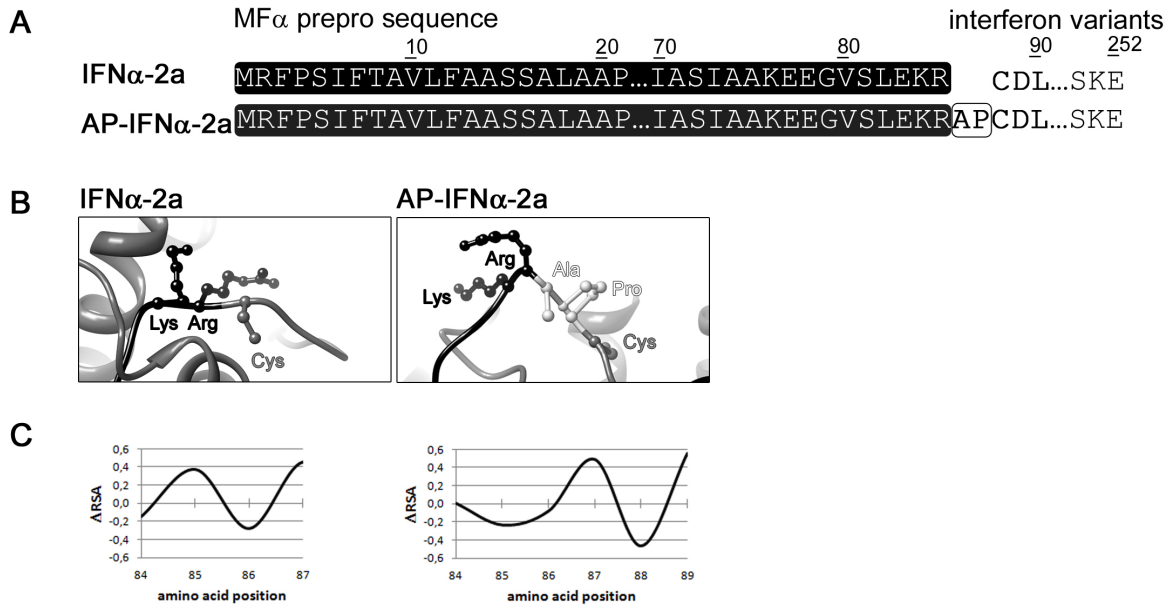


Figure 3.2: **Structural comparison of the Kex2p cleavage site exposition in dependency of adjacent amino acids.** The transitions between MFα prepro and protein of interest in dependency of Ala and Pro residues were visualized using *UCSF Chimera*. Side chains of relevant amino acids are displayed in ball and stick representation. The protein backbone is shown in a ribbon representation. In A and B the same color code is used. (A) Protein sequences of IFNα-2a and AP-IFNα-2a fused to the MFα prepro sequence. (B) Influence of Ala-Pro residues on Lys-Arg accessibility in the two MFα-IFNα-2a variants. (C) Determination of structural changes in dependency of drifts in surface volume. ΔRSA was calculated and plotted as a function of the amino acid position starting with Lys84

resulting in the strains KLA8-1/AP-IFNα-2a/ΔS15-16551-lx(Ls), KLA8-1/AP-IFNα-2a/ΔS9-15130-lx(Ls) and KLA8-1/AP-IFNα-2a/ΔS9-102261-lx(Ls).

Deletion of a putative Ste13p-homolog gene results in significantly decreased removal of Ala-Pro dipeptides

For each construct 12 transformants were subjected to small scale methanol induction and the culture supernatants were analyzed via HPLC for changes in processing of AP-IFNα-2a in comparison to KLA8-1/AP-IFNα-2a #4/1. Expression of MFα prepro IFNα-2a in KLA8-1 resulted in two isoforms of IFNα-2a one of which is the correctly processed from and the other one is prolonged by the Ala-Pro dipeptide (Fig. 3.3A). The deletion of the gene corresponding to HpS15-16551p resulted in the same processing pattern (B). The deletion of the gene corresponding to HpS9-15130p resulted in a different processing pattern (Fig. 3.3C). The amount of correctly processed material was significantly decreased from about 30% to 20% and simultaneously the amount of the prolonged isoform increased. A remaining fraction of 20% showed correct processing of MFα prepro AP-IFNα-2a to IFNα-2a. Additionally, an increase of the overall yield of about 80% was detected. The deletion of the gene corresponding to HpS9-102261p also resulted in an decrease of correctly processed material to about 25%, however, the overall yield of IFNα-2a in this sample was significantly lower (minus 20%), indicating a general problem in IFNα-2a expression after gene deletion.

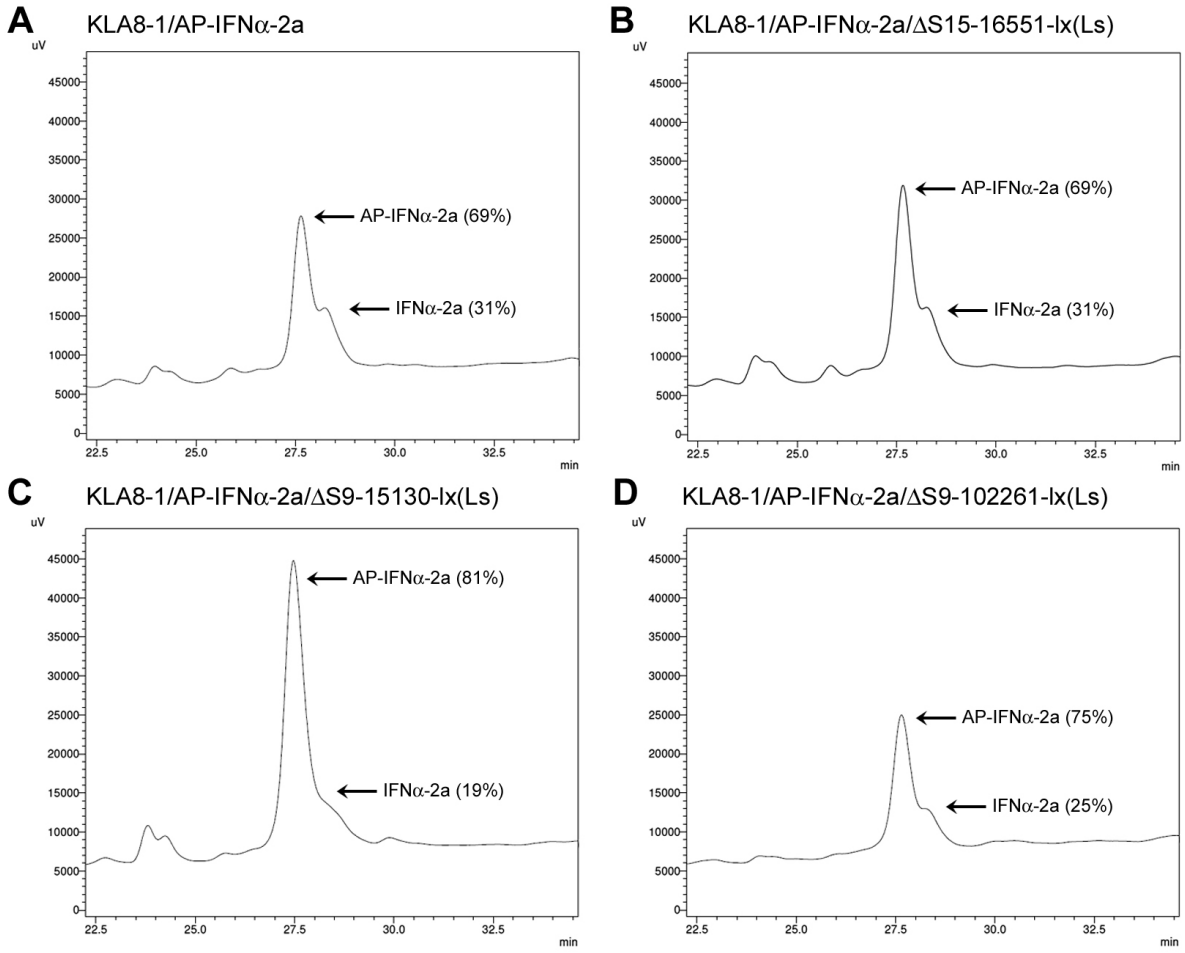
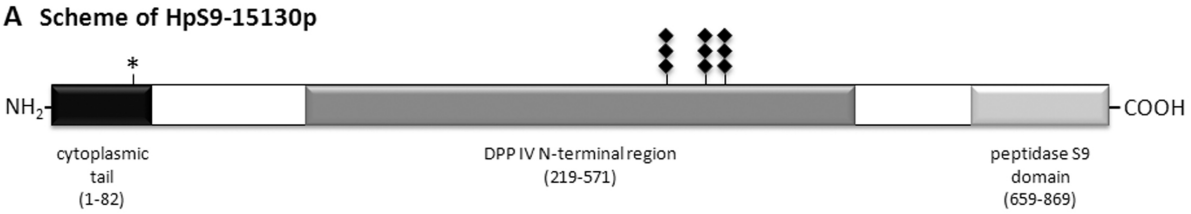


Figure 3.3: **HPLC analysis of KLA8-1/AP-IFNα-2a/ΔS9-15130-lx(Ls).** Deletion of *S9-15130* resulted in significant reduction of the Ala-Pro dipeptide at the AP-IFNα-2a N-terminus. Different isoforms are indicated. The amount of each isoform in the sample is indicated in percent.

Characterization of HpS9-15130p

Deletion of the open reading frame (ORF) corresponding to HpS9-15130p resulted in a reduction of Ala-Pro dipeptide removal during MFα prepro IFNα-2a processing. This ORF consists 2.615 kbp and encodes a polypeptide of 871 amino acids (aa) with a cytoplasmic tail at the N-terminus (aa 1 to 82). Moreover two conserved domains predicted for this protease were also found for the other putative Ste13p homologs. At the C-terminus a protease S9 domain was identified (aa 659 to 869). In the center part of the polypeptide a DPP IV N-terminal region was identified (aa 219 to 571). Three potential N-glycosylation sites are present within this DPP IV N-terminal region (aa positions 507, 539 and 545). A `YQRL` signal can be found at position 66 which might mediate localization of HpS9-15130p in the *trans*-golgi network (Humphrey et al., 1993).



B Amino acid sequence of HpS9-15130p

```
  1  MARNSDIGRL DADIELQNQT SRRNSFQLEE LNNFAIDNDD ADLVYHPTSY KGLNVVSGDL SGMSFYQRLH HFLSLPKTRR KITILGIVLF IVWIVLLISF
101  AHTGAFQNGN KTVYPGQNST SPVSTDESAL KDPQERIPMT LQDSRNGAFS VFSYDLEFVS IPVMSYDEDQ GYYLHRSGSY HFLKKASDSQ FSKKLYDRPE
201  FEYNGQSHRF SLLSLNMELN KALLVSEVEA QWRYSSYGVY WVLDIDSGAV SPVFEKDSKL PKISFAEFSP KGRYVTFVFE NNVYLLDLQT NEQIQITHDG
301  SSSISNGKPD WVYEEEVLSS DRAIYWCQDE SHFSFIRFDD TDVPAYDLEI FSDNPYPVTE KLKYPKPGYN NPIVSVHVYS VEGKTTHEVT HDNSRLGSDY
401  VVYQMAWTSG KKLIIKETDR TSRLVDIRVY DSETTKSTVV RSFDTKEYDG WYTNNGDIFV VPGGGYIDTV VYDYHDHLAY FESATDSEPK MITKGDWDVV
501  DGVAGYNPTA KLLYFIGTAG NALERHIFTV KLDGQDLRNL TNTENVSTYR LSVSSGGKYG FVQYKGPASP TDKIIEMDKF SEDRDEYVQA KNLNTANMVE
601  EALNRYQVPS KEYKTVTLED GTKIDVIEVK PFDFNPENKY PLFVSVYGGP GAQKLSSVFN YGFEEVLVSN LEIIVWYIDP RGTGGRGWKY KSFAKDKIGY
701  WEPRDIVEAT QRLIEMTDYI DMDLTAIWGW SYGGFTTLKT LEYDSGTTFK YGMAVAPVTN WLLYDSIYTE RYMDKPENNE NYQSVSKISD VDKFKQVTRF
801  LIMHGTADDN VHFQNTLQLL DAFDREGVEN YDMHVFPDSN HNINHNNANM VVYDKLFKWL EAAFDKKWVN I
```

Figure 3.4: **Scheme and amino acid sequence of HpS9-15130p from *H. polymorpha* NCYC 495 *leu1.1*.** Domains were determined using *Conserved Domains Database*. The ER-retention signal `YQRL` was identified using *PSORT II* and is indicated by an asterisk. Potential glycosylation sites for Asn-glycosylation are indicated by solid diamonds trees.

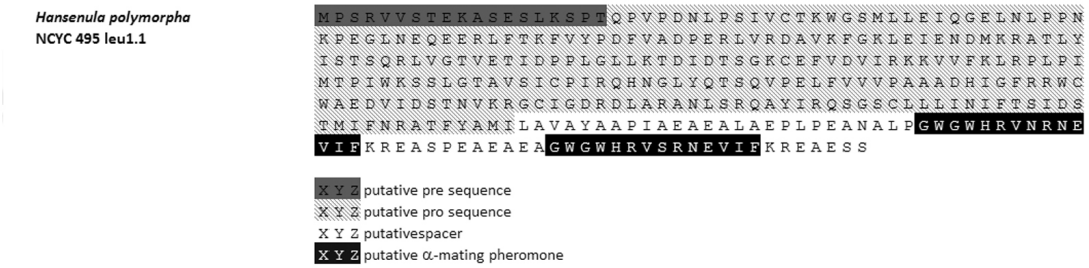


Figure 3.5: **Amino acid sequence of a putative α-mating pheromone precursor in *H. polymorpha* Protein ID 9879.** Structural elements are colored as indicated.

Identification of a putative α-mating pheromone precursor

The existence of Kex2p and Ste13p homologs may indirectly indicate the existence of an α-mating pheromone precursor in *H. polymorpha*. The genome transcripts of the strain NCYC 495 *leu1.1* were analyzed for a repeating Lys-Arg pattern within a peptide sequence. In addition the amino acid downstream of the Lys-Arg motif was set to be Asp, Asn, Glu, Ser or Val. These amino acids were chosen in analogy to separating spacer found in uneven-numbered positions in other yeasts (see Discussion, Fig. 3.7). We could identify one candidate in which the amino acid sequence contains two putative α-mating pheromone molecules separated by a Glu-Ala-Ser-Pro-Glu-Ala-Glu-Ala-Glu-Ala spacer (Protein ID 9879, Fig. 3.5). Both α-mating pheromone units were followed by a Kex2p recognition site and the first one was preceded by a prepro sequence like peptide and additional spacer-like sequence. We could detect a putative cleavage site for removal of the pre sequence, however, no Kex2p recognition site was found for removal of the pro peptide. We suggest that a protease unlike Kex2p is responsible for pro peptide removal. *protein-protein BLAST* analysis showed homology to three other putative α-mating pheromone precursors. The maximum identity was ranging between 30 and 42% (*Candida glabrata* XP_446929, *Nakaseomyces delphensis* AAO25615 and *Eremothecium*

cymbalariae XP_003645399). Interestingly, besides the Lys-Arg motifs also an Arg-Arg motif can be found within the precursors derived from *C. glabrata* and *E. cymbalariae*.

Phylogenetic Analysis

Homology search of Ste13p homologs in different yeasts using *DELTA-BLAST* resulted in maximum identities ranging from 32% to 50%. Even among yeasts closely related to *S. cerevisiae* no exact homologs have been identified yet. A phylogenetic tree (maximum likelihood) was constructed illustrating the relationship of those yeast in reference to *S. cerevisiae* (Fig. 3.6). The maximum identity for each homolog concluded from the *DELTA-BLAST* results is indicated behind the respective entry (in percent, gray lettering). The max. identity is defined as the similarity between the query and the matched sequences over the length of the coverage area. With *DELTA-BLAST* homology search a putative Ste13p homolog was identified for *H. polymorpha* DL-1 (max. identity 34%).

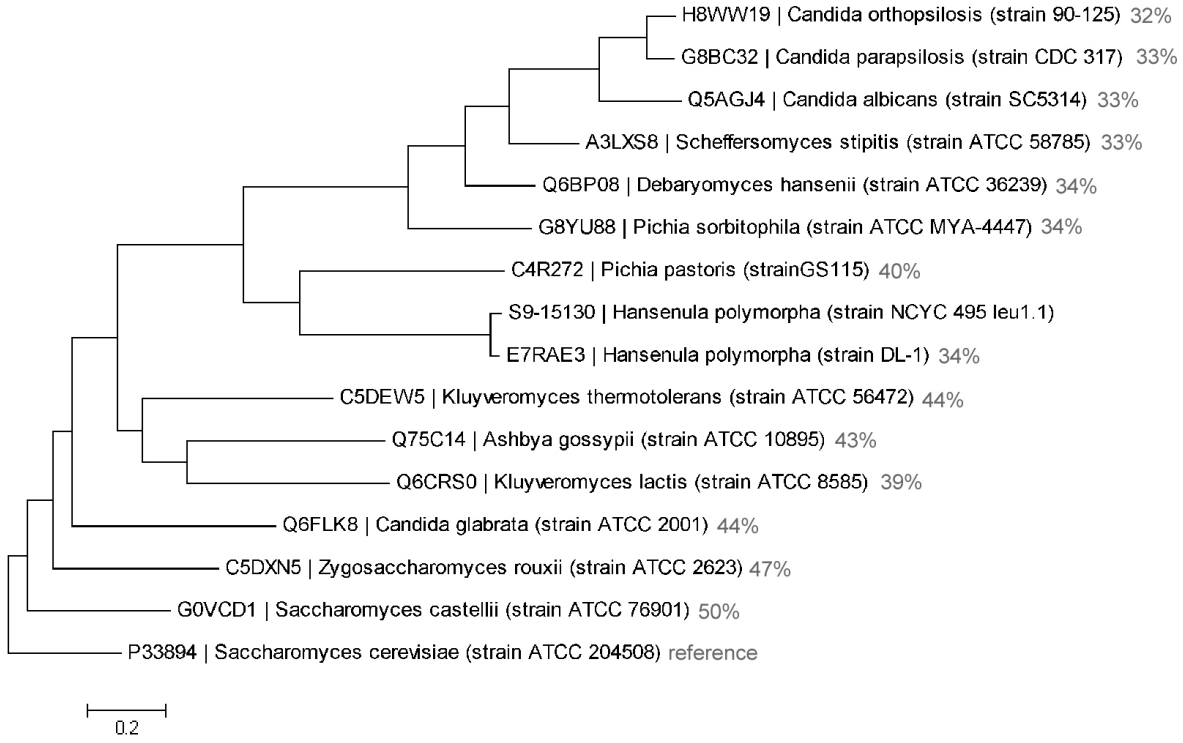


Figure 3.6: **Comparison of *S. cerevisiae* Ste13p and 15 putative Ste13p homologs derived from various yeasts by molecular phylogenetic analysis.** The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model. Initial tree(s) for the heuristic search were obtained automatically as follows. When the number of common sites was < 100 or less than one fourth of the total number of sites, the maximum parsimony method was used; otherwise BIONJ method with MCL distance matrix was used. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 16 amino acid sequences. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in *MEGA5*. Maximum identities were determined using *DELTA-BLAST* (grey lettering).

3.2.5 Discussion

To improve quality and quantity of a secreted heterologous protein strain engineering mainly focuses on four main topics (Idiris et al., 2010):
1. Enhancement of proper protein folding and engineering of the quality control system
2. Improvement of intracellular protein transport
3. Reduction of proteolytic degradation
4. Engineering of glycosylation patterns

In this study the transport from the cytosol through Golgi an ER to final secretion was subject of improvement. Recognition of cleavage sites for signal peptide removal is necessary for secretion of properly maturated protein. When the recognition site is not, incorrectly, or only partially cleaved by the specific protease partly un- or misprocessed material is secreted - if secretion occurs at all. In addition to cleavage site recognition the amount of processing proteases may also be limiting during high level expression of recombinant proteins (Zsebo et al., 1986).

Processing of the MFα prepro sequence is different in *H. polymorpha* strains RB11 and KLA8-1.

Several *H. polymorpha* strains find application in research. The strains DL-1, CBS4732 and NCYC495 or derivatives thereof are used for basic research as well as for the generation of strains producing a heterologous protein of interest. Belonging to the methylotrophic yeasts *H. polymorpha* shares several characteristics with other methanol-utilizing yeasts such as *P. pastoris* and *C. boidinii*. Recent analysis of the gene content and gene order of *H. polymorpha* strain DL-1 places it in a close relation to *Brettanomyces custersianus* (Eldarov et al., 2011). Notably, the relation to the other methylotrophic yeasts mentioned is not as close. Different processing between different strains of the same yeast has also been observed for MFα prepro INFα-2a. Expression of this fusion in *H. polymorpha* RB11 and KLA8-1 resulted in processing with different efficiency. Both strains are derived from the same parental strain CBS4732; the processing results in different quantities of isoforms of the protein. The strain RB11 was generated by application of standard ethylmethane sulphate method (Zurek et al., 1996). KLA8-1 was generated by *N*-methyl-*N*'-nitro nitrosoguanidine mutagenesis (Lahtchev et al., 2002). The generation of these strains apparently has led to differences in processing. The differences in processing may for example be caused by differences in cleavage site recognition, protein folding or glycosylation.

Differences in IL-6 processing have also been found between *H. polymorpha*, *S. cerevisiae*, *P. pastoris* and *Schizosaccharomyces pombe*. Phylogenetic analysis of the putative *S. cerevisiae* Ste13p homologs shows similarities of 32% (*P. pastoris*) and 40% (*H. polymorpha*). The cleavage specificity shows obvious differences (Tab. 3.2). *H. polymorpha* RB11 shares the processing of A28-IL-6 only with *S. cerevisiae*. In *P. pastoris* no removal of the Ala-Pro dipeptide occurs (Li et al., 2011). Furthermore differences in the processing pattern of V30-IL-6 were described for *S. cerevisiae* and *H. polymorpha*

strain HP102 (Steinborn et al., 2006). Summarizing these results, *P. pastoris* and *H. polymorpha* can be considered as the preferred systems for IL-6 expression. *P. pastoris* processed A28-IL-6 correctly, whereas *H. polymorpha* RB11 performed correct processing of V30-IL-6. Remarkably, processing of P29-IL-6 using the signal peptide P4 for secretion in *S. pombe* was correct. The signal peptide P4 is derived from the *S. pombe* α-mating factor with a Lys-Arg motif which is supposed to be recognized and cleaved by a Kex2p homolog (Giga-Hama et al., 1994). Hence, the *S. pombe* Kex2p homolog is able to cleave at the Lys-Arg cleavage site despite of proline being in the first position after the cleavage site.

H. polymorpha harbors an exoprotease with the ability to remove N-terminal Ala-Pro dipeptides. The protease which shows potential to be responsible for Ala-Pro removal was identified by deletion of the corresponding ORF.

A protease with the ability to remove Ala-Pro dipeptides was identified in *H. polymorpha* KLA8-1. Processing of P28- and V30-IL-6 results in the same mature form. Hence, in case of P28-IL-6 the N-terminal Ala-Pro dipeptide is removed. Processing of AP-IFNα-2a results in a mixture of correctly processed IFNα-2a as well as a longer form which is proposed to harbor additional Ala-Pro dipeptides at the N-terminus. There are two possibilities: the protease responsible for this processing may either be an endo- or an exoprotease. By evaluation of the IL-6 processing neither option can be excluded. However, the incomplete AP-IFNα-2a processing may indicate that this protease is an exoprotease which can only act after proper Kex2p cleavage.

To identify this putative exoprotease the ORFs of three proteases with predicted Xaa-Pro cleavage specificity were deleted. The deletion of the ORF corresponding to the protease HpS9-15130 resulted in a significantly altered processing pattern of MFα prepro AP-IFNα-2a in comparison to the wildtype host strain. The removal of the N-terminal dipeptide Ala-Pro was reduced indicating that this protease may be responsible for this processing effect. However, correctly processed material was also detected. There are two possibilities resulting in correct processing: 1. An additional protease with the ability to cleave the Ala-Pro dipeptide performs this removal, 2. The Ala-Pro dipeptide is removed by unspecific proteolysis. A second putative protease responsible for this processing may be HpS9-102261p. Further identification may be achieved by generation of a double deletion mutant. For a final conclusion, the role of these proteases for Ala-Pro removal during IL-6 processing needs to be further evaluated.

H. polymorpha possesses a protease with high similarities to _S. cerevisiae_ Ste13p (max. identity 34%).

The sequence of the protease HpS9-15130p shows significant homology to *S. cerevisiae* Ste13p as well as to other Ste13p homologs. The protein sequence evaluated here was derived from *H. polymorpha* strain NCYC 495 *leu1.1*. For a definitive identification of the protease the genomic sequence needs to be determined in strain KLA8-1 which may show differences to the already known sequence.

Table 3.2: **Comparison of processing of three IL-6 variants in *S. cerevisiae*, *P. pastoris*, *H. polymorpha* and *S. pombe*.**

	S. cerevisiae	***P. pastoris***	***H. polymorpha***	***S. pombe***
A28-IL-6	N-terminal truncation (2 aa) (Guisez et al., 1991)	correct processing (Li et al., 2011)	N-terminal truncation (2 aa) (Manuscript I)	-
P29-IL-6	misprocessed prepro sequence (Guisez et al., 1991)	-	misprocessed prepro sequence (Manuscript I)	correct processing (Giga-Hama et al., 1994)
V30-IL-6	N-terminal truncation (8 aa) (Steinborn et al., 2006)	-	N-terminal truncation (8 aa, strain HP102) (Steinborn et al., 2006), correct processing (strain RB11) (**this study**)	-

(-): not determined.

Subsequently, overexpression of the protease in the ΔS9-15130 deletion strain may be performed to confirm that the removal of Ala-Pro dipeptides is restored.

The sequences of the α-mating pheromone and its precursor have been identified in yeasts other than *S. cerevisiae*, amongst them *S. kluyveri* (Egel-Mitani and Hansen, 1987), *C. albicans* (Lockhart et al., 2003), *Kluyveromyces delphensis* (Wong et al., 2003), *Eremothecium cymbalariae* (Wendland et al., 2011) and *P. pastoris* (Rosales Rodriguez et al., 2011) (see Fig. 3.7). The α-factor precursors show significant similarities. They contain 2 to 9 mating pheromone molecules which are separated by spacer peptides. This sequential arrangement is preceded by a pre- (18 to 22 aa) and a pro-sequence (67 to 105 aa). In all variants the dibasic Lys-Arg motif which is recognized by Kex2p is highly conserved. However differences in the spacer sequences were identified. In all precursors the spacer contains at least 2 and a maximum of 9 dipeptides.

The high similarities of α-mating pheromone precursors suggest the existence of Kex2p and Ste13p homologs for processing of the precursor to mature α-pheromone molecules. The Kex2p cleavage site consisting of Lys-Arg is highly conserved in all cases described. In contrast, the motif recognized by putative Ste13p homologs varies in amino acid composition. Mostly a Xaa-Ala motif is found (for Xaa being Glu, Asp, Val or Ser). Xaa-Pro was detected for *S. kluyveri*. Thus the Ste13 protease is not limited to the cleavage of the Xaa-Ala from the Glu-Ala-Glu-Ala spacer of the natural α-mating factor, but is also able to cleave Xaa-Pro in other contexts. Cleavage of Ala-Pro dipeptides was also described for *S. cerevisiae* Ste13p during A28-IL-6 processing (Guisez et al., 1991).

After search for a repetitive Lys-Arg motif with adjacent amino acids homologous to the various Glu-Ala spacers in the *H. polymorpha* NCYC 495 *leu1.1* genome, we could identify a putative α-mating pheromone precursor. Two α-mating pheromone molecules are again separated by a spacer which shows high homology to the spacers of the other α-mating factor pheromone precursors described in literature. The α-mating pheromone units are followed by a Kex2p recognition site, respectively. However, upstream of the α-mating pheromone molecules no site for Kex2p cleavage was found which would enable removal of the pro peptide. Hence, we suggest that a so far unidentified protease may be responsible for removal of the putative prepro signal sequence. Another difference to other α-mating pheromones is the unusually long putative pro-peptide.

The putative Ste13p homolog and α-factor mating pheromone precursor cannot be solely inferred from sequence homology.

The putative Ste13p homolog HpS9-15130p was identified by prediction of functional domains and substrate specificity. For the candidates tested an Xaa-Pro specificity was predicted. However, the protease identified to be responsible for removal of Ala-Pro dipeptides during IFNα-2a expression only showed 34% sequence homology to the *S. cerevisiae* derived Ste13p. Nevertheless, both proteases share a similar substrate specificity with regard to Ala-Pro. In contrast, the *P. pastoris* derived Ste13p homolog showed 40% sequence homology to the *S. cerevisiae* derived Ste13p. Notably, the putative

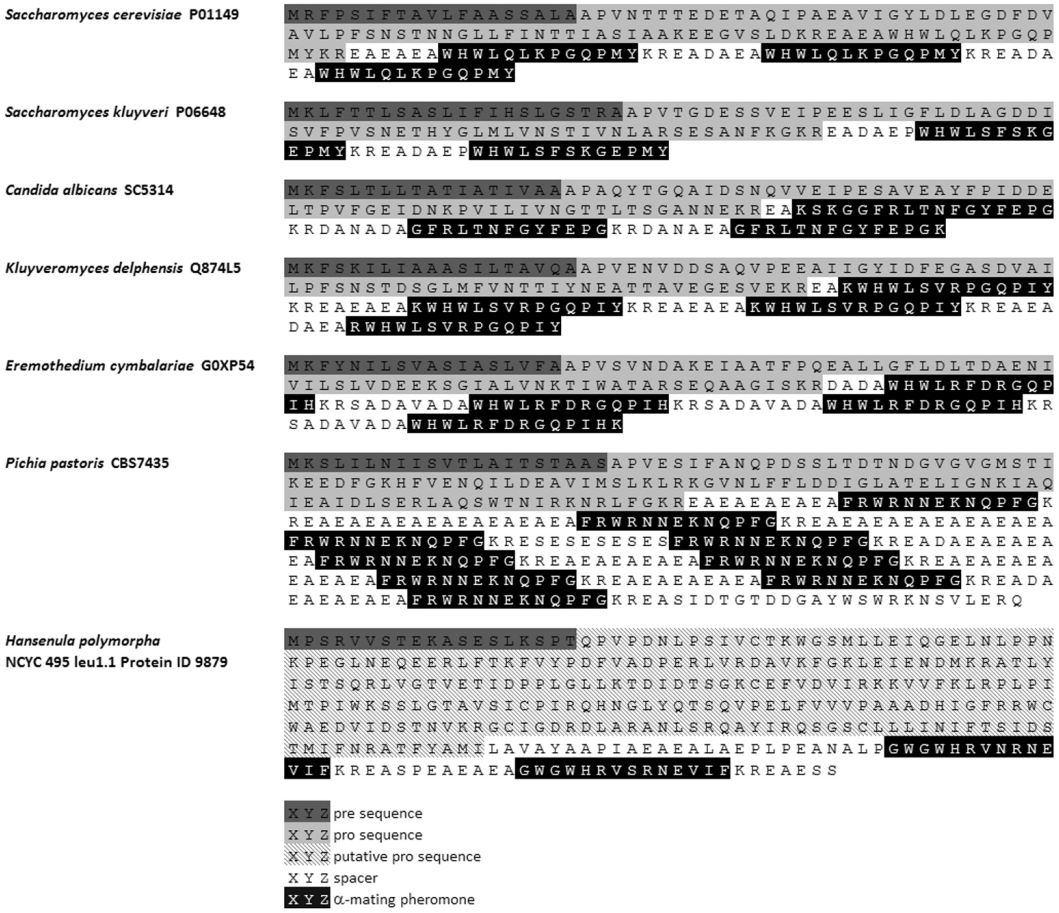


Figure 3.7: **Comparison of α-mating pheromone precursors derived from different yeast.** Structural elements are colored as indicated.

H. polymorpha and *P. pastoris* Ste13p homologs also show a close phylogenetic relation to each other (40% sequence homology). The *S. pombe* homolog shows only 32% sequence homology to the *S. cerevisiae* and the *H. polymorpha* homologs and 35% to the *P. pastoris* homolog. However, the substrate specificity of all three proteases seems to be different (see Tab. 3.2).

protein-protein BLAST analysis of the *H. polymorpha* NCYC 495 *leu1.1* genome did not reveal the existence of a α-mating pheromone precursor. A manual search algorithm was designed for the detection of characteristic processing sites. After search for a repetitive Lys-Arg motif with adjacent amino acids homologous to the various Glu-Ala spacers we could indeed identify a putative α-mating pheromone precursor (Fig. 3.7). Eldarov et al. (2011) found similar phylogenetic relations for concatenated mitochondrial proteins as we did for different Ste13p homologs (Fig. 3.8). Therefore, a close relationship does not necessarily infer a similar substrate specificity and a similar substrate specificity is also possible when two organisms are not closely related. We conclude, that a functional homology of this protein and the *S. cerevisiae* Ste13p can only be inferred from sequence homology concerning the active site.

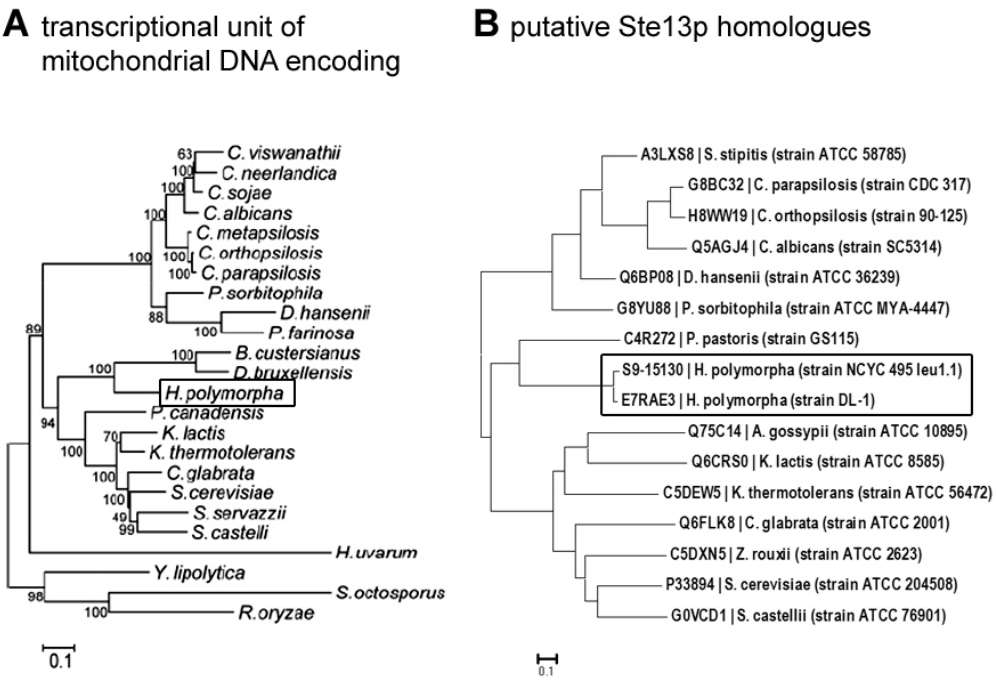


Figure 3.8: **Phylogenetic comparison of mitochondrial protein sequences and Ste13p homologs.** (A) Taken from Eldarov et al. (2011). The phylogenetic trees were calculated from the multiple sequence alignment of 7 concatenated mitochondrial DNA-encoded proteins. Topologies were inferred using Maximum Likelihood (ML). Numbers above the nodes indicate bootstrap support values from 100 replicates for the ML method. The trees are drawn to scale, with branch lengths measured by the number of substitutions per site. (B) For detailed description see Fig. 3.6.

.

The connection between MFα prepro sequence and the protein of interest is crucial for proper cleavage at the Kex2p recognition site.

The amino acids adjacent to the Kex2p cleavage site strongly influence the exposition thereof. It was shown that in dependency of the transition between MFα prepro sequence and protein of interest different processing may occur ranging from thoroughly misprocessing of the MFα prepro sequence over partial misprocessing to correct processing of the complete protein quantity. In *S. cerevisiae* the pheromone molecules are separated by Glu-Ala and Asp-Ala dipeptides. These dipeptides may enhance the hydrophilicity of the spacer which stands in contrast to the hydrophobic pheromone molecules. Zsebo et al. (1986) suggest that this discrepancy may position the spacer on the surface and consequently allow an enhanced accessibility of the Kex2p recognition sequence Lys-Arg. Schaefer and Plückthun (2012) proposed charge dependent cleavage by Kex2p in *P. pastoris*. They detected a misprocessing of the MFα prepro sequence 9 amino acids upstream of the Kex2p recognition site. The cleavage occurred behind a Ala-Lys dipeptide which was shown to have a similar charge distribution in comparison to the Lys-Arg cleavage site. Consequently the authors suggested that Kex2p may be responsible for this misprocessing.

Removal of the pre-sequence may occur upon ER-translocation (Waters et al., 1988), however Julius et al. (1984b) state that the pre-sequence is neither removed on ER-translocation nor afterwards. Ernst

(1988) however could show that the pre-sequence is properly removed and that the pre-sequence solely is sufficient for correct secretion. Therefore pre-sequence removal might also occur when the pro-sequence is removed as well which would be in the late Golgi (Julius et al., 1984b). For enhancement of Kex2p cleavage site recognition exposition of the cleavage site may also be influenced by different chaperones present in the ER. To increase this effect chaperone co-expression is often performed to improve processing of the heterologous protein. Examples for chaperones that conduct transport or folding in the ER are calnexin (Klabunde et al., 2007), Sec61 complex subunit beta (Seb1p) (Keränen et al., 2000), BiP and protein disulfide isomerase (Smith et al., 2004).

Proline in even numbered position to the Kex2p recognition site seems to have a positive effect towards exposition of the Kex2p cleavage site and can increase the rate of proper processing.

In *H. polymorpha*, A28-IL-6 is processed to 100% correctly. Its N-terminal sequence harbors two prolines at even numbered position (APVP). In addition to the correct processing by Kex2p the removal of the Ala-Pro dipeptide is positively influenced by the second proline. In comparison AP-IFNα-2a is only partially processed with regard to the Ala-Pro peptide due to unfavorable folding at this point of dipeptide removal. Kex2p processing however occurs to 100%.

Amongst all amino acids proline holds a unique position. It has no functional groups, hence, it does not take part in any bond-formation (Cunningham and O'Connor, 1997). However, prolines may have other stabilizing or regulatory effects on the protein. Hence, various proline motifs can be found throughout protein sequences. The presence of proline in a peptide chain often causes a bend which disrupts the secondary structure of the protein in that region (Cunningham and O'Connor, 1997). A proline within a helix cause kinks which are suggested to be of structural or functional importance (Barlow and Thornton, 1988). Theses conformational changes may indeed result in an exposition of a nascent cleavage site or in inaccessibility of the peptide bond involving the proline residue (Vanhoof et al., 1995). Consequently even peptidases with a broad specificity are often unable to cleave the latter peptide bond (Walter et al., 1980). Positioned at the N-terminus of a polypeptide chain proline may protect the protein against hydrolysis as well as against proteolysis (Yaron, 1987). Moreover, due to the unique structure proline may be extensively involved in biological regulation (Yaron and Naider, 1993). These kinds of N-terminally positioned prolines can often be found as Xaa-Pro motifs. These dipeptides are removed stepwise and can be found in a wide variety of organisms like yeasts, insects, frogs and humans (Kreil, 1990). Especially amongst cytokines a N-terminal Xaa-Pro motif seems to be conserved and thus indicates certain functionality (Vanhoof et al., 1995).

During the secretion of interleukin-1β in *S. cerevisiae* the N-terminus is correctly processed when using the MFα prepro sequence for secretory translocation (Ernst, 1988). In this case the authentic N-terminus starts with Ala-Pro. The Ala-Pro dipeptide was not removed by the Ste13p, which might indicate that the adjacent amino acids (AP-VR) strongly influence the removal of Ala-Pro dipeptides.

4 Manuscript II: The use of highly expressed FTH1 as carrier protein for cytosolic targeting in *Hansenula polymorpha*

Author's contribution

Journal: Journal of Biotechnology
Impact Factor: 3.045
1^{st} author
Author's contribution: 95%

- Experimental design (together with Manfred Suckow)
- Cloning and strain generation
- Cultivation (except fermentation)
- Protein qualification and quantification
- Authoring manuscript

Journal of Biotechnology 159 (2012) 172–176

Contents lists available at SciVerse ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

The use of highly expressed FTH1 as carrier protein for cytosolic targeting in *Hansenula polymorpha*

Eva Eilert[a], Cornelis P. Hollenberg[b], Michael Piontek[a], Manfred Suckow[a,*]

[a] *ARTES Biotechnology GmbH, Elisabeth-Selbert-Str. 9, Langenfeld (Rheinland), Germany*
[b] *Institut für Mikrobiologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, Düsseldorf, Germany*

ARTICLE INFO

Article history:
Received 7 October 2011
Received in revised form
15 December 2011
Accepted 19 December 2011
Available online 26 December 2011

Keywords:
Hansenula polymorpha
FTH1
Fusion protein
Cytosolic expression

ABSTRACT

The iron storage protein ferritin is a member of the non-heme iron protein family. It can store and release iron, therefore it prevents the cell from damage caused by iron-dioxygen reactions as well as it provides iron for biological processing. To study whether the human ferritin heavy chain (FTH1) can be expressed in *Hansenula polymorpha*, we integrated an expression cassette for FTH1 and analyzed the protein expression. We found very efficient expression of FTH1 and obtained yields up to 1.9 g/L under non-optimized conditions. Based on this result we designed a FTH1-PTH fusion protein to successfully express the parathyroid hormone fragment 1–34 (PTH) for the first time intracellular in *H. polymorpha*.

1. Introduction

Ferritin is the major iron storage protein found in most cell types of higher vertebrates including human and most other living organisms (Theil, 1987). The first function of ferritin is to keep iron in a metabolically accessible form for the synthesis of iron-containing proteins (Munro, 1993; Theil et al., 1997). Bioavailability of iron is solely given when iron is stored as Fe(II). In addition, the second function of ferritin is to prevent damage due to free radicals produced in iron-dioxygen interactions (Arosio et al., 1978; Theil et al., 1997).

All Fe (III) introduced to the cell needs to be converted to Fe (II) and stored for further utilization. Ferritin is able to reduce Fe (III) to Fe (II) and to bind Fe (II) in a spherical protein shell. This shell is composed of two different subunits (heavy and light chain), whereas the proportion of those subunits in ferritin is depending on the type of organism (Harrison and Arosio, 1996). When iron is absent, the spherical protein shell (called apoferritin) is still stable.

The human ferritin heavy chain (FTH1) has been expressed in several expression systems such as *Escherichia coli* (Santambrogio et al., 1993), *Saccharomyces cerevisiae* (Shin et al., 2001) and *Pichia pastoris* (Lee et al., 2003). It has been shown that heterologously expressed ferritin enhances the iron uptake of the studied organisms.

The methylotrophic yeast *Hansenula polymorpha* has been established as a highly efficient expression system in the past (Janowicz et al., 1991). Due to the strong promoters derived from the methanol metabolism pathway and high-copy mitotically stable integration, *H. polymorpha* has been established as an effective system for heterologous protein expression.

In this report we show high level production of functional FTH1 under control of the *FMD* promoter in *H. polymorpha*. Moreover we show that this protein can be used as a fusion partner for difficult-to-express proteins.

2. Materials and methods

2.1. Strains and culture conditions

For construction and propagation of plasmid DNA *E. coli* NEB10ß (*araD139* Δ(*ara-leu*)7697 *fhuA lacX74 galK* (ϕ80 Δ(*lacZ*)*M15*) *mcrA galU recA1 endA1 nupG rpsL* (Str^R) Δ(*mrr-hsdRMS-mcrBC*); New England Biolabs, Frankfurt a.M., Germany) was used. Cultivation took place in Luria Bertani media, when needed supplemented with 50 μg ampicillin per ml.

For protein expression *H. polymorpha* strain RB11 deficient in orotidine-5′-phosphate decarboxylase (*ura3*) (Suckow and Gellissen, 2005) was used. Cultivation took place in either yeast

Abbreviations: FTH1, ferritin heavy chain; FMD, formate dehydrogenase; PTH, parathyroid hormone.

* Corresponding author. Tel.: +49 0 2173 2758715; fax: +49 0 2173 2758777.
E-mail addresses: e.eilert@artes-biotechnology.com (E. Eilert), Cornelis.Hollenberg@uni-duesseldorf.de (C.P. Hollenberg), m.piontek@artes-biotechnology.com (M. Piontek), m.suckow@artes-biotechnology.com (M. Suckow).

0168-1656/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2011.12.014

extract–peptone–dextrose medium (YPD) or yeast nitrogen base medium (YNB). Fermentation was carried out in the synthetic medium SYN6 which had the following composition: 20 g/L glycerol, 13.3 g/L $NH_4H_2PO_4$, 3.3 g/L KCl, 3.0 g/L $MgSO_4 \times 7H_2O$, 0.33 g/L NaCl. After autoclaving of the basis medium 1 ml/L of the following stock solutions was added respectively: micro element stock solution: 6.65 g/L $(NH_4)_2Fe(SO_4)_2 \times 6H_2O$, 0.55 g/L $CuSO_4 \times 5H_2O$, 2.0 g/L $ZnSO_4 \times 7H_2O$, 2.65 g/L $MnSO_4 \times H_2O$, 6.65 g/L EDTA (Titriplex III); vitamin stock solution: 40 mg/L D-Biotin, 13.35 g/L thiamine chloride hydrochloride; trace elements stock solution: 65 mg/L $NiSO_4 \times 6H_2O$, 65 mg/L $CoCl_2 \times 6H_2O$, 65 mg/L H_3BO_3, 65 mg/L KJ, 65 mg/L $Na_2MoO_4 \times 2H_2O$; calcium stock solution: 100 g/L $CaCl_2 \times 2H_2O$. All cultivations were performed at 37 °C.

2.2. Vector construction

To generate the expression plasmid pFPMT-M-FTH1-H6, a synthetic open reading frame based on Universal Protein Resource Knowledgebase (UniProt ID: P02794) was inserted into with the polylinker of the vector pFPMT121 (Degelmann et al., 2002). A 6 His-tag was added C-terminally to the protein sequence. The cloned FTH1-H6 sequence with additional flanking recognition sites (5′ *Eco*RI and 3′*Bam*HI) was synthesized by GeneArt (Regensburg, Germany). To generate the expression plasmid pFPMT-M-FTH1-G3-Ek-PTH, a synthetic DNA fragment encoding a three-glycin-spacer (G3), an enterokinase cleavage site (Ek) and PTH (UniProt ID: P01270, amino acids 1–34) flanked by 5′ *Bst*EII and 3′ *Bam*HI sites was synthesized by GeneArt and ligated with the corresponding fragment of pFPMT-FTH1-H6. The structure of the encoded fusion protein is: FTH1-G3-Ek-target protein. *H. polymorpha* was transformed by electroporation as described in Faber et al. (1994). Afterwards strain generation was performed according to Guengerich et al. (2004). The passaging procedure lasted for four passages.

2.3. Protein expression and preparation of protein material

Transformant strains were subjected to derepression. Therefore, the cells were cultivated in 3 ml scale in YNB/2% glycerol for 48 h and subsequently in YNB/1% methanol for 24 h at 37 °C. After cultivation the cells were harvested by centrifugation. The supernatant was collected and stored at −20 °C. The total cell extract was obtained by glass bead cell disruption and stored at −20 °C.

2.4. Protein purification

For purification protein samples have been incubated for 15′ at 75 °C. Subsequently the samples have been centrifuged for 15′ at 16,000 × *g* to separate insoluble and soluble fraction.

2.5. Protein analysis

For molecular weight determination the samples were analyzed by SDS–PAGE according to Laemmli (Laemmli, 1970). For the electrophoretic separation 4–12% Bis–Tris acryl amide XT gels (BioRad, Hercules, CA, USA) were used. The gels were stained either with the Coomassie stain SimplyBlue (Invitrogen, Karlsruhe, Germany) or Prussian blue as described by Kim et al. (2003), or the protein patterns were was transferred to nitrocellulose-membranes. The protein bands were then visualized by immunoblotting. For immunological detection a 6 His-tag-specific antibody (Micromol, Karlsruhe, Germany) or a PTH-specific antibody (BIOTREND, Cologne, Germany) was applied. As secondary antibody a goat-anti-rabbit-AP-conjugate or a goat-anti-mouse-AP-conjugate (both BioRad) was applied respectively. Development was performed by using NBT/BCIP (Roche, Mannheim, Germany). Native PAGE samples were resolved on 5% Tris–HCl acryl amide separation gels (BioRad). Prior to adjustment to 1× native and non-reducing sample buffer conditions aliquots of the samples were adjusted to 1 mM ammonium Fe(II) sulphate followed by 15′ incubation at RT.

2.6. Copy number determination

The number of genomically integrated expression cassettes was determined by Southern blot analysis. Genomic DNA was prepared using the MasterPure Yeast DNA Purifcation Kit (Epicentre, Madison, WI, USA), restricted using *Bam*HI and *Kpn*I and precipitated afterwards. Dilutions of the genomic DNA from 1.2 μg to 19 ng were separated by gel electrophoresis. All further steps were carried out according to the Amersham ECL Direct Nucleid Acid Labelling And Detection System (GE Healthcare UK Limited, Buckinghamshire, UK). For detection of the expression cassette a *FMD* promoter probe was used. The copy numbers of the expression cassettes (FTH1: 1048 bp, FTH1-G3-Ek-PTH: 1156 bp) were determined in comparison to the homologous *FMD*-gene (FMD: 756 bp).

2.7. Fermentation

The fermentations were carried out in a 1.25 L Bioflo III fermenter (New Brunswick Scientific, Edison, NJ, USA). 100 ml YPD preculture were grown for 22.5 h. Fermentation was started by adjusting an initial optical density of OD_{600nm} = 1 in a start culture volume of 1 L.

The aeration rate was 1 vvm using a ring sparger. The cultivation temperature was adjusted to 37 °C. Agitation was performed by two stirrer blades (speed between 500 and 1000 rpm). The glycerol feed was started after 18 h with a fixed feed rate of 1 g/L/h and raised stepwise to a final rate of 2.5 g/L/h. During the derepression phase the pO_2 was kept at 40% by a stirrer cascade between 500 and 1000 rpm. The pH was adjusted and maintained at 4.5 using a 12.5% (v/v) ammonia solution. Sterile diluted antifoam agent (10% (v/v) Structol J 673) was added manually as appropriate. Feed was performed using a 75% (w/v) glycerol solution. The methanol induction was performed subsequently (after 123.25 h) by using a solution of 20% (v/v) glycerol and 80% (v/v) methanol according to the following scheme: 0 h/5 h: 1% (v/v); 8 h/12 h/15 h/18 h: 0.5% (v/v). Fermentation was terminated after 141.85 h.

During fermentation samples were taken and analyzed offline. The optical density (OD) was determined at 600 nm wavelength, dry cell weight was analyzed, and product was analyzed by Coomassie-stained SDS-protein gels. The total amount of the target proteins after the fermentation was estimated by using BioDocAnalyze (Biometra, Göttingen, Germany) in correlation to the optical density.

2.8. Enterokinase cleavage

FTH1-G3-Ek-PTH fusion proteins were treated with enterokinase to cleave off the target protein PTH. Enterokinase was supplied by New England Biolabs. The Enterokinase/fusion protein ratio in the reaction was 1:80,000.

3. Results

3.1. High level expression of human ferritin heavy chain in H. polymorpha

To study the expression of FTH1 in *H. polymorpha*, a synthetic FTH1 ORF was designed and inserted into the *H. polymorpha* expression vector pFPMT121. To provide the option of simple detection and purification, codons for a C-terminal His tag were added to

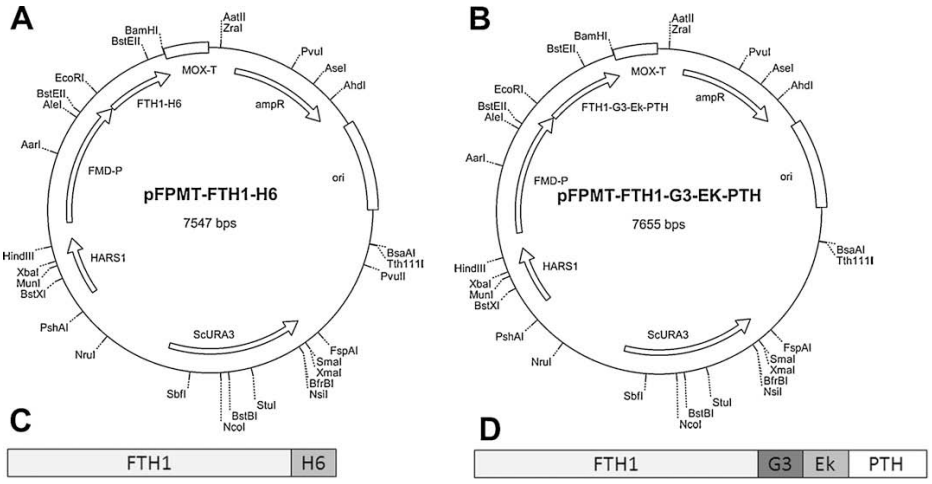


Fig. 1. Expression plasmids and fusion proteins. (A) pFPMT-FTH1-H6 (7.547 kb). (B) pFPMT-FTH1-G3-Ek-PTH (7.645 kb). (A) and (B) In both maps a set of single cutters and the cloning relevant site of the double cutter *Bst*EII are indicated. Features and abbreviations: FMD-P, promoter of the *H. polymorpha* formate dehydrogenase (*FMD*) gene; MOX-T, terminator of the *H. polymorpha* methanol oxidase (*MOX*) gene; ampR, ampicillin resistance gene; ori, origin of replication of pBR322; ScURA3, *URA3* gene of *S. cerevisiae* for selection in *H. polymorpha ura3* strains; HARS1, autonomously replicating sequence 1 of *H. polymorpha*; (C) and (D) composition of the FTH1-H6 and FTH1-G3-Ek-PTH fusion proteins encoded by the plasmids shown in (A) and (B), schematically. Abbreviations: FTH1, complete sequence of human ferritin heavy chain; H6, six-histidine tag; G3, three glycine residues as a flexible linker; Ek, enterokinase cleavage site (DDDDK); PTH, fragment of the human parathyroid hormone.

the FTH1 ORF. In the resulting expression plasmid pFPMT-FTH1-H6 (Fig. 1A and C) the FTH1-H6 ORF transcription is controlled by the strong carbon source regulated *FMD* promoter. To generate FTH1-H6 producer strains *H. polymorpha* RB11 (*ura3*) was transformed with pFPMT-FTH1-H6, and strains were cultivated through four passaging steps and two subsequent stabilisation steps. This procedure typically results in the isolation of strains harbouring multiple copies of the expression plasmid in a head-to-tail arrangement which can differ with regard to the exact copy number and the achieved expression rate of the heterologous gene (Gellissen and Hollenberg, 1997). A couple of strains resulting from this procedure were subjected to small scale methanol induction, and the total cell extracts were analyzed for the presence of FTH1-H6 gene products by SDS–PAGE/Western blot using a His-tag-specific primary antibody. Several strains displayed a band corresponding to the expected size of the FTH1-H6 protein (data not shown).

The strain with the highest productivity (FTH1-H6 # 2-1) was chosen for further analyses. Fig. 2 shows a Ponceau S stain and a Western blot analysis of the intracellular material of strain FTH1-H6 # 2-1 after derepression and MeOH induction. The FTH1-H6 fusion protein has a calculated molecular weight of 22.05 kDa. Fig. 2A shows the NC membrane after semi-dry transfer in Ponceau S stain; this method visualises all protein bands. Fig. 2B shows the corresponding immunodetection via a His-tag specific antibody. In both lanes corresponding to the ferritin strain (FTH1-H6) a band of very high intensity displaying an apparent MW of about 23 kDa was observed which is in agreement with the calculated MW. Due to the specific recognition by the antibody the C-terminal part of the protein must be intact. The supernatant of the FTH1-H6 # 2-1 culture did not display any ferritin which is consistent with intracellular targeting of the fusion protein (not shown). The vector control strain did not provide any His-tag-specific signals. Hence, the expression of the cytosolic FTH1-H6 provided high amounts of the heterologous protein.

Strikingly, FTH1-H6 is by far the dominant protein detected in the crude cell extract by Ponceau S stain (Fig. 2A) as well as Coomassie stain (not shown). The ferritin band even exceeded bands corresponding to proteins of the methanol pathway like Moxp and Fmdp (Strasser, 1988). We estimate that the amount of FTH1-H6 is in the order of at least 5 times the amount of Moxp. Based on these results FTH1-H6 has been chosen to test its function as a fusion partner for the expression of poorly and/or difficult-to-express proteins.

3.2. High level expression of a parathyroid hormone-fragment upon fusion to ferritin

The expression of the FTH1-H6 ORF in *H. polymorpha* revealed high-level synthesis and cytosolic accumulation. We therefore wondered whether FTH1 fusion proteins can be expressed in similar amounts in *H. polymorpha* which would be an option for difficult to express polypeptides.

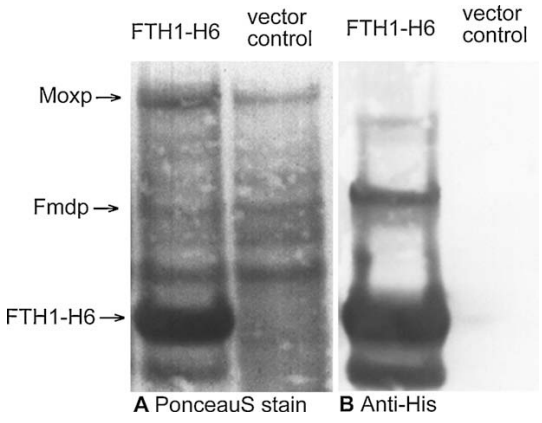


Fig. 2. SDS–PAGE of the total cell extract of a FTH1-H6 producing strain and a vector control strain. The indicated strains were subjected to a methanol induction at the 3 ml scale. The total cell extracts were prepared and analyzed by SDS–PAGE/Western Blot. Protein was visualised with Ponceau S stain (A) and with immunodetection via a His-tag specific antibody (B). The vector control strain used in this experiment contains genomically integrated copies of the pFPMT121 vector without insert, Moxp: methanol oxidase, Fmdp: formate dehydrogenase.

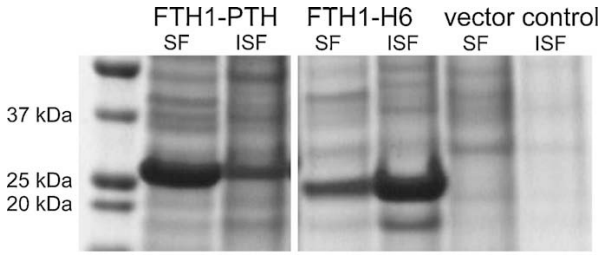


Fig. 3. SDS–PAGE of the intracellular soluble and insoluble fraction of *H. polymorpha* strains expressing FTH1-PTH and FTH1-H6. Intracellular soluble fractions and insoluble fractions of indicated strains derived from cultivation similar as described in Fig. 2. The samples were analyzed on a 4–12% Bis–Tris protein gel. The fraction containing the protein was identified by Coomassie Blue stain. SF: soluble fraction, ISF: insoluble fraction.

The N-terminal fragment of the human parathyroid hormone (1–34), furthermore referred to as PTH, has not yet been expressed in *H. polymorpha* as a stable, secreted peptide. Previous attempts to express and secrete PTH in fusion with the MFalpha-prepro sequence showed limited success (data not shown). A possible explanation is that PTH may be instable when introduced to the secretory pathway due to protease activity residing in ER and/or Golgi. It therefore appeared interesting to express PTH cytosolically. The purification of a 34 amino acid PTH fragment from cell extract may be difficult while the handling of a ferritin-PTH fusion protein (225 amino acids) may be easier. The scheme of the FTH1-PTH fusion protein is shown in Fig. 1d. FTH1 forms the N-terminal part of the fusion protein followed by a flexible three-glycine linker, an enterokinase cleavage site, and the PTH. A synthetic ORF encoding this FTH1-G3-Ek-PTH fusion protein was placed into pFPMT121; the resulting pFPMT-FTH1-G3-Ek-PTH is shown in Fig. 1b. The expression plasmid has been introduced into *H. polymorpha* RB11, and strain generation was accomplished as described above. FTH1-G3-Ek-PTH was expressed at very high levels in several strains. The levels were at least of the magnitude as those observed with FTH1-H6. The strain with the highest productivity was chosen for further analyses (FTH1-PTH # 20-1).

3.3. The fusion with PTH increases the solubility of ferritin

In the experiments described above total cell extracts have been analyzed. To analyze and compare the degree of solubility of *H. polymorpha*-produced FTH1-H6 and FTH1-G3-Ek-PTH the intracellular soluble and insoluble fractions of the producer strains were prepared upon methanol induction and subjected to SDS–PAGE and Western blot. As shown before only about 20% of FTH1-H6 were detected in the intracellular soluble fraction, the majority of 80% was found in the insoluble fraction (Fig. 3). Surprisingly, in the case of the FTH1-G3-Ek-PTH fusion protein the values were reciprocally exchanged: about 80% were detected in the intracellular soluble fraction and only 20% was found in the insoluble fraction. Thus, the C-terminal extension of the 34 amino acids residues of the PTH fragment to FTH1 causes a dramatic increase of the solubility of ferritin in the cytosol of *H. polymorpha*. This unexpected result may constitute an improvement for subsequent downstream processing which typically starts with the soluble intracellular fraction. A partial cleavage of the PTH part by digestion of the fusion protein with enterokinase was achieved but only below about 2% which was not enough for a recovery (not shown).

3.4. Both ferritin and ferritin-PTH are highly expressed in H. polymorpha

The total initial yields of intracellular soluble FTH1-H6 and FTH1-G3-Ek-PTH produced by *H. polymorpha* were determined by comparison of dilution series of each sample with a defined amount of protein standard (Interferon-α 2a) by SDS-PAGE. The analyzed material was derived from a non-optimized 1 L fermentation of the respective producer strains subjected to derepression and methanol induction (Fig. 4). The total amount of FTH1-H6 was estimated at 1.9 g/L (OD_{600nm} = 284.6), the amount of FTH1-G3-Ek-PTH was estimated at 1.3 g/L (OD_{600nm} = 254.4) corresponding to 0.4 g (FTH1-H6) and 1.0 g (FTH1-G3-Ek-PTH) present in the soluble fraction (see above). We analyzed the copy number of the expression cassette in both strains by Southertn blot and found 30 copies per haploid genome in both cases (data not shown).

3.5. The ferritin product can form iron-binding protein shells

The ability to bind iron and form functional ferritin spherical shells was analyzed for both FTH1-H6 and FTH1-G3-Ek-PTH. In order to find a suitable temperature for heat precipitation of ferritin or the ferritin-PTH fusion protein we tested the spectrum between 55 °C and 95 °C. Incubation for 15 min at 75 °C appeared to be the best condition. Partially purified samples were resolved by native PAGE. Subsequently, the protein was visualized by Coomassie Blue and Prussian blue stain (Fig. 5). The Coomassie stain indicated two different protein conformations for FTH1-H6 (A, bands 1 and 2). FTH1-G3-Ek-PTH showed only one of these protein conformations (A, band 1). After staining with Prussian Blue band 2 of the FTH1-H6 strain also provided an iron signal when the sample was preincubated with 1 mM ammonium Fe (II) sulphate (B). The protein bands indicated with 2 in both gels (A and B) comigrated exactly. The particle band occurring without treatment with ammonium Fe (II) sulphate did not display an iron signal by Prussian Blue staining. Thus, the Prussian Blue stain indicated a Fe (II) binding structure of the FTH1-H6 protein samples. This hints at the formation of a functional protein shell *in vitro*. The material of FTH1-G3-Ek-PTH

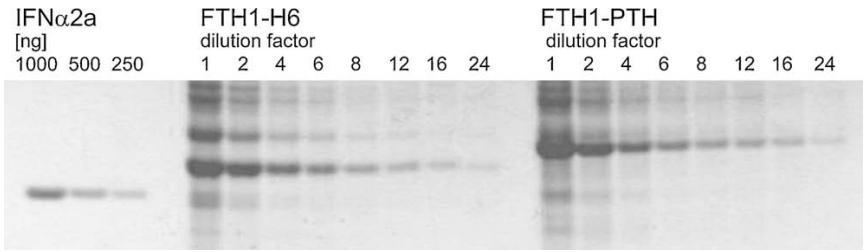


Fig. 4. Coomassie stain to estimate total amount of FTH-H6 and FTH-G3-Ek-PTH in whole cell extract by comparison with a standard protein in SDS–PAGE Total cell extracts of the indicated strains derived from fermentation in 1 L scale were diluted in defined steps (compare dilution factors). Interferon-α 2a has been used as a standard protein in defined amounts for comparison of protein amounts. The various samples have been analyzed on a 4–12% Bis–Tris protein gel. The gel has been stained with Coomassie Blue and scanned.

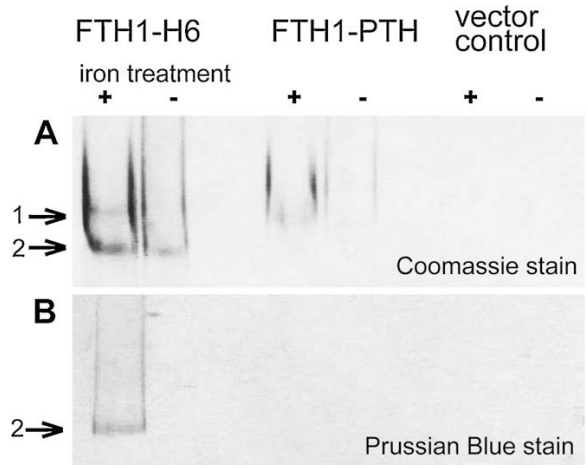


Fig. 5. Analysis of particle character and iron content of intracellular soluble FTH1-H6 and FTH1-PTH by native PAGE under non-reducing conditions. One gel has been stained with Coomassie Blue to show protein bands (A), one gel with Prussian Blue to indicate Fe(III) (B). The samples have been analyzed upon different treatments: + adjusted to 1 mM ammonium Fe (II) sulphate and subsequent 15′ incubation at RT, −: without any further treatment. 1: non-iron binding FTH1 structure, 2: iron binding FTH1 structure.

could not be stained with Prussian Blue, indicating that the fusion protein does not form iron binding structures *in vitro*.

4. Discussion

Functional FTH1 has been expressed for the first time as an intracellular located protein in the methylotrophic yeast *H. polymorpha*. The expression of the FTH1-H6 ORF in *H. polymorpha* revealed high-level synthesis and cytosolic accumulation of the target protein. Prussian blue staining indicated that the FTH1 expressed in this study appeared to assemble into a spherical shell with the ability to store iron. Coomassie Blue staining indicated that the formation of the spherical shell occurs even without addition of iron. FTH1 was expressed with yields up to 1.9 g/L. The yield of FTH1 was almost 100 times higher than reported for other expression systems like *P. pastoris* (25.7 mg/L) (Lee et al., 2003) underlining that *H. polymorpha* is an excellent host system for the production of particles.

FTH1-H6 has been chosen to test its possible function as a fusion partner for the expression of poorly and difficult-to-express proteins. The generation of a FTH1 fusion protein is the first approach for a carrier-protein-driven cytosolic targeting in *H. polymorpha*. The fusion protein was expressed by *H. polymorpha* with yields up to 1.3 g/L. The protein level is comparable to the amount of FTH1 reported for *H. polymorpha* in this study. The expressed fusion protein FTH1-G3-Ek-PTH had surprisingly a dramatically increased solubility in comparison to FTH1-H6.

Due to poor enterokinase cleavage, only a fraction of the fused PTH could be retrieved from FTH1 (data not shown). To improve recovery of the target protein different protease cleavage sites or spacer peptides may be tested and optimized for general function with any fusion partner. The fusion protein was very stable and PTH was not degraded neither by any present protease nor during heat purification. This implicates that cytosolic expression of PTH might be favourable over a secretory expression. In the case of *P. pastoris* the addition of EDTA was needed for the expression of intact PTH (Vad et al., 2005).

This study marks a successful proof-of-concept for the use of FTH1 as a fusion partner for difficult-to-express proteins like PTH for cytosolic expression. In addition the recovery of the fused PTH by application of enterokinase was shown to be possible. One other example for the use of a highly expressed protein as fusion partners in *H. polymorpha* is the amine oxidase, which has been used as a carrier for peroxisomal targeting in *H. polymorpha* (Faber et al., 1996). However use of a FTH1 fusion protein presented in this communication is the first approach for a carrier-protein-driven cytosolic targeting in *H. polymorpha*. To further optimize the concept of FTH1 to function as a fusion partner for improvement of the expression of difficult-to-express proteins further fusion partners will be tested.

Acknowledgements

We would like to thank Sylvia Denter and Michael Weniger for carrying out the fermentation and additional evaluations. Furthermore excellent technical assistance was provided by Manuela Bußmann, David Wilson and Nina Wingbermühle.

The research work was supported by grants from the German National Ministry of Education and Research (BMBF) within the ETB EuroTransBio program.

References

Arosio, P., Adelman, T.G., Drysdale, J.W., 1978. On ferritin heterogeneity. J. Biol. Chem. 253, 4451–4458.
Degelmann, A., Müller, F., Sieber, H., Jenzelewski, V., Suckow, M., Strasser, A.W., Gellissen, G., 2002. Strain and process development for the production of human cytokines in *Hansenula polymorpha*. FEMS Yeast Res. 2, 349–361.
Faber, K.N., Haima, P., Harder, W., Veenhuis, M.A.B G., 1994. Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*. Curr. Genet. 25, 305L 310.
Faber, K.N., Westra, S., Waterham, H.R., Keizer-Gunnink, I., Harder, W.A.B., Veenhuis, G.M., 1996. Foreign gene expression in *Hansenula polymorpha* A system for the synthesis of small functional peptides. Appl. Microbiol. Biotechnol. 45, 72–79.
Gellissen, G., Hollenberg, C.P., 1997. Application of yeasts in gene expression studies: a comparison of Saccharomyces cerevisae. Hansenula polymorpha and Kluyveromyces lactis – a review. Gene 190, 87–97.
Guengerich, L., Kang, H.A., Behle, B., Gellissen, G., Suckow, M., 2004. A platform for heterologous gene expression based on the methylotrophic yeast *Hansenula polymorpha*. In: Küeck, U. (Ed.), Genetics and Biotechnology. Springer Verlag, Berlin, Heidelberg.
Harrison, P.M., Arosio, P., 1996. The ferritins: molecular properties, iron storage function and cellular regulation. Biochim. Biophys. Acta 1275, 161–203.
Janowicz, Z.A., Melber, K., Merckelbach, A., Jacobs, E., Harford, N., Comberbach, M., Hollenberg, C.P., 1991. Simultaneous expression of the S and L surface antigens of hepatitis B, and formation of mixed particles in the methylotrophic yeast, *Hansenula polymorpha*. Yeast 7, 431–443.
Kim, H.J., Kim, H.M., Kim, J.H., Ryu, K.S., Park, S.M., Jahng, K.Y., Yang, M.S., Kim, D.H., 2003. Expression of heteropolymeric ferritin improves iron storage in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 69, 1999–2005.
Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.
Lee, J.L., Song, H.S., Kim, H.J., Park, J.H., Chung, D.K., Park, C.S., Jeung, D., Kim, H.Y., 2003. Functional expression and production of human H-ferritin in *Pichia pastoris*. Biotechnol. Lett. 25, 1019–1023.
Munro, H., 1993. The ferritin genes: their response to iron status. Nutr. Rev. 51, 65–73.
Santambrogio, P., Levi, S., Cozzi, A., Rovidaj, E., Albertini, A., Aros, P., 1993. Production and characterization of recombinant heteropolymers of human ferritin H and L chains. J. Biol. Chem. 268, 12744–12783.
Shin, Y.-M., Kwon, T.-H., Kim, K.-S., Chae, K.-S., Kim, D.-H., Kim, J.-H., Yang, M.-S., 2001. Enhanced iron uptake of *Saccharomyces cerevisiae* by heterologous expression of a tadpole ferritin gene. Appl. Environ. Microbiol. 67, 1280–1283.
Strasser, A.W.M., 1988. Industrielle Nutzung von Mikroorganismen. BioEngineering. Forschung und Praxis 2, 162–163.
Suckow, M., Gellissen, G., 2005. The expression platform based on *H. polymorpha* strain RB11 and its derivatives – history, status and perspectives. In: Gellissen, G. (Ed.), *Hansenula polymorpha*: Biology and Applications. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG.
Theil, E.C., 1987. Ferritin: structure, gene regulation, and cellular function in animals, plants, and microorganisms. Annu. Rev. Biochem. 56, 289–315.
Theil, E.C., Burton, J.W., Beard, J.L., 1997. A sustainable solution for dietary iron deficiency through plant biotechnology and breeding to increase seed ferritin control. Eur. J. Nutr. 51, 28–31.
Vad, R., Mafstad, E., Dahl, L.A., Gabrielsen, O.S., 2005. Engineering *Pichia pastoris* expression system for secretion of high amounts of intact human parathyroid hormone. J. Biotechnol. 116, 251–260.

Corrigendum

On page 172 paragraph 1, the iron uptake and storage by ferritin was described incorrectly. The 1st paragraph in the introduction should have read: All Fe(II) imported into the cell needs to be converted to Fe(III) and stored for further utilization. Ferritin is able to oxidize Fe(II) to Fe(III) and to bind (Fe(III) in a spherical shell. On page 175 section 3.5 Fe(II) and Fe(II) were interchanged. Hence, Prussian Blue staining was performed to indicate Fe(III) binding structures and not Fe(II).

4.2 Supplementary Data

4.2.1 Abstract

The manuscript is supplemented with further data regarding heat precipitation, copy number determination and enterokinase cleavage. Furthermore, a strain expressing a FTH1-G3-Ek-IFNα-2a fusion protein was generated to evaluate FTH1 as a suitable carrier protein to perform intracellular protein expression. Subsequently expression and solubility of the fusion protein was evaluated. The fusion protein was successfully expressed, but yields were estimated to be lower than secretory expression of IFNα-2a. Fusion of IFNα-2a to FTH1 resulted in the same distribution of the expressed protein among soluble and insoluble fraction as FTH1 alone. Overall, FTH1 was identified as a suitable carrier protein for the expression of intracellular targeted fusion proteins.

4.2.2 Introduction

To study whether the human ferritin heavy chain (FTH1) can be expressed in *H. polymorpha*, we integrated an expression cassette for FTH1-H6 and found high levels of protein expression. After MeOH induction the amount of FTH1 even exceeded the amounts of Fmdp and Moxp, two abundant enzymes of the methanol pathway. Due to this high expression level FTH1 was selected to be applied as a carrier protein for intracellular protein expression. For initial studies a model protein was chosen which could already be expressed in *H. polymorpha*. In Manuscript II, the expression of a FTH1-G3-Ek-PTH fusion protein is described. In that case PTH was chosen because it was found to be unstable when secreted to the medium. Prior to the study described in Manuscript II a proof-of-concept was performed using a protein which has already been expressed in *H. polymorpha* and which was found to be stable when secreted. On this account, IFNα-2a was chosen as the model protein resulting in successful intracellular expression of the fusion protein FTH1-G3-Ek-IFNα-2a.

4.2.3 Materials and methods

Vector construction

To generate the expression plasmids pFPMT-M-FTH1-G3-Ek-IFNα-2a, a synthetic DNA fragment of 0.193 kbp encoding a three-glycin-spacer (G3), an enterokinase cleavage site (Ek), and the initial codons of the IFNα-2a ORF flanked by 5' *Bst*EII and 3'*Bgl*II sites was synthesized by GeneArt and ligated with the 0.453 kbp fragment of *Eco*RI and *Bst*EII restricted pFPMT-FTH1-H6 and the 7.406 kbp fragment of *Bgl*II and *Eco*RI restricted pFPMT-MFa(D83E)-IFNa-2a. The structure of the encoded fusion protein is: FTH1-G3-Ek-target protein. The passaging procedure lasted for four passages.

Protein analysis

IFNα-2a as part of the fusion protein was detected as described in Manuscript I.

Heat precipitation

For sample preparation the crude cell extract was centrifuged for 15 min at 16,0000 x g to separate soluble and insoluble fraction. The insoluble fraction was dismissed and the soluble fraction was subjected to heat precipitation. For this the sample was placed in a preheated thermoblock. The temperature and the time are specified in the 'Results' section. Subsequently the sample was centrifuged for 15 min at 16,0000 x g and the soluble fraction was analyzed by Western blot.

Enterokinase cleavage

FTH1-G3-Ek-PTH fusion proteins were treated with enterokinase to cleave off the target protein PTH. Enterokinase was supplied by New England Biolabs. The Enterokinase/fusion protein ratio in the reaction was 1:80,000. The reaction was performed under the following buffer conditions: 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 2 mM $CaCl_2$ and 0.1% Triton X-100. Incubation was conducted at 23°C.

4.2.4 Results

Heat precipitation

Purification of a ferritin containing sample by heat precipitation has already been published by Linder and Munro (1972). Amongst other purification steps a homogenate from liver was treated for 10 min at 70°C. After centrifugation, 95% of the ferritin were recovered. In order to find a suitable temperature

for heat precipitation of FTH1-H6 and the FTH1-G3-Ek-PTH fusion protein we tested a spectrum between 55°C and 75°C. In addition two time durations were tested (15 min and 30 min). Precipitation was carried out using crude cell extract containing either FTH1-H6 or FTH1-G3-Ek-PTH. After incubation the samples were centrifuged and the soluble fractions were harvested. Subsequently the soluble fractions were separated via SDS-PAGE and blotted on a nitrocellulose membrane. Finally, the entire protein contained in the samples was detected via unspecific Ponceau S staining (see Fig. 4.1). Almost no contaminants were removed when samples were incubated at 55°C. In contrast, some contaminants were removed at 65°C and 75°C (see bands at ~39 kDa and ~75 kDa). At 75°C FTH1-G3-Ek-PTH showed significant signs of degradation after 30 min of incubation time. However, incubation for 15 min at this temperature appeared to be the condition with highest purification while only small amounts of the target protein were deprived.

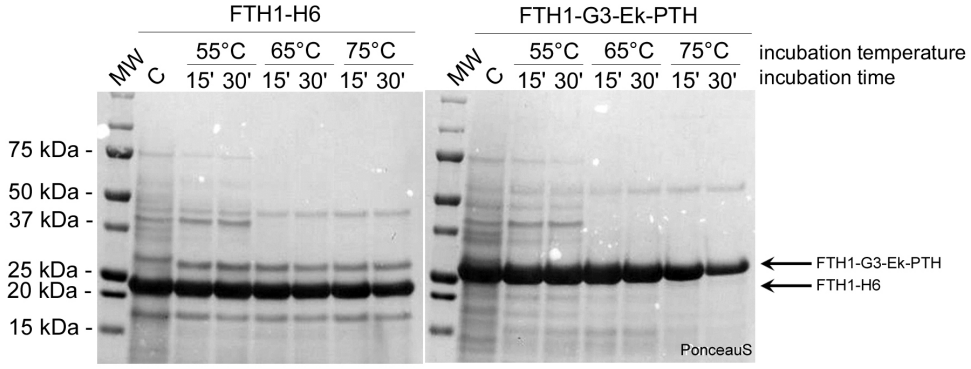


Figure 4.1: **Basic purification of FTH1-H6 and FTH1-G3-Ek-PTH from crude cell extracts by heat precipitation.** Samples of the crude cell extracts were incubated at the indicated temperatures for either 15 min or 30 min. After subsequent centrifugation the samples were separated on a 4-12% Bis-Tris protein gel. After transfer to a nitrocellulose membrane all proteins were visualized with unspecific Ponceau S staining. MW: molecular weight marker, C: initial sample (control).

Copy number determination

The expression level of a heterologous protein may be influenced by different factors such as the codon usage, gene dosage, promoter, integration locus, A+T composition of cDNA, endogenous protease activity and cultivation conditions (Sreekrishna et al., 1997). The expression levels of Moxp and Fmdp under MeOH induction exceed those of all other homologous proteins. This is an example for the impact of promoter and gene location since both genes are only present as one copy per haploid genome. It has however been shown, that the expression of a protein was highly influenced by the copy number (Weydemann et al., 1995; van der Heide, 2002). In Manuscript II we have shown that FTH1-H6 and FTH1-G3-Ek-PTH are highly expressed in *H. polymorpha* in gram per liter scale. Next we aimed to determine the copy number of the expression cassette integrated into the genome. To determine the copy number per haploid genome the expression cassette under control of the *FMD* promoter was detected using a probe which hybridizes within the promoter region (Gellissen et al., 1992). Genomic DNA was restricted using *Kpn*I/*Bam*HI and separated by gel electrophoresis. Subsequently the DNA pattern was transferred onto a nylon membrane. Probe hybridization and

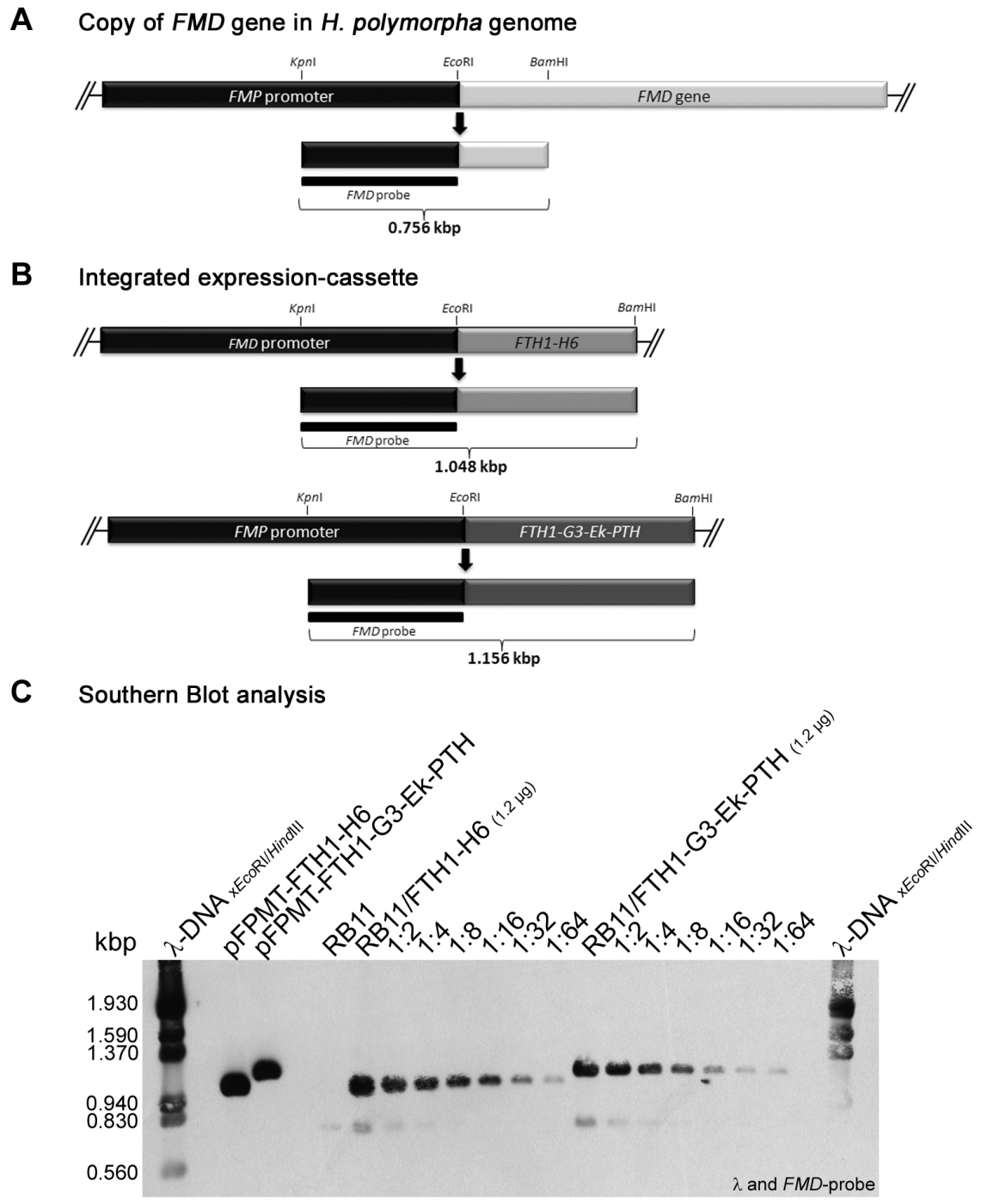


Figure 4.2: **Determination of copy number of expression cassette in FTH1-H6 and FTH1-PTH expression strains by Southern blot.** A + B: Schematic overview on resulting fragments for copy number determination and hybridization of the detection probe, C: Southern Blot analysis: detection of the expression cassettes in genome of RB11/pFPMT-FTH1-H6 and RB11/pFPMT-FTH1-G3-Ek-PTH. Genomic DNA was restricted using *Kpn*I and *Bam*HI. The expression cassettes were detected using a fragment of the *FMD* promoter as the probe labeled with horseradish peroxidase, the marker was detected using horseradish peroxidase labeled λ-DNA. For copy number determination the signal intensity for the expression cassette with unknown copy number was compared to the signal of the single copy of *FMD* per haploid genome. We estimated 30 copies per haploid genome in both cases. kbp: kilo base pairs

expected fragment size are shown in Fig. 4.2 A + B. The copy numbers of the expression cassettes (*FMD-FTH1-H6*: 1.048 kbp, *FMD-FTH1-G3-Ek-PTH*: 1.156 kbp) were determined in comparison to the homologous *FMD* gene (*FMD*: 0.756 kbp). Therefore, the signal intensity of the heterologous expression cassette was compared to the signal intensity of the homologous *FMD* gene. In both cases we found 30 copies per haploid genome (Fig. 4.2 C).

Recovery of PTH from the fusion protein

To generate mature PTH the carrier protein FTH1 part has to be removed. PTH is fused to FTH1 via a short three-glycine linker with an adjacent enterokinase recognition site (Asp-Asp-Asp-Asp-Lys, DDDDK). The enterokinase cleavage site was chosen because upon efficient cleavage PTH is released with authentic N-terminus. To remove the PTH moiety from the carrier protein a sample was treated with enterokinase for a total of 16 h. A reference sample was withdrawn at the beginning of the reaction immediately after addition of enterokinase, further samples were taken after 30 min, 1, 2, 4, 8, and 16 h. An additional sample was incubated in parallel without addition of enterokinase. These samples were analyzed for unspecific degradation or proteolysis of the fusion protein over time of the experiment (Fig. 4.3). Partial cleavage of PTH of about 2% was achieved after 1 to 4 h of digestion (see Fig. 4.3, lanes indicated by arrows below). After 6 h, degradation of the removed PTH occurred. The amount of PTH released from the fusion protein was very low, therefore, no purification of PTH was performed. The fusion protein remained stable throughout the entire experiment.

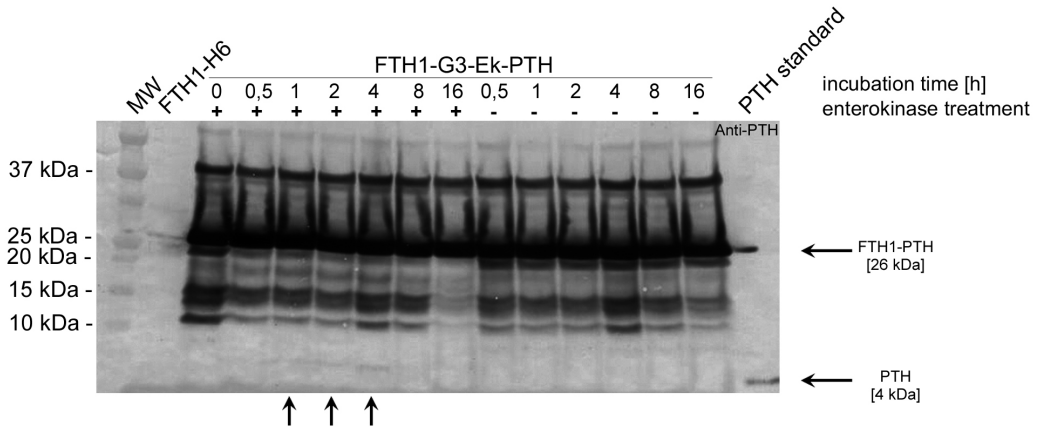


Figure 4.3: **Removal of the PTH moiety from the fusion partner FTH1 by enterokinase cleavage.** After treatment with enterokinase samples were separated on a 4-12% Bis-Tris protein gel. After transfer to nitrocellulose PTH and FTH1-G3-Ek-PTH were detected using a PTH specific antibody. Partial cleavage of PTH of about 2% was achieved after 1 to 4 h of digestion. After 4 h degradation of the removed PTH occurred. The fusion protein remained stable during the experiment. MW: molecular weight marker.

IFNα-2a as a model protein for evaluation of FTH1 as a carrier protein for intracellular protein expression

The fusion of PTH to FTH1 provided information on stabilization of PTH within the fusion protein. In addition high expression was achieved. Prior to PTH we had selected a different model protein for evaluation of FTH1 as a carrier protein for intracellular protein expression. For this purpose IFNα-2a was chosen since it had already been successfully expressed and secreted by *H. polymorpha*. When secreted to the medium IFNα-2a was found to be stable. Further information on IFNα-2a is provided in Manuscript I. The generated fusion protein FTH1-G3-Ek-IFNα-2a also contains a three-glycine linker and an enterokinase cleavage site. After strain generation expression of the fusion protein was analyzed. We found high amounts of the fusion protein within the crude cell extract (see Fig. 4.4). The distribution between soluble and insoluble fraction was comparable to the intracellular distribution of FTH1-H6 which was described earlier in Manuscript II. Hence, changes in solubility do not occur for every fused protein.

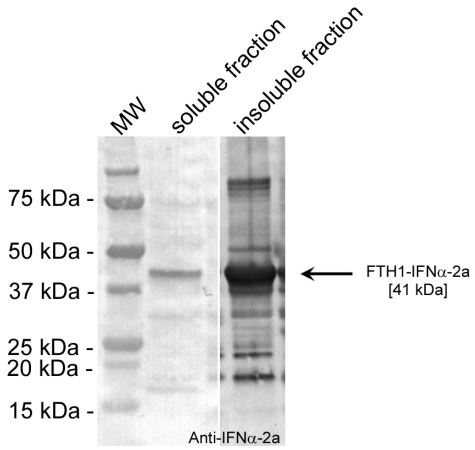


Figure 4.4: **Analysis of the solubility of the fusion protein FTH1-G3-Ek-IFNα-2a by Western blot.** Western blot of the intracellular soluble and insoluble fraction of *H. polymorpha* strains expressing FTH1-G3-Ek-IFNα-2a. The material of the dedicated strain was derived from cultivation similar as described in Fig. 2 of Manuscript II. The samples were analyzed on a 4-12% Bis-Tris protein gel. MW: molecular weight.

4.2.5 Discussion

The high intracellular expression of FTH1 is maintained for the fusion protein consisting of FTH1 and PTH.

The expression of small polypeptides is often challenging. In expression systems such as *E. coli* it has been reported that especially small, heterologous peptides are often victim of proteolytic degradation (for review see Marston, 1986). The 34 amino acid comprising PTH fragment is an example for such a polypeptide.

Bacillus subtilis, *S. cerevisiae* and *P. pastoris* were shown to be suitable for expression and subsequent secretion of PTH with *S. cerevisiae* being the most efficient expression system (Tab. 4.1). Nevertheless, the deletion of five yapsin genes was required to achieve these high yields in *S. cerevisiae* (Cho et al., 2010). The absence of various yapsins reduces degradation of the secreted PTH. In case of *P. pastoris* supplementation of the medium with the protease inhibitor EDTA and PDI coexpression, which is often used to improve the secretion of disulphide bridges containing proteins, had to be performed (Vad et al., 2005). The latter two examples of secretory PTH expression show that intensive optimization is required to obtain high PTH yields. PTH is only poorly expressed in *H. polymorpha* using the MFα prepro sequence and it is assumed that the expressed protein is rapidly degraded (M. Piontek, personal communication). It has been shown in various examples that the expression of this protein fused to a carrier protein results in enhanced the stability. Various fusion proteins have been designed to prevent proteolytic degradation in *E. coli* and *P. pastoris* (see Tab. 4.1). In many cases a spacer was introduced between carrier and target protein. Different proteases were employed for the removal of the target protein. Highest recovery was obtained in case of a GST fusion protein expressed in *E. coli* after initial cleavage with thrombin and subsequent cleavage with prolyl endopeptidase (Xiu et al., 2002). The authors reported a PTH yield of 600 mg/L.

In this study we performed the expression of PTH fused to the highly expressed FTH1. The fusion protein was expressed intracellularly to prevent proteolytic degradation in the medium. During fermentation of FTH1-G3-Ek-PTH yields up to 1.3 g/L within the whole cell extract were obtained. The fusion of PTH to the carrier protein FTH1 did not influence the FTH1 expression rate. The expression of the FTH1-G3-Ek-IFNα-2a also resulted in comparable amounts. Yet, the expression of this fusion protein did not exceed the amount generated by secretory expression of IFNα-2a and therefore this approach was not further pursued (M. Suckow, personal communication).

Removal of the PTH moiety from the fusion protein with enterokinase was of limited success.

On a molecular weight basis PTH accounts for 16% of the total amount of the FTH1-G3-Ek-PTH fusion protein. Due to the high expression rate of FTH1-G3-Ek-PTH (1.3 g/L), up to 200 mg/L PTH would be retrieved from this material after successful recovery. However, the removal of PTH was too inefficient. When the fusion protein was treated with enterokinase only a partial cleavage was achieved. In addition the released PTH showed only limited stability and was degraded completely after 6 h. Mendirette et al. (2012) also designed a fusion protein containing an enterokinase recognition site for PTH recovery. Unfortunately no attempt of recovery and thus efficiency was described.

The release of the target protein from the carrier protein needs to be improved. One approach for cleavage improvement is to increase the enterokinase cleavage efficiency. The amino acids adjacent to the cleavage site have been shown not to interfere with enterokinase cleavage in general (Hosfield and Lu, 1999). Furthermore, enterokinase is insensitive to a variety of detergents, which may be evaluated for PTH removal (Waugh, 2011). Cui et al. (2007) could also show that the enterokinase cleavage

Table 4.1: **Overview of PTH(1-34) production in various pro- and eukaryotic hosts.**

Fusion protein variants						
Host organism	**Carrier Protein**	**Spacer**	**Cleavage**	**Yield fusion protein [mg/L]**	**Yield PTH recovery [mg/L]**	**Reference**
E. coli	β-galactosidase (1-117)	HHHPGGSVKKR	Kex2-660	1040	500	(Suzuki et al., 1998)
E. coli	glutathione S-transferase (GST)	GSP	thrombin, prolyl endopeptidase	n.i.	600	(Xiu et al., 2002)
E. coli	glutathione S-transferase (GST)	none	Factor Xa		300	(Gangireddy et al., 2010)
E. coli	small ubiquitin-like modifier protein (SUMO)	GG	SUMO protease	20 - 40	3 - 4	(Bosse-Doenecke et al., 2008)
E. coli	Granulocyte colony-stimulating factor (G-CSF)	DDDDK	enterokinase	4620	n.i.	(Mendirette et al., 2012)
P. pastoris	human serum albumin (HSA)	n.i.	n.i.	n.i.	n.i.	(Chen et al., 2008)
H. polymorpha	FTH1	DDDDK	enterokinase	1300	n.m.	this study (Manuscript II)
Secretory variants						
Host organism	**Secretion Signal**				**PTH yield [mg/L]**	**Reference**
Bacillus subtilis	*Bacillus amyloliquefaciens* neutral protease				n.i.	(Saunders et al., 1991)
P. pastoris	MFα prepro				300	(Vad et al., 2005)
S. cerevisiae	MFα prepro				350	(Cho et al., 2010)

Data on fusion proteins as well as secretory variants are provided. n.i.: not indicated, n.m.: not measured.

reaction can be accelerated by application of hydrogen peroxide. A faster release of the PTH moiety may result in less degradation of the mature protein. An additional approach may be the alteration of the cleavage site. Indeed, enterokinase recognizes additional recognition sequences besides Asp-Asp-Asp-Asp-Lys (DDDDK) such as Asp-Asp-Asp-Asp-Arg (DDDDR) (Gasparian et al., 2011) and Leu-Lys-Gly-Asp-Arg (LKGDR) (Liew et al., 2005). In addition, the introduction of a recognition site of a different protease needs to be considered. Preferably a protease should be chosen which generates a native N-terminus (for example Factor Xa, a combination of thrombin and prolyl endopeptidase, or Kex2p). All of those proteases have already been applied for PTH removal from a fusion partner and it has been shown in these contexts that all of these proteases did not perform unspecific cleavage on PTH. Kex2p appears to be the most interesting candidate as it has already been expressed in our laboratory and used for *in vitro* processing as described in Manuscript I (for details on soluble Kex2p expression see Rolf, 2011). Apparently, the native PTH precursor is processed at a Lys-Arg motif and the mature PTH is released. A spacer derived from this PTH precursor was also found to mediate efficient processing by a Kex2p homolog during PTH removal from a β-galactosidase fusion partner (Suzuki et al., 1998).

FTH1-H6 and FTH1-G3-Ek-PTH show equally high thermostability. The addition of PTH to the C-terminus of FTH1 results in an increased solubility of the fusion protein in comparison to FTH1 without a fusion partner. The ability to bind iron was decreased.

When connected to the fusion partner PTH is fully stable in the whole cell extract. We could furthermore show that the fusion protein is also stable during heat treatment. Linder and Munro (1972) described heat precipitation as a method for FTH1 purification. In this study we were able to adapt this protocol to the purification of a FTH1-derived fusion protein. Heat precipitation was successfully performed with FTH1-G3-Ek-PTH to remove some protein contaminants from the whole cell extract.

Subsequently, FTH1-H6 and FTH1-G3-Ek-PTH samples were analyzed for iron binding ability. We detected assembly of the subunits and formation of a particle structure, but in both cases no binding of iron had occurred within the *H. polymorpha* cells *in vivo*. A similar finding was described by Shi et al. (2008). The authors expressed the H and L subunit of human ferritin in *S. cerevisiae* and could only detect small amounts of bound iron. However, they were able to identify a chaperone, which mediates iron transport to the assembled ferritin. To analyze whether the protein samples generated during this study were able to bind iron *in vitro* FTH1-H6 and FTH1-G3-Ek-PTH were incubated with ammonium Fe(II) sulphate. In this manner, we could show that only FTH1-H6 , but not FTH1-G3-Ek-PTH, was able to bind iron *in vitro*. Hence, the addition of the spacer and the PTH moiety seems to significantly change the ability of the protein to bind iron.

Similar observations were made by van Wuytswinkel et al. (1995) during expression of pea-seed ferritin in *E. coli*. The presence of a 24 amino acid extension peptide at the N-terminus of the mature protein strongly increased the solubility of the assembled protein. Furthermore, additional 47 amino

acids composing the transit peptide prevented assembly of the 24 subunits and no iron binding was detected. We suggest that major conformational changes may also occur during FTH1-G3-Ek-PTH expression and that the assembly of functional iron binding particle is hindered. These changes may also be connected with the different solubility behavior. However, concluding analysis of the particle formation needs to be performed to further clarify these assumptions.

5 Manuscript III: Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*

Author's contribution

Journal: Journal of Biotechnology
Impact Factor: 3.045
1^{st} author
Author's contribution: 90%

- Experimental design
- Cloning (design of single copy basic vectors together with Andreas Kranz)
- Strain generation and analytical PCR
- Small scale cultivation and fermentation
- Protein and metabolite analysis
- Authoring manuscript

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS

Journal of Biotechnology xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*

Eva Eilert[a,b], Andreas Kranz[a], Cornelis P. Hollenberg[a,b], Michael Piontek[a], Manfred Suckow[a,*]

[a] *ARTES Biotechnology GmbH, Elisabeth-Selbert-Str. 9, 40764 Langenfeld (Rheinland), Germany*
[b] *Institut für Mikrobiologie, Heinrich-Heine-Universität, Universitätsstr. 1, 40225 Düsseldorf, Germany*

ARTICLE INFO

Article history:
Received 17 December 2012
Received in revised form 7 February 2013
Accepted 11 February 2013
Available online xxx

Keywords:
Compatible solute
Ectoine
5-Hydroxyectoine
Hansenula polymorpha
Whole-cell biocatalysis
Halomonas elongata

ABSTRACT

Ectoine and 5-hydroxyectoine belong to the family of compatible solutes which are known to mainly contribute to the adaptation of the cell to osmotic stress by mediation of a constant turgor. In addition the cell's essential functions are maintained under stress conditions like high salinity, heat or aridity stress. *Hansenula polymorpha* was engineered to catalyze the transformation of monomeric substrates to 5-hydroxyectoine. For this purpose four genes encoding the enzymes of the 5-hydroxyectoine biosynthesis pathway of *Halomonas elongata*, EctA, EctB, EctC, and EctD, were inserted into the genome of *H. polymorpha*. Subsequently the syntheses of ectoine and 5-hydroxyectoine were analyzed and optimized. We showed that *H. polymorpha* is a suitable system for recombinant 5-hydroxyectoine synthesis in gram per liter scale (2.8 g L^{-1} culture supernatant, 365 μmol/g dcw) in which almost 100% conversion of ectoine to 5-hydroxyectoine without necessity of high salinity were achieved.

1. Introduction

Ectoine and 5-hydroxyectoine belong to the family of compatible solutes which are known to balance osmotic pressure to reestablish volume and turgor of the cell (Wohlfarth et al., 1990). Compatible solutes are synthesized under elevated saline conditions and then accumulate in the cell. In addition the cell's essential functions are maintained under stress conditions like high salinity, heat or aridity stress by supporting the stability of proteins (Lippert and Galinski, 1992). Ectoine and 5-hydroxyectoine as well as the precursor Nγ-acetyl-diaminobutyrate are also known to have protective properties toward nucleic acids, membranes and whole cells (Cánovas et al., 1999; Lentzen and Schwarz, 2006a). Ectoine was first discovered in *Ectothiorhodospira halochloris* (Galinski et al., 1985), later more halophilic or halotolerant organisms with the ability to produce ectoine, or in addition 5-hydroxyectoine, were identified (Inbar and Lapidot, 1988). It has been shown that other organisms without the ability to synthesize ectoine or 5-hydroxyectoine themselves are often able to accumulate those compounds from their environment and profit from their protective properties (for review see Pastor et al., 2010).

Among moderate halophilic bacteria, *Halomonas elongata* is considered one important model organism and has been studied extensively (García-Estepa et al., 2006). The synthesis of ectoine takes place in three steps starting with L-aspartate-semialdehyde (Peters et al., 1990) which is converted by EctB to L-2,4-diaminobutyrate. Next the Nγ-acetyl-2,4-diaminobutyrate is formed by the action of EctA. In the last step ectoine is generated by EctC by condensation. An additional hydroxylation step catalyzed by EctD converts ectoine into 5-hydroxyectoine (Bursy et al., 2007) (Fig. 1). The synthesis of ectoine and 5-hydroxyectoine in *H. elongata* only occurs under a variety of stress conditions, e.g. salt, whereby not the entire ectoine pool is converted to 5-hydroxyectoine (Inbar and Lapidot, 1988; Wohlfarth, 1990). In addition it is known that in *H. elongata* ectoine can also be hydrolyzed to Nγ- and Nα-acetyl-L-2,4-diaminobutyrate by the ectoine hydrolase (DoeA) resulting in a reduced yield (Schwibbert et al., 2010). Both ectoine and 5-hydroxyectoine can be utilized as an energy source (Vargas et al., 2006).

Due to their chiral carbons ectoine and 5-hydroxyectoine can only be efficiently produced via biological approaches (Schiraldi

Abbreviations: DABA, diaminobutyrate; EctA, DABA acyltransferase; EctB, DABA-2-oxoglutarate transaminase; EctC, ectoine synthase; EctD, ectoine hydroxylase; *FMD*, formate dehydrogenase; *MOX*, methanol oxidase; Ask, aspartate kinase; Asd, aspartate semialdehyde dehydrogenase; MeOH, methanol; dcw, dry cell weight.

* Corresponding author. Tel.: +49 0 2173 2758715; fax: +49 0 2173 2758777.
E-mail addresses: eva.eilert@uni-duesseldorf.de (E. Eilert), a.kranz@artes-biotechnology.com (A. Kranz), cornelis.hollenberg@uni-duesseldorf.de (C.P. Hollenberg), m.piontek@artes-biotechnology.com (M. Piontek), m.suckow@artes-biotechnology.com (M. Suckow).

0168-1656/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS

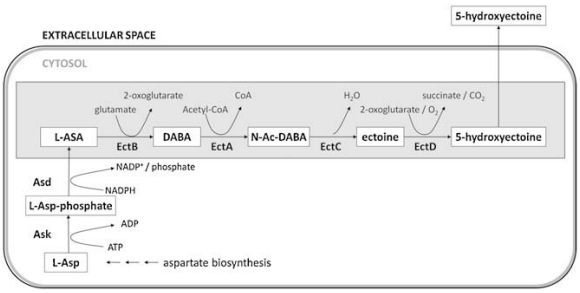


Fig. 1. Schematic overview of the 5-hydroxyectoine synthesis pathway of *H. elongata*. *Abbreviations*: L-Asp, L-aspartate; L-ASA, L-aspartate-b-semialdehyde; DABA, L-aspartate-diaminobutyrate; N-Ac, N-acetyl; Ask, aspartate kinase; Asd, aspartate-semialdehyde dehydrogenase; EctB, diaminobutyrate-2-oxoglutarate transaminase; EctA, L-2,4-diaminobutyrate acetyltransferase; EctC, ectoine synthase; EctD, ectoine hydroxylase. As a precursor for 5-hydroxyectoine synthesis L-Asp is provided by respective anabolic processes. Steps introduced in *H. polymorpha* are boxed in gray. For further explanations see text.

et al., 2006). Therefore we decided to generate a recombinant producer strain which synthesizes 5-hydroxyectoine independently of stress parameters such as high salt concentrations and therefore can grow without performing stress response reactions which may interfere with the production process.

Heterologous systems are of interest since these organisms not only accumulate the compatible solutes within the cells but can also export them. The relevant enzymes involved in ectoine biosynthesis have already been expressed in various microbial systems. *Escherichia coli* was utilized for the expression of the *ectABC* clusters derived from different species such as *Bacillus halodurans* (Anbu Rajan et al., 2008) or *Chromohalobacter salexigens* (Schubert et al., 2007), leading to yields of up to 6 g of ectoine/L. Another effective ectoine process reported using *E. coli* is 'bacterial milking' introduced by Sauer and Galinski (1998). Lentzen and Schwarz (2006b) developed the so far most effective 'permanent-milking' process using the genuine producer *H. elongata* with yields up to 20 g ectoine/L, a process which finds application in industrial production.

The 5-hydroxyectoine biosynthesis was optimized recently using the homologous producer *Pseudomonas stutzeri* and recombinant *E. coli* expressing the *P. stutzeri ectABCD-ask* cluster (Seip et al., 2011). These approaches resulted in almost complete transformation of ectoine to 5-hydroxyectoine. In addition, *Marinococcus* M52 was developed for 5-hydroxyectoine production, however, this process did not yet find industrial application due to several limitations (Frings et al., 1995; Schiraldi et al., 2006). 5-Hydroxyectoine production using *H. elongata* is also possible by modification of the 'bacterial milking' process (Lentzen and Schwarz, 2006a). However, Seip et al. (2011) stated that even after optimization of the production process by increased salinity and elevated temperature 5-hydroxyectoine synthesis did not exceed 50%. Therefore, due to the inavailability of an efficient heterologous production system both compounds are currently biotechnologically produced solely using the homologous system *H. elongata* (Reshetnikov et al., 2011).

The methylotrophic yeast *Hansenula polymorpha* has been established as a highly efficient expression system (review Gellissen and Hollenberg (1997)). Due to the strong promoters derived from the methanol metabolism pathway and the possibility of mitotically stable high-copy integration, *H. polymorpha* has been developed as an effective system for heterologous protein expression (Janowicz et al., 1991; Mayer et al., 1999; Weydemann et al., 1995; Eilert et al., 2012) as well as a whole-cell biocatalyst (Gellissen et al., 1996).

In this study, we report the development of a stable whole-cell biocatalyst on the basis of *H. polymorpha* which catalyze the transformation of monomeric substrates to 5-hydroxyectoine. We inserted the four genes of the 5-hydroxyectoine biosynthesis pathway of *H. elongata* into the genome of *H. polymorpha*. Subsequently the synthesis and export of ectoine and 5-hydroxyectoine were analyzed and optimized.

2. Materials and methods

2.1. Strains and culture conditions

For construction and propagation of plasmid DNA, *E. coli* NEB10β (araD139 Δ(ara-leu)7697 fhuA lacX74 galK (ϕ80 Δ(lacZ)M15) mcrA galU recA1 endA1 nupG rpsL (Str^R) Δ(mrr-hsdRMS-mcrBC); New England Biolabs, Frankfurt a. M., Germany) was used. Cultivation took place in Luria Bertani media, when needed supplemented with 50 μg ampicillin per mL. For protein expression *H. polymorpha* strain ALU3 (*ade1 leu2 odc1*) was used as host (Lahtchev et al., 2002). Cultivation took place in either yeast extract-peptone-dextrose medium complemented with 100 μg adenine per mL (YPD-Ade) or yeast nitrogen base medium complemented with 80 μg uracil per mL (YNB-Ura). All cultivations were performed at 37 °C if not indicated otherwise.

2.2. Plasmid constructions

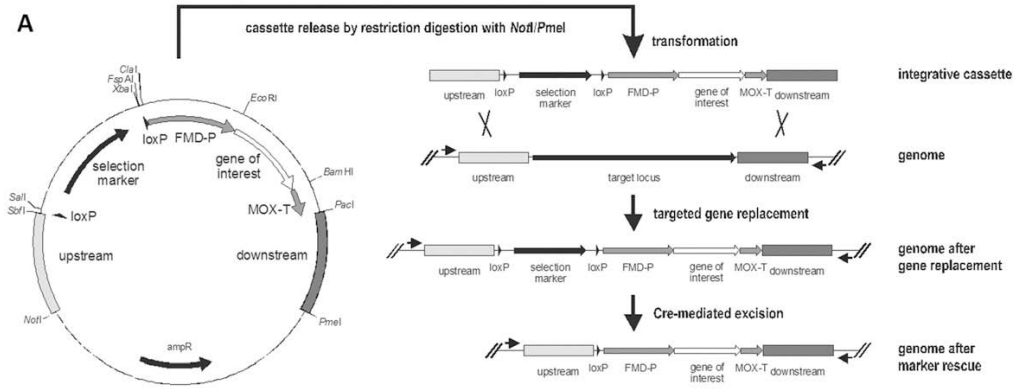
To generate the expression plasmids, synthetic open reading frames for the enzymes EctB, EctA, EctC and EctD from *H. elongata* ATCC 33173 based on Universal Protein Resource Knowledgebase were inserted into four different vectors containing the up- and downstream fragment for the target locus as well as a unique selection marker, respectively (see Section 2.3). The ORFs were codon-optimized for expression in *H. polymorpha* by GeneArt (Regensburg, Germany). The sequences of the codon-optimized genes have been submitted to NCBI GenBank and can be retrieved under the following accession numbers: *ectA-myc* KC470074, *ectB-T7* KC470075, *ectC-HA* KC470076, *H6-ectD* KC470077. Each ORF has been fused with the codons of a C- or N-terminal tag allowing individual detection. The various combinations of genes, tags, loci, and selection markers in those plasmids are shown in Fig. 2B. The cloned sequences with additional flanking tags and restriction sites (5′*Eco*RI and 3′*Bam*HI) were synthesized by GeneArt. *H. polymorpha* was transformed by electroporation as described in Faber et al. (1994).

2.3. Single copy integration system

In *H. polymorpha*, introduced circular plasmids are initially episomal before integrated as high-copy head-to-tail repeats into the genome (Gellissen et al., 1992). Plasmid integration and amplification are stochastic processes, and the number of copies finally integrated is not predictable. For the defined expression of several genes of interest we developed a modular system for integration of single expression cassettes into the genome by targeted gene replacement. The significant qualities of that system are as follows (Fig. 2A): (1) *FMD* promoter and *MOX* terminator as transcriptional elements with a multiple cloning site in between to confer high-level expression of any gene of interest; (2) interchangeable yeast selection markers flanked by loxP sites suitable for Cre-mediated marker rescue; (3) both selection marker and expression cassette are flanked by up- and downstream target sequences of approximately 1 kb size to confer double-homologous recombination at the target locus with high frequency; (4) all modules are flanked with unique restriction sites for convenient cloning; and (5) the integrative cassette can be isolated via restriction with *Not*I/*Pme*I.

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS
E. Eilert et al. / Journal of Biotechnology xxx (2013) xxx–xxx



B

Step no.	Name of integration cassette	Protein of interest	UniProt ID	Immuno tag	Target locus	Selection marker	Genotype relevant for selection
1	**dl-lx(As)-F-EctB-T7**	EctB	O52250	T7 (C-terminal)	*leu2*	*ADE1*	*ADE1 leu2 odc1 HIS4*
2	Removal of selection marker *ADE1* by Cre-mediated excision (marker rescue)						*ade1 leu2 odc1 HIS4*
3	**dh4-lx(As)-F-EctA-myc**	EctA	O52249	c-myc (C-terminal)	*HIS4*	*ADE1*	*ADE1 leu2 odc1 his4*
4	**du-lx(H4)-F-EctC-HA**	EctC	O52251	HA (C-terminal)	*odc1*	*HIS4*	*ADE1 leu2 odc1 HIS4*
5	**da-lx(Ls)-F-H6-EctD**	EctD	E1VA04	H6 (N-terminal)	*ade1*	*LEU2*	*ADE1 LEU2 odc1 HIS4*

Fig. 2. Overview of targeted single copy integration of *ect* genes into the genome of *H. polymorpha*. (A) Schematic representation of the plasmids harboring elements for integration of single copy expression cassettes and their integration by targeted gene replacement as well as the genomic structure upon marker rescue. Primer hybridization sites for analytical PCR are symbolized by solid black arrows adjacent to the + and − strands of the DNA fragments. The PCR fragment lengths provided by the wild type genome, the correct single copy integration, or the subsequent marker rescue were clearly distinguishable (as schematically indicated by the different distances between the arrows symbolizing the forward and reverse primers of the different situations). For details regarding the modular plasmid design see text. (B) Succession of integration events leading to strain ALU3/EctBACD #1 with indication of source of the *ect* gene, immuno tag used for detection, target locus for single copy integration, selection marker, and genotype relevant for selection. Genes were integrated into the genome of *H. polymorpha* ALU3 by the indicated integration order. For removal of the selection marker within step no. 2 the Cre-recombinase was transiently expressed.

After transformation with the cassette and selection of transformants, colonies displaying the expected prototrophy were screened by a locus-specific analytical PCR for the expected recombination event. The primer hybridization sites were chosen in way that a PCR product length dimorphism shows the presence of the target locus with or without the integrated cassette, respectively (see Fig. 2A; primer hybridization sites are indicated as arrows).

We used the Cre-loxP system for site-specific removal of selection marker genes (Sauer, 1987). For transient expression of the Cre-recombinase the plasmid pFPMT-2m-Cre was constructed. This derivate of pFPMT121 (Degelmann et al., 2002) contains the Cre gene from pSH47 (Güldener et al., 1996) under control of the *FMD* promoter. Transformants were selected on agar plates containing 2% glycerol for immediate expression of the recombinase; subsequently, isolated colonies were picked and transferred onto plates with nonselective medium to allow plasmid loss. Cells having lost both selection marker and Cre expression plasmid were identified by replica plating onto selective agar plates.

2.4. Protein expression and preparation of protein material

For protein expression the recombinant strains were subjected to derepression. Cells were cultivated in 3 mL scale in YNB containing 1–2% (w/v) of a derepressing carbon source (glycerol or sorbitol) for 24–48 h at 37 °C, then methanol (MeOH) was added to 1% (v/v) and cultivation was continued for additional 24–48 h with repetition of the MeOH induction every 24 h if applicable. After cultivation (72 h) the cells were harvested by centrifugation. Cells and culture supernatant were separately collected and stored at −20 °C. The total cell extract was prepared by glass bead cell disruption. Subsequently the samples were centrifuged for 15′ at 16,000 × *g* to separate the insoluble and soluble fractions.

2.5. Protein analysis

For molecular weight determination the samples were analyzed by SDS-PAGE according to Laemmli (1970). For the electrophoretic separation 4–12% Bis–Tris acryl amide XT gels (BioRad, Hercules, CA, USA) were used. The protein band patterns were transferred to nitrocellulose membranes. The protein bands were then visualized by immunoblotting. For immunological detection a His tag-specific antibody (MicroMol, Karlsruhe, Germany), a HA-specific antibody (Sigma–Aldrich, St. Louis, MO, USA), a c-myc-specific antibody (BioLegend, San Diego, CA, USA), or a T7-specific antibody (MBL, Nagoya, Japan) was applied depending on the protein to be detected (see Table 1). As secondary antibodies a goat-anti-mouse-AP-conjugate or a goat-anti-rabbit-AP-conjugate (both BioRad) were used. Development was performed by use of NBT/BCIP (Roche, Mannheim, Germany).

2.6. HPLC detection of ectoine and 5-hydroxyectoine

For HPLC analysis samples were diluted with the HPLC eluent (acetonitrile/phosphate buffer 71/29, pH 7.2) to an approximate concentration of 10–50 ng/mL and centrifuged (13,000 rpm, 5 min) with subsequent filtration of the supernatant with a cutoff at 0.22 μm. HPLC analyses were carried out using a Nucleodur 100-5 NH2-RP CC 125/4 column (Macherey & Nagel, Düren, Germany)

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS

Table 1
5-Hydroxyectoine/ectoine yields provided by published non-continuous production processes in comparison with those obtained by the *H. polymorpha*-based process shown in this study. For a better comparison the yields are indicated as μmol/g dcw. Please note that the *H. polymorpha*-based process does not require any salt stress. n.i., not indicated.

Organism (strain/vector)	5-Hydroxyectoine (μmol/g dcw)	Ectoine (μmol/g dcw)	Ratio	Salinity (% NaCl)	Reference
Marinococcus sp. strain M52	847	0	100:0	10	Frings et al. (1995)
Marinococcus sp. strain M52	225	n.i.	–	10	Schiraldi et al. (2006)
E. coli DH5α/pSB01	500	∼64	∼90:10	2	Seip et al. (2011)
P. stutzeri DSM5190^T	480	∼15	∼97:3	5	Seip et al. (2011)
H. polymorpha ALU3/EctBACD	365	6	98:2	0	This study

at a flow rate of 1 mL/min at 25 °C. Ectoine and 5-hydroxyectoine were measured by absorbance at 210 nm using a UV/VIS detector. Standard components (ectoine, 5-hydroxyectoine and Nγ-acetyl-diaminobutyrate) were kindly provided by the bitop AG.

2.7. Fermentation

The fermentation was carried out in a 1.25 L Bioflo III fermenter (New Brunswick Scientific, Edison, NJ, USA). 100 mL YPD preculture was grown for 20 h. Fermentation was started by adjusting an initial optical density of $OD_{600\,nm}$ = 1 in a start culture volume of 1 L using SYN6 medium, as previously described in Eilert et al. (2012), supplemented with 2% (w/v) sorbitol and 500 mg/L uracil.

The aeration rate was 1 vvm using a ring sparger. The cultivation temperature was adjusted to 37 °C. Agitation was performed by two stirrer blades (speed between 500 and 1000 rpm). The sorbitol/MeOH feed was started after 25 h with a feed rate of $2\,g\,L^{-1}\,h^{-1}$; adjustments of the sorbitol/MeOH feed are described in Section 3. During the derepression phase the pO_2 was kept at 40% by a stirrer cascade between 500 and 1000 rpm. pH 5 was adjusted and maintained using a 12.5%, v/v ammonia solution. Sterile diluted antifoam agent (10% (v/v) structol J 673) was added manually as appropriate. Feed was performed using a 37.5% (w/v) sorbitol/37.5% (v/v) MeOH solution or a 15% (w/v) sorbitol/60% (v/v) MeOH solution. Fermentation was terminated after 189 h.

During fermentation samples were taken and analyzed offline. The optical density (OD) was determined at 600 nm wavelength, dry cell weight (dcw) was analyzed, and products were analyzed by HPLC.

3. Results

3.1. H. polymorpha is a suitable candidate for 5-hydroxyectoine synthesis

It was shown that ectoine cannot only serve as an osmolyte but that some organisms can utilize it as a nutrient (Ono et al., 1999; Vargas et al., 2008). Therefore, before introduction of the genes of the 5-hydroxyectoine biosynthesis pathway into *H. polymorpha*, a possible ability of *H. polymorpha* ALU3 to utilize ectoine had to be excluded because degradation of the intermediate ectoine would counterproductively interfere with the desired 5-hydroxyectoine synthesis. For this purpose *H. polymorpha* ALU3 was cultivated in YNB medium without carbon- or nitrogen-source but supplemented with 1% (w/v) ectoine. Determination of the $OD_{595\,nm}$ did not show any growth of the yeast strain (data not shown) supporting that *H. polymorpha* ALU3 did either not utilize ectoine as a carbon- or a nitrogen-source or is not able to import this molecule.

The synthesis of 5-hydroxyectoine is based on the synthesis of amino acids of the aspartate family (Pastor et al., 2010). L-Aspartate is initially converted to L-aspartate-phosphate by the aspartate kinase (Ask) and then further converted to L-aspartate-β-semialdehyde by the L-aspartate-β-semialdehyde dehydrogenase (Asd). Database analysis suggested the existence of enzymes with Ask- and Asd-activity in *H. polymorpha* (*H. polymorpha* NCYC 495 leu1.1, Ask – protein ID 17263, Asd – protein ID 76258). In addition *H. polymorpha* does not seem to harbor the gene for ectoine hydrolase (DoeA), and therefore degradation of ectoine is unlikely. For that reason, for 5-hydroxyectoine synthesis the heterologous expression of the enzymes EctA, EctB, EctC and EctD in *H. polymorpha* should be sufficient without further strain modifications.

3.2. Generation of recombinant H. polymorpha strains co-expressing genes of the H. elongata 5-hydroxyectoine pathway

Synthetic ORFs encoding EctA, EctB, EctC and EctD were designed and inserted into *H. polymorpha* single copy expression vectors with different selection markers and integration loci (Fig. 2B). To provide the option of single detection of each protein, codons for a C- or N-terminal tag were added to the ORFs (resulting enzymes: EctA-myc, EctB-T7, EctC-HA, H6-EctD). All ORF transcriptions are controlled by the strong carbon source regulated *FMD* promoter. *H. polymorpha* ALU3 (*ade1 leu2 odc1*) was transformed with *Not*I/*Pme*I fragments harboring the respective integration cassette and up- and downstream fragments for homologous recombination. The plates were incubated for 6 days. This procedure typically results in the isolation of strains harboring one single copy of the expression cassette at a defined integration locus. Proper integration was determined by analytical PCR using flanking PCR primers (see Fig. 2 for more details). A couple of strains resulting from this procedure were subjected to small scale MeOH induction, and the total cell extracts were analyzed for the presence of gene products by SDS-PAGE/Western blot using tag-specific primary antibodies (data not shown). After the first integration step the selection marker was removed using the cre/loxP recombinase system. The gene integration procedure was repeated until all four expression cassettes were correctly integrated into the desired loci of *H. polymorpha* ALU3 co-expressed. The co-expression was analyzed via SDS-PAGE/Western blot resulting in the strain ALU3/EctBACD #1 (*odc1*) (Fig. 3).

3.3. 5-Hydroxyectoine is quantitatively synthesized from ectoine and is efficiently exported

After integration of all genes and expression of EctB, EctA, EctC and EctD we analyzed whether all enzymes were active. We cultivated the strain ALU3/EctBACD #1 under derepressing/MeOH induction conditions (YNB 2% (w/v) glycerol, induction after 48 h with 1% (v/v) MeOH final concentration, total cultivation time 72 h). Analysis of the culture supernatant by HPLC showed that both compounds were synthesized by this strain in a ratio of 5-hydroxyectoine to ectoine of about 60:40 (data not shown). This ratio under non-optimized conditions was already higher than that in the homologous system *H. elongata*, where only a 50:50 ratio could be reached (Seip et al., 2011).

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS
E. Eilert et al. / Journal of Biotechnology xxx (2013) xxx–xxx

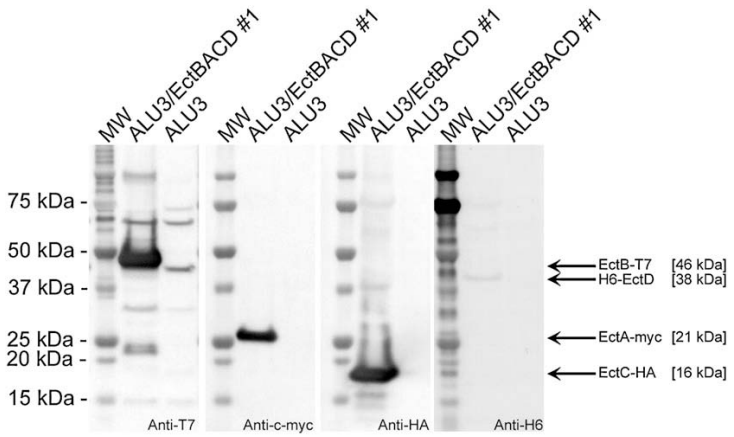


Fig. 3. Western blot analysis of total cell extract of strain ALU3/EctBACD #1 coexpressing EctB, EctA, EctC and EctD. Fermentation was carried out as described in Section 2 at 37 °C in SYN6 at pH 5. After 96 h a sample was taken and total cell extract has been prepared and analyzed by SDS-PAGE/Western blot under reducing conditions. The Ect proteins were then detected individually in the cell extract by four different Western blots using antibodies specific for the individual tags, respectively. The arrows indicated the positions of the bands corresponding to the various *ect* genes, the corresponding calculated molecular weights are indicated in brackets. MW, molecular weight marker; ALU3, wildtype strain.

3.4. Efficient 5-hydroxyectoine synthesis is dependent on media pH

The optimal pH-values of EctA, EctB and EctC are within the range of 8.2–9.0 (Ono et al., 1999); EctD has its pH optimum at 7.5 (Bursy et al., 2007). Hence, to further improve the 5-hydroxyectoine/ectoine ratio a variety of buffered YNB-based media of different pH-values from 3 to 7 were tested. Glycerol was applied as derepressing carbon source. Subsequent HPLC analysis of the culture supernatant showed a broad range of synthesis yields (Fig. 4). For further optimization of the culture conditions YNB pH 5 was chosen since in this medium an increase of 5-hydroxyectoine yield (136 mg/L), a low ectoine yield (10 mg/L) in a ratio of about 90:10 were combined. The media with pH 3, 6, and 7 showed also a high potential toward 5-hydroxyectoine synthesis, however, in addition to the target compounds several not further characterized side products were detected (data not shown).

3.5. 5-Hydroxyectoine formation is dependent on carbon source

Cultivations without MeOH induction did not result in any conversion of ectoine to 5-hydroxyectoine (data not shown). Thus MeOH induction seems to be crucial for ectoine hydroxylation. Unclear is whether this is solely due to the EctD expression or to an additional effect caused by the MeOH metabolism (see Section 4). Besides glycerol, sorbitol was used as derepressing carbon source in combination with MeOH (Fig. 5). In addition glucose was tested since at low concentrations it can also be considered a derepressing carbon source. After depletion of the carbon sources from the batch, the *ect* genes were induced with MeOH. The applied scheme is shown in Fig. 5.

In general we could observe a strong dependency of the synthesis yield on the MeOH induction pattern. MeOH feeding from cultivation start or early induction pulses starting after 24 h led to higher 5-hydroxyectoine synthesis yields (up to 440 mg/L for

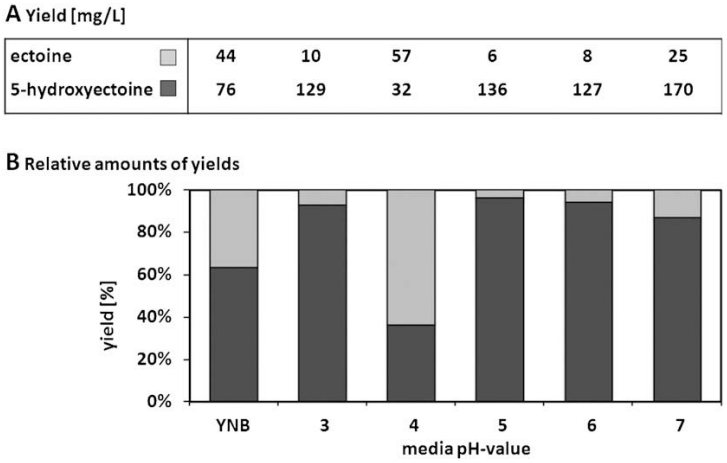


Fig. 4. Determination of ectoine and 5-hydroxyectoine conversion yields of ALU3/EctBACD #1 in dependence of media pH-value during cultivation in 3 ml scale. 5-Hydroxyectoine producer ALU3/EctBACD #1 was cultivated at 37 °C for a total of 72 h in six different YNB-based medium which were buffered at different pH-values ranging from 3 to 7 as indicated. Unbuffered YNB-medium with an initial pH-value of 6 was used as control medium. In all cases glycerol (2%, w/v) was used as derepressing carbon source ("YNB"). After consumption of glycerol (after 48 h) a MeOH induction was performed by adding MeOH to a final concentration of 1% (v/v), respectively and cultivation was continued for additional 24 h. (A) Absolute values of ectoine and 5-hydroxyectoine in the culture supernatants as measured by HPLC. Light gray: ectoine, dark gray: 5-hydroxyectoine. (B) Comparison of 5-hydroxyectoine/ectoine ratios. The sum of 5-hydroxyectoine and ectoine has been defined as 100%, respectively, without considering the absolute values (which are shown in (A)). The values correspond to a single measurement each.

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS

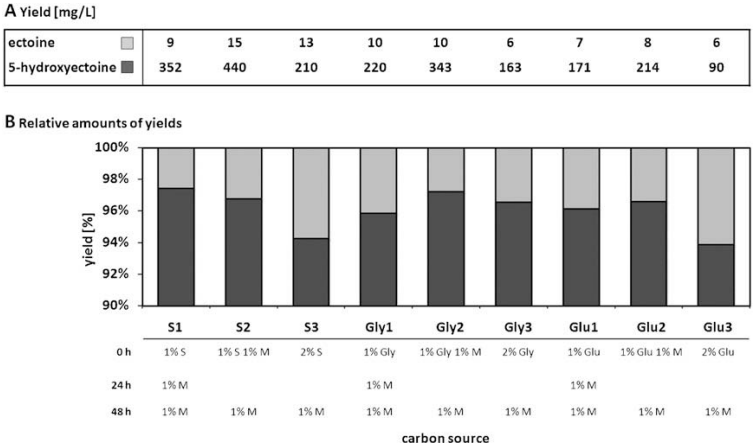


Fig. 5. Determination of ectoine and 5-hydroxyectoine conversion yields during cultivation of strain ALU3/EctBACD #1 in 3 mL scale in dependence of carbon source feeding mode. 5-Hydroxyectoine producer ALU3/EctBACD #1 was cultivated at 37 °C for a total of 72 h. All carbon source combinations were tested on the basis of buffered YNB medium, pH 5 (see also Fig. 3). The applied carbon source feeding profiles are indicated at the bottom of B. (A) Absolute values of ectoine and 5-hydroxyectoine in the culture supernatants as measured by HPLC. Light gray: ectoine, dark gray: 5-hydroxyectoine. (B) Comparison of 5-hydroxyectoine/ectoine ratios (for scheme and symbols see Fig. 3B). Below the bars the course of the respective carbon source feeding mode and the respective carbon source(s) added at 0 h, 24 h, and 48 h (v/v, respectively) is (are) indicated. Feeding with 2% glycerol (G3) already used in the cultivations shown in Fig. 3 was the control condition. The values correspond to a single measurement each. *Abbreviations*: S, sorbitol; Gly, glycerol; Glu, glucose; M, MeOH.

condition S2). Induction starting after 48 h resulted in significantly lower synthesis yields (compare conditions S3, Gly3 and Glu3). In addition, the 5-hydroxyectoine/ectoine ratio could be further improved to >95% of 5-hydroxyectoine. In this regard condition S2 resulted in the highest 5-hydroxyectoine synthesis yield. Consequently, a fermentation process based on this condition was designed for upscaling of the 5-hydroxyectoine production in *H. polymorpha*.

3.6. 5-Hydroxyectoine is exported in gram per liter scale

Based on condition S2, we used the strain ALU3/EctBACD #1 for fermentation in 1 L scale. In contrast to the previously optimized cultivation conditions we used SYN6 medium, a synthetic minimal medium for yeast that is similar to YNB but leads to higher cell densities. During cultivation the pH was maintained at 5 as in the best test tube conditions. For initial growth the medium was supplemented with 2% sorbitol. After 25 h a mixed carbon source feed containing 1 g L^{-1} h^{-1} MeOH and 1 g L^{-1} h^{-1} sorbitol was started. After 45 h the feed was switched to 1.6 g L^{-1} h^{-1} MeOH and 0.4 g L^{-1} h^{-1} sorbitol for further activation of 5-hydroxyectoine synthesis. The carbon source feed was increased during fermentation (Fig. 6A and B). Fermentation was stopped after 190 h. The dry cell weight (dcw) and yields for ectoine and 5-hydroxyectoine are shown in Fig. 6C. Efficient conversion of ectoine to 5-hydroxyectoine arises with the mixed carbon source feed. The 5-hydroxyectoine formation thus clearly depends on MeOH induction. After 144 h of fermentation the cells reached the stationary phase, and the synthesis yield did not increase further. Analysis of the intracellular soluble fraction of the yeast cells showed that only amounts of ectoine and 5-hydroxyectoine were retained in the cell and thus no accumulation occurred (data not shown).

Strain ALU3/EctBACD #1 produced 2.8 g L^{-1} 5-hydroxyectoine (364.5 μmol/g dcw at 144 h) in the culture supernatant with only 44 mg/L ectoine (6.3 μmol/g dcw at 144 h) as a side product. Thus the very high 5-hydroxyectoine yield of >95% could be transferred from the test tube scale to the 1 L scale.

3.7. The homologous H. polymorpha Ask is feedback regulated

To further investigate options for future optimization of the 5-hydroxyectoine synthesis we analyzed the availability of the precursor L-aspartate-semialdehyde via the aspartate pathway which depends on Asd and Ask activity in the cell (Viola, 2001). Ask has been found to be feedback sensitive to methionine, lysine and threonine which are products of the aspartate pathway (Kikuchi et al., 1999). As a result the Ask activity is downregulated and L-aspartate-semialdehyde cannot be provided in sufficient amounts. We therefore evaluated whether the homologous *H. polymorpha* Ask activity may be downregulated by intracellular methionine, lysine, and/or threonine. Sensitivity was determined by cultivation of the strain ALU3/EctBACD #1 in presence of 1 mM methionine, lysine, or threonine. In addition we tested whether supplementation with 1 mM aspartate or glutamate may increase the synthesis yield (Fig. 7). While aspartate and glutamate did not have any effect the addition of methionine, lysine, and threonine to the culture medium resulted in a decrease of the 5-hydroxyectoine synthesis yield. Highest inhibition was caused by methionine, where only 50% of the initial yield were obtained. Therefore, we propose that *H. polymorpha* ALU3/EctBACD #1 may have a feedback sensitive Ask influencing 5-hydroxyectoine synthesis.

4. Discussion

Whole-cell biocatalysis provides an economic technique to produce compounds of chemical or pharmaceutical interest. Ectoine and 5-hydroxyectoine are biologicals of high value and mainly find application in skin care and medical products. In this study we show that *H. polymorpha* promises to be an excellent system for the production of 5-hydroxyectoine.

For further optimization of the 5-hydroxyectoine synthesis in *H. polymorpha* several points need consideration. The precursor compound L-aspartate-semialdehyde is synthesized from aspartate by the two enzymes aspartatekinase (Ask) and L-aspartate-semialdehyde dehydrogenase (Asd). However, it is known that variants of Ask can be feedback sensitive to compounds of the aspartate pathway (methionine, lysine and threonine) (Kikuchi

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS
E. Eilert et al. / Journal of Biotechnology xxx (2013) xxx–xxx

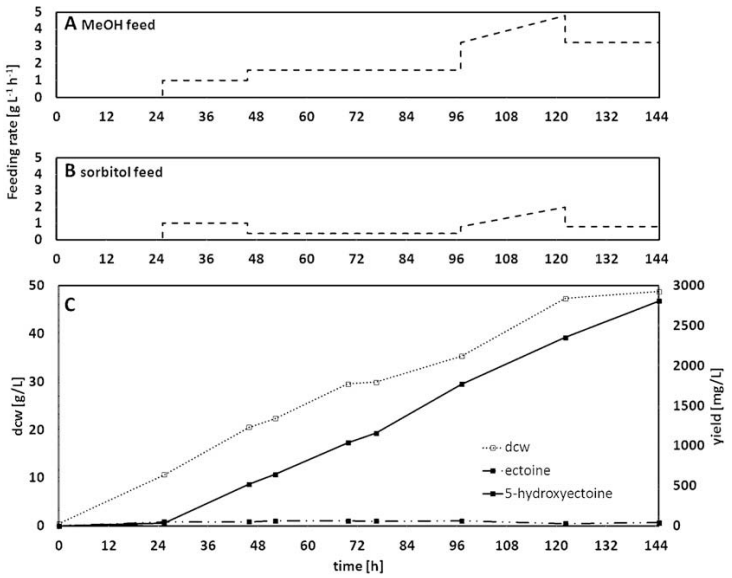


Fig. 6. Fermentation of strain ALU3/EctBACD #1 at 1 L scale. The fermentation was carried out as described in Section 2 at 37 °C in SYN6 at pH 5. (A and B) Feeding modi. 30 g L^{-1} sorbitol were provided as batch for initial cell growth. After consumption of the carbon source (assumed after 24 h) a [37.5% (w/v) sorbitol/37.5% (v/v) MeOH] solution has been applied as a mixed feed with 2 g L^{-1} h^{-1} was started. After 45 h the feeding solution was exchanged with a [15% (w/v) sorbitol/60% (v/v) MeOH] solution without altering of the feeding rate. Between 96 h and 122 h the feeding rate with the latter feeding solution has been continuously increased from 4 to 6 g L^{-1} h^{-1}, as shown in (A) and (B). After 122 h the feeding rate has been reduced to 4 g L^{-1} h^{-1}. The different feeding solutions for the feed are indicated. (C) Corresponding courses of dry cell weight (dcw), and ectoine or 5-hydroxyectoine (mg/L) in the culture supernatant are shown as measured by HPLC. The values correspond to a single measurement each.

et al., 1999). In *E. coli* deletion of Ask and complementation with a feedback resistant variant derived from *Corynebacterium glutamicum* resulted in an increased ectoine production when co-expressed with EctA, EctB and EctC (Bestvater et al., 2008). Expression of the *ectABCD_Ask* cluster derived from *P. stutzeri* had a similar effect (Seip et al., 2011). This Ask_Ect derived from *P. stutzeri* is located in the same cluster as the genes encoding for EctA, EctB, EctC and EctD, and it was shown to be only sensitive to lysine (Stöveken et al., 2011). The existence of a *H. polymorpha* Ask is supported by our finding that the 5-hydroxyectoine synthesis was downregulated when methionine, lysine or threonine were added during cultivation. Deletion of the putative feedback regulated Ask and expression of a feedback independent variant is an option for process improvement. Another interesting candidate for overexpression in *H. polymorpha* might be Ask_Lys derived from *H. elongata* of which the regulation is yet unknown but resulting 5-hydroxyectoine synthesis yields are high (Schwibbert et al., 2010).

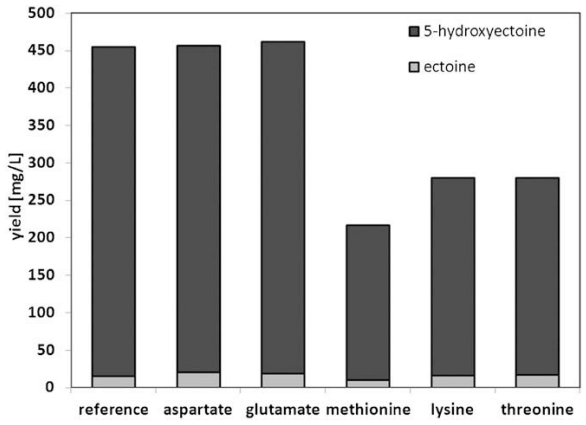


Fig. 7. Analysis of possible feedback sensitivities of the predicted homologous *H. polymorpha* Ask. Strain ALU3/EctBACD #1 was cultured in the 3 mL scale in YNB 1% (w/v) sorbitol buffered to a pH of 5. After 48 h a MeOH induction has been performed (1% (v/v) MeOH final concentration, respectively). The cultures were supplemented with aspartate, glutamate, methionine, lysine or threonine. The control culture without supplementation is termed "reference". 5-Hydroxyectoine and ectoine synthesis yields in the various culture supernatants were analyzed via HPLC. The bars indicate the absolute values.

Surprisingly, we found that in *H. polymorpha* strain ALU3/EctBACD #1 not the ectoine, but solely the 5-hydroxyectoine formation was MeOH dependent while the ectoine synthesis performed well also under derepressing culture conditions. There are different possible explanations for this observation. (1) The expression of EctD under derepressing conditions is too low to provide detectable conversion of ectoine into 5-hydroxyectoine. Only methanol induction leads to an adequate expression level and activity of ectoine hydroxylase. (2) The hydroxylation reaction is influenced by the metabolism of methanol. Indications for the second possibility come from the data from Bursy et al. (2007) showing that addition of catalase enhances EctD activity *in vitro*. Catalase expression in *H. polymorpha* is highly increased during MeOH assimilation since hydrogen peroxide is formed during oxidation of MeOH to formaldehyde by the MeOH oxidase (Veenhuis et al., 1983). Catalase is thus needed for detoxification and conversion of hydrogen peroxide to water and oxygen.

Interestingly we could show that during fermentation the majority of ectoine as well as 5-hydroxyectoine were found in the culture supernatant with only a small amount remaining inside the cell. The lack of an increase in salt resistance in the 5-hydroxyectoine producing *H. polymorpha* strains is in agreement with the finding that both ectoine and 5-hydroxyectoine are not accumulated in the cell (data not shown). There are two possible causes for a release of ectoine and 5-hydroxyectoine to the medium: (1) leakage of both compounds through the plasma membrane; (2) active export via a transport system. In ascomycetous yeasts glycerol, arabitol and mannitol are accumulated as

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS

a reaction to osmotic stress, however glycerol was found to be the most osmotically active solute (van Eck et al., 1993). For transporters involved in adaptation to osmotic stress it has been described, that not only export (mediated by the major intrinsic protein channel Fps1p in *S. cerevisiae* [Tamás et al., 1999]) but also uptake or recapturing is conducted (for example in *Debaryomyces hansenii* [Adler et al., 1985] and *Zygosaccharomyces rouxii* [van Zyl et al., 1990]). For *H. polymorpha* NRRL Y-5441 no active glycerol import under osmotic stress was described (Lages et al., 1999). For genuine ectoine or 5-hydroxyectoine producers several types of ectoine transporters have been described (Grammann et al., 2002; Vermeulen and Kunte, 2004; Jebbar et al., 2005). These transporting mechanism with regard to import, export or recapturing also differ among the different organisms. Yet, in contrast to the release conducted by the genuine producers, the release of ectoine and 5-hydroxyectoine conducted by *H. polymorpha* seems to be salt-independent.

In summary we could demonstrate the successful synthesis of 5-hydroxyectoine in gram per liter scale by using *H. polymorpha* as an eukaryotic host without application of stress factors (e.g. high salinity) as for example necessary for *H. elongata* (Reshetnikov et al., 2011). Export of the product and a low rate of ectoine as a side product require a downstream process which is significantly simpler than for the other systems described so far (Table 1). Even without consideration of further possible options for improval of synthesis yield, *H. polymorpha* is a promising candidate for 5-hydroxyectoine production in industrial scale.

Acknowledgements

The authors thank Dr. Stefanie Huppert for generation of the EctD multi-copy plasmid and the respective control strain, Dr. Erik Böer (IPK Gatersleben, Germany) for ectoine-utilization experiments, and Sylvia Denter, Michael Weniger and David Wilson for expertise and excellent assistance on perfoming fermentation. The authors thank especially Nina Wingbermühle and Renske Klassen for expert technical assistance and Dr. Markus Neumann (bitop AG; Witten, Germany) for additional HPLC measurements on 5-hydroxyectoine detection. The *H. elongata* EctD-gene sequence and HPLC protocol for ectoine detection were kindly provided by the bitop AG.

The research work was supported by Grants from the German National Ministry of Education and Research (BMBF).

References

Adler, L., Blomber, A., Nilsson, A., 1985. Glycerol Metabolism and osmoregulation in salt-tolerant yeast *Debaryomyces hansenii*. Journal of Bacteriology 162 (1), 300–306.

Anbu Rajan, L., Joseph, T., Thampuran, N., James, R., Ashok Kumar, K., Viswanathan, C., Bansal, K., 2008. Cloning and heterologous expression of ectoine biosynthesis genes from *Bacillus halodurans* in *Escherichia coli*. Biotechnology Letters 30 (8), 1403–1407.

Bestvater, T., Louis, P., Galinski, E., 2008. Heterologous ectoine production in *Escherichia coli*: by-passing the metabolic bottle-neck. Saline Systems 4, 12.

Bursy, J., Pierik, A., Pica, N., Bremer, E., 2007. Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. Journal of Biological Chemistry 282 (43), 31147–31155.

Cánovas, D., Borges, N., Vargas, C., Ventosa, A., Nieto, J., Santos, H., 1999. Role of Nγ-acetyldiaminobutyrate as an enzyme stabilizer and an intermediate in the biosynthesis of hydroxyectoine. Applied and Environment Microbiology 65 (9), 3774–3779.

Degelmann, A., Müller, F., Sieber, H., Jenzelewski, V., Suckow, M., Strasser, A., Gellissen, G., 2002. Strain and process development for the production of human cytokines in *Hansenula polymorpha*. FEMS Yeast Research 2 (3), 349–361.

Eilert, E., Hollenberg, C.P., Piontek, M., Suckow, M., 2012. The use of highly expressed FTH1 as carrier protein for cytosolic targeting in *Hansenula polymorpha*. Journal of Biotechnology 159 (3), 172–176.

Faber, K.N., Haima, P., Harder, W., Veenhuis, M.A.B.G., 1994. Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*. Current Genetics 25 (4), 305–310.

Frings, E., Sauer, T., Galinski, E., 1995. Production of hydroxyectoine: high cell-density cultivation and osmotic downshock of *Marinococcus* strain M52. Journal of Biotechnology 43, 53–61.

Galinski, E., Pfeiffer, H.P., Trüper, H., 1985. 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. European Journal of Biochemistry 149 (1), 135–139.

García-Estepa, R., Cánovas, D., Iglesias-Guerra, F., Ventosa, A., Csonka, L., Nieto, J., Vargas, C., 2006. Osmoprotection of *Salmonella enterica* serovar Typhimurium by Nγ-acetyldiaminobutyrate, the precursor of the compatible solute ectoine. Systematic and Applied Microbiology 29 (8), 626–633.

Gellissen, G., Hollenberg, C.P., 1997. Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis*—a review. Gene 190 (1), 87–97.

Gellissen, G., Janowicz, Z.A., Weydemann, U., Melber, K., Strasser, A.W., Hollenberg, C.P., 1992. High-level expression of foreign genes in *Hansenula polymorpha*. Biotechnology Advances 10 (2), 179–189.

Gellissen, G., Piontek, M., Dahlems, U., Jenzelewski, V., Gavagan, J.E., DiCosimo, R., Anton, D.L., Janowicz, Z.A., 1996. Recombinant *Hansenula polymorpha* as a biocatalyst: coexpression of the spinach glycolate oxidase (GO) and the *S. cerevisiae* catalase T (CTT1) gene. Applied Microbiology and Biotechnology 46 (1), 46–54.

Grammann, K., Volke, A., Kunte, H., 2002. New type of osmoregulated solute transporter identified in halophilic members of the bacteria domain: TRAP transporter TeaABC mediates uptake of ectoine and hydroxyectoine in *Halomonas elongata* DSM 2581(T). Journal of Bacteriology 184 (11), 3078–3085.

Güldener, Heck, U., Fielder, S., Beinhauer, T., Hegemann, J.J.H., 1996. A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Research 24 (13), 2519–2524.

Inbar, L., Lapidot, A., 1988. The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ^{13}C- and ^{15}N-labeled L-glutamate and ^{13}C and ^{15}N NMR spectroscopy. Journal of Biological Chemistry 263 (31), 16014–16022.

Janowicz, Z.A., Melber, K., Merckelbach, A., Jacobs, E., Harford, N., Comberbach, M., Hollenberg, C., 1991. Simultaneous expression of the S and L surface antigens of hepatitis B, and formation of mixed particles in the methylotrophic yeast, *Hansenula polymorpha*. Yeast 7, 431–443.

Jebbar, M., Sohn-Bösser, L., Bremer, E., Bernard, T., Blanco, C., 2005. Ectoine-induced proteins in *Sinorhizobium meliloti* include an Ectoine ABC-type transporter involved in osmoprotection and ectoine catabolism. Journal of Bacteriology 187 (4), 1293–1304.

Kikuchi, Y., Kojima, H., Tanaka, T., 1999. Mutational analysis of the feedback sites of lysine-sensitive aspartokinase of *Escherichia coli*. FEMS Microbiology Letters 173 (1), 211–215.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (5259), 680–685.

Lages, F., Silva-Graça, M., Lucas, C., 1999. Active glycerol uptake is a mechanism underlying halotolerance in yeasts: a study of 42 species. Microbiology 145, 2577–2585.

Lahtchev, K., Semenova, V., Tolstorukov, I., van der Klei, I., Veenhuis, M., 2002. Isolation and properties of genetically defined strains of the methylotrophic yeast *Hansenula polymorpha* CBS4732. Archives of Microbiology 177 (2), 150–158.

Lentzen, G., Schwarz, T., 2006a. Extremolytes: natural compounds from extremophiles for versatile applications. Applied Microbiology and Biotechnology 72 (4), 623–634.

Lentzen, G., Schwarz, T., 2006b. Kompatible Solute: Mikrobielle Herstellung und Anwendung. In: Antranikian, G (Ed.), Angewandte Mikrobiologie. Springer-Verlag, Berlin, Heidelberg.

Lippert, K., Galinski, E., 1992. Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. Applied Microbiology and Biotechnology 37, 61–65.

Mayer, A.F., Hellmuth, K., Schlieker, H., Lopez-Ulibarri, R., Oertel, S., Dahlems, U., Strasser, A.W., van Loon, A.P., 1999. An expression system matures: a highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*. Biotechnology and Bioengineering 63 (3), 373–381.

Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., Shinmyo, A., Takano, M., Murooka, Y., 1999. Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. Journal of Bacteriology 181 (1), 91–99.

Pastor, J., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L., Iborra, J., Vargas, C., Nieto, J., Cánovas, M., 2010. Ectoines in cell stress protection: uses and biotechnological production. Biotechnology Advances 28 (6), 782–801.

Peters, P., Galinski, E., Trüper, H., 1990. The biosynthesis of ectoine. FEMS Microbiology Letters 71, 157–162.

Reshetnikov, A., Khmelenina, V., Mustakhimov, I., Trotsenko, Y., 2011. Genes and enzymes of ectoine biosynthesis in halotolerant methanotrophs. Methods in Enzymology 495, 15–30.

Sauer, B., 1987. Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. Molecular and Cellular Biology 7 (6), 2087–2096.

Sauer, T., Galinski, E., 1998. Bacterial milking: a novel bioprocess for production of compatible solutes. Biotechnology and Bioengineering 57 (3), 306–313.

Schiraldi, C., Maresca, C., Catapano, A., Galinski, E., de Rosa, M., 2006. High-yield cultivation of *Marinococcus* M52 for production and recovery of hydroxyectoine. Research in Microbiology 157 (7), 693–699.

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

G Model
BIOTEC-6289; No. of Pages 9
ARTICLE IN PRESS
E. Eilert et al. / Journal of Biotechnology xxx (2013) xxx–xxx

Schubert, T., Maskow, T., Benndorf, D., Harms, H., Breuer, U., 2007. Continuous synthesis and excretion of the compatible solute ectoine by a transgenic, nonhalophilic bacterium. Applied and Environment Microbiology 73 (10), 3343–3347.

Schwibbert, K., Marin-Sanguino, A., Bagyan, I., Heidrich, G., Lentzen, G., Seitz, H., Rampp, M., Schuster, S.C., Klenk, H.-P., Pfeiffer, F., Oesterhelt, D., Kunte, H., 2010. A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581. Environmental Microbiology 13 (8), 1973–1994.

Seip, B., Galinski, E., Kurz, M., 2011. Natural and engineered hydroxyectoine production based on the *Pseudomonas stutzeri* ectABCD-ask gene cluster. Applied and Environment Microbiology 77 (4), 1368–1374.

Stöveken, N., Pittelkow, M., Sinner, T., Jensen, R., Heider, J., Bremer, E., 2011. A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *Pseudomonas stutzeri* A1501. Journal of Bacteriology 193 (17), 4456–4468.

Tamás, M.J., Luyten, K., Sutherland, C.W., Hernandez, A., Albertyn, J., Valadi, H., Li, H., Prior, B.A., Kilian, S.G., Ramos, J., Gustafsson, L., Thevelein, J.M., Hohmann, S., 1999. Fps1p controls the accumulation and release of compatible solute glycerol in yeast osmoregulation. Molecular Microbiology 31 (4), 1087–1104.

van Eck, J.H., Prior, B.A., Brandt, E.V., 1993. The water relations of growth and polyhydroxyl alcohol production by ascomycetous yeasts. Journal of General Microbiology 139, 1047–1054.

van Zyl, P.J., Kilian, S.G., Prior, B.A., 1990. The role of an active transport mechanism in glycerol accumulation during osmoregulation by *Zygosaccharomyces rouxii*. Applied Microbiology and Biotechnology 34, 231–235.

Vargas, C., Jebbar, M., Carrasco, R., Blanco, C., Calderón, M.I., Iglesias-Guerra, F., Nieto, J.J., 2006. Ectoines as compatible solutes and carbon and energy sources for the halophilic bacterium *Chromohalobacter salexigens*. Journal of Applied Microbiology 100 (1), 98–107.

Vargas, C., Argandoña, M., Reina-Bueno, M., Rodríguez-Moya, J., Fernández-Aunión, C., Nieto, J., 2008. Unravelling the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with broad salinity tolerance. Saline Systems 4, 14.

Veenhuis, M., van Dijken, J.P., Harder, W., 1983. The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts. Advances in Microbial Physiology 24, 1–82.

Vermeulen, V., Kunte, H.J., 2004. *Marinococcus halophilus* DSM 20408T encodes two transporters for compatible solutes belonging to the betaine–carnitine–choline transporter family: identification and characterization of ectoine transporter EctM and glycine betaine transporter BetM. Extremophiles 8 (3), 175–184.

Viola, R.E., 2001. The central enzymes of the aspartate family of amino acid biosynthesis. Accounts of Chemical Research 34 (5), 339–349.

Weydemann, U., Keup, P., Piontek, M., Strasser, A.W., Schweden, J., Gellissen, G., Janowicz, Z.A., 1995. High-level secretion of hirudin by *Hansenula polymorpha* – authentic processing of three different preprohirudins. Applied Microbiology and Biotechnology 44 (3–4), 377–385.

Wohlfarth, A., Severin, J., Galinski, E.A., 1990. The spectrum of compatible solutes in heterotrophic halophilic eubacteria of the family *Halomondaceae*. Journal of General Microbiology 136, 705–712.

Please cite this article in press as: Eilert, E., et al., Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J. Biotechnol. (2013), http://dx.doi.org/10.1016/j.jbiotec.2013.02.005

5.2 Supplementary Data

5.2.1 Abstract

The manuscript is supplemented with further data regarding fermentation strategies. Prior to the final fermentation described in Manuscript III, two additional fermentation procedures were tested to evaluate the impact of MeOH on 5-hydroxyectoine synthesis. We could identify a correlation between the amount of MeOH in the mixed feed and the velocity of 5-hydroxyectoine formation. Based on the results of these two initial fermentations an improved fermentation strategy was designed and performed as described in Manuscript III.

5.2.2 Introduction

The main function of ectoine and 5-hydroxyectoine production is to counteract the outflow of water from the cell that occurs under high osmolarity conditions (Wohlfarth et al., 1990). This function was also shown for the precursor Nγ-acetyl-diaminobutyrate (Cánovas et al., 1997). The addition of 5-hydroxyectoine and its precursors ectoine and Nγ-acetyl-diaminobutyrate was shown to enhance salt tolerance of various organisms like *E. coli* (Jebbar et al., 1997) or *Salmonella entericia* serovar Typhimurium (García-Estepa et al., 2006).

5.2.3 Materials and methods

All other experiments were performed as described earlier in Manuscript III.

Buffering solutions

The following buffering solutions were used throughout this study: pH 3: citric acid/NaOH, pH 4: glycine/NaOH, pH 5: citric acid/Na_2HPO_4, pH 6: 2-(N-morpholino)ethanesulfonic acid(MES)/HCl or NaOH and pH 7: Na_2HPO_4/NaH_2PO_4. 1 M stock solutions of the first component were prepared, while the second component was used to adjust the pH of the buffering solution. Buffer was added to the medium to a final concentration of 100 mM after the autoclaving.

5.2.4 Results

Initial strategies of 5-hydroxyectoine production in 1 liter scale

Sorbitol was identified as the derepressing carbon source with highest 5-hydroxyectoine yield in small scale cultivation (described in Manuscript III). In addition MeOH induction at an early point of the cultivation was found to be crucial to enhance productivity of the strain ALU3/EctABCD #1. The

best conditions found during small scale cultivation were used as the basis for the design of two fermentations in 1 L scale.

As described in Manuscript III we used SYN6 medium, a synthetic minimal medium for yeast that is similar to YNB but leads to higher cell densities. During cultivation the pH was maintained at 5. For initial growth the medium was supplemented with 2% sorbitol. After 22 h a mixed carbon source feed was started. Two strategies were tested: 1.6 g L^{-1} h^{-1} MeOH and 0.4 g L^{-1} h^{-1} sorbitol (HPEct_01, Fig. 5.1) or 1 g L^{-1} h^{-1} MeOH and 1 g L^{-1} h^{-1} sorbitol (HPEct_02, Fig. 5.2) . After 142 h the feed was switched to 2 g L^{-1} h^{-1} glucose for reactivation of the cell metabolisms (Fig. 5.1 A, B and C (HPEct_01) and Fig. 5.2 A, B and C (HPEct_02)). Fermentation was stopped after 160 h. The time courses of dry cell weight (dcw) and yields of ectoine and 5-hydroxyectoine in the culture supernatant are shown in Fig. 5.1 D and Fig. 5.2 D. During the batch phase only slight ectoine and 5-hydroxyectoine syntheses occurred. After the start of the mixed carbon source feed after 22 h the ectoine amount decreases while the 5-hydroxyectoine amount increases suggesting that EctD now efficiently converts all ectoine to 5-hydroxyectoine. The 5-hydroxyectoine concentration in the culture supernatant further increases throughout the fermentation in both cases. The ectoine concentration remains at a low level.

The 5-hydroxyectoine formation clearly depends on MeOH induction. After about 138 h of fermentation the cells reached the stationary phase, and the synthesis rate was slowing down. The mixed feed was exchanged for a glucose feed in both cases. Consequently cell growth was stimulated. In case of HPEct_01 the 5-hydroxyectoine concentration increased again. In case of HPEct_02 no increment in 5-hydroxyectoine concentration was detected. However, an increase in ectoine synthesis was detected. Analysis of the intracellular soluble fraction of the yeast cells showed that only small amounts of ectoine and 5-hydroxyectoine were retained in the cell and thus no significant accumulation occurred (data not shown). Yields of both conditions were comparable. Strain ALU3/EctBACD #1 produced about 2 g/L 5-hydroxyectoine (243 μmol/g dcw at 160 h) in the culture supernatant with only 70 mg/L ectoine (9 μmol/g dcw at 160 h) as a side product.

The fermentation HPEct_02 was impaired when accidentally the gassing was interrupted after 76 h (indicated by the asterisk in Fig. 5.2D). Gassing was restored after 93 h. The cells recovered from the shortage of oxygen. Despite of this incident the maximum of 5-hydroxyectoine yield was reached earlier in comparison to HPEct_01, which however resulted in a higher dcw.

We concluded that a high MeOH percentage in the mixed feed could contribute to a further optimization of the 5-hydroxyectoine production in 1 L scale. Based on these results a third fermentation was performed. As an optimization based on the two strategies described here, we decided to increase the feeding rate of the mixed feed throughout the fermentation and to increase the MeOH concentration within the mixed feed solution to achieve an enhanced 5-hydroxyectoine synthesis proportional to the MeOH amount. This final fermentation is described in detail in Manuscript III.

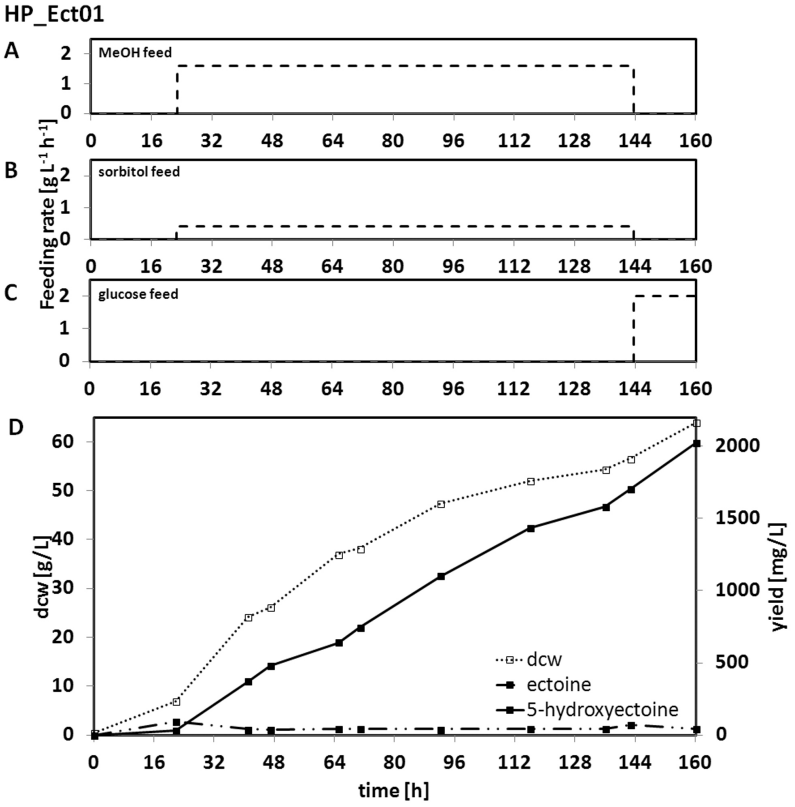


Figure 5.1: **Fermentation condition HPEct_01: 37.5% (w/v) sorbitol/37.5% (v/v) MeOH mixed feed.** The fermentation was carried out as described in Materials & Methods at 37°C in SYN6 at pH 5. (A) + (B) + (C) Feeding modi. 20 g/L sorbitol were provided as batch for initial cell growth. After consumption of the carbon source (assumed after 23 h) a [37.5% (w/v) sorbitol/37.5% (v/v) MeOH] solution has been applied as a mixed feed with 2 g L^{-1} h^{-1}. After 142 h the feeding solution was exchanged with a 75% (w/v) glucose solution without alteration of the feeding rate. Fermentation was terminated after 160 h. (D) Corresponding courses of dry cell weight (dcw), and ectoine or 5-hydroxyectoine (mg/L) in the culture supernatant as measured by HPLC. The values correspond to a single measurement each.

5.2.5 Discussion

H. polymorpha is suitable for staggered directed integration of single copy expression cassettes resulting in the expression of a complete heterologous biosynthesis pathway.

Site-directed integration of a single expression cassette in *H. polymorpha* was first described by Faber et al. (1992). The authors performed various integrations into the *MOX* locus. Later on the same working group performed studies on site-directed integration using the amino oxidase (*AMO*) locus (Faber et al., 1993) and also Agaphonov et al. (1995) performed site-directed integration using the *TRP3* locus.

In this study directed integration of single expression cassettes was chosen for the integration of the genes encoding the enzymes of the 5-hydroxyectoine biosynthesis pathway. This was performed using flanking sequences of approximately 1 kbp with homology to the target locus. The integration was independent of the condition of the target locus, it was performed targeting both intact and defective

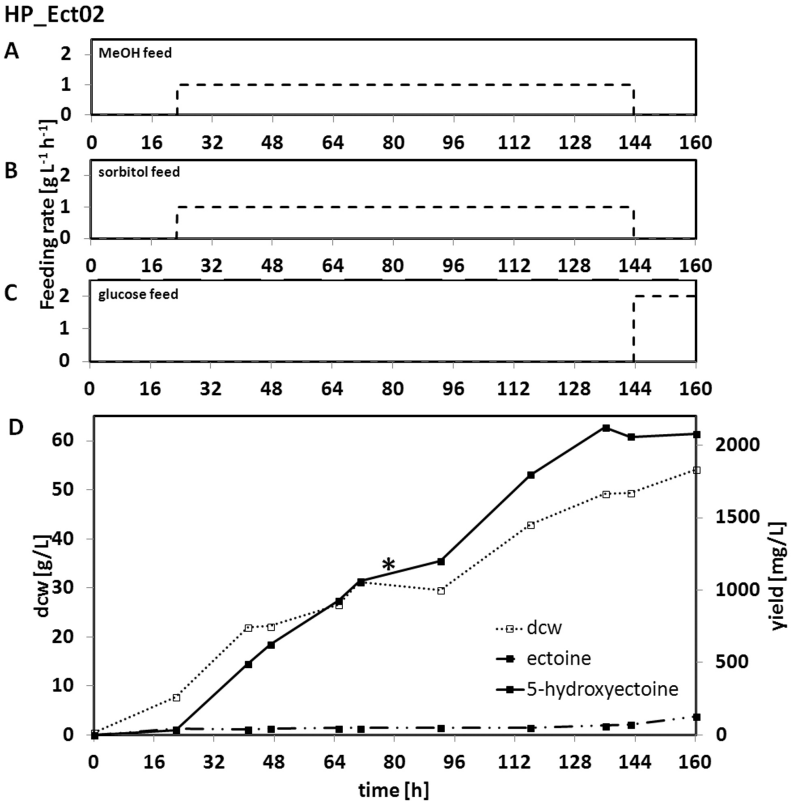


Figure 5.2: **Fermentation condition HPEct_02: 15% (w/v) sorbitol/60% (v/v) MeOH mixed feed..** The fermentation was carried out as described in Materials & Methods at 37°C in SYN6 at pH 5. (A) + (B) + (C) Feeding modi. 20 g/L sorbitol were provided as batch for initial cell growth. After consumption of the carbon source (assumed after 23 h) a [15% (w/v) sorbitol/60% (v/v) MeOH] solution has been applied as a mixed feed with 2 g L^{-1} h^{-1}. After 76 h gassing was disrupted by a technical defect (indicated by an asterisk), it was restored after 93 h. After 142 h the feeding solution was exchanged with a 75% (w/v) glucose solution without alteration of the feeding rate. Fermentation was terminated after 160 h. (D) Corresponding courses of dry cell weight (dcw), and ectoine or 5-hydroxyectoine (mg/L) in the culture supernatant as measured by HPLC. The values correspond to a single measurement each.

gene loci. Successful integration correlated with the following measurable criteria: 1. Integration disrupted the intact target locus resulting in a corresponding minus phenotype. 2. Integration of the cassette was performed using a selection marker resulting in a corresponding plus phenotype. 3. Integration was proven by PCR which could distinguish the recombinant expression cassette from the native target locus.

By integration of an expression cassette into an intact target locus a new auxotrophy was generated. In this study the staggered integration of a single copy of each gene was chosen to enable the use of more selection markers than initially available for the host strain. In addition to the *ADE1syn*, *LEU2syn* and *ODC1* the *HIS4* marker was used in the strain ALU3 (*ade1 leu2 odc1*).

Although all genes were shown to be integrated in only one copy, the expression levels of all enzymes were very diverse. Estimation by Western blot analysis showed that EctB was expressed at the highest levels, whereas EctD was only detected in small amounts. As discussed earlier in Manuscript II (Supplementary Data) the expression level of a heterologous protein may be influenced by different

factors such as the codon usage, gene dosage, promoter, integration locus, A+T composition of cDNA, translation rate, endogenous protease activity and cultivation conditions (Sreekrishna et al., 1997). In Western blot analyses signal intensities may also differ due to different sensitivities of the antibodies used. For a more detailed evaluation of the impact of the target locus and selection marker on the expression of these proteins, all genes should be integrated in each locus in combination with every selection marker used in this study. From the data of this study no further assumptions can be made.

Efficient conversion of ectoine to 5-hydroxyectoine as well as overall yield is dependent on media pH and carbon source. Noteworthy, conversion is highly increased by usage of MeOH as carbon source.

Protein expression, protein folding and activity of an enzyme may be strongly influenced by the pH value of the medium. In this study 5-hydroxyectoine synthesis occurred within a broad range of pH values. Highest conversion yields in small scale cultivation were obtained at pH 3, 5 and 6. Since the pH value can only be adjusted outside the cells, it remains unknown how the pH inside the cell is influenced by this pH adjustment.

A pH optimum could not be determined for the 5-hydroxyectoine synthesis. The data suggest that the conversion yield may also be influenced by the different buffer components in small scale cultivations. The buffer components may have an even stronger influence on the conversion yield than the pH. The following buffer components were used: pH 3: citric acid/NaOH, pH 4: glycine/NaOH, pH 5: citric acid/Na_2HPO_4, pH 6: MES and pH 7: Na_2HPO_4/NaH_2PO_4. Two of the buffer solutions resulting in high 5-hydroxyectoine synthesis contained citric acid, whereas one contained MES. Those two components were not part of any other buffer solution used in the study. As a precursor of α-ketoglutarate within the citric acid cycle, citric acid is involved in the biosynthesis of various amino acids. Amongst these amino acids glutamate can be found, which is required during 5-hydroxyectoine synthesis. The positive effect of MES on 5-hydroxyectoine synthesis could not be explained.

When *H. polymorpha* cells are starved before feeding with MeOH, accumulation of glutamate, glutamine, alanine and formate was observed (Jones, 1991). After the first growth phase during fermentation, methanol induction was started when cells showed signs of carbon source starvation. Hence, elevated levels of intracellular glutamate may be present which might increase the ectoine and 5-hydroxyectoine synthesis rate.

The conversion of ectoine to 5-hydroxyectoine is almost complete.

The homologous producer *H. elongata* could only be optimized to convert 50% of the ectoine into 5-hydroxyectoine (Seip et al., 2011). In this study almost full conversion was achieved by expression of the genes derived from *H. elongata*. Even though EctD is able to catalyze the forward and reverse reaction between ectoine and 5-hydroxyectoine the ectoine level remained low while conversion towards

5-hydroxyectoine was favored during fermentation. These four enzymes are able to catalyze an almost full conversion. The 5-hydroxyectoine yield may be increased further by alteration of the gene copy number (e.g. introduction of an additional *ectA, ectB, ectC* or *ectD* copies). The application of different promoters may be an additional option to further improve the synthesis yield. To evaluate the dependency of the 5-hydroxyectoine synthesis on the MeOH-inducible *FMD* promoter, it would also be interesting to evaluate the heat-inducible *TPS1*-promoter which shows a strength comparable to that of the *FMD* promoter (Amuel et al., 2000). In addition the *PMA1* promoter (Cox et al., 2000) or the *Arxula adeninivorans*-derived *TEF1*-promoter (Terentiev et al., 2004) may be tested as these are constitutive promoters functional in *H. polymorpha*. The impact of MeOH induced protein expression on the synthesis yield of 5-hydroxyectoine has to be studied in more detail.

6 Concluding Summary and Outlook

In this study we aimed to explore new approaches for heterologous gene expression in *H. polymorpha* with the main goal to improve extra- and intracellular protein expression. First, we developed a new strategy for protein secretion by optimization of the processing of the MFα prepro sequence (Manuscript I). In a second approach we developed a system for high level intracellular protein expression based on a fusion protein strategy. A protein expressed by this approach was not only generated in high levels but showed also a highly increased stability in comparison to a secretory approach (Manuscript II). Last, we established a new metabolic pathway for the production of 5-hydroxyectoine in H. polymorpha by multi-step integration of four heterologous genes encoding enzymes of the salt tolerant bacterium *H. elongata* (Manuscript III).

We found that the insertion of Ala-Pro between the MFα prepro sequence and the protein of interest leads to an improved Kex2p-dependent processing of the signal sequence in case of IL-6 and IFNα-2a. The improved cleavage could be explained by modeling which indicates an Ala-Pro-dependent improved accessibility of the Kex2p cleavage site. Studies concerning the influence of other adjacent amino acids on cleavage at the Lys-Arg motif may help to further optimize the processing of this heterologous signal sequence in *H. polymorpha*. In this context we found that the choice of the strain background had a surprisingly significant impact: the processing was affected differently in different *H. polymorpha* host strains. Evaluation of further strains may therefore be recommendable in cases of incomplete or incorrect processing of a target protein fused to theMFα prepro sequence. With regard to the N-terminal Ala-Pro a partial or even complete removal of this dipeptide was observed. A putative protease responsible for the processing of Ala-Pro dipeptides was identified by gene deletion. The protein sequence shows significant similarity to the *S. cerevisiae* Ste13p. Co-overexpression of this putative protease may provide quantitative removal of the N-terminal Ala-Pro dipeptide after Kex2p-dependent processing. Without removal of the Ala-Pro dipeptide an application of the MFα-Ala-Pro prepro sequence may still be useful for the expression of technical enzymes or other cases in which an authentic N-terminus is not strictly necessary.

This study marks the first successful approach of carrier-protein-driven cytosolic targeting in *H. polymorpha*. The strongly expressed cytosolic FTH1 was analyzed with regard to its potential as fusion partner for difficult-to-express target proteins. The fusion of PTH and FTH1 significantly enhanced PTH stability. We could show that the expression level of the fusion protein was at least as high as that of FTH1 without fusion partner. The solubility of the fusion protein was increased in comparison

to FTH1 alone and no degradation of PTH occurred when C-terminally fused to FTH1. To further optimize the concept other target proteins shall be tested and the recovery of the protein of interest has to be optimized for the use this concept in production processes.

We could demonstrate that the expression of the enzymes EctB, EctA, EctC and EctD in *H. polymorpha* resulted in the production of 5-hydroxyectoine in gram per liter scale with almost complete conversion of ectoine to 5-hydroxyectoine. We showed that the high level 5-hydroxyectoine formation requires the feeding with MeOH during cultivation. 5-Hydroxyectoine synthesis was downregulated by the addition of methionine, lysine or threonine indicating the existence of a feedback regulated aspartate kinase in *H. polymorpha*. A further optimization of the 5-hydroxyectoine synthesis by recombinant *H. polymorpha* may be achieved by optimizing the feeding strategy during fermentation. In addition, the supply of precursor molecules to the cell may be a target of further optimization. The data presented in this study show that *H. polymorpha* is a highly promising candidate for 5-hydroxyectoine production in an industrial scale.

List of Figures

List of Tables

References

Abad, S., Nahalka, J., Bergler, G., Arnold, S.A., Speight, R., Fotheringham, I., Nidetzky, B., Glieder, A., 2010. Stepwise engineering of a *Pichia pastoris* D-amino acid oxidase whole cell catalyst. Microbial cell factories 9, 24.

Achstetter, T., 1989. Regulation of alpha-factor production in *Saccharomyces cerevisiae*: a-factor pheromone-induced expression of the *MF alpha 1* and *STE13* genes. Molecular and cellular biology 9, 4507–4514.

Agaphonov, M.O., Beburov, M.Y., Ter-Avanesyan, M.D., Smirnov, V.N., 1995. A disruption-replacement approach for the targeted integration of foreign genes in *Hansenula polymorpha*. Yeast 11, 1241–1247.

Agaphonov, M.O., Sokolov, S.S., Romanova, N.V., Sohn, J.H., Kim, S.Y., Kalebina, T.S., Choi, E.S., Ter-Avanesyan, M.D., 2005. Mutation of the protein-O-mannosyltransferase enhances secretion of the human urokinase-type plasminogen activator in *Hansenula polymorpha*. Yeast 22, 1037–1047.

Akeboshi, H., Chiba, Y., Kasahara, Y., Takashiba, M., Takaoka, Y., Ohsawa, M., Tajima, Y., Kawashima, I., Tsuji, D., Itoh, K., Sakuraba, H., Jigami, Y., 2007. Production of recombinant beta-hexosaminidase A, a potential enzyme for replacement therapy for Tay-Sachs and Sandhoff diseases, in the methylotrophic yeast *Ogataea minuta*. Applied and environmental microbiology 73, 4805–4812.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research 25, 3389–3402.

Amuel, C., Gellissen, G., Hollenberg, C.P., Suckow, M., 2000. Analysis of heat shock promoters in *Hansenula polymorpha*: The *TPS1* promoter, a novel element for heterologous gene expression. Biotechnology and Bioprocess Engineering 5, 247–252.

Anbu Rajan, L., Joseph, T.C., Thampuran, N., James, R., Ashok Kumar, K., Viswanathan, C., Bansal, K.C., 2008. Cloning and heterologous expression of ectoine biosynthesis genes from *Bacillus halodurans* in *Escherichia coli*. Biotechnology letters 30, 1403–1407.

Barlow, D.J., Thornton, J.M., 1988. Helix geometry in proteins. Journal of molecular biology 201, 601–619.

Bautista-Muñoz, C., Hernández-Rodríguez, C., Villa-Tanaca, L., 2005. Analysis and expression of *STE13ca* gene encoding a putative X-prolyl dipeptidyl aminopeptidase from *Candida albicans*. FEMS immunology and medical microbiology 45, 459–469.

Bellu, A.R., van der Klei, I.J., Rechinger, K.B., Yavuz, M., Veenhuis, M., Kiel, J.A.K.W., 1999. Characterization of the *Hansenula polymorpha CPY* gene encoding carboxypeptidase Y. Yeast 15, 181–189.

Böer, E., Steinborn, G., Matros, A., Mock, H.P., Gellissen, G., Kunze, G., 2007. Production of interleukin-6 in *Arxula adeninivorans*, *Hansenula polymorpha* and *Saccharomyces cerevisiae* by applying the wide-range yeast vector (CoMed) system to simultaneous comparative assessment. FEMS yeast research 7, 1181–1187.

Boratyn, G.M., Schäffer, A.A., Agarwala, R., Altschul, S.F., Lipman, D.J., Madden, T.L., 2012. Domain enhanced lookup time accelerated BLAST. Biology direct 7, 12.

Bosse-Doenecke, E., Weininger, U., Gopalswamy, M., Balbach, J., Knudsen, S.M., Rudolph, R., 2008. High yield production of recombinant native and modified peptides exemplified by ligands for G-protein coupled receptors. Protein expression and purification 58, 114–121.

Brake, A.J., Merryweather, J.P., Coit, D.G., Heberlein, U.A., Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P., Barr, P.J., 1984. Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces* cerevisiae. Proceedings of the National Academy of Sciences of the United States of America 81, 4642–4646.

Brenner, C., Fuller, R.S., 1992. Structural and enzymatic characterization of a purified prohormone-processing enzyme: secreted, soluble Kex2 protease. Proceedings of the National Academy of Sciences of the United States of America 89, 922–926.

Brickner, J.H., Fuller, R.S., 1997. *SOI1* encodes a novel, conserved protein that promotes TGN-endosomal cycling of Kex2p and other membrane proteins by modulating the function of two TGN localization signals. The Journal of cell biology 139, 23–36.

Bryant, N.J., Stevens, T.H., 1997. Two separate signals act independently to localize a yeast late Golgi membrane protein through a combination of retrieval and retention. The Journal of cell biology 136, 287–297.

Buckholz, R.G., 1993. Yeast systems for the expression of heterologous gene products. Current opinion in biotechnology 4, 538–542.

Bursy, J., Pierik, A.J., Pica, N., Bremer, E., 2007. Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. The Journal of biological chemistry 282, 31147–31155.

Cánovas, D., Vargas, C., Iglesias-Guerra, F., Csonka, L.N., Rhodes, D., Ventosa, A., Nieto, J.J., 1997. Isolation and characterization of salt-sensitive mutants of the moderate halophile *Halomonas elongata* and cloning of the ectoine synthesis genes. The Journal of biological chemistry 272, 25794–25801.

Chaudhuri, B., Latham, S.E., Stephan, C., 1992. A mutant Kex2 enzyme with a C-terminal HDEL sequence releases correctly folded human insulin-like growth factor-1 from a precursor accumulated in the yeast endoplasmic reticulum. European journal of biochemistry 210, 811–822.

Chen, J., Sun, H.Y., Yang, Y., Wang, X.F., Chen, S.Q., 2008. Construction, expression and characterization of recombinant fusion protein HSA-PTH (1-34) in *Pichia pastoris*. Zhejiang da xue xue bao. Yi xue ban = Journal of Zhejiang University. Medical sciences 37, 126–133.

Chen, R.R., 2007. Permeability issues in whole-cell bioprocesses and cellular membrane engineering. Applied microbiology and biotechnology 74, 730–738.

Chen, X., Gao, B., Shi, W., Li, Y., 1992. Expression and secretion of human interferon alpha A in yeast *Kluyveromyces lactis*. Yi chuan xue bao = Acta genetica Sinica 19, 284–288.

Chin-Joe, I., Straathof, A.J.J., Pronk, J.T., Jongejan, J.A., Heijnen, J.J., 2001. Influence of the ethanol and glucose supply rate on the rate and enantioselectivity of 3-oxo ester reduction by baker's yeast. Biotechnology and bioengineering 75, 29–38.

Cho, E.Y., Cheon, S.A., Kim, H., Choo, J., Lee, D.J., Ryu, H.M., Rhee, S.K., Chung, B.H., Kim, J.Y., Kang, H.A., 2010. Multiple-yapsin-deficient mutant strains for high-level production of intact recombinant proteins in *Saccharomyces cerevisiae*. Journal of biotechnology 149, 1–7.

Clare, J.J., Romanos, M.A., Rayment, F.B., Rowedder, J.E., Smith, M.A., Payne, M.M., Sreekrishna, K., Henwood, C.A., 1991. Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. Gene 105, 205–212.

Cox, H., Mead, D., Sudbery, P., Eland, R.M., Mannazzu, I., Evans, L., 2000. Constitutive expression of recombinant proteins in the methylotrophic yeast *Hansenula polymorpha* using the *PMA1* promoter. Yeast (Chichester, England) 16, 1191–1203.

Cui, T., Puah, C.M., Liew, O.W., Lee, S.H., 2007. Method for catalysing a fusion polypeptide undergoing a cleavage reaction WO Patent 2007/149053 A1.

Cunningham, D.F., O'Connor, B., 1997. Proline specific peptidases. Biochimica et biophysica acta 1343, 160–186.

van Dijk, R., Faber, K.N., Kiel, J.A.K.W., Veenhuis, M., van der Klei, I.J., 2000. The methylotrophic yeast *Hansenula polymorpha*: a versatile cell factory. Enzyme and microbial technology 26, 793–800.

Dmochowska, A., Dignard, D., Henning, D., Thomas, D.Y., Bussey, H., 1987. Yeast *KEX1* gene encodes a putative protease with a carboxypeptidase B-like function involved in killer toxin and alpha-factor precursor processing. Cell 50, 573–584.

Egel-Mitani, M., Hansen, M.T., 1987. Nucleotide sequence of the gene encoding the *Saccharomyces kluyveri* alpha mating pheromone. Nucleic acids research 15, 6303.

Eggeling, L., Sahm, H., 1978. Derepression and partial insensitivity to carbon catabolite repression of the methanol dissimilating enzymes in *Hansenula polymorpha*. European Journal of Applied Microbiology and Biotechnology 5, 197–202.

Eggeling, L., Sahm, H., 1980. Regulation of alcohol oxidase synthesis in *Hansenula polymorpha*: oversynthesis during growth on mixed substrates and induction by methanol. Archives of microbiology 127, 119–124.

Egli, T., van Dijken, J.P., Veenhuis, M., Harder, W., Fiechter, A., 1980. Methanol Metabolism in Yeasts: Regulation of Synthesis of Catabolic Enzymes. Archives of microbiology 124, 115–121.

Eldarov, M.A., Mardanov, A.V., Beletsky, A.V., Ravin, N.V., Skryabin, K.G., 2011. Complete sequence and analysis of the mitochondrial genome of the methylotrophic yeast *Hansenula polymorpha* DL-1. FEMS yeast research 11, 464–472.

Ernst, J.F., 1988. Efficient secretion and processing of heterologous proteins in *Saccharomyces cerevisiae* is mediated solely by the pre-segment of alpha-factor precursor. DNA 7, 355–360.

Faber, K.N., Haima, P., Harder, W., Veenhuis, M., AB, G., 1994. Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*. Current genetics 25, 305–310.

Faber, K.N., Haima, P., de Hoop, M.J., Harder, W., Veenhuis, M., AB, G., 1993. Peroxisomal amine oxidase of *Hansenula polymorpha* does not require its SRL-containing C-terminal sequence for targeting. Yeast 9, 331–338.

Faber, K.N., Swaving, G.J., Faber, F., AB, G., Harder, W., Veenhuis, M., Haima, P., 1992. Chromosomal targeting of replicating plasmids in the yeast *Hansenula polymorpha*. Journal of general microbiology 138, 2405–2416.

Faber, K.N., Westra, S., Waterham, H.R., Keizer-Gunnink, I., Harder, W., Veenhuis, G.A., 1996. Foreign gene expression in *Hansenula polymorpha*. A system for the synthesis of small functional peptides. Applied microbiology and biotechnology 45, 72–79.

Frings, E., Sauer, T., Galinski, E.A., 1995. Production of hydroxyectoine: high cell-density cultivation and osmotic downshock of *Marinococcus* strain M52. Journal of biotechnology 43, 53–61.

Fuller, R.S., Sterne, R.E., Thorner, J., 1988. Enzymes required for yeast prohormone processing. Annual review of physiology 50, 345–362.

Galinski, E.A., Pfeiffer, H.P., Trüper, H.G., 1985. 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. European journal of biochemistry 149, 135–139.

Gangireddy, S.R., Madhavi, R.D., Ravikanth, K.R., Reddy, P.K., Konda, V.R., Rao, K.R.S.S., Gunwar, S., 2010. High yield expression of human recombinant PTH (1-34). Current Trends in Biotechnology and Pharmacy 4, 568–577.

García-Estepa, R., Cánovas, D., Iglesias-Guerra, F., Ventosa, A., Csonka, L.N., Nieto, J.J., Vargas, C., 2006. Osmoprotection of *Salmonella enterica* serovar Typhimurium by Ngamma-acetyldiaminobutyrate, the precursor of the compatible solute ectoine. Systematic and applied microbiology 29, 626–633.

Gasparian, M.E., Bychkov, M.L., Dolgikh, D.A., Kirpichnikov, M.P., 2011. Strategy for improvement of enteropeptidase efficiency in tag removal processes. Protein expression and purification 79, 191–196.

Gellissen, G., Hollenberg, C.P., 1997. Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. Gene 190, 87–97.

Gellissen, G., Janowicz, Z.A., Weydemann, U., Melber, K., Strasser, A.W.M., Hollenberg, C.P., 1992. High-level expression of foreign genes in *Hansenula polymorpha*. Biotechnology Advances 10, 179–189.

Gellissen, G., Piontek, M., Dahlems, U., Jenzelewski, V., Gavagan, J.E., DiCosimo, R., Anton, D.L., Janowicz, Z.A., 1996. Recombinant *Hansenula polymorpha* as a biocatalyst: coexpression of the spinach glycolate oxidase (GO) and the *S. cerevisiae* catalase T (CTT1) gene. Applied microbiology and biotechnology 46, 46–54.

Ghosalkar, A., Sahai, V., Srivastava, A., 2008. Secretory expression of interferon-alpha 2b in recombinant *Pichia pastoris* using three different secretion signals. Protein expression and purification 60, 103–109.

Gidijala, L., Kiel, J.A.K.W., Douma, R.D., Seifar, R.M., van Gulik, W.M., Bovenberg, R.A.L., Veenhuis, M., van der Klei, I.J., 2009. An engineered yeast efficiently secreting penicillin. PloS one 4, e8317.

Giga-Hama, Y., Tohda, H., Okada, H., Owada, M.K., Okayama, H., Kumagai, H., 1994. High-level expression of human lipocortin I in the fission yeast Schizosaccharomyces pombe using a novel expression vector. Bio/technology 12, 400–404.

Gödecke, S., Eckart, M., Janowicz, Z.A., Hollenberg, C.P., 1994. Identification of sequences responsible for transcriptional regulation of the strongly expressed methanol oxidase-encoding gene in *Hansenula polymorpha*. Gene 139, 35–42.

Goto, N., Prins, P., Nakao, M., Bonnal, R., Aerts, J., Katayama, T., 2010. BioRuby: bioinformatics software for the Ruby programming language. Bioinformatics (Bioinformatics) 26, 2617–2619.

Gough, S., Deshpande, M., Scher, M., Rosazza, J.P.N., 2001. Permeabilization of *Pichia pastoris* for glycolate oxidase activity. Biotechnology letters 23.

Gould, S.J., Keller, G.A., Hosken, N., Wilkinson, J., Subramani, S., 1989. A conserved tripeptide sorts proteins to peroxisomes. The Journal of cell biology 108, 1657–1664.

Guengerich, L., Kang, H.A., Behle, B., Gellissen, G., Suckow, M., 2004. A platform for heterologous gene expression based on the methylotrophic yeast *Hansenula polymorpha*, in: Kueck, U. (Ed.), Genetics and Biotechnology. Springer Verlag, Berlin Heidelberg.

Guisez, Y., Tison, B., Vandekerckhove, J., Demolder, J., Bauw, G., Haegeman, G., Fiers, W., Contreras, R., 1991. Production and purification of recombinant human interleukin-6 secreted by the yeast *Saccharomyces cerevisiae*. European journal of biochemistry 198, 217–222.

Gunkel, K., van Dijk, R., Veenhuis, M., van der Klei, I.J., 2004. Routing of *Hansenula polymorpha* alcohol oxidase: an alternative peroxisomal protein-sorting machinery. Molecular biology of the cell 15, 1347–1355.

Hansen, H., Didion, T., Thiemann, A., Veenhuis, M., Roggenkamp, R., 1992. Targeting sequences of the two major peroxisomal proteins in the methylotrophic yeast *Hansenula polymorpha*. Molecular & general genetics : MGG 235, 269–278.

van der Heide, M., 2002. *Hansenula polymorpha* as cell factory: secretion of flavo-proteins: Chapter 3: Secretion of *Aspergillus niger* glucose oxidase by *Hansenula polymorpha*: an analysis of secretion efficiency. Ph.D. thesis. (PhD thesis) Rijksuniversiteit Groningen.

van der Heide, M., Hollenberg, C.P., van der Klei, I.J., Veenhuis, M., 2002. Overproduction of BiP negatively affects the secretion of *Aspergillus niger* glucose oxidase by the yeast *Hansenula polymorpha*. Applied microbiology and biotechnology 58, 487–494.

Hermann, T., 2003. Industrial production of amino acids by coryneform bacteria. Journal of biotechnology 104, 155–172.

Hollenberg, C.P., Janowicz, Z.A., 1995. DNA-molecules coding for FMDH control regions and structural gene for a protein having FMDH-activity and their use thereof U.S. Patent No. 5389525.

Hosfield, T., Lu, Q., 1999. Influence of the amino acid residue downstream of $(Asp)_4Lys$ on enterokinase cleavage of a fusion protein. Analytical biochemistry 269, 10–16.

Humphrey, J.S., Peters, P.J., Yuan, L.C., Bonifacino, J.S., 1993. Localization of TGN38 to the trans-Golgi network: involvement of a cytoplasmic tyrosine-containing sequence. The Journal of cell biology 120, 1123–1135.

Idiris, A., Tohda, H., Kumagai, H., Takegawa, K., 2010. Engineering of protein secretion in yeast: strategies and impact on protein production. Applied microbiology and biotechnology 86, 403–417.

Inbar, L., Lapidot, A., 1988. The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ^{13}C- and ^{15}N-labeled L-glutamate and ^{13}C and ^{15}N NMR spectroscopy. The Journal of biological chemistry 263, 16014–16022.

Janowicz, Z.A., Melber, K., Merckelbach, A., Jacobs, E., Harford, N., Comberbach, M., Hollenberg, C.P., 1991. Simultaneous expression of the S and L surface antigens of hepatitis B, and formation of mixed particles in the methylotrophic yeast, *Hansenula polymorpha*. Yeast 7, 431–443.

Jebbar, M., von Blohn, C., Bremer, E., 1997. Ectoine functions as an osmoprotectant in *Bacillus subtilis* and is accumulated via the ABC-transport system OpuC. FEMS microbiology letters 154, 325–330.

Jones, E.W., 1991. Three proteolytic systems in the yeast *Saccharomyces cerevisiae*. The Journal of biological chemistry 266, 7963–7966.

Julius, D., Blair, L., Brake, A., Sprague, G., Thorner, J., 1983. Yeast alpha factor is processed from a larger precursor polypeptide: the essential role of a membrane-bound dipeptidyl aminopeptidase. Cell 32, 839–852.

Julius, D., Brake, A., Blair, L., Kunisawa, R., Thorner, J., 1984a. Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-alpha-factor. Cell 37, 1075–1089.

Julius, D., Schekman, R., Thorner, J., 1984b. Glycosylation and processing of prepro-alpha-factor through the yeast secretory pathway. Cell 36, 309–318.

Keränen, S., Toikkanen, J., Tieaho, V., Söderlund, H., 2000. Increased production of secreted proteins by recombinant yeast cells. European Patent No. 0994955B1.

Kern, A., Tilley, E., Hunter, I.S., Legisa, M., Glieder, A., 2007. Engineering primary metabolic pathways of industrial micro-organisms. Journal of biotechnology 129, 6–29.

Kim, S.Y., Sohn, J.H., Pyun, Y.R., Choi, E.S., 2002. A cell surface display system using novel GPI-anchored proteins in *Hansenula polymorpha*. Yeast 19, 1153–1163.

Kim, S.Y., Sohn, J.H., Pyun, Y.R., Choi, E.S., 2007. Variations in protein glycosylation in *Hansenula polymorpha* depending on cell culture stage. Journal of microbiology and biotechnology 17, 1949–1954.

Klabunde, J., Kleebank, S., Piontek, M., Hollenberg, C.P., Hellwig, S., Degelmann, A., 2007. Increase of calnexin gene dosage boosts the secretion of heterologous proteins by *Hansenula polymorpha*. FEMS yeast research 7, 1168–1180.

Kreil, G., 1990. Processing of precursors by dipeptidylaminopeptidases: a case of molecular ticketing. Trends in biochemical sciences 15, 23–26.

Kurjan, J., Herskowitz, I., 1982. Structure of a yeast pheromone gene (*MF alpha*): a putative alpha-factor precursor contains four tandem copies of mature alpha-factor. Cell 30, 933–943.

La Vallie, E.R., Riggs, P., 1994. Enzymatic and chemical cleavage of fusion proteins: Supplement 28, in: Ausubel, F.M. (Ed.), Current Protocols in Molecular Biology Online. Current Protocols and John Wiley & Sons, Incorporated, Hoboken.

Lahtchev, K.L., Semenova, V.D., Tolstorukov, I.I., van der Klei, I.J., Veenhuis, M., 2002. Isolation and properties of genetically defined strains of the methylotrophic yeast *Hansenula polymorpha* CBS4732. Archives of microbiology 177, 150–158.

Ledeboer, A.M., Edens, L., Maat, J., Visser, C., Bos, J.W., Verrips, C.T., Janowicz, Z.A., Eckart, M., Roggenkamp, R., Hollenberg, C.P., 1985. Molecular cloning and characterization of a gene coding for methanol oxidase in *Hansenula polymorpha*. Nucleic acids research 13, 3063–3082.

Lee, S.Y., Choi, J.H., Xu, Z., 2003. Microbial cell-surface display. Trends in biotechnology 21, 45–52.

Lentzen, G., Schwarz, T., 2006a. Extremolytes: Natural compounds from extremophiles for versatile applications. Applied microbiology and biotechnology 72, 623–634.

Lentzen, G., Schwarz, T., 2006b. Kompatible Solute: Mikrobielle Herstellung und Anwendung, in: Antranikian, G. (Ed.), Angewandte Mikrobiologie. Springer-Verlag Berlin Heidelberg, Berlin and Heidelberg.

Li, H., Wang, Y., Xu, A., Li, S., Jin, S., Wu, D., 2011. Large-scale production, purification and bioactivity assay of recombinant human interleukin-6 in the methylotrophic yeast *Pichia pastoris*. FEMS yeast research 11, 160–167.

Liew, O.W., Ching Chong, J.P., Yandle, T.G., Brennan, S.O., 2005. Preparation of recombinant thioredoxin fused N-terminal proCNP: Analysis of enterokinase cleavage products reveals new enterokinase cleavage sites. Protein expression and purification 41, 332–340.

Lin, S., Houston-Cummings, N.R., Prinz, B., Moore, R., Bobrowicz, B., Davidson, R.C., Wildt, S., Stadheim, T.A., Zha, D., 2012. A novel fragment of antigen binding (Fab) surface display platform using glycoengineered *Pichia pastoris*. Journal of immunological methods 375, 159–165.

Linder, M.C., Munro, H.N., 1972. Assay of tissue ferritin. Analytical biochemistry 48, 266–278.

Lippert, K., Galinski, E.A., 1992. Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. Applied microbiology and biotechnology 37, 61–65.

Liu, Y., Fujita, Y., Kondo, A., Fukuda, H., 2000. Preparation of high-activity whole cell biocatalysts by permeabilization of recombinant yeasts with alcohol. Journal of bioscience and bioengineering 89, 554–558.

Lockhart, S.R., Zhao, R., Daniels, K.J., Soll, D.R., 2003. Alpha-pheromone-induced "shmooing" and gene regulation require white-opaque switching during *Candida albicans* mating. Eukaryotic cell 2, 847–855.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H., 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic acids research 39, D225–9.

Marston, F.A., 1986. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. The Biochemical journal 240, 1–12.

Matoba, S., Ogrydziak, D.M., 1989. A novel location for dipeptidyl aminopeptidase processing sites in the alkaline extracellular protease of *Yarrowia lipolytica*. The Journal of biological chemistry 264, 6037–6043.

Mayer, A.F., Hellmuth, K., Schlieker, H., Lopez-Ulibarri, R., Oertel, S., Dahlems, U., Strasser, A.W.M., van Loon, A.P., 1999. An expression system matures: a highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*. Biotechnology and bioengineering 63, 373–381.

Mendirette, S.K., Patel, P.R., Bandyopadhyay, S., Saraswat, V., Singh, S.K., 2012. Fusion protein systems suitable for high expression of peptides U.S. Patent No. 20120142891.

Meyer, D.I., 1988. Preprotein conformation: the year's major theme in translocation studies. Trends in biochemical sciences 13, 471–474.

Müller, S., Sandal, T., Kamp-Hansen, P., Dalbøge, H., 1998. Comparison of expression systems in the yeasts *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Klyveromyces lactis*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Cloning of two novel promoters from *Yarrowia lipolytica*. Yeast 14, 1267–1283.

Munro, S., Pelham, H.R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. Cell 48, 899–907.

Nakai, K., Horton, P., 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends in biochemical sciences 24, 34–36.

Nakano, A., Luini, A., 2010. Passage through the Golgi. Current opinion in cell biology 22, 471–478.

Nguyen, H.T.T., Nevoigt, E., 2009. Engineering of *Saccharomyces cerevisiae* for the production of dihydroxyacetone (DHA) from sugars: a proof of concept. Metabolic Engineering 11, 335–346.

Nothwehr, S.F., Roberts, C.J., Stevens, T.H., 1993. Membrane protein retention in the yeast Golgi apparatus: dipeptidyl aminopeptidase A is retained by a cytoplasmic signal containing aromatic residues. The Journal of cell biology 121, 1197–1209.

Ogata, K., Nishikawa, H., Ohsugi, M., 1969. A Yeast Capable of Utilizing Methanol. Agricultural and Biological Chemistry 33, 1519–1520.

Otzen, M., Wang, D., Lunenborg, M.G.J., van der Klei, I.J., 2005. *Hansenula polymorpha* Pex20p is an oligomer that binds the peroxisomal targeting signal 2 (PTS2). Journal of cell science 118, 3409–3418.

Park, E., Rapoport, T.A., 2012. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. Annual Review of Biophysics 41.

Pelham, H.R., Hardwick, K.G., Lewis, M.J., 1988. Sorting of soluble ER proteins in yeast. The EMBO journal 7, 1757–1762.

Peters, P., Galinski, E.A., Trüper, H.G., 1990. The biosynthesis of ectoine. FEMS microbiology letters 71, 157–162.

Popa, O., Babeanu, N., Vamanu, A., Cornea, P., Livadariu, O., Vamanu, E., Drugulescu, M., 2007. Production of optically active compounds by microbial biotransformation. Bulletin USAMV-CN 64.

Porro, D., Sauer, M., Branduardi, P., Mattanovich, D., 2005. Recombinant protein production in yeasts. Molecular biotechnology 31, 245–259.

Prabha, L., Govindappa, N., Adhikary, L., Melarkode, R., Sastry, K., 2009. Identification of the dipeptidyl aminopeptidase responsible for N-terminal clipping of recombinant Exendin-4 precursor expressed in *Pichia pastoris*. Protein expression and purification 64, 155–161.

Rapoport, T.A., 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature 450, 663–668.

Rodriguez-Moya, J., Argandona, M., Iglesias-Guerra, F., Nieto, J.J., Vargas, C., 2013. Temperature- and salinity-decoupled overproduction of hydroxyectoine by *Chromohalobacter salexigens*. Applied and Environmental Microbiology 79, 1018–1023.

Roggenkamp, R., Hansen, H., Eckart, M., Janowicz, Z.A., Hollenberg, C.P., 1986. Transformation of the methylotrophic yeast *Hansenula polymorpha* by autonomous replication and integration vectors. Molecular Genetics & Genomics 202, 302–308.

Rolf, T., 2011. Untersuchungen zur heterologen Synthese von löslichen Varianten der Kex2-Protease in der methylotrophen Hefe *H. polymorpha*. Ph.D. thesis. (Bachelor thesis) Fachhochschule Aachen. Jülich.

Rosales Rodriguez, I., Ahmad, M., Pichler, H., Schwab, H., 2011. Identification of the alpha mating factor from improved draft genome sequence of *Pichia pastoris* CBS7435 and investigating its ability for secretory heterologous protein expression. Prokagenomics (Göttingen, Germany) - Conference Poster.

Rucktäschel, R., Girzalsky, W., Erdmann, R., 2011. Protein import machineries of peroxisomes. Biochimica et biophysica acta 1808, 892–900.

Sakai, Y., Murdanoto, A.P., Konishi, T., Iwamatsu, A., Kato, N., 1997. Regulation of the formate dehydrogenase gene, *FDH1*, in the methylotrophic yeast *Candida boidinii* and growth characteristics of an *FDH1*-disrupted strain on methanol, methylamine, and choline. Journal of bacteriology 179, 4480–4485.

Sauer, T., Galinski, E.A., 1998. Bacterial milking: A novel bioprocess for production of compatible solutes. Biotechnology and bioengineering 57, 306–313.

Saunders, C.W., Pedroni, J.A., Monahan, P.M., 1991. Optimization of the signal-sequence cleavage site for secretion from *Bacillus subtilis* of a 34-amino acid fragment of human parathyroid hormone. Gene 102, 277–282.

Schaefer, J.V., Plückthun, A., 2012. Engineering aggregation resistance in IgG by two independent mechanisms: lessons from comparison of *Pichia pastoris* and mammalian cell expression. Journal of molecular biology 417, 309–335.

Schiraldi, C., Maresca, C., Catapano, A., Galinski, E.A., de Rosa, M., 2006. High-yield cultivation of *Marinococcus* M52 for production and recovery of hydroxyectoine. Research in microbiology 157, 693–699.

Schmidt, O., Pfanner, N., Meisinger, C., 2010. Mitochondrial protein import: from proteomics to functional mechanisms. Nature reviews. Molecular cell biology 11, 655–667.

Schreuder, M.P., Brekelmans, S., van den Ende, H., Klis, F.M., 1993. Targeting of a heterologous protein to the cell wall of *Saccharomyces cerevisiae*. Yeast 9, 399–409.

Schroer, K., Luef, K.P., Hartner, F.S., Glieder, A., Pscheidt, B., 2010. Engineering the *Pichia pastoris* methanol oxidation pathway for improved NADH regeneration during whole-cell biotransformation. Metabolic Engineering 12, 8–17.

Schubert, T., Maskow, T., Benndorf, D., Harms, H., Breuer, U., 2007. Continuous synthesis and excretion of the compatible solute ectoine by a transgenic, nonhalophilic bacterium. Applied and environmental microbiology 73, 3343–3347.

Seeboth, P.G., Heim, J., 1991. In-vitro processing of yeast alpha-factor leader fusion proteins using a soluble yscF (Kex2) variant. Applied microbiology and biotechnology 35, 771–776.

Seip, B., Galinski, E.A., Kurz, M., 2011. Natural and engineered hydroxyectoine production based on the *Pseudomonas stutzeri* ectABCD-ask gene cluster. Applied and environmental microbiology 77, 1368–1374.

Shi, H., Bencze, K.Z., Stemmler, T.L., Philpott, C.C., 2008. A cytosolic iron chaperone that delivers iron to ferritin. Science 320, 1207–1210.

Smith, J.D., Tang, B.C., Robinson, A.S., 2004. Protein disulfide isomerase, but not binding protein, overexpression enhances secretion of a non-disulfide-bonded protein in yeast. Biotechnology and bioengineering 85, 340–350.

Sreekrishna, K., Brankamp, R.G., Kropp, K.E., Blankenship, D.T., Tsay, J.T., Smith, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A., 1997. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene 190, 55–62.

Steinborn, G., Böer, E., Scholz, A., Tag, K., Kunze, G., Gellissen, G., 2006. Application of a wide-range yeast vector (CoMed™) system to recombinant protein production in dimorphic *Arxula adeninivorans*, methylotrophic *Hansenula polymorpha* and other yeasts. Microbial cell factories 5, 33.

Straathof, A.J.J., Panke, S., Schmid, A., 2002. The production of fine chemicals by biotransformations. Current opinion in biotechnology 13, 548–556.

Suwannarangsee, S., Kim, S., Kim, O.C., Oh, D.B., Seo, J.W., Kim, C.H., Rhee, S.K., Kang, H.A., Chulalaksananukul, W., Kwon, O., 2012. Characterization of alcohol dehydrogenase 3 of the thermotolerant methylotrophic yeast *Hansenula polymorpha*. Applied microbiology and biotechnology 96, 697–709.

Suzuki, Y., Yabuta, M., Ohsuye, K., 1998. High-level production of recombinant human parathyroid hormone 1-34. Applied and environmental microbiology 64, 526–529.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution 28, 2731–2739.

Tanaka, T., Kita, H., Murakami, T., Narita, K., 1977. Purification and amino acid sequence of mating factor from *Saccharomyces cerevisiae*. Journal of biochemistry 82, 1681–1687.

Terentiev, Y., Pico, A.H., Böer, E., Wartmann, T., Klabunde, J., Breuer, U., Babel, W., Suckow, M., Gellissen, G., Kunze, G., 2004. A wide-range integrative yeast expression vector system based on *Arxula adeninivorans*-derived elements. Journal of industrial microbiology & biotechnology 31, 223–228.

Terpe, K., 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Applied microbiology and biotechnology 60, 523–533.

Vad, R., Nafstad, E., Dahl, L.A., Gabrielsen, O.S., 2005. Engineering of a *Pichia pastoris* expression system for secretion of high amounts of intact human parathyroid hormone. Journal of biotechnology 116, 251–260.

Vanhoof, G., Goossens, F., de Meester, I., Hendriks, D., Scharpé, S., 1995. Proline motifs in peptides and their biological processing. The FASEB Journal 9, 736–744.

Varshavsky, A., 1997. The N-end rule pathway of protein degradation. Genes to cells 2, 13–28.

de Vries, B., Kiel, J.A.K.W., Scheek, R., Veenhuis, M., van der Klei, I.J., 2007. A conserved alpha helical domain at the N-terminus of Pex14p is required for PTS1 and PTS2 protein import in *Hansenula polymorpha*. FEBS letters 581, 5627–5634.

Walter, R., Simmons, W.H., Yoshimoto, T., 1980. Proline specific endo- and exopeptidases. Molecular and cellular biochemistry 30, 111–127.

Waters, M.G., Evans, E.A., Blobel, G., 1988. Prepro-alpha-factor has a cleavable signal sequence. The Journal of biological chemistry 263, 6209–6214.

Waugh, D.S., 2011. An overview of enzymatic reagents for the removal of affinity tags. Protein expression and purification 80, 283–293.

Wendel, F., 2009. Untersuchungen zur Expression von Interferon alpha in der methylotrophen Hefe *Hansenula polymorpha*. Ph.D. thesis. (Diploma thesis) Fachhochschule Aachen. Jülich.

Wendland, A., Dünkler, A., Walther, A., 2011. Characterization of alpha-factor pheromone and pheromone receptor genes of *Ashbya gossypii*. FEMS yeast research 11, 418–429.

Weydemann, U., Keup, P., Piontek, M., Strasser, A.W.M., Schweden, J., Gellissen, G., Janowicz, Z.A., 1995. High-level secretion of hirudin by *Hansenula polymorpha* - authentic processing of three different preprohirudins. Applied microbiology and biotechnology 44, 377–385.

Williams, C., Bener Aksam, E., Gunkel, K., Veenhuis, M., van der Klei, I.J., 2012. The relevance of the non-canonical PTS1 of peroxisomal catalase. Biochimica et biophysica acta 1823, 1133–1141.

Wohlfarth, A., Severin, J., Galinski, E.A., 1990. The spectrum of compatible solutes in heterotrophic halophilic eubacteria of the family *Halomondaceae*. Journal of General Microbiology 136, 705–712.

Wohlgemuth, R., 2010. Biocatalysis - key to sustainable industrial chemistry. Current opinion in biotechnology 21, 713–724.

Wong, S., Fares, M.A., Zimmermann, W., Butler, G., Wolfe, K.H., 2003. Evidence from comparative genomics for a complete sexual cycle in the 'asexual' pathogenic yeast *Candida glabrata*. Genome biology 4, R10.

van Wuytswinkel, O., Savino, G., Briat, J.F., 1995. Purification and characterization of recombinant pea-seed ferritins expressed in *Escherichia coli*: influence of N-terminus deletions on protein solubility and core formation in vitro. The Biochemical journal 305 (Pt 1), 253–261.

Xiu, Z., Li, M., Zhou, S., Dou, H., Zhou, H., Chen, C., 2002. A new method for the preparation of human parathyroid hormone 1-34 peptides. Biotechnology and applied biochemistry 36, 111–117.

Yamada, H., Shimizu, S., 1988. Microbial and enzymatic processes for the production of biologically and chemically useful compounds. Angewandte Chemie International Edition in English 27, 622–642.

Yaron, A., 1987. The role of proline in the proteolytic regulation of biologically active peptides. Biopolymers 26 Suppl, S215–22.

Yaron, A., Naider, F., 1993. Proline-dependent structural and biological properties of peptides and proteins. Critical reviews in biochemistry and molecular biology 28, 31–81.

Yurimoto, H., Oku, M., Sakai, Y., 2011. Yeast methylotrophy: metabolism, gene regulation and peroxisome homeostasis. International Journal of Microbiology 2011, 1–8.

Zsebo, K.M., Lu, H.S., Fieschko, J.C., Goldstein, L., Davis, J., Duker, K., Suggs, S.V., Lai, P.H., Bitter, G.A., 1986. Protein secretion from *Saccharomyces cerevisiae* directed by the prepro-alpha-factor leader region. The Journal of biological chemistry 261, 5858–5865.

Zurek, C., Kubis, E., Keup, P., Hörlein, D., Beunink, J., Thömmes, J., Kula, M.R., Hollenberg, C.P., Gellissen, G., 1996. Production of two aprotinin variants in *Hansenula polymorpha*. Process Biochemistry 31, 679–689.

van Zutphen, T., Baerends, R.J.S., Susanna, K.A., de Jong, A., Kuipers, O.P., Veenhuis, M., van der Klei, I.J., 2010. Adaptation of *Hansenula polymorpha* to methanol: a transcriptome analysis. BMC genomics 11, 1.

Publications

Results described in this dissertation have been submitted for publication in the following journals or have been presented at conferences:

Articles

Eilert, E., Hollenberg, C.P., Piontek, M., Suckow, M., 2012. The use of highly expressed FTH1 as carrier protein for cytosolic targeting in *Hansenula polymorpha*. Journal of Biotechnology 159(3): 172-176.

Eilert, E., Rolf, T., Heumaier, A., Hollenberg, C.P., Piontek, M.; Suckow, M, 2012. Improved processing of secretory proteins in *Hansenula polymorpha* by sequence variation near the processing site of the alpha mating factor prepro sequence. Journal of Biotechnology, In Press, http://dx.doi.org/10.1016/j.jbiotec.2012.08.024.

Eilert, E., Kranz, A., Hollenberg, C.P., Piontek, M.; Suckow, M., 2013. Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. Journal of Biotechnology, In Press, http://dx.doi.org/10.1016/j.jbiotec.2013.02.005.

Patent applications

Suckow, M., Eilert, E., Piontek, M., 2012. Verfahren zum Herstellen eines Ziel-Proteins im Verbund mit einem heterologen Ferritin-Protein unter Verwendung von Hefe-Wirtszellen. EP2484766.

Suckow, M., Eilert, E., Piontek, M., 2012. A method for producing ectoine or a derivative thereof and a yeast cell for use as a host cell in such a method. EP12196916.6, Patent application.

Poster presentation

Eilert, E., Rolf, T., Hollenberg, C.P., Piontek, M., Suckow, M., 2012. Improved processing of the alpha mating factor in *Pichia angusta* by sequence variation. Pichia 2012 Conference. Alpbach, Austria.

Danksagung

Nach einer spannenden Zeit ist es nun daran einigen Leuten zu danken, die nicht unwesentlich dazu beigetragen haben, dass diese Dissertation so nun zu Stande gekommen ist, und mir mit Rat und Tat zur Seite gestanden haben.

Ich danke...

... , allen vorran, besonders Prof. Dr. Cornelis P. Hollenberg für die Übernahme des Referats, die stete Diskussionsbereitschaft, die vielen konstruktiven Anregungen sowie die ausgezeichnete Betreuung.

... ebenfalls besonders Prof. Dr. Joachim F. Ernst für die Übernahme des Referats und Prof. Dr. Michael Bott für die Übernahme des Korreferats.

... desweiteren der Firma ARTES Biotechnology GmbH, im Besonderen Dr. Michael Piontek, für die Überlassung des Themas und die Möglichkeit die Arbeiten in den dortigen Laboren durchführen zu können. Ich möchte mich hierbei besonders bei Dr. Manfred Suckow für die Betreuung sowie für die interessanten Gespräche, spannenden Gedankenexperimente und stete Diskussionsbereitschaft bedanken.

... all meinen Kollegen bei ARTES, die mich mehr oder weniger lang während meiner Promotionszeit begleitet haben. Danke für die gute Zusammenarbeit und die angenehme Arbeitsatmosphäre. Danke auch für all die lustigen Zeiten und euer jederzeit offenes Ohr. Insbesondere möchte ich mich bei Nina Wingbermühle, Manuela Bußmann, Theresa Rolf, Andreas Kranz, Sylvia Denter, Michael Weniger, David Wilson, Patrick Glaser und Renske Klassen für die technische Unterstützung sowie die sehr gute Zusammenarbeit bedanken.

... Andreas Heumaier, für Hilfestellung zur Modellierung und Analyse von Transkriptdatenbanken.

... unseren Kooperationspartner im CLIB 2012 'BioCat' Projekt, bitop AG und m2p labs GmbH, für die vielen guten Ratschläge, insbesondere Dr. Markus Neumann für HPLC Messungen zum Nachweis von 5-Hydroxyectoine, und Oliver Born für Kultivierungsexperimente.

... allen Beteiligten des CLIB Graduate Cluster Industrial Biotechnologie für die Aufnahme als Kollegiatin, insbesondere Dr. Sonja Meyer zu Berstenhorst für ihr großes Engagement für die Doktoranden.

... meine Herzallerliebsten: Ina, Alex, Anna und Nina. Danke, dass ihr so seid wie ihr seid. Gute Freunde muss man festhalten.

... von Herzen danke ich meinen Eltern und meiner Schwester für die Unterstützungen jeglicher Art - sei es vor Ort oder aus der Ferne.

Erklärung zur Promotion

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfsmittel angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Ort, Datum Eva Eilert