

# The secretome of *Pichia pastoris* in fed-batch cultivations is largely independent of the carbon source but changes quantitatively over cultivation time

Jonas Burgard,<sup>1,2,†</sup> Clemens Grünwald-Gruber,<sup>1,3</sup> Friedrich Altmann,<sup>1,3</sup> Jürgen Zanghellini,<sup>1,2,4</sup>  Minoska Valli,<sup>1,2</sup> Diethard Mattanovich<sup>1,2,\*</sup>  and Brigitte Gasser<sup>1,2</sup>

<sup>1</sup>Austrian Centre of Industrial Biotechnology (acib), Vienna, Austria.

<sup>2</sup>Department of Biotechnology, BOKU-University of Natural Resources and Life Sciences, Vienna, Austria.

<sup>3</sup>Department of Chemistry, BOKU-University of Natural Resources and Life Sciences, Vienna, Austria.

<sup>4</sup>Austrian Biotech University of Applied Sciences, Tulln, Austria.

## Summary

The quantitative changes of the secretome of recombinant *Pichia pastoris* (*Komagataella phaffii*) CBS7435 over the time-course of methanol- or glucose-limited fed-batch cultures were investigated by LC-ESI-MS/MS to define the carbon source-specific secretomes under controlled bioreactor conditions. In both setups, no indication for elevated cell lysis was found. The quantitative data revealed that intact and viable *P. pastoris* cells secrete only a low number of endogenous proteins (in total 51), even during high cell

density cultivation. Interestingly, no marked differences in the functional composition of the *P. pastoris* secretome between methanol- and glucose-grown cultures were observed with only few proteins being specifically affected by the carbon source. The ‘core secretome’ of 22 proteins present in all analysed carbon sources (glycerol, glucose and methanol) consists mainly of cell wall proteins. The quantitative analysis additionally revealed that most secretome proteins were already present after the batch phase, and depletion rather than accumulation occurred during the fed-batch processes. Among the changes over cultivation time, the depletion of both the extracellularly detected chaperones and the only two identified proteases (Pep4 and Yps1-1) during the methanol- or glucose-feed phase appear as most prominent.

## Introduction

*Pichia pastoris* (syn. *Komagataella* spp.) is a methylotrophic yeast, which is commonly used as host platform for the production of recombinant secretory proteins of pharmaceutical (Weinacker, *et al.*, 2013) and industrial interest (Rabert, *et al.*, 2013; Love, *et al.*, 2017; Zhu, *et al.*, 2019). *P. pastoris* has been reclassified to the genus *Komagataella* in the recent years, which currently consists of seven species (including *K. pastoris* and *K. phaffii*). It is able to metabolize several carbon sources, including methanol, glucose and glycerol. In industrial processes, both the methanol-inducible promoter of alcohol oxidase 1 ( $P_{AOX1}$ ) and the constitutive glycolytic glyceraldehyde-3-phosphate dehydrogenase promoter ( $P_{GAP}$ ) are routinely applied and were proven to be suitable for obtaining high production levels of different protein classes up to the  $g\ l^{-1}$  scale (Clare, *et al.*, 1991; Yu, *et al.*, 2003; Damasceno, *et al.*, 2004; Advitiya, *et al.*, 2016). One advantage for secretory production is that *P. pastoris* has been shown to secrete a relatively low number of endogenous proteins compared to other hosts, which simplifies downstream processing and purification of the recombinant product. However, there are discrepancies regarding the actual number of natively secreted proteins. The secretome of *P. pastoris* was investigated in glucose-limited chemostat cultivations of the *K. pastoris* wild-type strain DSMZ70382 at a specific growth rate of  $0.1\ h^{-1}$  (Mattanovich, *et al.*,

Received 17 January, 2019; accepted 7 October, 2019.

\*For correspondence. Email Diethard.mattanovich@boku.ac.at; Tel. +43 1 47654-79042; Fax +43 1 47654-79009.

†Present address: Octapharma Pharmazeutika Produktionsgesellschaft m.b.H, Vienna, Austria.

*Microbial Biotechnology* (2020) 13(2), 479–494  
doi:10.1111/1751-7915.13499

## Funding Information

This work has been supported by the Austrian Federal Ministry of Digital and Economic Affairs (bmdw), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol, the Government of Lower Austria and ZIT—Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG. Further support by BIOMIN Research Center, Boehringer Ingelheim RCV GmbH & Co KG, Lonza AG, Biocrates Life Sciences AG, Validogen (formerly VTU Technology) GmbH, and Sandoz GmbH is acknowledged. EQ BOKU VIBT GmbH is acknowledged for providing fermentation equipment, LC-MS/MS and flow cytometry instrumentation through the VIBT-Extremophile Center and the BOKU Core Facilities Cellular Analysis and Biomolecular & Cellular Analysis (BmCA).

[Correction added on 22 November 2019, after first online publication: The name of the author Clemens Grünwald-Gruber has been corrected in this version.]

© 2019 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

2009) as well as during glycerol batch and methanol induction of a recombinant schistosomiasis vaccine producing *K. phaffii* strain and its non-producing parent strain X-33 (Huang, *et al.*, 2011). Thereby, 20 and 75 proteins were found in the supernatants, respectively. While the lower number of identified proteins in the chemostat cultivation might be attributed to the continuous dilution of the culture, the latter study assumed that cell lysis of methanol-grown cells contributed to the higher number of detected proteins (Huang, *et al.*, 2011), compared to the < 1% cell lysis that was determined for the wild type cells grown in the glucose chemostat (Mattanovich, *et al.*, 2009). Reduced culture viability resulting in degradation of the recombinant product by the release of intracellular proteases upon cell lysis was reported as one major disadvantage of methanol-based processes (Sinha, *et al.*, 2005; Xiao, *et al.*, 2006). It was often traced back to the main vacuolar protease Pep4; nevertheless, there are several cases where the use of a *pep4* knockout strain could not circumvent product degradation (reviewed by Puxbaum, *et al.*, 2015; Zahrl, *et al.*, 2019). Identification of the responsible proteases would thus be of interest to prevent product loss.

Furthermore, it remains unclear if the differences in the reported secretomes are due to the different *P. pastoris* strains, the different cell densities, the different carbon sources or the different modes of cultivation applied. Moreover, no quantitative data on the transitions of the secretome over the course of a whole cultivation process are available so far.

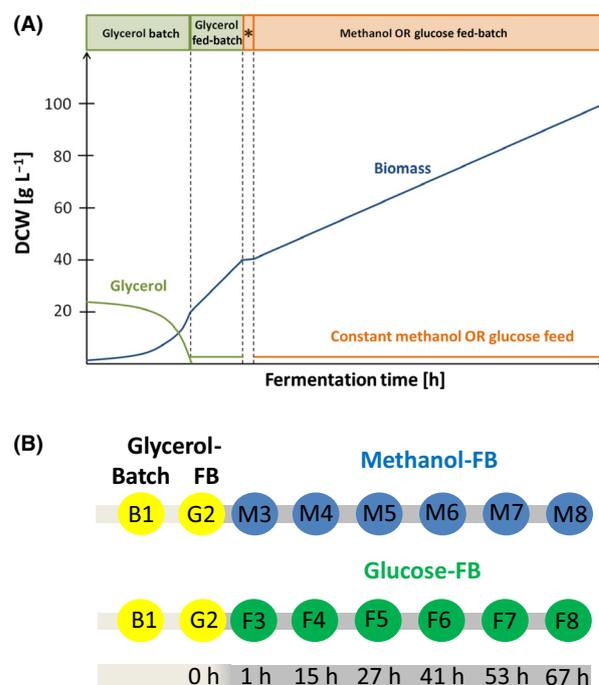
To close this gap, we analysed a *P. pastoris* strain (*K. phaffii* CBS7435) producing porcine pro-carboxypeptidase B (CpB) under control of  $P_{GAP}$  over the time-course of glucose- and methanol-limited fed-batch cultivations, respectively. As  $P_{GAP}$  is about equally active on glucose and methanol, strains producing the same recombinant protein under control of  $P_{GAP}$  are ideal tools to study whether the different substrates have a main influence on the size and composition of the secretome.

## Results

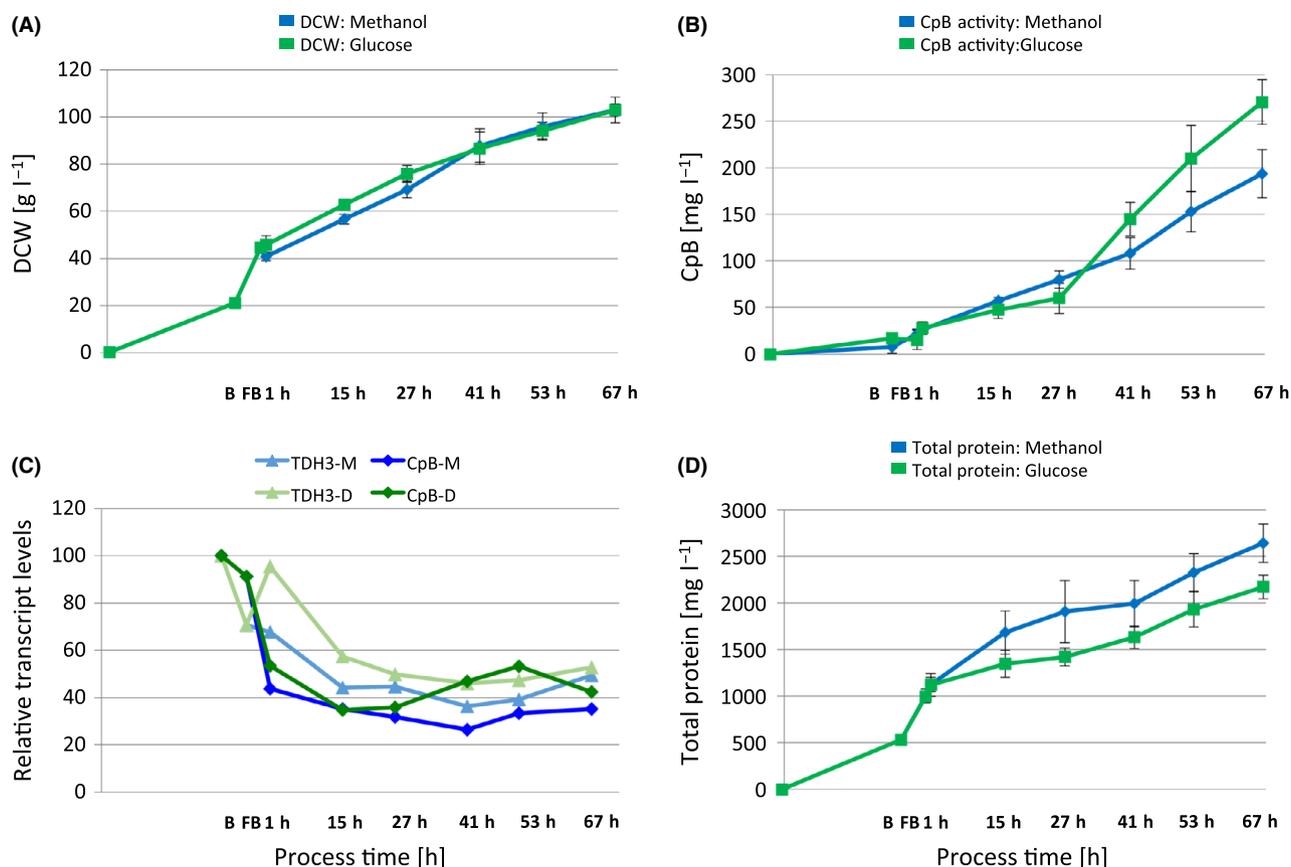
### High cell density fed-batch does not lead to extensive cell lysis and proteolytic activity in culture supernatants

The aim of the present study was to compare the secretomes of glucose- and methanol-grown recombinant *P. pastoris* in carbon-limited high cell density fed-batch cultivations and to quantitatively analyse the abundance of the proteins in the supernatant over time. To focus on the influence of the substrate, we decided to use the same *P. pastoris* strain producing porcine pro-carboxypeptidase under control of  $P_{GAP}$  and apply a feed strategy designed for obtaining equal specific growth rates and biomass development in both set-ups (Burgard, *et al.*, 2017). In

order to achieve comparable process conditions, the feed profile of the glucose-based process was intended to closely resemble our standard methanol fed-batch protocol, considering different biomass on substrate yield coefficients for both carbon sources (Hohenblum, *et al.*, 2004; Burgard, *et al.*, 2017). It included 5 h of a common glycerol-limited fed-batch before changing the carbon source to start the methanol- or glucose-limited feed respectively (Fig. 1). Both processes were performed in triplicates and characterized for growth, recombinant protein production and changes in the secretome over the course of the whole process. As contrived, on both carbon sources the processes had highly comparable specific growth rates and biomass development (Fig. 2A). The dry cell weight (DCW) reached  $21 \pm 1 \text{ g l}^{-1}$  after the glycerol batch and  $44 \pm 2 \text{ g l}^{-1}$  after the glycerol fed-batch. At the end of the fed-batch (67 h), methanol-grown cultures reached  $103 \pm 5 \text{ g l}^{-1}$  DCW and glucose-grown cultures  $103 \pm 3 \text{ g l}^{-1}$ . During the processes, the specific growth



**Fig. 1.** Schematic representation of the methanol- or glucose-based fed-batch (FB) cultivations for secretome analysis of *P. pastoris*. A. The processes comprised a glycerol batch phase, followed by a glycerol-limited fed-batch for 5 h. Then, methanol-grown cultures received a methanol pulse, while glucose-grown cultures were slowly fed with glucose to accumulate the same amount of biomass during this period (\*). Afterwards, a constant methanol or glucose feed was applied for 67 h. See material and methods section for details of media composition and feed rates. Fed-batch cultivations were performed in triplicates. B. Sampling scheme for the analysis of the methanol and glucose fed-batch processes. Samples were taken at the end of the glycerol batch phase (B1), the glycerol fed-batch phase (G2), and at six time points during the methanol- or glucose-limited fed-batch phase (M3-M8 and F3-F8 respectively).



**Fig. 2.** Biomass generation and amounts of secreted recombinant and native proteins in methanol- and glucose-limited fed-batch (FB) cultivations of the *P. pastoris* strain producing CpB under control of  $P_{GAP}$ .

A. Biomass generation.

B. Secreted CpB concentrations determined by enzyme activity assay. Average values and standard deviations of the three biological replicates per cultivation condition are shown.

C. Relative transcript levels of *TDH3* and CpB in glucose- (D) or methanol-grown cultures (M). Normalized mean expression data from microarray analysis of the three biological replicates per cultivation condition are shown (data taken from Burgard *et al.* (2017)). All data points are relative to the batch condition.

D. Total secreted protein content in the supernatants during the fed-batch processes.

rate decreased from 0.2 to 0.01 h<sup>-1</sup> upon the increasing carbon limitation.

CpB activity in the culture supernatants constantly increased during the fed-batch (FB) processes (Fig. 2B). As expected, final CpB product titers were in a similar range, reaching 271 ± 24 mg l<sup>-1</sup> in glucose-grown cultures at the end of the FB cultivation, and 194 ± 26 mg l<sup>-1</sup> in methanol-grown cultures. The approximately 30% lower CpB levels in the methanol FB can be attributed to CpB transcript levels (microarray data from Burgard *et al.* (2017)), which were on average 20–30% lower during the last four sampling time points of the methanol-limited FB in comparison with the glucose-limited FB, similarly to the transcript levels of the native glyceraldehyde-3-phosphate dehydrogenase gene *TDH3* (Fig. 2C). The total protein content in culture supernatants was shown to increase over the time-course of the fed-batch cultivation. In contrast to the recombinant product levels, the total extracellular

protein content (Fig. 2D) was approximately 20% higher in methanol-grown cultures (2644 ± 207 µg ml<sup>-1</sup>) than in glucose-grown cultures (2174 ± 128 µg ml<sup>-1</sup>) at the end of the cultivation (67 h).

In order to differentiate whether these differences were due to active secretion or due to cell lysis, cellular viability and DNA content in the supernatant were determined. Viability remained high throughout the processes. At the end of the fed-batch, the cultures grown on glucose had slightly higher viability (> 99%) than the methanol-grown cultures (> 96%). Host cell DNA (hcDNA) in culture supernatants was used as a measure of cell lysis, as described in Matanovich *et al.* (2009). Culture supernatants of all sampling points and biological replicates were analysed using the Qubit fluorescence assay specific for DNA, and due to assay sensitivity, quantitative values could only be obtained at the end of the fed-batch processes. Corresponding to the viability measurements, hcDNA levels in supernatants of

methanol-grown cultures ( $0.23 \pm 0.05 \mu\text{g ml}^{-1}$ ) were higher than in glucose-grown cultures ( $0.16 \pm 0.01 \mu\text{g ml}^{-1}$ ), but this difference was only marginal. In both cases, the quantified hcDNA in culture supernatants was equivalent to cell lysis of < 0.2% (methanol) or 0.1% (glucose) of DCW accounting for maximally 4% of total measured protein in the supernatants according to Carnicer *et al.* (2009). Proteolytic activity in culture supernatants remained at a low level over the whole process and was rather similar at the end of the methanol and glucose fed-batch processes (Table 1). It was also proven that autoactivated CpB did not contribute to the measured proteolytic activity in the culture supernatants.

*The P. pastoris secretome contains only a low number of proteins even in high cell density cultures and has a similar size on methanol and glucose as substrate*

For qualitative analysis of proteins in culture supernatants, SDS-PAGE was performed from samples taken throughout the whole fed-batch process. The recombinant product CpB was identified as the dominant band at a molecular weight of 46 kDa (proCpB) in the Coomassie-stained SDS-PAGE (Fig. 3A) and confirmed by Western blot analysis (Fig. 3B). The band patterns were very similar between supernatants of methanol- and glucose-grown cultures and comprised a low number of different bands. Only one distinct band below the recombinant product became visible in samples at the end of the cultivation, which was identified as autoactivated CpB (CpB: 35 kDa). To get a more detailed picture of the secretome, proteomics analyses of culture supernatants were performed using LC-ESI-MS/MS to investigate the secretome over the time-course of the methanol- and glucose-based production processes. Therefore, supernatants of the glycerol batch, glycerol fed-batch and three time points (1, 27 and 67 h) during the fed-batch processes were analysed, which have been shown to be representative for early, middle and late fed-batch before (Burgard, *et al.*, 2017). Proteins were identified by MASCOT or MaxQuant and were considered to be present at a sampling point when they were found in at least two of the analysed biological replicates. In total, 51 different *P. pastoris* proteins were detected in the supernatants. In

**Table 1.** Proteolytic activity in culture supernatants over the course of the methanol and glucose fed-batch processes.

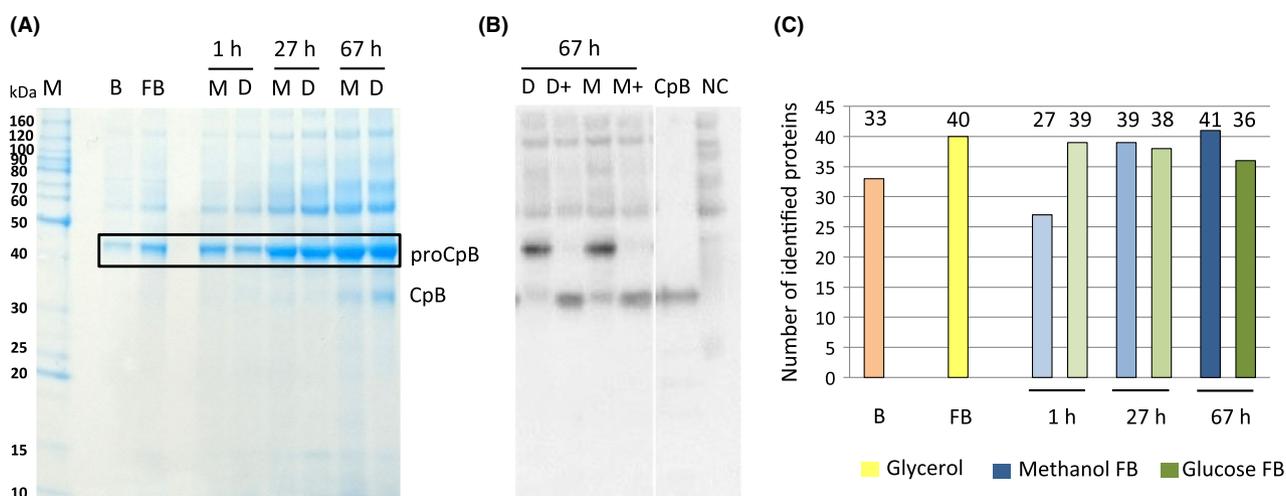
	Proteolytic activity [ $\mu\text{g ml}^{-1}$ ]		
	Glycerol	Methanol	Glucose
Glycerol batch	$0.04 \pm 0.009$		
Glycerol fed-batch	$0.05 \pm 0.01$		
1 h fed-batch		$0.04 \pm 0.003$	$0.05 \pm 0.002$
27 h fed-batch		$0.07 \pm 0.005$	$0.08 \pm 0.007$
67 h fed-batch		$0.07 \pm 0.004$	$0.10 \pm 0.005$

the glycerol batch 33 and in glycerol fed-batch phase 40 secreted proteins were identified (Fig. 3C). Unexpectedly, the number of identified secreted proteins did not increase during the fed-batches. At the end of the methanol-grown culture, 41 proteins were identified, while on glucose a maximum of 39 proteins were found 1 h after the feed start, which slightly decreased to 36 until the end of the fed-batch (Fig. 3C). A summary of all identified proteins is given in Table 2, while the underlying data such as MASCOT scores and identified peptides can be found as supporting information (Appendix S1 and S2). The quantitative results of the individual proteins obtained from MaxQuant are given in Appendix S3: Table S1.

As can be seen in Table 2, more than 70% of the identified proteins are predicted to contain a signal peptide (36/51), which makes them likely to reach the supernatant by secretion and not to be released upon cell lysis during cultivation. Approximately half of them (17/36) also contain a predicted GPI anchoring site. In accordance with previous publications (Mattanovich, *et al.*, 2009; Huang, *et al.*, 2011), 'cell wall and cell wall-associated proteins' such as glucanases, glucanosyl-transferases and chitin-modifying enzymes were highly enriched among the identified secretory proteins. Also, flocculins (members of the cell surface adhesion protein family) were detected in the secretome, as were the secretory proteins putatively involved in sterol transport (Pry2, Epx1). Additionally, five endoplasmic reticulum (ER)-resident folding factors, the chaperone Kar2, its co-chaperone Lhs1, protein disulphide isomerase Pdi1 and the two peptidyl-prolyl isomerases Cpr5 and Fpr2 were found. Similar to the previous studies, only two proteins with protease domains were detected in the supernatant, the main vacuolar peptidase Pep4 and the yapsin Yps1-1, a GPI-anchored aspartic protease. Sixteen proteins found in the supernatants of some of the samples were without predicted signal peptide. These are mainly proteins involved in metabolism or associated with ribosomes and translation. According to the quantitative data (Appendix S3: Table S1), they are among the low abundant proteins in supernatants but highly transcribed, and highly abundant in the cytosol, making it possible that they are just contaminants due to cell lysis. Alternatively, they might reach the supernatant as cargo of so-called extracellular vesicles (Brown, *et al.*, 2015; Peres da Silva, *et al.*, 2015). Overall, the functional distribution of the identified proteins is very similar in supernatants of methanol- or glucose-grown cultures with some exceptions.

*The 'core secretome' of P. pastoris cultivated in glycerol, glucose or methanol fed-batch is enriched for cell wall proteins*

Twenty-two proteins were present in all analysed conditions and were thus defined as the 'core secretome'.



**Fig. 3.** Analysis of proteins present in the culture supernatant.

A. Exemplary SDS-PAGE of culture supernatants of the glycerol batch phase (B), glycerol fed-batch (FB) and different time points of the glucose (D) and methanol (M) fed-batch processes (1, 27 and 67 h; A).

B. Anti-CpB Western Blot of cultures (1:10 diluted culture supernatants, reduced and denatured), probed with monoclonal HRP-conjugated anti-CpB antibody (1:1500) and detected by chemiluminescence; + indicates that samples were treated with trypsin to activate pro-CpB to mature CpB.

C. Number of identified proteins in culture supernatants of the corresponding samples shown in A.

Proteins of the core secretome are marked with 'core' in Table 2.

The functional distribution, based on GO-term analysis, shows that the core secretome is enriched for cell wall- or cell wall-associated proteins such as glucanases or chitinases. Glucanases and chitinases are required for cell wall plasticity during cell growth and division (Adams, 2004). Another prominent protein fraction belonged to the group of flocculins, which can induce filamentous growth, cell aggregation or cell adhesion on surfaces. Especially Flo5-1 and Flo5-2 were present in all analysed conditions, while Flo11, which is described as the major flocculin in *Saccharomyces cerevisiae*, was only found in the later stages of the fed-batch.

#### *Changes of the secretome during the fed-batch are not carbon source-dependent*

For identification of process dependent changes in the composition of the secretome, UpSet plots (Lex *et al.*, 2014) were drawn (protein lists for each intersection can be found in Appendix S3: Table S2). The composition was analysed after changing the carbon source (1 h), in the middle of the fed-batch (27 h) and at the end of the fed-batch (67 h), and compared to the glycerol fed-batch as reference (Fig. 4A). The majority of proteins were already present in the supernatant after the glycerol batch or fed-batch phase, and there is only a low number of carbon source-specific secretory proteins that only appear in the fed-batches (in total 7 proteins, Fig. 4B). These include the chitin-modifying enzymes Cda2 and

Crh1, the flocculin Flo11 and the uncharacterized protein PP7435\_Ch3-1225. From the early FB to the end of the fed-batch, we observed that the intersection representing the shared secretome of all conditions slightly increased (from 27 to 32 proteins) and the glucose-specific section decreased (from 3 to 0), whereas the methanol-specific section rose from 0 to 3 proteins. Interestingly, the similarity between methanol and glycerol or methanol and glucose slightly increased, whereas it decreased between glucose and glycerol. Noticeably, some of the ER proteins and chaperones (Lhs1, Cpr5, Fpr2) that were already present in the supernatants of the glycerol batch and fed-batch got depleted over the course of the methanol or glucose fed-batch (still present at 1 h of feed, but not at 27 h and/or 67 h).

#### *Changes in secretome abundance over the course of the fed-batches by absolute quantification of proteins using MaxQuant*

In order to investigate whether the secretome protein levels remain constant or show changes in abundance over the time-course of the fed-batch cultivations, label-free quantification of the identified proteins of the MS-analysis was performed using MaxQuant. Based on the MaxQuant iBAQ-values, highly abundant proteins (Appendix S3: Table S3) in the methanol or glucose fed-batch and the changes over the respective process time compared to the glycerol fed-batch (strongest accumulation/depletion, Appendix S3: Tables S3 and S4) were investigated. In this context, it became visible that the



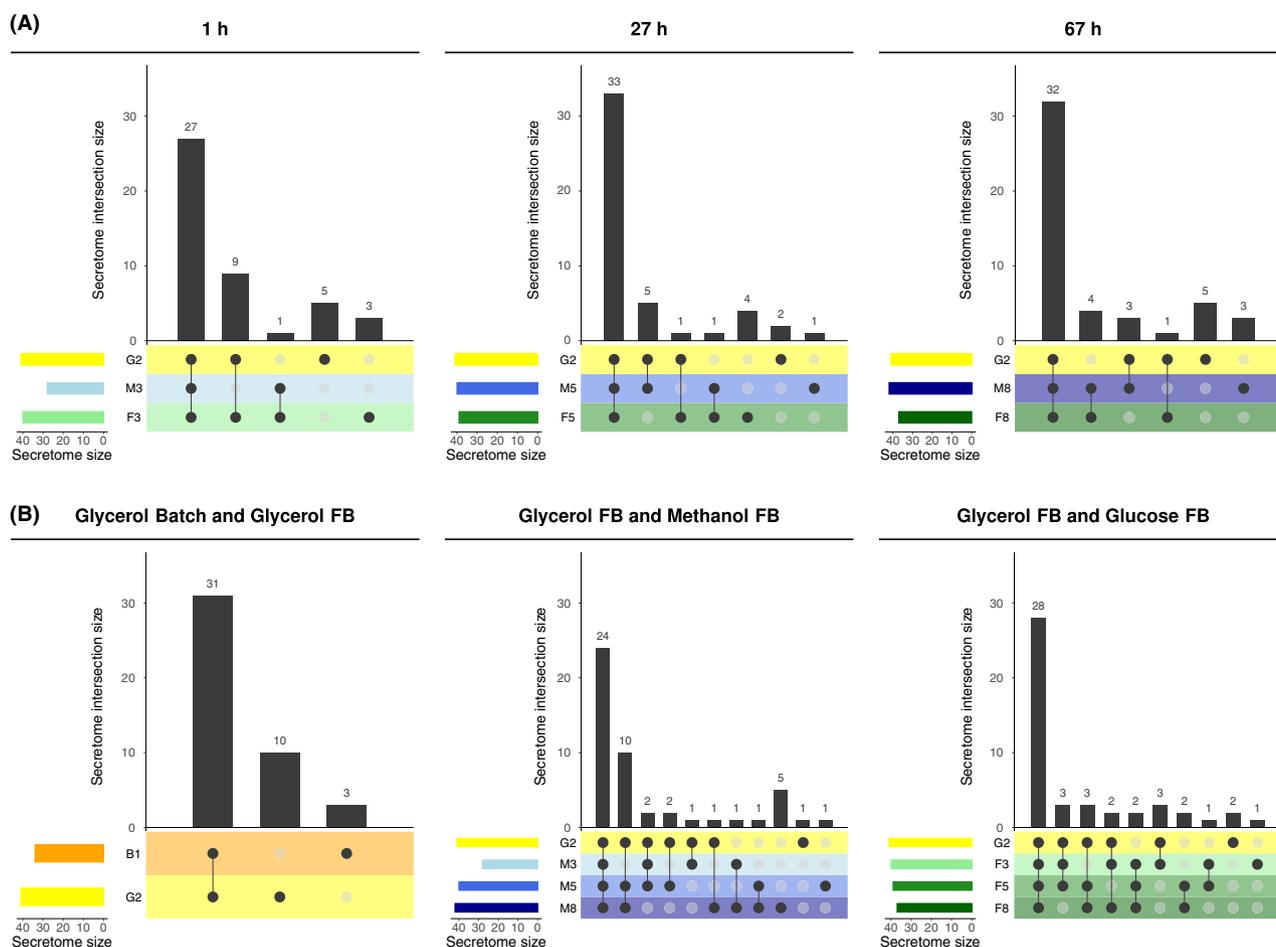
Table 2. (Continued)

Short Name <i>P. pastoris</i>	Core proteome	Accession	Glycerol batch/ fed batch					Methanol fed batch					Glucose fed batch					Description	Predicted protein parameters			
			B1	G2	M3	M5	M8	F3	F5	F8	F8	F5	F3	M8	M5	M3	B1		Protein function	MW/pI	Signal peptide	GPI anchor
Gcs1		F2QUT9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	ADP-ribosylation factor GTPase activating protein (ARF GAP), involved in ER-Golgi transport	4.0/5.8	N		nuc1: 15.5, cyto_nuc1: 11.5	
Hsp12		F2QYN4	1	1	1	1	1	1	1	1	1	1	1	1	1	Plasma membrane protein involved in maintaining membrane organization in stress conditions	12.5/4.8	N		nuc1: 14, cyto_nuc1: 9.8		
Hyp2		F2QUL4	1	1	1	1	1	1	1	1	1	1	1	1	1	Translation elongation factor eIF-5A	17.0/5.0	N		cyto: 14.5, cyto_nuc1: 14.5, extr: 15, E.R.: 9		
Kar2	core	F2QTW5	1	1	1	1	1	1	1	1	1	1	1	1	1	Molecular chaperone of the endoplasmic reticulum lumen	7.8/4.7	Y		extr: 1, E.R.: 8		
Lhs1		F2QLU	1	1	1	1	1	1	1	1	1	1	1	1	1	Molecular chaperone of the endoplasmic reticulum lumen	99.5/5.0	Y				
Msb2		F2QNL1	1	1	1	1	1	1	1	1	1	1	1	1	1	Mucin family member involved in signalling	86.5/3.8	Y	Highly probable	extr: 1, cyto: 8		
Pdi1	core	F2QY5	1	1	1	1	1	1	1	1	1	1	1	1	1	Protein disulphide isomerase.	55.4/4.6	Y		cyto: 11, E.R.: 7		
Pep4	core	F2QUG8	1	1	1	1	1	1	1	1	1	1	1	1	1	multifunctional protein of ER lumen	42.0/4.7	Y		extr: 13, cyto: 8.5		
Pir1	core	F2QS11	1	1	1	1	1	1	1	1	1	1	1	1	1	Vacuolar aspartyl protease (proteinase A)	27.5/4.2	Y		extr: 22, cyto: 3.5		
Pir2	core	F2QZM1	1	1	1	1	1	1	1	1	1	1	1	1	1	Cell wall protein with internal repeats	31.7/4.4	Y		extr: 27		
PP7435_Chr2-752	core	F2QSQ9	1	1	1	1	1	1	1	1	1	1	1	1	1	Hypothetical protein not conserved (domain: syndecan)	23.7/6.0	Y	Highly probable	extr: 25, mito: 1		
PP7435_Chr3-1213	core	F2QXM5	1	1	1	1	1	1	1	1	1	1	1	1	1	Hypothetical protein not conserved (domain: Bacterial adhesins)	6.9/4.3	Y	Highly probable	extr: 18, E.R.: 3		
PP7435_Chr3-1225		F2QXN7	1	1	1	1	1	1	1	1	1	1	1	1	1	Hypothetical protein conserved (domain: PASTA, Colicin E3 immunity protein)	23.2/4.1	Y		extr: 26		
PP7435_Chr4-69		F2QZL9	1	1	1	1	1	1	1	1	1	1	1	1	1	Protein with similarity to YMR244W (Putative protein of unknown function)	36.6/4.9	Y	Highly probable	extr: 24, mito: 1		
Pry2	core	F2QNB4	1	1	1	1	1	1	1	1	1	1	1	1	1	Putative sterol binding protein involved in the export of acetylated sterols	29.3/4.1	Y	Highly probable	extr: 22, mito: 2		
Pst1		F2QR22	1	1	1	1	1	1	1	1	1	1	1	1	1	Cell wall protein that contains a putative GPI-attachment site, secreted by regenerating protoplasts	41.5/3.8	Y	Highly probable	extr: 26		
Rce3	core	F2QYL8	1	1	1	1	1	1	1	1	1	1	1	1	1	Endo-glucanase similar to <i>R. oryzae</i> RCE3 with 3 fungal cellulose binding domains	62.0/4.0	Y		extr: 26		
Rpl4A		F2QSW4	1	1	1	1	1	1	1	1	1	1	1	1	1	Ubiquitin-ribosomal 6S subunit protein L4A fusion protein	14.6/9.9	N		nuc1: 18, cyto_nuc1: 16, mito: 16, nuc1: 5		
Rps17B		F2QS1	1	1	1	1	1	1	1	1	1	1	1	1	1	Ribosomal protein 51 (rp51) of the small (4s) subunit	15.7/1.3	N				
Scw1	core	F2QNG1	1	1	1	1	1	1	1	1	1	1	1	1	1	Cell wall protein with similarity to glucanases	34.0/4.9	Y	Highly probable	extr: 24, cyto: 2		
Scw11	core	F2QU52	1	1	1	1	1	1	1	1	1	1	1	1	1	Cell wall protein with similarity to glucanases	46.7/4.4	Y		extr: 22, cyto: 2.5		

Table 2. (Continued)

Short Name <i>P. pastoris</i>	Core proteome	Accession	Glycerol batch/ fed batch				Methanol fed batch				Glucose fed batch				Description	Predicted protein parameters		
			B1	G2	M3	M5	M8	F3	F5	F8	F3	F5	F8	F3		F5	F8	MW/pl
Sod1		F2QY66	1			1	1	1	1	1	1	1	1	1	1	15.7/5.9	N	cyto: 23.5, cyto_nucl: 14 extr: 11, E.R.: 4 extr: 26
Ssp12		F2RE9	1		1	1	1	1	1	1	1	1	1	1	1	25.7/4.8	Y	extr: 11, E.R.: 4 extr: 26
Sun4	core	F2QQT7	1		1	1	1	1	1	1	1	1	1	1	1	43.2/4.3	Y	
Tdh3		F2QT12	1		1	1	1	1	1	1	1	1	1	1	1	35.6/6.2	N	pero: 16, cyto: 8
Tef2		F2QQG8	1		1	1	1	1	1	1	1	1	1	1	1	5.1/9.1	N	cyto: 26.5, cyto_nucl: 14 extr: 26
Tos1	core	F2QQJ	1		1	1	1	1	1	1	1	1	1	1	1	43.1/5.0	Y	
Ubp11		F2QNR5	1		1	1	1	1	1	1	1	1	1	1	1	18.8/8.0	N	nucl: 2, cyto: 5
Ufh1	core	F2RE1	1		1	1	1	1	1	1	1	1	1	1	1	34.0/4.8	Y	extr: 26
Ulr2		F2QP19	1		1	1	1	1	1	1	1	1	1	1	1	48.8/3.9	Y	Highly probable extr: 23, golg: 2
Vps72		F2QX78	1		1	1	1	1	1	1	1	1	1	1	1	84.0/4.9	N	nucl: 21, cyto: 5
Ynk1		F2QQT	1		1	1	1	1	1	1	1	1	1	1	1	17.0/6.1	N	cyto: 19.5, cyto_nucl: 13.3 extr: 24, mito: 1
Yps1-1	core	F2QYS7	1		1	1	1	1	1	1	1	1	1	1	1	6.5/4.4	Y	

Molecular weight/isoelectric point predicted by ExPASy ProtParam.  
 Predicted by Signal P4.1 – Y; signal peptide predicted, N, not predicted  
 Predicted by PredGPI.  
 Predicted by Wolf-PSORT.

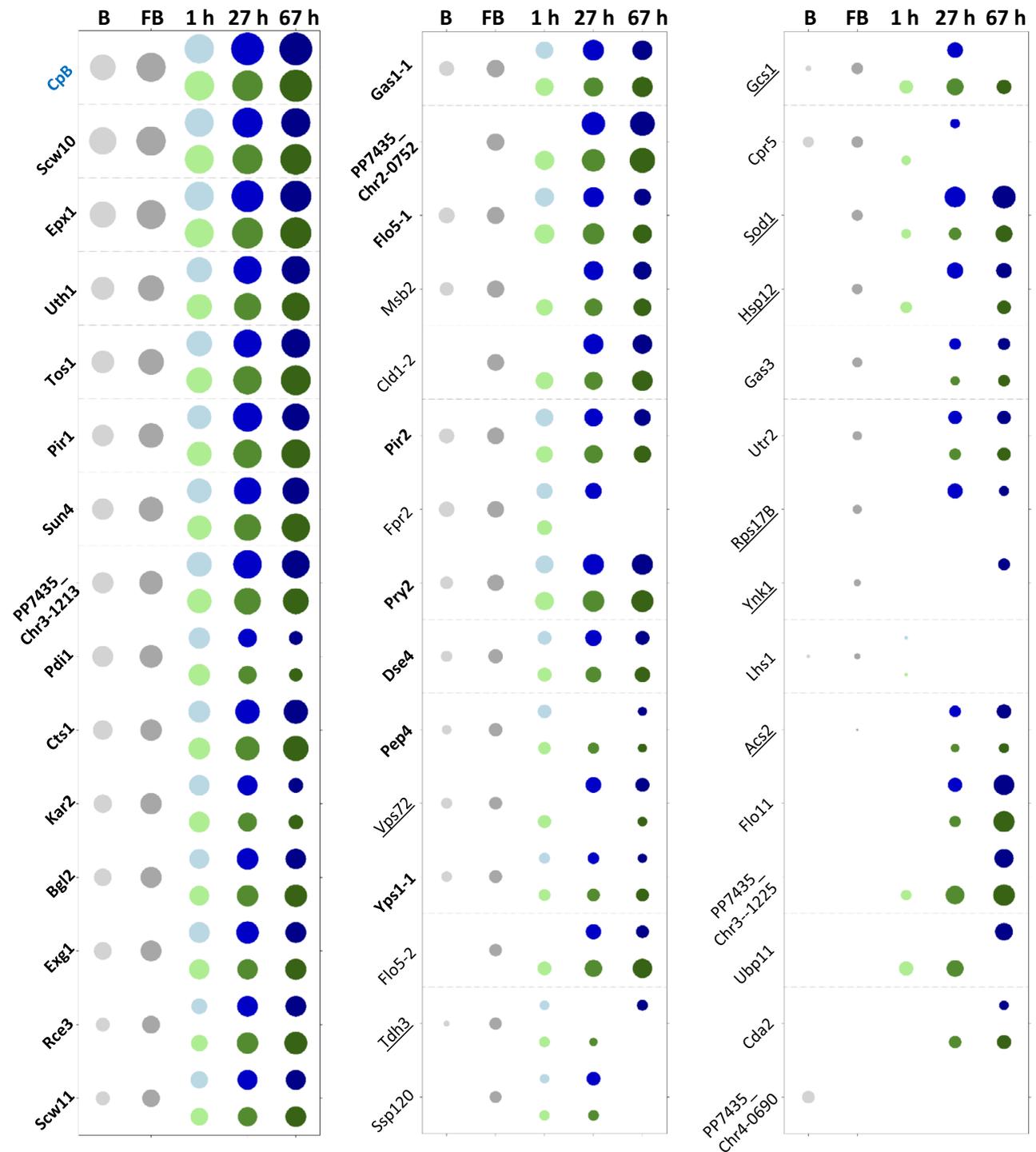


**Fig. 4.** UpSet plots (Lex and Gehlenborg, 2014) visualizing the intersections of the secretome (numbers of proteins present in the secretome) across different conditions as function of time (panel A) and the intersections of the secretome across time as function of the cultivation process (panel B). Samples are labelled as in Figure 1 (glycerol batch and fed-batch (FB) and the early (1 h), middle (27 h) and late stage (67 h) of the methanol and glucose FB processes). In each subfigure, intersections are displayed in a matrix layout. Each column corresponds to a specific section in a traditional Venn diagram. Each row corresponds to the secretome at the indicated time and condition. The number of proteins in the respective secretome is indicated by horizontal bars. Connected, dark circles in the matrix indicate secretomes that share the same proteins. The numbers of shared proteins in each intersection are visualized by vertical bars. Empty intersections are not displayed. For instance, the first column in the leftmost subfigure of panel A indicates that the same 27 proteins are expressed at the end of the glycerol FB and after the first hour of the methanol and glucose FB. Please note that the proteins that are identified in the glycerol FB condition is the same for all diagrams of panel A, just the number of proteins in the intersections with the glucose FB and/or methanol FB differ.

recombinant product represented the largest fraction of the proteins in the supernatants and only very few others could be detected in a similar abundance. The uncharacterized extracellular protein Epx1 (Heiss, *et al.*, 2013) was clearly verified as the most abundant native secreted protein in both carbon sources, followed by the cell wall protein Scw10, which had the highest abundance in the glycerol batch and fed-batch (Fig. 5 and Appendix S3: Table S1). Besides, other rather abundant proteins were cell wall- or cell wall-associated proteins (flocculins, chitins and glucanases). While the ER chaperones Kar2 and Pdi1 were highly abundant in the glycerol batch and fed-batch, they gradually decreased over the course of the fed-batch. The group of proteins with low abundance in the supernatant differed between the

individual conditions, and as stated before, comprised of many of the proteins without a signal peptide. Three proteins were only identified in the glycerol batch sample (Atp2, Pst1 and PP7435\_Ch4-0690/YMR244W), which are all proteins of low abundance.

One could assume that proteins generally accumulate over the course of fed-batch processes. In contrast, our quantitative secretome analysis shows that some proteins accumulate but others get strongly depleted during the process compared to the glycerol fed-batch. In the methanol-based process, the strongest accumulation over time (slope, Appendix S3: Table S4) was observed for PP7435\_Ch2-0752, PP7435\_Ch3-1213, Tos1, Sod1 and Pir1, whereas the strongest depletion was observed for Scw10, Pdi1, Epx1, Kar2 and Exg1. In the glucose-



**Fig. 5.** Relative quantities of proteins present in the supernatants of the methanol- and glucose-based fed-batch (FB) processes using Max-Quant. Bubble size indicates protein abundance in the respective culture condition related to the protein in lowest detectable concentration (Acs2 in the glycerol FB) in a log scale sorted for their abundance in the glycerol FB sample. The bubble colour is representing the glycerol batch (B) in light grey, the glycerol FB in dark grey, the methanol FB in shades of blue and the glucose FB in shades of green. Proteins of the core secretome are written in bold letters, and underlined proteins do not contain a predicted signal peptide.

based process Tos1, Pir1, Sun4, PP7435\_Ch2-0752 and Epx1 became enriched, while Pdi1, Kar2, Egx1, Fpr2 and Flo5-1 declined. Compared to the glycerol fed-batch,

the ER proteins Pdi1, Kar2, Cpr5 and Fpr2, the proteases Pep4 and Yps1-1, and the cytosolic enzyme Tdh3 (GAP) were the most depleted proteins in both methanol- and

glucose-based fed-batches. In the glucose-based process, the uncharacterized proteins Vps72 and Ssp120 were additionally depleted. This indicated that the ER-resident proteins were secreted during the conditions when cells were growing at their maximum specific growth rate (batch phase). Also, the protease Pep4, which should be located in the vacuole, followed a similar time-course. The strongest difference between the methanol- and glucose-based processes (Appendix S3: Tables S3 and S4) was a stronger representation of superoxide dismutase Sod1, Vps72 and acetyl-coA synthase Acs2 (the latter both putatively involved in histone modifications) in the supernatant of the methanol FB, whereas lower amounts of the cell wall or cell wall-associated proteins (Flo5-2, Cda2, Vps1-1) and PP7435\_Chr3-1225 were detected in the methanol FB compared to the glucose FB. Interestingly, levels of the glucanase Sun4 and the major secreted protein Epx1 remained stable or slightly declined in the methanol fed-batch, while they accumulated over the course of the glucose fed-batch. In contrast, glucanase Bgl2 and cell wall protein Scw10 behaved *vice versa*. Although showing increased abundance at the end of all the bioreactor cultivations, enrichment of the cell wall-associated proteins Cts1, Flo5-2, Gas1, Uth1, Pir1, Pry2, PP7435\_Chr3-1225, Rce3 and Tos1 was stronger in the glucose fed-batch.

## Discussion

The secretome is defined as proteins secreted into the extracellular space (culture supernatant) of cells (Agrawal, *et al.*, 2010) and is a crucial measure to evaluate process performance and downstream processing requirements during recombinant protein production. Although bioinformatics prediction data for the secretome are available, the results of the prediction tools often do not reflect the experimentally identified proteins in the culture supernatants as they simply judge if a protein is able to enter the secretory pathway, but do not take into account condition-specific expression changes. Different growth conditions have been shown to influence or alter the secretome in the yeasts *S. cerevisiae* (Giardina, *et al.*, 2014), *Kluyveromyces lactis* (Swaim, *et al.*, 2008; Madinger, *et al.*, 2009) and *Candida utilis* (Buerth, *et al.*, 2011). However, no systematic or quantitative analysis of the secretomes of glucose- or methanol-grown *P. pastoris* fed-batch cultures have been performed so far.

*Intact and viable P. pastoris in controlled bioreactor fed-batch cultures secrete only a low number of native proteins, even during high cell density cultivation*

Contrary to previous studies (Mattanovich, *et al.*, 2009; Huang, *et al.*, 2011), we performed comparative

quantitative proteomics analyses of culture supernatants of a recombinant *P. pastoris* strain that was cultivated under controlled bioreactor conditions to investigate the secretome composition over the time-course of glycerol batch and glucose- or methanol-based fed-batch processes. Thereby, we could show that even in high cell density fed-batch cultures (up to 100 g l<sup>-1</sup> dry cell weight) only a low number of proteins was secreted by *P. pastoris* besides the recombinant product (in total 51 proteins) and that many of the host cell proteins are only present at low levels.

Proteins can reach the extracellular space by active secretion facilitated by a secretion signal, by passive secretion due to co-localization with other secretory proteins, or by cell lysis. The presence of signal peptides should make it possible to distinguish between active or passive secretion and release from the cell by cell lysis. The respective analysis revealed that more than 70% of the identified proteins possess a predicted signal peptide, which makes them likely to reach the supernatant by secretion and not via cell lysis. The smaller fraction without predicted signal peptide had predicted cellular localizations in the cytosol, mitochondria, nucleus, cytoskeleton or peroxisome. Most of these proteins, which are involved in metabolism or associated with ribosomes and translation (Tdh3, Acs2, Fdh1, Hyp2 and Tef2), were detected in the supernatants with only low abundance despite having high transcript and intracellular protein levels (see Appendix S3: Tables S1). ER proteins such as Kar2, Pdi1 or Cpr2 and the vacuolar protease Pep4, which showed higher abundance during the batch phase and got depleted during the fed-batch, were most probably colocalized with the recombinant protein and became cargoed into secretory vesicles.

Unexpectedly, only few differences were detected between the secretomes of the methanol- or glucose-grown cells. Moreover, most of the proteins were already present after the glycerol batch and/or glycerol fed-batch and the number of identified secreted proteins did not further increase but rather declined during the fed-batch, independent of the used substrate.

As proteolytic activity remained low throughout the processes, the observed decline cannot be attributed to proteolytic degradation of secreted proteins, but rather indicates that their active secretion was ceased and the levels decreased below the detection limit by dilution.

*The secretomes of methanol- and glucose-based fed-batch processes are highly similar*

By analysing the proteomics data of the different culture conditions, we were able to define a core secretome, comprising 22 proteins identified in all three carbon sources in our study. This core secretome mainly

consists of cell wall proteins. Despite being found in all carbon sources, the individual abundance of these proteins changes over cultivation time. Also besides the core secretome, the functional distribution of proteins identified in supernatants of methanol- or glucose-grown cultures was rather similar to some exceptions (Fig. 4). These minor differences are in contrast to *K. lactis* and *S. cerevisiae* where a strong impact of the carbon source on the secretome was observed (Madinger, et al., 2009; Giardina, et al., 2014). The different behaviour could be explained by different native habitats of these species and a potential physiological role of secreted proteins.

It has been reported that reduced viability in methanol-based expression systems can be problematic due to the release of intracellular proteases upon cell lysis, which is detrimental for the secreted recombinant product (Sinha, et al., 2005; Xiao, et al., 2006). Contrary to these reports and to other yeast species, for example, *Candida albicans* (Vargas, et al., 2015), proteases are not prominent in the supernatants of *P. pastoris* in controlled bioreactor cultures. Only two proteases were detected, the vacuolar protease Pep4 and the cell wall-associated yapsin Yps1-1. Both were already detected in glycerol batch and fed-batch, and their abundance declined over the course of the fed-batches independently whether the carbon source was glucose or methanol.

This correlates with the high viability and the low proteolytic activity detected in our fed-batch processes, even when methanol was used as a carbon source. Based on the total DNA content of the supernatants, a maximum of 0.2% lysed cells was estimated, accounting for maximally 4% of total protein in the supernatant. This emphasizes that the culture conditions are crucial for cell physiology and viability, thereby also influencing the quality and purity of the secreted product. In this respect, it should be noted that all of the 13 signal peptide-containing proteins from Mattanovich et al. (2009) were also identified in the present study, while from the fraction of proteins without predicted signal peptide only Tdh3 was shared (shared proteins are highlighted in bold in the column DSMZ70382 in Appendix S3: Table S1). This effect was even more pronounced when comparing our results to the ones of Huang et al. (2011), where the authors speculated that cell lysis occurred upon methanol induction, in particular in the recombinant strain, as they found 34 proteins without a signal peptide. In fact, 28 of the 41 proteins with signal peptide in Huang et al. (2011) were identified also in our study, compared to only 3 of the 34 proteins without signal peptide (Rpl40A, Sod1 and Ynk1; all shared proteins are highlighted in bold in the column 'GS115 RefSeq Protein' in Appendix S3: Table S1).

#### Quantitative changes of secretome proteins during methanol and glucose fed-batch

Quantitative data further revealed that in *P. pastoris* the most abundant proteins represent cell wall or cell wall-associated proteins and also chaperones like Kar2 and Pdi1, which could be co-secreted with the recombinant protein. Secretion of ER-resident chaperones has been described for *S. cerevisiae* and *P. pastoris* under conditions where the unfolded protein response (UPR) is induced (Chaudhuri, et al., 1991; Belden and Barlowe, 2001; Liu, et al., 2005; Guerfal, et al., 2010; Samuel, et al., 2013; Roth, et al., 2018). It is assumed that during the UPR either the ER retention and Erd2-dependent retrieval of ER proteins is saturated, thus leading to secretion of Kar2 and other HDEL-containing proteins. Activation of the UPR has been reported for most *P. pastoris* strains secreting heterologous proteins (Gasser, et al., 2007; Resina, et al., 2007) and is also occurring to a minor extent in the CpB producing strain used in our study.

As the abundance of extracellular Kar2 and other ER chaperones was highest in the glycerol batch and fed-batch and declined over glucose- or methanol-feed time, our results indicate that such minor UPR in recombinant cells does not aggravate over time in bioreactor cultivations. The fact that the highest levels of chaperones like Kar2 and Pdi1 or the vacuolar protease Pep4 were already present after the batch phase in our study shows that cells growing near their maximum specific growth rate and producing recombinant proteins probably overwhelm their intercellular protein sorting capacity, leading to leakage of proteins otherwise retained within the secretory pathway. It seems that the phenomenon of leakage of ER-resident proteins is restricted to cells with secretion stress (induced UPR) and we can speculate that a similar logic also applies to secreted Pep4.

#### Conclusion

The high similarity of the secretomes on glycerol, glucose and methanol allowed to define a core secretome of *P. pastoris*, which comprised of 22 proteins mainly involved in cell wall biogenesis and related functions. Most secretome proteins were already present after the glycerol batch, and especially, proteases and chaperones were depleted from the batch towards the end of the fed-batches. Altogether, our results indicate that recombinant proteins should be preferentially produced by inducible expression systems such as methanol-based P<sub>AOX1</sub> or methanol-independent P<sub>GTH1</sub> to ensure short exposure time to proteases and facilitate downstream processing.

## Experimental procedures

### Strain

*Pichia pastoris* (*Komagataella phaffii*) CBS7435 (CBS-KNAW culture collection, NL) producing porcine pro-carboxypeptidase B (CpB) as model protein under the control of the  $P_{GAP}$  and the *S. cerevisiae* alpha mating factor secretion leader were used (Burgard, *et al.*, 2017).

### Fed-batch cultivations

The fed-batch processes and the cultivation media were described previously (Burgard, *et al.*, 2017). Bioreactor cultivations were performed in triplicates at 25°C in 2.7 l DASGIP bioreactors with a batch volume of 1.225 l (Eppendorf, Germany). Dissolved oxygen was controlled at 20%, and pH was adjusted to 5.0 by 25%  $NH_3$  upon demand. Foam formation was prevented by the addition of Glanapon 2000 (5%; Bussetti & Co GmbH, Vienna, Austria) upon demand using a level probe. The fed-batch (FB) process comprised four phases, starting with a glycerol batch phase (inoculated to OD 1). The cultivations were performed using the following set-up: after reaching the batch end at about 24 h, a carbon-limited glycerol fed-batch (GLY01) for 5 h was started with a feed rate of  $5.5 \text{ g l}^{-1} \text{ h}^{-1}$  to reach a biomass of  $40 \text{ g l}^{-1}$  dry cell weight (DCW). Methanol-grown cultures were then supplemented with 0.5% (w/v) methanol for the induction of methanol metabolism. After consumption of the methanol pulse, a methanol fed-batch (MET01) was started and cultures were grown for 67 h to reach a final biomass of  $100 \text{ g l}^{-1}$ . The glucose fed-batch (GLU04) was started with an initial feed rate of  $1.9 \text{ g l}^{-1} \text{ h}^{-1}$  for 4 h to generate the same biomass as the methanol shot-cultures. The feed rate of the methanol or glucose fed-batch medium during the fed-batch was increased after 14 h (methanol FB: initial feed rate  $6.8 \text{ g l}^{-1} \text{ h}^{-1}$ , feed increase after 12 h to  $9.0 \text{ g l}^{-1} \text{ h}^{-1}$ ; glucose FB: initial feed rate  $9.9 \text{ g l}^{-1} \text{ h}^{-1}$ , feed increase after 14 h to  $13.2 \text{ g l}^{-1} \text{ h}^{-1}$ ). Due to insolubility of some salts in the methanol fed-batch medium, methanol- and glucose-grown cultures were supplemented with respective amounts of salt shots (see below) representing a surplus for the generation of about 10 g DCW. Samples were taken at the end of the glycerol batch; the end of the glycerol fed-batch; and 1 h, 15 h, 27 h, 41 h, 53 h and 67 h after starting the methanol or glucose fed-batches.

### Biomass determination and total protein content

Dry cell weight (DCW) was determined in triplicates by sampling of 2-ml cell suspension into pre-weighed and

pre-incubated (1 d, 100°C) reaction tubes and subsequent washing of cell pellet. DCW was determined after 3 days of drying at 100°C. Protein content of culture supernatants was quantified using the colorimetric BCA assay (Pierce™ BCA Protein Assay Kit, Thermo Fisher Scientific, Vienna, Austria).

### Viability measurements

Culture viability was determined using a flow cytometry-based method (BD™ Cell Viability kit). Bioreactor samples were stored on ice and cells were then separated by ultrasonication ( $3 \times 6 \text{ s}$ , 85% amplitude) and stained with propidium iodide according to Hohenblum *et al.* (2003) exhibiting fluorescence at 480 nm and emission maximum at 630 nm. Flow cytometric analysis was performed at a Gallios™ Flow Cytometer (Beckman Coulter, Vienna, Austria).

### Transcript levels

Transcript levels of the genes encoding the secreted proteins were excerpted from the microarray data in Burgard *et al.* (2017), which were performed on samples taken from the same fed-batch cultivations as the secretome. All experimental details on RNA preparation, microarray analysis and bioinformatics data evaluation are given in Burgard *et al.* (2017). Expression strengths in Fig. 2C and Appendix S3: Table S1 are based on the mean intensity values of the microarray signals.

### Analyses of culture supernatants

**Enzymatic CpB assay.** Quantification of recombinant product in culture supernatants was done using an enzymatic assay detecting CpB activity. Prior to activity determination, supernatants were desalted using Zeba™ Spin columns (Thermo Fisher Scientific, Vienna, Austria; 2 ml, 7K MWCO) and pro-CpB was activated by trypsin (30 min, 37°C). Due to the very sensitive assay principle, reactions were performed in cuvettes and CpB activity was measured by the spectrophotometric method of Folk *et al.* (1960) where the reaction velocity is determined by an increase in absorbance at 254 nm resulting from the hydrolysis of hippuryl-L-arginine. One unit CpB causes the hydrolysis of one  $\mu\text{mole}$  of hippuryl-L-arginine per minute at 25°C and pH 7.65. All samples were measured in triplicates, and activity was correlated to product concentrations using a CpB standard ranging from 0 to  $69 \text{ mg l}^{-1}$ .

**SDS-PAGE.** Qualitative analysis of culture supernatants was performed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) using the NuPAGE Bis-Tris system (MOPS buffer). Samples

were heat-denatured (10 min, 95°C) and reduced. A BenchMark™ Protein Ladder (Thermo Fisher Scientific) was used. Proteins were separated (180 V, 70 min) and stained using PageBlue™ Protein Staining Solution (Thermo Fisher Scientific). After electrophoresis, the gel was washed 3 times with deionized water, followed by staining with Coomassie Brilliant Blue and destained with deionized water. For the anti-CpB, Western blot of 9.75 µl of 1:10 diluted culture supernatants after 96 h of total process time (reduced and denatured) was applied before and after trypsin treatment (see Enzymatic CpB assay), and CpB (either as pro-CpB or mature CpB) was detected with the monoclonal HRP-conjugated anti-CpB antibody (1:1500) and chemiluminescence detection.

**Proteolytic activity.** For quantification of proteolytic activity in culture supernatants, a colorimetric assay was used (QuantiCleave; Pierce, Thermo Scientific, Vienna, Austria) as described in Marsalek *et al.* (2017). This assay uses fully succinylated casein as a substrate and measures the activity of any protease that cleaves casein. Hydrolysis of this substrate in the presence of proteases results in the release of peptide fragments with free amino-terminal groups. These peptides are reacted with trinitrobenzene sulphonic acid (TNBSA), followed by measurement of the absorbance increase at 450 nm that is due to the formation of yellow-coloured TNB-peptide adducts. Protease concentration is given as tryptic equivalents ( $\mu\text{g ml}^{-1}$ ).

**Host cell DNA.** Host cell DNA (hcDNA) was purified from culture supernatants using the Blood&Tissue Kit (Qiagen, Vienna, Austria). Host cell DNA was then quantified from the eluate using the fluorescence-based Qubit assay (Life Technologies, Thermo Scientific, Vienna, Austria) according to the manufacturer's instructions. Other, more sensitive assays, were not applicable due to fluorescent low molecular weight substances (probably flavins) present especially in methanol-grown cultures, which emitted strongly at the same wavelength as the DNA-intercalating dye.

**Proteomics analyses of the secretome by LC-ESI-MS/MS.** Culture supernatants of four biological replicates of glycerol batch and glycerol fed-batch, and 3 biological replicates each for 3 sampling time points during methanol or glucose FB were incubated for 10 min at 99°C for protease inactivation, and proteins were subsequently precipitated using precooled acetone ( $-20^{\circ}\text{C}$ ). After 1 h incubation at  $-20^{\circ}\text{C}$ , the precipitate was spun down (15 000 g) and the supernatant was removed. After acetone was evaporated at room temperature, the precipitate was resuspended in ammonium hydrogen carbonate buffer (100 mM).

The samples were then reduced (15 mM DTT), carbamidomethylated (55 mM iodoacetamide) and

methanol/chloroform precipitated (Botelho, *et al.*, 2010). The pellet was re-dissolved in 0.1 M ammonium bicarbonate buffer and digested overnight with trypsin at a 1:50 enzyme to substrate ratio (sequencing grade trypsin, Promega) at 37°C (Dragosits, *et al.*, 2009; Klug, *et al.*, 2014).

The peptide mixture was analysed using a Dionex Ultimate 3000 system, directly linked to a QTOF MS (Bruker maxis 4G ETD) equipped with the standard ESI source in the positive ion, DDA mode (= switching to MSMS mode for eluting peaks). MS-scans were recorded (range: 150–2200 Da), and the 6 highest peaks were selected for fragmentation. Instrument calibration was performed using ESI calibration mixture (Agilent). Thermo BioBasic C18 column (BioBasic-18, 150 x 0.32 mm, 5 µm, Thermo Scientific) using 65 mM ammonium formate buffer as the aqueous solvent was used. A gradient from 5% B (B: 80% ACCN) to 42% B in 45 min was applied, followed by a 15 min gradient from 42% B to 90% B that facilitates elution of large peptides, at a flow rate of 6 µl min<sup>-1</sup>.

The analysis files were converted (using Data Analysis, Bruker) to XML files, which are suitable for performing a MS/MS ion search with ProteinScape (MASCOT). The tandem mass spectrometry data were assigned to protein sequences of databases (the used database was UniProt —*P. pastoris*). Alternatively, the analysis files were analysed using MaxQuant (Cox and Mann, 2008; Cox, *et al.*, 2014) for 'absolute' quantification of selected proteins. For qualitative analysis, the iBAQ (intensity-based absolute quantification) values calculated from MaxQuant were used. These give a good approximation of the actual concentration in the sample (Krey, *et al.*, 2014).

The criteria for valid identification in MASCOT were score > 30 and identification of at least two peptides. Further data processing included filtering of proteins that were not found in at least two biological replicates. Protein content of analysed culture supernatants was then corrected for the increasing biomass fraction of the culture broth during the fed-batch process. The dilution resulting from constant media feed and requested base for pH control was considered by using a volume-dependent correction factor.

**Prediction of signal peptides and cellular localization.** Prediction of signal peptides and GPI-anchor was performed using SignalP 4.1 (<http://www.cb.s.dtu.dk/services/SignalP-4.0/>) (Petersen, *et al.*, 2011)) and PredGPI (<http://gpcr2.biocomp.unibo.it/gpipe/index.htm>) (Pierleoni, *et al.*, 2008)) respectively. The cellular localization of proteins was predicted using WoLF PSORT (<http://wolfpsort.hgc.jp/>).

**Categorization of the secretome.** A categorization of the proteins identified in the secretome was performed using

Slim GO mapper (<https://www.yeastgenome.org/cgi-bin/GO/goSlimMapper.pl>), using the annotations of the homologs in *S. cerevisiae*), based on their biological function or cellular compartment.

### Acknowledgements

The authors thank Markus Buchetics for his excellent help during bioreactor cultivations.

### Conflict of interest

None declared.

### References

- Adams, D.J. (2004) Fungal cell wall chitinases and glucanases. *Microbiology* **150**: 2029–2035.
- Adivitiya Dagar, V.K., Devi, N., and Khasa, Y.P. (2016) High level production of active streptokinase in *Pichia pastoris* fed-batch culture. *Int J Biol Macromol* **83**: 50–60.
- Agrawal, G.K., Jwa, N.S., Lebrun, M.H., Job, D., and Rakwal, R. (2010) Plant secretome: unlocking secrets of the secreted proteins. *Proteomics* **10**: 799–827.
- Belden, W.J., and Barlowe, C. (2001) Deletion of yeast p24 genes activates the unfolded protein response. *Mol Biol Cell* **12**: 957–969.
- Botelho, D., Wall, M.J., Vieira, D.B., Fitzsimmons, S., Liu, F., and Doucette, A. (2010) Top-down and bottom-up proteomics of SDS-containing solutions following mass-based separation. *J Proteome Res* **9**: 2863–2870.
- Brown, L., Wolf, J.M., Prados-Rosales, R., and Casadevall, A. (2015) Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nat Rev Microbiol* **13**: 620–630.
- Buerth, C., Heilmann, C.J., Klis, F.M., de Koster, C.G., Ernst, J.F., and Tielker, D. (2011) Growth-dependent secretome of *Candida utilis*. *Microbiology* **157**: 2493–2503.
- Burgard, J., Valli, M., Graf, A.B., Gasser, B., and Mattanovich, D. (2017) Biomarkers allow detection of nutrient limitations and respective supplementation for elimination in *Pichia pastoris* fed-batch cultures. *Microb Cell Fact* **16**: 117.
- Carnicer, M., Baumann, K., Töplitz, I., Sánchez-Ferrando, F., Mattanovich, D., Ferrer, P., and Albiol, J. (2009) Macromolecular and elemental composition analysis and extracellular metabolite balances of *Pichia pastoris* growing at different oxygen levels. *Microb Cell Fact* **8**: 65.
- Chaudhuri, B., Helliwell, S.B., and Priestle, J.P. (1991) A Lys27-to-Glu27 mutation in the human insulin-like growth factor-1 prevents disulfide linked dimerization and allows secretion of BiP when expressed in yeast. *FEBS Lett* **294**: 213–216.
- Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K., and Romanos, M.A. (1991) High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Biotechnology (N Y)* **9**: 455–460.
- Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**: 1367–1372.
- Cox, J., Hein, M.Y., Lubner, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**: 2513–2526.
- Damasceno, L.M., Pla, I., Chang, H.J., Cohen, L., Ritter, G., Old, L.J., and Batt, C.A. (2004) An optimized fermentation process for high-level production of a single-chain Fv antibody fragment in *Pichia pastoris*. *Protein Expr Purif* **37**: 18–26.
- Dragosits, M., Stadlmann, J., Albiol, J., Baumann, K., Maurer, M., Gasser, B., et al. (2009) The effect of temperature on the proteome of recombinant *Pichia pastoris*. *J Proteome Res* **8**: 1380–1392.
- Folk, J.E., Piez, K.A., Carroll, W.R., and Gladner, J.A. (1960) Carboxy-peptidase B. 4. Purification and characterization of the porcine enzyme. *J Biol Chem* **235**: 2272–2277.
- Gasser, B., Maurer, M., Rautio, J., Sauer, M., Bhat-tacharyya, A., Saloheimo, M., et al. (2007) Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions. *BMC Genom* **8**: 179.
- Giardina, B.J., Stanley, B.A., and Chiang, H.L. (2014) Glucose induces rapid changes in the secretome of *Saccharomyces cerevisiae*. *Proteome Sci* **12**: 9.
- Guerfal, M., Ryckaert, S., Jacobs, P.P., Ameloot, P., Van Craenenbroeck, K., Derycke, R., and Callewaert, N. (2010) The *HAC1* gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins. *Microb Cell Fact* **9**: 49.
- Heiss, S., Maurer, M., Hahn, R., Mattanovich, D., and Gasser, B. (2013) Identification and deletion of the major secreted protein of *Pichia pastoris*. *Appl Microbiol Biotechnol* **97**: 1241–1249.
- Hohenblum, H., Borth, N., and Mattanovich, D. (2003) Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. *J Biotechnol* **102**: 281–290.
- Hohenblum, H., Gasser, B., Maurer, M., Borth, N., and Mattanovich, D. (2004) Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol Bioeng* **85**: 367–375.
- Huang, C.J., Damasceno, L.M., Anderson, K.A., Zhang, S., Old, L.J., and Batt, C.A. (2011) A proteomic analysis of the *Pichia pastoris* secretome in methanol-induced cultures. *Appl Microbiol Biotechnol* **90**: 235–247.
- Klug, L., Tarazona, P., Gruber, C., Grillitsch, K., Gasser, B., Trozsmuller, M., et al. (2014) The lipidome and proteome of microsomes from the methylotrophic yeast *Pichia pastoris*. *Biochim Biophys Acta* **1841**: 215–226.
- Krey, J.F., Wilmarth, P.A., Shin, J.B., Klimek, J., Sherman, N.E., Jeffery, E.D., et al. (2014) Accurate label-free protein quantitation with high- and low-resolution mass spectrometers. *J Proteome Res* **13**: 1034–1044.
- Lex, A., Gehlenborg, N., Strobel, H., Vuillemot, R., and Pfister, H. (2014) UpSet: Visualization of Intersecting

- Sets. *IEEE Trans Vis Comput Graph* **20**(12): 1983–92. <https://doi.org/10.1109/TVCG.2014.2346248>.
- Liu, Y.Y., Woo, J.H., and Neville, D.M. (2005) Overexpression of an anti-CD3 immunotoxin increases expression and secretion of molecular chaperone BiP/Kar2p by *Pichia pastoris*. *Appl Environ Microbiol* **71**: 5332–5340.
- Love, K.R., Dalvie, N.C., and Love, J.C. (2017) The yeast stands alone: the future of protein biologic production. *Curr Opin Biotechnol* **53**: 50–58.
- Madinger, C.L., Sharma, S.S., Anton, B.P., Fields, L.G., Cushing, M.L., Canovas, J., et al. (2009) The effect of carbon source on the secretome of *Kluyveromyces lactis*. *Proteomics* **9**: 4744–4754.
- Marsalek, L., Gruber, C., Altmann, F., Aleschko, M., Mattanovich, D., Gasser, B., and Puxbaum, V. (2017) Disruption of genes involved in CORVET complex leads to enhanced secretion of heterologous carboxylesterase only in protease deficient *Pichia pastoris*. *Biotechnol J* **12**: 1600584. <https://doi.org/10.1002/biot.201600584>.
- Mattanovich, D., Graf, A., Stadlmann, J., Dragosits, M., Redl, A., Maurer, M., et al. (2009) Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb Cell Fact* **8**: 29.
- Peres da Silva, R., Puccia, R., Rodrigues, M.L., Oliveira, D.L., Joffe, L.S., Cesar, G.V., et al. (2015) Extracellular vesicle-mediated export of fungal RNA. *Sci Rep* **5**: 7763.
- Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**: 785–786.
- Pierleoni, A., Martelli, P.L., and Casadio, R. (2008) PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* **9**: 392.
- Puxbaum, V., Mattanovich, D., and Gasser, B. (2015) Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. *Appl Microbiol Biotechnol* **99**: 2925–2938.
- Rabert, C., Weinacker, D., Pessoa, A. Jr, and Farias, J. G. (2013) Recombinants proteins for industrial uses: utilization of *Pichia pastoris* expression system. *Braz J Microbiol* **44**: 351–356.
- Resina, D., Bollok, M., Khatri, N.K., Valero, F., Neubauer, P., and Ferrer, P. (2007) Transcriptional response of *P. pastoris* in fed-batch cultivations to *Rhizopus oryzae* lipase production reveals UPR induction. *Microb Cell Fact* **6**: 21.
- Roth, G., Vanz, A.L., Lünsdorf, H., Nimtz, M., and Rinas, U. (2018) Fate of the UPR marker protein Kar2/Bip and autophagic processes in fed-batch cultures of secretory insulin precursor producing *Pichia pastoris*. *Microb Cell Fact* **17**: 123.
- Samuel, P., Prasanna Vadhana, A.K., Kamatchi, R., Antony, A., and Meenakshisundaram, S. (2013) Effect of molecular chaperones on the expression of *Candida antarctica* lipase B in *Pichia pastoris*. *Microbiol Res* **168**: 615–620.
- Sinha, J., Plantz, B.A., Inan, M., and Meagher, M.M. (2005) Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast *Pichia pastoris*: case study with recombinant ovine interferon-tau. *Biotechnol Bioeng* **89**: 102–112.
- Swaim, C.L., Anton, B.P., Sharma, S.S., Taron, C.H., and Benner, J.S. (2008) Physical and computational analysis of the yeast *Kluyveromyces lactis* secreted proteome. *Proteomics* **8**: 2714–2723.
- Vargas, G., Rocha, J.D., Oliveira, D.L., Albuquerque, P.C., Frases, S., Santos, S.S., et al. (2015) Compositional and immunobiological analyses of extracellular vesicles released by *Candida albicans*. *Cell Microbiol* **17**: 389–407.
- Weinacker, D., Rabert, C., Zepeda, A.B., Figueroa, C.A., Pessoa, A., and Farias, J.G. (2013) Applications of recombinant *Pichia pastoris* in the healthcare industry. *Braz J Microbiol* **44**: 1043–1048.
- Xiao, A., Zhou, X., Zhou, L., and Zhang, Y. (2006) Improvement of cell viability and hirudin production by ascorbic acid in *Pichia pastoris* fermentation. *Appl Microbiol Biotechnol* **72**: 837–844.
- Yu, Z., Wu, X., Li, D., Yang, S., Zhou, Z., Cai, J., and Yuan, Z. (2003) Enhancement of the production of SAM by overexpression of SAM synthetase in *Pichia pastoris*. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* **35**: 127–132.
- Zahl, R.J., Gasser, B., Mattanovich, D., and Ferrer, P. (2019) Detection and elimination of cellular bottlenecks in protein-producing yeasts. *Methods Mol Biol* **1923**: 75–95.
- Zhu, T., Sun, H., Wang, M., and Li, Y. (2019) *Pichia pastoris* as a versatile cell factory for the production of industrial enzymes and chemicals: Current status and future perspectives. *Biotechnol J* **14**: e1800694.

### Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Appendix S1.** List of proteins identified by MASCOT are presented. The MASCOT score, number of identified peptides and sequence coverage are shown for each sample.

**Appendix S2.** List of identified proteins by MaxQuant are presented. In the file, the number of identified peptides, the sequence coverage, the theoretical molecular weight and the iBAQ (Intensity Based Absolute Quantification) values are shown.

**Appendix S3. Table S1:** List of proteins being present or absent in the fed-batch samples including protein function, protein accession numbers, and predicted protein features. Microarray-derived transcript levels (mean expression values) of genes encoding for the identified proteins in the methanol- or glucose-based processes (derived from Burgard et al. (2017)). **Table S2:** List of proteins distributed to the different UpSet plot sections and intersections from the glycerol fed-batch as reference and the methanol or glucose fed-batch (1h, 27 h, 67 h). **Table S3:** List of proteins with different abundances (log<sub>2</sub> fold changes) in the methanol- and glucose-based processes compared to the glycerol fed-batch or between methanol and glucose fed-batch based on normalized MAXQUANT iBAQ-values. **Table S4:** Proteins which show strong accumulation or depletion during the methanol or glucose fed-batch phase compared to the glycerol fed-batch. Evaluation is based on normalized MAXQUANT iBAQ-values.