FEMS Yeast Research, 20, 2020, foaa009

doi: 10.1093/femsyr/foaa009
Advance Access Publication Date: 3 February 2020
Minireview

MINIREVIEW

Yeast synthetic biology for designed cell factories producing secretory recombinant proteins

Eun Jung Thak[†], Su Jin Yoo[†], Hye Yun Moon and Hyun Ah Kang[*,‡]

Laboratory of Molecular Systems Biology, Department of Life Science, Chung-Ang University, Seoul 06974, South Korea

[*]**Corresponding author:**Laboratory of Molecular Systems Biology, Department of Life Science, Chung-Ang University, Seoul 06974, South Korea. Tel: +82-2-820-5863; Fax: +82-2-825-5206; E-mail: hyunkang@cau.ac.kr

One sentence summary: Here, we discuss recent developments in synthetic biology approaches to the construction of yeast cell factories with enhanced capacities of protein folding, secretion, and designed post-translational modification functions.

[†]These authors contributed equally.

Editor: Jin Hou

[‡]Hyun Ah Kang, http://orcid.org/0000-0002-3722-525X

ABSTRACT

Yeasts are prominent hosts for the production of recombinant proteins from industrial enzymes to therapeutic proteins. Particularly, the similarity of protein secretion pathways between these unicellular eukaryotic microorganisms and higher eukaryotic organisms has made them a preferential host to produce secretory recombinant proteins. However, there are several bottlenecks, in terms of quality and quantity, restricting their use as secretory recombinant protein production hosts. In this mini-review, we discuss recent developments in synthetic biology approaches to constructing yeast cell factories endowed with enhanced capacities of protein folding and secretion as well as designed targeted post-translational modification process functions. We focus on the new genetic tools for optimizing secretory protein expression, such as codon-optimized synthetic genes, combinatory synthetic signal peptides and copy number-controllable integration systems, and the advanced cellular engineering strategies, including endoplasmic reticulum and protein trafficking pathway engineering, synthetic glycosylation, and cell wall engineering, for improving the quality and yield of secretory recombinant proteins.

Keywords: yeast cell factory; secretory recombinant proteins; protein trafficking pathway engineering; synthetic glycosylation pathway

INTRODUCTION

Developments in recombinant DNA technology have provided alternate routes for protein production in heterologous systems to overcome the difficulties in obtaining satisfactory yields from natural sources. There has been a consistent demand for enhanced yield and improved quality of recombinant proteins, necessitating the development of novel expression systems. Currently, various yeast expression systems are utilized to produce recombinant proteins with a wide range of industrial and medical applications (Nicaud *et al.* 2002; Çelik and Çalik 2012; Ergün *et al.* 2019). Because of their robust growth on simple media in large-scale bioreactors combined with their capacity for eukaryotic post-translational modifications and feasibility in genetic manipulations, yeasts are practical production hosts (Mattanovich *et al.* 2012; Baghban *et al.* 2019). Baker's yeast, *Saccharomyces cerevisiae*, is one of the best-characterized eukaryotes and most widely used host for recombinant protein production since the early days of genetic engineering. During the last decades, non-conventional yeast species, including *Hansenula polymorpha* (*Ogataea polymorpha*), *Pichia pastoris*

Received: 6 December 2019; **Accepted:** 30 January 2020

(*Komagataella phaffii*), *Yarrowia lipolytica* and *Kluyveromyces lactis*, have gained popularity as alternative hosts for the production of recombinant proteins due to several superior physiological properties (Vieira Gomes *et al.* 2018; Manfrao-Netto, Gomes and Parachin 2019). These non-conventional yeasts are mostly Crabtree-negative, allowing higher biomass yields in fermentation processes compared to *S. cerevisiae*, and utilize broad range of carbon sources (Kim, Yoo and Kang 2015). The methylotrophic yeasts *H. polymorpha* and *P. pastoris* can grow using methanol as a sole carbon source. Different from other methylotrophic yeasts, *H. polymorpha* shows peculiar resistance to high temperature up to 50°C, heavy metals and oxidative stress. The oleaginous dimorphic *Y. lipolytica* is known for its ability to utilize hydrophobic compounds such as fatty acids, alkanes and oils as carbon sources. The lactose-assimilating yeast *K. lactis* has the ability to metabolize xylitol, cellobiose and lactose in addition to pentose sugars containing xylose and arabinose. Compared to *S. cerevisiae*, non-conventional yeasts show generally much lesser extent of hypermannosylation (Thak *et al.* 2018). Moreover, the secretory pathway of these non-conventional yeasts is reported to be more similar to higher eukaryotes in contrast to *S. cerevisiae* in aspects of the structure of endoplasmic reticulum (ER) exit sites and Golgi apparatus (Ergün *et al.* 2019).

Protein secretion encompasses numerous complex steps mediated by several hundred cellular proteins. Core secretion functions, including translocation through the ER membrane, primary glycosylation, folding and quality control and vesicle-mediated secretion, are similar from yeasts to higher eukaryotes (Delic *et al.* 2013). Therefore, yeasts are preferred microbial hosts for the secretory production of recombinant proteins, particularly for proteins that originated from higher eukaryotes (Yoo, Moon and Kang 2019). Moreover, the limited endogenous protein secretion from yeast cells simplifies the downstream purification process. Thus, for large-scale industrial production, secretory recombinant protein production in yeasts offers the advantage of simple and efficient downstream purification that avoids costly cell rupture, denaturation and refolding. However, despite several evident advantages of yeasts as host cells, there are certain limitations to their use as expression systems, including inefficient secretion, improper folding, hyperglycosylation and aberrant proteolytic processing of proteins (Fig. 1). Each step of the protein secretion pathway can be engineered to improve protein yield and quality. Recently, novel methods and tools based on synthetic biology have attracted attention as they can facilitate the rapid and efficient engineering of yeast cells of various species for secretory recombinant protein production (Kim, Yoo and Kang 2015; Wagner and Alper 2016; Zahrl *et al.* 2017). Several synthetic biological approaches have allowed the improvement of secretion efficiency and folding capacity, which affect the final yield and quality of recombinant proteins in the culture supernatant.

Moreover, systems biology approaches have greatly contributed to understanding of underlying cellular mechanisms for efficient recombinant protein production, therefore helping in choosing targets and designing yeast cell factories. For example, a genome-scale metabolic model was used to predict gene targets for deletion or overexpression for enhanced recombinant protein production in *P. pastoris* (Nocon *et al.* 2014). As a more advanced *in silico* model, a dynamic genome-scale metabolic model of *P. pastoris* was developed to predict cell behavior for the rational design of bioprocesses and validated for its facility to improve the performance of a *P. pastoris* strain producing a secretory recombinant protein, human serum albumin (Saitua *et al.* 2017). RNA-seq analysis on the genome-wide transcriptional response to protein secretion in *S. cerevisiae* mutant strains, having higher protein secretion capacity for a recombinant protein, showed that tuning of energy metabolism for increased fermentation, as well as balancing of amino acid biosynthesis and reduced thiamine biosynthesis, is important to support efficient protein secretion (Huang *et al.* 2017). The advanced synthetic biology tools and modules combined with continued progress in systems biology approaches are expected as powerful strategies for holistic and rational engineering of cell factories (Davy, Kildegaard and Andersen 2017; Pena *et al.* 2018), which would facilitate systematically targeting various steps of protein secretion pathway and multilevel modifications of yeast host cells to improve secretory recombinant protein quality and yield.

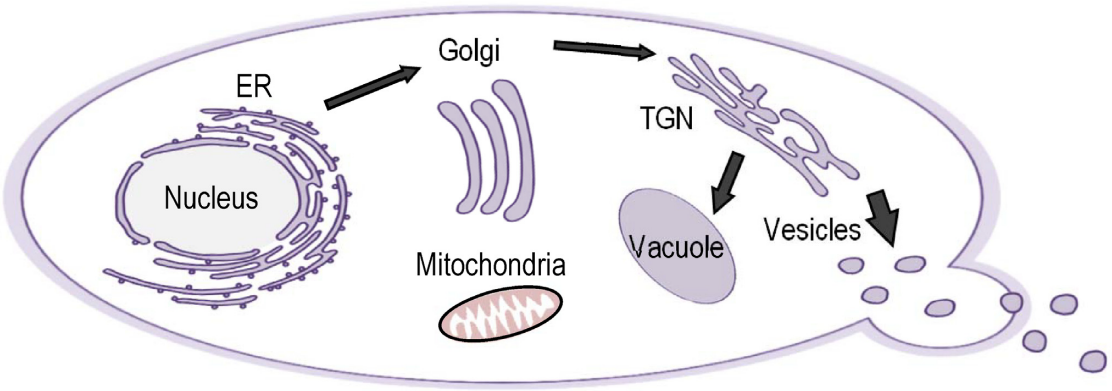
GENETIC TOOLBOXES FOR OPTIMIZED SECRETORY PRODUCTION OF PROTEINS

Several factors, such as the efficiency of translocation to the ER lumen, affect the early steps of the secretion pathway. Further, the level of protein production has to be optimized to avoid saturation of the secretory capacity of the host cells. Overloading of ER folding and secretion capacity can lead to the accumulation of mis- or unfolded proteins and, consequently, cause cellular stress and low protein yields. New copy-number control tools and synthetic secretion signal sequences have been developed as biological components and devices to optimize the expression levels and improve the secretion efficiency; these will be discussed below.

Codon-optimized synthetic genes for improved expression

Codon usage varies among different organisms, that is, a phenomenon known as codon usage bias. There are subsets of codons, called optimal codons, which are decoded by abundant transfer ribonucleic acids (tRNAs). The optimal codons are efficiently translated and used nearly exclusively in many highly expressed genes. Thus, codon optimization of heterologous sequences to match the host codon bias is considered as an initial strategy to achieve effectively high levels of recombinant protein expression. It was prevailingly considered that slow translation of codons decoded by rare tRNAs is the main mechanism to reduce translation elongation and efficiency. However, recent studies have strongly indicated that tRNA availability is not the sole determinant of rate, rather interactions between adjacent codons and wobble base pairing are critical (Chevance, Le Guyon and Hughes 2014). Specific codons and codon combinations were also indicated as important factors in modulating ribosome speed and facilitating protein folding (Brule and Grayhack 2017).

The traditional approach for codon optimization has been based on the individual codon usage (ICU) bias, which refers to the usage preference of synonymous codons in coding sequences. In recent years, another approach based on codon-pair context (CC) bias, which refers to the preferences for specific codon pairs observed in each organism, has drawn increasing attention as an efficient strategy for synthetic gene design to improve protein expression (Chung and Lee 2012). A recent paper reported the condition-specific codon optimization approach based on systems level information and codon context for the condition under which heterologous genes are being expressed, employing a probabilistic algorithm to generate multiple variants of a given gene (Lanza *et al.* 2014). The improved



	Nucleus	ER	Golgi	Vacuole	Post-Golgi
Key processes	• Transcription	• Signal recognition • Folding/modification • Glycosylation	• Protein processing • Glycosylation	• Degradative processes	• Post-Golgi process • Cell wall passage
Putative bottlenecks	• Vector Stability • Inefficient transcription	• Poor translocation • Misfolding • ER proteolytic degradation	• Incorrect processing • Trafficking defect • Hypermannosylation	• Mis-sorting	• Proteolysis • Retention by cell wall
Solutions	• Multiple integration • Synthetic promoters	• Synthetic signals • ER engineering • Glycoengineering	• Protease deficiency • Vesicle engineering • Glycoengineering	• Protease deficiency • Vesicle engineering	• Protease deficiency • Cell wall engineering

Figure 1. Genetic engineering strategies to overcome bottlenecks in the yeast protein secretion pathway. Protein secretion in yeast involves multiple complex steps, such as protein translocation, folding, post-translational modification and vesicle trafficking between several membrane organelles and plasma membranes. The secretion of proteins synthesized inside cells can be hampered by low secretion efficiency, abnormal post-translational modifications, retention within the secretion pathway or the cell wall space as a cell-associated form. The development of engineering strategies targeted to each step of the secretion pathway in a modular fashion has gained increasing attention in yeast synthetic biology to design cell factories producing secretory recombinant proteins.

translational efficiency using this approach was demonstrated with two heterologous genes, the fluorescent protein-encoding enhanced green fluorescent protein (eGFP) and the catechol 1,2-dioxygenase gene *CatA*, expressed in *S. cerevisiae*. In a more recent study on the secretory production of CAL-B in *P. pastoris*, three alternative strategies, ICU, CC and the trade-off between ICU and CC in a multiobjective fashion, were employed and compared for codon optimization of the alpha-mating factor (αMF) prepro-leader sequence to enhance secretory production of *Candida antarctica* lipase B (Ahn *et al.* 2016). The results indicated CC as the most relevant parameter for the codon optimization and no synergistic effect was achieved by combining both ICU and CC.

When performing codon optimization, several factors including average guanine-cytosine (GC) content of host genome, the secondary structure of messenger ribonucleic acids (mRNA), the codon adaption index and the tRNA pool should be taken into account to avoid unpredictable effects of changing coding sequence on expression (Zhao *et al.* 2014). Moreover, based upon a genome-wide analysis of mRNA stability, codon usage is implicated as the major determinant of mRNA stability and may affect protein activity (Brule and Grayhack 2017). Up to now, a number of case studies on the enhanced expression of synthetic codon-optimized genes have been reported in several yeast species (Kim, Yoo and Kang 2015; Yang and Zhang 2018). Both ICU- and CC-based parameters have been applied to codon optimization of many genes in yeast, in many cases resulting in improved levels of protein expression. Overall, compared to the conventional practice of using ICU fitness as a key element of synthetic gene design, CC fitness appears to be a more relevant design parameter for optimizing sequences for enhanced heterologous protein expression.

Copy-number-controllable integration systems for expression optimization

Generally, the gene constructs for recombinant protein expression can be introduced into yeast cells by employing two distinct strategies, namely, plasmid-based systems and chromosomal integration. Due to plasmid genetic instability, including segregation instability and structural instability, which significantly affect the ultimate yield of target products, particularly during large-scale and long-term industrial fermentation, integration of expression constructs into host chromosomes is a preferable approach over plasmid-based expression (Li *et al.* 2019). However, DNA integration into the host genome by non-homologous recombination or untargeted ectopic integration could cause the risk of unwanted side effects, such as altered gene expression due to interfering regulatory elements, or gene disruption following integration into protein-coding regions of the genome. To avoid the disadvantage of non-targeted integration, it is preferred to insert the genetic constructs into the gene locus whose

deletion does not affect the growth of yeast. For example, the highly efficient and stable secretory production of endo-1,4-β-glucanase was achieved by the expression of *Aspergillus niger eng1* from the vector integrated into the *HO* locus of the *S. cerevisiae* chromosome, since deletion of the *HO* locus does not affect the growth of yeast (Taipakova *et al.* 2015). When choosing the integration site for expression constructs, it should be kept in mind that protein expression levels significantly vary. Distinct position effects have been previously demonstrated by introducing foreign reporter genes into different loci of chromosomes in *S. cerevisiae* (Chen and Zhang 2016; Wu *et al.* 2017). To identify accessible integration sites for efficient expression without impairing cell growth, a set of genomic loci was screened using a GFP expression cassette in *Y. lipolytica* (Schwartz *et al.* 2017). There are several factors affecting gene expression in different chromosomal loci, such as level of DNA compaction, distance to DNA replication initiation site and availability of regulatory factors with effects on gene expression. Thus, further accumulated information on fundamental principles and characteristics of genomic integration by systematic examination of position effects will be helpful in guiding the design of integration sites for expression vectors to maximize production of recombinant proteins.

A more practical strategy to avoid the disadvantages of non-targeted integration and position effect is to integrate the genetic constructs for recombinant protein expression into a non-coding genomic position of yeast cells, such as the non-transcribed spacer (NTS) of the ribosomal DNA (rDNA) unit (Moon *et al.* 2016) and terminal repeat delta sites of Ty retrotransposon (Shi *et al.* 2016) in *S. cerevisiae*. To achieve stable and high-level heterologous protein expression, the introduction of DNA cassettes into multiple sites in yeast chromosomes, including the repeated ribosomal DNA, delta and sigma element sequences, is a preferred strategy over the integration into single insertion site. The correlation between expression cassette copy number and protein yield, particularly for the secretory production of recombinant proteins, is not always linear (Aw and Polizzi 2013). For example, studies on human serum albumin expression in *H. polymorpha* and *P. pastoris* showed that increasing the gene copy number did not result in higher production levels from a certain copy number, due to secretion saturation (Kang *et al.* 2001; Whyteside *et al.* 2011). Sha *et al.* (2013) reported enhancement of lipase production in *P. pastoris* by regulating gene dosage. Thus, using controllable multiple integration vectors would be a good strategy to introduce optimal copy numbers for high-level secretory recombinant protein expression. Combination of the multiple integration approach with the use of a defective selection marker has been put forward as a helpful strategy to modulate the copy number of an integrated DNA construct in several yeast expression systems. By exploiting the NTS of rDNA in combination with defective *URA3*, *LEU2*, *HIS3* and *TRP1* selection markers, which are designed to have truncated promoters of different lengths, a new set of ribosomal DNA-NTS-based multiple integrative cassettes was created for the development of antibiotic-marker-free recombinant yeasts in *S. cerevisiae* (Moon *et al.* 2016). The integration numbers of NTS-based expression cassettes, ranging from 1 to ~30, increased proportionally with the decrease in auxotrophic marker expression. In *Y. lipolytica*, a set of 26S ribosomal DNA-based multiple integrative vectors was used in combination with a series of defective *Ylura3* with truncated promoters as selection markers (Luu *et al.* 2017). In *H. polymorpha*, besides multiple integration systems based on defective auxotrophic markers, rapid and copy-number-controlled selection systems have been developed using antibiotic resistance markers as dominant selection markers, which allow the selection of multiple integrants with different copy numbers on plates containing different concentrations of G418 or hygromycin B (Yoo, Moon and Kang 2019). Such copy-number-controllable integration tools will be useful for the tuning of gene expression to maintain stable, high-level secretory production.

Modular toolkits for secretion signal libraries

The first step of secretion is the transfer of a protein through the ER membrane, termed translocation (Fig. 2A). Secretory signal peptide (SP) sequences are required to translocate proteins into the ER; thus, the efficiency of such SP sequences is one of the critical factors in determining the final protein yield. Protein secretion can be directed by either the native signal peptide of the heterologous protein or that of an endogenous yeast protein. Accurate SP selection or design is considered key to a high secretory protein production rate. Native prepro-leader sequences derived from yeast secretory proteins, such as *S. cerevisiae* αMF, and their modified versions are generally exploited to facilitate secretory expression of heterologous proteins in various yeast expression systems (Lin-Cereghino *et al.* 2013).

The secretion efficiency of a recombinant protein fused to a given SP sequence can strongly vary even for proteins with only small differences in their sequence, structure or gene expression regulation. None of the SP sequences reported to date can serve as an omnipotent partner for all proteins. Therefore, in recent studies, a hierarchical assembly method was popularly exploited to generate large and diverse secretion libraries to screen a wide range of secretion constructs and identify protein-specific SPs. Bae *et al.* (2015) developed an efficient genome-wide fusion-partner screening system to identify the best translation fusion partner to direct optimal secretion in *S. cerevisiae*. A translation fusion partner library was constructed from a genomic and truncated cDNA library, using the invertase-based signal sequence trap technique (Lee and Rose 2012). This screening technique was used to identify various secretion signals and fusion partners from the translation fusion partner library for the secretion of individual proteins, facilitating the search for an optimal secretion signal and fusion partner. A modular toolkit for generating *P. pastoris* secretion libraries was developed by employing standardized regulatory elements specific for *P. pastoris*, which allowed the tuning of gene expression by six different promoters (p*AOX1*, p*GAP1*, p*TPI1*, p*ENO1*, p*TEF1* and p*PGK1*) and a choice of 10 different protein secretion tags (αMF and nine modified derivatives) (Obst, Lu and Sieber 2017). These results highlight the importance of generating diverse secretion libraries when searching for optimal expression conditions.

Using *in silico* analysis based on the D-score of a given sequence segment, five endogenous SPs of *P. pastoris* were selected and their efficiency was investigated using recombinant human growth hormone as the model protein (Massahi and Çalık 2016). A newly identified SP23 (MKILSALLLFTLAFA) showed the highest efficiency among the tested endogenous SPs, comparable to *S. cerevisiae* αMF. However, the examined endogenous SPs did not show clear correlation between secretion efficiency and major physicochemical properties, including hydropathy index, isoelectric point and aliphatic index. The results indicate that other factors such as protein N-terminus effect, length of the SP, secondary structure of the SP and

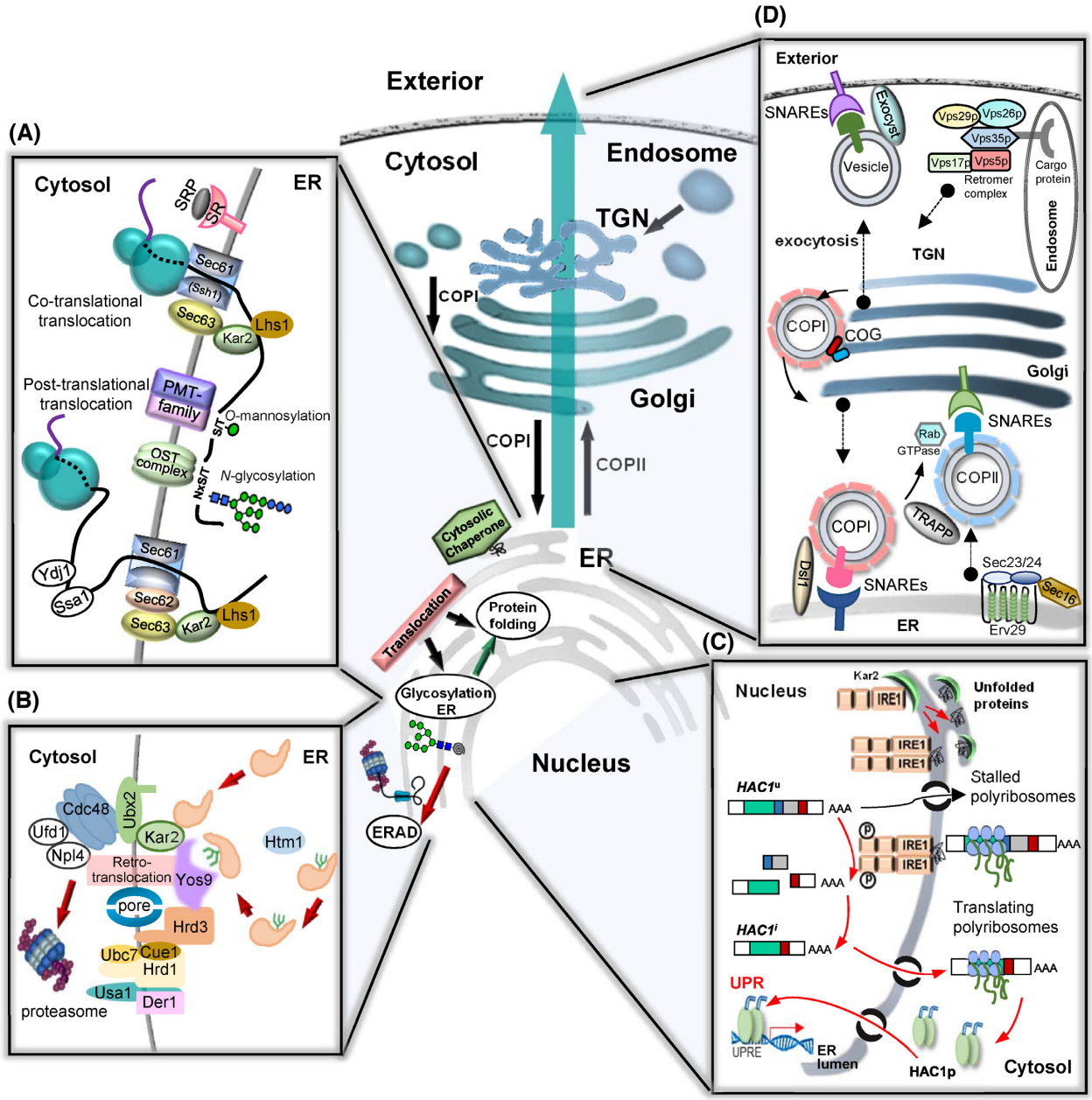


Figure 2. Major steps in canonical protein trafficking, ER-associated degradation pathway (ERAD) and the unfolded protein response (UPR) in yeasts. **(A)** Protein translocation across the ER membrane. In co-translational translocation, the signal sequence of the protein is recognized by the signal recognition particle (SRP), whereas the protein is synthesized on the ribosome. The ribosome–protein–SRP complex binds to the SRP receptor on the ER membrane, and the nascent protein is inserted into either the Sec61p or the Ssh1p translocation complex that pushes it through the ER membrane. In post-translational translocation, the mature protein is inserted through the Sec61p translocation pore into the ER after translation has been completed in the cytosol. **(B)** Protein folding, glycosylation and degradation in the ER. After translocation into the ER lumen, proteins are folded with the assistance of chaperones and can be subjected to *N*-/*O*-glycosylation. Mis- and unfolded proteins become targets of the ERAD, where they are exported from the ER, ubiquitinated and hereby targeted for degradation by the proteasome. **(C)** Induction of the UPR by accumulation of misfolded proteins in the ER. Correctly folded proteins enter the secretory pathway, whereas protein misfolding leads to the activation of the UPR, which induces the activation of numerous cellular processes. The accumulation of misfolded proteins results in increased binding of Kar2p to these proteins and, consequently, dimerization of the UPR sensor Ire1p to initiate the UPR by activating Hac1p synthesis through *HAC1* mRNA splicing. **(D)** Protein trafficking from the ER to the Golgi and exocytosis. Folded proteins are transported by COPII vesicles to the Golgi, whereas COPI vesicles are responsible for retrotransport to *cis*-Golgi and ER. Proteins destined for the exterior are transported by secretory vesicles to the cell membrane.

interactions of the selected properties should also be investigated in evaluating the parameters that influence SP efficiency *in silico*. In a recent work in *Y. lipolytica*, five novel SPs for recombinant protein secretion were identified by genomic and transcriptomic data mining (Celińska *et al.* 2018). The secretory potential of these new SPs was assessed in comparison with that of five known, widely exploited SPs in *Y. lipolytica*, including the native SP from LIP2 (lipase) and its engineered variant (LIP2pre-3xLA), preXPR2 (alkaline extracellular protease) and heterologous SPs derived from SoAMY (*Sitophilus oryzae* alpha-amylase) and TlGAMY (*Thermomyces lanuginosus* glucoamylase). Based on a functional screening of large libraries using the 10

SPs and computational analysis of the SP characteristics, a consensus sequence for potentially robust synthetic SP for use in secretory overexpression in *Y. lipolytica* was suggested (Celińska *et al.* 2018). Besides the optimization of SP signal sequences, the overexpression of translocon components and stabilization of cytosolic chaperones can be considered alternative strategies to enhance translocation efficiency (Tang *et al.* 2015).

ENGINEERING OF THE PROTEIN SECRETION PATHWAY

Following translocation to the ER, the next step in the secretory pathway is protein folding (Fig. 2B). One of the main limiting factors in secretory recombinant protein production is the protein-folding capacity in the yeast ER. In particular, for proteins with complex tertiary structures, such as antibodies, a major hindrance lies in the yeast ER as it is not equipped for efficient folding of overexpressed complex proteins originating from high eukaryotes. Compared to animal cells, yeast cells possess a lower number of chaperones and display limited extension of ER membrane, hindering their use as a practical host for industrial production of full-length antibodies. Misfolding of nascent peptides or overload of secretory proteins leads to a luminal burden, resulting in ER stress, which activates the unfolded protein response (UPR), which, in turn, induces multiple protective cellular events involved in proper protein folding and misfolded protein degradation (Fig. 2C) and ER-associated degradation (ERAD) of misfolded proteins (Payne *et al.* 2008). Thus, correct protein folding in the yeast ER is very important as both the UPR and ERAD govern the flux of proteins through the ER to the Golgi apparatus for secretion.

ER engineering by modulating UPR and ERAD activity

Several approaches, including the overexpression of protein-folding factors (chaperones) or redox enzymes in the ER, have been evaluated to manipulate the ER luminal environment with the aim of enhancing secretion capacities in *S. cerevisiae*, *P. pastoris* and *K. lactis* (Kim, Yoo and Kang 2015). For example, overexpression of chaperone BiP (Kar2p) or protein disulfide isomerase Pdi1p has been used to assist protein folding. Moreover, activation of the entire UPR pathway through the overexpression of Hac1p, a master transcription factor regulating this pathway, has been attempted in several yeast species. It was expected that the systematic manipulation of the UPR by Hac1p would be an excellent strategy to develop yeast cell factories for the secretory production of correctly folded and processed recombinant proteins. Hac1p overexpression did increase the expression of a set of ER chaperones, thus improving the secretory production of some heterologous proteins (Valkonen, Penttila and Saloheimo 2003) and secreted, surface-displayed and membrane proteins in *P. pastoris* (Guerfal *et al.* 2010). However, depending on the target protein, Hac1p expression either decreased the amount of heterologous proteins or had only a modest to no effect on the expression level, suggesting that the usability of Hac1p overexpression needs to be evaluated on a case-by-case basis (Valkonen, Penttila and Saloheimo 2003; Gasser *et al.* 2006; Guerfal *et al.* 2010). Notably, in *H. polymorpha*, overexpression of active *Hp*Hac1p significantly increased the *N*-linked glycosylation efficiency and tunicamycin resistance, suggesting that the modulation of Hac1p activity is an attractive strategy for efficient glycosylation in this yeast species (Moon *et al.* 2015). Modulation of the UPR by exploiting a regulator upstream of Hac1p has also been considered for ER engineering. Ire1p is an essential component of the UPR pathway important for sensing and responding to ER stress. As a transmembrane protein, Ire1p mediates the UPR by regulating Hac1p synthesis through *HAC1* mRNA splicing (Fig. 2C). In a recent study, *IRE1* overexpression in a mutant strain lacking Ypt32p, which mediates intra-Golgi traffic and the budding of post-Golgi vesicles, increased HBsAg expression by 2.12-fold, compared to the level in the wild-type strain (Sheng, Flick and Feng 2017), suggesting *IRE1* as another target gene for modulating UPR activity in host-cell engineering.

In an effort to reduce the clearance of a recombinant protein by ERAD, deletion of *HTM1*, *YOS9*, *HRD1*, *HRD3* or *UBC7* (Fig. 2B) in the presence and absence of Ire1p has been evaluated. Based on the observation that Δ*yos9*Δ*ire1* and Δ*htm1*Δ*ire1* yeast strains showed a delayed IgG clearance from the cells, it was proposed that removal of components involved in the generation and recognition of the glycan signal needed for ERAD-mediated protein degradation might be a strategy to increase the IgG ER residence time (de Ruijter and Frey 2015). Schuck *et al.* (2009) previously reported a noticeable new strategy to induce ER expansion in yeast by deleting the negative lipid-regulatory gene *OPI1*, which resulted in upregulation of *de novo* lipid synthesis and, consequently, ER expansion. In a recent study, a folding factor overexpression library in the background of the Δ*opi1* ER-enlarged yeast strain was screened to identify proteins that enhance the secretion of antibodies (de Ruijter, Koskela and Frey 2016). Out of six proteins identified, peptidyl-prolyl isomerase Cpr5p exerted the strongest beneficial effect on specific product yield, whereas Pdi1p, Ero1p, Kar2p, Lhs1p and Sil1p had limited or even negative effects on antibody secretion efficiency. By combining the Δ*opi1* strain with *CPR5* overexpression, the specific antibody product yield was boosted by a factor of 10 relative to the non-engineered strain. This is a good example of an effective strategy for ER engineering by combining rational strain design and high-throughput screening applications. A very recent study reported a remarkable improvement in the secretion efficiency by combining several engineering strategies of the early secretory pathway in *S. cerevisiae* (Besada-Lombana and Da Silva 2019). First, the efficient co-translational translocation into the ER was achieved via secretion peptide engineering by fusing the signal peptide of the alpha subunit from the oligosaccharyltransferase (OST) complex of the ER lumen (Ost1p) to the pro-region of the MFα1 leader sequence, generating a hybrid Ost1-pro-MFα1 leader. To further optimize the early secretory pathway, two additional strategies for ER engineering were combined: (i) increasing secretory pathway capacity by expanding the ER via deletion of *PAH1*, which encodes a cytoplasmic phosphatidate phosphatase required for lipid droplet formation and (ii) limiting ER-associated degradation via deletion of *DER1* and overexpression of *ERV29*. *DER1* encodes an ERAD transmembrane protein responsible for inserting aberrant proteins into the ER membrane for degradation, whereas *ERV29* encodes the ER multispanning transmembrane receptor for efficient ER exit and protein packing into coat protein complex II (COPII) vesicles. Finally, novel terminators derived from the 3′ untranslated region (3′UTR) of Pmp1 and Pmp2, which are co-translationally translocated proteins, were employed to enhance translocation process via SRP pre-recruitment. The optimal combination of pathway interventions was shown to be varied for each of three different proteins, a bacterial endoglucanase, a fungal β-glucosidase and a single-chain antibody fragment, with the highest overall increases in the range from 5.8 to 11-fold. This is a good example of multilevel approaches to

improve multiple critical steps of the early secretory pathway, focusing on ER transit and processing.

Engineering of the intracellular protein trafficking pathway

Accumulating recent evidence strongly indicates that the vesicle trafficking capacity plays an important role in the expression of membrane or secretory proteins. Modulation of vesicular trafficking, such as the upregulation of proteins involved in ER-to-Golgi and Golgi-to-plasma membrane vesicle trafficking, has been shown to improve protein secretion in yeast. Soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins are essential components of the yeast protein trafficking machinery. The SNARE proteins are required in the majority of membrane fusion events in the cell, where they facilitate fusion between the protein transport vesicles, the various membrane-enclosed organelles and, ultimately, the plasma membrane (Fig. 2D). Overexpression of native *S. cerevisiae* ER-to-Golgi SNARE genes increased heterologous cellulase secretion (Van Zyl, Den Haan and Van Zyl 2016).

Generally, ER-to-*cis*-Golgi transport of correctly folded precursor proteins is mediated by COPII-coated vesicles, which are composed mainly of Sar1p, Sec16p, Sec23p/Sec24p and Sec13p/Sec31p, and bud at the ER exit sites, whereas COPI vesicles are responsible for retrotransport to *cis*-Golgi and ER (Fig. 2D). It was shown that the moderate expression of *SEC16* enhances protein secretion by increasing the number of ER exit sites, thus decreasing ER stress (Bao *et al.* 2017). This change increases COPII formation, facilitating the translocation of α–amylase from ER to Golgi. The moderate overexpression of *SEC16* was also applied for the increase of other recombinant protein production, such as endoglucanase I from *Trichoderma reesei* and glucan-1,4-α-glucosidase from *Rhizopus oryzae*, demonstrating the feasibility as a general strategy for increasing the secretion of recombinant proteins. Meanwhile, a study using microfluidic screening and whole-genome sequencing identified several genes involved in the secretory and trafficking pathways, including *ECM3*, *EMC1*, *ERV29*, *GOS1*, *VPS5*, *TDA3*, *COG5* and *SNC2*, mutation of which greatly affected the secretion capacity (Huang *et al.* 2015). As expected, the deletion of trafficking genes, such as *ERV29* and *COG5*, the products of which are associated with COPII secretion vesicles, decreased α-amylase secretion. Noticeably, deletion of *GOS1*, involved in the retrograde Golgi traffic, increased the secretion efficiency. This finding clearly indicated that anterograde trafficking in the Golgi also represents a potential pathway to be engineered for achieving balanced trafficking within the Golgi. More interestingly, a follow-up study showed that deletion of *VPS5* and *VPS17*, the products of which are involved in the transport between the Golgi and endosome, substantially enhanced heterologous protein secretion, while reducing intracellular heterologous protein retention (Huang *et al.* 2018). Vps5p forms a retromer subcomplex with Vps17p and interacts with another subcomplex consisting of Vps26p, Vps29p and Vps35p, all involved in protein trafficking from the late endosome to the late Golgi. The increase in amylase production as a consequence of *VPS5* and *VPS17* deletion supported the notion that modification of endosome-to-Golgi trafficking can effectively reduce protein retention besides increasing protein secretion.

A different approach to engineer vesicle trafficking was reported by Tang *et al.* (2017), in which the components of the vesicle trafficking from ER to Golgi (Sec12p, Sec13p, Erv25p and Bos1p) and those from Golgi to the plasma membrane (Sso1p, Snc2p, Sec1p, Exo70p, Ypt32p and Sec4p) were overexpressed to improve the extracellular activity and surface display efficiency of cellulases. While overexpressing, the targeted components in the ER to Golgi enhanced the extracellular activity of *Clostridium thermocellum* endoglucanase (CelA); only Sec13p overexpression led to the increased secretion of *Saccharomycopsis fibuligera* β-glucosidase (BGL1). In contrast, overexpression of the components in the Golgi to plasma membrane exhibited better performance in increasing BGL1 secretion and extracellular activity, indicating that engineering vesicle trafficking has protein-specific effects. Intriguingly, the engineering of the vesicle trafficking components from ER to Golgi improved the display efficiency of CelA and BGL1 fused with a-agglutinin. These results strongly support engineering of the vesicle trafficking pathway as a promising avenue for the design of efficient cell factories for protein production both for secretion and display.

Engineering of the vacuolar sorting pathways

Several recent studies reported that *vps* mutants impaired in vacuolar sorting can be used for more efficient production of various recombinant proteins, indicating that prevention of mislocalization into the vacuole improves secretion. In particular, disruption of subunits of the class C core vacuole/endosome tethering complex enhanced the secretion levels of recombinant proteins in several yeasts (Marsalek *et al.* 2017). Simultaneous disruption of a vacuolar sorting pathway and vacuolar proteases, such as deletion of *PRB1*—encoding the major cellular serine protease—in Δ*vps8* and Δ*vps21* strains, was shown to prevent missorting of recombinant proteins, resulting in higher secreted heterologous protein levels in *P. pastoris*. Disruption of genes encoding vacuolar sorting components of the homotypic fusion and protein sorting complex, especially *YPT7*, also led to enhanced secretion of recombinant proteins in *P. pastoris*. Additional overexpression of Sbh1p, a subunit of the ER translocation pore, significantly increased extracellular Fab levels by 4-fold in the Δ*ypt7* strain (Marsalek *et al.* 2019). These findings strongly indicated that vacuolar missorting of recombinant proteins occurs in the Golgi-to-vacuole sorting pathways and thus, likely affects correctly folded active proteins in yeast cells. Therefore, disruption of these pathways alone or in combination with modification of early steps in the vacuolar pathway might improve heterologous protein secretion. Furthermore, it was recently reported that mutation of Mtc6p, a protein involved in autophagy regulation, attenuates autophagy and improves secretory expression of heterologous proteins in *Kluyveromyces marxianus* (Liu *et al.* 2018b). These results suggest that attenuating or interdicting autophagy would be another strategy to dramatically improve protein yields in yeast.

GLYCOSYLATION PATHWAY ENGINEERING IN YEAST

Among various post-translational modifications, glycosylation is the most complex modification. Although yeasts can perform *N*- and *O*-glycosylation of proteins, the final glycosylation patterns of yeast proteins significantly differ from those of human proteins, which can lead to potentially allergenic glycan structures of recombinant proteins destined for therapeutic application. Yeasts only synthesize *N*-oligosaccharide chains of the

high mannose type and are unable to synthesize oligosaccharide chains of the mammalian complex type. High-mannose-type glycans and terminal α-1,3-mannose linkages associated with glycoproteins produced in *S. cerevisiae* are known to cause short serum half-lives or even immunogenicity of therapeutic proteins (De Pourcq, De Schutter and Callewaert 2010; Walsh 2010). Recent studies on *N*-glycan structures in various yeast species have revealed that the *N*-glycosylation patterns in non-conventional yeasts generally are much simpler than those in the traditional yeast *S. cerevisiae* (Thak *et al.* 2018). While *S. cerevisiae* shows a tendency to hyperglycosylate proteins by the addition of >40 mannose residues, non-*Saccharomyces* yeasts, including *P. pastoris, H. polymorpha* and *Y. lipolytica*, produce glycoproteins with an average mannose chain length of 8 to 14 without hyperimmunogenic terminal α-1,3-linked mannose residues. The structures of *O*-glycans are quite diverse among eukaryotic organisms and even within yeast species. In most yeasts, an α-1,2-linked mannotriose forms the major common core of *O*-glycans, which is differentially expanded depending on the species (Thak *et al.* 2018). Intensive attempts have been made to develop intelligent yeast strains endowed with the ability to synthesize human-type glycans by deleting yeast-specific genes involved in glycan processing and, subsequently, introducing heterologous genes involved in human-type glycan biosynthesis in various yeast species (Table 1); these will be discussed in detail below.

Glycoengineering for humanized N-glycosylation pathways in yeast

Attempts to implement humanized *N*-glycosylation pathways have been reported in not only *S. cerevisiae* but also various non-conventional yeast species (Fig. 3A). As the first step, remodeling of the glycosylation pathway in yeast to produce human compatible high-mannose type *N*-glycans ($Man_5GlcNAc_2$) was achieved by deleting *OCH1* for the removal of yeast-specific outer-chain biosynthesis of *N*-glycans in *S. cerevisiae* (Chiba *et al.* 1998), *H. polymorpha* (Kim *et al.* 2006), *K. lactis* (Liu *et al.* 2009), *P. pastoris* (Choi *et al.* 2003) and *Y. lipolytica* (Song *et al.* 2007), and subsequent introduction of α-1,2 mannosidase from *Aspergillus saitoi* (Chiba *et al.* 1998; Kim *et al.* 2006; Cheon *et al.* 2012), *Caenorhabditis elegans* (Choi *et al.* 2003; Hamilton *et al.* 2003) and *T. reesei* (De Pourcq *et al.* 2012b) in the ER for mannose trimming. In *S. cerevisiae* and *K. lactis*, additional deletion of *MNN1* was required to block the addition of immunogenic α-1,3-mannose residues at the end of *N*-glycans (Chiba *et al.* 1998; Liu *et al.* 2009; Piirainen *et al.* 2016). To prevent yeast-specific mannosylphosphorylation of *N*-linked oligosaccharides, disruption of *MNN4* or its homologs (*PNO1*, *MPO1*) was carried out as optional modification in *S. cerevisiae* (Chiba *et al.* 1998), *P. pastoris* (Hamilton *et al.* 2006) and *Y. lipolytica* (Park *et al.* 2011). In *P. pastoris*, additional deletion of four genes from β-mannosyltransferase family (*BMT1,2,3,4*) was carried out to eliminate *P. pastoris*-specific β-mannose structure (Hopkins *et al.* 2011). Subsequently, genes encoding β-1,2-*N*-acetylglucosaminyltransferase I and II (GnTI, GnTII), along with the gene encoding mannosidase II (MnsII) for the mannose trimming of *N*-glycans in the Golgi, were introduced for the addition of *N*-acetylglucosamine in the Golgi to produce human type *N*-glycans containing *N*-acetylglucosamine as a terminal sugar in the glycoengineered *H. polymorpha* (Cheon *et al.* 2012) and *P. pastoris* (Choi *et al.* 2003). For efficient supply of uridine diphosphate *N* -acetylglucosamine (UDP-GlcNAc) into the Golgi lumen, the introduction of *K. lactis* gene coding for UDP-GlcNAc transporter was optionally performed (Choi *et al.* 2003; Hamilton *et al.* 2003; Wang *et al.* 2013; Piirainen *et al.* 2016). To implement the next galactose addition step, the galactosyltransferase (GalT) gene was introduced into the glycoengineered *P. pastoris* to produce human compatible complex-type glycans containing galactose as a terminal sugar (Hamilton *et al.* 2006). For optimal expression of these genes, precise spatial positioning along the ER-Golgi secretory pathway is crucial for successful assembly of the foreign glycosylation pathway. For functional targeted expression at the yeast Golgi membrane, semi-synthetic glycosyltransferase and glycosidase libraries composed of several combinations of C-terminal catalytic domains and transmembrane fragments were constructed, and these combinatorial libraries were used for high-throughput screening of *H. polymorpha* and *P. pastoris*-engineered strains harboring optimal foreign glycosylation pathways (Choi *et al.* 2003; Cheon *et al.* 2012). Finally, pioneering glycoengineering work to produce completely humanized glycoproteins with terminal sialic acids was performed in *P. pastoris* by the additional introduction of a *Mus musculus* α-2,6-sialyltransferase (SAT) gene along with five synthetic genes to implement the sialic acid biosynthesis pathway (Hamilton *et al.* 2006) (Table 1). This synthetic sialylation pathway in yeast requires the introduction of heterologous genes for the conversion to cytidine monophosphate (CMP)-sialic acid from endogenous UDP-GlcNAc by sequential reactions mediated by UDP- *N*-acetylglucosamine-2-epimerase/*N*-acetylmannosamine kinase (GNE), *N*-acetylneuraminate-9-phosphate synthase (SPS), sialylate-9-P phosphatase (SPP) and CMP-sialic acid synthase (CSS). Subsequently, CMP-sialic acid is transported to the Golgi lumen via CMP-sialic acid transporter (CST) and added to *N*-glycans assembled on proteins by SAT (Hamilton *et al.* 2006).

As a strategy to remove the requirement of MnsII for the trimming of mannose residues in *N*-glycan humanization in yeast, the *ALG3* gene, encoding a dolichylphosphate-mannose-dependent α-1,3 mannosyltransferase in the ER, was deleted in *P. pastoris*, *H. polymorpha* and *Y. lipolytica* (Davidson *et al.* 2004; Cheon *et al.* 2012; De Pourcq *et al.* 2012a). The resulting artificial glycosylation pathway generates protein-bound $Man_3GlcNAc_2$, which serves as a direct substrate for GnTI. Introduction of the Golgi apparatus-targeted heterologous glycosyltransferases GnTI, GnTII and GalT in *P. pastoris* led to successful production of complex humanized glycoproteins with a terminal galactose (Bobrowicz *et al.* 2004). *ALG3* deletion in an *och1* mutant background resulted in the secretion of homogeneously glycosylated proteins bearing a predicted $Man_5GlcNAc_2$ structure, which was used as a substrate by GnTI to produce glycoproteins bearing human hybrid-type *N*-glycans with a terminal *N*-acetylglucosamine (Cheon *et al.* 2012). Furthermore, *ALG3/ALG11* double deletion has been used in *S. cerevisiae* to generate a trimannosyl core glycan in humanized *N*-glycan structures, which was further processed by human GnTI and GnTII to produce glycoproteins bearing human hybrid-type *N*-glycans with a terminal *N*-acetylglucosamine (Nasab *et al.* 2013). A similar glycosylation pathway modification was made in the *ALG3/ALG11* deletion background to express glycoproteins with complex human-like glycans by eliminating *OCH1* in *H. polymorpha* (Wang *et al.* 2013). Subsequent introduction of three foreign glycosyltransferases genes, including human GnTI, rat GnTII and human GalTI, into this *H. polymorpha* $\Delta alg3/\Delta alg11/\Delta och1$ strain allowed the production of glycoproteins bearing complex human-like glycans with a terminal galactose. Piirainen *et al.* (2016) deleted *MNN1* in the *S. cerevisiae* $\Delta alg3/\Delta alg11$ strain to eliminate interfering structures, generating $\Delta alg3/\Delta alg11/\Delta mnn1$, and addi-

Table 1. Target genes for humanized *N*-/*O*-glycosylation pathway engineering in yeasts.

Category		Gene	Protein (function)	Purpose	Ref.
N-glycan engineering	Deletion	*ALG11*	α-1,2-mannosyltransferase in the ER	Blocking lipid-linked *N*-glycan biosynthesis	Nasab *et al.* (2013), Wang *et al.* (2013), Piirainen *et al.* (2016)
		ALG3	α-1,3-mannosyltransferase in the ER	Blocking lipid-linked *N*-glycan biosynthesis	Bobrowicz *et al.* (2004), Davidson *et al.* (2004), Cheon *et al.* (2012), De Pourcq *et al.* (2012a), Nasab *et al.* (2013), Wang *et al.* (2013), Piirainen *et al.* (2016)
		OCH1	α-1,6-mannosyltransferase in the Golgi	Blocking outer chain biosynthesis of *N*-glycans	Chiba *et al.* (1998), Choi *et al.* (2003), Hamilton *et al.* (2003), Bobrowicz *et al.* (2004), Kim *et al.* (2006), Song *et al.* (2007), Liu *et al* (2009), Cheon *et al.* (2012), De Pourcq *et al.* (2012b), Wang *et al.* (2013)
		MNN4	Mannosylphosphate transferase	Blocking phosphorylation of *N*-/*O*-linked glycans	Chiba *et al.* (1998), Hamilton *et al.* (2006), Park *et al.* (2011)
		MNN1	α-1,3-mannosyltransferase in the Golgi	Blocking addition of immunogenic terminal α-1,3-mannose	Chiba *et al.* (1998), Liu *et al* (2009), Hamilton *et al.* (2013), Piirainen *et al.* (2016)
	Addition	*Leishmania major STT3D*	Oligosaccharyltransferase	Increasing *N*-glycosylation efficiency	Choi *et al.* (2012), Nasab *et al.* (2013)
		Aspergillus saitoi msdS	α-1,2-mannosidase (α-1,2-MNS)	Trimming alpha-1,2-linked mannose residues	Chiba *et al.* (1998), Kim *et al.* (2006), Cheon *et al.* (2012)
		Caenorhabditis elegans C52E4.5			Choi *et al.* (2003), Hamilton *et al.* (2003)
		Trichoderma reesei TRIREDRAFT_64 285			De Pourcq *et al.* (2012a, 2012b)
		Kluyveromyces lactis YEA4	UDP-GlcNAc transporters in the Golgi	Increasing abundance of a complex-type glycans	Choi *et al.* (2003), Hamilton *et al.* (2003), Wang *et al.* (2013), Piirainen *et al.* (2016)
		Homo sapiens MGAT1	N-acetylglucosaminyl transferase I (GnT-I)	Generating hybrid *N*-glycans	Choi *et al.* (2003), Hamilton *et al.* (2003), Cheon *et al.* (2012), Nasab *et al.* (2013), Wang *et al.* (2013)
		Drosophila melanogaster alpha-Man-Ia	Mannosidase II (MnsII) in the Golgi	Trimming α-1,2-linked mannose residues	Hamilton *et al.* (2003), Bobrowicz *et al.* (2004)
		Rattus norvegicus or H. sapiens MGAT2	β-1,2-N-acetyl glucosaminyltransferase II (GnT-Il) in the Golgi	Generating complex *N*-glycans	Hamilton *et al.* (2003), Bobrowicz *et al.* (2004), Nasab *et al.* (2013), Wang *et al.* (2013)
		Schizosaccharomyces pombe UGE1	UDP-galactose 4-epimerase	UDP-galactose pool in the cytosol	Bobrowicz *et al.* (2004), Wang *et al.* (2013)
		D. melanogaster	UDP-Gal transporter in the Golgi	UDP-galactose pool in the Golgi lumen	Hamilton *et al.* (2006)
		H. sapiens B4GALT1	β-1,4-galactosyltransferase (GalT) in the Golgi	Producing *N*-glycans containing galactose	Wang *et al.* (2013)
O-glycan engineering	Deletion	*PMT1-PMT6*	Protein-*O*-mannosyltransferase	Blocking yeast specific *O*-mannosylation	Amano *et al.* (2008)
	Addition	*Bacillus subtilis GalE*	UDP-galactose 4-epimerase	UDP-Gal/GalNAc pool in the cytosol	Amano *et al.* (2008)
		Arabidopsis thaliana MUR1, FX/GER1	GDP-mannose 4,6 dehydratase 2, GDP-L-fucose synthase 1	UDP-fucose pool in the cytosol	Chigira *et al.* (2008)
		H. sapiens UGT2	UDP-Gal transporter in the Golgi	UDP-galactose pool in the Golgi	Amano *et al.* (2008)
		H. sapiens GFR	GDP-fucose transporter in the Golgi lumen	UDP-fucose pool in the Golgi	Chigira *et al.* (2008)

Table 1. Continued

Category		Gene	Protein (function)	Purpose	Ref.
		Mus musculus Pomgnt1	Protein-O-linked-mannose β-1,2-*N*-acetyl glucosaminyltransferase1 (POMGnT1) in the Golgi	Production of sialylated α-dystroglycan type O-glycans	Hamilton *et al.* (2013)
		D. melanogaster C1GalTA	Core1-β-1,3-galactosyl transferase (core β1,3-GalT) in the Golgi	Generating core1 O-glycan structures	Amano *et al.* (2008)
		H. sapiens POFUT1	GDP-fucose protein O-fucosyltransferase 1 (O-FucT-1) in the Golgi	Generating fucose O-mannosylation	Chigira *et al.* (2008)
		H. sapiens ppGalNAcT	Polypeptide:N-acetyl galactosaminyltransferase 1,2 or 3 in the Golgi	Generating mucin-type-O-glycans	Amano *et al.* (2008)
		H. sapiens MFNG	β-1,3-*N*-acetyl-glucosaminyltransferase (Fringe) in the Golgi	Elongation of O-fucose glycans	Chigira *et al.* (2008)
Sialylation pathway	Addition	*H. sapiens GNE*	UDP-N-acetyl glucosamine-2epimerase/N-acetylmannosamine kinase (GNE)	ManNAc-6-P pool in the cytosol	Hamilton *et al.* (2013)
		H. sapiens NANS	N-acetylneuraminate-9-phosphate synthase (SPS)	Sialic acid pool in the Golgi	Hamilton *et al.* (2006, 2013)
		H. sapiens CMAS	CMP-sialic acid synthase (CSS)	Sialic acid pool in the Golgi	Hamilton *et al.* (2006, 2013)
		M. musculus Slc35a1	CMP-sialic acid transporter (CST) in the Golgi	Production of sialylated α-dystroglycan type O-glycan	Hamilton *et al.* (2006, 2013)
		M. musculus ST6Gal1	α-2,6-sialyltransferase (SAT) in the Golgi	Generation of sialylated complex-type N- and α-dystroglycan-type O-glycans	Hamilton *et al.* (2006, 2013)

tionally expressed the *K. lactis* UDP-*N*-acetylglucosamine transporter to increase the import of UDP-*N*-acetylglucosamine into the Golgi apparatus, which remarkably improved the glycan pattern of the engineered strain in terms of both glycoform homogeneity and complex-type glycosylation efficiency. The glycoengineered yeast strains, particularly *P. pastoris* strains based on *OCH1* deletion, have been employed to produce various therapeutic glycoproteins, such as monoclonal antibody targeting human epidermal growth factor receptor 2 (Herceptin) and human interferon (IFN) alpha 2b (Liu *et al.* 2018a; Baghban *et al.* 2019; Katla *et al.* 2019). The results validate the advantage of glycoengineered yeast strains over mammalian cell lines as the host for the production of recombinant glycoproteins to generate glycoforms with greater homogeneity in higher titers and lower cost.

In silico yeast models for N-glycosylation engineering

With the increasing contribution of computational models, which describe *in silico* metabolic pathways in yeast cells, to the identification of new targets or strategies for further metabolic engineering (Kerkhoven, Lahtvee and Nielsen 2015), the development of *in silico* yeast models for *N*-glycosylation engineering has also been attempted. An initial metabolic model for *P. pastoris N*-glycosylation was constructed by incorporation of a subset of the metabolic reaction present in *P. pastoris* only and used for the flux analysis for glycosylated erythropoietin production (Eskitoros, Ata and Çalık 2014). The more advanced *in silico* model, *ihGlycopastoris,* was developed by extension of the previous genome-scale metabolic model of *P. pastoris*, iLC915, with manual curating to render the model more robust and subsequent incorporation of both *P. pastoris* native and humanized *N*-glycosylation pathways into the updated model (Irani *et al.* 2016). The *in silico* glycoengineered *P. pastoris, ihGlycopastoris,* was used to predict the production yield of humanized glycosylated proteins. The reduction in theoretical yield was predicted in proportion to the size of the recombinant protein and their number of *N*-glycosylation sites; smaller proteins and more glycosylation sites result in a larger reduction in yield, while the precise effect is dependent on *N*-glycosylation level (number of sites/number of amino acids). Simulation using different carbon sources also indicated that theoretical protein yields vary depending on carbon source. Furthermore, *ihGlycopastoris* was used to predict the potential amplification targets for increased protein yield, identifying *N*-glycan metabolism and the biosynthetic pathway of precursors UDP-GlcNAc and GDP-Man as the targets to increase production.

The current models represent a step toward a more complete construction of *in silico* humanized secretory pathway in yeast, which will be powerful tools to design and construct synthetic secretory pathways with improved secretion and glycosylation capacity in yeast cells. Further improvement of *in silico* yeast models for glycoprotein production is expected by iterative development of this model by incorporation of other processes in the secretory pathway, such as protein folding, degradation and targeting, with consideration of the nature of the protein including its hydrophobicity, arrangement and composition of amino acids.

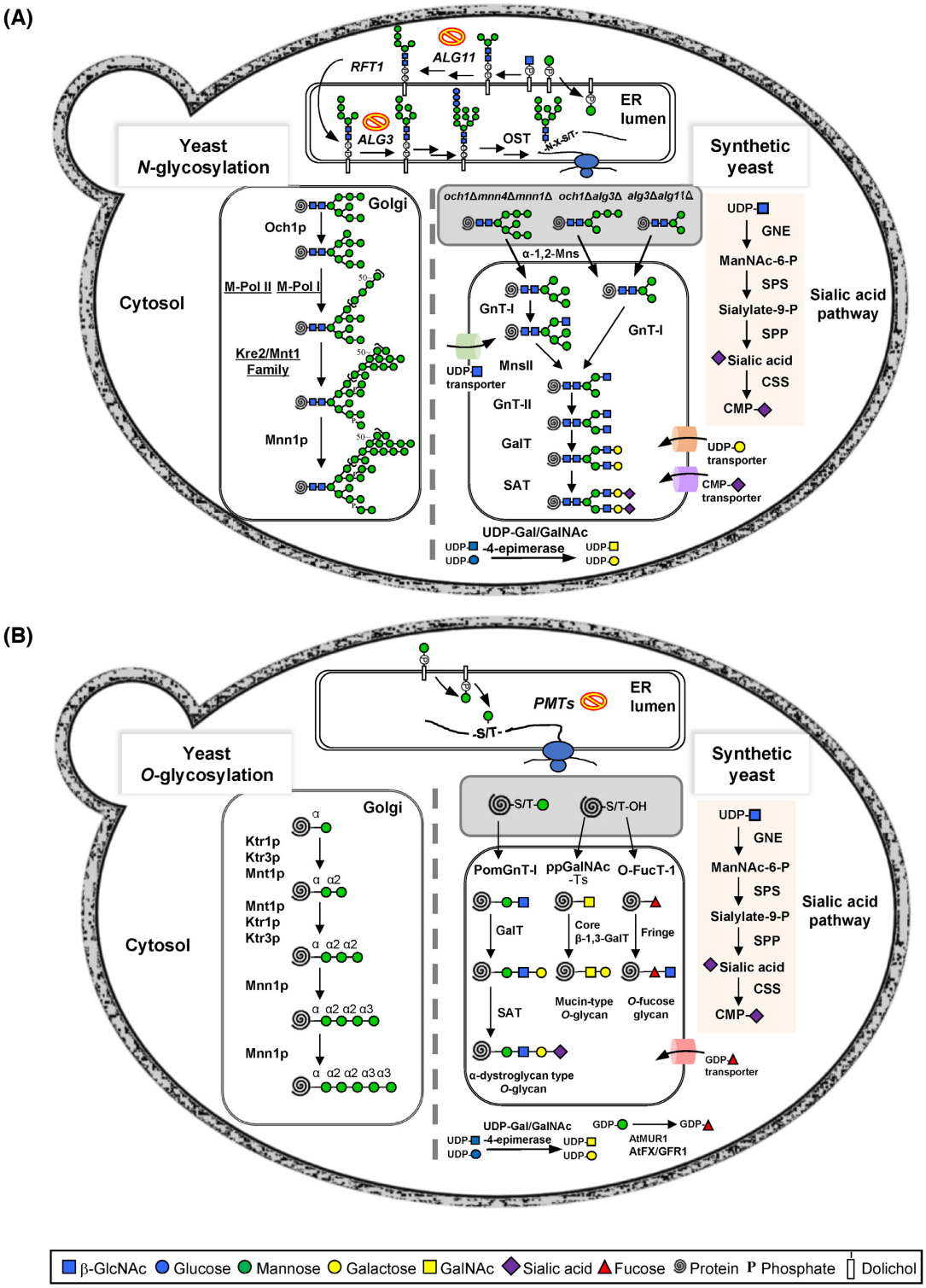


Figure 3. Synthetic protein glycosylation pathways to produce humanized glycans in yeast. **(A)** Endogenous and humanized *N*-glycosylation pathways in yeast. Dolichol-linked oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$-PP-Dol) is synthesized by the serial attachment of monosaccharides mediated by the *ALG* gene family and attached to a protein in the ER by OST. The outer chain of *N*-glycans assembled on proteins in the Golgi is initiated by Och1p and, subsequently, elongated by several glycosyltransferases, resulting in hypermannosylation. For humanized *N*-glycosylation, yeast-specific genes responsible for outer chain biosynthesis, *OCH1*, *MNN4* and *MNN1* in *S. cerevisiae*, are eliminated, or genes involved in lipid-linked oligosaccharide biosynthesis, *ALG3* and *ALG11*, are deleted to generate truncated core *N*-glycans in the ER. Further modification toward humanized *N*-glycans is achieved by expressing several heterologous glycosyltransferases. **(B)** Endogenous and humanized *O*-glycosylation pathways in yeast. Biosynthesis of *O*-glycans is initiated in the ER by the *PMT* family genes. In the Golgi, *O*-glycans are extended as linear short mannose chains by several mannosyltransferases. To mimic human *O*-glycosylation to produce mucin-type *O*-glycans and *O*-fucose glycans, several glycosyltransferase genes specific for each type of glycosylation were introduced into yeast cells in which the *PMT* family genes were disrupted. To produce dystroglycan-type *O*-glycan, the single *O*-linked mannose added in the ER is extended and capped with sialic acid in glycoengineered yeast cells equipped with the synthetic sialylation pathway.

Glycoengineering for humanized O-glycosylation pathways in yeast

Glycoengineering of yeasts has also been targeted toward humanized O-glycosylation, which is characterized by relatively short, but varied, glycan structures with the addition of several different sugars to the serine or threonine (S/T) residues (Fig. 3B). Glycoengineered *S. cerevisiae* capable of producing mucin-type O-glycosylation was constructed by introducing a set of heterologous enzymes, absent in yeast, such as polypeptide *N*-acetylgalactosaminyltransferase 1 (ppGalNAc-T1), core1 β1–3GalT, UDP-Gal/GalNAc transporter and UDP-Gal/GalNAc 4-epimerase (Amano *et al.* 2008) into the mutant strains generated by deletion of *PMT* genes encoding protein O-mannosyltransferase (Pmt) family. The *pmt* deletion strains are blocked in the first addition step of a mannose residue from Dol-P-Man to S/T residue (Amano *et al.* 2008). An artificial O-fucosylation system in *S. cerevisiae* was also constructed by introducing genes that encode protein O-fucosyltransferase 1 (O-FucT-1), along with genes whose protein products convert cytoplasmic GDP-mannose to GDP-fucose from intrinsic GDP-mannose in the yeast cytosol and transport GDP-fucose into the Golgi lumen, in the *pmt* deletion background (Chigira *et al.* 2008). Efficient conversion of GDP-mannose to GDP-fucose in *S. cerevisiae* was achieved by co-expression of *A. thaliana MUR1*, which encodes GDP-mannose-4,6-dehydratase, and *AtFX/GER1*, which encodes GDP-4-keto-6-deoxy-mannose-3,5-epimerase/4-reductase. Subsequent introduction of Fringe facilitated the addition of *N*-acetylglucosamine, resulting in O-fucose glycosylation type. For production of human-like sialylated O-linked glycans structurally similar to α-dystroglycan-type glycoforms, foreign genes encoding an α-1,2-mannosidase, for trimming of mannose residues, and an enzyme protein-O-linked-mannose β-1,2-*N*-acetylglucosaminyltransferase 1 (PomGnT1), to attach an *N*-acetylglucosamine residue to O-mannose chains, were introduced into a glycoengineered *P. pastoris* strain (Δ*och1*, Δ*pno1*, Δ*mnn4b*, Δ*bmt1–4*) equipped with a heterologous sialic acid biosynthesis pathway and harboring heterologous enzymes GalT and SAT for addition of galactose and sialic acid (Hamilton *et al.* 2013). The successful production of various types of O-linked glycans in yeasts extends their utility as production platforms beyond *N*-linked glycosylated biotherapeutics to include molecules possessing O-linked glycans.

Synthetic toolboxes to improve N-glycan site occupancy

N-glycan sites on recombinant glycoproteins produced in yeast can be underglycosylated by incomplete occupation of *N*-glycan sites, particularly when glycoproteins are overexpressed. The transfer of the oligosaccharide from the donor lipid onto the nascent polypeptide in the ER is mediated by oligosaccharyltransferase. Yeast OST complex consists of multiple subunits, including Ost1p, Ost2p, Ost3p, Ost4p, Ost5p, Ost6p, Stt3p, Swp1p and Wbp1p (Fig. 2A). The *N*-glycan site occupancy of secretory glycoproteins was greatly improved by overexpressing a single-unit protozoan OST, *Leishmania major* STT3D, in *P. pastoris* (Choi *et al.* 2012). The novel synthetic *N*-glycosylation pathways generated by *ALG3* or *ALG3/ALG11* deletion produce glycoproteins containing a truncated core glycan, which are poor substrates, posing a problem of decreased *N*-glycan site occupancy. Furthermore, truncated lipid-linked *N*-glycans are built up on the cytoplasmic side of the ER due to low flipping efficiency. Improvement of the site occupancy of the Δ*alg3*/Δ*alg11* strain was achieved by introducing artificial flippase and protozoan OSTs, which possess altered specificities for the oligosaccharide as well as for the protein acceptor site (Nasab *et al.* 2013). These results clearly indicate that flipping of the lipid-linked oligosaccharide into the ER lumen can be improved by an artificial flippase and that the transfer of the oligosaccharide from the donor lipid onto the protein can be optimized by a protozoan oligosaccharyltransferase (Piirainen *et al.* 2014).

Glycoengineered yeast strains for mannose-6-phosphate (M6P) modification

As M6P glycan is important for cellular uptake and lysosomal targeting, many therapeutic enzymes for lysosomal storage diseases require a high content of M6P glycan. As the bi-mannosylphosphorylated (mannose-1-phosphate-6-O-mannose) residues are generated in yeasts and can be converted to contain M6P by uncapping of the outer mannoses, several trials have been made to generate recombinant proteins with high M6P glycan content using glycoengineered yeast strains. The level of yeast-produced bi-mannosylphosphorylated glycoproteins was enhanced by overexpressing Mnn4p in Δ*och1* strains of *P. pastoris* and *Y. lipolytica* producing human-type high-mannose $Man_{8-9}GlcNAc_2$ *N*-glycans (Tiels *et al.* 2012). Recently, *Y. lipolytica MPO1*, encoding a Mnn4p homolog, was overexpressed in glycoengineered *S. cerevisiae* in which immunogenic glycan synthesis was abolished by deleting *OCH1* and *MNN1* (Gil *et al.* 2015). The heterologous overexpression of *Y. lipolytica MPO1* enhanced the level of *N*-glycans containing M6P-Man in *S. cerevisiae*. Through *in vitro* uncapping of the outer mannose residue, the bi-mannosylphosphorylated structure was changed to a bi-phosphorylated structure with high affinity for the M6P receptor. The glycopeptides containing M6P glycans derived from this glycoengineered yeast were further processed by uncapping and α(1,2)-mannosidase digestion and, then, conjugated to a recombinant acid α-glucosidase, which was shown to be successfully targeted to the lysosome for the treatment of lysosomal storage diseases (Kang *et al.* 2018). The superior ability of overexpressed yeast Mnn4p to increase bi-mannosylphosphorylated glycan levels in yeast shows promise for the production of therapeutic enzymes with improved lysosomal targeting capability in glycoengineered yeast cells.

CELL WALL ENGINEERING FOR ENHANCED PRODUCT RECOVERY

The yeast cell wall, a highly complex structure composed mainly of glucan, mannan and chitin, determines the shape of the yeast cell and protects it from its ever-changing environment (Orlean 2012). However, this thick and rigid cell wall is a major obstacle in yeast cell content recovery. Mutation of major cell wall biosynthetic and signal transduction pathway genes represents an alternative method to mechanical/chemical disruption, which are cheaper and more energy efficient for the recovery of yeast cell contents (Fig. 4). Downregulation of the expression of genes involved in the cell wall integrity pathway, such as *SRB1/PSA1*, *MPK1/SLT2* and *PKC1*, has been attempted by using the repressive promoter, p*MET3* (Zhang *et al.* 1999a, 1999b). *Saccharomyces cerevisiae* with conditional downregulation of *SRB1* and *PKC1* can serve as an oral delivery vehicle for therapeutic proteins in the human gut by using gene promoters whose transcript abundance is reduced under human gut conditions (Omara *et al.* 2010).

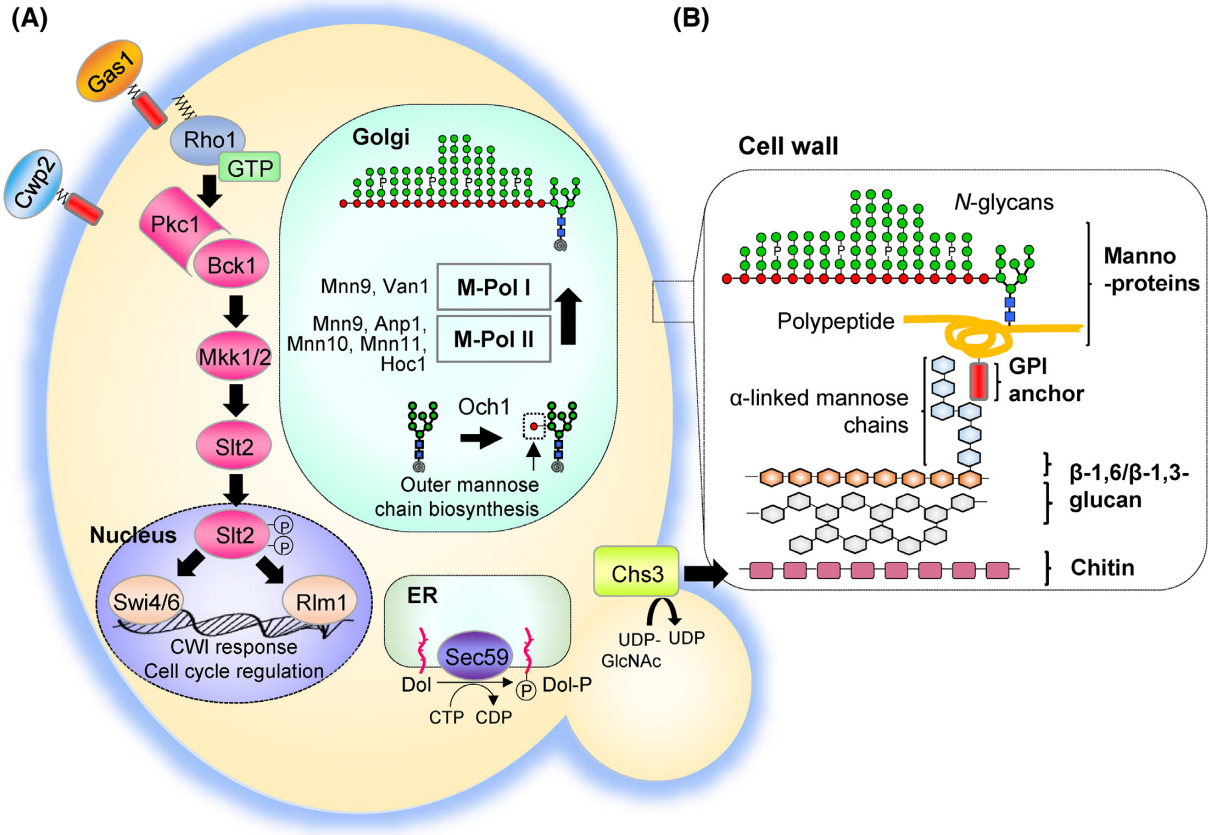


Figure 4. Engineering targets to modulate the biosynthesis of cell wall components. **(A)** Cell wall integrity pathways. Upon cell wall stress, Rho1p-GTP interacts with the sole yeast protein kinase C (Pkc1p), resulting in the activation of a conserved mitogen-activated protein kinase cascade. The transcription factor Rlm1p activates genes for the maintenance of cell wall integrity and Swi4/6p regulates cell cycle progression. **(B)** Structure of the yeast cell wall and its components. The cell wall consists of an external layer of cell wall proteins, mostly mannoproteins carrying glycosylphosphatidylinositol (GPI) anchors, and an internal layer of glucan (β-1,3- and β-1,6-glucan polymer) and chitin (an *N*-acetyl-glucosamine polymer). Dolichol phosphate (Dol-P) generated from dolichol (Dol) by Sec59p is used for the biosynthesis of the core *N*-glycan moiety in the ER. Outer chain biosynthesis of *N*-glycans on proteins is initiated by Och1p and the chains are extended by various mannosyltransferases. The cell wall chitin (90%) is synthesized by chitin synthase III, encoded by *CHS3*. GPI-anchored cell wall proteins, Gas1p and Cwp2p, play key roles in cell wall assembly and integrity.

Modulation of cell wall integrity by the removal of cell wall components

GAS1 encodes a β-1,3-glucanosyltransferase, a glycoprotein that is anchored to the plasma membrane by a glycosylphosphatidylinositol anchor and is necessary for correct assembly of cell wall polymers in *S. cerevisiae*. Due to changes in the composition and organization of cell wall constituents, the Δ*gas1* mutant exhibits morphogenetic defects, including atypical shape, defects in bud maturation and cell separation, and an altered growth rate (Popolo and Vai 1999). The use of *S. cerevisiae* Δ*gas1* allowed improved secretion of recombinant human insulin-like growth factor 1 (rhIGF1) by >7-fold in a fed-batch bioreactor (Vai *et al.* 2000). This result strongly indicated that alterations in the cell wall architecture, such as decreased cell wall binding of proteins, allow better accumulation of the heterologous protein in the medium. In accordance herewith, Li *et al.* (2020) recently reported that disruption of *CWP2*, encoding a cell wall mannoprotein, led to improved production of cellobiohydrolase in *S. cerevisiae*. Despite decreases in cell wall thickness and density of the outer cell wall, the *CWP2* disruption strain showed no difference in growth, but remarkably increased extracellular cellobiohydrolase activity when compared to the parental strain. Global transcriptome analysis revealed that genes involved in translation, such as tRNA genes, and ribosome biogenesis were upregulated, along with some genes related to the secretory pathway, including ER targeting, Golgi trafficking and vesicle transport. Moreover, deletion of *CWP2* appeared to alleviate ER stress and thus, suppress the UPR. Thus, this study demonstrated that yeast cell wall proteins can be good manipulation targets for heterologous protein production.

Cell wall engineering by modulating protein glycosylation

Findings in several studies strongly supported that changes in the cell wall structure and mannan composition could significantly affect the last stages of protein secretion. A recent study in *S. cerevisiae* reported that disruption of mannosyltransferases involved in outer chain biosynthesis of *N*-glycans, such as Och1p, Mnn9p and Mnn1p, improved the secretion of recombinant proteins, including β-glucosidase, endoglucanase and cellobiohydrolase (Tang *et al.* 2016). The cell wall integrity was disrupted in these strains, contributing to secretory protein release. Interestingly, certain key genes in the secretory pathway, including protein folding and vesicular trafficking, were upregulated

in these mutant strains. To recover the severe cell growth retardation and cell aggregation in *ScOCH1* deletion mutants, *RHO1* and *PKC1*, which are involved in the cell wall integrity signaling pathway, were additionally overexpressed (Tang *et al.* 2016). A very recent study reported the development of conditional lysis mutants of *S. cerevisiae* by expressing *OCH1* and *CHS3*, which are required for the synthesis of cell wall chitin, under the control of the *MET3* promoter (P_{MET3}) (Fig. 4) (Luu *et al.* 2019). Osmotic lysis of the conditional mutants cultivated with methionine was sufficient to release the intracellular recombinant protein, nodavirus capsid protein, with up to 60% efficiency, compared to glass bead breakage. In addition, the mutant strains achieved 3-fold enhanced secretion of a recombinant extracellular glycoprotein, *S. fibuligera* β-glucosidase, with markedly reduced hypermannosylation, particularly in the P_{MET3}-*OCH1* strain. Compared to the Δ*och1* strain, the conditional P_{MET3}-*OCH1* and P_{MET3}-*OCH1*/P_{MET3}-*CHS3* mutants displayed a substantially less severe growth defect and did not aggregate in the absence of methionine. Thus, even under the repressed condition, the conditional mutants exhibited more robust growth and less cell aggregation, which is evidently advantageous over *och1* deletion in terms of long-term sustainability of the host strains and high cell density cultivation for industrial production. However, further studies are required to examine the potential of these strains as practical hosts by analyzing their performance in fed-batch cultivation, and other stringent promoters that are regulated by chemical less expensive than methionine (e.g. *tetO*, which is repressed by tetracycline) should be considered for future applications.

Investigations into enhanced protein secretion through the isolation of cell wall integrity mutants have also been reported in *K. lactis*. Screening of genes involved in a supersecretion phenotype of *K. lactis* for recombinant *Bacillus amyloliquefaciens* α-amylase had previously identified *KlMNN10*, encoding a mannosyltransferase responsible for the majority of the α-1,6-polymerizing activity of the mannosyltransferase complex (Bartkeviciute and Sasnauskas 2004). Disruption of *KlMNN10* in *K. lactis*, as well as of *MNN10* and *MNN11* in *S. cerevisiae*, resulted in supersecretion of *B. amyloliquefaciens* α-amylase, human growth hormone and invertase. A recent study based on genome sequencing of *K. lactis* mutant strains showing significant functional differences in the secretion of recombinant proteins identified the mutated gene responsible for the supersecretion phenotype in the *K. lactis* MD2/1–9 mutant as a *SEC59* homolog of *K. lactis* (Ziogiene *et al.* 2019). Dolichol kinase, encoded by *SEC59*, is an essential protein that catalyzes the cytidine triphosphate (CTP)-dependent phosphorylation of dolichol in the terminal step of *de novo* dolichyl monophosphate biosynthesis in the mevalonate pathway. Compared to wild-type strain, the *K. lactis* MD2/1–9 mutant carries two mutations, G405S and I419S in Sec59p, both of which are necessary for supersecretion capability. While *S. cerevisiae sec59* mutant is temperature sensitive and shows reduced *N*-glycosylation and secretory efficacy, the *K. lactis* dolichol kinase mutations had fewer effect on glycosylation, temperature sensitivity and cell wall integrity. Interestingly, the *K. lactis* mutant strain showed 2–3-fold enhanced secretion of recombinant *B. amyloliquefaciens* α-amylase, *S. cerevisiae* Kex2 proteinase lacking the transmembrane domain, and human growth hormone fused with α-factor signal sequence when compared with the wild-type strain. This study suggested that enhanced secretion related to lower activity of mutant dolichol kinase in *K. lactis* results from mild changes in glycosylation levels that affect the activity of other proteins in the secretory pathway.

Perspectives

Optimization of protein secretion in yeast is a major challenge because of the multiple complex steps involved in this process, including protein translocation, folding, post-translational modification and vesicle trafficking between several organelles and the plasma membrane. Thus, the development of efficient engineering strategies targeted to each step of the secretion pathway in a modular fashion has gained considerable interest in recent decades. Many genetic toolkits have proved to be valuable assets for the creation of efficient yeast cell factories for secretory production of recombinant proteins. Particularly, modular integrated secretory system engineering is emerging as a powerful synthetic biology approach to enhance secretion capacity with designed function. As a good example, engineering of the secretion pathway of *P. pastoris* by integrating four modules, including UPR induction, co-expression of protein-specific chaperone, glycoengineering and introducing protease deficiency, was shown to substantially enhance the expression of membrane proteins with designed glycan structure (Claes *et al.* 2016). However, notably, accumulating studies strongly indicate that the effects of co-expressed secretion-enhancing factors partly depend on the protein to be produced. Therefore, integration of genetic engineering with systems biology analysis, such as transcriptome analysis, is being increasingly used to identify genes specifically causing enhanced secretory capacity for a protein of interest (Love *et al.* 2012). A comprehensive understanding of the main factors and events implicated in protein secretion, folding and glycosylation can be obtained by combining data from multiple omics analyses, such as genomics, transcriptomics, proteomics and glycomics, which will provide valuable information for the engineering of artificial secretion pathways with improved secretion capacity and specifically designed function.

Novel genome-wide and high-throughput screening systems exploiting advanced bionanotechnology are expected to facilitate the identification of new gene targets, especially of those that have synergistic effects with previously identified targets. Notably, a novel approach based on microfluidic screening combined with whole-genome sequencing has allowed the identification of several genes involved not only in secretory and trafficking pathways but also in other cellular processes, such as histone deacetylation and carbohydrate metabolism, as targets for improving protein secretion (Huang *et al.* 2015). A state-of-the-art approach for the modulation of gene expression, using RNAi combined with high-throughput microfluidic single-cell screening, identified novel targets for improving protein secretion, including genes involved in cell metabolism and cell cycle, as well as protein modification and degradation (Wang *et al.* 2019). Particularly, this RNAi screening approach is an efficient strategy to identify targets that can enhance the production of proteins with intermediate expression levels, unlike screenings using overexpression and knockout libraries and, thus, allows for the identification of essential genes as candidates for protein production optimization. These findings strongly suggest fine tuning of expression levels of genes involved in the protein secretion pathway as a superior approach compared to simple deletion or overexpression of target genes. Along with such remarkable advances in novel high-throughput screening techniques, which will expand the repertoire of target genes for synthetic biology, combinatorial genetic manipulation of newly identified target genes using versatile synthetic biology toolboxes will facilitate the successful exploitation of yeast as a cell factory optimized for secretory production of correctly folded

and processed recombinant proteins in both research and industrial fields.

FUNDING

This research was supported by the National Research Foundation of Korea (grant no. NRF2018R1A5A1025077, Advanced Research Center Program). EJ Thak was the recipient of the Chung-Ang University *Excellent Student Scholarship* in 2014.

Conflicts of interest. None declared.

REFERENCES

Ahn J, Jang MJ, Ang KS *et al.* Codon optimization of *Saccharomyces cerevisiae* mating factor alpha prepro-leader to improve recombinant protein production in *Pichia pastoris*. *Biotechnol Lett* 2016;**38**:2137–43.

Amano K, Chiba Y, Kasahara Y *et al.* Engineering of mucin-type human glycoproteins in yeast cells. *Proc Natl Acad Sci USA* 2008;**105**:3232–7.

Aw R, Polizzi KM. Can too many copies spoil the broth? *Microb Cell Fact* 2013;**12**:128–36.

Bae JH, Sung BH, Kim HJ *et al.* An efficient genome-wide fusion partner screening system for secretion of recombinant proteins in yeast. *Sci Rep* 2015;**5**:12229–43.

Baghban R, Farajnia S, Rajabibazl M *et al.* Yeast expression systems: overview and recent advances. *Mol Biotechnol* 2019;**61**:365–84.

Bao J, Huang M, Petranovic D *et al.* Moderate expression of *SEC16* increases protein secretion by *Saccharomyces cerevisiae*. *Appl Environ Microb* 2017;**83**:e03400–16.

Bartkeviciute D, Sasnauskas K. Disruption of the *MNN10* gene enhances protein secretion in *Kluyveromyces lactis* and *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2004;**4**:833–40.

Besada-Lombana PB, Da Silva NA. Engineering the early secretory pathway for increased protein secretion in *Saccharomyces cerevisiae*. *Metab Eng* 2019;**55**:142–51.

Bobrowicz P, Davidson RC, Li H *et al.* Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. *Glycobiology* 2004;**14**:757–66.

Brule CE, Grayhack EJ. Synonymous codons: choose wisely for expression. *Trends Genet* 2017;**33**:283–97

Çelik E, Çalik P. Production of recombinant proteins by yeast cells. *Biotechnol Adv* 2012;**30**:1108–18.

Celinska E, Borkowska M, Bialas W *et al.* Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags. *Appl Microbiol Biotechnol* 2018;**102**:5221–33.

Chen X, Zhang J. The genomic landscape of position effects on protein expression level and noise in yeast. *Cell Syst* 2016;**2**:347–54.

Cheon SA, Kim H, Oh DB *et al.* Remodeling of the glycosylation pathway in the methylotrophic yeast *Hansenula polymorpha* to produce human hybrid-type *N*-glycans. *J Microbiol* 2012;**50**:341–8.

Chevance FF, Le Guyon S, Hughes KT. The effects of codon context on *in vivo* translation speed. *PLos Genet* 2014;**10**:e1004392.

Chiba Y, Suzuki M, Yoshida S *et al.* Production of human compatible high mannose-type ($Man_5GlcNAc_2$) sugar chains in *Saccharomyces cerevisiae*. *J Biol Chem* 1998;**273**:26298–304.

Chigira Y, Oka T, Okajima T *et al.* Engineering of a mammalian *O*-glycosylation pathway in the yeast *Saccharomyces cerevisiae*: production of *O*-fucosylated epidermal growth factor domains. *Glycobiology* 2008;**18**:303–14.

Choi BK, Bobrowicz P, Davidson RC *et al.* Use of combinatorial genetic libraries to humanize *N*-linked glycosylation in the yeast *Pichia pastoris*. *Proc Natl Acad Sci USA* 2003;**100**:5022–7.

Choi BK, Warburton S, Lin H *et al.* Improvement of *N*-glycan site occupancy of therapeutic glycoproteins produced in *Pichia pastoris*. *Appl Microbiol Biotechnol* 2012;**95**:671–82.

Chung BK, Lee DY. Computational codon optimization of synthetic gene for protein expression. *BMC Syst Biol* 2012;**6**:134–47.

Claes K, Vandewalle K, Laukens B *et al.* Modular integrated secretory system engineering in *Pichia pastoris* to enhance G-protein coupled receptor expression. *ACS Synth Biol* 2016;**5**:1070–5.

Davidson RC, Nett JH, Renfer E *et al.* Functional analysis of the *ALG3* gene encoding the Dol-P-Man: $Man_5GlcNAc_2$-PP-Dol mannosyltransferase enzyme of *P. pastoris*. *Glycobiology* 2004;**14**:399–407.

Davy AM, Kildegaard HF, Andersen MR. Cell factory engineering. *Cell Syst* 2017;**4**:262–75.

Delic M, Valli M, Graf AB *et al.* The secretory pathway: exploring yeast diversity. *FEMS Microbiol Rev* 2013;**37**:872–914.

De Pourcq K, De Schutter K, Callewaert N. Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Appl Microbiol Biotechnol* 2010;**87**:1617–31.

De Pourcq K, Tiels P, Van Hecke A *et al.* Engineering *Yarrowia lipolytica* to produce glycoproteins homogeneously modified with the universal $Man_3GlcNAc_2$ *N*-glycan core. *PLoS One* 2012a;**7**:e39976.

De Pourcq K, Vervecken W, Dewerte I *et al.* Engineering the yeast *Yarrowia lipolytica* for the production of therapeutic proteins homogeneously glycosylated with $Man_8GlcNAc_2$ and $Man_5GlcNAc_2$. *Microb Cell Fact* 2012b;**11**:53–64.

de Ruijter JC, Frey AD. Analysis of antibody production in *Saccharomyces cerevisiae*: effects of ER protein quality control disruption. *Appl Microbiol Biotechnol* 2015;**99**:9061–71.

de Ruijter JC, Koskela EV, Frey AD. Enhancing antibody folding and secretion by tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum. *Microb Cell Fact* 2016;**15**:87–104.

Ergün GB, Hüccetoğulları D, Öztürk S *et al.* Established and upcoming yeast expression systems. *Methods Mol Biol* 2019;**1923**:1–74.

Eskitoros MŞ, Ata Ö, Çalık P. Metabolic reaction network of *Pichia pastoris* with glycosylation reactions: flux analysis for erythropoietin production. *J Chem Technol Biotechnol* 2014;**89**:1675–85.

Gasser B, Maurer M, Gach J *et al.* Engineering of *Pichia pastoris* for improved production of antibody fragments. *Biotechnol Bioeng* 2006;**94**:353–61.

Gil JY, Park JN, Lee KJ *et al.* Increased mannosylphosphorylation of *N*-glycans by heterologous expression of *YlMPO1* in glyco-engineered *Saccharomyces cerevisiae* for mannose-6-phosphate modification. *J Biotechnol* 2015;**206**:66–74.

Guerfal M, Ryckaert S, Jacobs PP *et al.* 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* 2010;**9**:49–60.

Hamilton SR, Bobrowicz P, Bobrowicz B *et al.* Production of complex human glycoproteins in yeast. *Science* 2003;**301**:1244–6.

Hamilton SR, Cook WJ, Gomathinayagam S *et al*. Production of sialylated O-linked glycans in *Pichia pastoris*. *Glycobiology* 2013;**23**:1192–203.

Hamilton SR, Davidson RC, Sethuraman N *et al*. Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* 2006;**313**:1441–3.

Hopkins D, Gomathinayagam S, Rittenhour AM *et al*. Elimination of beta-mannose glycan structures in *Pichia pastoris*. *Glycobiology* 2011;**21**:1616–26.

Huang M, Bai Y, Sjostrom SL *et al*. Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proc Natl Acad Sci USA* 2015;**112**:E4689–96.

Huang M, Bao J, Hallström BM *et al*. Efficient protein production by yeast requires global tuning of metabolism. *Nat Commun* 2017;**8**:1131–42.

Huang M, Wang G, Qin J *et al*. Engineering the protein secretory pathway of *Saccharomyces cerevisiae* enables improved protein production. *Proc Natl Acad Sci USA* 2018;**115**:E11025–32.

Irani ZA, Kerkhoven EJ, Shojaosadati SA *et al*. Genome-scale metabolic model of *Pichia pastoris* with native and humanized glycosylation of recombinant proteins. *Biotechnol Bioeng* 2016;**113**:961–9.

Kang HA, Kang W, Hong WK *et al*. Development of expression systems for the production of recombinant human serum albumin using the MOX promoter in *Hansenula polymorpha* DL-1. *Biotechnol Bioeng* 2001;**76**:175–85.

Kang JY, Shin KK, Kim HH *et al*. Lysosomal targeting enhancement by conjugation of glycopeptides containing mannose-6-phosphate glycans derived from glyco-engineered yeast. *Sci Rep* 2018;**8**:8730–43.

Katla S, Yoganand KNR, Hingane S *et al*. Novel glycosylated human interferon alpha 2b expressed in glycoengineered *Pichia pastoris* and its biological activity: N-linked glycoengineering approach. *Enzyme Microb Technol* 2019;**128**: 49–58.

Kerkhoven EJ, Lahtvee PJ, Nielsen J. Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Res* 2015;**15**:1–13.

Kim H, Yoo SJ, Kang HA. Yeast synthetic biology for the production of recombinant therapeutic proteins. *FEMS Yeast Res* 2015;**15**:1–16.

Kim MW, Kim EJ, Kim JY *et al*. Functional characterization of the *Hansenula polymorpha HOC1*, *OCH1*, and *OCR1* genes as members of the yeast *OCH1* mannosyltransferase family involved in protein glycosylation. *J Biol Chem* 2006;**281**:6261–72.

Lanza AM, Curran KA, Rey LG *et al*. A condition specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae*. *BMC Syst Biol* 2014;**8**:33–42.

Lee SJ, Rose JK. A yeast secretion trap assay for identification of secreted proteins from eukaryotic phytopathogens and their plant hosts. *Methods Mol Biol* 2012;**835**:519–30.

Li J, Zhang MM, Wan C *et al*. Improved cellulase production in recombinant *Saccharomyces cerevisiae* by disrupting the cell wall protein-encoding gene *CWP2*. *J Biosci Bioeng* 2020;**129**:165–71.

Li L, Liu X, Wei K *et al*. Synthetic biology approaches for chromosomal integration of genes and pathways in industrial microbial systems. *Biotechnol Adv* 2019;**37**:730–45.

Lin-Cereghino GP, Stark CM, Kim D *et al*. The effect of alpha-mating factor secretion signal mutations on recombinant protein expression in *Pichia pastoris*. *Gene* 2013;**519**: 311–7.

Liu B, Gong X, Chang S *et al*. Disruption of the *OCH1* and *MNN1* genes decrease N-glycosylation on glycoprotein expressed in *Kluyveromyces lactis*. *J Biotechnol* 2009;**143**:95–102.

Liu CP, Tsai TI, Cheng T *et al*. Glycoengineering of antibody (Herceptin) through yeast expression and *in vitro* enzymatic glycosylation. *Proc Natl Acad Sci USA* 2018a;**115**:720–5.

Liu Y, Mo WJ, Shi TF *et al*. Mutational Mtc6p attenuates autophagy and improves secretory expression of heterologous proteins in *Kluyveromyces marxianus*. *Microb Cell Fact* 2018b;**17**:144–57.

Love KR, Politano TJ, Panagiotou V *et al*. Systematic single-cell analysis of *Pichia pastoris* reveals secretory capacity limits productivity. *PLoS One* 2012;**7**:e37915.

Luu VT, Moon HY, Hwang JY *et al*. Development of recombinant *Yarrowia lipolytica* producing virus-like particles of a fish nervous necrosis virus. *J Microbiol* 2017;**55**:655–64.

Luu VT, Moon HY, Yoo SJ *et al*. Development of conditional cell lysis mutants of *Saccharomyces cerevisiae* as production hosts by modulating *OCH1* and *CHS3* expression. *Appl Microbiol Biotechnol* 2019;**103**:2277–93.

Manfrao-Netto JHC, Gomes AMV, Parachin NS. Advances in using *Hansenula polymorpha* as chassis for recombinant protein production. *Front Bioeng Biotechnol* 2019;**7**:94–106.

Marsalek L, Gruber C, Altmann F *et al*. Disruption of genes involved in CORVET complex leads to enhanced secretion of heterologous carboxylesterase only in protease deficient *Pichia pastoris*. *Biotechnol J* 2017;**12**:1–12.

Marsalek L, Puxbaum V, Buchetics M *et al*. Disruption of vacuolar protein sorting components of the HOPS complex leads to enhanced secretion of recombinant proteins in *Pichia pastoris*. *Microb Cell Fact* 2019;**18**:119–34.

Massahi A, Çalık P. Endogenous signal peptides in recombinant protein production by *Pichia pastoris*: from *in silico* analysis to fermentation. *J Theor Biol* 2016;**408**:22–33.

Mattanovich D, Branduardi P, Dato L *et al*. Recombinant protein production in yeasts. *Methods Mol Biol* 2012;**824**:329–58.

Moon HY, Cheon SA, Kim H *et al*. *Hansenula polymorpha* Hac1p is critical to protein N-glycosylation activity modulation, as revealed by functional and transcriptomic analyses. *Appl Environ Microb* 2015;**81**:6982–93.

Moon HY, Lee DW, Sim GH *et al*. A new set of rDNA-NTS-based multiple integrative cassettes for the development of antibiotic-marker-free recombinant yeasts. *J Biotechnol* 2016;**233**:190–9.

Nasab FP, Aebi M, Bernhard G *et al*. A combined system for engineering glycosylation efficiency and glycan structure in *Saccharomyces cerevisiae*. *Appl Environ Microb* 2013;**79**:997–1007.

Nicaud J-M, Madzak C, van den Broek P *et al*. Protein expression and secretion in the yeast *Yarrowia lipolytica*. *FEMS Yeast Res* 2002;**2**:371–9.

Nocon J, Steiger MG, Pfeffer M *et al*. Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. *Metab Eng* 2014;**24**:129–38.

Obst U, Lu TK, Sieber V. A modular toolkit for generating *Pichia pastoris* secretion libraries. *ACS Synth Biol* 2017;**6**:1016–25.

Omara WA, Rash BM, Hayes A *et al*. Conditional cell-wall mutants of *Saccharomyces cerevisiae* as delivery vehicles for therapeutic agents *in vivo* to the GI tract. *J Biotechnol* 2010;**147**:136–43.

Orlean P. Architecture and biosynthesis of the *Saccharomyces cerevisiae* cell wall. *Genetics* 2012;**192**:775–818.

Park JN, Song Y, Cheon SA *et al*. Essential role of *YlMPO1*, a novel *Yarrowia lipolytica* homologue of *Saccharomyces cerevisiae MNN4*, in mannosylphosphorylation of N- and O-linked glycans. *Appl Environ Microb* 2011;**77**:1187–95.

Payne T, Hanfrey C, Bishop AL *et al*. Transcript-specific translational regulation in the unfolded protein response of *Saccharomyces cerevisiae*. *FEBS Lett* 2008;**582**:503–9.

Peña DA, Gasser B, Zanghellini J *et al*. Metabolic engineering of *Pichia pastoris*. *Metab Eng* 2018;**50**:2–15.

Piirainen MA, Boer H, de Ruijter JC *et al*. A dual approach for improving homogeneity of a human-type N-glycan structure in *Saccharomyces cerevisiae*. *Glycoconj J* 2016;**33**:189–99.

Piirainen MA, de Ruijter JC, Koskela EV *et al*. Glycoengineering of yeasts from the perspective of glycosylation efficiency. *N Biotechnol* 2014;**31**:532–7.

Popolo L, Vai M. The Gas1 glycoprotein, a putative wall polymer cross-linker. *Biochim Biophys Acta* 1999;**1426**:385–400.

Saitua F, Torres P, Pérez-Correa JR *et al*. Dynamic genome-scale metabolic modeling of the yeast *Pichia pastoris*. *BMC Syst Biol* 2017;**11**:27–47.

Schuck S, Prinz WA, Thorn KS *et al*. Membrane expansion alleviates endoplasmic reticulum stress independently of the unfolded protein response. *J Cell Biol* 2009;**187**:525–36.

Schwartz C, Shabbir-Hussain M, Frogue K *et al*. Standardized markerless gene integration for pathway engineering in *Yarrowia lipolytica*. *ACS Synth Biol* 2017;**6**:402–9.

Sha C, Yu XW, Lin NX *et al*. Enhancement of lipase r27RCL production in *Pichia pastoris* by regulating gene dosage and co-expression with chaperone protein disulfide isomerase. *Enzyme Microb Technol* 2013;**53**:438–43.

Sheng J, Flick H, Feng X. Systematic optimization of protein secretory pathways in *Saccharomyces cerevisiae* to increase expression of hepatitis B small antigen. *Front Microbiol* 2017;**8**:875–84.

Shi S, Liang Y, Zhang MM *et al*. A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. *Metab Eng* 2016;**33**:19–27.

Song Y, Choi MH, Park JN *et al*. Engineering of the yeast *Yarrowia lipolytica* for the production of glycoproteins lacking the outer-chain mannose residues of N-glycans. *Appl Environ Microb* 2007;**73**:4446–54.

Taipakova SM, Smekenov IT, Saparbaev MK *et al*. Characterization of *Aspergillus niger* endo-1,4-β-glucanase ENG1 secreted from *Saccharomyces cerevisiae* using different expression vectors. *Genet Mol Res* 2015;**14**:6439–52

Tang H, Bao X, Shen Y *et al*. Engineering protein folding and translocation improves heterologous protein secretion in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 2015;**112**:1872–82.

Tang H, Song M, He Y *et al*. Engineering vesicle trafficking improves the extracellular activity and surface display efficiency of cellulases in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2017;**10**:53–65.

Tang H, Wang S, Wang J *et al*. N-hypermannose glycosylation disruption enhances recombinant protein production by regulating secretory pathway and cell wall integrity in *Saccharomyces cerevisiae*. *Sci Rep* 2016;**6**:25654–66.

Thak EJ, Kim J, Lee DJ *et al*. Structural analysis of N-/O-glycans assembled on proteins in yeasts. *J Microbiol* 2018;**56**:11–23.

Tiels P, Baranova E, Piens K *et al*. A bacterial glycosidase enables mannose-6-phosphate modification and improved cellular uptake of yeast-produced recombinant human lysosomal enzymes. *Nat Biotechnol* 2012;**30**:1225–31.

Vai M, Brambilla L, Orlandi I *et al*. Improved secretion of native human insulin-like growth factor 1 from *gas1* mutant *Saccharomyces cerevisiae* cells. *Appl Environ Microb* 2000;**66**:5477–9.

Valkonen M, Penttila M, Saloheimo M. Effects of inactivation and constitutive expression of the unfolded-protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microb* 2003;**69**:2065–72.

Van Zyl JH, Den Haan R, Van Zyl WH. Overexpression of native *Saccharomyces cerevisiae* ER-to-Golgi SNARE genes increased heterologous cellulase secretion. *Appl Microbiol Biotechnol* 2016;**100**:505–18.

Vieira Gomes AM, Souza Carmo T, Silva Carvalho L *et al*. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* 2018;**6**:38–60.

Wagner JM, Alper HS. Synthetic biology and molecular genetics in non-conventional yeasts: current tools and future advances. *Fungal Genet Biol* 2016;**89**:126–36.

Walsh G. Post-translational modifications of protein biopharmaceuticals. *Drug Discov Today* 2010;**15**:773–80.

Wang G, Bjork SM, Huang M *et al*. RNAi expression tuning, microfluidic screening, and genome recombineering for improved protein production in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 2019;**116**:9324–32.

Wang H, Song HL, Wang Q *et al*. Expression of glycoproteins bearing complex human-like glycans with galactose terminal in *Hansenula polymorpha*. *World J Microbiol Biotechnol* 2013;**29**:447–58.

Whyteside G, Alcocer MJ, Kumita JR *et al*. Native-state stability determines the extent of degradation relative to secretion of protein variants from *Pichia pastoris*. *PLoS One* 2011;**6**:e22692.

Wu XL, Li BZ, Zhang WZ *et al*. Genome-wide landscape of position effects on heterogeneous gene expression in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2017;**10**:189–98.

Yang Z, Zhang Z. Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: a review. *Biotechnol Adv* 2018;**36**:182–95.

Yoo SJ, Moon HY, Kang HA. Screening and selection of production strains: secretory protein expression and analysis in *Hansenula polymorpha*. *Methods Mol Biol* 2019;**1923**:133–51.

Zahrl RJ, Pena DA, Mattanovich D *et al*. Systems biotechnology for protein production in *Pichia pastoris*. *FEMS Yeast Res* 2017;**17**:1–15.

Zhang N, Gardner DC, Oliver SG *et al*. Down-regulation of the expression of *PKC1* and *SRB1/PSA1/VIG9*, two genes involved in cell wall integrity in *Saccharomyces cerevisiae*, causes flocculation. *Microbiology* 1999a;**145**:309–16.

Zhang N, Gardner DC, Oliver SG *et al*. Genetically controlled cell lysis in the yeast *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 1999b;**64**:607–15.

Zhao H, Blazanovic K, Choi Y *et al*. Gene and protein sequence optimization for high-level production of fully active and aglycosylated lysostaphin in *Pichia pastoris*. *Appl Environ Microb* 2014;**80**:2746–53.

Ziogiene D, Valaviciute M, Norkiene M *et al*. Mutations of *Kluyveromyces lactis* dolichol kinase enhances secretion of recombinant proteins. *FEMS Yeast Res* 2019;**19**:1–13.