Production of biopharmaceutical proteins by yeast Advances through metabolic engineering

Bioengineered

ISSN: 2165-5979 (Print) 2165-5987 (Online) Journal homepage: www.tandfonline.com/journals/kbie20

Production of biopharmaceutical proteins by yeast

Advances through metabolic engineering

Jens Nielsen

To cite this article: Jens Nielsen (2013) Production of biopharmaceutical proteins by yeast, Bioengineered, 4:4, 207-211, DOI: 10.4161/bioe.22856

To link to this article: https://doi.org/10.4161/bioe.22856

Bioengineered 4:4, 207–211; July/August 2013; © 2013 Landes Bioscience

Production of biopharmaceutical proteins by yeast

Advances through metabolic engineering

Jens Nielsen

Novo Nordisk Foundation Center for Biosustainability; Chalmers University of Technology; Gothenburg, Sweden; Novo Nordisk Foundation Center for Biosustainability; Technical University of Denmark; Hørsholm, Denmark; Science for Life Laboratory; Solna, Sweden

Keywords: *Saccharomyces cerevisiae*, systems biology, secretory pathway, insulin, industrial biotechnology

Production of recombinant proteins for use as pharmaceuticals, so-called biopharmaceuticals, is a multi-billion dollar industry. Many different cell factories are used for the production of biopharmaceuticals, but the yeast *Saccharomyces cerevisiae* is an important cell factory as it is used for production of several large volume products. Insulin and insulin analogs are by far the dominating biopharmaceuticals produced by yeast, and this will increase as the global insulin market is expected to grow from USD12B in 2011 to more than USD32B by 2018. Other important biopharmaceuticals produced by yeast are human serum albumin, hepatitis vaccines and virus like particles used for vaccination against human papillomavirus. Here is given a brief overview of biopharmaceutical production by yeast and it is discussed how the secretory pathway can be engineered to ensure more efficient protein production. The involvement of directed metabolic engineering through the integration of tools from genetic engineering, systems biology and mathematical modeling, is also discussed.

The introduction of genetic engineering by Cohen and Boyer in 1973[1] laid the fundament for the current biotech industry, which is based on using microorganisms or cell cultures for production of proteins that can serve as pharmaceuticals, often referred to as biopharmaceuticals.[2,3] A few years later, researchers at Genentech cloned the genes for human insulin and growth hormone, and expressed them in *Escherichia coli*,[4] hereby demonstrating the utility and applicability of genetic engineering in creating genetically engineered bacteria that produce these two human proteins. In 1982 this led to marketing of the first biopharmaceutical, human insulin, by Eli Lilly, who licensed the technology from Genentech. In 1985 Genentech received FDA approval to market their own first product, Protropin®, the human growth hormone to be used for children with growth hormone deficiency. In 1987 this was followed by the tissue-plasminogen activator (t-PA, Activase®), another Genentech product, an enzyme that can resolve blood clots in patients with acute myocardial infarction. Also in 1987 Novo (now Novo Nordisk), a major insulin producing company, launched human insulin produced by the yeast *Saccharomyces cerevisiae* as a replacement for their human insulin enzymatically derived from porcine insulin. Shortly following these early developments many other products were launched and today there are more than 300 biopharmaceutical proteins and antibodies on the market with sales exceeding USD100B,[5,6] with monoclonal antibodies representing the majority (> USD18B) followed by hormones (> USD11B) and growth factors (> USD10B).[7] Furthermore, biopharmaceuticals have the fastest growth in the marked with an annual growth of about 19%,[8] and there are currently more than 240 monoclonal antibodies and 120 recombinant proteins in clinical trials.[9]

About 40% of the biopharmaceuticals are currently being produced by mammalian cell cultures, mainly using Chinese Hamster Ovarian cell lines (CHO cells), as these allow for production of proteins with very similar glycosylation patterns as human proteins.[10,11] *E. coli* is used as cell factory for production of another 30% of the biopharmaceuticals whereas 20% are being produced by *S. cerevisiae*.[10,11] The dominant biopharmaceuticals produced by *S. cerevisiae* are insulin (and insulin analogs), human serum albumin, hepatitis vaccines and virus like particles, e.g., for vaccination against human papillomavirus. The advantages of using yeast *S. cerevisiae* as a cell factory for the production of biopharmaceuticals are that this eukaryal model system enables production and proper folding of many human proteins. Furthermore, the proteins can be secreted to the extracellular medium and this facilitates subsequent purification. A further advantage is that in many cases yeast can perform proper post-translational modifications of the protein, including proteolytic processing of signal peptides, disulfide bond formation, subunit assembly, acylation and glycosylation.[12] *S. cerevisiae* is also widely used as an eukaryal model organism[13,14] and there is therefore much information available about this organism through high-throughput studies,[15] databases, sequenced genomes and extensive toolbox for molecular modification, which provides an extensive knowledge base for further engineering of this organism. One of the limitations with the use of yeast is, however, that it performs high-mannose type N-glycosylation. This confers a short half-life of the modified protein in vivo, which then can have a reduced efficacy for therapeutic use.[16] Much work has been performed on engineering yeasts, both *S. cerevisiae* and *Pichia pastoris*, so that they can carry out human-like N-glycosylation patterns that even includes terminal addition of sialic acid to the

Correspondence to: Jens Nielsen; Email: nielsenj@chalmers.se
Submitted: 10/04/12; Revised: 11/09/12; Accepted: 11/12/12
http://dx.doi.org/10.4161/bioe.22856

glycoprotein.[16-18] This has opened for an even wider use of *S. cerevisiae* as a cell factory for production of biopharmaceuticals and there is therefore much interest in further engineering of yeast for ensuring efficient production of recombinant proteins.

More than 40 different recombinant proteins have been expressed, produced and secreted by *S. cerevisiae*.[12] This includes several biopharmaceuticals and **Table 1** provides an overview of some of these products, i.e., the protein name, their therapeutic application, leader sequence used and the titer reported in the publically available literature. As the table illustrates there are basically used three different types of leader sequences to ensure efficient secretion of the protein through the secretory pathway. *S. cerevisiae* only secretes few proteins to the extracellular medium, with the α-factor (a yeast hormone involved in mating) being the most studied and therefore most frequently used for efficient secretion of recombinant proteins. To further improve protein secretion in yeast Kjeldsen and coworkers at Novo Nordisk developed a synthetic leader that has been shown to be very efficient in protein secretion.[12,19] As illustrated in **Table 1**, yeast can, however, also secrete human proteins that are expressed with their native leader sequences.

The secretory pathway in yeast is quite complex (see **Fig. 1** for a schematic overview) as it involves more than 160 proteins that are responsible for different post-translational processes, e.g., folding and glycosylation. The secretory pathway handles more than 550 proteins that have a signal peptide in the yeast proteome, but only very few of these proteins are secreted to the extracellular matrix as all proteins targeted to the endoplasmic reticulum (ER), Golgi, vacuole and cytoplasmic membrane are also processed through the secretory pathway. Protein folding in the ER is of key importance for the secretory pathway as the accumulation of mis-folded proteins results in ER stress that is handled then by the unfolded protein response (UPR). Activation of the UPR results in transcriptional change of about 400 genes,[20] many of which are under regulation of the Hac1p transcription factor.[21] A result of this regulation is upregulation of chaperones and foldases as well as ER associated degradation (ERAD). Based on studies of the UPR many targets for improving protein secretion have been identified and implemented,[12] and it is generally believed that any factor that reduces ER stress and its downstream damage caused by heterologous protein production, results in improved secretion of the produced protein.

Due to the complexity of the protein secretion pathway there has traditionally been focus on transcriptional regulation of protein production. A large number of different promoters have been evaluated for driving expression of the heterologous genes, and among the most widely used promoters are strong glycolytic promoters like *pTDH3*, *pPGK1* or *pTPI1*, *pADH1*, or galactose induced promoters like *pGAL1*, *pGAL7* or *pGAL10*.[12] However, even though changing expression may improve production of one protein, it does not imply there would be improvement in production of another protein, as clearly demonstrated in our recent study on production of insulin precursor and α-amylase.[22] In this study we used an expression system originally developed by Novo for their insulin production. This expression system involves deletion of the *TPI1* gene in the chromosome and use of the corresponding gene, *POT1*, from *Schizosaccharomyces pombe* as a plasmid marker (see **Fig. 2**). This gives a very stable construct as all cells losing the plasmid will be deficient in triose-phosphate isomerase activity, a key glycolytic enzyme. A particular strength of this method is that it is stable also with use of complex media containing amino acids and nucleotides, which is not the case with standard yeast auxotrophy markers. Using this vector system we evaluated two different promoters (*pTPI1* and *pTEF1*) as well as two different leader sequences (synthetic leader and the α-factor leader) for production of insulin precursor and α-amylase, and found that there was considerable differences in production of the two proteins with the different expression systems evaluated, i.e., with low gene expression α-amylase was produced at higher levels, whereas for high gene expression system insulin precursor was produced at higher levels.[22] This pointed to very different protein processing in the ER, i.e., the larger and more complex α-amylase is more challenging for ER-processing when the flux is high. This hypothesis was confirmed by a measured increase in ER stress (by genome-wide transcription analysis) in the α-amylase producing strains. Based on this analysis it was hypothesized that engineering of the downstream secretion pathway may be able to improve the secretion of amylase, and indeed overexpressing regulators of the so-called SNARE complex, Sec1p and Sly1p, resulted in improved protein secretion.[23] Moreover, it was found that overexpression of *SEC1*, that is involved in regulating vesicle trafficking from Golgi to the cell membrane, resulted in improved production of both insulin precursor and α-amylase, whereas overexpression of *SLY1*, that is involved in regulating the vesicle fusion from ER to Golgi, increased only the α-amylase production.[23] Through combined overexpression of the Sec1p and Sly1p the overall secretion of α-amylase could be improved by about 70%, whereas insulin precursor production was increased by about 30%, and the study therefore clearly demonstrates that engineering of the secretory pathway can result in significant improvement of recombinant protein production.

There are many other studies that clearly demonstrate that engineering of the secretory pathway can result in improved

Table 1. Overview of some biopharmaceuticals produced by *S. cerevisiae*[12]

Type	Protein	Therapeutic application	Leader sequence	Titer
Blood related	Human Serum Albumin	Surgery (plasma expander)	Native	3 g/L
	Hirudin	Blood coagulation disorders	α-Factor	460 mg/L
	Human transferrin	Anemia	Native	1.8 g/L
Hormones	Insulin Precursor	Diabetes	Synthetic	80 mg/L
	Glucagon	Diabetes	α-Factor	17.5 mg/L
Antigen	Hepatitis surface antigen	Hepatitis vaccination	Native	19.4 mg/L

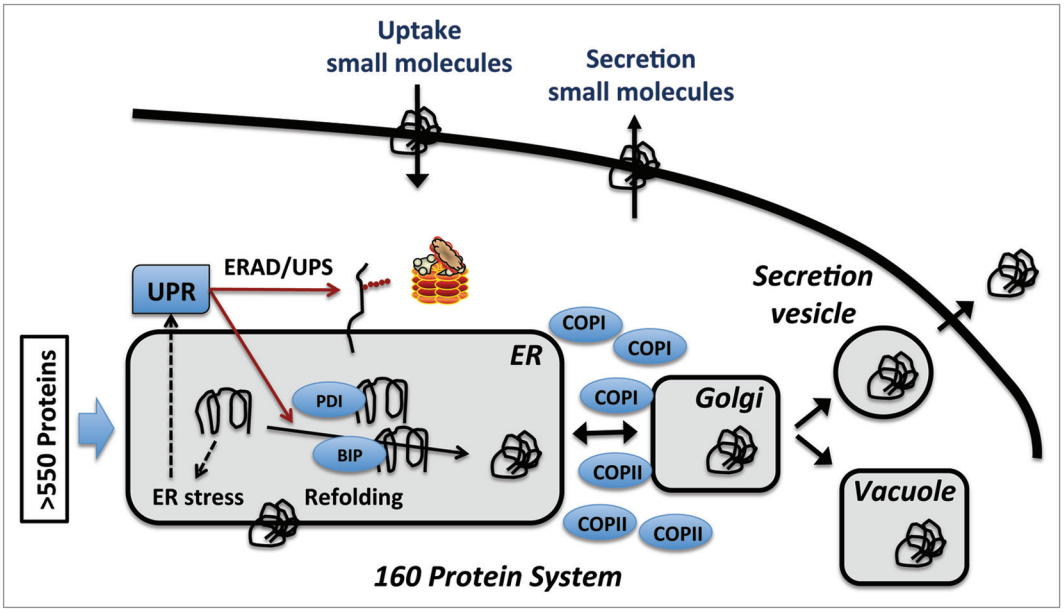

Figure 1. Schematic overview of the secretory pathway in yeast. Proteins targeted for secretion enter the endoplasmic reticulum (ER). If they fold correctly they can enter the secretory pathway, whereas misfolded protein cause ER stress leading to the activation of the unfolded protein response (UPR) that results in activation of a very large number of cellular processes, including activation of chaperones and foldases (like BIP and PDI) that assist with refolding. UPR is also upregulating ER-associated degradation (ERAD) where the unfolded proteins are exported from the ER, ubiquitinated and hereby targeted for degradation by the proteasome (ubiquitin-proteasome system, UPS). Correctly folded proteins can be exported to the Golgi for further processing (including additional glycosylation). The COPI- and COPII-complexes facilitate the ER-Golgi transfer, and from the Golgi the protein may be secreted via the endosome or be targeted to the vacuole for storage and/or degradation. Different colors represent different types of vesicular compartments of the secretory pathway.

protein production,[12] but a general finding is that most targets are quite specific, depending which particular protein is being overexpressed. Ideally, there would be one (or few) efficient yeast platform strain(s) that can serve as production host for a wide range of different biopharmaceuticals. There is therefore a need for a more rational approach to engineering yeast for improved protein secretion. In such an approach detailed models of the secretory pathway can be used for design, similarly as it has been done with engineering of metabolic pathways, where the use of genome-scale metabolic models has been shown to be of great importance.[15,24,25] This approach is referred to as metabolic engineering, and the workflow, often called the metabolic engineering cycle,[26] is illustrated in **Figure 3**.

As illustrated, the workflow involves detailed modeling, often based on detailed analysis of the cellular metabolism and physiology using high-throughput experimental techniques developed in the field of genomics, and the concept of quantitatively describing cellular processes with mathematical models is at the core of systems biology.[27] The metabolic engineering cycle is therefore very similar to the workflow of many systems biology studies where perturbation of the cellular system is performed using genetic engineering, e.g., by overexpression or deletion of specific genes, followed by detailed analyses that can be used to define a mathematical model of the biological system. However, there is a major difference in the sense that systems biology is a fundamental science where the primary objective is to gain novel insights, whereas metabolic engineering is an applied science with the primary objective to obtain an improved cell factory. Clearly the process

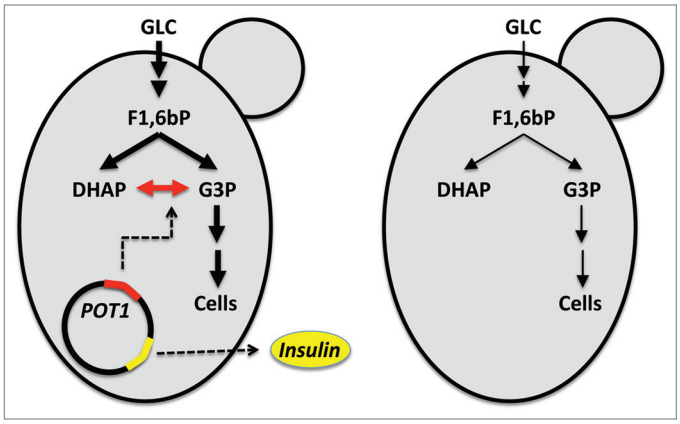

Figure 2. Illustration of the stable expression system with a glycolytic gene as the selection marker. One of the glycolytic enzymes is used as a marker for plasmid presence: the endogenous gene encoding triosephosphate isomerase (*TPI1*) is deleted and the corresponding gene (*POT1*) from *Schizosaccharomyces pombe* is expressed from a plasmid. The same plasmid carries the gene for the heterologous gene to be expressed (here demonstrated with a gene encoding human insulin). If the plasmid is lost the cells lack a key glycolytic enzyme and the glycolytic flux is therefore reduced dramatically resulting in impaired growth. Cells that are replicating the plasmid in high copy numbers and expressing the genes from the plasmid therefore have an inherent growth advantage.

of metabolic engineering also results in improved insight of the cellular metabolism and physiology, but generally to a less extend

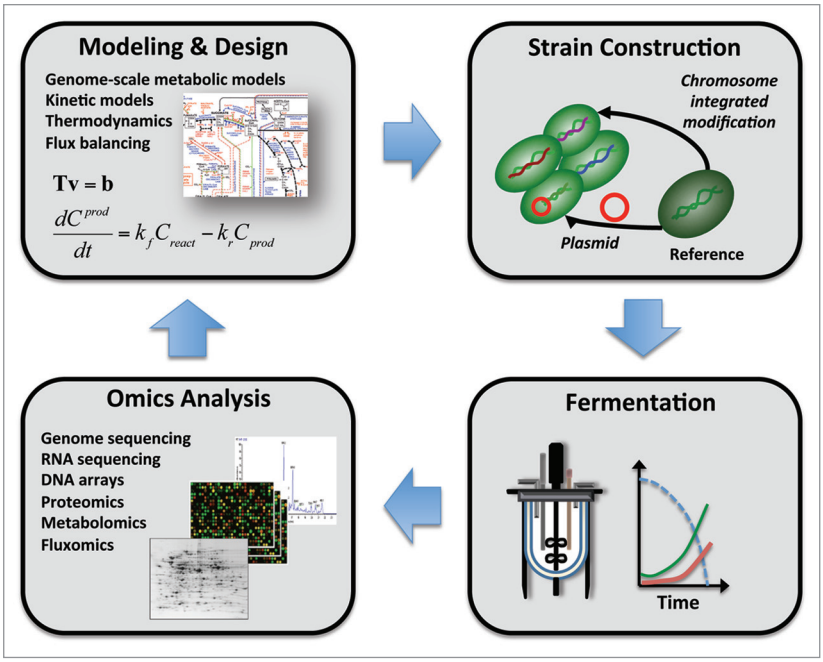

Figure 3. Schematic overview of the metabolic engineering cycle where systems biology tools are implemented for design of improved cell factories. Through advanced modeling novel targets for genetic engineering can be identified, e.g., to ensure improved protein secretion. These targets are evaluated by characterizing the strains in bioreactors where rates of biomass growth, sugar consumption, and product formation are quantified. Fermentation analysis may be combined with high-throughput analyses, or omics analyses, where the transcriptome, proteome, metabolome and fluxome are measured. Omics analyses may provide new insights into the cellular metabolism and physiology, and this may be used to improve the models, that can hence be used for further design. The metabolic engineering cycle is very similar to the workflow of many systems biology studies where perturbation of the cellular system is performed using genetic engineering, e.g., by overexpression or deletion of specific genes, followed by detailed analysis that can be used to define a mathematical model for the biological system.

than in a systems biology study due to the differences in study design resulting from the different objectives. A major limitation in using the rational Metabolic Engineering approach, with model based design, for improving protein secretion, is the lack of detailed models of the protein secretion pathway. Even though there have been described several kinetic models,[28,29] these describe only part of the pathway. There is therefore an obvious need for a genome-scale model for protein secretion, in analogy to what has been done for metabolism.[30] Such detailed models will not only enable rational design of improved protein secretion routes they will also enable improved integrative analysis of this complex pathway using previously developed tools (for studying metabolism[31]) and hereby lead to more insights into how the secretory pathway operates as a system.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

I would like to thank my students and post docs working in the field of protein production by yeast for fruitful discussions and collaborations: Dr Jin Hou, Dr Keith Tyo, Dr Jose Martinez, Dr Mingtao Huang, Zihe Liu, Lifang Liu, Tobias Österlund and Amir Feizi. I also would like to thank Prof Dina Petranovic for excellent collaboration and constructive comments on this paper. Research in the area of yeast protein production in my group is funded by European Research Council (grant no. 247013) and the Novo Nordisk Foundation.

References

1. Cohen SN, Chang ACY, Boyer HW, Helling RB. Construction of biologically functional bacterial plasmids *in vitro.* Proc Natl Acad Sci U S A 1973; 70:3240-4; PMID:4594039; http://dx.doi.org/10.1073/pnas.70.11.3240.
2. Walsh G. Biopharmaceutical benchmarks. Nat Biotechnol 2000; 18:831-3; PMID:10932150; http://dx.doi.org/10.1038/78720.
3. Walsh G. Biopharmaceuticals: Biochemistry and Biotechnology. 1998. J. Wiley & Sons, Ltd., Chichester, UK.
4. Chang CN, Rey M, Bochner B, Heyneker H, Gray G. High-level secretion of human growth hormone by *Escherichia coli.* Gene 1987; 55:189-96; PMID:3311882; http://dx.doi.org/10.1016/0378-1119(87)90279-4.
5. Langer ES. Biomanufacturing outsourcing outlook. BioPharm International. 2012; 25:15-6.
6. Goodman M. Market watch: Sales of biologics to show robust growth through to 2013. Nat Rev Drug Discov 2009; 8:837; PMID:19876035; http://dx.doi.org/10.1038/nrd3040.
7. Aggarwal S. What's fueling the biotech engine—2010 to 2011. Nat Biotechnol 2011; 29:1083-9; PMID:22158359; http://dx.doi.org/10.1038/nbt.2060.
8. Schröder M. Engineering eukaryotic protein factories. Biotechnol Lett 2008; 30:187-96; PMID:17885737; http://dx.doi.org/10.1007/s10529-007-9524-1.
9. Walsh G. Biopharmaceutical benchmarks 2010. Nat Biotechnol 2010; 28:917-24; PMID:20829826; http://dx.doi.org/10.1038/nbt0910-917.
10. Ferrer-Miralles N, Domingo-Espín J, Corchero JL, Vázquez E, Villaverde A. Microbial factories for recombinant pharmaceuticals. Microb Cell Fact 2009; 8:17; PMID:19317892; http://dx.doi.org/10.1186/1475-2859-8-17.
11. Martinez JL, Liu L, Petranovic D, Nielsen J. Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation. Curr. Opionion Biotechnol. 2012; In press; http://dx.doi.org/10.1016/j.copbio.2012.03.011.
12. Hou J, Tyo KEJ, Liu Z, Petranovic D, Nielsen J. Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae.* FEMS Yeast Res 2012; 12:491-510; PMID:22533807; http://dx.doi.org/10.1111/j.1567-1364.2012.00810.x.
13. Petranovic D, Nielsen J. Can yeast systems biology contribute to the understanding of human disease? Trends Biotechnol 2008; 26:584-90; PMID:18801589; http://dx.doi.org/10.1016/j.tibtech.2008.07.008.
14. Petranovic D, Tyo K, Vemuri GN, Nielsen J. Prospects of yeast systems biology for human health: integrating lipid, protein and energy metabolism. FEMS Yeast Res 2010; 10:1046-59; PMID:20977625; http://dx.doi.org/10.1111/j.1567-1364.2010.00689.x.
15. Kim I-K, Roldão A, Siewers V, Nielsen J. A systems-level approach for metabolic engineering of yeast cell factories. FEMS Yeast Res 2012; 12:228-48; PMID:22188344; http://dx.doi.org/10.1111/j.1567-1364.2011.00779.x.
16. Wildt S, Gerngross TU. The humanization of *N*-glycosylation pathways in yeast. Nat Rev Microbiol 2005; 3:119-28; PMID:15685223; http://dx.doi.org/10.1038/nrmicro1087.
17. Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang Y, Rios S, et al. Humanization of yeast to produce complex terminally sialylated glycoproteins. Science 2006; 313:1441-3; PMID:16960007; http://dx.doi.org/10.1126/science.1130256.
18. Li H, Sethuraman N, Stadheim TA, Zha D, Prinz B, Ballew N, et al. Optimization of humanized IgGs in glycoengineered *Pichia pastoris.* Nat Biotechnol 2006; 24:210-5; PMID:16429149; http://dx.doi.org/10.1038/nbt1178.
19. Kjeldsen T, Hach M, Balschmidt P, Havelund S, Pettersson AF, Markussen J. Prepro-leaders lacking *N*-linked glycosylation for secretory expression in the yeast *Saccharomyces cerevisiae.* Protein Expr Purif 1998; 14:309-16; PMID:9882564; http://dx.doi.org/10.1006/prep.1998.0977.
20. Tyo KEJ, Liu Z, Petranovic D, Nielsen J. Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress. BMC Biol 2012; 10:16; PMID:22380681; http://dx.doi.org/10.1186/1741-7007-10-16.

21. Patil C, Walter P. Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. Curr Opin Cell Biol 2001; 13:349-55; PMID:11343907; http://dx.doi.org/10.1016/S0955-0674(00)00219-2.
22. Liu Z, Tyo KEJ, Martínez JL, Petranovic D, Nielsen J. Different expression systems for production of recombinant proteins in Saccharomyces cerevisiae. Biotechnol Bioeng 2012; 109:1259-68; PMID:22179756; http://dx.doi.org/10.1002/bit.24409.
23. Hou J, Tyo KEJ, Liu Z, Petranovic D, Nielsen J. Engineering of vesicle trafficking improves heterologous protein secretion in *Saccharomyces cerevisiae*. Metab Eng 2012; 14:120-7; PMID:22265825; http://dx.doi.org/10.1016/j.ymben.2012.01.002.
24. Bro C, Regenberg B, Förster J, Nielsen J. *In silico* aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. Metab Eng 2006; 8:102-11; PMID:16289778; http://dx.doi.org/10.1016/j.ymben.2005.09.007.
25. Tyo KEJ, Kocharin K, Nielsen J. Toward design-based engineering of industrial microbes. Curr Opin Microbiol 2010; 13:255-62; PMID:20226723; http://dx.doi.org/10.1016/j.mib.2010.02.001.
26. Nielsen J. Metabolic engineering. Appl Microbiol Biotechnol 2001; 55:263-83; PMID:11341306; http://dx.doi.org/10.1007/s002530000511.
27. Nielsen J, Jewett MC. Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. FEMS Yeast Res 2008; 8:122-31; PMID:17727659; http://dx.doi.org/10.1111/j.1567-1364.2007.00302.x.
28. Umaña P, Bailey JE. A mathematical model of N-linked glycoform biosynthesis. Biotechnol Bioeng 1997; 55:890-908; PMID:18636599; http://dx.doi.org/10.1002/(SICI)1097-0290(19970920)55:6<890::AID-BIT7>3.0.CO;2-B.
29. Krambeck FJ, Betenbaugh MJ. A mathematical model of N-linked glycosylation. Biotechnol Bioeng 2005; 92:711-28; PMID:16247773; http://dx.doi.org/10.1002/bit.20645.
30. Österlund T, Nookaew I, Nielsen J. Fifteen years of large scale metabolic modeling of yeast: developments and impacts. Biotechnol Adv 2012; 30:979-88; PMID:21846501; http://dx.doi.org/10.1016/j.biotechadv.2011.07.021.
31. Patil KR, Nielsen J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proc Natl Acad Sci U S A 2005; 102:2685-9; PMID:15710883; http://dx.doi.org/10.1073/pnas.0406811102.