Pacific Journal of Health

Volume 4 | Issue 1 Article 2

2021

Overcoming Obstacles in Protein Expression in the Yeast Pichia pastoris: Interviews of Leaders in the Pichia Field

Zoe Ingram
University of the Pacific, z_ingram1@u.pacific.edu

Abha Patkar
University of the Pacific, a_patkar@u.pacific.edu

Dahoon Oh
University of the Pacific, d_oh9@u.pacific.edu

Kevin K. Zhang
University of the Pacific, k_zhang17@u.pacific.edu

Christina Chung
University of the Pacific, c_chung15@u.pacific.edu

Overcoming Obstacles in Protein Expression in the Yeast Pichia pastoris: Interviews of Leaders in the Pichia Field

Abstract

The yeast *Pichia pastoris* (also known as *Komagataella pastoris*) has been used for over 30 years to produce thousands of valuable, heterologous proteins, such as insulin to treat diabetes and antibodies to prevent migraine headaches. Despite its success, there are some common, stubborn problems encountered by research scientists when they try to use the yeast to produce their recombinant proteins. In order to provide those working in this field with strategies to overcome these common obstacles, nine experts in *P. pastoris* protein expression field were interviewed to create a written review and video (https://www.youtube.com/watch?v=Q1oD6k8CdG8). This review describes how each respected scientist addressed a specific challenge, such as identifying high expression strains, improving secretion efficiency and decreasing hyperglycosylation. Their perspective and practical advice can be a tool to help empower others to express challenging proteins in this popular recombinant host.

Keywords

Pichia pastoris, recombinant protein expression, protein secretion, protein glycosylation, high throughput screening, clonal variation

Authors

Zoe Ingram, Abha Patkar, Dahoon Oh, Kevin K. Zhang, Christina Chung, Joan Lin-Cereghino, and Geoff P. Lin-Cereghino

Introduction

Interviewee: James Cregg, PhD, Professor Emeritus, Keck Graduate Institute, Claremont, California

In the late 20th century, a number of hosts for heterologous protein expression were developed and tested. In 1969, the scientist Koichi Ogata paved the way for the study and use of *Pichia pastoris* when he discovered that certain species of yeast could use methanol as a sole carbon and energy source. Fifteen years later, the company SIBIA (Salk Institute Biotechnology/Industrial Associates) was tasked with developing *P. pastoris* as an organism for recombinant protein expression. Dr. Jim Cregg was a major player in this research effort.

At the time, *Saccharomyces cerevisiae* (baker's yeast) was the most popular yeast expression system. However, there were serious limitations in using this host. At a meeting sponsored by the Upjohn Company in the early 1980's, Dr. Cregg listened to several speakers detail these problems. One problem was *S. cerevisiae's* strong tendency to ferment sugars to ethanol, a product that quickly reached toxic levels at low culture density. A second problem was the lack of a strong inducible promoter in this yeast. At this meeting, Dr. Cregg had an epiphany. He realized that because *P. pastoris* did not have these problems, "making an expression system with this yeast was a helluva good idea."

Two developments were required to turn *P. pastoris* into an expression system: 1) a means to insert vector DNAs stably into the host (a DNA transformation system) and 2) a strong inducible promoter to drive expression of heterologous genes. In 1985, SIBIA published an article describing the first *P. pastoris* transformation system [1]. The method involved a series of steps to prepare the yeast spheroplasts with lyticase for the subsequent introduction of foreign DNA using calcium chloride and polyethylene glycol followed by a recovery in agar containing sorbitol.

Since this procedure, more efficient methods have been developed for transformation, including heat shock and electroporation [2,3]; however, the significance of this original methodology has not diminished. Second, to regulate expression of foreign genes in *P. pastoris*, the inducible *AOX1* promoter was isolated and engineered [4]. Expression of the *AOX1* gene is controlled at the level of transcription. In methanol-grown cells, about 5% of poly(A) mRNA is from *AOX1*; however, in cells grown on most other carbon sources, *AOX1* message is undetectable. Simply transferring cells from medium containing glucose to methanol induces transcription from the *AOX1* promoter by approximately one thousand fold.

Moreover, as Dr. Cregg pointed out, there are a number of additional advantages to *P. pastoris* as a protein expression system [5]. For example, because *P. pastoris* does not like to ferment, it is able to reach extremely high-cell densities. Furthermore, homologous recombination can occur for transformed DNA. Finally, *P. pastoris* is capable of performing most higher eukaryotic post-translational modifications of secreted proteins.

Once *P. pastoris* was found to be an effective protein expression system, it became employed in both basic and applied science. *P. pastoris* has contributed to the development of drugs for many diseases [6]. For example, *P. pastoris* has been utilized to express Hepatitis B antigen, a subunit vaccine currently on the market in South America [7]. As a second example, monoclonal antibodies, including ones for rheumatoid arthritis and for migraine headaches, have recently been developed [8].

Although the capabilities of *P. pastoris* are vast, Dr. Cregg cautioned that there are certain limitations to this yeast as a protein expression system. For one, it cannot always express any gene of interest: while some proteins have no issues, others may encounter certain obstacles involving RNA stability, glycosylation, protein folding, or secretion. Despite these shortcomings, however,

P. pastoris as an expression system has stood the test of time, and new developments are constantly being introduced in this field. For instance, the *AOX1* promoter, which was once only activated by methanol, has been modified by VALIDOGEN (Grambach, Austria) to express strongly in methanol-free environments.

For this review, the authors, who are undergraduates performing research in the Lin-Cereghino lab, interviewed experts in the *P. pastoris* field to solicit their solutions to some of the most common challenges encountered when trying to produce recombinant proteins in this expression system, such as identifying high expression strains, improving secretion efficiency and decreasing hyperglycosylation. These scientists provided their insights about identifying the cause of each problem as well as their approaches to overcoming these obstacles. The guidance provided by these well-respected authorities should empower both novice and experienced *P. pastoris* users to save time and find greater success in obtaining high yields of active heterologous proteins.

Picking a Secretion Signal

Interviewees: Benjamin S. Glick, PhD, Professor of Molecular Genetics and Cell Biology, University of Chicago and Juan J. Barrero, PhD, Department of Chemical, Biological, and Environmental Engineering, Autonomous University of Barcelona, Cerdanyola del Vallès

Inefficient protein secretion can have several causes in *P. pastoris*. The GFP-HDEL test can confirm if protein translocation from the cytoplasm into the endoplasmic reticulum (ER) is the reason for decreased secretion within 2 to 3 days [9]. GFP-HDEL represents a green fluorescent protein (GFP) that has an HDEL tetrapeptide on the C-terminus for ER retention. The GFP-HDEL test allows researchers to visualize the location of the fluorescent fusion protein, thereby determining whether the translocons in the ER are obstructed, which lead to the presence of GFP-

HDEL in the cytoplasm. According to Dr. Benjamin Glick and Dr. Juanjo Barrero, who developed and were interviewed about this system, the GFP-HDEL test is an inexpensive yet highly efficient method to establish whether translocation from the cytoplasm into the ER is an issue in protein secretion. If this translocation step is the problem, then a new pre-signal sequence named Ost1 can be utilized to resolve this issue and may drastically improve protein secretion.

The inspiration behind using the Ost1 pre-signal sequence was when Dr. Glick recognized that the α-factor secretion signal inhibited the secretion of a monomeric superfolder GFP within both *S. cerevisiae* and *P. pastoris* [10]. Since the α-factor secretion signal—consisting of a 19 amino acid pre-signal and 66 amino acid pro domain—directs post-translational translocation across the ER membrane, it may work well for only some proteins [11]. The nature of the signal sequence can control whether the protein travels via a co-translational or a post-translational pathway [12]. With this knowledge, the traditional pre-signal of the α-factor secretion signal was replaced by the contemporary Ost1 pre-signal sequence, which drives co-translational translocation across the ER membrane. The result was a hybrid secretion signal consisting of the Ost1 pre-signal and the α-factor pro domain [10].

To assess the efficiency of the Ost1 pre-signal sequence, the Glick lab engineered a regulatable, secretory fluorescent protein E2-Crimson, so that its secretory passage through the endoplasmic reticulum could be observed [9]. After discovering that E2-Crimson could become trapped in the secretory pathway, two modifications were made by the Glick lab to resolve the aggregation issue: (1) the use of the Ost1 pre-signal sequence and (2) an allelic variant of a single amino acid (Ser42) in the pro region. The key successes of the experiment showcased that there was improved E2-Crimson secretion with the Ost1 pre-signal pro-α-factor(I) construct. Limited cytoplasmic fluorescence was seen with the pre-Ost1-pro-αf(I) variant, and Leu42Ser was

confirmed as the ideal point mutation to improve protein translocation into the ER. In addition, the researchers further explored the effects of the Ost1 signal sequence beyond model fluorescent proteins and verified that the secretion of BTL2 lipase was also increased with the improved secretion signal. When using the improved Ost1 pre-signal sequence, the stability of any respective secreted protein remains completely normal because the secretion signal is clipped before the protein is secreted.

There has been no evidence yet that suggests that using the Ost1 pre-signal sequence would be any less satisfactory than using the α-factor secretion signal. However, depending on the characteristics of the model protein used, there could be possible limitations when utilizing the Ost1 pre-signal sequence that have yet to be determined. If the chosen protein for secretion contains many disulfide bonds, then the protein may not fold quickly due to its extensive need for post-translational modification, lessening the chance of exhibiting improvement when using the Ost1 pre-signal sequence. In contrast, for proteins that contain fewer disulfide bonds or that oligomerize together, the chances of enhanced secretion are more likely when using the Ost1 pre-signal sequence.

When reflecting back on attempted experiments prior to the reviewed publication, Dr. Barrero mentioned that he had attempted to slow down the translation process by inserting specific amino acids, so that proper folding of the pre-Ost1 signal sequence could be achieved, which was expected to improve translocation [13]. However, with this procedure, Dr. Barrero obtained only a 15% improvement in protein secretion—an insufficient increase that prompted Dr. Barrero to discontinue this line of experimentation. Dr. Barrero disclosed that his most recent publication [14] illustrated that the Ost1 pre-signal sequence and the Ser42 variant displayed significant improvement not only in protein production but also in cell growth rate in a 5 liter bench top

reactor. The Ost1 pre-signal sequence reduced the burden that the cells endured during protein secretion, allowing the cells to grow better overall.

In conclusion, the GFP-HDEL test can indicate whether translocation from the cytoplasm into the ER is the underlying cause of impaired secretion. If this is indeed the problem, Dr. Glick and Barrero's Ost1 pre-signal sequence, which is available in some vectors provided by Addgene (Watertown, MA) and BioGrammatics (Carlsbad, CA), may be the solution.

Overexpressing Chaperones to Enhance Secretion

Interviewee: Carl Batt, PhD, Professor of Food Science, Cornell University

P. pastoris has strong secretory capacities while only secreting low amounts of endogenous proteins [5]. However, just like all other protein expression systems, *P. pastoris* is not without its disadvantages. Overexpression of heterologous proteins can lead to the saturation of its secretory pathway, elevating levels of unfolded proteins [15]. Because a proofreading mechanism prevents these proteins from leaving the endoplasmic reticulum, the efficiency with which a cell is able to fold its secreted proteins is a major factor to increasing the rate of transport of proteins from the ER to the Golgi apparatus and to the extracellular region [16].

Immunoglobulin binding protein (BiP) and protein disulfide isomerase (PDI) are two abundant ER resident proteins that aid folding and therefore facilitate transport of a diverse array of secreted proteins [17]. BiP is classified as an Hsp70 class heat shock protein which binds to the hydrophobic amino acid stretches of immature proteins, consequently stabilizing them. PDI is one of the most abundant proteins in the ER and is an essential protein that is involved in oxidation, reduction, and isomerization of disulfide bonds. PDI inhibits the aggregation of misfolded proteins, even those lacking disulfide bonds, therefore behaving also as a chaperone.

Prior studies on overexpressing BiP and/or PDI in *Saccharomyces cerevisiae* in order to increase protein secretion have yielded promising results. Specifically, overexpression of either BiP or PDI has led to a twofold increase of single-chain antibody fragment (ScFv) in *S. cerevisiae,* and the co-overexpression of both BiP and PDI has resulted into an astounding eightfold increase of protein secretion [18]. Drawing inspiration from this previous research, Dr. Carl Batt and his colleagues attempted to overexpress BiP and PDI with the goal of enhancing secretion of ScFvs in *P. pastoris*.

The Batt group found that, in *P. pastoris,* overexpression of BiP led to a threefold increase in ScFv secretion over the control strain, which was in-line with the results with *S. cerevisiae.* However, unlike baker's yeast, PDI overexpression did not have a positive impact on secretion [16]. Furthermore, unlike the observations made in *S. cerevisiae*, co-overexpression of BiP and PDI did not lead to a synergistic effect on protein secretion. However, PDI overexpression alone led to a threefold increase in BiP levels, suggesting that overproduction of a chaperone can burden the secretory system to a degree that triggers Unfolded Protein Response (UPR) induction, which may have led in turn to lower levels of secreted proteins. Moreover, this study demonstrated the potential for large-scale application: Dr. Batt's team was able to achieve a twofold increase in protein secretion with overexpression of BiP using a fermenter.

Today, as Dr. Batt noted in his interview, the use of *P. pastoris* strains that overexpress BiP and/or PDI to secrete recombinant proteins has become an almost universal standard protocol, but there are several limitations to this strategy which could apply to many recombinant proteins produced in *P. pastoris*. First, the synergistic effect that was expected from co-overexpression of BiP and PDI was not observed. This may be due to the fact that the role of PDI in unfolding and refolding misfolded proteins was nullified by the increase in intracellular BiP levels triggered by

activation of the UPR, which ultimately led to ER-associated degradation (ERAD). Second, despite their many similarities, *P. pastoris* and *S. cerevisiae* are different yeasts. Therefore, as Dr. Batt warned, a particular genetic modification in one organism with promising results may not yield the same outcome in the other. Third, the negative effects of overexpression of PDI on secretion of ScFv may hinge on the structural characteristics of the protein itself, not the inefficacy of the foldase. This is most likely the case, as overexpression of PDI has been proven to be effective in improving secretion of other proteins such as human secretory leukocyte protease inhibitor (SLPI) [19].

Looking ahead, this study suggests several alterations that can be made in this strategy. A weaker promoter on BiP and/or PDI could allow for the right amount of BiP/PDI in the ER that would bring about the desired enhancement of secretion without triggering UPR. Second, the constitutive expression of the gene encoding HAC1, a UPR-regulating transcription factor, has been shown to lessen the unfolded protein burden by attenuating protein synthesis and increasing ER protein efficiency. Taken together, the constitutive overexpression of HAC1, in combination with the co-overexpression of lower levels of BiP/PDI, might be a successful strategy to enhance secretion in *P. pastoris* [20]. Thus, enhanced production of BiP and/or PDI is a strategy to increase protein secretion that is worthy of consideration, but it must be noted that overexpression of both or either of these proteins is not guaranteed to deliver the desired effect for all recombinant proteins.

Screening Transformants Efficiently

Interviewee: Roland Weis, PhD, Head of Operations, VALIDOGEN GmbH

Screening for the best *P. pastoris* strain to secrete a protein of interest is a crucial preliminary step to successful recombinant protein expression [21]. The recent 2020 *Pichia* Expression Kit User Guide (Invitrogen Carlsbad, CA)—a user manual designed to assist scientists that are new to the *Pichia* field—includes protocols describing screening methods using 250mL shake flasks. Although this is a classic screening method, there are other options available that can make the *P. pastoris* screening process much less time consuming and expensive. Dr. Roland Weis produced the 96 deep well plate screening assay: an efficient, low cost screening technique that allows researchers to find a few suitable *P. pastoris* strains that he considers "consistently and reproducibly better than others." The key features of this screening assay include a basic microplate with 96 flat bottom, square wells (acting like baffles) to reduce cell pelleting—which the yeast is prone to—, a reliable laboratory shaker with high speed and humidity control, as well as a plate reading system to determine the amount of secreted protein in the medium [22,23]. (Insider tip: If a lab does not have a humidity-controlled shaker and wants to be more cost effective, it can place water into the perimeter wells of the plate, parafilm the sides, and use any standard shaker incubator).

The idea of the 96 deep well plate screening assay for *P. pastoris* was born when a colleague of Dr. Weis was researching at the California Institute of Technology applying such a cultivation and screening scale with *E. coli*. They wanted to devise a high throughput, microscale screening system for *P. pastoris* that was similar to the already established methods for *E. coli*. Determining the perfect shape and size of the deep well plate along with the ideal shaker speed and volume to fill the wells was very exhaustive but necessary work. Once the design parameters were optimized, Dr. Weis tested this screening assay, which involved growing up strains in glucose medium to accumulate biomass and inducing recombinant expression from an *AOX1* promoter in

methanol medium. Unlike the time-consuming shake flask method, the 96 deep well plate screening assay does not require centrifugation and removal of glucose medium prior to addition of methanol. Instead, this microscale screening technique allows for total depletion of the glucose medium (250µL) by cellular consumption, followed by addition of methanol-containing medium (250µL) to induce protein expression. The induction process takes place for approximately 3 days (including several further additions of methanol-containing medium at lower volume), the cells are pelleted, and secreted recombinant protein is quantitated in the harvested extracellular supernatant. The results were most consistent and reproducible across countless numbers of plates, users, *P. pastoris* strains and target proteins when biomass accumulation was done in 1% glucose media. Stable dissolved oxygen levels and highest recombinant protein expression were also observed when the media had a glucose concentration of 1%. Furthermore, the top performing strains selected by the 96 deep well plate method exhibited success in large scale bioreactors as well, allowing researchers to skip the shake flask screening step altogether.

Although it is typically assumed that increasing *P. pastoris* biomass by using higher glucose concentration in the media will increase recombinant protein productivity [24], this was not the case in the 96 deep well plate screening assay. As expected, higher glucose concentrations resulted in increased yeast biomass, however, protein yield and, most importantly, comparability across the entire 96-well plate was reduced with >1% glucose concentrations. The higher the glucose concentration in the media, the worse the productivity was, indicating that the 1% glucose concentration was the "oxygen preferred state," while 2% and even 1.5% glucose media resulted in substantial necrosis and apoptosis in the wells. This screening method is not for obtaining enormous yields of cells or product; instead, it is, according to Dr. Weis, an ideal system for "comparing clones in a fair way with a methodology or protocol that can be enabled by any lab

simply." He continues to use this procedure routinely at VALIDOGEN to successfully screen strains.

Other screening systems such as the 48 well FlowerPlate, created by m2p-labs (Baesweiler, Germany), allows for controlled feeding, pH, and temperature, as well as the ability to enhance dissolved oxygen within the medium to reduce cell death and increase biomass, but this method is more expensive and laborious. Additionally, other media such as glycerol and sorbitol were tested in the biomass accumulation stage in this screening assay, but glucose proved most effective. Strains featuring recombinant expression driven by constitutive promoters, such as *GAP*, can be screened by the deep well plate assay by maintaining cells on glucose instead of shifting to methanol. In the future, Dr. Weis aims to develop a complex media screening method for proteins that will not secrete efficiently on minimal media along with liquid glucose time-release methods to increase protein yield.

When asked if researchers not utilizing the 96 deep well plate screening assay may be missing out on finding the most productive strains, Dr. Weis replied with a resounding "YES!" Overall, if a lab is struggling to produce a protein of interest, it is possible that a different *P. pastoris* strain, which is superior for secreting that particular protein, is actually present on a transformation plate. The high throughput, inexpensive 96 deep well plate screening assay is a more consistent and efficient method compared to classic shake flask screening procedures to identify that optimal strain and enable secreted production of the protein of interest, as well as increase recombinant protein yield.

Overcoming Clonal Variation

Interviewee: Rochelle Aw, PhD, Research Associate, Faculty of Engineering, Department of Chemical Engineering, Imperial College London

Clonal variation is defined as when there are observed differences in expression levels of a heterologous protein amongst strains that have the same copy number and the same insertion site of the recombinant gene. Clonal variation in *P. pastoris* makes it difficult to attain maximum efficiency in protein secretion. In contrast, this is not a problem in *Escherichia coli* or *Saccharomyces cerevisiae*. While *P. pastoris* was described as showing clonal variation over 30 years ago [25], only two other articles carried out comprehensive investigations into the yeast's clonal variation [26,27].

A possible solution to clonal variation would be to identify a specific biomarker that characterizes high-level secretors. If a biomarker was classified, the process of screening thousands of colonies to identify a high-level secretor would be streamlined into a timely and cost-effective process. Furthermore, if a specific quality of a high secretor was pinpointed, then genetic engineering could be utilized to greatly improve the secretion efficiency of other transformants. Dr. Rochelle Aw undertook the challenge to identify the underlying cause of clonal variation and to identify a biomarker that characterized a high-secretor [28]. As she stressed in her interview, Dr. Aw wanted not only "to ask the fundamental biological questions behind clonal variation" but "make it easier to find that great secreter...quickly, like in a couple of hours, without the laborious technique of expressing the protein."

To investigate clonal variation, three high-, medium-, and low-level human serum albumin (HSA) secreting strains were isolated by Dr. Aw [28]. The copy number and the insertion sites of the expression plasmids, which relied on *AOX1* promoter-based expression, were the same in all nine strains. The HSA-secreting cultures were induced on methanol and measured for the secretion

of only fully folded, intact protein, not total protein, which would have included misfolded or partially degraded species. The results were analyzed with titer analysis, qPCR, flow cytometry, transcriptomic analysis, gene expression, and pathway analysis.

Several factors were ruled out as the cause of clonal variation in *P. pastoris* among the HSA secreting strains. First, their results suggested that the unfolded protein response (UPR) did not factor into clonal variation. Specifically, protein degradation via the ER-associated degradation (ERAD) pathway did not contribute to the low protein expression that was observed. Second, a transcriptomic analysis revealed that the high secreting strains did not contain the highest mRNA levels, implying that there was no correlation between mRNA levels and high protein secretion levels. Third, cell viability did not correlate with protein export; in fact, most of the high secretors showed an increased proportion of dead cells compared to the other stains when grown on methanol, which was a surprising finding.

The positive findings of Dr. Aw's study included that increased oxidative phosphorylation levels demonstrated a correlation with high levels of secretion of fully folded, intact HSA. Due to the increased demand for recombinant protein synthesis, the cells could be increasing their oxidative phosphorylation levels to produce more ATP. Additionally, the *SKP1* gene, which is involved in oxidative phosphorylation, was upregulated in high-secreting strains, making it a possible candidate for a biomarker that relates to high secretors. Even with the promising discovery of the *SKP1* gene, however, at this point there has been no distinctive biomarker established that concretely characterizes high-level *Pichia* secretors.

Pondering the underlying cause, Dr. Aw suggested that non-homologous recombination may factor into clonal variation. She mentioned that a researcher must be careful to purify linearized plasmid prior to transformation of *P. pastoris* because the presence of contaminating

E.coli DNA (from a plasmid miniprep) may lead to integration of the bacterial genomic fragments into random loci of the *P. pastoris* genome, which could contribute to clonal variation. Furthermore, she affirmed the commonly held practice of banking a high secreting strain in a minus 80 degrees Celsius freezer immediately after isolation because further nonhomologous recombination integration may occur while it is metabolically active [29].

Because protein secretion efficiency is dependent on the structure of the individual protein, Dr. Aw recommended that these experiments should be repeated with more proteins beyond HSA, with the addition of using RNA-sequencing and proteomics to analyze the results. To conclude, although a high-secretor characterizing biomarker has not been confirmed at this point, Dr. Aw produced compelling information on which factors did not cause clonal variation, encouraging further research on this novel topic within the *Pichia* community.

Improving Glycosylation

Interviewee: Tom Chappell, PhD, Co-founder and Principal, BioGrammatics

Glycosylation is an integral form of posttranslational modification of secreted and membrane-associated eukaryotic proteins. In order to attain efficient folding of certain glycoproteins, cells usually must perform the correct N-glycosylation on these peptides as they transit the secretory organelles [30]. Because N-glycosylation is not found in most prokaryotic hosts, such as *E. coli*, *P. pastoris* is often used as a protein expression system for recombinant proteins requiring this post-translational modification [31].

Despite the ability of *P. pastoris* to provide N-glycosylation, genetic modifications to the yeast are necessary because the wild type yeasts often display heterogeneous hyperglycosylation of some recombinant proteins [32]. This is problematic in 2 ways: 1) *P. pastoris* modifies its

proteins with mannose while humans produce more complex N-glycan structures such as galactose and sialic acid, and 2) *P. pastoris* tends to produce a diverse population of the secreted protein that vary in the number of sugar units added to the given protein. Some these glycosylated proteins are still functional and suitable for their intended application. However, results of these differences may include a decreased half-life of the desired protein in the body, increased immunogenicity, and potential negative effects on protein function and folding [33].

To combat this problem, scientists have attempted to convert the *P. pastoris* high mannose N-glycosylation into mammalian N-glycosylation [34]. This pathway engineering consists of two parts: 1) inactivation of the main *P. pastoris* glycosyltransferases which lead to the accumulation of a glycan precursor structure that is common with the human pathway, such as Man8GlcNAc2, and 2) introduction of heterologous glycosidases and glycosyltransferases which convert the common precursor to the desired structures (often mammalian-type high-mannose, hybrid or complex N-glycans).

Many researchers use the *Pichia* GlycoSwitch® system to alter the *P. pastoris'* N-glycosylation mechanisms [35]. Seven unique strains of *P. pastoris* were originally created for this purpose, with each strain exhibiting unique results with respect to secretion and robustness [36]. For one such strain, M8, homologous recombination was used to knock out the endogenous *OCH1* gene to yield secreted recombinant proteins with Man8GlcNac2 as the major glycosylation structure. A human α1,2 mannosidase gene was then added to remove all α1,2 linked mannosidases to produce Man5GlcNAc2 on the desired proteins, generating the M5 strain. The remaining 5 strains, which produce proteins with hybrid/complex types of glycosylation, were developed with the subsequent addition of more glycosylation modifying genes.

According to senior scientist Dr. Tom Chappell, who was interviewed about his experience with the *Pichia* GlycoSwitch® system at BioGrammatics, the use of the M5 strain host is a good starting point for trying to produce recombinant human proteins with human-like glycosylation. The proteins produced in M5 must be analyzed to determine the identity and homogeneity of the glycosylation structures on the recombinant proteins. Although there have been successes, further research is necessary to perfect the desired glycosylation in *P. pastoris*. For instance, the proteins generated from M5 have exhibited protein degradation, and there is rarely homogeneity among the population of secreted proteins [37]. One reason is that the new N-glycan structures inadvertently might be substrates for one or more of the many endogenous glycosyltransferases, resulting in the formation of novel, undesired glycan structures. In the *Pichia* GlycoSwitch® M5 strain, Man5GlcNAc2 typically makes up >80 % of all N-glycans on proteins [33]. In the case of mouse interleukin-22 (mIL-22) produced in the M5 strain, glycan profiling revealed two major species: Man5GlcNAc2 and an unexpected, partially α-mannosidase-resistant structure which was Man5GlcNAc2 modified with a Glcα-1,2-Manβ-1,2-Manβ-1,3-Glcα-1,3-R tetra-saccharide [37].

Dr. Chappell admitted that the current *Pichia* GlycoSwitch® strains are not perfect, but several strategies can be tried to combat some of these setbacks. One way of dealing with interference by endogenous glycosyltransferases is to mutate the amino acid residues that are glycosylated, with the hope that function is not compromised by the substitution. Another solution is to modify fermentation conditions so that more homogeneous N-glycan profiles can be obtained. A third alternative is to outcompete these endogenous enzymes with overexpression of heterologous glycosyltransferases with a similar acceptor substrate specificity. For example, the overexpression of the human N-acetylglucosaminyltransferase I (hGnT-I) was able to suppress the presence of undesired glycan species in the mIL-22-secreting, M5 strain [37]. However, Dr.

Chappell warns that there is a limit to the genetic modifications that can be tolerated in the secretion apparatus of *P. pastoris* because disruption of too many host glycosylation enzyme genes leads to slower growing, less productive strains. The glycosylation changes affect the function of the endogenous cell wall and membrane proteins needed for viability. Thus, engineering the N-glycosylation pathway in *P. pastoris* to produce particular human-type structures is still a work in progress; however, as time goes on, researchers remain optimistic about the future prospects of this important process.

Going Large Scale for Protein Production

Interviewee: Kerry Love, PhD, Technical Program Manager, Love lab, Koch Institute for Integrative Cancer Research. Massachusetts Institute of Technology and CEO, Sunflower Therapeutics

Without an automated manufacturing system providing consistent and reproducible protocols, developing potent and pure protein biologics at high yields in *P. pastoris* can pose a very difficult task [38,39]. Inventing a dependable and adaptable system for laboratory-scale production of recombinant biopharmaceuticals in *P. pastoris* is a daunting task, but fortunately, much of the work has already been done. A group including Dr. Kerry Love spent three years developing the Integrated Scalable Cyto-Technology (InSCyT) system in J. Christopher Love's lab at the Koch Institute of Massachusetts Institute of Technology. The InSCyT system is described as "an automated, benchtop, multiproduct manufacturing system…for end-to-end production of hundreds to thousands of doses of clinical-quality protein biologics" [40]. Unlike other technologies available for on-demand production of biomolecules, the InSCyT system yields products which have the safety, purity, and potency to meet the standards required for clinical use.

In her interview, Dr. Love stated that the idea of the InSCyT system was inspired by a DARPA-sponsored program called Biologically Derived Medicines On-Demand, which was developed by Colonel Jeff Ling, a United States Army medical doctor who treated patients in various regions of the Middle East. Due to routine drug shortages, Dr. Ling conceived the idea of on-demand manufacturing of medicines to "meaningfully address the local needs of the people he was treating." This sentiment impelled the research team including Dr. Love to "create new manufacturing solutions that increase accessibility and reduce the cost of biologic drugs." Thus, InSCyT was born, offering a compact system—with an approximate length of 5 feet (150cm) and a height of 3 feet (90cm)—for laboratory use that is geared toward research and process development, and is useful for "[organizations] that want to create a biologic product for themselves."

The fully integrated and automated aspects of the InSCyT system provide continuous production of biologic products with minimal human intervention, thus reducing user-induced error and processing time. Implementation of in-tank perfusion within the benchtop bioreactor of the InSCyT system allows for control of pH, temperature, dissolved oxygen, impeller speed as well as input and output flow rates. This perfusion-adapted bioreactor permits inline pH adjustments prior to chromatographic separations, thus eliminating the need for lengthy purification processes that require intermediate adjustments between steps. Additionally, the product purification process does not require affinity tags, which could affect protein function and lead to harmful impurities, and provides the option of straight-through chromatography, in which the eluate from one column passes directly onto another column without adjustment [41]. A tangential flow-filtration system is also incorporated in the InSCyT system, permitting the formulation of the biologics to a final dosage, liquid form. Together, these efficient automated

modules make the InSCyT system more consistent and reproducible than other options, such as microfluidics or in vitro transcription and translation.

The InSCyT system has proven itself to be successful for the production of recombinant therapeutic proteins for pre-clinical trials, specifically for the production of human growth hormone (hGH), interferon-α (IFNα-2b), and granulocyte-colony stimulating factor (G-CSF) [40]. Each of these biologics produced in the InSCyT system were comparable to those currently on the market for parenteral use, meeting the strict regulations implemented by agencies like the FDA and EMA. Given InSCyT's design, the same hardware can be modified to yield multiple different products. This was highlighted by the ability to adapt the InSCyT system to manufacture IFNα-2b shortly after the production of hGH, with only minor adjustments to the purification processes and without changing the core system. When making these adaptations to the production and purification protocols, it is crucial, notes Dr. Love, to understand the biology behind the "critical quality attributes" of the desired product. For example, hGH possesses specific residues that are targets of oxidation and deamidation, which can vastly affect the purity of the product. Awareness of how these characteristics are affected by the bioprocess impacts the clinical outcome of the product, and thankfully the InSCyT system process can be minimally modified to prevent these undesirable effects. In the case of hGH, simply switching the cells to a defined growth medium and adjusting DO levels resulted in a significant reduction of impurities. As Dr. Love emphasized, "if you want a high-quality product, then you need a high-quality bioprocess."

Dr. Love currently serves as the CEO and President of Sunflower Therapeutics, which was awarded an $8.1M grant from the Bill & Melinda Gates Foundation to adapt the InSCyT system for commercial use. The InSCyT system has been used to develop the Daisy^{TM} manufacturing system prototype, with the primary aim to "support low- and middle-income countries in the

discovery, development and manufacturing of protein-based biologics" (https://sunflowertx.com/sunflower-therapeutics-awarded-8-1m-grant/). The Daisy™ system is expected to be commercially available in the near future. With its compact size and powerful capabilities, this system should be considered by any lab that aims to produce high quality protein biologics and biopharmaceuticals on-demand, as well as those conducting critical vaccine research.

Seeking Company Help: The Business of *Pichia pastoris*

Interviewee: Knut Madden PhD, CEO, BioGrammatics

Heterologous protein secretion using *P. pastoris* as the host organism is faced with many challenges. The main obstacles include the identification of an effective secretion signal, expression plasmid or genetically modified strain [42]. This is not straightforward; for instance, a certain secretion signal may be effective at secreting a particular protein but not others [43,44]. During the last 30 years, great leaps have been made with *P. pastoris* technology, leading to greater success in expressing different types of proteins. However, such discoveries do not translate to efficient production of all proteins, and, because *P. pastoris* is not that common, one cannot always find an experienced user to consult with among colleagues or even in a professional network.

Due to the complexity of this issue, it may be wise to reach out to various companies that have extensive experience working with this yeast. The companies can provide advice and tools associated with vector design, strain development, and expression strategies as well as perform the actual recombinant protein production. These companies include Validogen (https://www.validogen.com), BISY (https://www.bisy.at), Lonza (https://www.lonza.com), and BioGrammatics (https://www.BioGrammatics.com). BioGrammatics Inc., founded by Drs.

Thomas Chappell, Knut Madden, Ilya Tolstorukov and Jim Cregg, offers reagents and contract services to expand the use of *P. pastoris* as a host system. According to the CEO, Dr. Madden, there are "no concrete rules" in developing a strategy for a specific protein due to the intricate nature of heterologous protein production. For instance, to identify which secretion signal to use, he may suggest the use of the native secretion signal, followed by trying out a heterologous leader peptide, such as α-factor secretion leader, with or without its pro peptide, or a different secretion signal altogether.

Furthermore, Dr. Madden feels that an advantage of seeking help from a company is that the company is led by individuals who are committed to staying up-to-date with the latest innovations in the *P. pastoris* field. Therefore, many of the strains and vectors that are offered by BioGrammatics are often newer and more suitable for their intended purposes than those referenced in scientific literature. For instance, the company's expression plasmid choices include the more novel Ost1 and Aga2 signal peptides along with the well-established α-factor secretion leader. In addition, BioGrammatics holds the licensure for distribution of the SuperMan5 strain, part of the *Pichia* GlycoSwitch® technology that provides more human-like glycosylation for secreted proteins [45]. In fact, Dr. Madden pointed out that BioGrammatics assisted in production of the HSV2 vaccine in Superman5, which highlights the use of *Pichia* GlycoSwitch® technology.

While companies offer tools and services for *P. pastoris* expression to both academic and industrial users, Dr. Madden emphasizes that the potential client must understand the policies surrounding their use. For example, all consultations between the client and BioGrammatics are considered confidential. BioGrammatics will provide technical support and troubleshooting for their products, with the promise to refund money if the clients are not satisfied. Furthermore, there are no licensing fees required for academic purposes, but these fees apply if the recombinant

protein is intended for profit-making purposes. However, every company that offers *P. pastoris* licenses has their own rules regarding licensing. For example, each commercial entity that purchases a *Pichia* Classic Kit or a *Pichia* GlycoSwitch® Kit from Research Corporation Technologies is allowed a single, one year-long evaluation at no cost, but at the end of this period, the user will need to acquire a commercial license to continue working with the system (https://pichia.com/licensing/licensing-faqs/). On the other hand, BioGrammatics does not require a client who hires them for contract services to buy a license from them, but as Dr. Madden explains, "BioGrammatics works hard so that the client will have commercial success and then want to buy a license to use the expression strain BioGrammatics has developed for them. BioGrammatics will grant a commercial license for less than it would cost the client to rebuild the strain with similar technology from another company that offers licenses."

The future is bright for companies and researchers working with *P. pastoris* to produce recombinant proteins. These interactions place an emphasis on the advancement of science and the optimization of contemporary techniques. Thus, reaching out to these companies is a worthy investment of time. Collaboration between these companies and researchers can result in a synergistic effect, bringing both parties together and ultimately leading to a deeper knowledge of protein expression in *P. pastoris*.

Conclusion

Because the success of *P. pastoris* is dependent on the structural features of the individual protein, system users must realize that some strategies may improve recombinant protein production while others will have no effect or even decrease expression levels. However, if

researchers are committed to and persist in exploring the options, such as those detailed in this review, they will most likely attain higher yields of their biologically active proteins.

References

1. Cregg JM, Barringer KJ, Hessler AY, Madden KR (1985) *Pichia pastoris* as a host system for transformations. Mol Cell Biol 5:3376-3385

2. Cregg JM, Russell KA (1998) Transformation. In: Higgins DR, Cregg JM (eds) *Pichia* Protocols, vol 103. Humana Press, Totowa, New Jersey, pp 27-39

3. Cregg JM (2007) DNA-mediated transformation. Methods Mol Biol 389:27-42. doi:10.1007/978-1-59745-456-8_3

4. Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR (1987) Expression of the *lacZ* gene from two methanol-regulated promoters in *Pichia pastoris*. Nucleic Acids Res 15 (9):3859-3876

5. Cereghino JL, Cregg JM (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol Rev 24 (1):45-66

6. Weinacker D, Rabert C, Zepeda AB, Figueroa CA, Pessoa A, Farías JG (2013) Applications of recombinant *Pichia pastoris* in the healthcare industry. Braz J Microbiol 44 (4):1043-1048. doi:10.1590/s1517-83822013000400004

7. Wang M, Jiang S, Wang Y (2016) Recent advances in the production of recombinant subunit vaccines in *Pichia pastoris*. Bioengineered 7 (3):155-165. doi:10.1080/21655979.2016.1191707

8. Pellesi L, Guerzoni S, Pini LA (2017) Spotlight on Anti-CGRP Monoclonal Antibodies in Migraine: The Clinical Evidence to Date. Clin Pharmacol Drug Dev 6 (6):534-547. doi:10.1002/cpdd.345

9. Barrero JJ, Casler JC, Valero F, Ferrer P, Glick BS (2018) An improved secretion signal enhances the secretion of model proteins from *Pichia pastoris*. Microb Cell Fact 17 (1):161. doi:10.1186/s12934-018-1009-5

10. Fitzgerald I, Glick BS (2014) Secretion of a foreign protein from budding yeasts is enhanced by cotranslational translocation and by suppression of vacuolar targeting. Microb Cell Fact 13 (1):125. doi:10.1186/s12934-014-0125-0

11. Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, Urdea MS, Valenzuela P, Barr PJ (1984) Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 81 (15):4642-4646

12. Ng DT, Brown JD, Walter P (1996) Signal sequences specify the targeting route to the endoplasmic reticulum membrane. J Cell Biol 134 (2):269-278. doi:10.1083/jcb.134.2.269

13. Pechmann S, Chartron JW, Frydman J (2014) Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. Nat Struct Mol Biol 21 (12):1100-1105. doi:10.1038/nsmb.2919

14. Barrero JJ, Pagazartaundua A, Glick BS, Valero F, Ferrer P (2021) Bioreactor-scale cell performance and protein production can be substantially increased by using a secretion signal that drives co-translational translocation in *Pichia pastoris*. N Biotechnol 60:85-95. doi:10.1016/j.nbt.2020.09.001

15. Damasceno LM, Huang CJ, Batt CA (2012) Protein secretion in *Pichia pastoris* and advances in protein production. Appl Microbiol Biotechnol 93 (1):31-39. doi:10.1007/s00253-011-3654-z

16. Damasceno LM, Anderson KA, Ritter G, Cregg JM, Old LJ, Batt CA (2007) Cooverexpression of chaperones for enhanced secretion of a single-chain antibody fragment in *Pichia pastoris*. Appl Microbiol Biotechnol 74 (2):381-389. doi:10.1007/s00253-006-0652-7

17. Brodsky JL, Skach WR (2011) Protein folding and quality control in the endoplasmic reticulum: Recent lessons from yeast and mammalian cell systems. Curr Opin Cell Biol 23 (4):464-475. doi:10.1016/j.ceb.2011.05.004

18. Shusta EV, Raines RT, Pluckthun A, Wittrup KD (1998) Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. Nature Biotechnology 16 (773-777)

19. Li Z, Moy A, Gomez SR, Franz AH, Lin-Cereghino J, Lin-Cereghino GP (2010) An improved method for enhanced production and biological activity of human secretory leukocyte protease inhibitor (SLPI) in *Pichia pastoris*. Biochem Biophys Res Commun 402 (3):519-524. doi:10.1016/j.bbrc.2010.10.067

20. Lin XQ, Liang SL, Han SY, Zheng SP, Ye YR, Lin Y (2013) Quantitative iTRAQ LC-MS/MS proteomics reveals the cellular response to heterologous protein overexpression and the regulation of HAC1 in *Pichia pastoris*. J Proteomics 91:58-72. doi:10.1016/j.jprot.2013.06.031

21. Olsen M, Iverson B, Georgiou G (2000) High-throughput screening of enzyme libraries. Curr Opin Biotechnol 11 (4):331-337. doi:10.1016/s0958-1669(00)00108-7

22. Weis R, Luiten R, Skranc W, Schwab H, Wubbolts M, Glieder A (2004) Reliable high-throughput screening with *Pichia pastoris* by limiting yeast cell death phenomena. FEMS Yeast Res 2:179-189

23. Weis R (2019) High-Throughput Screening and Selection of *Pichia pastoris* Strains. Methods Mol Biol 1923:169-185. doi:10.1007/978-1-4939-9024-5_7

24. Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, Fotiadis D, Kovar K (2015) Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. Biotechnol Adv 33 (6 Pt 2):1177-1193. doi:10.1016/j.biotechadv.2015.05.008

25. Cregg JM, Madden KR, Barringer KJ, Thill GP, Stillman CA (1989) Functional characterization of the two alcohol oxidase genes from the the yeast *Pichia pastoris*. Mol Cell Biol 9:1316-1323

26. Viader-Salvadó JM, Cab-Barrera EL, Galán-Wong LJ, Guerrero-Olazarán M (2006) Genotyping of recombinant *Pichia pastoris* strains. Cell Mol Biol Lett 11 (3):348-359. doi:10.2478/s11658-006-0029-z

27. Schwarzhans JP, Wibberg D, Winkler A, Luttermann T, Kalinowski J, Friehs K (2016) Integration event induced changes in recombinant protein productivity in *Pichia pastoris* discovered by whole genome sequencing and derived vector optimization. Microb Cell Fact 15:84. doi:10.1186/s12934-016-0486-7

28. Aw R, Barton GR, Leak DJ (2017) Insights into the prevalence and underlying causes of clonal variation through transcriptomic analysis in *Pichia pastoris*. Appl Microbiol Biotechnol 101 (12):5045-5058. doi:10.1007/s00253-017-8317-2

29. Sunga AJ, Tolstorukov I, Cregg JM (2008) Posttransformational vector amplification in the yeast *Pichia pastoris*. FEMS Yeast Res 8 (6):870-876

30. Aebi M (2013) N-linked protein glycosylation in the ER. Biochim Biophys Acta 1833 (11):2430-2437. doi:10.1016/j.bbamcr.2013.04.001

31. Bretthauer RK, Castellino FJ (1999) Glycosylation of *Pichia pastoris*-derived proteins. Biotechnol Appl Biochem 30:193-200

32. Gong B, Cukan M, Fisher R, Li H, Stadheim TA, Gerngross T (2009) Characterization of N-linked glycosylation on recombinant glycoproteins produced in *Pichia pastoris* using ESI-MS and MALDI-TOF. Methods Mol Biol 534:213-223. doi:10.1007/978-1-59745-022-5_16

33. Jacobs PP, Inan M, Festjens N, Haustraete J, Van Hecke A, Contreras R, Meagher MM, Callewaert N (2010) Fed-batch fermentation of GM-CSF-producing glycoengineered *Pichia pastoris* under controlled specific growth rate. Microb Cell Fact 9:93. doi:10.1186/1475-2859-9-93

34. Choi BK, Bobrowicz P, Davidson RC, Hamilton SR, Kung DH, Li H, Miele RG, Nett JH, Wildt S, Gerngross TU (2003) Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. Proc Natl Acad Sci 100 (9):5022-5027

35. Laukens B, De Wachter C, Callewaert N (2015) Engineering the *Pichia pastoris* N-Glycosylation Pathway Using the GlycoSwitch Technology. Methods Mol Biol 1321:103-122. doi:10.1007/978-1-4939-2760-9_8

36. Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N (2009) Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. Nat Protoc 4 (1):58-70. doi:nprot.2008.213 [pii] 10.1038/nprot.2008.213

37. Laukens B, Jacobs PP, Geysens K, Martins J, De Wachter C, Ameloot P, Morelle W, Haustraete J, Renauld JC, Samyn B, Contreras R, Devos S, Callewaert N (2020) Off-target glycans encountered along the synthetic biology route toward humanized N-glycans in *Pichia pastoris*. Biotechnol Bioeng 117 (8):2479-2488. doi:10.1002/bit.27375

38. Tripathi NK, Shrivastava A (2019) Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. Front Bioeng Biotechnol 7:420. doi:10.3389/fbioe.2019.00420

39. Love KR, Dalvie NC, Love JC (2018) The yeast stands alone: the future of protein biologic production. Curr Opin Biotechnol 53:50-58. doi:10.1016/j.copbio.2017.12.010

40. Crowell LE, Lu AE, Love KR, Stockdale A, Timmick SM, Wu D, Wang YA, Doherty W, Bonnyman A, Vecchiarello N, Goodwine C, Bradbury L, Brady JR, Clark JJ, Colant NA, Cvetkovic A, Dalvie NC, Liu D, Liu Y, Mascarenhas CA, Matthews CB, Mozdzierz NJ, Shah KA, Wu SL, Hancock WS, Braatz RD, Cramer SM, Love JC (2018) On-demand manufacturing of clinical-quality biopharmaceuticals. Nat Biotechnol. doi:10.1038/nbt.4262

41. Crowell LE, Rodriguez SA, Love KR, Cramer SM, Love JC (2021) Rapid optimization of processes for the integrated purification of biopharmaceuticals. Biotechnol Bioeng. doi:10.1002/bit.27767

42. Yang Z, Zhang Z (2018) Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: A review. Biotechnol Adv 36 (1):182-195. doi:10.1016/j.biotechadv.2017.11.002

43. Aw R, McKay PF, Shattock RJ, Polizzi KM (2018) A systematic analysis of the expression of the anti-HIV VRC01 antibody in *Pichia pastoris* through signal peptide optimization. Protein Expr Purif 149:43-50. doi:10.1016/j.pep.2018.03.013

44. Neiers F, Belloir C, Poirier N, Naumer C, Krohn M, Briand L (2021) Comparison of Different Signal Peptides for the Efficient Secretion of the Sweet-Tasting Plant Protein Brazzein in. Life (Basel) 11 (1). doi:10.3390/life11010046

45. Ahmad M, Hirz M, Pichler H, Schwab H (2014) Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol 98 (12):5301-5317. doi:10.1007/s00253-014-5732-5