Universitat Autònoma de Barcelona

Multidisciplinary approach for recombinant protein production bioprocess design with classic and novel expression systems in Pichia pastoris

Javier Garrigós Martínez

Multidisciplinary approach for recombinant protein production bioprocess design with classic and novel expression systems in *Pichia pastoris*

Memoria para optar al Grado de Doctor por la Universidad Autónoma de Barcelona dentro del Programa de Doctorado en Biotecnología, bajo la dirección de los doctores:

Francisco Valero Barranco, José Luis Montesinos Seguí y Xavier Garcia Ortega

Javier Garrigós Martínez

Bellaterra, 2020

Francisco Valero Barranco, Catedrático, José Luis Montesinos Seguí, Profesor Titular, miembros del Departamento de Ingeniería Química, Biológica y Ambiental de la Universidad Autónoma de Barcelona y Xavier Garcia Ortega, Profesor Agregado del Departamento de Biociencias de la Universidad de Vic – Universidad Central de Cataluña

CERTIFICAMOS:

Que el biotecnólogo Javier Garrigós Martínez, ha realizado, bajo nuestra dirección, en los laboratorios del Departamento de Ingeniería Química, Biológica y Ambiental de la Universidad Autónoma de Barcelona, el trabajo con el título: "Multidisciplinary approach for recombinant protein production bioprocess design with classic and novel expression systems in *Pichia pastoris"* que se presenta en esta memoria, la cual constituye su Tesis para optar al Grado de Doctor por la Universidad Autónoma de Barcelona dentro del programa de doctorado de Biotecnología.

Y para dejar constancia y que tenga los efectos que corresponda, firmamos la presente en Bellaterra, 18 de Noviembre de 2020

Dr. Francisco Valero Barranco Dr. José Luis Montesinos Seguí

Dr. Xavier Garcia Ortega

AGRADECIMIENTOS

Como lo bueno si breve, dos veces bueno, no me explayaré demasiado. Me gustaría agradecer a mis directores de tesis Paco, José Luis y Xavi por todo lo que me han enseñado en este tiempo. Creo que de esta etapa salgo mucho más formado, y eso se lo debo fundamentalmente a ellos. Respecto al ámbito personal, el buen rollo que generamos en los pasillos del departamento, sin duda, favoreció que el trabajo de todos se hiciese mucho más ameno. Por eso mismo, le doy mil gracias a los chavales, por conseguir que mi estancia en ese lugar se hiciera tan divertida.

The present thesis is focused on the characterization of alternative expression systems used for recombinant protein production (RPP) in the methylotrophic yeast *Pichia pastoris*. Over this whole work, the integration of results from different fields —bioprocess engineering and gene regulation— is attempted in order to fill the gaps that frequently come up when designing a RPP bioprocess.

In the first chapter of the thesis, the classical P_{AOX1}-based expression system is thoroughly studied. Specifically, through a set of chemostat cultivations of two clones expressing the *Candida rugosa* lipase 1 (Cr1l) with different gene dosage, it was determined the interrelation of three important factors in RPP processes such as specific growth rate (μ), heterologous gene relative transcript levels (RTL) and specific product generation rate (q_p). Moreover, the expression of a crucial transcription factor of the methanol utilization pathway (*MIT1*) was also determined in chemostat cultivations. Once the optimal operation conditions were identified in steady state conditions, fed-batch cultivations were conducted to validate the clones' behaviour observed previously in terms of both physiological state and Crl1 production kinetics.

The inherent drawbacks of using the powerful P_{AOX1}-based expression system —mainly derived from the use of methanol as carbon source, electron source and RPP inducer— has forced the *Pichia* community to investigate and develop alternative methanol-free expression systems. Due to that, in the second chapter of the thesis, two novel expression systems, based on the P_{PDF} and P_{UPP}, were similarly characterized in order to determine whether they can compete in terms of protein production with the widely used P_{GAP}, usually considered the methanol-free reference, for the production of the Lipase B from *Candida antarctica* (CalB). All the three expression systems performance was firstly compared in chemostat cultivations, which enabled to shed light on the influence of μ and *CALB* expression on the CalB production kinetics. Additionally, the clones harboring the

novel expression system were cultivated in fed-batch mode at the optimal conditions observed in chemostat in order to test their potential bioprocess scalability.

Finally, in the last chapter of the thesis, the production of active whole cell biocatalyst based on the human cytochrome P450 2C9 (CYP2C9) in *P. pastoris* was afforded. The coexpression of the protein along with its redox partner (cytochrome P450 reductase, CPR) was achieved by means of a bidirectional promoter system. After a screening phase in which up to 8 promoters were tested for CYP2C9/CPR production, the combination that provided the best balance was selected for subsequent bioprocess optimization experiments. In this way, the influence of important bioprocess parameters — pH, μ and methanol addition— on active CYP2C9/CPR whole cell biocatalyst production was determined. Finally, the efficiency of *P. pastoris* whole cell biocatalyst based on CYP2C9/CPR was tested in a proof of concept reaction of interest, in which ibuprofen is hydroxylated into its oxidized derivatives.

La presente tesis doctoral se centra en la caracterización de sistemas de expresión utilizados para la producción de proteínas recombinantes (RPP) en la levadura metilotrófica *Pichia pastoris.* A lo largo de todo este trabajo se integran los resultados de diferentes campos —ingeniería de bioprocesos y regulación génica— con el fin de ampliar la información sobre los sistemas de expresión analizados y, así, reducir la incertidumbre al diseñar un bioproceso RPP.

En el primer capítulo de la tesis, se estudia el sistema clásico de expresión basado en el $P_{AOX1.}$ En un primer paso, se cultivaron en quimiostato dos clones productores de la Lipasa 1 de *Candida rugosa* (Cr1l) con diferente dosis génica. De esta manera, se determinó la interrelación de tres factores importantes en los bioprocesos de RPP tales como la velocidad específica de crecimiento (μ), los niveles de expresión relativa del gen heterólogo (RTL) y la velocidad específica de generación de producto (q_p). Además, también se monitorizó la expresión de un factor de transcripción importante para la vía de utilización del metanol (*MIT1*). Una vez establecidas las condiciones óptimas de operación en continuo, se llevaron a cabo cultivos en *fed-batch* para validar y comparar el comportamiento de los clones productores observado, tanto desde el punto de vista de su estado fisiológico como de la cinética de producción de Crl1.

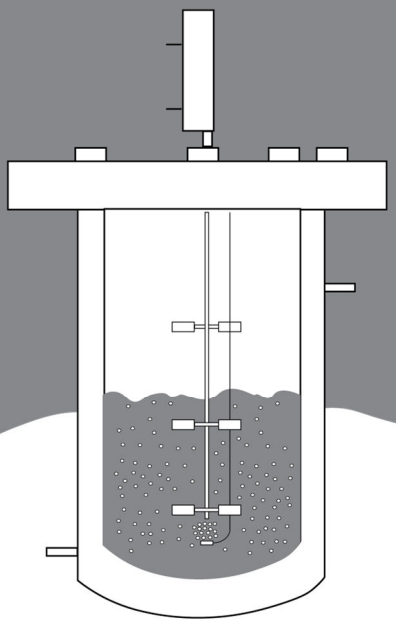
Los inconvenientes del uso del potente sistema de expresión basado en P_{AOX1}— principalmente derivado del uso del metanol como fuente de carbono, fuente de electrones e inductor de RPP— han llevado a la comunidad de *Pichia* a investigar y desarrollar sistemas de expresión alternativos que no dependan de la adición de metanol. Debido a ello, en el segundo capítulo de la tesis, se caracterizaron dos nuevos sistemas de expresión, basados en el P_{PDF} y $P_{UPP,}$ para determinar si pueden competir con el ampliamente utilizado P_{GAP} en la producción de la Lipasa B de *Candida antarctica* (CalB). El comportamiento de los tres sistemas de expresión se comparó, en primer lugar,

en quimiostato. Esto permitió obtener información sobre la influencia de la μ y la expresión de *CALB* en la cinética de producción de la proteína. Además, los clones productores de la proteína recombinante bajo los nuevos sistemas de expresión se cultivaron en *fed-batch*, en las condiciones óptimas observadas previamente, con el fin de probar la potencial escalabilidad del bioproceso.

En el último capítulo de la tesis se optimizó, utilizando células enteras de *P. pastoris* (*whole cells*), la producción del citrocromo P450 2C9 humano (CYP2C9). La coexpresión de esta enzima de interés industrial junto al dominio donador de electrones (*cytochrome P450 reductase*, CPR) se realizó mediante un novedoso sistema de promotores bidireccionales. Después de una fase de *screening* en la que se probaron hasta 8 promotores para la producción de CYP2C9/CPR, se seleccionó la combinación que proporcionó la mayor actividad oxidasa, y se realizaron una serie de cultivos para optimizar el bioproceso. De este modo, se determinó la influencia de parámetros importantes del bioproceso —μ, pH y estrategia de adición de metanol— en la producción de biocatalizador. Finalmente, la eficiencia del biocatalizador obtenido se testó en la reacción de hidroxilación del ibuprofeno.

Table of contents

INTRODUCTION



1

Introduction content

Biotechnology: definition and bioproducts demand

Biotechnology has been defined in multiple ways over years, but a general one was given by The United Nations Convention on Biological Diversity, which states that biotechnology is "any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use." Since this description could be considered rather general and lack of preciseness, biotechnology has been divided into different branches and a colour has been assigned to each of them. Among them, the most important biotechnology branches are described below:

- Red biotechnology, which is related to medicine and human health. It studies discovery of new drugs, construction of artificial organs, production of vaccines and antibiotics, regenerative therapies and new diagnostics (Sasson, 2005)
- Green biotechnology, which considers the development of agriculture. It aims to progress in the discovery of new fertilizers and biopesticides as well as to develop plant varieties resistant to stresses (Laursen, 2010).
- Yellow biotechnology, which aims to improve food in terms of elimination of potential allergens and fortification with health-promoting components. The obtaining of functional food is the goal of yellow biotechnology (Siró *et al.*, 2008).
- Blue biotechnology, which exploits ocean resources to create products and applications of industrial interest. The most-known application of blue biotechnology is the use of photosynthetic microalgae for next-generation biofuels using ocean raw materials (Chemistry *et al.*, 2012).
- White biotechnology, which focuses on the recombinant production of chemical compounds considered useful for its application in different field purposes

(Barcelos *et al.*, 2018). Specifically, a lot of resources are being put on the use of renewable substrates and environmentally friendly processes, that is why the white biotechnology is considered a part of the so-called green chemistry. Lastly, the biorefinery concept has gained attention and currently is one of the most attractive trends of white biotechnology (Kamm *et al.*, 2008). Biorefineries are specific facilities which aims to transform renewable substrates —sugars, oils— into added value chemical compounds.

In this sense, research focused on recombinant production optimization of high-valuable compounds is necessary since approximately 60% of anticancer compounds and 75% of drugs for infectious diseases are either natural products or natural product derivatives, for instance (Newman *et al.*, 2003; Cragg *et al.*, 2012). These kinds of molecules are under the scope of white biotechnology.

First attempts to obtain the mentioned compounds were based on direct extraction from the natural source. However, the concentration of biologically active chemicals in the natural source is usually low, which makes the straight purification both unsustainable and unfeasible, especially for bulk chemicals (Kolewe *et al.*, 2008). Thus, active chemical demand is unreachable for natural extraction techniques. Similarly, some fungal metabolites are interesting owing their antimicrobial activity. However, the growing need of new antimicrobials would force to cultivate these fungi in large-scale. Therefore, problems regarding dense mycelium clumps must be addressed (Krivoruchko and Nielsen, 2015). In consequence, during the past couple of decades, research into natural products obtention has experienced a steady global decline (McChesney *et al.*, 2007).

Biotechnology perspectives for high added value compounds production

In order to achieve market requirements regarding biologically active chemical compounds, two different but complementary disciplines have been extensively used: chemical engineering and biotechnology.

On the one hand, chemical engineering provides a very high productivity, reaching high substrate to product conversion values. Moreover, product purification costs are quite low, and no reactor sterilization is needed, reducing process cost and enabling the possibility to perform continuous processes. However, in general, substrates are petroleum derivatives and large amounts process residues, which are tough to treat, are generated. Furthermore, chemical catalyst usually requires high temperature and pressure conditions, which leads to high energy requirements. (Chen, 2012). Finally, due to the complex structures of some of these compounds, their chemical synthesis becomes tricky (Krivoruchko and Nielsen, 2015).

On the other hand, biotechnology processes are much less efficient in terms of product concentration, conversion and rates. Importantly, contamination is considered one of the biggest concerns of this discipline, forcing to perform sterilization cycles before bioprocessing in cases which a pure culture is needed. Nevertheless, bioprocesses are normally carried out at mild conditions, in some cases can use as substrates the residues from other industrial processes and the residues are often easy to treat (Chen, 2012). Altogether, bioprocesses have an obvious lack of productivity. However, owing to our increasing concerns on the environment issues, the efforts to produce chemicals and materials from renewable biomass has exponentially increased (Lee *et al.*, 2011). Also, chemical engineering has problems for synthesis of enantiomerically pure pharmaceutical

intermediates and fuels in an environmentally friendly manner (Chen and Kazlauskas, 2011). Taking into consideration the advantages and drawbacks of biotechnology, its competitive chances are in the "high volume and low price" products, also known as high added value products. The comparison between the two disciplines are summarized in Table 1.1.

Table 1.1 Comparison of industrial biotechnology and chemical technology. Taken from (Chen, 2012)

Items	Industrial Biotechnology	Chemical technology
Reaction time	Slow: production takes days	Fast: production takes hours
Substrates	Renewable products	Petroleum or its derivatives
Conversion of substrates to products	Low: e.g. PHB/glucose ≈33wt%	High: e.g. polyethylene/ethylene≈100
Medium	Water	Mostly organic solvents
Reaction conditions	30-40˚C, atmospheric pressure	Generally >100˚C, high pressure
Product concentration	Low: several mg L^{-1} to 100 g L^{-1}	High
Product recovery cost	Very high	Low to medium
Processing	Normally discontinuous	Continuous for large scale processes
Sterilization	Often necessary	No need
Production facility cost	Very high	Low to high (explosive proof)
Wastewater	No toxic, easier to treat	Generally toxic, difficult to treat

A representative example of biotechnology use for high added value compounds production is the microbial synthesis of biopolymers such as the ones shown in Figure 1.1. These chemicals compounds can serve as building blocks for biodegradable material production and offers an opportunity to commercialize biotechnology applications in

ready-made markets that have been historically using petrochemicals (Erickson *et al.*, 2012). The most desirable way for biopolymers production is to use a microorganism that naturally produce them. Some microorganisms not only have the biosynthetic pathways for endogenous polymer production but also the production could be enhanced under certain cultivation conditions (Lee *et al.*, 2011). However, the fast growth of metabolic engineering integrated with systems biology has enabled to both produce biopolymers that are not naturally produced because cells do not have respective biosynthetic pathways and, also, for increasing native biopolymers production rates through metabolic engineering. These developments are usually based on pathway genes identification, optimal induction of enzymes in the target pathway, elimination of competing pathways, redirection of central metabolism towards the target pathway, supplementation with necessary cofactors and modulation of redox potential (Sang *et al.*, 2005; Park *et al.*, 2008). For instance, adipic acids, isoprene, and 3-carboxymuconic acid are biomonomers that cannot be synthesized by natural microorganisms due to the absence of the respective biosynthetic pathways. Therefore, genes required for the eventual biochemical production pathway must be firstly identify in the natural environment (Park *et al.*, 2008; Blazeck and Alper, 2010) —which is called bioprospecting. Once identified, codifying genes could be cloned in a microbial host in order to characterize the resulting enzymes biocatalytically. In addition, resulting enzymes can be improved with specific techniques, such as gene transfer, gene shuffling or directed evolution in order to, for example, increase the enzyme stability in exiting industrial conditions or even broaden the substrate specificity so one enzyme can work in multi-step biocatalysis of chemicals (Erickson *et al.*, 2012).

Once the heterologous genes have been identified and the potentially improved, they have to be cloned in a production host for assembling the synthetic pathway that allows to

produce a particular compound (Lee *et al.*, 2011). It is the case of glucaric acids production in *E. coli*. (Moon *et al.*, 2009).

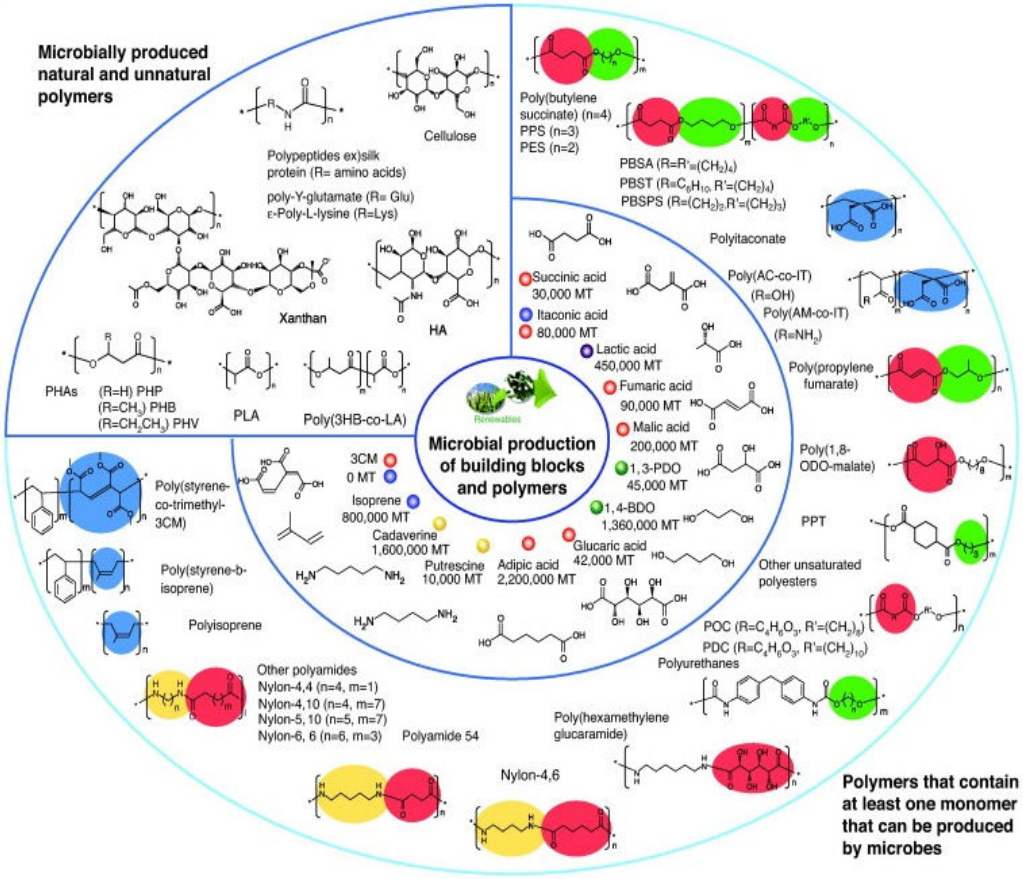


Figure 1.1. Microbially produced natural or unnatural building block chemicals used for polymer synthesis as well as polymers that can be directly produced *in vivo*. Numbers below each chemical name in the inner circle designate the amount of total annual production where MT represents metric ton. Taken from (Lee *et al.*, 2011).

The glucaric acid is just a case of a building block derivative production in the broad range of biopolymers synthesized microbially (Figure 1.1). In this sense, certain microorganisms are able to produce some polymers directly through cultivation in determined conditions —cellulose, xanthan, polysaccharides, for instance. However, in the majority of cases, only monomers could be purified from the fermentation broth and a subsequent step of polymerization must be added after the microorganism cultivation. The last one is an interesting case of potential convergence between biotechnology and

chemical engineering. A well-studied case is the polylactide (PLA) production. The monomer —lactic acid —could be obtained from anaerobic fermentation in a very efficient process (J. Wang *et al.*, 2016). Afterwards, monomers could be chemically polymerized to polylactide (Hu *et al.*, 2016). Nevertheless, one step microbial production of polymers is a preferable approach since no environmentally harmful substrates and solvents would be used and, additionally, it enables the accurate control of polymers composition, standardizing the quality of the material (Lee *et al.*, 2011).

Another expanding market is the plant-derived molecules. Up to now, 25% of medicines are obtained from nature and the most promising pharmaceuticals have been discovered in plants, leading to a value beyond $25 billion per year (Raskin *et al.*, 2002). As occurred with biopolymers, bioprospecting techniques have been used to elucidate plant biosynthetic pathways which results in a medically active compound. Through metabolic engineering and systems biology, the desired pathways could be built in microbial systems for enhancing the active compounds production (De Luca *et al.*, 2012).

Above all potential added value products, the biotechnology workhorse is the recombinant protein production (RPP), due to the absence of competitiveness from chemical engineering. Proteins are complex molecules, containing usually hundreds of monomers —aminoacids. Moreover, a strict order is needed in the monomer polymerization, otherwise the protein would not be active. Consequently, protein synthesis by chemical catalysts is highly unlikely. The global market of industrial enzymes is evolving quickly, from $2.9 billion in 2008 to $5.5 billion in 2018 and expecting to reach $7.0 billion by 2023 (Global Markets for Enzymes in Industrial Applications). Among the different protein types, technical enzymes such as lipases, proteases or amylases, were valued over $1.5 billion in 2015 whereas the value of food and beverage enzyme was estimated around $1.3 billion by 2015 (Erickson *et al.*, 2012).

Remarkably, the proteins that could bring major benefits to biotechnology are the biopharmaceuticals —therapeutic proteins, vaccines and diagnostic reagents. Both the worldwide population and age is continuously increasing, so there is a raising need therapeutic protein production at million people scale. Lately, major steps in the field of protein production have been made, both boosting the production yields and rates and reducing costs. Therefore, it is expected that the increasing demand of biopharmaceuticals will be affordable by the industrial biotechnology. However, despite it is important to reach as much protein amount as possible, it is also mandatory to keep a constant protein quality in order to standardize the process, to guarantee the final product quality and the reproducibility between batches. In this sense, industrial companies have implemented the concept Quality by Design (QbD). QbD is a bioprocess-related principle in which the final aim focused on achieving protein quality is integrated to the manufacturing process. An increased process understanding leads to and optimized and robust production process (Yu *et al.*, 2014; Kumar *et al.*, 2019). As shown in Figure 1.2, QbD includes a characterization of critical protein attributes, an identification of critical material attributes, a study of the critical production process parameters, a control of the different steps of manufacturing and the process continual improvement (Yu *et al.*, 2014).

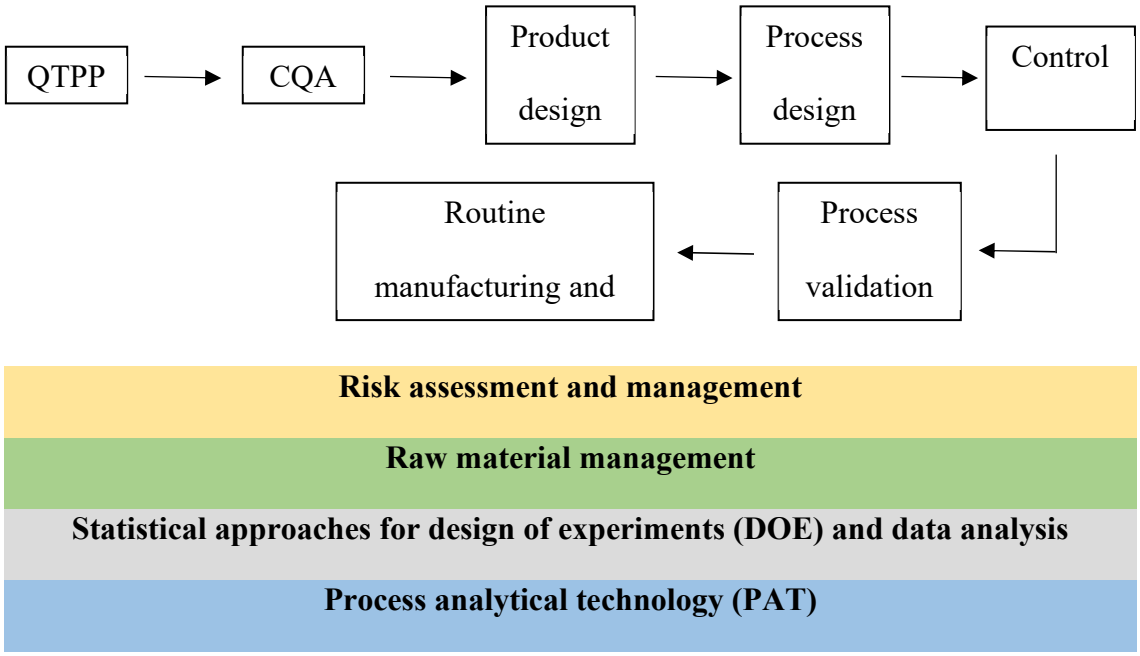


Figure 1.2. Roadmap for QbD implementation for a biotechnology product. Risk assessment and management, raw material management, statistical approaches and process analytical technology (PAT) provide foundational support to the implementation. QTPP: quality target product profile. CQA: critical quality attributes. Figure adapted from (Rathore, 2009).

The control of the different manufacturing steps is made by means of process analytical technologies (PAT). Process parameters that affect the critical protein quality attributes have to be tightly controlled in order to assure a standard product quality (Tripathi and Shrivastava, 2019). PAT include different types of process measurements such as on-line —integrated into the system, at-line —manual sampling and analysis, and in-line — directly connected to the system (Esmonde-White *et al.*, 2017; Fischer *et al.*, 2018).

Recombinant Protein Production hosts

The selection of a host for RPP depends essentially on the protein complexity and the protein amount demand (Sørensen, 2010). Briefly, if the protein does not require a complex protein processing machinery, microbial systems —bacteria, yeast— should be selected as host for RPP. On the other hand, if the protein demands specific features for delivering a biological activity, such as complex PTM —concrete glycosylation patterns, polymeric proteins— it might be necessary to use superior organisms such as mammalian cell lines, plant cell lines, insect cells, transgenic plants or even transgenic animals. Moreover, the culture medium, bioprocess strategy and physical parameters could also have an impact in the final product conformation, so in its biological activity (Tripathi and Shrivastava, 2019).

The use of transgenic animals for RPP is only related to very uncommon cases which no other host can be used. Despite being a unpopular host for RPP, protein biopharmaceuticals such as monoclonal antibodies (mAbs), vaccines, hormones and growth factors have been produced in transgenic animals in the past 20 years (Maksimenko *et al.*, 2013). Mainly, the recombinant production processes lead to a protein secretion in milk and eggs (Moura *et al.*, 2011). The main advantage of this system is that it possesses similar PTM than human organisms, so its marketplace is in the mAbs and human recombinant proteins (hRP) production (Maksimenko *et al.*, 2013). Nevertheless, the protein titers obtained are quite low and, importantly, using animals as bioreactors is not considered ethic by most of the society. Furthermore, protein extracts obtained from transgenic animals might contain zoonotic pathogens. Due to its inherent drawbacks, using transgenic animals is normally avoided.

Alternatively, plant-based systems can also be used as host for RPP. Plant production systems include microalgae, cell lines, roots, moss, and whole plants (Xu *et al.*, 2016). Several hurdles regarding protein quality and quantity make this expression host unsustainable for large-scale RPP processes (Park and Wi, 2016). In terms of protein quality, the plant glycosylation pattern differs from human one, specifically in the β-(1,2)-xylose and α-(1,3)-fucose residues added by plants but not by mammalian cells. Therefore, concerns about biosafety and human health have to be addressed (Yao *et al.*, 2015). In terms of protein amount, plant-based systems provide very low protein concentration in resulting extracts. Therefore, downstream processing and purification of products is currently tedious and costly (Yao *et al.*, 2015). In order to address the worldwide pharmaceutical demands for instance, it would require a huge raise of production capacity that is not affordable by the current plant-based RPP systems (Schillberg *et al.*, 2019). In addition, another barrier for plant-derived recombinant protein commercialization is the high cost associated with the regulatory approval of the new plant varieties created (Park and Wi, 2016). Consequently, despite promising profit margins, the commercialization of therapeutic proteins becomes time-consuming and expensive (Tripathi and Shrivastava, 2019).

Until recent years, the RPP using insect cells as hosts —mostly Sf9, S2 and High Five cell lines (Yee *et al.*, 2018)— is based on the baculovirus expression vector system (BEVS). Through this method, an infective virus is transformed with the target DNA cassette. The integrative cassette must harbour the heterologous gene which codifies the recombinant protein of interest. Moreover, this approach allows to add further DNA sequences to the baculovirus. Since the insect cells lines are not able to carry out N-glycosylation, codifying sequences for mammalian glycosyltransferases could be also included in order to co-express both the gene of interest and the enzymes needed for

performing human-like glycosylation patterns (Yee *et al.*, 2018; Owczarek *et al.*, 2019). In general, the insect culture is grown until a determined cell density and they are then infected with the baculovirus, triggering the heterologous gene expression. However, the inherent nature of the virus leads to the lysis of the insect cells, releasing a considerable amount of cellular debris, recombinant protein, endogenous proteases and baculovirus molecules. As a consequence, the downstream processes for the purification of the protein of interest turns more complicated (Jarvis *et al.*, 1990; Vicente *et al.*, 2011). In order to prevent virus-associated cell lysis, virus-less strategies have been developed —mainly in *Drosophila* and Bombyx cell lines, getting a dramatically reduced insect cell lysis and simplifying downstream processes (Yee *et al.*, 2018). In conclusion, insect cell lines can clearly compete with microbial hosts —bacteria, yeast— for production of relatively complex proteins e.g. VLPs, due to the superior PTM capabilities. On the other hand, despite several advantages of insect cells respect to mammalian cell lines —no CO_2 requirement, reduced temperature incubations, lower biosafety concerns (Ikonomou *et al.*, 2003; Kost *et al.*, 2005), its distinct glycosylation pattern makes it unable for direct production of pharmaceuticals targeted for humans, if no mammalian glycosyltransferases genes are also transformed in the producer insect cell line.

Currently, among the approved protein-based pharmaceuticals by the regulatory agencies, most of them have been produced with mammalian cell lines —84% of approved protein-based therapeutics have been produced with the Chinese Hamster Ovary (CHO) cell line. Their high eukaryotes characteristics enable them to efficiently produce complex recombinant proteins such as full mAbs, clotting factors, hormones, and cytokines (Dumont *et al.*, 2016). In addition, mammalian cell lines have notable tolerance to temperature, pH and oxygen shifts (Owczarek *et al.*, 2019), although a tight control of these physico-chemical parameters is recommended for process standardization.

However, medium composition design is considered a limiting step in bioprocess optimization because mammalian cell lines often require the addition of supplementary costly compounds such as aminoacids, growth factors or vitamins (Zhang, 2010) in order to prevent cultivation in serum-based medium. The medium optimization, which is usually required for each bioprocess, is time-consuming and expensive due to the supplementation of such a costly molecule. Furthermore, the transfection and subsequent positive clone selection is still too slow and laborious, although new transfection methods for effective heterologous DNA introduction are constantly under investigation (Gupta and Shukla, 2017). CHO cell line has been widely used for biopharmaceutical production. Nevertheless, the glycosylation pattern of this cell line does not perfectly match with the human one. Therefore, human-derived cell lines such as HEK293, HKB11, PER.C6, HeLa, and CAP cells have been also proved as host for biopharmaceuticals production, enhancing the expression of proteins with human-like glycosylation patterns (Bandaranayake and Almo, 2014; Dyson, 2016; Gupta *et al.*, 2019).

Owing to its huge metabolic wealth, filamentous fungi has been the source of a wide range of bioactive molecules and technical enzymes used currently in industry (Hoffmeister and Keller, 2007). In addition, this microorganism is also commonly used as host for RPP due to several positive features. It includes a relative fast growth under minimal medium and a low nutritional requirement. Both characteristics leads to a straightforward and cheap design of culture medium (Magaña-Ortíz *et al.*, 2018; Owczarek *et al.*, 2019). The most distinctive feature of this microorganisms is their enormous capacity for secreting endogenous proteins to the extracellular medium. It has been reported *Trichoderma reesei* bioprocesses in which concentrations up to 100 g L^{-1} of endogenous proteins were reached extracellularly (Havlik *et al.*, 2017). Even though filamentous fungi can perform eukaryotic PTM, their glycosylation pattern does not

match with the human one. Thus, the biopharmaceutical recombinant production has not be commonly carried out with this expression host. Moreover, apart from protein glycoform, the fungi highly active endogenous proteases could alter the final product quality (Owczarek *et al.*, 2019).

In contrast to the other microbial hosts, the DNA cassette transformation in fungi is still not efficient and the construction of recombinant strains is usually time-consuming. One of the important reasons is that the complex morphology of filamentous fungi differs from the different species. Therefore, the transformation protocols must be optimized for each strain (Van Den Berg, 2016; Li *et al.*, 2017). From the scalability potential of fungi-based bioprocesses, the mycelium morphological conformation adopted by fungi results in the increase of the medium viscosity. It leads to mass transfer limitations and thus, to pO_2, nutrients and pH gradients formation, potentially affecting the final overall product yield (Pollard *et al.*, 2007).

In general terms, bacterial host system is, *a priori*, the preferable host for RPP (Tripathi and Shrivastava, 2019). Concretely, *Escherichia coli* has been extensively used as host for RPP due to the huge amount of genetic and biochemical information gathered during decades, availability of several *E. coli* expression vectors and strains, fast growth, low cost and high titer and productivity rates achieved (Baeshen *et al.*, 2015; Gupta and Shukla, 2016). As a gram-negative bacterium, this RPP host have also several drawbacks such as production bacterial endotoxins, formation inclusion bodies (IB) and lack of capacity to perform eukaryotes PTM. Despite of mentioned drawbacks, around 30% of approved biopharmaceuticals are produced in *E. coli* (Tripathi and Shrivastava, 2019). In general, the approved biopharmaceuticals produced in *E. coli* does not require mammalian-like PTM —such as glycosylation, phosphorylation— for performing their biological activity. The massive research load in *E. coli* regarding bioprocess engineering

and specially, genetics and biochemical features is leading to the progressive minimization of the previously mentioned bacterial limitation. The appearance of protein IB could be avoided by co-expressing the heterologous gene along with chaperones (Gupta *et al.*, 2019), by adding fusion tags to the recombinant protein (M. Liu *et al.*, 2019) or by lowering the cultivation temperature and, therefore, the biomass growth and translational rates (Gupta and Shukla, 2016). Interestingly, protein production could be directed to the bacterial periplasmic space. This approach has been proved as beneficial in terms of protein yields, folding and solubility (Baeshen *et al.*, 2015). Recently, the exponential advance of bacterial technologies for RPP have led to the possibility to produce very complex proteins in *E. coli*. In particular, *E. coli* strains have been engineered to perform human-like glycosylation, raising the scope of RPP in this host. Hence, full-length glycosylated antibodies have been produced in this expression host (Valderrama-Rincon *et al.*, 2012; Gupta and Shukla, 2016). Even, the advanced methods for *E. coli* genetic manipulation allows to modify a bacterial strain for fitting a particular recombinant protein (Tripathi and Shrivastava, 2019).

Finally, although *E. coli* has been widely used, other bacterial host are showing up for their application in RPP such as *Bacilli* species, *Pseudomonas fluorescens*, *Corynebacterium glutamicum*, among others. In order to promote the use of these new bacterial hosts for RPP, a set of shuttle expression vectors are available for ease transfer between bacteria species (Nakata, 2017; Rosano *et al.*, 2019).

In the recent years, yeast species have emerged as a promising host for RPP. These organisms combine the positive features of microbial organisms —fast growth, low nutritional demands, easy genome manipulation— and the eukaryotes —ability to perform eukaryotic-like PTM, possibility to secrete proteins to the extracellular medium (Kurjan and Herskowitz, 1982; Çelik and Çalik, 2012; Fidan and Zhan, 2015). The yeast

research load was focused on *Saccharomyces cerevisiae* due to its annotated genome, easy genetic manipulation and biochemical characterization (Chao *et al.*, 2015; Vieira Gomes *et al.*, 2018). In terms of genetic manipulation, well-standardized tools for genome modification have been developed, raising the rate of new clone generation. Specifically, the massive amount of biological information gathered of *S. cerevisiae* has enabled to build a wide variety of genetic manipulation tools based on its particular homologous recombination (HR) efficiency (Shao *et al.*, 2009; Jakočiūnas *et al.*, 2018). Importantly, among their PTMs capabilities, yeasts are able to perform both N- and O-linked glycosylation. However, although the glycan processing is quite broad among the yeast species, all of them tends to produce hyperglycosylated proteins. This yeast characteristic is detrimental for producing biopharmaceuticals since this overglycosylation pattern reduces the therapeutic half-life in human organism and increases the allergenic response (Baghban *et al.*, 2019). Among the diverse yeast species, *S. cerevisiae* has the most hyper-glycosylation pattern. Moreover, it often produces proteins with a N-linked glycosylation terminated in α-1,3-linked mannose residues, which has been proved as allergenic (Jigami and Odani, 1999; Gellissen, 2005).

These *S. cerevisiae*-associated issues have led to broaden the yeast species scope. Thus, alternative yeast host for RPP have been investigation, including *Pichia pastoris*, *Hansenula polymorpha*, *Kluyveromyces lactis*, *Schizosaccharomyces pombe* or *Yarrowia lipolytica* (Çelik and Çalik, 2012). Above the other yeast species, *P. pastoris*, recently renamed as *Komagataella phaffii*, has been studied for years for its application in RPP processes (Macauley-Patrick *et al.*, 2005; Li *et al.*, 2007; Çelik and Çalik, 2012; Looser *et al.*, 2014). In contrast to *S. cerevisiae*, *P. pastoris* is a Crabtree-negative yeast, which means that the carbon source is not fermented to any intermediate in normoxia conditions —e.g. ethanol. Thus, a higher biomass is reached and potentially more protein titer can

be achieved (Mattanovich *et al.*, 2012). Furthermore, the endogenous glycosylation pattern of *P. pastoris* leads to a much lesser overglycosylation. In *P. pastoris*, the mannose residues addition is around 10-fold lower than in *S. cerevisiae*, reducing differences with the human pattern. In addition, *P. pastoris* strains have been glycoengineered in order to reach recombinant proteins with human-like N-glycan structures (Jacobs *et al.*, 2009).

Pichia pastoris as robust host for RPP

Methanol utilization (MUT) pathway as source of promoters for RPP

P. pastoris is a methylotrophic yeast that was firstly isolated by the company Phillips Petroleum in California in 1960 decade. The capability of *P. pastoris* to consume methanol efficiently was the key for researchers to start studying this yeast as a potential host for RPP (Ogata *et al.*, 1969). The methanol consumption pathway is shared by the methylotrophic yeasts and have been described in detail (Hartner and Glieder, 2006) (Figure 1.3). In the first step of the methanol utilization pathway (MUT), the Alcohol oxidase (Aox) catalyses the methanol oxidation to formaldehyde and hydrogen peroxide (Johnson *et al.*, 1999; Stewart *et al.*, 2001; Sudbery, 2003). On the one hand, the hydrogen peroxide is then detoxified by catalase (Cat), resulting in the release of molecular oxygen and water. On the other hand, formaldehyde could follow to different MUT branches. Briefly, in the dissimilatory branch, formaldehyde suffers two consecutive oxidations, generating formate and carbon dioxide. In both dehydrogenase reactions, NADH is generated, which is considered to be used for energy production (Lee *et al.*, 2002). In the MUT assimilatory branch, formaldehyde is condensed with xylulose 5-phosphate in a reaction catalysed by the dihydroxyacetone synthase (Das). It results in the generation of dihydroxyacetone (DHA) and glyceraldehyde 3-phosphate (G3P), which then enter in the

glycolysis pathway and generating biomass (Roggenkamp *et al.*, 1984; Sakai *et al.*, 1998; Stewart *et al.*, 2001). Regarding, the pathway cellular compartmentalization, Aox, Cat and Das enzymes are located in the peroxisomes, while the rest of reactions included in the two formaldehyde metabolism branches are located in the cytosol (Sakai *et al.*, 1997; Lee *et al.*, 2002).

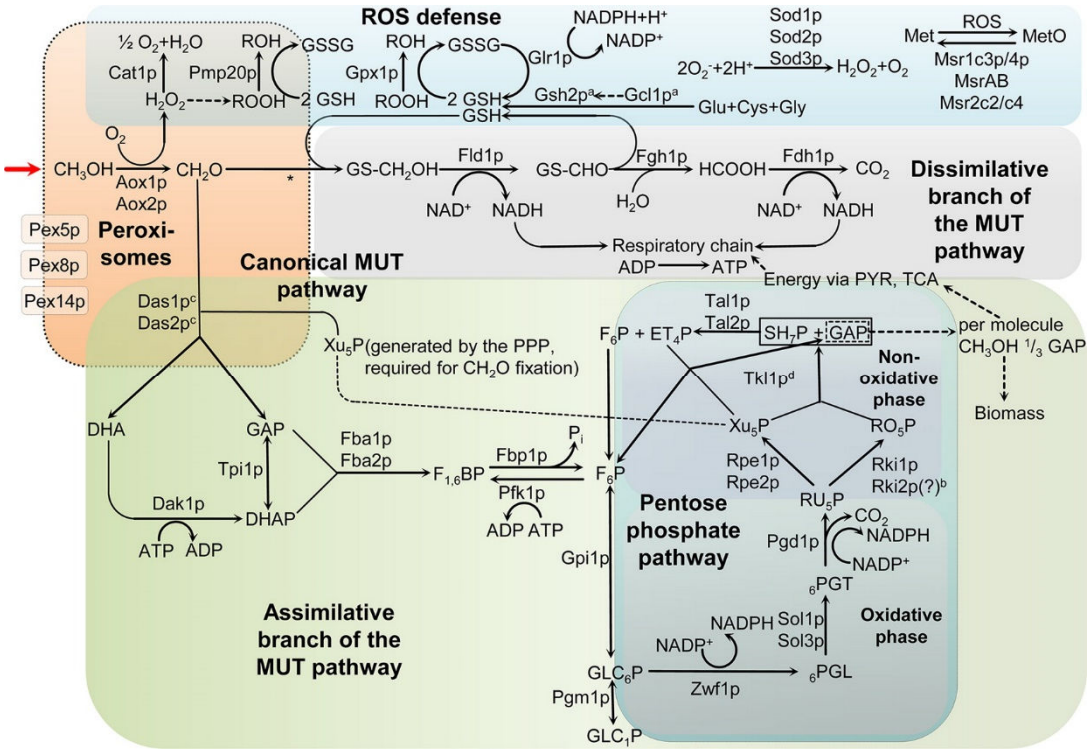


Figure 1.3. Canonical and noncanonical parts of the *P. pastoris* methanol utilization pathway. The canonical MUT pathway is based on the most recent annotation and physiological studies. Figure shows the interconnection between MUT pathway, the pentose phosphate pathway and the Reactive Oxygen Species (ROS) defense. Taken from (Vogl, Hatzl, Gerstmann, Pitzer, Wagner, G. G. Thallinger, *et al.*, 2016).

Interestingly, the tight regulation of the genes involved in the MUT pathway was demonstrated since it was observed that many of the enzymes required for the methanol metabolism were only present when culture was grown on methanol, but not other carbon sources (Veenhuis *et al.*, 1983; Hartner and Glieder, 2006). Among all the methanol

consumption pathway enzymes, the intracellular levels of Aox and Dha were significantly high in *P. pastoris* cultivations under methanol as carbon source. Even, Aox levels reached up to 30% of total soluble proteins in cells cytoplasm (Baratti and Couderc, 1980). Consequently, the peroxisomes can account for up to 80% of the cytosolic space upon methanol induction (Gleeson and Sudbery, 1988).

In the past few years, researchers have put a lot of efforts to shed light on the methanol catabolism gene regulation. Since Aox enzyme has been shown as the most predominant intracellular protein in cells growing on methanol, the regulation of the *AOX1* gene expression was investigated. The regulation of this gene could be divided in three parts: repression, derepression and induction. In each part, a set of transcription factors (TF) are involved. In presence of carbon sources such as glycerol or glucose, the MUT pathway is tightly repressed. This part of the *AOX1* regulation is less studied and would require further research, but it is known that in presence of either glucose or glycerol, the TFs Mig1, Mig2 or Nrg1 take part on *AOX1* repression by binding one or some of the TFs involved in *AOX1* induction (Shi *et al.*, 2018). When the carbon sources responsible of the *AOX1* catabolite repression are depleted —specially glucose, the promoter gets depressed. In this regulation phase, and important TF called methanol expression regulator 1 (Mxr1) has a crucial role. This TF was the first one involved in the MUT pathway and also in the peroxisome biogenesis (Lin-Cereghino *et al.*, 2006). Lin-Cereghino *et al.* first studied the Mxr1 putative cis-sites in the *AOX1* promoter (P_{AOX1}). After that, other researchers proved that Mxr1 could be bounded in P_{AOX1} in 6 different cis regions. Moreover, deletions of theses binding regions led to a significant decrease of *AOX1* in vivo promoter activity (Kranthi *et al.*, 2009). Furthermore, additional binding sites were found in another MUT genes promoter —*DAS*, and even in peroxisomal

biogenesis gene —*PEX8* (Kranthi *et al.*, 2010), confirming that Mxr1 is a master transcriptional regulator involved in expression of, at least, several MUT related genes.

In addition, a recent study reported a reciprocal regulation between Mxr1 and the Glycerol Transporter 1 (Gt1). When growing cells in glycerol medium, Gt1 gets overexpressed in order to internalize the glycerol to the cell cytoplasm. The overexpression of this gene leads to an inhibition of Mxr1 and, thus, of Aox1 as well. On the other hand, when the glycerol content in the culture medium is relatively low —e.g. in C-limiting bioprocesses, Mxr1 could bind in the *GT1* promoter (P_{GT1}) and repress its transcription. Hence, the glycerol content in cytoplasm is reduced, relieving the glycerol catabolite repression on P_{AOX1} (Zhan *et al.*, 2017). Remarkably, the Mxr1 subcellular compartmentalization has been also described. It was shown that this TF is located in the cytoplasm, when cells are growing on glucose as sole carbon source, whereas its location changes to the nucleus when the substrate is glycerol, ethanol or methanol (Lin-Cereghino *et al.*, 2006). This could explain why Mxr1can not perform its MUT deregulation function in glucose-based medium.

During the methanol induction, a cascade of transduction signal is produced. Mxr1 transmit the induction signal to Prm1. *PRM1* is constitutively expressed under glucose, glycerol and methanol. Therefore, it is hypothesized that Mxr1 provokes a conformational change in Prm1 that enables it to bind Mit1 promoter (P_{MIT1}), overexpressing *MIT1*. Mit1 is responsible to regulate P_{AOX1} (X. Wang *et al.*, 2016). Although it has been proved that Mxr1 could also bind to P_{AOX1}, its direct function in this promoter is considered much lesser than the Mit1 one.

Summarizing, P_{AOX1}-based expression system fulfils the two main features for being applied in RPP: strong and tightly regulated. For decades, it has taken advantage of this expression system since a vast number of recombinant proteins have been produced under

the regulation of P_{AOX1} (Yang and Zhang, 2018). For instance, recombinant protein concentrations with this expression system has reached up to 22 g L^{-1} of intracellular hydroxynitrile lyase from *Hevea brasiliensis* (Hasslacher *et al.*, 1997) and up to 18 g L^{-1} of an extracellular cellulase (Mellitzer *et al.*, 2014). Apart from *AOX1*, the promoters of other MUT pathway genes have been also investigated. However, the information about promoter cis-sequences and TFs that could be involved in promoter regulation are much less studied that de P_{AOX1} (Vogl and Glieder, 2013). Despite that, some interesting genes showed distinct features that makes them candidates to develop innovative expression systems.

On the one hand, Das in the enzyme responsible for detoxifying formaldehyde by combine it with xylulose-5-phosphate. It results in the synthesis of DHA and G3P. Therefore, this enzyme has a crucial role because of it is able to remove the toxic formaldehyde and, moreover, redirects it to the assimilative branch of MUT pathway, resulting in the biomass generation. Consequently, due to its important function, its regulation and strength has been also studied. Das enzyme is encoded by two genes, *DAS1* and *DAS2*, which differs in just 9% of its coding sequence. In terms of promoter strength, both *DAS1* and *DAS2* promoters ($P_{DAS1;}$ P_{DAS2}) showed similar induction levels. Even, this study indicated that P_{DAS2} even outperform the P_{AOX1} induction strength (Tschopp *et al.*, 1987). Concerning its regulation, these promoters share the same regulation with P_{AOX1}. In fact, several Mxr1 binding sites have been found in the promoters of genes encoding both Das isoforms (Kranthi *et al.*, 2010). Nevertheless, this powerful expression system requires the use of higher concentration of methanol in order to ensure as higher protein yields as possible. This inducer is a hazardous compound due its flammability and toxicity, representing a relevant increased cost for large-scale bioprocesses (Shi *et al.*, 2018). In addition, methanol consumption leads to a high oxygen demand and high heat

production (Heyland *et al.*, 2010; Çalık *et al.*, 2015). Both factors further also increase the bioprocess costs since the heat excess must be removed and for some cases, pure oxygen needs to be supplied. Owing to its similar characteristics to P_{AOX1}, P_{DAS1} and P_{DAS2} lack of interest since they do not present alternative features that allows to overcome the P_{AOX1} performance.

Towards the reduction of methanol utilization in *P. pastoris* bioprocesses

MUT alternative phenotypes

In order to mitigate the methanol-related drawbacks of this expression system, some solutions have been proposed. Firstly, alternative *P. pastoris* phenotypes which present different methanol catabolism behaviour have been constructed. In *P. pastoris*, Aox enzyme is encoded by two different genes, *AOX1* and *AOX2*. It was described that the coding sequences match around at 90% and the regulation of both genes was similar. However, the *AOX2* expression is controlled by a much weaker promoter than *AOX1*. Consequently, on average, only the 15% of total Aox protein belongs to the *AOX2* gene expression (Cregg *et al.*, 1989). The *P. pastoris* wild type (WT) strains harbour both genes encoding Aox enzyme. Regarding its methanol consumption properties, it is considered to have a Methanol utilization plus (Mut^+) phenotype. Other *P. pastoris* strains have been constructed, in which *AOX1* has been knocked-out. Consequently, these variants present a slower methanol catabolism, which has been classified as a Mut^S phenotype. Lastly, alternative *P. pastoris* strains in which both genes encoding Aox are deleted have been also built. These last strains are not able to metabolize the methanol and have been classified as Mut^- phenotype.

Briefly, Mut^+ strains are able to consume methanol at rates significantly higher than Mut^S strains. Since, the P_{AOX1} induction is methanol-dependent, expression is growth-coupled. Therefore, some works have reported better RPP productivities when working with Mut+ strains (Chiruvolu *et al.*, 1997; Kim *et al.*, 2009). However, as mentioned previously, more methanol consumption implies higher oxygen uptake and heat production, and it could be derived into important operational troubles. On the other hand, some researchers have also reported Mut^S relevant benefits regarding specific recombinant protein productivity. Krainer *et al.* observed that methanol uptake rate of a Mut^+ strain doubled the one observed in the Mut^S strain when both strains produced the horseradish peroxidase (HRP) regulated by P_{AOX1} (Krainer *et al.*, 2012). Despite the methanol consumption differences, Mut^S strain showed a 3-fold higher HRP specific and volumetric productivity in comparison with the Mut^+ strain. Moreover, they also calculated the product-to-substrate yield ($Y_{P/S}$) which provide information about the strain efficiency to convert the medium substrate into the desired product. Remarkably, the Mut^S strain delivered a 7-fold higher $Y_{P/S}$ than Mut^+ strain (Krainer *et al.*, 2012). Similar results were obtained by Orman. *et al.* when studying the production of a recombinant human growth factor (rhGH) with the two different phenotype strains. Among a set of methanol concentrations tested, the best rhGH titer was obtained with the Mut^S strain and adjusting the initial methanol concentration at 2%. The Mut^S rhGH production overcame the Mut^+ one by 3-fold (Orman *et al.*, 2009).

MUT deregulation

Taking advantage of the increasing information regarding MUT pathway gene regulation, other alternative expression systems based on MUT deregulation has been investigated by some researchers in order to avoid the use of methanol as RPP inducer while still exploiting its excellent regulation features. Shen. *et al.* put the focus on the *P. pastoris*

kinases which could potentially induce/repress the P_{AOX1} (Shen *et al.*, 2016). After a high-throughput knock-out experiments, they identified two kinase genes —glycerol kinase and dihydroxyacetone kinase (Dak)— whose deletion led to an activation of P_{AOX1} under glycerol as sole carbon source. Therefore, they integrated P_{AOX1}-driven expression cassettes into these kinase-deficient clones, aiming to trigger the gene expression in methanol absence. Particularly, three recombinant proteins were tested —amylase (Amy) from *Geobacillus sp*, glucose oxidase (God) from *Aspergillus niger* and hepatitis B small surface antigen (HBsAg) from human. Among them, the best results were obtained with the Dak deletion clone, in which around 50% of P_{AOX1} induction —using DHA as inducer— was obtained when comparing with the standard induction with methanol (Shen *et al.*, 2016).

Similarly, Wang *et al.* decided to knock-out the known P_{AOX1} repressors, instead of finding new ones by means of high-throughput methods (Wang *et al.*, 2017). The target repressors to delete were Mig1, Mig2 and Nrg1. Thus, they constructed *P. pastoris* mutants, in which there were single gene mutations and combinatorial ones. By cultivating mutants with glycerol as sole carbon source, they observed that the highest Aox activities were found in double —Δ*MIG1*Δ*MIG*— and triple mutants —Δ*MIG1*Δ*MIG*Δ*NRG1*—. Importantly, *AOX* expression was tightly repressed in all mutants in presence of glucose, confirming the repression power of this compound, beyond the glycerol one. On the other hand, they also checked the effect of deleting the mentioned repressors on the induction of the recently described Mit1. The knockout of Mig2 provided no changes in Mit1 expression. However, both the double and triple mutant led to a drastic increase of Mit1 induction on glycerol medium, indicating that one or both repressors have a direct impact on *MIT1* transcriptional regulation and thus, on Mit1 expression. Combining the triple mutant and a heterologous constitutive overexpression of Mit1 gave the best P_{AOX1} induction results,

since the GFP production under this promoter reached the 77% of induction rate comparing with the WT one, induced by methanol. Remarkably, RPP experiments in bench top bioreactors provided some practical information regarding the new mutant clone, as the oxygen consumption and heat production were reduced by 56% and 52%, respectively in comparison with a standard bioprocess induced with methanol. Therefore, despite a 23% less protein was produced, the energy and oxygen supply saved makes the deregulated clone a good alternative to the classical one (Wang *et al.*, 2017). Accordingly, Vogl *et al.* reported that overexpressing either *MIT1* or *MXR1* was enough to reach the 44% and 25% of the WT P_{AOX1} induction by methanol, respectively (Vogl, Sturmberger, *et al.*, 2018). These values of P_{AOX1} methanol-free induction are not that high as the ones describes in the work on Wang and colleagues, confirming that MUT pathway presents a complex regulation, where both repressor and inducers TFs are involved. Accordingly, combining repressor deletions and inducers overexpression provides better improvements.

Methanol-free promoters

Formaldehyde could also be detoxified via the dissimilatory branch of the MUT pathway. The first step of this branch consists in the spontaneous formaldehyde condensation with glutathione, generating S-hydroxymethylglutathione. This compound is subsequently oxidized to S-formylglutathione, in a reaction catalysed by the formaldehyde dehydrogenase (Fld). As important step in the MUT pathway, the *FLD1* promoter (P_{FLD1}) is induced by methanol. However, the distinctive feature of this promoter is that it could be also induced by methylamine as sole nitrogen source (Shen *et al.*, 1998). Therefore, a methanol-free bioprocess strategy could be designed, but avoiding C-sources such as glucose or glycerol, since it was reported that an excess of these widely used substrates inhibit the P_{FLD1}. Alternatively, other C-sources such as sorbitol have been proposed

(Resina *et al.*, 2004). In this sense, a novel bioprocess strategy was designed using the P_{FLD1}-based expression system for the *Rhizopus oryzae* lipase (Rol) production. In this case, sorbitol and methylamine were used as carbon and nitrogen sources, respectively. P_{FLD1}-based expression system showed up as an attractive alternative to the classical P_{AOX1}, as the specific and volumetric Rol productivities reached with the P_{FLD1} outperformed the results obtained with P_{AOX1} (Resina *et al.*, 2005). Likewise, Shen *et al.* reported that the β-galactosidase production levels obtained with the P_{FLD1}-based expression systems —induced by methylamine— were rather similar compared to ones obtained with the P_{AOX1}—induced by methanol (Shen *et al.*, 1998).

In order to further investigate the promoters that drive the expression of the MUT and detoxification pathway genes, a broad characterization of MUT-related promoters was carried out (Vogl, *et al.*, 2016). This study provided a detailed characterization of up to different 45 promoters in terms of regulation and expression strength. Among them, the promoter that triggers the expression of *CAT1* (P_{CAT1}) delivered a remarkable performance. *CAT1* is the gene that encodes for the catalase, which detoxifies the hydrogen peroxide produced in the methanol oxidation, among other metabolic processes. As expected, it showed a significant induction under methanol and a tight repression under glucose excess. Even, under methanol induction, P_{CAT1} expression system provided similar yields than P_{AOX1} to produce the horseradish peroxidase (HRP) and outperformed P_{AOX1} for CalB production. In addition, this promoter is deregulated when the glucose reaches values near zero in the culture medium. So, the P_{CAT1} starts to show some induction when C-source is depleted, and it could be further induced with methanol. It leads to a high degree of freedom in its regulation and tunability, which is considered also a very valuable feature (Rajamanickam *et al.*, 2017). In line with the depressible promoters, recently, Vogl *et al*. reported the outstanding performance of the formate

deshydrogenase promoter from *Hansenula polymorpha* (P_{HpFMD}) for RPP bioprocesses in *P. pastoris* (Vogl *et al.*, 2020). In this work, it was shown that the derepressed expression of HRP and CalB with the P_{HpFMD} outperformed their expression with the P_{CAT} in the same culture conditions. Moreover, it was noticed that the P_{HpFMD}-driven expression could be further boosted by means of adding methanol as carbon source. Interestingly, in presence of methanol, the P_{HpFMD} outperformed the P_{AOX1} up to 2.5-fold. The performance of this promoter for driving *P. pastoris* RPP bioprocess is studied in the Chapter II of Results.

Another source of promising promoters that may be an alternative to P_{AOX1} and, so, to prevent the use of methanol is the glycolysis pathway genes (Figure 1.4). Among the glycolysis-derived promoters, the most extensively used is the glyceraldehyde-3-phosphate dehydrogenase promoter (P_{GAP}) (Waterham *et al.*, 1997). Concerning its regulation, this promoter is constitutively expressed, even though the expression strength depends on the carbon source used. Both glucose and glycerol stimulate the P_{GAP}-driven expression, but the highest P_{GAP}-driven expression levels are reached with glucose as sole C-source (Waterham *et al.*, 1997; Çalık *et al.*, 2015). Although it has been widely described that the expression levels of P_{GAP} could nearly achieve the ones reached by P_{AOX1}, it highly depends on the recombinant protein and the bioprocess strategy followed (Cos *et al.*, 2006). Its constitutive expression provokes a major problem for producing toxic proteins (Li *et al.*, 2007; Potvin *et al.*, 2012; Çalık *et al.*, 2015). Therefore, if the target recombinant protein interferes with any metabolic essential process on *P. pastoris*, no colonies will be obtained in the early phases of strain construction —clone selection. Owing its remarkable advantages to be applied on large-scale RPP processes, exhaustive bioprocess optimization has been done. In this sense, recently, García-Ortega and co-workers reviewed the last advances performed in this field using P_{GAP} has the heterologous gene trigger (García-Ortega *et al.*, 2019). The worst drawback of P_{GAP} with

respect to P_{AOX1} is the lack of information about the promoter cis-elements and TFs that could potentially alter the promoter expression levels. In this regard, an exhaustive study was carried out recently (Ata *et al.*, 2017), in which some putative TFs were identified. In this work, deletions/duplications of TF binding sites (TFBS) and TFs overexpression were performed. By duplicating the TFBS of a *GAL4*-like TF, an outstanding increase of P_{GAP} -driven rhGH up to 2.4-fold was found. In addition, when combining the cis-site duplication with the *GAL4*-like overexpression, a further 1.8-fold production improvement was achieved. Furthermore, the *GAL4*-like TF knock out led to a decrease in both growth and protein production in a range of 35-85% comparing with the wild-type P_{GAP} -driven strain. Altogether, these results indicate that the *GAL4*-like TF has a crucial role of P_{GAP} activation, at least (Ata *et al.*, 2017). Additionally, it suggests that there is still a big room for P_{GAP} optimization through promoter engineering. Despite of that, the use of the P_{GAP} as constitutive promoter for RPP in *P. pastoris* is at risk since the study provided by Liang *et al*. (Liang *et al.*, 2013). In their work, the performance of the widely used P_{GAP} was compared with the also constitutive glycosyl phosphatidyl inositol (GPI)-anchored protein promoter (P_{GCW14}). Clones expressing an eGFP were cultivated in shake-flasks with either glucose or glycerol as carbon sources. Surprisingly, the eGFP production provided by the P_{GCW14} was about 10-fold higher than P_{GAP} on glycerol and 5-fold higher on glucose. On the contrary, Wang *et al*. reported only a 19% of improvement in the production of a thermo-alkaline lipase from *Thermomyces dupontii* (Wang *et al.*, 2019). Due to this disparity of results regarding the P_{GCW14}, a thorough characterization of P_{GCW14} as RPP promoter is carried out in the Chapter II of Results.

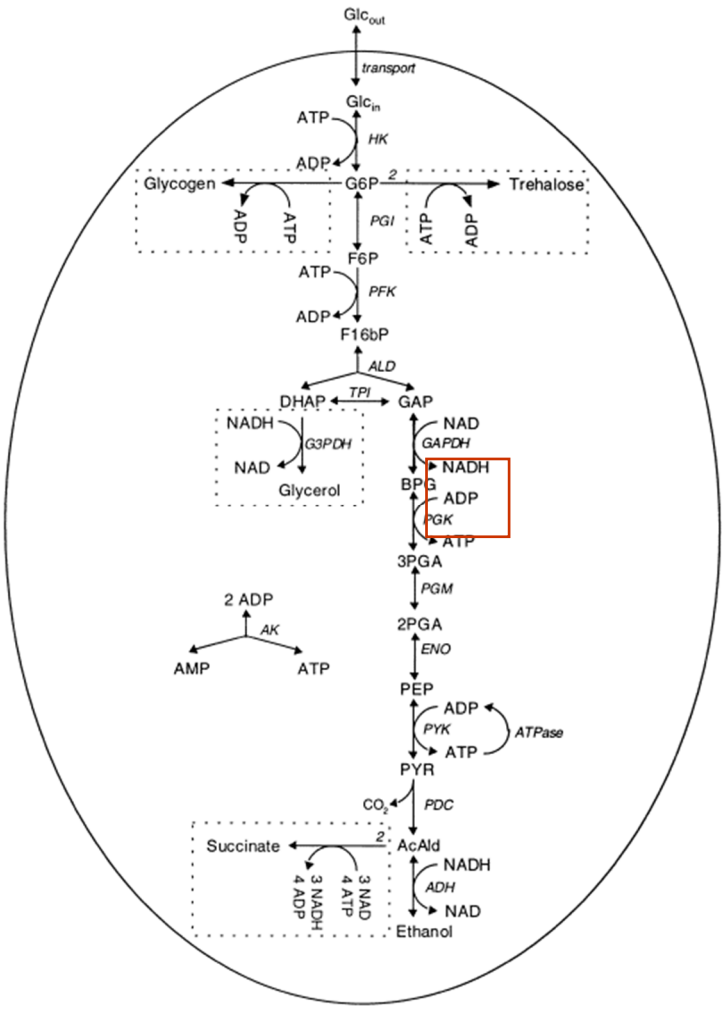


Figure 1.4. Glycolysis pathway. Adapted from (Teusink *et al.*, 2000). A red edge square points out the reaction catalysed by the enzyme expressed under the P_{GAP} control (GAPDH).

In the glycolysis pathway (Figure 1.4), the G3P dehydrogenase catalyzed reaction is followed by a dephosphorylation of the 1.3-biphosphoglycerate to 3-phosphoglycerate. This last reaction is catalyzed by the phosphoglycerate kinase (Pgk). The constitutive promoter of the encoding gene (P_{PGK1}) delivered also a good performance in heterologous gene expression, as showed in a work made by Robert and coworkers (Robert *et al.*, 2017). They reported that CalB production under this promoter showed similar values that those from P_{GAP}-driven expression, although significantly lower than those obtained with P_{AOX1} —around 90% lower. Since bioprocesses driven by P_{PGK1} are still under-investigated, they stated that the absence of bioprocess optimization might be one

important cause of such a big difference with the P_{AOX1}-driven CalB production. Furthermore, interestingly, crude glycerol extracted from biodiesel production processes was used as C-source for *P. pastoris* growth.

In the lower glycolysis, concretely in the pyruvate node, two additional interesting enzymes can be found. They are the pyruvate kinase (Pyk) and pyruvate decarboxylase (Pdc). The promoters that drive the expression of the respective genes (P_{PYK}; *PYK* and P_{PDC}; *PDC*) have also been studied as alternatives to P_{GAP} for methanol-free RPP bioprocesses (Massahi and Çalık, 2018). To assess the performance of these two pyruvate node-related promoters, the production of rhGH was carried out in both expression systems and also, in a P_{GAP} -based, as control. The rhGH productivities under the novel promoters reached similar values as the widely used P_{GAP}. Even, P_{PYK}-based expression system showed slightly higher productivity value than P_{GAP}. Remarkably, P_{GAP}-based expression system suffered from higher protease activity from the half until the end of the bioprocess. This detrimental effect was prevented by using the novel expression systems, leading to a significantly higher rhGH productivities in this bioprocess range of time. A summary of the aforementioned promoters is depicted in Table 1.2.

Table 1.2. Summary of the most used strong promoters for RPP in *P. pastoris*

Promoter name	Source	Regulation
P_{AOX1} P_{DAS1} P_{DAS2}	MUT pathway	Methanol inducible; tightly repressed by glucose/glycerol
P_{FLD1}	MUT pathway	Methanol/methylamine inducible
P_{CAT1}	MUT pathway	Derepressible; further inducible by methanol
P_{GAP} P_{PGK}	Upper glycolysis	Constitutive
P_{PYK} P_{PDC}	Lower glycolysis	
P_{GCW14}	Orthologous	Constitutive
P_{HpFMD}		Derepressible; further inducible by methanol

Adjusting the balance between heterologous gene expression and protein processing capabilities

Transcription rate raise

Theoretically, the ideal promoter properties should combine a tight regulation and a strong induction/derepression levels. Consequently, there is a continuous search of new regulated promoters that allows a strong protein expression induction, especially for being applied in promising host for RPP such as *P. pastoris*. However, instead of searching new expression systems, to modify the existing ones is another option for boosting the heterologous gene transcription rate. In this regard, the most popular procedure is raising the heterologous gene dosage. Several attempts to increase recombinant protein yields by raising heterologous gene dosage has been made in *P. pastoris* and other methylotrophic

yeasts (Kobayashi *et al.*, 2000; Kang *et al.*, 2001; Marx *et al.*, 2009; Zhu *et al.*, 2009; Sha, X. W. Yu, *et al.*, 2013; H. Yang *et al.*, 2016; Shu *et al.*, 2016). Zhu and colleagues created a set of strains whose differences was the number of integrations of the heterologous cassette, composed of the precursor of porcine insulin gene (PIP), the P_{AOX1} and de *AOX1* gene terminator (Zhu *et al.*, 2011). By means of a sequential integration of double digested plasmid into single digested original plasmids and a subsequent clone selection, it was created a set of Pip producing clones, containing from 1 to 52 cassette copies. They observed that the relationship between Pip production and gene copy number presented a bell-shape trend, where the maximum in Pip production was found at 12 copies of the heterologous cassette. Interestingly, for the same culture conditions, both the final biomass concentration and the μ started to decrease linearly when the gene dosage reached values exceeded 6. In the same way, Shu *et al.* followed a similar procedure than the previous work for studying the heterologous gene dosage impact on the production of a serine protease (Sptk) under the control of P_{AOX1} (Shu *et al.*, 2016). Nevertheless, in this study, interesting analyses such as the heterologous gene transcript levels were also performed. Linking the transcription rate with the secreted protein concentration helps to understand the effect of raising the heterologous gene dosage in the yeast. In this case, their study comprises just 4 clones, harbouring from 1 to 4 heterologous cassette copies. Among the results described by the authors, they observed that the *SPTK* relative mRNA levels followed a linear trend along the gene dosage whereas the secreted protein concentration reached its maximum when 2 cassette copies are integrated in the *P. pastoris* genome. Beyond 2 cassette copies, the Sptk levels in the extracellular medium are significantly lower than the maximum, suggesting the presence of a bottleneck through the protein processing processes —mRNA translation, folding and secretion. As mentioned, one important advantage of yeasts as host for RPP, is that its protein

processing machinery is able to manage the satisfactory production of a wide range of protein, from simple to very complex ones (Damasceno *et al.*, 2004; Hackel *et al.*, 2006; Braren *et al.*, 2007; Shi *et al.*, 2007; Shukla *et al.*, 2007). Even, the yeast protein processing pathways allow the protein maturation, including disulphide bond formation and glycosylation among others PTMs.

Unfolded protein response (UPR)

The first step in the protein processing pathway is the protein translocation and folding. This process takes place in the endoplasmic reticulum (ER), which is a subcellular compartment that provides the optimal environment for a proper protein folding — including disulphide bond formation. In order to contribute to the protein folding, several intra-ER chaperones assists the process. Afterwards, an ER quality control determines if proteins are properly folded. In particular, N-linked glycosylation in the ER plays an important role in protein quality. In this case, proteins are guided to subsequent trafficking organelles, called Golgi apparatus, where a protein maturation process takes place (Helenius *et al.*, 1992). An excessive gene expression (e.g. heterologous gene expression) might lead to an unfolded protein overload in the ER lumen, causing a heavy ER-related stress. In order to mitigate the stress, eukaryote cells activate the unfolded protein response (UPR, Figure 1.5). The main UPR sensor in *P. pastoris* is Kar2. By binding them, this ER-resident chaperone prevents the action of three UPR activators, the PKR-like ER protein kinase/translation initiation factor 2 (eIf2α) kinase (Perk/eIf2α); the activating transcription factor 6 (Atf6), and the inositol-requiring enzyme 1 (Ire1). However, when a high amount of misfolded proteins is presented in the ER-lumen, Kar2 binds misfolded proteins in order to avoid aggregation between them. This provokes the release of the 3 UPR activators and the transduction signal is transmitted. As a result, an intron in removed from the *HAC1* mRNA. The translation of this *HAC1* mRNA results

in a powerful UPR deployment (Guerfal *et al.*, 2010; Damasceno *et al.*, 2012; Puxbaum *et al.*, 2015) (Figure 1.5), although its concrete target genes remain still unclear. Among the UPR effects, the mRNA translation rate is decreased, the ER organelle proliferates, ER-associated degradation (ERAD) is activated to degrade the excess of misfolded proteins and, importantly, the expression of chaperone genes such as *KAR2*, *PDI*, *ERO1* among others are induced (Bernales *et al.*, 2006; Wu and Kaufman, 2006).

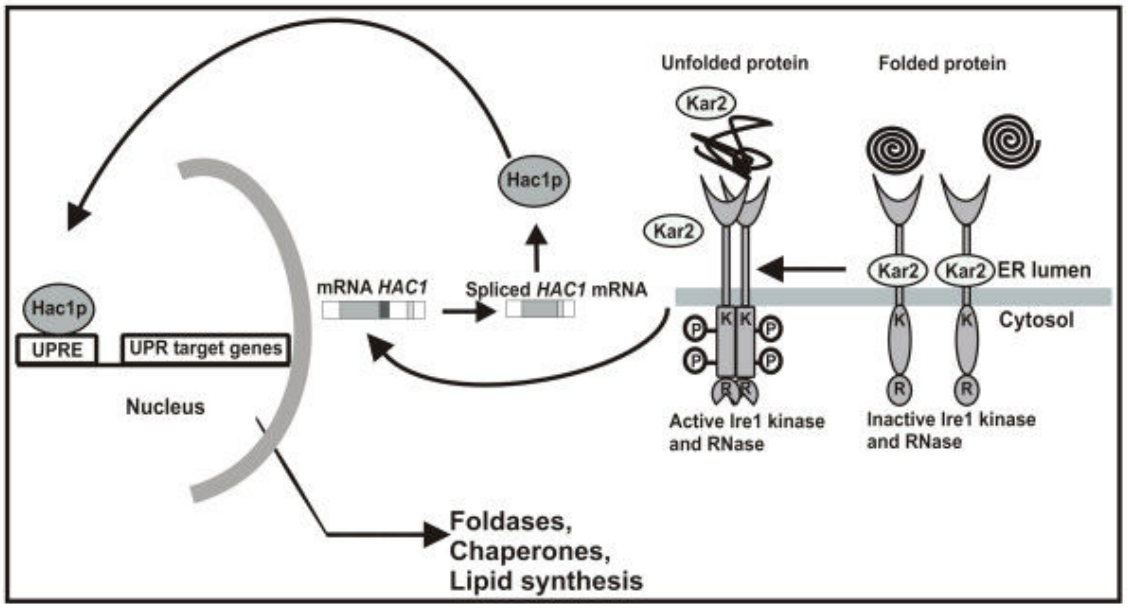


Figure 1.5. The unfolded protein response in the yeast *S. cerevisiae*. Taken from (Guerfal *et al.*, 2010).

The UPR induction in *P. pastoris* has been proved by means of diverse omics analyses. Hesketh *et al.* studied the *P. pastoris* transcriptional response to the production of two differently misfolded variants of a human lysozyme, in chemostat cultivations (Hesketh *et al.*, 2013). Strikingly, it was observed a constitutive *HAC1* mRNA splicing, regardless the producer clone or in the heterologous gene induction conditions. Accordingly, *KAR2* and *PDI* genes were not significantly expressed between pre- and post-induction steady states cultivations. Taking that into consideration, they hypothesized that the C-limiting conditions strongly stimulates UPR and the impact of both lysozyme variants and

induction conditions is negligible. Conclusions obtained by Vanz *et. al* followed the same direction (Vanz *et al.*, 2014). By following the UPR and ERAD-related proteins profiles, they aimed to notice their differential production between the non-induced phase —batch, glycerol as C-source— and the recombinant protein induction phase —methanol feeding phase— in fed-batch cultivations. A not expected decrease of UPR and ERAD-related proteins profiles was noted in the induction phase in comparison with the batch phase. They stated that the excess of glycerol used in the batch phase, is responsible for the high UPR and ERAD high basal induction. Consequently, batch phase preconditioned cells for managing the high misfolded protein burden produced in the induction phase.

In contrast, Yu and co-workers performed transcriptomic analyses of two clone producing the phospholipase A (Pla2) from *Streptomyces violaceoruber*. The sole difference between both clones were the *PLA2* gene dosage —1 cassette copy and 12 cassette copies. Transcriptomic analyses revealed a set of 150 genes that were differentially expressed in 12-copies clone comparing with single copy clone. Those genes were mainly involved in DNA replication, protein refolding, stress response and ATPase activity regulation (Yu *et al.*, 2017). In the same direction, the *PIP* gene dosage raise led to an expected *PIP* mRNA levels increase (Zhu *et al.*, 2011). Consequently, it was observed an upregulation of both *KAR2* —up to 6.5-fold— and *PDI1* —up to 3-fold— with respect to the negative control clone which no heterologous gene was integrated. As explain, the ER-resident chaperones overexpression are cellular attempts to alleviate the misfolded protein-derived stress.

The contradictory results obtained by different researchers suggest there is not a common UPR/ERAD induction pattern. In fact, it seems that the operational mode and culture conditions could also induce UPR by themselves. Additionally, the strength of the RPP-derived UPR has showed to be highly protein-dependent.

Engineering of protein processing pathways

Summarizing, boosting the heterologous gene transcription rate might not have the expected impact in protein production due to the limited *P. pastoris* protein processing machinery. Even, it has been described that if the protein processing is the limiting step in a RPP process, a strong promoter might be counter-productive (Hohenblum *et al.*, 2004; Mattanovich *et al.*, 2004; Ruth *et al.*, 2010). In order to prevent the folding/secretion limitations, several works have been focused in increasing the *P. pastoris* protein processing capacity through distinct approaches. Since some works reported a positive effect of overexpressing helper genes for RPP, numerous works have co-express these helper proteins along with the target recombinant protein. Nevertheless, this approach has demonstrated to not be fully effective for boosting the extracellular recombinant protein production. Zhang *et al.* reported positive effects of overexpressing *S. cerevisiae* helper genes for the production of a hormone-like protein (G-CSF) in *P. pastoris*. Concretely, the independent overexpression of Kar2, Sec63, Pdi, Ydj and Ssa1 led to a G-CSF production increases up to 5.6, 2.9, 4.0, 3.6 and 6.8, respectively (Zhang *et al.*, 2006). Inan and co-workers studied the potential production improvements of several combinations of heterologous and helper gene dosage (Inan *et al.*, 2006). The recombinant protein was the *Necator americanus* secretory protein (Na-ASP1) and the helper protein was Pdi, an ER-resident chaperone. They found their maximum relative protein production improvements with a producer clone harbouring 4 copies of both heterologous and helper protein, demonstrating that an optimal balance between both genes is needed for achieving the maximum product amount. Interestingly, Pdi availability seems to be limiting in this case, since significant improvements are observed from 3 *PDI* copies. In contrast, Azoun *et al.* reported that overexpression of single *PDI* copy under P_{GAP} was enough to dramatically increase rabies virus glycoprotein (RABV-

G) amount up to 8-fold in a clone bearing 7 RABV-G copies. Further *PDI* copies did not provide significant RABV-G titer increase (Ben Azoun *et al.*, 2016). Balancing the helper and the heterologous gene expression was also considered crucial for Shen and co-workers. In their work, the best result in terms of HSA fusion protein (IH) production was found with a clone harbouring 3 copies of each gene (Shen *et al.*, 2012). The overexpression of the master UPR regulator *HAC1* has also provided recombinant protein production raises. Guerfal *et al.* proved that the expression regulation of the helper protein is a crucial factor, since the constitutive expression of *HAC1* did not provide major improvements in mouse interleukin-10 (mIL-10) secretion. However, inducible *HAC1* overexpression led to an increased mIL-10 and trans-sialidase (TS) production by 2.2-fold and 2.1-fold, respectively (Guerfal *et al.*, 2010). Likewise, the inducible *HAC1* overexpression under P_{AOX1} led to a 2.2-fold increase of HRP volumetric activity (Krainer *et al.*, 2016).

Nevertheless, despite some works reported beneficial effects obtained by overexpressing helper proteins such as ER-resident chaperones or UPR regulators, some other reports noted that this strategy had no significant effect and even, it could be detrimental for RPP processes. In the case of Shen and colleagues, they pointed out a positive effect of overexpressing *PDI*, but it was noticed a decrease IH protein levels when *KAR2* was overexpressed (Shen *et al.*, 2012). In contrast, the coexpression of *KAR2* along with an antibody fragment (scFv) resulted in a 3-fold increase with respect to the control clone. However, the *PDI* overexpression did not provide any further improvement with respect to the control clone. All helper and heterologous genes overexpression were driven by the P_{AOX1} (Damasceno *et al.*, 2007). In agreement, *PDI* overexpression did not increase the amount of total human A2a adenosine receptor in *S. cerevisiae* with respect to the parental producer clone (Butz *et al.*, 2003).

From the literature, no final conclusions can be stated regarding the strategy of overexpressing helper proteins along with the heterologous gene. Some researchers have speculated about the detrimental effects of a high presence of chaperones in the ER. Ideally, proteins should be transient the ER until achieving their correct folded conformation. Whether a high chaperone levels are presented in the ER lumen, the interaction of these helpers with the nascent proteins could be excessive (Hendershot *et al.*, 1987; Sagt *et al.*, 1998), Thus, triggering ERAD (Knittler *et al.*, 1995), instead of promoting a proper protein secretion.

Altogether, the chaperones specific activities are still under investigated. Therefore, published works aiming to increase protein production through overexpressing ER-resident chaperones are still based in co-expression of heterologous and helper genes dosage. Further investigation in this field is required in order to shed light on the interactions between recombinant protein expression, helper protein contribution and effect of the cultivation conditions.

Apart from the protein folding in the ER, other steps in protein processing pathway have also been improved. As mentioned previously, the most used signal peptide for secreting recombinant proteins to the extracellular medium in *P. pastoris*, is the *S. cerevisiae* α-mating factor pheromone (Brake *et al.*, 1984). This peptide consists of two functional regions: the pre-region, that directs the nascent protein to the ER; and a pro-region, that is involved in the protein translocation from ER to Golgi apparatus. Thus, in order to by-pass a potential bottleneck related with the secretion mechanisms, some works have attempted to modify the signal peptide to increase the recombinant secretion rate. By analysing the secondary structure of the peptide, rationale mutagenesis was performed (Lin-Cereghino *et al.*, 2013). Remarkably, by performing an entire α-helix removal of the pro-region, could be achieved a HRP secretion increase of 50% and a CalB secretion

increase up to 100%. Taking into consideration that the signal peptide have been obtained directly from *S. cerevisae*, Ahn *et al.* adjusted the peptide codon usage of *P. pastoris* following two design parameters (Ahn *et al.*, 2016). In this way, they proved that it is a successful approach, as the CalB secretion was improved up to 3-fold.

Interestingly, in a work of our research group, Barrero and co-workers hypothesized that the translocation from cytosol to the ER-lumen might be a major bottleneck in protein processing pathway, especially for proteins that could be (partially) folded prior their ER entry (Barrero *et al.*, 2018). Therefore, they proposed to substitute the α-mating factor pre-region to the *S. cerevisiae* Ost1 signal sequence in order to promote the cotranslational translocation instead of the posttranslational translocation, and thus preventing that proteins could be misfolded in the cytosol. This pre-region modification together with two-point mutations in pro-region provided huge improvements in the expression of the fluorescent E2-Crimson and *Bacillus thermocatenulatus* lipase 2 (Btl2)—up to 20 and 10-fold, respectively. Marsalek *et al.* put the focus downstream in the secretory pathway. In concrete, they observed a mis-sorting of proteins from Golgi apparatus to the vacuolar pathways, instead of secretory endosomes. The non-desirable protein sorting to vacuoles was prevented by disrupt them along with some vacuolar proteases. These two modifications led to secretion improvements of a Fab and a carboxylesterase up to 80% (Marsalek *et al.*, 2019).

Genetic engineering tools advances in *Pichia pastoris*

The rate for obtaining *P. pastoris* recombinant clones has been exponentially increased during the past five years due to the massive steps made on developing innovative genetic engineering tools (Fischer and Glieder, 2019). The most important could be considered the adaptation of the widely used tool CRISPR/Cas9 for genome editing (Weninger *et al.*,

2016). However, firsts attempt to apply this promising genetic engineering tool was not as efficient as desired due to the *P. pastoris* feature of using non-homologous end-joining (NHEJ) instead of homologous recombination (HR) for DNA repair, in contrast to other yeasts such as *S. cerevisiae*. In order to increase the tool efficiency, a new *P. pastoris* strain with deleted *KU70* (*Δku70*) gene was developed. The ku70 is a protein involved in NHEJ DNA repair, so its deletion leads to a reduction of this DNA repair tool, favouring the HR one (Weninger *et al.*, 2018). In addition, it was demonstrated that the introduction of two simultaneous double-strands breaks, along with the integration of the donor DNA in the intermittent region, further optimized the recombination system (Weninger *et al.*, 2018). The advanced genome editing tool as well as the cassette targeted integration require accurate information regarding genome sequence and gene annotation. In this sense, recently it was published a refined *P. pastoris* genome sequence which added over 500 corrected sites as well as revised gene annotations (Sturmberger *et al.*, 2016a), which is expected to help to avoid off-targeting modifications.

Model proteins used in this study: from soluble to membrane-anchored proteins

The eukaryote-like mechanisms for protein folding and maturation has enabled to successfully produce a wide range of fully active soluble proteins in *P. pastoris* such as Fabs (Damasceno *et al.*, 2004, 2007; Zarei *et al.*, 2014), human-like interferons (Shi *et al.*, 2007; Wu *et al.*, 2011; Khan *et al.*, 2014), amylases (Nakano *et al.*, 2006), xylanases (Karaoglan *et al.*, 2016; Yang and Zhang, 2017; Raschmanová *et al.*, 2019), phytases (Xiong *et al.*, 2006; Zhang *et al.*, 2020) and oxidoreductases (Anasontzis *et al.*, 2014; Gu *et al.*, 2014; Krainer *et al.*, 2016). Among the soluble interesting enzymes from the industrial point of view, lipases recombinant production has been thoroughly studied

(Barrigón *et al.*, 2013; Ponte *et al.*, 2016; de Macedo Robert *et al.*, 2019; Nieto-Taype *et al.*, 2019). Lipases are serine hydrolases defined as triacylglycerol acylhydrolases (E.C. 3.1.1.3) capable of hydrolyzing carboxyl esters of long-chain acylglycerol (>10 carbon atoms). The length of the substrate chain was used as a reference criterion for distinguish lipases and esterases (Casas-Godoy *et al.*, 2012; Filho *et al.*, 2019). Although they belong to the hydrolases class, in thermodynamic favorable conditions, they also can catalyze synthesis reactions such as esterification and transesterification (Reis *et al.*, 2009). An esterification consists in binding a fatty acid and an alcohol with a covalent bound whereas transesterification includes alcoholysis, acidolysis, aminolysis and interesterification reactions. The mentioned synthesis reactions take place in a reaction medium with low thermodynamic water activity, e.g. in an organic solvent (Casas-Godoy *et al.*, 2012). Lipases are extensively used in industry owing their high stability in organic media, their high substrate specificity and product chemo-, regio- and enantioselectivity. Remarkably, lipases are able to perform the reaction of interest without the addition of expensive cofactors. Due to their features, lipases are the third most commercialized group of enzymes, representing one billion dollar a year (Hasan *et al.*, 2006) and are used in several industrial application such as:

- Biofuels industry: transesterification of vegetable/animal oils for the production of lubricants, biodiesel, and biokerosene (Jaeger and Eggert, 2002), which are mainly methyl-or other short-chain alcohol esters.
- Food industry: involved in cheese, baby food and structured lipids production. Furthermore, lipases are used as emulsifiers of animal feeding additives. Additionally, they are also used for flavor modification and fragrance compounds production (Houde *et al.*, 2004; Casas-Godoy *et al.*, 2012).

- Biopolymers industry: lipases and esterases are used for synthesis of polymeric materials. Their high chemo-, regio- , and enantioselectivity under mild reaction conditions is a major advantage for biopolymers production such as polyphenols and polysaccharides (Jaeger and Eggert, 2002).
- Detergent industry: added to detergent mixes along with other enzymes such as amylases. The high stability at high temperature and alkaline pH of lipases make them essential for the production of soap, washing products and cleaning solvent, for instance (Pandey *et al.*, 1999; Jaeger and Eggert, 2002).
- Fine chemicals: their high selectivity enable lipases to solve racemic mixtures to produce pure enantiomers. Chiral molecules are applied in cosmetic industry as surfactants and emollients and also, for medical purposes, such as anti-inflammatory drugs, prostaglandins, cephalosporines, among others (Jaeger and Eggert, 2002).
- Paper industry: for triglycerides and waxes removal; and for increasing paper whiteness.
- Bioremediation: specially for degradation of organic debris rich in oils (Hasan *et al.*, 2006).

Apart from the vast number of soluble proteins that have been produced in *P. pastoris* (Damasceno *et al.*, 2012; Looser *et al.*, 2014; Çalık *et al.*, 2015; Puxbaum *et al.*, 2015), recently the field of membrane anchored protein production has been opened in this yeast (Suades *et al.*; Ramón and Marín, 2011). However, the knowledge about the mechanisms of membrane protein production are still far away from the soluble protein ones. Therefore, obtaining high yields on membrane proteins is currently considered still challenging. Despite that, some approaches for increasing membrane protein titer have been successfully carried out. Among them, lowering temperature, media

supplementation with specific ligands, addition of chemical chaperones, UPR induction by Hac1 overexpression have been proved to be beneficial (Damasceno *et al.*, 2012).

For instance, successful attempts of displaying lipases in the *P .pastoris* cell membrane have been done in order to apply them as whole cell biocatalyst in biodiesel production processes (Jin *et al.*, 2013; Yan *et al.*, 2014; Y. Liu *et al.*, 2019). Recently, *P. pastoris* has shown up as suitable host for cytochrome P450 (CYP) production. Cytochrome P450s (CYPs) are a super-family of heme-containing membrane-associated monooxygenases capable of hydroxylating C-H bonds, epoxidate C=C bonds, N- and S- oxidations, N-, O-, S-dealkylations, C-C bond cleavage, phenolic coupling and Baeyer-Villiger-type oxidations (Bernhardt and Urlacher, 2014). Consequently, they can be involved in biotransformations of a wide range of different substrates, from simple ones such as monoterpenes, saturated fatty acids and alkanes, to very complex ones such as vitamins, steroids, antibiotics, eicosanoids (Guengerich, 2015; Bhattacharya and Yadav, 2017) and the clearance of drugs and xenobiotics in humans (Bernhardt, 2006; Schroer, Kittelmann, *et al.*, 2010). Drug metabolism generates intermediates that must be also toxicologically tested before its application approval in humans. These kinds of studies are typically elucidated by analysis of tissue samples (e.g. human liver homogenate) or microsomal preparations thereof (Guengerich, 2015; Rinnofner *et al.*, 2019), because their synthesis is unfeasible by chemical engineering, as hydroxylation of non-activated carbon atoms is highly difficult (Geier *et al.*, 2012). Consequently, recombinant production of human CYP (hCYP) has gained attention since a feasible production of xenobiotic metabolism intermediates could be performed. However, it is important to take into consideration that these human enzymes are difficult to produce in microbial hosts due to their low stability, efficiency and scalability. Alternatively, several attempts to engineer microbial CYPs

have been made in order to mimic the human metabolism. However, the control of enzyme regioselectivity still remains a challenge (Urlacher and Girhard, 2019).

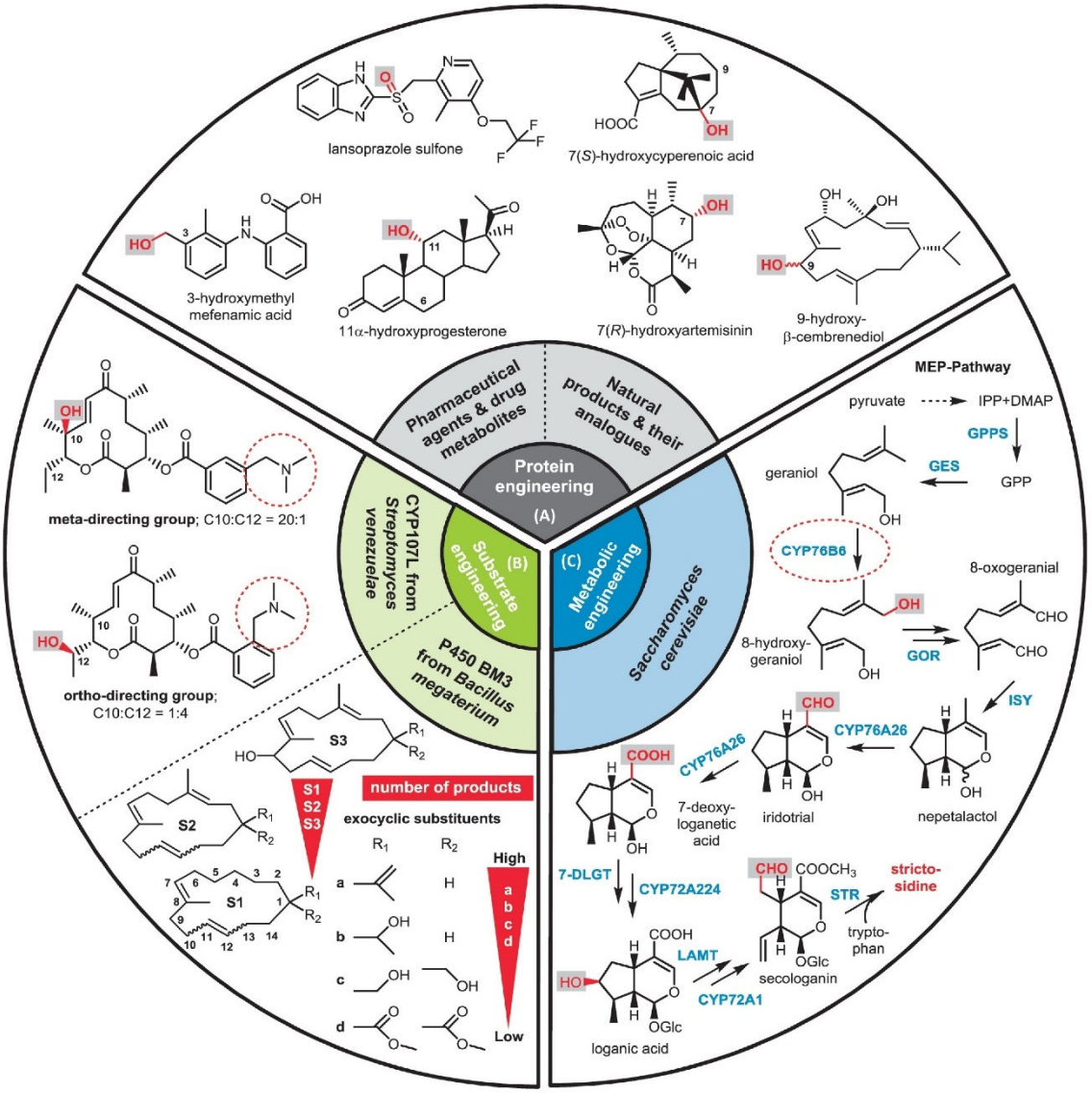


Figure 1.6. Regio- and stereoselective oxidations of pharmaceutical agents, drugs, and complex natural products achieved by (A) means of protein engineering (hydroxylation sites are marked with red colour and a grey box), and (B) means of substrate engineering (red triangles indicate increasing stereoselectivity). (C) Metabolic engineering of *S. cerevisiae* for production of strictosidine. Figure was taken from (Dong *et al.*, 2018).

In addition to the broad range of substrate types, CYPs show regio-, chemo- and stereoselective oxidations. These catalytic features offer a promising tool for selective molecule transformations that are difficult to afford by synthetic chemistry (Figure 1.6)

(Dong *et al.*, 2018). In order to further improve their properties, CYP enzymes have been modified through protein engineering and directed evolution tools, not only for increasing substrate scope, activity or stability, but also for catalysing abiotic reactions. The wide room for improvement of CYPs offers new prospects for synthetic biology (Urlacher and Girhard, 2019).

Based on their versatile applications, there is an increasing interest and high demand for hCYPs in pharmaceutical industry (Urlacher and Girhard, 2012). Regarding their catalytical activity, these enzymes require an electro transport system, which provides electrons to CYPs for oxygen activation and substrate oxidation (Nebert *et al.*, 2013). Consequently, hCPYs activity rely on the presence of a CYP reductase (CPR), which takes electrons from a cofactor NAD(P)H. Therefore, these reduced cofactors are stoichiometrically required for driving the reaction to the desired product. As the manual addition of this cofactors is not economically feasible, an efficient NAD(P)H recycling system is needed (Weckbecker *et al.*, 2010). Cofactor regeneration enzymes such as formaldehyde dehydrogenase and D-glucose-6- phosphate dehydrogenase can be co-expressed ensuring the cofactor supply. However, as an additional advantage of whole cell biocatalysis, natural yeast metabolism can be used as NAD(P)H *in vivo* regeneration system (Fischer and Glieder, 2019). Even, many efforts have been put in enhancing the *P. pastoris* NAD(P)H regeneration through altering central metabolism pathways such as MUT pathway (Schroer, Peter Luef, *et al.*, 2010; Geier *et al.*, 2015).

Protein production rate improvements through bioprocess approaches

Compromise between process control and clone screening size

The success of the strain modifications for improving the RPP performance highly depends on the cultivation conditions. Over the years, culture strategies based on substrate pulses-based shake flasks and microtiter plates have been used for testing the potential strain engineering approaches discussed in the previous sections. Although these platforms are cost-effective and time-sparing, the cultivation conditions are not accurately controlled, including physicochemical parameters —pH, dissolved oxygen (DO)— and specific growth rate, which may eventually become into a source of noise when comparing clones. Consequently, results obtained by means of these cultivations could not be considered conclusive and must be confirmed through more reliable cultivation platforms. Due to the necessity of continuous advance for the RPP industry and the lack of models that aids researchers to predict how to improve their expression systems, several high-throughput cultivation devices (HTCD) have been developed in order to broaden the strain modification success probability (Tripathi and Shrivastava, 2019). Strikingly, there is a trade-off between process control and size of clone group. Process control and monitoring become more difficult as the clone number increases, so choosing the right HTCD is essential in early stages of RPP process development (Betts and Baganz, 2006). In order to regulate the substrate consumption rate, controlled substrate delivery systems have been integrated in the previously mentioned cultivation platforms —microtiter plates, shake flasks. As a result, growth rate control can be controlled in a specific range, but pH and DO tension decrease are still considered main problems (Panula-Perälä *et al.*, 2008; Krause *et al.*, 2010).

Therefore, the imperative necessity to control the bioprocess leads to the use of bioreactors. And, in addition, as it is desirable to test as much clones as possible, the

working volume of the cultivations has to be decreased until, at least, millilitre scale. In this sense, lots of efforts has been put in developing mini-scale bioreactors (MSB) (Ali *et al.*, 2012; Soley *et al.*, 2012; Klein *et al.*, 2013; Hemmerich *et al.*, 2014) (Figure 1.7). This platform enables a relatively high throughput workflow, but at the same time, allows to mimic large-scale cultivation conditions. Optical non-invasive sensors are integrated in these systems, so some physicochemical parameters can be measured and controlled online (Long *et al.*, 2014). Importantly, Gill *et al.* checked the reliability of this platforms. In their work, they demonstrated that a 10 mL-bioreactor presented similar oxygen transfer coefficient and mixing time than a standard 7-L bench top fermenter. In addition, a high degree of similarity between the two scales was found regarding the growth and production kinetics of an *E. coli* recombinant strain producing a cyclohexanone monooxygenase. Similarly, Ali *et al.* cultivated a recombinant *E. coli* producing a Fab in both a 20-L bioreactor and a 25-mL MSB. It was found a rather similar biomass and Fab concentration at the end of bioprocesses, confirming the reliability provided by their miniaturized bioreactor system.

The screening cultivation vessels can be further reduced to microliter scale in order to increase even more the number of clones being tested in parallel. In this context, the liquid dosing for accurately controlling the bioprocess becomes a challenge. This challenge could be addressed by using microfluidics technologies (Gebhardt *et al.*, 2011) including micro-/nano- chambers, channels, valves and pumps (Wu *et al.*, 2008; Dai *et al.*, 2013). This methodology allows to manipulate nano-volumes of liquids, possibility to process automation and capability to carry out a vast number of experiments at one run. However, this technology is still not prepared for reaching high biomass concentration (Long *et al.*, 2014). Therefore, it should be used as scale down model until a certain biomass concentration value.

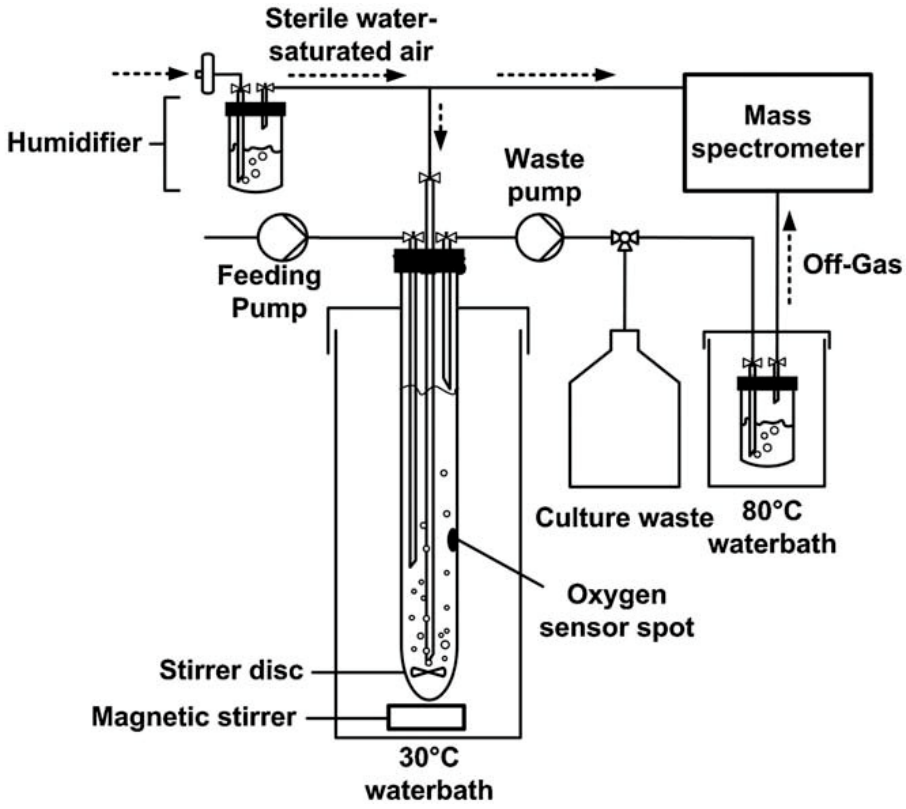


Figure 1.7. Schematic setup of a miniaturized bioreactor. Shown is one unit out of eight. Figure was taken from (Klein *et al.*, 2013).

Some biotechnology standard studies do not require a large number of cultures running in parallel, e.g. in new promoter performance characterization or DoE experiments. In these cases, the reliability of the results prevails over the number of cultivations performed in parallel. Consequently, a few of parallel bioreactor with vessel working volumes of 100 > V > 1000 mL would be ideal. In this sense, the ambr® 250 system stands out, since it enables from 12 to 24 fully automated with working volume vessels from 100 to 250 mL (Manahan *et al.*, 2019). The robustness of this parallel mini-bioreactor system has been widely tested by many researchers, as results obtained with it were rather like those acquired in bench-top fermenters or even in pilot plant fermenters (up to 1000 L) (Bareither *et al.*, 2013; Janakiraman *et al.*, 2015; Xu *et al.*, 2017).

Bioprocess operational mode

Regardless the bioreactor system chosen, bioprocesses can be performed in different operational modes. On the one hand, batch mode offers simplicity and cost-effectiveness. However, throughout the cultivation there are severe alterations in μ values, affecting the protein production kinetics. Briefly, in the early stages of the bioprocess, the high substrates concentration can potentially provoke catabolite repression, substrate inhibition or activate undesirable metabolic pathways. Both phenomena alter the cell homeostasis, and therefore, batch cultivations should not be recommended to achieve accurate kinetical characterization of strains of interest. Nevertheless, this operational mode is considered to reach medium, pH or temperature optimization (Looser *et al.*, 2014).

On the other hand, continuous and fed-batch are much more suitable operational modes for accurate recombinant strain characterization. In continuous cultivations, there are constant inlet and outlet flows. The inlet flow includes the feeding medium, containing nutrients needed for *P. pastoris* normal growth. The outlet flow include remainder feeding nutrients, biomass and ideally the product of interest (Ahmad *et al.*, 2014; Yang and Zhang, 2018). Interestingly, inlet and outlet flow rates are equal, so at a certain point of a chemostat cultivation, the macroscopic growth parameters and the metabolic ones keep constant. This leads to the so-called steady state, which is the most valuable feature of this operational mode. As the bioprocess kinetics are kept constant over time, this operational mode enables to characterize an expression system in depth, integrating kinetic values and omics-like information (Graf *et al.*, 2009). A comprehensive example that shows the suitability of chemostat mode for strain characterization is the *P. pastoris* multi-level study behaviour under different oxygen availability conditions (Baumann *et al.*, 2010). The control of the culture state enabled to acquire reliable data from different

omics analyses such as proteomics, transcriptomics and fluxomics, and their integration with the macroscopic kinetic values. From this thorough multi-level study, important insights regarding *P. pastoris* response to oxygen availability were revealed. In hypoxic conditions, a strong regulation central carbon metabolism of *P. pastoris* was pointed out. Specifically, glycolytic genes and the genes of the non-oxidative branch of pentose phosphate pathway were significantly upregulated whereas the Tricarboxylic Acid (TCA) pathway was markedly downregulated. Importantly, the central carbon metabolism regulation was positively correlated with proteomic and metabolomic data (Baumann *et al.*, 2010).

Remarkably, in chemostat cultivations the product quality is also kept constant as well. Therefore, it leads to much more reproducible process and homogeneous products (Zydney, 2015). In fact, in the FDA's recent Regulatory Science Strategic Plan, it is specified the continuous manufacturing as one of the main areas that would support the quality of both the bioprocess and the product (Strategic Plan for Regulatory Science | FDA). However, even though it is being intensively investigated as an alternative to fed-batch cultivations, full continuous manufacturing is not affordable nowadays by most companies since it requires coordination between bioprocess and downstream processes. Other challenging aspects are the contamination risks and the strain genetic stability over long periods of time (Zydney, 2015).

In fed-operations, continuous or periodic feeding of medium is pumped in the culture broth. This strategy allows to control the culture substrate consumption rate and, thus, the specific growth rate. Therefore, a high degree of control is reached through this feeding strategy. Moreover, controlling the substrate consumption rate assures to keep the substrate concentration below growth inhibitory values (Doran, 2013). Furthermore, fed-batch mode is closer to the biotechnology companies current technologies, so its

application is often considered more straightforward, in contrast to the continuous mode. In addition, as there is not any outlet flow, the product and the biomass are accumulated, and it allows to reach higher concentration values than in continuous cultivations. Achieving high product titer is essential for easing downstream processes, especially for high added value products.

Carbon source selection according to the expression system used

One of the most interesting advantage of *P. pastoris* as host for RPP processes at industrial scale is that do not have complex nutrient requirements, so can be grown in cheap, mineral media (Maurer *et al.*, 2006) and to reach biomass concentrations beyond 100 g L^{-1}. Therefore, medium composition depends mainly on the expression system used for triggering the RPP. In the case of using glycolytic constitutive promoters —P_{GAP}, P_{PGK1}, P_{PYK}, P_{PDC}; both glycerol and glucose have been used as sole C-sources. Waterham *et al.* first observed that the *GAP* mRNA amount in presence of glycerol was 33% less than in cells grown on glucose. Besides, the P_{GAP}-driven β-lactamase production was superior on *P. pastoris* cells grown on glucose by 27% (Waterham *et al.*, 1997). Pal and co-workers gathered similar results in terms of carbon source comparison, since the human granulocyte-macrophage colony-stimulation factor (hGM-CSF) production was boosted up to 75% using glucose as carbon source, although rich medium was employed (Pal *et al.*, 2006). Slight glutathione production differences were found by Fei *et al.* when using either glucose or glycerol as carbon source (Fei *et al.*, 2009). García-Ortega *et al.* went beyond in the C-source study since their work pointed out the benefits of each C-source in the different bioprocess phases (Garcia-Ortega *et al.*, 2013). In C-source excess, glucose led to a higher secretion of fermentative by-products, affecting the final biomass reached. Consequently, glycerol was considered more suitable for this bioprocess phase. However, this substrate increases the culture oxygen demand and the heat production in

comparison with glucose, raising the process costs. Therefore, glucose was preferred for the fed-batch phase in C-limiting conditions, since process costs regarding oxygen supply and cooling requirements are reduced and, in addition, by-products generation could be prevented by the C-limiting operational strategy. On the other hand, expressing the heterologous gene under the P_{AOX1} requires the presence of methanol as the inducer during the fed-batch phase. The *P. pastoris* strain MUT phenotype has a huge impact on the carbon source selection. Mut^+ strains are able to consume methanol relatively fast, so this compound could be used as sole C-source for RPP induction and to provide C for biomass growth (Barrigón *et al.*, 2013; Ponte *et al.*, 2016, 2018). However, Mut^S strain lacks the *AOX1* gene, whose expression represents the 90% of the peroxisomal Aox. Consequently, the yeast growth rate under methanol is around 5-fold lower than a Mut^+ strain. Therefore, Mut^S bioprocesses using methanol as sole carbon source would be time-consuming which is detrimental for protein productivities. In order to avoid these productivities loses, mixed substrate strategies are often followed, where either glycerol or sorbitol have been used as substrate along with methanol (Arnau *et al.*, 2011). Interestingly, derepressed promoters obtained from the MUT pathway broaden the substrate utilization possibilities. In general, these promoters are characterized by being derepressed when carbon sources such as glycerol or glucose are depleted and, also, can be further induced by methanol. Therefore, if maximum promoter expression is needed, a cosubstrate bioprocess strategy should be designed (Rajamanickam *et al.*, 2017). Nevertheless, sometimes maximizing the promoter activity is not recommended because it might overload the protein folding/processing pathway. In this case, a C-limiting process using either glucose or glycerol could be considered enough to reach high level of protein titer, thus avoiding also the use of the hazardous methanol as inducer.

Medium feeding as strategy for bioprocess control

Although substrate selection is nowadays quite established for RPP under the standard expression systems —P_{GAP}, P_{AOX1}, several substrates feeding strategies have been tried out for optimizing bioprocesses performance in fed-batch mode. The simplest feeding strategy is the pulses-based one. Basically, a substrate pulse of non-inhibitory concentration is injected in the bioreactor once the substrate is depleted in the culture broth. The substrate concentration can be measured by either on-line, in-line, at-line methods. Furthermore, an alternative and usually easier option is to detect indirectly the substrate depletion by means of sharp changes in DO tension, pH, temperature, CO_2 and O_2 molar fraction in the off-gas or respiratory quotient (RQ) (Çalık *et al.*, 2015). This feeding strategy has been widely followed by several researchers regardless expression system used. In a bioprocess designed for hydrophobin HFB1 production under P_{GAP}, Kottmeier *et al.* measured the glucose concentration of the cultivation through the widely known DNS reagent protocol (Kottmeier *et al.*, 2012). Once glucose was depleted, two subsequent pulses were administered. Likewise, Hu *et al.* aimed to optimize the glycerol feeding procedure, in order to improve the production of a methionine adenosyltransferase driven by the same promoter. Among some feeding strategies proved, a two-pulse strategy was followed, in which 20 g L^{-1} pulses were injected when C-source drops to near zero values, indicated by sudden DO peaks (Hu *et al.*, 2008).

In the case of P_{AOX1}-driven RPP bioprocesses, the substrate changes from batch —glucose, glycerol— to induction phase —methanol— must be taken into consideration since *P. pastoris* usually requires a period of time to adapt its metabolism machinery for consuming methanol. This is the so-called transition phase. The transition phase of most of the works are triggered by simple low concentration methanol pulse (Dietzsch *et al.*, 2011; Hesketh *et al.*, 2013). However, other researchers reported an alternative transition

phase based on mixed glycerol-methanol feeding (Minning *et al.*, 2001). Briefly, they stated that by stepwise decreasing a glycerol constant flow rate and, simultaneously, increasing a methanol constant flow rate, the specific consumption rate in the beginning of the induction phase was enhanced. Therefore, the global protein productivity was higher than using the standard pulse-based transition procedure.

After the transition phase, culture can be induced with methanol as sole carbon source or by a co-substrate feeding. By adding up to 4 consecutive pulses of 1.0% (v/v), Dietzsch *et al.* calculated the maximum specific growth rate and the maximum specific methanol consumption rate of 5 Mut^S different strains (Dietzsch *et al.*, 2011). Nevertheless, this feeding strategy leads to a very unstable bioprocesses, as μ_{max} and no growth periods are alternated depending in the substrate availability. Consequently, bioprocesses reproducibility is low, which is detrimental for the product quality standardization. Therefore, this kind of strategies should be avoided.

In order to control the μ into a narrower range, stepwise/linear increasing feeding flow rates could be applied (Goodrick *et al.*, 2001; Gasser *et al.*, 2006; Tolner *et al.*, 2006; Fei *et al.*, 2009; Mellitzer *et al.*, 2012). Tolner *et al.* wrote a general protocol in which some recommendations regarding methanol feeding are given (Tolner *et al.*, 2006). This feeding strategy mainly consists in applying a constant methanol flow rate, which is increased every two hours in order to adapt it to the culture growth. Likewise, Mellitzer *et al.* implemented a similar methanol feeding strategy for the production of three lignocellulosic enzymes driven by the P_{AOX1} in *P. pastoris* (Mellitzer *et al.*, 2012). Concretely, induction phase was started with a methanol flow rate of 2 $g_{Meth}\ L^{-1}\ h^{-1}$, which was stepwise increased up to a methanol flow rate of 6 $g_{Meth}\ L^{-1}\ h^{-1}$ at the end of the cultivation. This strategy does not enable an accurate control of cell growth but at least the μ is kept in a relatively constricted range. The same work also carried out the

lignocellulosic enzymes production under synthetic promoters which are derepressed in absence of C-source in the culture broth. Fed-batch cultivations using these novel expression systems were performed with constant feeding of glycerol. In contrast to the previous case mentioned, a glycerol flow rate of 6 g_{Meth} L^{-1} h^{-1} was maintained throughout the entire bioprocess. Consequently, the μ covered a wide range of values during bioprocess, from high at the early stages to low at the late ones.

A P_{AOX1}-driven widely used alternative to feed-forward μ control, is the implementation of a closed-loop methanol control system (Yamawaki *et al.*, 2007; Gurramkonda *et al.*, 2010; Barrigón *et al.*, 2013; Gu *et al.*, 2014; Ponte *et al.*, 2016). Interestingly, Barrigón *et al.* compared both bioprocess control strategies in the production of Rol driven by the P_{AOX1}. Concretely, Rol production in C-limiting conditions at different μ set point values —0.015, 0.025 and 0.045 h^{-1}— were compared with those obtained in C-excess at different concentrations —1, 2, 3, 5 and 10 g L^{-1}. On the one hand, increasing μ was proved to be detrimental for Rol production, since among C-limiting conditions, the lowest μ delivered the best q_p value. On the other hand, the μ values showed in C-excess conditions was similar to the highest μ tested in C-limiting conditions —0.045 h^{-1}, so it might be considered the μ_{max} of this recombinant strain under methanol consumption. The q_s values showed in C-excess were slightly higher than in C-limiting conditions —around 10% higher. Despite the rather similar μ and q_s values, the best C-excess condition —[MeOH]=3 g L^{-1}— exceeded the q_p value of the higher C-limiting condition tested in which μ_{sp} was set to 0.045 h^{-1} by almost 10-fold (Barrigón *et al.*, 2013). Therefore, in this specific case, the methanol q_s and the q_p are not directly related as occurred in previous P_{AOX1}-driven RPP bioprocesses commented (Lee *et al.*, 2003). As mentioned above, the relationship q_s-μ- q_p is protein dependent and must be exhaustively studied in each case. Remarkably, bioprocess strategies based on a closed-loop control of methanol allow the

cells to grow next to its μ_{max}. However, the transition phase becomes extremely important. Whether cells are not perfectly adapted to methanol once the induction phase is started, it could alter the methanol consumption capabilities of the yeast, risking the reproducibility of the bioprocesses.

In this sense, the feeding strategy that enables the simplest bioprocess control and satisfactory reproducibility is the pre-programmed exponential feeding rate. Consequently, many RPP bioprocesses have been implemented at μ-stat controlled conditions (Potvin *et al.*, 2012; Looser *et al.*, 2014; Çalık *et al.*, 2015). The amount of substrate required for the cells at each moment of the bioprocess is calculated in advance through the following equation from the mass balance:

$$F(t) = \frac{\mu X_0 V_0}{(S_0 - S) Y_{X/S}} exp\ (\mu_0 t)$$

Where F is the volumetric flow rate of substrate, μ is the pre-determined or set-point for the specific growth rate, V_0 is the initial volume of the fermentation medium, medium, X_0 is the initial cell concentration, S_0 is the feeding substrate concentration, S is the substrate concentration in the culture broth and $Y_{X/S}$ is the overall biomass-to-substrate yield.

Previously to apply this feeding strategy, it is essential to determine μ_{max} of the recombinant strain. The μ setpoint must be lower to μ_{max}, thus the specific substrate consumption requirement will be always lower than the strain $q_{s,max.}$ Therefore, the added substrate is immediately consumed, assuring C-limiting conditions during the whole bioprocess. The C-limiting conditions allows full control of cell growth, since cells can only grow as fast as the feeding substrate flow rate (Looser *et al.*, 2014; Çalık *et al.*, 2015).

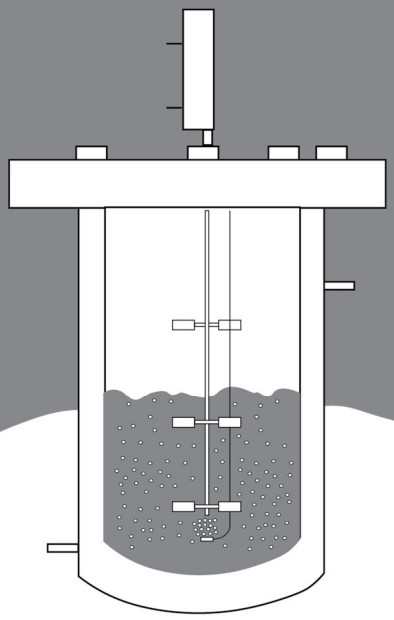
In chemostat cultivations, the broth volume is maintained constant by adjusting the inlet and outlet flow rates. The μ is thus constant and a steady state is reached when

concentrations in the system adjust themselves to the feed rate and keep constant over time. The so-called production kinetics is the relationship between μ and q_p and has been thoroughly studied both in fed-batch and continuous cultivations. Many research works have demonstrated the crucial impact of μ on recombinant producer strains. Hence, from the industrial point of view, the relationship between these two parameters must be determined for each expression system and for the specific target recombinant protein. To date, a positive relationship between both parameters has been commonly found in P_{GAP} and P_{AOX1}-driven RPP bioprocesses (Jungo *et al.*, 2006; Garcia-Ortega *et al.*, 2013, 2016; Rebnegger *et al.*, 2014; Nieto-Taype *et al.*, 2019). It seems obvious that, since both endogenous genes play crucial roles in the glycolysis and MUT pathway, respectively, the growth rate and the production rate should be coupled. Furthermore, Rebnegger *et al.* reported interesting gene expression regulation in response to μ increases. Concretely, among the upregulated genes with increasing μ, there were found many UPR-associated genes, other related with protein translocation, folding and glycosylation. Remarkably, the UPR master regulator *HAC1* was also overexpressed, which led to an enhanced expression of some cytosolic chaperones. These global overexpression of UPR might contribute to the higher q_p values found in higher μ (Rebnegger *et al.*, 2014).

Strikingly, Nieto-Taype *et al.* reported that the RPP kinetics trend is promoter strength dependent (Nieto-Taype *et al.*, 2019). It was demonstrated by the P_{GAP}-driven production of the *Candida rugosa* lipase 1 (Crl1) by two recombinant clones harboring different heterologous gene dosage. The single copy clone showed a linear production kinetics. However, increasing the *CRL1* expression by adding 4 extra copies resulted in a production kinetics saturated pattern. Authors hypothesize that the significant raise of heterologous gene expression led to a bottleneck in at least one of the subsequent protein processing steps, resulting in a rather low impact of μ effect on q_p (Nieto-Taype *et al.*,

2019). On the other hand, despite recombinant production under P_{AOX1} should be considered growth-coupled when using methanol as sole carbon source, many researchers described bell shaped (Yamawaki *et al.*, 2007; Hang *et al.*, 2008; Potgieter *et al.*, 2010) or even negatively growth related (Kobayashi *et al.*, 2000; Min *et al.*, 2010) production kinetics profiles in RPP bioprocesses driven by the P_{AOX1}. For instance, Canales *et al.* noticed that the q_p presented a maximum at μ around 0.06 h^{-1} in chemostat cultivations for Rol production under this methanol-inducible promoter (Canales *et al.*, 2015b). As commented previously, one of the first experiments that must be performed with new expression system/recombinant protein is indeed shedding light on the μ impact on production rates. Prielhofer and co-workers noticed a bell-shaped production kinetics in the recombinant production of i-bodies under an engineered version of a glucose transporter promoter (P_{GTH1}). Even, the production kinetics study enabled them to design an optimized fed-batch cultivation strategy based on stepwise decreasing μ over bioprocess time (Prielhofer *et al.*, 2018).

STATE OF THE ART



2

Biotechnology has found its marketplace in the field of high added value compounds, such as recombinant proteins and microbial metabolite production (Chen and Kazlauskas, 2011; Chen, 2012; Erickson *et al.*, 2012). Among the microbial candidates for hosting the production of the mentioned molecules, *P. pastoris* has emerged as a promising microbial host because it combines the positive features of bacteria, such as fast growth, minimal nutrients requirements and available strong and tightly regulated expression systems; and also, the eukaryote protein processing mechanisms such as the ability to perform post-translational modifications (PTMs), which are often essential for the molecules biological activity (Cereghino and Cregg, 2000; Çelik and Çalik, 2012; Fidan and Zhan, 2015).

Up to now, the most used expression system in *P. pastoris* has been the methanol inducible P_{AOX1}-based due to its expression strength and its tight regulation (Cregg *et al.*, 1989; Looser *et al.*, 2014; Yang and Zhang, 2018). Furthermore, valuable information regarding MUT pathway regulation has been provided by many research works (Lin-Cereghino *et al.*, 2006; Kranthi *et al.*, 2009; Zhan *et al.*, 2017). These studies have even given ground to the possibility to deregulate this pathway in order to prevent the use of the hazardous methanol as carbon source for growth and heterologous gene induction, which present several operational drawbacks (Shen *et al.*, 2016; Vogl, Sturmberger, *et al.*, 2018; Shirvani *et al.*, 2019). Among the vast information provided regarding MUT pathway regulation, it was suggested a limitation in the TFs amount involved in P_{AOX1} induction, mostly in clones harboring several P_{AOX1}-based cassettes (Cámara *et al.*, 2017a). Deregulating TFs involved in P_{AOX1} induction has been proved this hypothesis as through overexpression of Mxr1 and Mit1 led to an enhanced growth, methanol consumption and recombinant protein productivity in a recombinant multicopy clone (Cámara *et al.*, 2019). This result would explain why increasing the heterologous gene dosage might become detrimental when working with high target-gene dosage producer.

Often, the studies concerning MUT pathway regulation and, specifically, *AOX1* regulation have been made only in single growth rate conditions. However, μ impact on this crucial gene regulation have not been tested yet, despite it has been widely proved that q_p highly depends on μ conditions in P_{AOX1}-driven RPP bioprocesses. This effect has been also reported in other *P. pastoris* expression systems such as the constitutive *GAP* promoter. Furthermore, *AOX1* regulation in conditions with limiting MUT-related TFs —e.g. multicopy clones— and its effect on q_p are still underinvestigated.

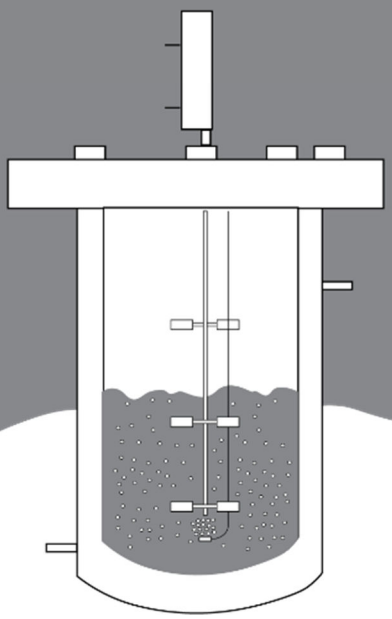
Moreover, alternative promoters involved in the MUT pathway gene have been also tested for driving the heterologous gene expression. On the one hand, those that present the same regulation type that P_{AOX1} (Tschopp *et al.*, 1987)—e.g. P_{DAS1}, P_{DAS2}— lack of interest since methanol is still essential for driving the RPP process. On the other hand, promoters that depict a deregulation pattern —e.g. P_{CAT1}, P_{FDH}— are gaining importance because they enable a tight repression on glucose/glycerol excess conditions whereas they turn active when these substrates are depleted. The derepressed promoter activity under C-limiting conditions are shown to be significantly lower that the widely-known P_{AOX1}, but their activity could be further induced by the addition of methanol as co-substrate. The tunability of these promoters activity is an additional advantage of the MUT-related derepressed promoters, since it allows to adjust the heterologous gene expression levels according to the protein amount demands e.g. for balancing multigene coexpression in synthetic biology applications (Vogl, Hatzl, Gerstmann, Pitzer, Wagner, G. G. Thallinger, *et al.*, 2016). Alternatively, in order to prevent methanol utilization for driving RPP bioprocesses, constitutive systems based glycolytic promoters can be used—mostly P_{GAP} (Waterham *et al.*, 1997; Potvin *et al.*, 2012; Vogl and Glieder, 2013). Their constitutive expression represents a hurdle as their activity could not be tuned by means of substrate-derived bioprocess strategies. Anyway, plenty of works have reported very interesting

results based on P_{GAP} for RPP, yielding protein titer in the range of P_{AOX1} (Looser *et al.*, 2014; Çalık *et al.*, 2015; García-Ortega *et al.*, 2019). Nevertheless, the knowledge about P_{GAP} regulation is still poor, although basic early promoter analysis pointed out potential major promoter improvements by rational promoter engineering approaches (Ata *et al.*, 2017). Due to the inherent drawbacks associated with the widely used P_{AOX1} and P_{GAP}, the novel expression system discovery workload is exponentially increasing in order to find strong and fine-tunable promoters that could potentially overcome the classical ones.

On the other hand, protein processing and secreting machinery in *P. pastoris* has shown to be often limiting factors in RPP. A misfolded protein overload provokes the UPR induction, which involves overexpression of helper proteins as well as ERAD enhancement, due to the inability to fold and secrete high misfolded burden (Guerfal *et al.*, 2010; Puxbaum *et al.*, 2015). Consequently, protein processing pathway has been an optimization target for several years in *P. pastoris* (Damasceno *et al.*, 2012; Puxbaum *et al.*, 2015), by means of exploiting the recent progress achieved in the field of advanced genetic engineering tools, such as CRIPSR/Cas9 (Weninger *et al.*, 2016), genome annotation (Sturmberger *et al.*, 2016a) and controlled high-throughput screenings (Tripathi and Shrivastava, 2019). RPP in *P. pastoris* have been mainly focused on soluble proteins, taking advantage of the secretion capability of this yeast (Cereghino and Cregg, 2000; Ahmad *et al.*, 2014; Looser *et al.*, 2014). However, in recent years, *P. pastoris* has shown up as a candidate for hosting membrane-anchored protein production (Suades *et al.*; Ramón and Marín, 2011). The membrane protein production as whole cells biocatalysts prevents costly protein purification procedures and increases the stability of the protein, which represents interesting operational advantages. Furthermore, the yeast metabolism could be used as cofactor recycling system for biocatalysts that require cofactor in stoichiometrically amounts to drive the target reaction, such as CYPs (Geier

et al., 2015). Nevertheless, bioprocess optimization of membrane-anchored protein production has been barely investigated, so the membrane protein production kinetics has still much more room for improvement.

OBJECTIVES



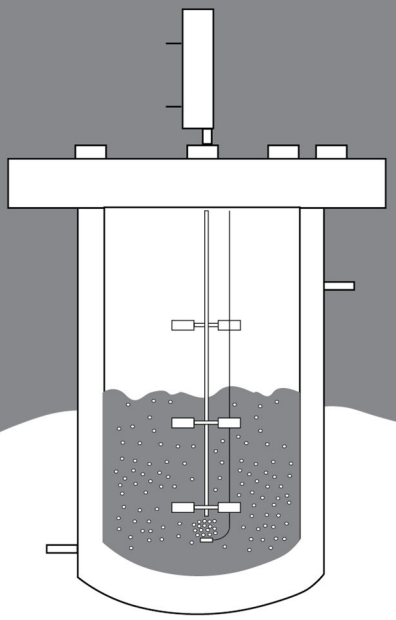
3

In general terms, the main objective of this thesis has been to establish a platform to characterize expression systems for RPP bioprocesses in the yeast *P. pastoris*. This platform consists in the cultivation of recombinant clones in bench-top fermenters, assuring a tight control of cultivation conditions. The robustness of this cultivation system enables to gather reliable clone information from both the macrokinetic and transcriptional levels. Moreover, the clone behavior under different culture conditions and operational modes —continuous, fed-batch— has also been studied in this thesis. This detailed characterization strategy, at both microkinetic and transcriptional levels, have been applied to different case studies producing enzymes of industrial interest, allowing to split the main objective into three partial ones:

- To study the influence of specific growth rate and gene dosage on MUT regulation and, thus, on P_{AOX1}-driven recombinant *Candida rugosa* lipase 1 (Crl1) production, on continuous and fed-batch cultivations. Furthermore, the performance comparison of P_{AOX1} and P_{GAP} as promoters for *CRL1* gene expression is also analyzed.
- Once analyzed the suitability of the proposed strategy for characterization of the classical P_{AOX1} and P_{GAP} in Crl1 production bioprocesses, a deep characterization of the regulation patterns and performance of two novel promoters for *P. pastoris* RPP bioprocesses, called P_{PDF} and P_{UPP}, is pursued. Moreover, their promising features as expression systems for RPP in *P. pastoris* are evaluated in comparison with the widely used P_{GAP}, for *Candida antarctica* lipase B (CalB) production.
- To evidence the applicability of the previously studied P_{PDF}, in a dual-gene recombinant protein system in *P. pastoris*. This strain comprises the recombinant production of an industrially interesting Cytochrome P450 (CYP2C9) with a proper reductase partner, cytochrome P450 reductase (CPR), forming an active

whole cell biocatalyst capable of oxidize ibuprofen. The CYP2C9/CPR production enhancement includes an optimization of the specific growth rate as well as the impact of methanol use as co-substrate.

METHODS



4

Methods content

Expression vectors construction and recombinant strains generation

Recombinant strains of *P. pastoris* expressing *CRL1* gene under the regulation of P_{AOX1} were constructed by using the pPICZαA plasmid (Invitrogen, Carlsbad, CA, US) assembled with the codon-optimized synthetic open reading frame (ORF) encoding the *CRL1* gene sequence (GeneScript, Piscataway, NJ, USA). Ten *P. pastoris* X-33 cells (Invitrogen, Carlsbad, CA, US) were transformed with chimeric vector under the conditions described elsewhere (Cámara *et al.*, 2019). Briefly, 100 ng of PmeI-linearized pPICZαA_*CRL1* was introduced into *P. pastoris* competent cells by electroporation by means of a Gene Pulser X cell™ Electroporation System (Bio-Rad, Hercules, CA). Electroporation conditions were 1,500 V, 25 µF, and 200Ω. Transformants were isolated on yeast extract peptone glycerol plates (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glycerol, 2% (w/v) bacteriological agar) containing 100 µg/mL Zeocin (Invitrogen, San Diego, CA). Positive clones were confirmed by PCR (after genomic DNA extraction with Wizard® Genomic DNA Purification Kit; Promega, Madison, WI).

Recombinant clones expressing *CALB* gene under the control of P_{GAP}, P_{PDF} and P_{UPP}; as well as CYP2C9/CPR expressing clone under the bidirectional P_{PDF}/P_{DC} were kindly provided by Bisy (Graz, Austria).

Gene dosage determination

The number of *CRL1* expression cassettes integrated into the genome was determined by droplet digital PCR (ddPCR) according to Cámara *et al.* (Cámara *et al.*, 2016). In short, genomic DNA (gDNA) of positive clones were restricted with BamHI and EcoRI (ThermoScientific™, Waltham, MA) in order to obtain DNA fragments smaller than 5kb,

assuring separation of *CRL1*-containing cassettes and preventing DNA overestriction. 0.4 ng of restricted gDNA belonging to each sample were added to reaction mixture, which is equivalent to 40000 copies of haploid *P. pastoris* genome. Each PCR was performed in a 20-µL reaction mixture containing 12.5 µL of ddPCR™ Supermix (Bio-Rad, Hercules, CA), 300 nM of each primer and the required amount of nuclease-free water. Droplets containing PCR reaction mixtures were built by the Droplet Generator (Bio-rad, Hercules, CA) and transferrer into a 96-well plate. PCR reactions placed in the 96-well plate were started by incubation at 95˚C for 10 min, followed by 40 cycles of denaturation-annealing-extension steps. DNA denaturation was carried out at 94˚C for 30 seconds, and the annealing-extension cycles were performed at 57.4˚C for 1 min. Positive droplets were detected using the QX100 Droplet Digital PCR System and the software QuantaSoft v.1.5.38 (Bio-Rad, Hercules, CA). Gene copy number was determined by calculating the ratio between positive copies of *CRL1* and the housekeeping gene *ACT1*. The specific primers used are presented in Table 4.1.

Total RNA extraction

Chemostat samples for RNA isolation were collected according to Landes *et al.* (Landes *et al.*, 2016). Pellets from 1 mL culture broth samples were resuspended in 1 mL of TRIzol™ reagent (Waltham, Massachusetts, USA) and lysed with glass beads (425–600 µm, Sigma-Aldrich, St. Louis, MO, USA) for mechanical disruption. Cell lysis was attempted by alternating cycles of 30-s of vortexing and cooling. All further steps were performed according to the manufacturer's instructions. RNA integrity was checked by 1% agarose (wt/vol) electrophoresis and the RNA concentration was determined with a Nanodrop 2000 instrument from Thermo Scientifc™ (Waltham, MA, US).

Synthesis of cDNA and determination of transcriptional levels

cDNA was synthetized with the iScript™ cDNA Synthesis kit (Bio-Rad, Hercules, CA, USA), following the manufacturer's instructions. For qPCR, a set of primers were designed for specific target cDNA. The set of selected genes comprised:

Table 4.1. Primer pairs used for gene dosage analyses and relative transcription levels determination by means of ddPCR and qPCR, respectively. Gene dosage determination was carried out only for heterologous genes.

#	Name	Product	Function in analysis	Sequence
1	CRL1_FW	*CRL1*	Heterologous gene	CCTGAGGGTACTTACGAAG
2	CRL1_RE			CCAGGTGGTCTAACAACG
3	CALB_FW	*CALB*	Heterologous gene	GATCAACCACTGGTCAGGCAAGATC
4	CALB _RE			CGCAATTCTGTTTAGGACCTGCAACG
5	MTH1_FW	*MTH1*	Housekeeping gene	TACGACATGGTTCCTCCCCCTTG
6	MTH1_RE			CGTTTCCTCTTGGACTCGTCTATCGTC
7	AOX1_FW	*AOX1*	MUT gene	GACATTCACGGTTTCGAAGG
8	AOX1_RE			CCTCAAGAAGTCCTGGCAAAC
9	CALB_FW	*MIT1*	MUT gene	CGTGAATCTGCAACAACAGC
10	CALB _RE			CGGATCTGAACTGCCAGAG
11	KAR2_FW	*KAR2*	UPR reporter gene	GGTAGTCATAACGCCACCAGTAGTCTC
12	KAR2_RE			GGAATTAACCCAGATGAAGCTGTCGC
13	HAC1_FW	*HAC1*	UPR reporter gene	GGTTGGAAGCCTTAGGTGGTACCG
14	HAC1_RE			CCTCAGTCAAAGATCTGCGAGTGG
15	ERO1_FW	*ERO1*	UPR reporter	GGAATGGTGATGAGGGATTCTGCAAG
16	ERO1_RE			GAGGACAGCTCATTTTCATCTTGGCCC

- Chapter 1 of results: *CRL1* (heterologous gene); *AOX1*, the alcohol oxidase 1 native gene; and *MIT1*, which codifies a key transcription factor of the methanol-induced transcription.

- Chapter 2 of results: *CALB* (heterologous gene); *ERO1*, which is involved in the oxidative protein folding machinery in the ER and in the UPR.
- Common in both 1 and 2 result chapters: *KAR2*, which codifies for two ER-resident chaperones; and *HAC1*, which is widely considered as UPR master regulator. Expression of both genes are known to be regulated by UPR.

For qPCR, reactions were done with SYBR™ Select Master Mix (Thermo Scientific™,Waltham, MA, US). Additionally, and as suggested by the manufacturer to assure the maximum accuracy, the reaction mixes were made by EpMotion® 5075 robot (Eppendorf, Germany). The amplification program was run on a QuantStudio 12 K Flex Real-Timer from Thermo Scientific™ (Waltham, MA, US), following the manufacturer's instructions. The annealing extension temperature was set at 57.4 °C. Relative transcript levels (RTLs) were determined by using *MTH1* as a housekeeping gene as it showed basal expression across the conditions tested.

Chemostat cultivations

Chemostat cultivations were performed in a 2L Biostat B Bioreactor (Braun Biotech, Melsungen, Germany). The working volume was 1L. The batch phase lasted around 24 h. When this phase was finished, feeding and outlet were plugged in, starting the continuous cultivation. In case of chemostat cultivations with methanol as sole carbon source, two methanol pulses of 2 g L^{-1} were injected in the culture broth before feeding initiation in order to adapt cells to the new substrate. Feeding flow rate was calculated for fulfilling the selected dilution rate (D) conditions in each case:

$$D = \frac{F}{V}$$

Where *D* is the dilution rate (h^{-1}), F is the feeding flow (L h^{-1}) rate and V (L) is the fermentation working volume. Media salts stoichiometry was previously described (Maurer *et al.*, 2006), although C-source and concentrations were different depending on the recombinant clone used:

- Batch media (per liter): 1.8 g citric acid, 12.6 g $(NH_4)_2HPO_4$, 0.02 g $CaCl_2 \cdot 2H_2O$, 0.9 g KCl, 0.5 g $MgSO_4 \cdot 7H_2O$, 40 g glycerol, 4.6 mL PTM1 trace salts stock solution.
- Continuous (per liter):
  - Crl1 production cultivations: 0.9 g citric acid, 4.4 g $(NH_4)_2HPO_4$, 0.01 g $CaCl_2 \cdot 2H_2O$, 1.7 g KCl, 0.65 g $MgSO_4 \cdot 7H_2O$, 50 g methanol, 0.2 mL antifoam, 1.6 mL PTM1 trace salts stock solution.
  - CalB and CYP2C9 production cultivations: 0.46 g citric acid, 2.2 g $(NH_4)_2HPO_4$, 0.005 g $CaCl_2 \cdot 2H_2O$, 0.85 g KCl, 0.33 g $MgSO_4 \cdot 7H_2O$, 25 g glycerol, 0.2 mL antifoam, 0.8 mL PTM1 trace salts stock solution.
- PTM1 trace salts stock solution (per liter): 6.0 g $CuSO_4 \cdot 5H_2O$, 0.08 g NaI, 3.0 g $MnSO_4 \cdot H_2O$, 0.2 g $Na_2MoO_4 \cdot 2H_2O$, 0.02 g H_3BO_3, 0.5 g $CoCl_2$, 20.0 g $ZnCl_2$, 65.0 g $FeSO_4 \cdot 7H_2O$, 0.2 g biotin and 5.0 mL H_2SO_4 (95%–98%).

Chemostat conditions were monitored and controlled at the following set points:

Table 4.2. Culture conditions of the chemostat cultivations conducted throughout the thesis

	Chemostat conditions		
	Crl1	**CalB**	**CYP2C9**
pH	6.0	5.0	5.0-7.0
Dilution rate (h^{-1})	0.025-0.095	0.05-0.15	0.10
Temperature (˚C)	25.0	25.0	25.0
Pressure (bar)	1.2	1.2	1.2
pO_2 (%, air saturation)	Above 25	Above 25	Above 25
Stirring rate (rpm)	700	700	700
Inlet gas flow (vvm)	0.80	0.80	0.80

pH was controlled by means of 15% NH_4OH (w/v) addition. Inlet gas flow was controlled by means of mass-flow controllers (Bronkhorst Hi-Tech, Ruurlo, The Netherlands). Feeding flow rate was controlled by means of an IPC peristaltic pump (Ismatec, Germany). A condenser was integrated in the exhaust gas line to minimize water losses by evaporation. Under each set of conditions, operation was conducted at least up to five residence times. To ensure that a steady state was reached, samples were analyzed from the third residence time, to check the stability of the parameters of interest.

Fed-batch cultivations

Fed-batch cultivations were carried out in a 5 L Biostat B Bioreactor (Sartorius, Melsungen, Germany). The batch phase lasted around 24 h. Once batch phase was finished, feeding was started with 2 L as initial volume. Feeding salt composition was already described (Garcia-Ortega *et al.*, 2013), although C-source was selected according to the recombinant clone used.

- Fed-batch feeding (per liter)
  - Crl1 production cultivations: 0.35 g $CaCl_2 \cdot 2H_2O$, 10 g KCl, 6.45 g $MgSO_4 \cdot 7H_2O$, 400 g methanol, 30 mL PTM1 trace salts stock solution, 0.4 mL antifoam.
  - CalB and CYP2C9 production cultivations: 0.35 g $CaCl_2 \cdot 2H_2O$, 10 g KCl, 6.45 g $MgSO_4 \cdot 7H_2O$, 400 g glycerol, 30 mL PTM1 trace salts stock solution, 0.4 mL antifoam. 100% methanol (w/v) was added as cosubstrate in CYP2C9 production fed-batches for testing its impact on protein production.
- PTM1 trace salts stock solution. Same as used in chemostat cultivations.

In fed-batch bioprocesses, a tight control of μ was achieved by applying pre-programmed exponential feeding rate profiles. Mass balance equations enabled the calculation of the initial feeding rate applied right after batch phase end:

$$F(t) = \frac{\mu X_0 V_0}{(S_0 - S) Y_{X/S}} exp\ (\mu_0 t)$$

Where F is the volumetric flow rate (L h^{-1}), μ is the specific growth rate set-point (h^{-1}), V_0 is the initial volume of the fermentation medium (L), X_0 is the initial cell concentration (g L^{-1}), S_0 is feed substrate concentration (g L^{-1}), S is the substrate concentration in the culture broth (g L^{-1}) and $Y_{X/S}$ is the overall biomass-to-substrate yield (g_x g_s^{-1}).

The control of the μ allows to reach pseudo steady state culture conditions, enhancing bioprocess reproducibility and easing the exhaustive characterization of recombinant strains. Automatic MicoBU-2031 microburettes were used for precisely substrate feeding into the culture broth.

The impact of methanol on the CYP2C9 production was studied by feeding this compound as cosubstrate together with glycerol-based medium described previously. Two different methanol feeding strategies were followed:

- Periodic methanol pulse additions: 3 g L^{-1} methanol pulses were added when methanol was consumed by cells, identified by a sharp increase on pO_2.
- Feed-back control: methanol concentration was determined by a Methanol Sensor System (Raven Biotech Inc., Vancouver, Canada), and controlled by following a predictive-PI control strategy. Furthermore, regular off-line methanol determinations by HPLC were used to validate the control system.

Fed-batch conditions were monitored and controlled at the following set points:

Table 4.3. Culture conditions of the fed-batch cultivations conducted throughout the thesis

	Fed-batch conditions		
	Crl1	**CalB**	**CYP2C9**
pH	6.0	5.0	5.0
μ (h^{-1})	0.025-0.095	0.05-0.15	0.10
Temperature (˚C)	25.0	25.0	25.0
Pressure (bar)	1.0	1.0	1.0
pO_2 (%, air saturation)	25	25	25
Stirring rate (rpm)	700-1200	700-1200	700-1200
Inlet gas flow (vvm)	1.0	1.0	1.0

pH was controlled by means of 20% NH_4OH (w/v) addition. Inlet gas flow was controlled by means of mass-flow controllers (Bronkhorst Hi-Tech, Ruurlo, The Netherlands). A condenser was integrated in the exhaust gas to minimize water losses by evaporation. Dissolved oxygen (pO_2) was controlled by a cascade control, involving both stirring rate and air-oxygen mixtures (up to 25% of pure oxygen flow rate with respect to the total gas flow rate), but maintaining 1.0 vvm of total inlet gas.

Analytical methods

Biomass concentration determination

Biomass concentrations were measured as Dry Cell Weight (DCW) as previously described (Cos *et al.*, 2005). 2-mL from fresh samples were filtered through pre-weighted 0.7 µm glass microfiber filters (Whatman GF/F, Maidstone, UK). Then, filters were washed twice with 0.09% NaCl (w/v) and dried at 105˚C until constant weight. Determinations were performed by triplicate and the relative standard deviation (RSD) was about 3%.

Quantification of the carbon source and byproducts concentration

The concentration of feeding carbon sources as well as potential culture byproducts were measured by means of HPLC (Dionex Ultimate 3000) analysis using an anionic exchange column (ICSepICE-COR-EGEL87H3, Transgenomic) as previously described (Jordà *et al.*, 2014). The mobile phase was 6 mM sulfuric acid, and the sample injection volume was 20 µL. Chromeleon software (Dionex) was used for data treatment. Determinations were performed by triplicate and RSD was always less than 1%.

Off-gas analyses

A BlueInOne Cell gas analyzer (BlueSens, Herten, Germany) was used in both chemostat and fed-batch cultivations. The CO_2 and O_2 gas molar fractions were recorded online with provision for off-gas pressure and humidity. Data obtained were used to calculate the oxygen uptake rate (OUR), carbon dioxide production rate (CPR), their corresponding specific rates (q_{O2} and q_{CO2}) and respiratory quotient (RQ). RSD was less than 5% in all cases.

Enzymatic activity assays

Crl1 activity assay

An enzymatic p-nitrophenyl butyrate (pNPB) based assay was selected to determine Crl1 activity by using a procedure described elsewhere (Chang *et al.*, 2006), albeit with slight modifications. Thus, 20 µL of samples supernatant were mixed with 980 µL of reaction buffer, which contained 1 mM pNPB, 50 mM phosphate buffer (pH 7.0) and 4% (v/v) acetone. The absorbance at 348 nm was monitored at 30 °C by using a Specord 200 Plus spectrophotometer from Analytic Jena (Jena, Germany). One activity unit was defined as the amount of enzyme needed to release 1 mmol of p-nitrophenol per minute under assay conditions. RSD was less than 1%.

CalB activity assay

CalB activity was determined using a p-nitrophenyl butyrate (pNPB) based assay, as for Crl1 determinations. Nonetheless, for CalB activity measuring, 100 µL of samples supernatant samples were mixed with 900 µL of reaction buffer, which contained 2 mM pNPB, 300 mM Tris buffer (pH 7) and 4% (v/v) acetone. The absorbance at 405 nm was monitored at 30 °C by using a Specord 200 Plus spectrophotometer from Analytic Jena (Jena, Germany). One activity unit was defined as the amount of enzyme needed to release 1 mmol of p-nitrophenol per minute under assay conditions. RSD was less than 1%.

CYP2C9 activity assay

For bioreactor samples, the enzymatic assays were performed in 50 mL falcon tubes. In this case, the washed cells were diluted to an equivalent OD 5 with 100 mM phosphate buffer pH 7.4. Then, 2 mL of cells were taken and mixed with diclofenac (2 mM final concentration) and reactions were carried at 28 °C and 200 rpm. After 20 h of incubation,

50 µL were taken and mixed with 450 µL of MetOH/acetonitrile (1:1), to stop the reaction and to extract both the substrate (diclofenac) and the product (4'OH-diclofenac). Then, samples were centrifuged at 12000 x g for 2 min to remove cells and supernatants were filtered (0.45 µm) and analyzed by HPLC-UV, to quantify both substrate and product concentration. The product-related results of the fermentations are presented in terms of activity units (AU). One AU is defined as the amount of enzyme which catalyzes the conversion of 1 µmol of substrate (diclofenac) along the reaction time (20 h) under the assay conditions. Thus, this parameter provides an enzyme titer value in terms of (AU mL^{-1}). Since the product is anchored to the cell membrane, it is initially referred also to the amount of biomass (AU OD_{600}^{-1}).

For whole cell screening samples, CYP2C9 activity measuring were carried out in 96-well plates. Two different methods were used:

- Diclofenac-based activity assay: 4 µL of a 100 mM diclofenac stock solution in methanol (MetOH) were added to each well and conversions were performed for 20 h at 28°C and 320 rpm. Reactions were stopped by addition of 200 µL of MetOH/acetonitrile (1:1) and vigorously mixed. After centrifugation (10 min, 3200xg) 200 µl of supernatant were transferred into a fresh 96 well microtiter plate (polypropylene, V-shaped), sealed and analyzed by HPLC-MS. One activity unit was defined as the amount of enzyme needed to oxidize 1 µmol of diclofenac per minute under assay conditions.
- 7-methoxy-4-(trifluoromethyl)-2H-chromen-2-one (MFC)-based assay: 95 µL of the cell suspension diluted in assay solution (100 mM phosphate buffer pH 7.4) were mixed with 5 µL of 1 mM MFC (final concentration 50 µM) in a black microtiter plate. Fluorescent metabolite formation was quantified measuring the fluorescence of the product 7-hydroxy-4-(trifluoromethyl)-2H-chromen-2-one

(HFC). The fluorescence signal of the product formed (ex 410 nm / em 538 nm) was recorded every minute for a period of 1 h using a SynergyMX plate reader (BioTek Instruments Inc, USA). During the reaction cells were kept at 30°C. One activity unit was defined as the amount of enzyme needed to oxidize 1 µmol of MFC per minute under assay conditions.

Process parameters determination

Chemostat

For the calculation of the main macrokinetic parameters in the steady state, equations derived from mass balances were used, as describe elsewhere (Garcia-Ortega *et al.*, 2016):

$$\begin{bmatrix} \mu \\ q_s \\ q_p \\ q_{o_2} \\ q_{co_2} \end{bmatrix} XV = \begin{bmatrix} F_{out}X \\ -FS_0 + F_{out}S \\ F_{out}P \\ OUR\,V \\ CPR\,V \end{bmatrix} \quad (1)$$

Where μ is the specific growth rate (h^{-1}); q_s, specific substrate uptake rate (g g_x^{-1} h^{-1}); q_p, specific production rate (AU g_x^{-1} h^{-1}); q_{O_2}, specific oxygen uptake rate (mol O_2 g_x^{-1} h^{-1}); q_{CO_2}, specific carbon dioxide production rate (mol CO_2 g_x^{-1} h^{-1}); F, substrate feeding rate (L h^{-1}); F_{out}, outlet flow rate (L h^{-1}); X, biomass concentration in the broth (g L^{-1}); V, volume of broth in the reactor (L); S_0, substrate feeding concentration (g L^{-1}); S, broth substrate concentration (g L^{-1}); OUR, oxygen uptake rate (mol O_2 L^{-1} h^{-1}); CPR, carbon dioxide production rate (mol CO_2 L^{-1} h^{-1}), considered equal to CER.

Fed-batch

Likewise, mass balance equations concerning fed-batch mode can be formulated:

$$\frac{d}{dt}\begin{bmatrix} XV \\ SV \\ PV \\ O_2V \\ CO_2V \end{bmatrix} = \begin{bmatrix} \mu \\ q_s \\ q_p \\ q_{o_2} \\ q_{co_2} \end{bmatrix} XV + \begin{bmatrix} -F_{out}X \\ FS_0 - F_{out}S \\ -F_{out}P \\ OTR\,V - F_{out}O_2 \\ -CTR\,V - F_{out}CO_2 \end{bmatrix} \quad (2)$$

Where μ is the specific growth rate (h^{-1}); q_s, specific substrate uptake rate (g g_x^{-1} h^{-1}); q_p, specific production rate (AU g_x^{-1} h^{-1}); q_{O_2}, specific oxygen uptake rate (mol O_2 g_x^{-1} h^{-1}); q_{CO_2}, specific carbon dioxide production rate (mol CO_2 g_x^{-1} h^{-1}); F, substrate feeding rate (L h^{-1}); F_{out}, outlet flow rate (L h^{-1}); X, biomass concentration in the broth (g L^{-1}); V, volume of broth in the reactor (L); P, total enzyme activity (AU); S_0, substrate feeding concentration (g L^{-1}); S, broth substrate concentration (g L^{-1}); OTR, oxygen transfer rate (mol O_2 L^{-1} h^{-1}); CTR, carbon dioxide transfer rate (mol CO_2 L^{-1} h^{-1}). OTR and CTR are considered approximately equal to OUR and CER, respectively.

By applying the mass balances in fed-batch mode enables to calculate the mean of the main macrokinetic parameters, as previously described (Garcia-Ortega *et al.*, 2013):

$$\int_{(XV)_0}^{(XV)} d(XV) = \mu_{mean} \int_{t_0}^{t} XV\, dt \quad (3)$$

$$\int_{(PV)_0}^{(PV)} d(PV) = q_{p_{mean}} \int_{t_0}^{t} XV\, dt \quad (4)$$

$$S_{feed} \int_0^t F_{feed} dt - \int_{(SV)_0}^{(SV)} d(SV) = q_{s_{mean}} \int_{t_0}^{t} XV\, dt \quad (5)$$

$$q_{O_2\,mean} \int_{t_0}^{t} (XV)\, dt = \int_{t_0}^{t} (OUR \cdot V)\, dt \quad (6)$$

$$q_{CO_2 mean} \int_{t_0}^{t} (XV)\, dt = \int_{t_0}^{t} (CER \cdot V)\, dt \quad (7)$$

Where μ_{mean} is the mean specific growth rate (h^{-1}); q_{smean}, mean specific substrate uptake rate (g g_x^{-1} h^{-1}); q_{pmean}, mean specific production rate (AU g_x^{-1} h^{-1}); q_{O2mean}, mean specific oxygen uptake rate (mol O_2 g_x^{-1} h^{-1}); $q_{CO2mean}$, mean specific carbon dioxide production rate (mol CO_2 g_x^{-1} h^{-1}); F_{feed}, substrate feeding rate (L h^{-1}); X, biomass concentration in the broth (g L^{-1}); P, total enzyme activity (AU); V, volume of broth in the reactor (L); S_{feed}, substrate feeding concentration (g L^{-1}); S, broth substrate concentration (g L^{-1}); OTR, oxygen transfer rate (mol O_2 L^{-1} h^{-1}); CTR, carbon dioxide transfer rate (mol CO_2 L^{-1} h^{-1}). OTR and CTR are considered approximately equal to OUR and CER, respectively.

Consistency check and data reconciliation

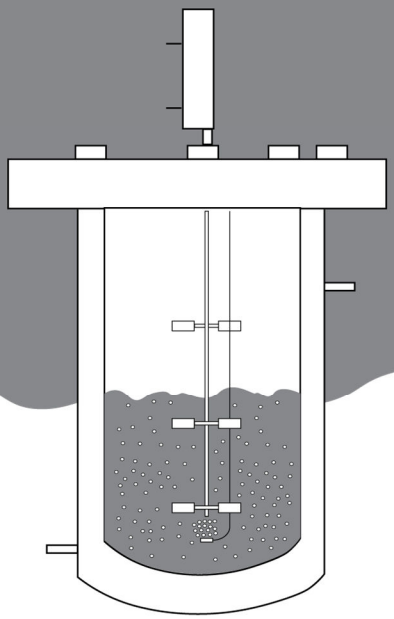
Carbon and redox balances were used as constraints to perform consistency check of the macrokinetic parameters obtained, except q_p, as detailed previously (Ponte *et al.*, 2016). Consistency check could be summarized in the following equation:

$$h = \varepsilon P^{-1} \varepsilon$$

where h is the statistical parameter that determines if measurements contains significant errors. If h exceeds the threshold value of χ^2-test with a confidence of 95%, it means that the parameter contains significant error, so it cannot be considered in balance equations. Moreover, a standard parameter reconciliation procedure (van der Heijden *et al.*, 1994) was carried out taking into consideration h values and the fulfilment of mass balance equations.

RESULTS I

Specific growth rate governs *AOX1* gene expression, affecting the production kinetics of *Pichia pastoris* (*Komagataella phaffii*) P_{AOX1}-driven recombinant producer strains with different target gene dosage



5

Chapter published as a research article in Microbial Cell Factories Journal

Garrigós-Martínez, J., Nieto-Taype, M.A., Gasset-Franch, A., Montesinos-Seguí, JL., Garcia-Ortega, X., Valero, F. Specific growth rate governs *AOX1* gene expression, affecting the production kinetics of *Pichia pastoris* (*Komagataella phaffii*) P_{AOX1}-driven recombinant producer strains with different target gene dosage.

Microb Cell Fact 18, 187 (2019). https://doi.org/10.1186/s12934-019-1240-8

Results I content

Background

In the last two decades, *Komagataella phaffii*, which was formerly known as *Pichia pastoris*, has emerged as a promising host for recombinant protein production (RPP) (Cereghino *et al.*, 2002; Vogl *et al.*, 2013; Ahmad *et al.*, 2014; Looser *et al.*, 2014; Juturu and Wu, 2018; Yang and Zhang, 2018). Also, it has lately been increasingly used for metabolite production. Mattanovich *et al.* have summarized the main uses of *P. pastoris* for metabolite production (Peña *et al.*, 2018). The potential of *P. pastoris* for hosting the production of recombinant proteins is increased by its ability to grow at high cell densities (ca. 100 g L^{-1} dry cell weight) on defined media, the availability of strong protein expression systems, the possibility to secrete the target proteins to the extracellular medium, its enabling eukaryotic post-translational modifications (Cereghino and Cregg, 2000; Daly *et al.*, 2005) and a reference genome sequence (Sturmberger *et al.*, 2016b).

The alcohol oxidase 1 promoter (P_{AOX1}) expression system has been widely used for RPP on *P. pastoris*. In terms of regulation, P_{AOX1} is strongly inducible by methanol and repressible by both glucose and glycerol. Its tight regulation allows bioprocess decoupling into a first phase of biomass generation and a second phase of where heterologous gene expression is induced by the addition of methanol. Properly designing the induction phase is crucial to obtain acceptable amounts of recombinant protein (Barrigón *et al.*, 2013; Vogl and Glieder, 2013; Ahmad *et al.*, 2014; Looser *et al.*, 2014). P_{AOX1} typically allows large amounts of proteins to be obtained (Barrigón *et al.*, 2015; Ponte *et al.*, 2016, 2018; Yang and Zhang, 2018); however, the need to use methanol leads to some drawbacks related to plant safety, high oxygen consumption and also high heat production (Heyland *et al.*, 2010; Çalık *et al.*, 2015).

In the literature, recent relevant advances in P_{AOX1} regulation can be found (Yang and Zhang, 2018). Thus, promoter sequence analysis has allowed several binding sites for transcription factors (TFs) to be identified. Most such TF were previously known and have been related to stress response, glucose repression and oxygen consumption (Hartner *et al.*, 2008). Three of them (Mig1, Mig2 and Nrg1) have emerged as strong repressors of genes involved in methanol uptake (Wang *et al.*, 2017), whereas three others (Mxr1, Mit1 and Prm1) have proved crucial triggers of MUT genes expression (Lin-Cereghino *et al.*, 2006; X. Wang *et al.*, 2016; Zhan *et al.*, 2017). The increasing information gathered about MUT gene expression has allowed some researchers to develop methanol-free expression systems based on MUT machinery (Shen *et al.*, 2016; Wang *et al.*, 2017; Vogl, Sturmberger, *et al.*, 2018). Such systems do not need methanol to trigger MUT genes because their TF genes have been derepressed by genetic engineering.

Some researchers have focused on the relationship between heterologous gene dosage and protein production rate. As previously reported, in P_{AOX1}-driven expression systems, gene dosage and protein production are usually positively correlated, albeit with a relatively small number of copies (2 or 3) only (Sha, X.-W. Yu, *et al.*, 2013; Shu *et al.*, 2016; Cámara *et al.*, 2017a; Yu *et al.*, 2017). However, producer clones integrating high gene of interest (GOI) expression cassettes are often subject to folding and secretion restrictions that result in oxidative stress in the endoplasmic reticulum, thereby having a direct impact on protein production. Also, producer clones containing large numbers of copies have been found to possess a limited transcription efficiency (Sha, X.-W. Yu, *et al.*, 2013; Cámara *et al.*, 2017a). According to Cámara *et al.* (Cámara *et al.*, 2017a), the main limitation in strains with a large number of GOI copies occurs at the transcriptional level rather than in folding or secretion processes. Interestingly, both P_{AOX1}-driven *Rhizopus oryzae* lipase (*ROL*) gene and MUT genes (*AOX1* included) have been found to

be downregulated in clones with many GOI copies, a limitation that results in decreased Rol production and methanol accumulation in chemostat cultivations.

Furthermore, specific growth rate (μ) has been confirmed as a key parameter that affects the specific protein production rate (q_p). To date, many attempts to correlate both parameters have been successfully made. Thus, a positive relationship between them was observed when producing different proteins under the P_{GAP} (Garcia-Ortega *et al.*, 2013, 2016; Rebnegger *et al.*, 2014) and P_{AOX1} control (Jungo *et al.*, 2006). As the *P. pastoris* endogenous genes controlled by these promoters play crucial roles in glycolysis and methanol metabolism, respectively, the protein production driven by these expression systems are coupled to cell growth. By contrast, other authors point out the presence of a maximum in the q_p-μ curve. Thus, Prielhofer et.al (Prielhofer *et al.*, 2018), observed a bell-shaped relationship between q_p and μ when expressing single domain antibody-like molecules of human origin (i-bodies) under the control of an improved glucose-repressible P_{GTH1} promoter. These results led them to devise an optimized bioprocess strategy in which μ was drastically reduced at the end of their -batch experiments. Canales *et al.* studied the effect of glycerol:methanol mixtures in the chemostat feeding stream and the specific growth rate on Rol production under P_{AOX1} promoter (Canales *et al.*, 2015a). They found μ to be much more influential on q_p than was the methanol fraction in the feeding.

In this work, the integrated effect of μ and gene dosage on *AOX1* gene regulation and production kinetics of *Candida rugosa* lipase 1 (Crl1) driven by P_{AOX1} in *P. pastoris* was studied for designing a rational approach to optimize the operating conditions. For this purpose, a single-copy clone (SCC) and a multi-copy clone (MCC) were cultivated under chemostat conditions to establish the relationship between μ, *CRL1* relative transcript levels (RTL) and q_p. This correlation has allowed determining the operational strategy

that maximizes Crl1 production for each clone according to the production kinetic profile. Additionally, transcriptional analyses of two key genes involved in methanol metabolism —*AOX1* and *MIT1*— were used in order to establish whether this pathway might be limited under specific conditions. Finally, fed-batch cultivations were used to confirm the q_p–μ profile pattern observed with chemostat cultivations to validate this experimental platform for the standard industrial operation mode used in *P. pastoris* cell factory.

Results and discussion

Effect of increasing *CRL1* gene dosage on culture physiological state

Increasing the dosage of heterologous genes is known to affect homeostasis in *P. pastoris* cultivations through restrictions in protein processing (Puxbaum *et al.*, 2015; Ben Azoun *et al.*, 2017). Also, P_{AOX1}-driven expression systems have been found to exhibit attenuated MUT gene expression (Cámara *et al.*, 2017a), thereby affecting the methanol uptake rate (q_s) of producer strains and potentially reducing their ability to grow (Zhu *et al.*, 2011; Ben Azoun *et al.*, 2017; Yu *et al.*, 2017).

Figure 5.1A shows the variation of the specific substrate uptake rate (q_s) and overall biomass-to-substrate yield ($Y_{X/S}$*) over a wide range of dilution rates (D) (0.020–0.095 h^{-1}) in chemostat cultivations of SCC and MCC. No methanol accumulation was observed under any conditions, but no D values above 0.095 h^{-1} were used in order to avoid cell washout. In addition, the carbon and electron balances were checked and deviations prior to data reconciliation found to be less than 5% in all cases. With both clones, q_s increased linearly across the D range, and q_s values at equivalent D values were rather similar for both clones. As a result, intrinsic substrate-to-biomass yield ($Y_{S/X}$), and their respective maintenance coefficients (m_s), were very similar (Table 5.1). Interestingly, both clones

had mean $Y_{S/X}$ values around 2.2 $g_{MetOH}\ g_X^{-1}$. This value is similar to the yield for the wild-type strain (Tomàs-Gamisans *et al.*, 2019) and a slightly lower than reported for an important number of recombinant protein producer strains, which $Y_{S/X}$ ranges 2-3 $g_{MetOH}\ g_X^{-1}$. However, for the recombinant production of other target proteins $Y_{S/X}$ can reach higher values (Looser *et al.*, 2014). For instance, $Y_{S/X}$ reached in the production of Rol under the same expression system was two-fold higher than those obtained in the present work (Ponte *et al.*, 2016). About the value of m_s reached for both clones, no significant statistical differences were observed. Results fell within the range of values reported in the literature, from 0.007 to 0.042 $g_{MetOH}\ g_X^{-1}\ h^{-1}$ (Looser *et al.*, 2014).

Other factors affecting the physiological state of the cell factory are related to respiration parameters, such as the specific O_2 uptake rate (q_{O2}), the specific CO_2 production rate (q_{CO2}), the corresponding intrinsic yields ($Y_{i/X}$) and maintenance coefficients ($m_{i/X}$), and the respiratory quotient (RQ). As can be seen from Figure 5.1B and Table 5.1, these factors exhibited identical trends and similar values across the *D* range regardless of the gene dosage.

$Y_{O2/X}$ values obtained in this work —0.07 $mol_{O2}\ g_X^{-1}$— were similar to those obtained in the recombinant production of the human serum albumin (Liang and Yuan, 2006), antibody (Potgieter *et al.*, 2010) and human chymotrypsinogen B (Curvers *et al.*, 2001) between 0.08 and 0.09 $mol_{O2}\ g_X^{-1}$. However, highest values around 0.15 $mol_{O2}\ g_X^{-1}$ were obtained for *Rhizopus oryzae* lipase production (Ponte *et al.*, 2016). The higher values of $Y_{O2/X}$ are directly related with the lower $Y_{X/S}$* obtained. The results suggest that the values of the different yields are a function of the target protein expressed.

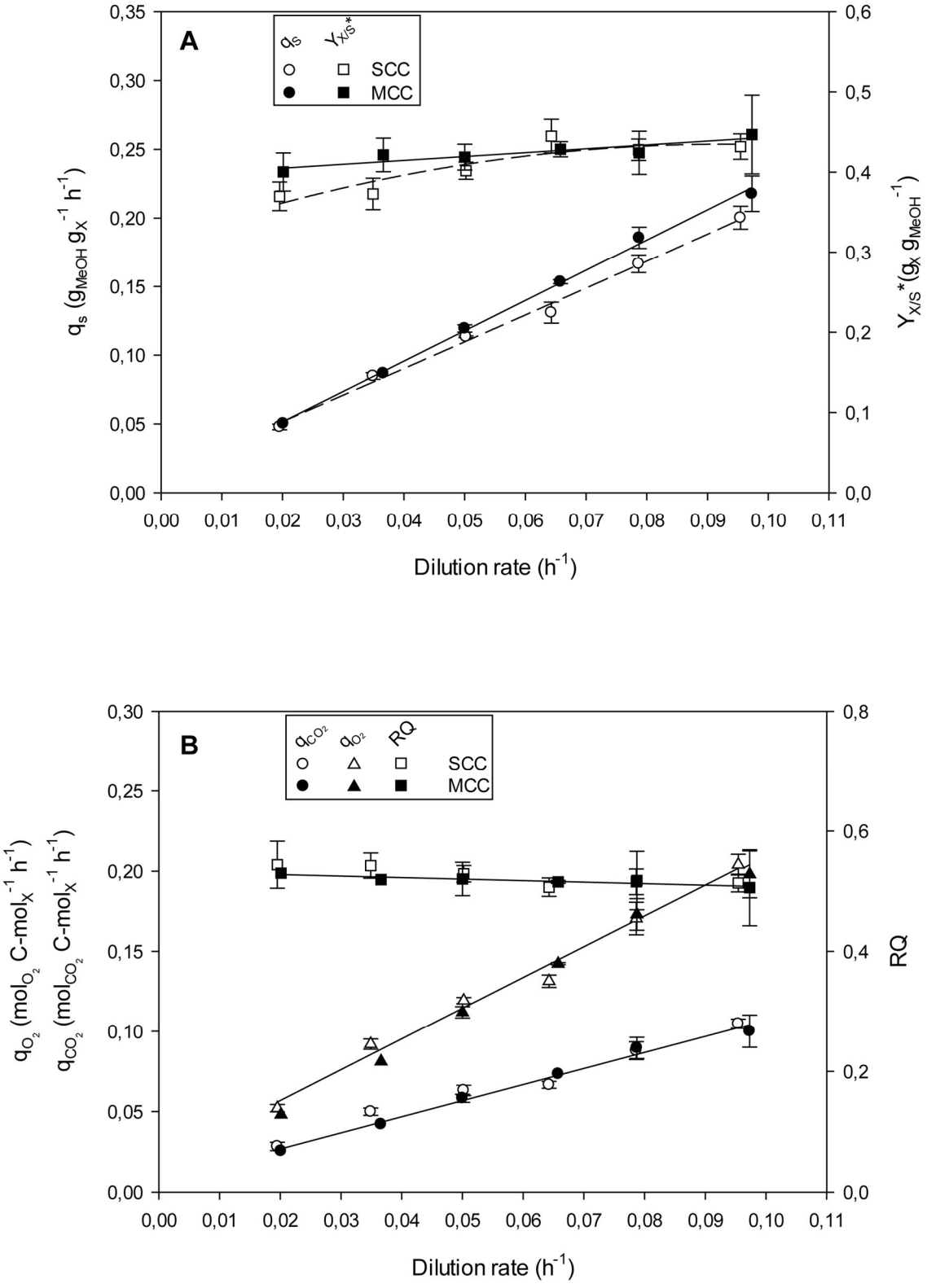


Figure 5.1. *Pichia pastoris* physiological response to an increase in *CRL1* gene dosage in chemostat cultivations. **A**) Specific methanol consumption rate (q_s), overall biomass-to-substrate yield ($Y_{X/S}$*). **B**) Specific oxygen uptake rate (q_{O2}), specific carbon dioxide production rate (q_{CO2}) and respiratory quotient (RQ). Error bars represent the standard deviation of two biological replicates.

Table 5.1. Intrinsic yields ($Y_{i/X}$) and maintenance coefficients ($m_{i/X}$) for biomass growth obtained from chemostat cultivations.

	Single-copy clone	**Multi-copy clone**
$Y_{S/X}$ ($g_{MeOH}\ g_X^{-1}$)	2.16 ± 0.08	2.21 ± 0.05
m_S ($g_{MeOH}\ g_X^{-1}\ h^{-1}$)	0.014 ± 0.005	0.007 ± 0.004
$Y_{O2/X}$ ($mol_{O2}\ C\text{-}mol_X^{-1}$)	1.92 ± 0.11	1.99 ± 0.07
m_{O2} ($mol_{O2}\ C\text{-}mol_X^{-1}\ h^{-1}$)	0.018 ± 0.007	0.010 ± 0.005
$Y_{CO2/X}$ ($mol_{CO2}\ C\text{-}mol_X^{-1}$)	0.95 ± 0.07	1.00 ± 0.05
m_{CO2} ($mol_{CO2}\ C\text{-}mol_X^{-1}\ h^{-1}$)	0.012 ± 0.005	0.007 ± 0.003

± Indicates standard error (SE) from regression analysis.

The fact that the major macrokinetic parameters related to the physiological status were very similar for both clones suggests that increasing the *CRL1* cassette dosage from one to three in *P. pastoris* genome had no effect on the physiological performance of the yeast in chemostat cultivations. On the other hand, a marked influence of the operational mode on such parameters was observed in fed-batch (FB) cultivations of both clones. As can be seen in Figures 5.2A and 5.2B, the evolution of q_S, q_{CO2} and q_{O2} across the range of μ tested were different for both clones. These rates followed a nearly linear trend for SCC, similar to the behaviour observed on chemostat cultivations. In contrast, a saturated curve trend was obtained for MCC, thus indicating that the higher MCC metabolic burden caused by an increase heterologous gene dosage affects the clone capabilities for both methanol and O_2 consumption, as well as CO_2 production in this operational mode. Consequently, accumulation of methanol on MCC in fed-batch cultivation was observed at the highest μ tested (0.08 h^{-1}). Thus, maximum μ tested for MCC was decreased to

0.065 h^{-1} in order to maintain carbon-limiting conditions. Although significant differences in q_{CO2} and q_{O2} were observed, RQ was quite similar in both operational modes irrespective of μ.

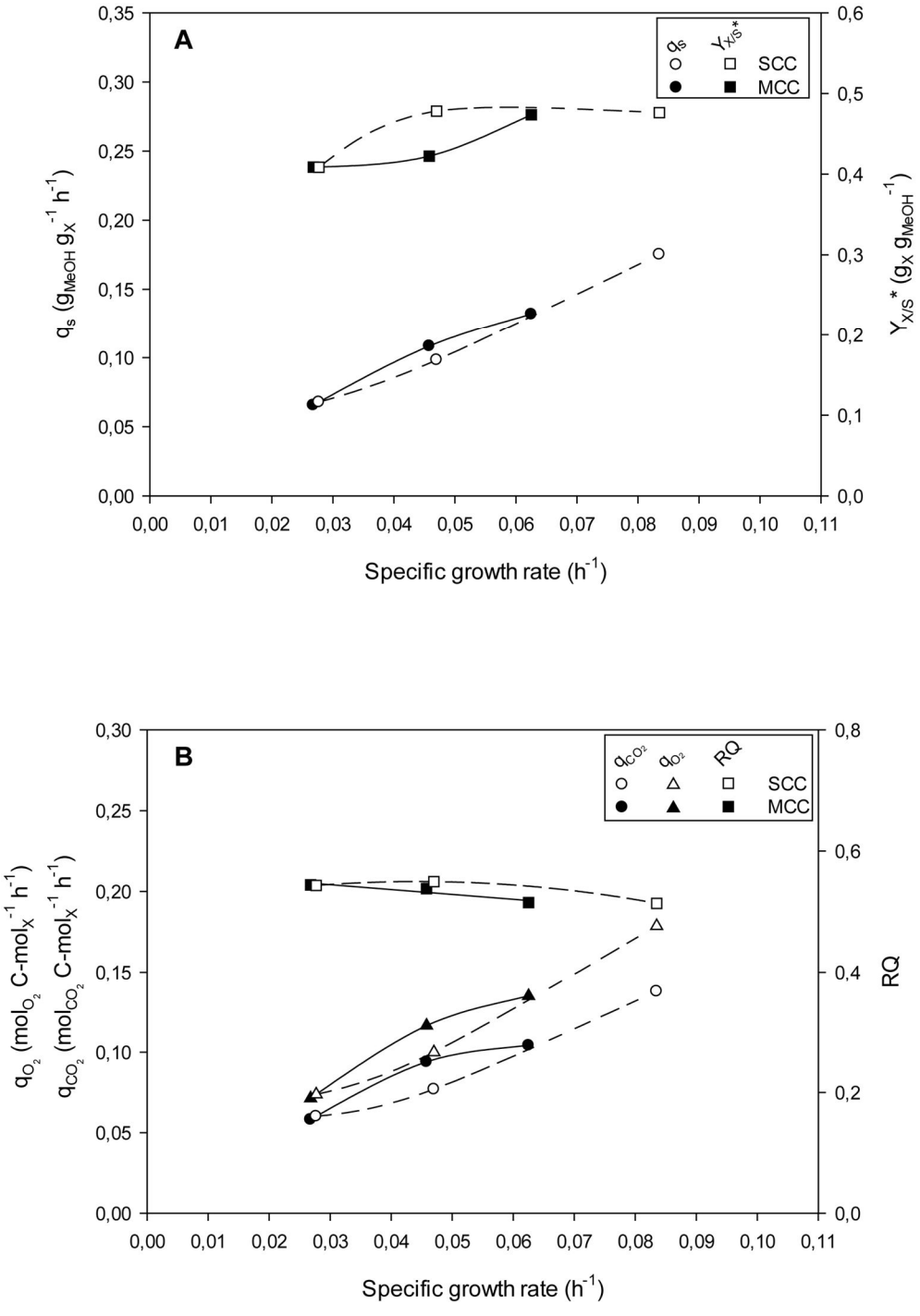


Figure 5.2. *Pichia pastoris* physiological response to an increase in *CRL1* gene dosage in fed-batch (FB) cultivations. **A**) Specific methanol consumption rate (q_s), overall biomass-to-substrate yield ($Y_{X/S}$*). **B**) Specific oxygen uptake rate (q_{O2}), specific carbon dioxide production rate (q_{CO2}) and respiratory quotient (RQ).

Relationship between Mit1 limitation and a decreased *AOX1* relative expression (RE)

The induction of strains with multiple copies of a P_{AOX1}-driven heterologous gene with methanol has been reported to result in transcriptional limitation of MUT genes (Cámara *et al.*, 2017a). Therefore, transcriptional analysis of key genes such as *AOX1*, *CRL1* and the methanol-induced transcription factor 1 (*MIT1)* genes were performed in order to examine their impact on the first step of methanol metabolism (see Figures 5.3A and 5.3B).

As can be seen from Figure 5.3A, and consistent with previous results (Zhu *et al.*, 2011), *AOX1* gene expression was in average 2-fold higher in SCC than it was in MCC, whichever the dilution rate. It indicates that the resources needed to trigger transcription of P_{AOX1}-driven genes may be shared among them —heterologous gene cassettes and the endogenous *AOX1* gene included. These analyses were only carried out in chemostat cultivations, since this operational mode works steady-state conditions, thus allows a to increase to robustness of the results.

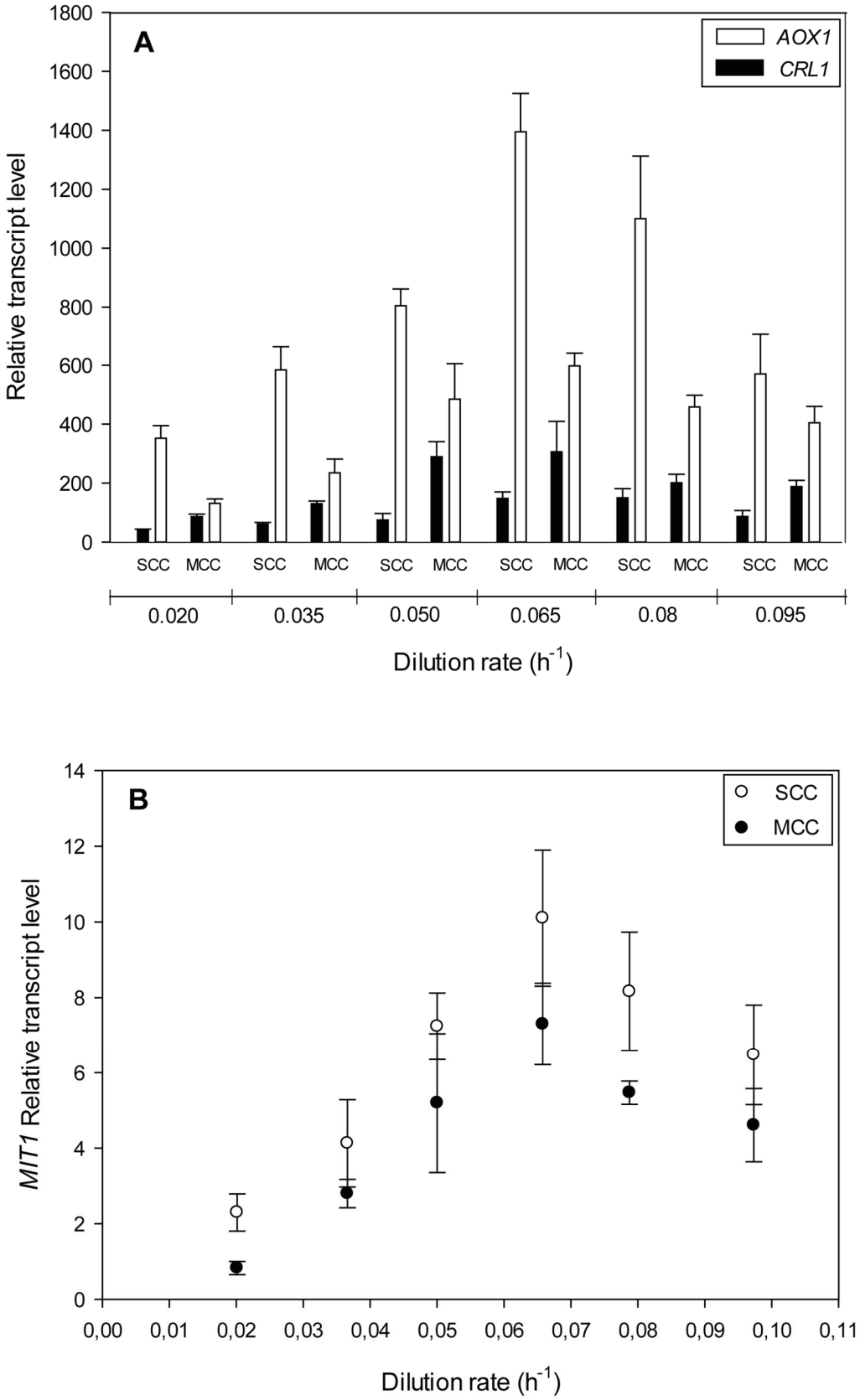


Figure 5.3. Influence of the dilution rate on relative gene transcription levels (RTLs) in chemostat cultivations. **A**) Genes *AOX1* and *CRL1*. **B)** Gene *MIT1*. *MTH1* was used as housekeeping gene for the analysis. Error bars represent the standard deviation of two biological replicates.

This phenomenon was further studied by analysing transcriptional levels of *MIT1*, a crucial TF for P_{AOX1} induction (X. Wang *et al.*, 2016), in both clones. As can be seen from Figure 5.3B, *MIT1* relative transcript levels (RTL) were not significantly different comparing both clones across the *D* range tested, except for two dilution rates (0.02 h^{-1} and 0.08 h^{-1}). Therefore, as expected, increasing heterologous gene dosage has not led to a proportional increase in *MIT1* transcription rate. One should therefore hypothesize that the Mit1 pool is a limited resource, all the genes whose expression depends on P_{AOX1} promoter would compete with one another for this TF —and hence for being transcribed. Consequently, the *AOX1* gene was less strongly expressed in MCC than it was in SCC owing to competition with three *CRL1* copies for the equivalent Mit1 resources. This hypothesis of Mit1 limitation is reinforced by the work of Cámara *et al.* where the overexpression of Mit1 is enough to reverse the transcriptional limitation derived from increasing heterologous gene dosage (Cámara *et al.*, 2019). Moreover, deregulating the expression of some MUT-related TFs increased protein production driven by the P_{AOX1} expression system even in absence of methanol (Wang *et al.*, 2017; Shi *et al.*, 2018; Vogl, Sturmberger, *et al.*, 2018; Cámara *et al.*, 2019).

Although these insights were previously obtained from a heterologous gene dosage comparison, no similar studies had examined a potential correlation of MUT-related genes RTL with the μ. A positive proportional relationship between μ, *MIT1* RTL and P_{AOX1}-driven transcription rate should be expected since the more methanol was fed to the culture, the greater was the amount of *AOX1* enzyme needed to consume it. However, as can be seen in Figures 5.3A and 5.3B, the correlation of the MUT-related genes RTL with *D* was bell-shaped for both clones, which suggests a close relationship between P_{AOX1}-driven genes expression and *MIT1*. Further research at transcriptional level would be

needed to elucidate why *P. pastoris* decreases its methanol consumption resources when it approaches its maximum specific growth rate, μ_{max}.

Overall, the previous results show that the *MIT1* RTL is governed by μ, affecting the transcription rate of P_{AOX1}-driven genes. One could hypothesize that a similar phenomenon could also take place for to other MUT-related TFs such as Mxr1 and Prm1, since the regulation of their expression in presence of methanol must be coupled (X. Wang *et al.*, 2016). Although *AOX1* RTLs were low in MCC relative to SCC, *AOX1* was expressed strongly enough to produce the minimum amount needed to catabolize all methanol fed to chemostat cultivations. However, as noted in the previous section, growing FB cultivations of MCC at the maximum μ level reached by SCC (0.08 h^{-1}) led to methanol accumulation during early stages of the feeding phase.

Influence of the operational mode on production-related parameters

The primary aim of this work was to elucidate the Crl1 production kinetics for both clones in chemostat and FB cultivations. Furthermore, transcriptional analysis provided valuable information, which also should be related with both the growth and the recombinant protein production. Since Crl1 production in this cell factory is governed by the P_{AOX1} promoter, it was expected to be coupled to growth because the sole carbon source used was methanol (Jungo *et al.*, 2006).

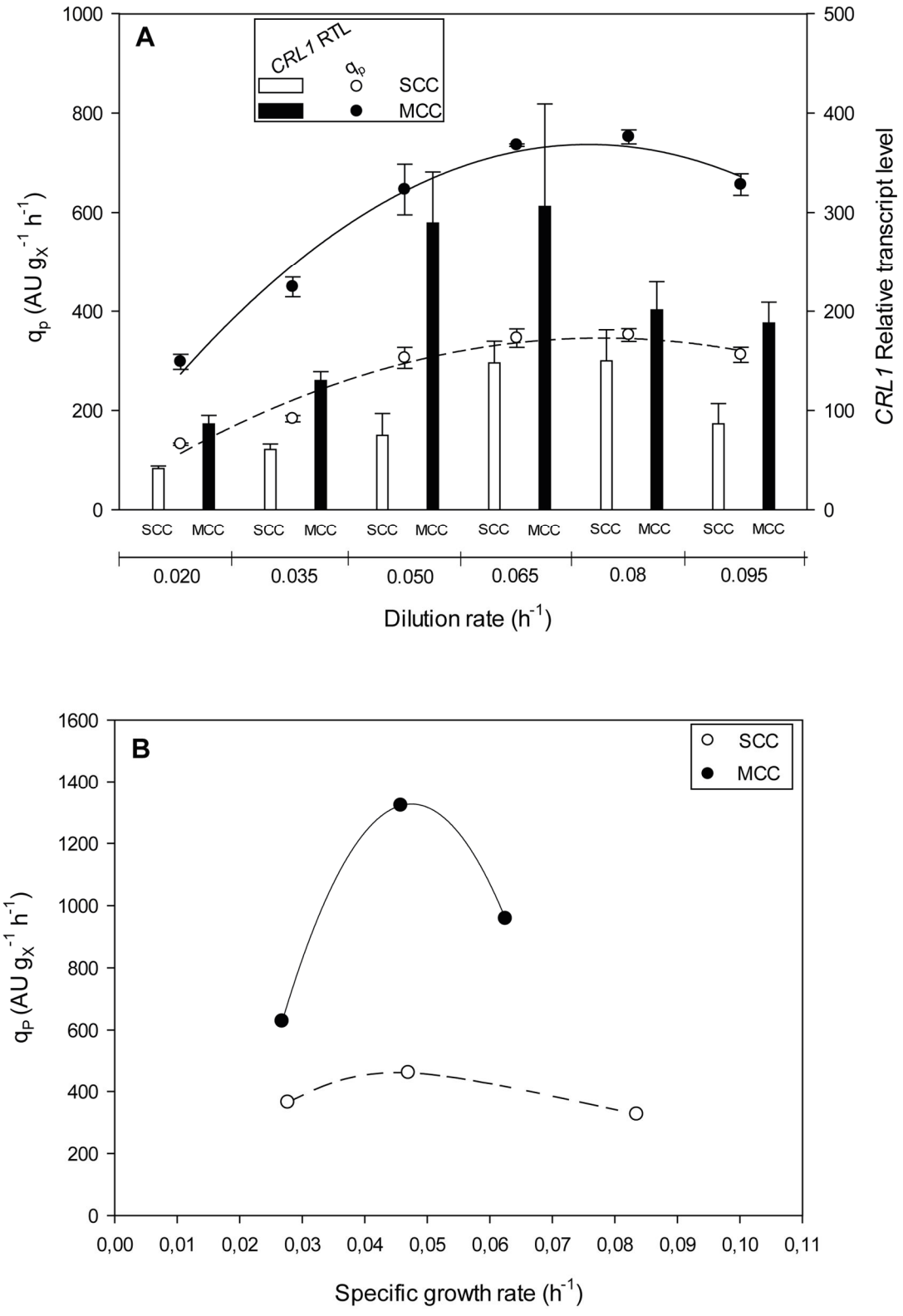


Figure 5.4. Comparison of SCC and MCC Crl1 production kinetics and its relationship with *CRL1* relative transcript levels. The specific Crl1 production rate (q_p) was calculated for chemostat (**A**) and fed-batch cultivations (**B**). *CRL1* transcriptional analyses were also done on chemostat cultivations (**A**). *MTH1* was used as housekeeping gene for *CRL1* RTL calculations. Error bars represent the standard deviation of two biological replicates.

However, in contrast to the P_{GAP}-based examples mentioned in the introduction of the chapter, q_p was not linearly related to D or μ in either operational mode. Rather, both chemostat and FB cultivations exhibited a bell-shaped trend in both clones, production being optimal at D = 0.08 h^{-1} in chemostat cultivations (Figure 5.4A) and μ = 0.045 h^{-1} in FB cultivations (Figure 5.4B). In FB cultivations, MCC presented a more pronounced bell-shape curve than did SCC. Therefore, MCC would require a more precise control of μ because a slight deviation from the optimal setpoint would result in a marked decrease of q_p. Consequently, the optimum differences in μ should be considered in designing bioprocesses for recombinant protein production.

The overall product-to-biomass and product-to-substrate yield ($Y_{P/X}$ and $Y_{P/S}$, respectively) exhibited a linear decreasing trend in chemostat cultivations of both clones (Figure 5.5A). Although SCC behaved identically in FB cultivations, the MCC exhibited a maximum value at an intermediate μ level (0.045 h^{-1}; Figure 5.5B). These yields are important inasmuch as they are closely related to product titer (Figure 5.6), which is a parameter susceptible to be optimized in industry due to its influences on downstream processing costs. Therefore, MCC would be the strain of choice for optimum Crl1-related yields and titer when cultivated at intermediate μ values. Irrespective of gene dosage, q_p and product-related yields were more than twice greater in FB cultivations than they were in chemostat cultivations.

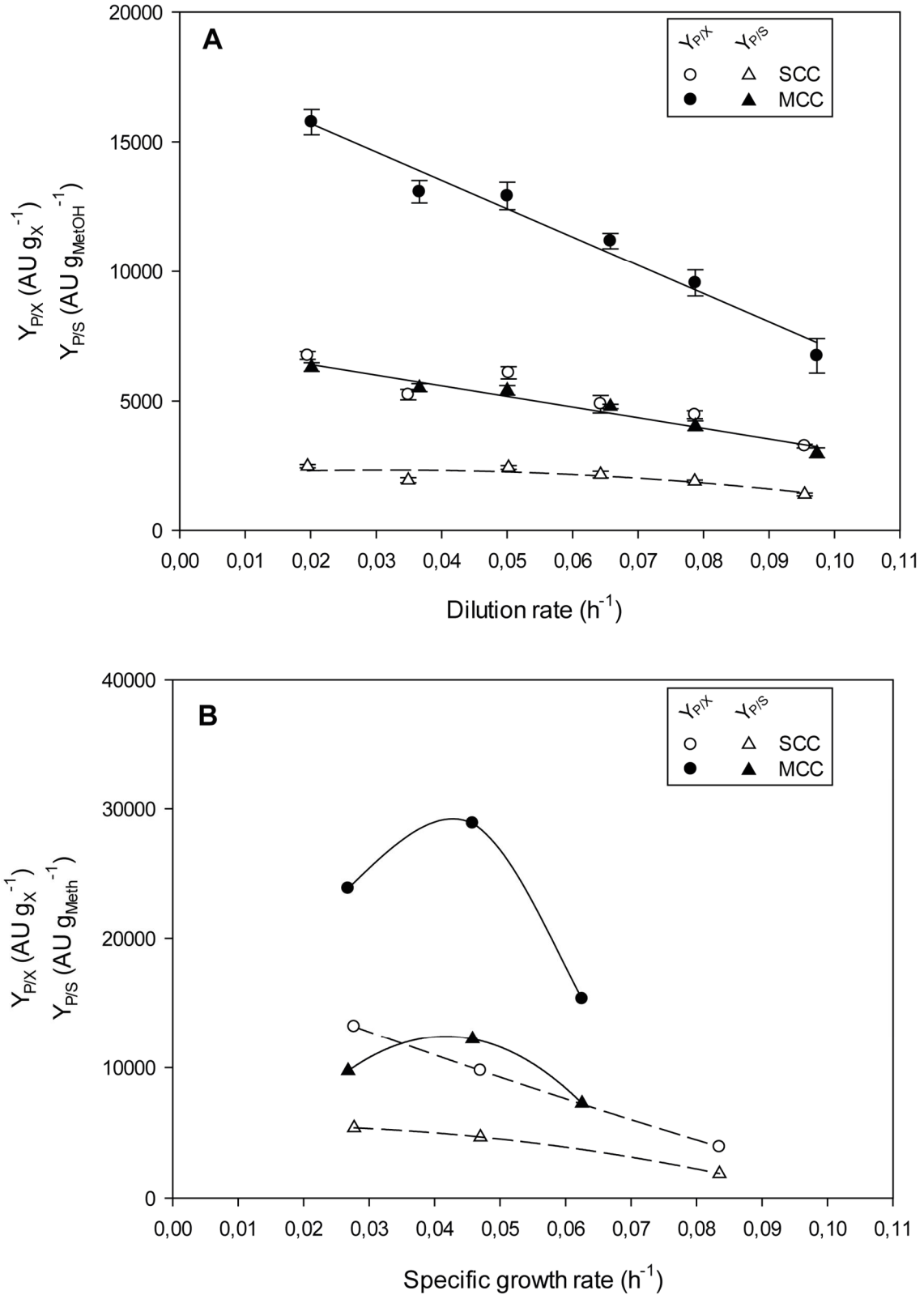


Figure 5.5. Comparison of SCC and MCC for Crl1-related yields. Overall product-to-biomass yield ($Y_{P/X}$) and product-to-substrate yield ($Y_{P/S}$). **A**) Chemostat cultivations. **B**) Fed-batch cultivations. In Figure A, error bars represent the standard deviation of two biological replicates.

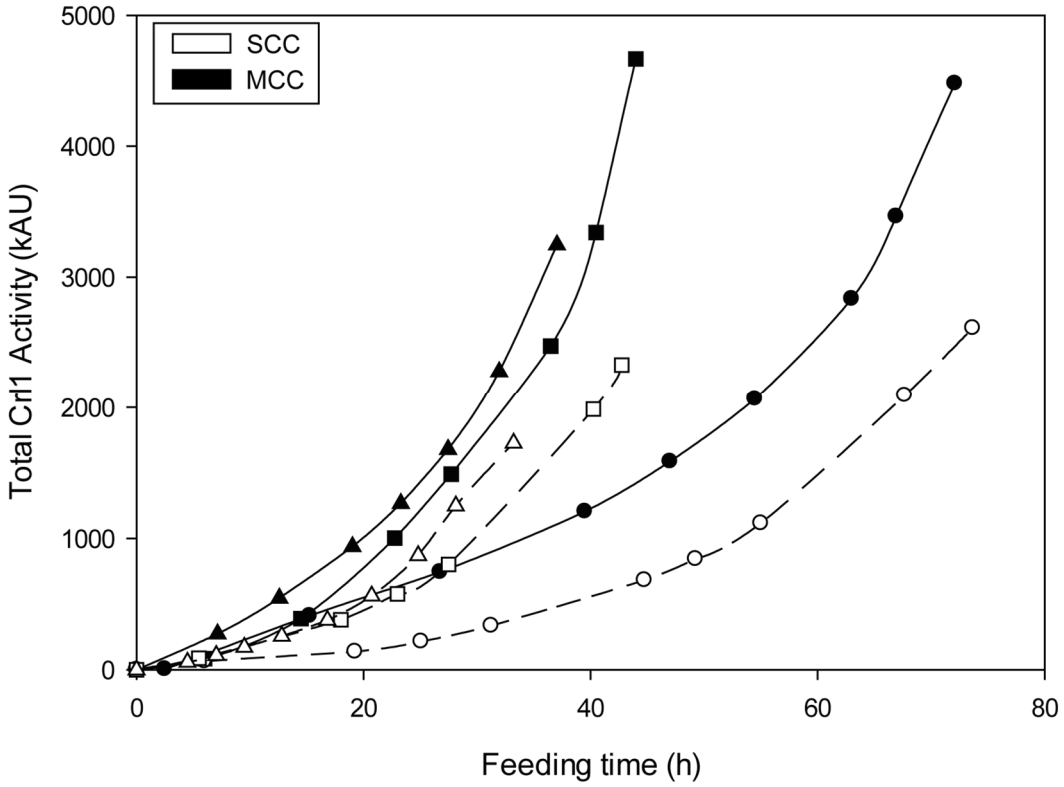


Figure 5.6. Crl1 production time evolution expressed as total activity units in fed-batch cultivations at different μ: (●,○), 0.028 h^{-1}; (■,□), 0.047 h^{-1}; (▲), 0.063 h^{-1}; (△), 0.084 h^{-1}.

Increasing *CRL1* gene dosage boosts protein production

Although a transcriptional limitation had been proved in at least two MUT-related genes, it was necessary to quantify to what extent the GOI transcription rate was affected by an increase in gene dosage, and hence how it influenced q_p.

In chemostat cultivations, *CRL1* RTL was on average 2.2-fold higher in MCC than in SCC across the *D* range (Figure 5.7), which was also reflected on that q_p raises between 2.1 and 2.4 (Table 5.2).

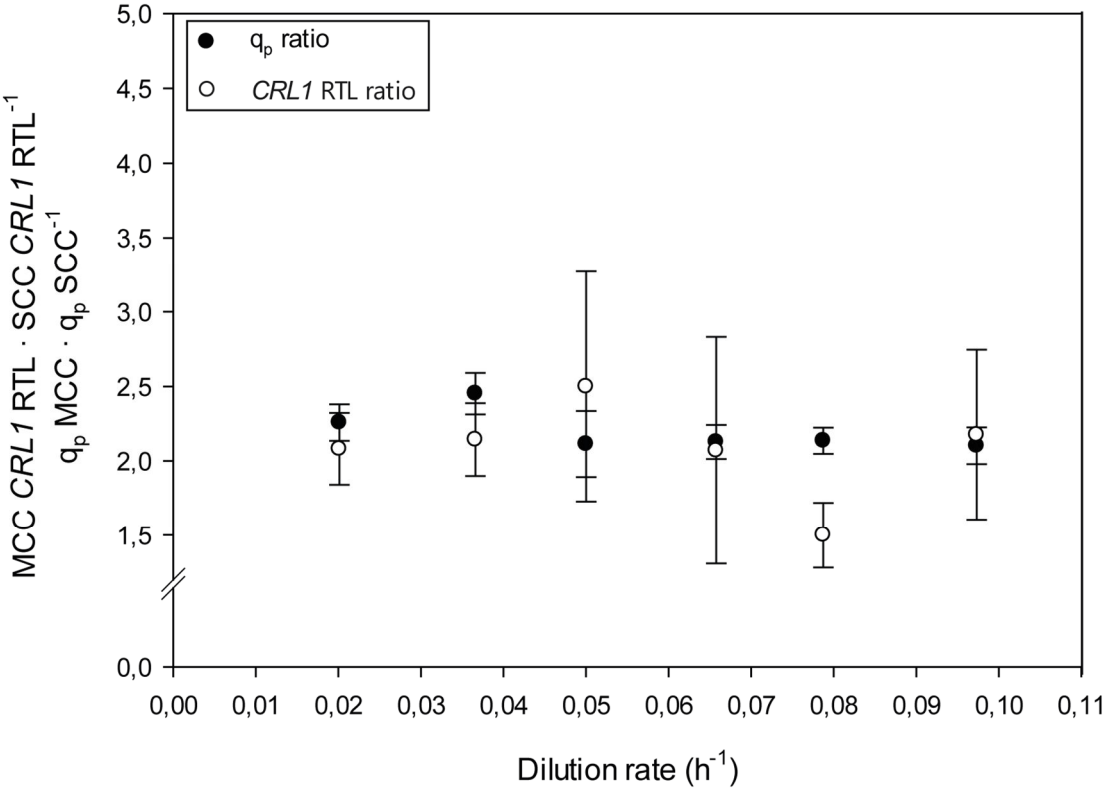


Figure 5.7. Effect of dilution rate on the *CRL1* relative transcript levels and specific production rate ratios between MCC and SCC. Error bars represent the standard deviation of q_p and RTL ratios.

This correlation between *CRL1* RTL and q_p ratios suggests the absence of a bottleneck in further protein processing–secretion steps, and hence in overall Crl1 production rate. This hypothesis was supported by conducting a transcriptional analysis of UPR-related genes such as *KAR2* and *HAC1*, the expression levels of which were rather constant across the *D* range (results not shown). Likewise, raises of other product-related parameters were also around 2.0 and 2.8-fold higher when comparing equivalent *D* conditions (Table 5.2). However, the improvement of product-related parameters could be considered slightly lower than expected, since MCC harbour three *CRL1* expression cassette copies. This fact has been widely reported previously, increasing the heterologous GOI dosage in the genome need not lead to a proportional increase in protein production rates (Sha, X.-W. Yu, *et al.*, 2013; Shu *et al.*, 2016; Cámara *et al.*, 2017a; Yu *et al.*, 2017). In this case, the

low *CRL1* transcription efficiency –*CRL1* RTL was just 2.2-fold higher on MCC– is the responsible of the lower than expected product-related parameters values.

In the following tables, a comparison of main product-related parameters of producer clones is presented.

Table 5.2. Comparison of Crl1 production-related parameters for chemostat cultivations.

	Single copy clone						**Multi copy clone**					
D (h^{-1})	0.019	0.035	0.050	0.064	0.079	0.095	0.020	0.036	0.050	0.066	0.079	0.097
Product Titer (AU mL^{-1})	124	101	118	107	93	68	321	267	274	243	212	155
Product Titer ratio	-------	------	------	------	------	------	2.6	2.6	2.3	2.3	2.3	2.3
q_P (AU g_X^{-1} h^{-1})	132	183	306	346	352	312	297	449	646	735	751	656
q_P **ratio**	-------	-------	------	------	------	------	2.2	2.4	2.1	2.1	2.1	2.1
Q_P (kAU L^{-1} h^{-1})	2.43	3.51	5.91	7.59	7.31	6.51	6.46	9.81	13.72	16.02	16.71	15.15
Q_P **ratio**	-------	-------	------	------	------	------	2.7	2.8	2.3	2.1	2.3	2.3
$Y_{P/S}$ (kAU g_S^{-1})	2.50	1.96	2.45	2.17	1.92	1.41	6.31	5.52	5.41	4.79	4.05	3.01
$Y_{P/S}$ **ratio**	-------	-------	-------	------	------	------	2.5	2.8	2.2	2.2	2.1	2.1
$Y_{P/X}$ (kAU g_X^{-1})	6.76	5.25	6.09	4.88	4.48	3.27	15.77	13.08	12.92	11.18	9.55	6.74
$Y_{P/X}$ **ratio**	-------	-------	-------	------	------	------	2.3	2.5	2.1	2.3	2.1	2.0

The ratios between the MCC and SCC were calculated by dividing the MCC parameter values to SCC ones at similar dilution rate (*D*).

Table 5.3. Comparison of Crl1 production-related parameters for fed-batch cultivations.

	Single-copy clone			Multi-copy clone		
μ (h^{-1})	0.028	0.047	0.084	0.027	0.046	0.063
Product Titer (AU mL^{-1})	769	660	261	1386	1542	1033
Product Titer ratio	-------	------	-------	1.80	2.32	3.95
q_P (AU g_X^{-1} h^{-1})	364	460	326	628	1322	958
q_P **ratio**	-------	-------	-------	1.73	2.25	2.93
Q_P (kAU L^{-1} h^{-1})	7.87	10.89	5.53	14.66	21.81	18.01
Q_P **ratio**	-------	-------	-------	1.86	2.00	3.26
$Y_{P/S}$ (kAU g_S^{-1})	5.37	4.67	1.86	9.74	12.19	7.25
$Y_{P/S}$ **ratio**	-------	-------	-------	1.81	2.74	3.90
$Y_{P/X}$ (kAU g_X^{-1})	13.16	9.76	3.91	23.84	28.88	15.32
$Y_{P/X}$ **ratio**	-------	-------	-------	1.81	2.56	3.91

The ratios between the MCC and SCC were calculated by dividing the MCC parameter values to SCC ones at similar specific growth rate (μ).

The increase in protein production derived from increasing the *CRL1* gene dosage in FB cultivations was similar to those obtained in chemostat cultivations. In fed-batch cultures, however, the ratios between clones increased with increasing μ (Table 5.3). Thus, at low μ levels, product-related parameters such as titer, q_p, Q_p (volumetric productivity) and

product yields were about 1.8-fold higher in MCC than they were in SCC. At intermediate μ levels, the previous parameters were roughly 2.3-fold higher in MCC. At the highest μ level, however, the comparison might not considered accurate enough since the culture conditions were not totally equivalent. Specifically, the μ set-point used with MCC had to be adapted to avoid an eventual methanol accumulation.

Overall, increasing the *CRL1* gene dosage resulted in increased protein production in both chemostat and FB cultivations.

Transcriptional efficiency differences between *AOX1* and *CRL1* genes, and their impact on Crl1 production

Regarding the balance between the transcription levels of the *AOX1* and the *CRL1* genes, being both P_{AOX1}-driven, an unexpected ratio between *AOX1* RTL and *CRL1* RTL was found in chemostat cultivations.

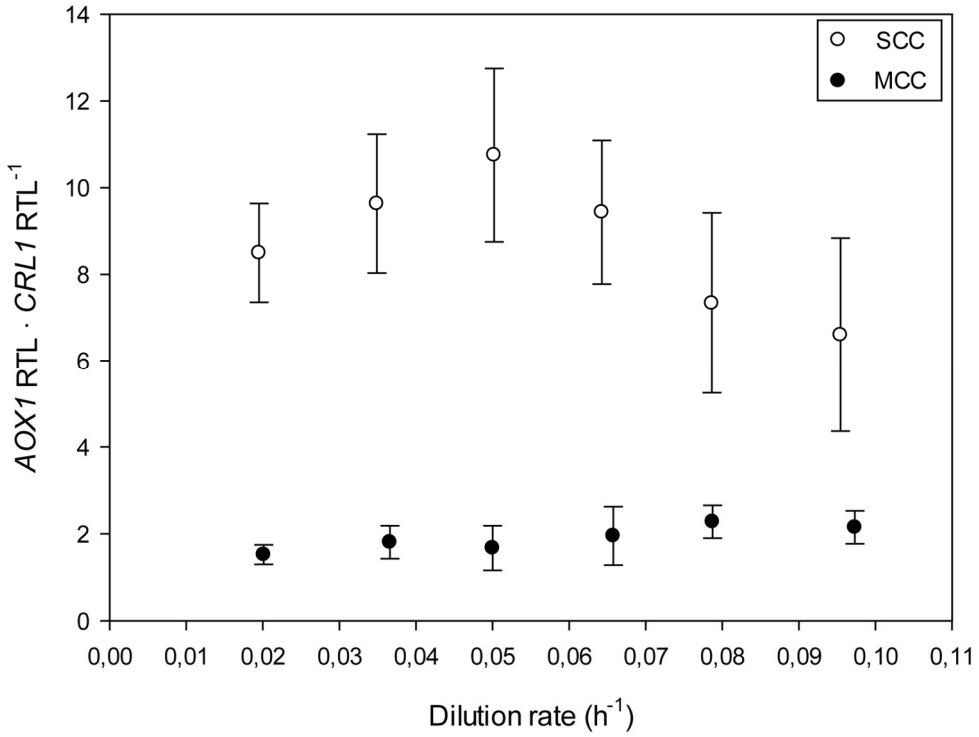


Figure 5.8. Effect of dilution rate on the *AOX1*-*CRL1* relative transcript levels ratio between MCC and SCC. Error bars represent the standard deviation of RTL ratios.

As can be seen from Figure 5.8, *AOX1* RTL was considerably greater than *CRL1* RTL in SCC, the difference ranging from 11 times at low *D* values to 6 times at the highest one. Even in MCC, which harbour three *CRL1* expression cassettes —versus only one of *AOX1*—, *AOX1* RTL exceeded clearly the *CRL1* RTL. The amount of mRNA a given gene contains is known to depend on the balance between transcription rate and mRNA decay (Cheng *et al.*, 2017). Therefore, since both coding sequences were flanked by the same promoter (P_{AOX1}) and transcription terminator (*AOX1*), one would expect the transcription rate to be similar. Hence, the differences in mRNA between *CRL1* and *AOX1* might be related with mRNA stability and hence with mRNA degradation.

As shown in Figure 5.7, *CRL1* RTL was closely correlated with q_p as a result of the absence of stacks in folding, trafficking and secretion processes. The *AOX1* RTL/*CRL1* RTL ratio was thus identified as a crucial parameter. As noted earlier, the pool of MUT-related TFs that can be shared by all the P_{AOX1}-driven genes is expected to be limited. As a result, increasing the number of *CRL1* cassettes in the genome should gradually increase the ratio up to a point where *AOX1* expression would not be enough to consume all the methanol fed in the culture. This hypothesis was confirmed in those cases where methanol accumulation was substantial. In this expression system, increasing the *CRL1* gene dosage to three expression cassettes (MCC) reduced the *AOX1*-to-*CRL1* RTL ratio from 10–7 in SCC to 2 in MCC; as a result, q_p was increased by a factor of 1.7–3.0 without relevantly affecting the ability to metabolize methanol fed to the culture. Therefore, the SCC expression system could be considered inefficient for producing Crl1 because the cell factory *P. pastoris* expresses higher levels of *AOX1* than the essentially needed, which is detrimental to Crl1 production.

Further increases of the number of *CRL1* cassettes is therefore the way of identifying the optimum *AOX1* RTL/*CRL1* RTL ratio for maximal Crl1 production without detracting from the physiological capabilities of the yeast. However, potential bottlenecks arising from an increased protein production should also be considered.

P_{AOX1}-based expression system rises above P_{GAP} for Crl1 production

Recombinant protein production levels highly depend on the expression system used in *P. pastoris* (Cos *et al.*, 2006; Vogl and Glieder, 2013; Vogl, Hatzl, Gerstmann, Pitzer, Wagner, G. G. Thallinger, *et al.*, 2016). Owing the drawbacks derived from the methanol utilisation in P_{AOX1}-triggering bioprocesses, alternative promoters have been also proposed in order to avoid the use of the mentioned hazardous compound. Among them, the constitutive P_{GAP} has been extensively used in RPP bioprocesses in *P. pastoris* (Waterham *et al.*, 1997; Looser *et al.*, 2014; Çalık *et al.*, 2015). The performance of P_{GAP} in comparison with P_{AOX1} has been a matter of discussion for decades and no definitive conclusions could be stated yet since it relies upon the target recombinant protein and the bioprocess strategy used (Cos *et al.*, 2006). To shed light on the most suitable promoter for Crl1 production, a homolog set of experiments in continuous cultivations were designed with an isogenic recombinant single copy clone expressing *CRL1* under the P_{GAP}. The deep characterization of this recombinant clone is part of a recently published work (Nieto-Taype *et al.*, 2019). Thus, a production performance comparison could be made with the SCC expressing the heterologous gene under P_{AOX1}. Cultivations designed for producing Crl1 under P_{GAP} were carried out using glucose as sole C-source instead of methanol. Therefore, wider range of μ could be covered —0.025-0.150 h^{-1}.

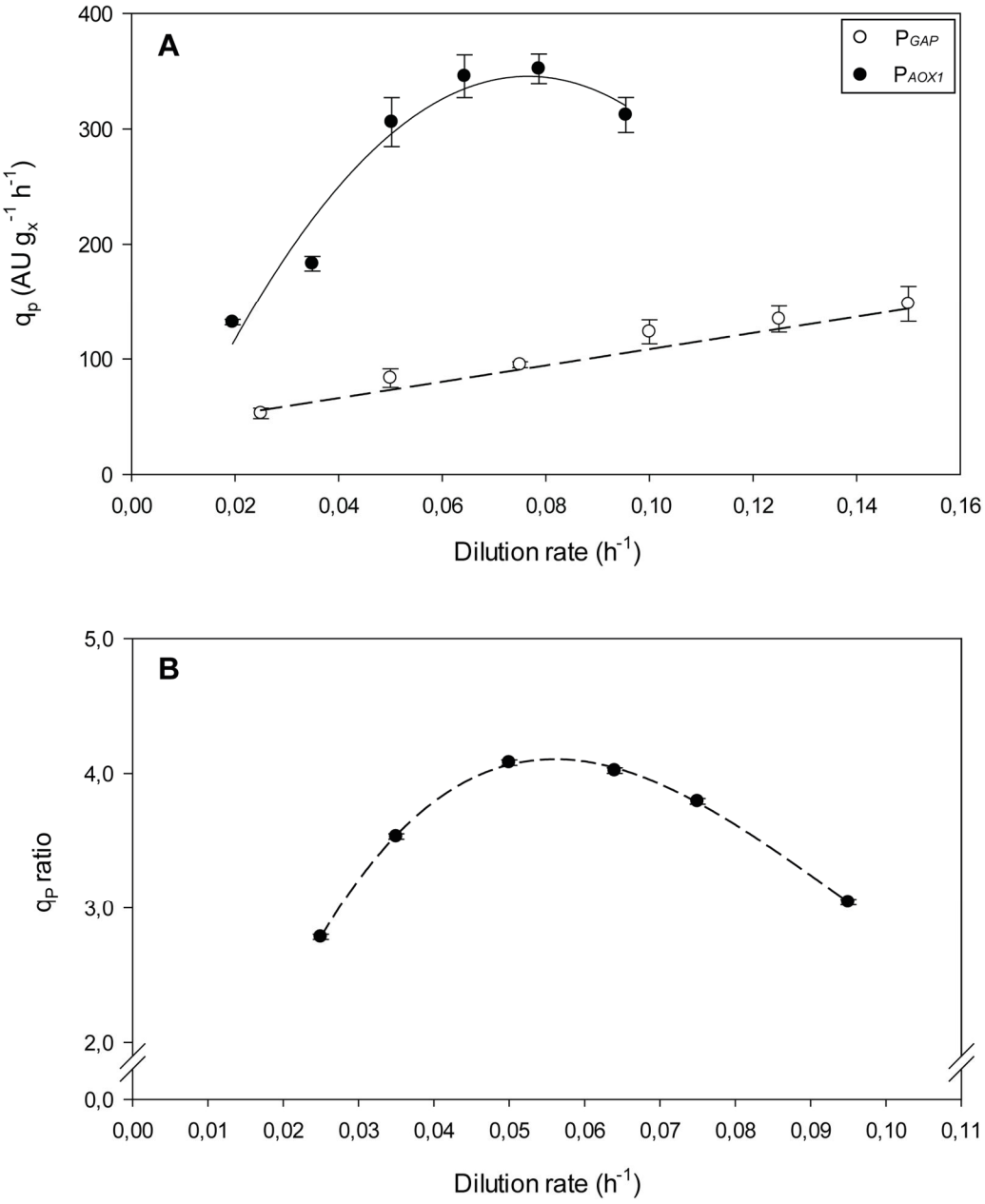


Figure 5.9. Production kinetics of single copy *CRL1*-expressing strains under P_{GAP} and P_{AOX1}. **A)** Specific Crl1 production rate (q_p). **B)** q_p ratio between P_{AOX1} and P_{GAP}-driven expression. q_p ratios were only calculated in the common *D* range of both expression systems. Lipase activity in liquid phase was used for q_p calculation. In figure A, error bars represent the standard deviation of two biological replicates.

The figure 5.9A shows the Crl1 production kinetics obtained with both P_{GAP} and P_{AOX1}. As pointed out previously, the C-source selection determined the $D_{washout}$ and, therefore, the *D* range covered. Regarding Crl1 production kinetics, two important observations must be remarked. On the one hand, the q_p values delivered by the P_{AOX1}-based expression

system outplayed the P_{GAP} ones by 2.8-4.0 times (Figure 5.9B), pointing out the significant benefits of using the methanol-inducible expression system for Crl1 production. On the other hand, a clear difference in production kinetics shape could be observed among the two expression systems. In previous results section, it was showed that μ takes over the expression of the *MIT1*, involved in P_{AOX1} induction. As a consequence, the expression of *AOX1* and *CRL1*, as well as the q_p followed a bell-shaped trend. Nonetheless, P_{GAP} has been widely recognized as a constitutive and growth coupled promoter. This regulation pattern led to a straight production kinetics trend, as reported elsewhere (Garcia-Ortega *et al.*, 2013, 2016; Nieto-Taype *et al.*, 2019). Importantly, the degree of tunability showed by the P_{AOX1} was much more pronounced that the P_{GAP}. It entails a considerable advantage since recombinant production strength could be adjusted according to the particular protein requirements.

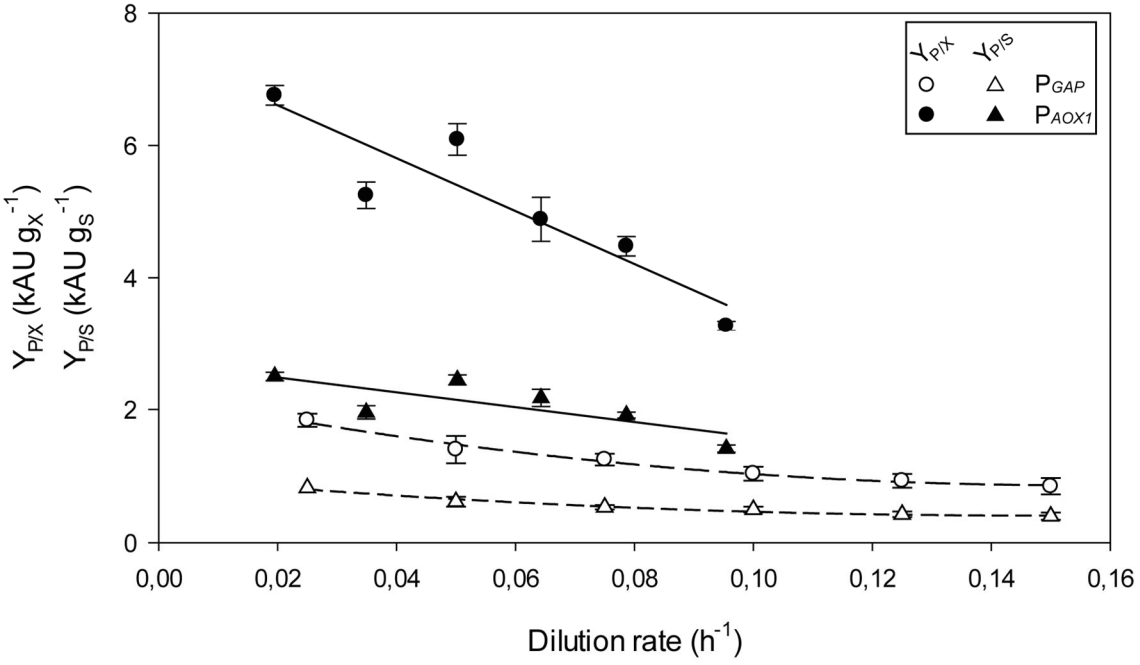


Figure 5.10. Production-related yields of single copy *CRL1*-expressing strains under P_{GAP} and P_{AOX1}. Lipase activity in liquid phase was used for product-related yields calculation. Error bars represent the standard deviation of two biological replicates.

Concerning the product-related yields, they followed a decreasing trend over D (Figure 5.10), suggesting that lower specific growth rates are beneficial for obtaining the highest protein levels per cell and substrate. Similarly to q_p, P_{AOX1}-based expression system exhibited significantly higher product-related yield values than P_{GAP} —between 3.3 and 4.0 times, confirming the superior performance of this expression system for Crl1 production.

Conclusions

In this work, the influence of the heterologous gene dosage was used to expose the high importance of μ on the transcription of MUT genes, production kinetics and culture physiological status in recombinant clones expressing *CRL1* under the control of P_{AOX1}. An operational mode impact on recombinant clone physiological state was noted when increasing the *CRL1* gene dosage. While both clones showed the same physiological behaviour in chemostat cultivations, the MCC physiological profiles departed from the expected linear trends seen in chemostat mode.

According to the results presented, μ seemingly determines the expression of *MIT1*, which have a key role in triggering transcription of P_{AOX1}-driven genes, thus influencing the amount of protein of interest that is produced at the end of the process. Also, increasing the number of *CRL1* expression cassettes integrated in the genome from one to three in *P. pastoris* effectively boosted production —$Y_{P/X}$ and q_p— without significantly altering the physiological status of the yeast. Furthermore, since increasing the *CRL1* dosage strongly reduced *AOX1* expression, one could hypothesize an eventual limitation of the TF *MIT1* pool, which is supported by our results. Finally, the strong correlation between *CRL1* RTL and specific *CRL1* production rate in both clones suggests the

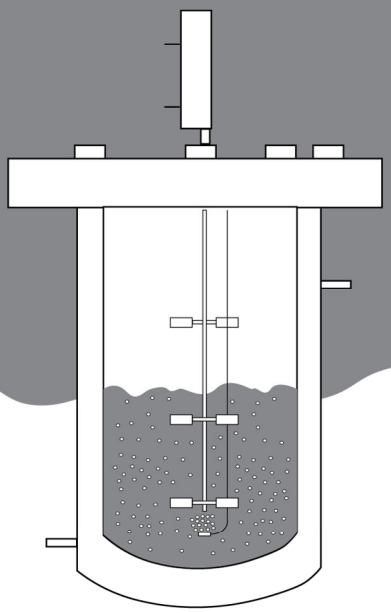
absence of bottlenecks in protein processing processes for this expression system in chemostat cultivations.

Furthermore, P_{AOX1}-based expression system has been proved to be superior to the constitutive P_{GAP}, according to the work published by Nieto and coworkers (Nieto-Taype *et al.*, 2019).

The outcome of these experiments expects to provide a wealth of knowledge for designing a rational approach to optimizing the operating conditions. Although the production patterns are expected to be similar for different proteins of interest to be expressed, the outcome usually depends on the expression regulation system as well as the target protein. Therefore, similar experiments should be conducted in each case, not only to maximize production rates, but also to identify the most suitable conditions for testing other strains with industrial potential.

RESULTS II

Bioprocess performance of novel methanol-independent promoters for recombinant protein production with *Pichia pastoris*



6

Chapter submitted as a research article in Microbial Cell Factories Journal

Garrigós-Martínez, J., Vuoristo, K., Nieto-Taype, M.A., Tähtiharju, J., Uusitalo, J., Tukiainen, P., Schmid, C., Tolstorukov, I., Madden, K., Penttilä, M., Montesinos-Seguí, JL., Valero, F., Glieder, A., Garcia-Ortega, X. Bioprocess performance of novel methanol-independent promoters for recombinant protein production with *Pichia pastoris*.

Results II content

Background

The non-conventional yeast *Komagataella phaffii,* widely known under the former name *Pichia pastoris*, is a distinguished host for recombinant protein production (RPP) (Cereghino *et al.*, 2002; Cregg *et al.*, 2009; Vogl *et al.*, 2013; Ahmad *et al.*, 2014; Looser *et al.*, 2014; Juturu and Wu, 2018; Yang and Zhang, 2018) and metabolite production (Peña *et al.*, 2018). Among the many positive features that makes *P. pastoris* a good choice for RPP, historically one of the most relevant ones is strong and tightly regulated expression based on the alcohol oxidase 1 promoter (P_{AOX1}) (Barrigón *et al.*, 2013, 2015; Ponte *et al.*, 2016, 2018). When using the P_{AOX1} promoter, induction occurs in the presence of methanol, whereas glycerol or glucose fully repress expression (Vogl and Glieder, 2013). Derepression is not enough for significant gene expression. Therefore, a simple recombinant protein production process is typically divided into two phases. First, a glucose/glycerol-based batch phase, where a relatively high amount of biomass is generated without recombinant protein production. Subsequently, the methanol feeding phase triggers strong P_{AOX1}-driven protein production. However, such tightly controlled induction and strong expression levels by P_{AOX1} causes operational drawbacks due to the use of methanol, including high oxygen requirements and heat production, as well as increased costs derived from methanol storage and handling (Heyland *et al.*, 2010; Çalık *et al.*, 2015). To counteract against these challenges, mutated promoter variants (Hartner *et al.*, 2008) or co-substrate feeding strategies had been employed (Berrios *et al.*, 2017).

In order to open new opportunities, innovative alternatives are constantly being developed, evaluated and implemented. In terms of promoter strength, the other methanol inducible promoters, *DAS1* and *DAS2* ($P_{DAS1;}$ P_{DAS2}), showed similar strength (Tschopp *et al.*, 1987). In addition to the numerous attempts that had been made to modify P_{AOX1}

regulation by mutagenesis or synthetic fusions (Portela *et al.*, 2018), the co-expression of transcription factors was demonstrated as an interesting alternative to induction by methanol. In addition, numerous methanol-independent expression systems have been developed and tested with promising results such as P_{GTH1}, P_{CAT1}, P_{THI11} P_{HpFMD} or P_{TEF}, among others (Landes *et al.*, 2016; Rajamanickam *et al.*, 2017; Prielhofer *et al.*, 2018; Shirvani *et al.*, 2019; Wang *et al.*, 2019; Vogl *et al.*, 2020).

Historically, RPP improvements have been mainly obtained through strain and promoter system engineering (Ben Azoun *et al.*; Hartner *et al.*, 2008; J. Yang *et al.*, 2016; Vogl, Sturmberger, Kickenweiz, Wasmayer, Schmid, A.-M. Hatzl, *et al.*, 2016; Liu *et al.*, 2020). Multiple clones with different expression cassettes or random integration variants with the same expression cassette are tested and compared in parallel in shake flasks or microtiter plates. This approach is considered fast and cost-effective; however, most testing platforms do not allow control of key bioprocess parameters such as dissolved oxygen, pH, and growth and feed rates. Since these parameters affect target protein expression, selection of best performing clones might not always be optimal. Accordingly, the performance of the production clones candidates need to be compared in cultivation platforms such as bench-top bioreactors (Landes *et al.*, 2016; Cámara *et al.*, 2017b, 2019), and/or alternative systems that allow controlled substrate delivery. Using bioreactors, production processes can be carried out applying optimal ranges of the key bioprocess parameters (Panula-Perälä *et al.*, 2008; Krause *et al.*, 2010). Chemostat systems, where cultures are maintained at non-dynamic, steady-state conditions, have become a valuable tool for process optimization (Nieto-Taype *et al.*, 2020). In this way, a full kinetic characterization of the candidate cell factories can be performed. Furthermore, interestingly, sampling for ’omics’ analyses can be reliably carried out on

cells from the chemostat from cultures having constant key process parameters, and the cell population is highly homogeneous (Nieto-Taype *et al.*, 2020).

Studies including precise strain characterization by chemostat cultivations, have revealed how the specific growth rate (μ) significantly affects the RPP rates (Potvin *et al.*, 2012; Looser *et al.*, 2014; Rebnegger *et al.*, 2014; Çalık *et al.*, 2015; Garcia-Ortega *et al.*, 2016; García-Ortega *et al.*, 2019; Garrigós-Martínez *et al.*, 2019; Nieto-Taype *et al.*, 2019). Importantly, the relationship between μ and q_p, also called production kinetics, is dependent on both the expression system used and the recombinant protein expressed. In previous studies, García-Ortega *et al.*(Garcia-Ortega *et al.*, 2016), and Nieto-Taype *et al.* (Nieto-Taype *et al.*, 2019) described a linear μ-q_p relationship when producing a human Fab, as well as the *Candida rugose* lipase 1 (Crl1), respectively, both under the control of the constitutive P_{GAP}. The same trend was observed for the production of Lipase B from *Candida antarctica* using the also constitutive PGK promoter (de Macedo Robert *et al.*, 2019). These authors concluded that since the constitutive P_{GAP} has a pivotal role in the growth-associated glycolysis, one should expect the RPP to be growth-coupled. On the other hand, curved/non-linear μ-q_p trends (Canales *et al.*, 2015a; Prielhofer *et al.*, 2018; Garrigós-Martínez *et al.*, 2019) were observed, suggesting non-coupled transcriptional regulation, or bottlenecks in the protein processing pathway. In particular, Garrigós-Martínez *et al.* remarked that the μ-q_p bell-shaped trend observed in the P_{AOXI} regulated production of Crl1 was probably caused by an alternative transcriptional regulation. This was suggested because at different μ, the target protein production profiles and the relative transcripts did not present the usual linearity of growth-coupled expression systems (Garrigós-Martínez *et al.*, 2019).

In this work, the performance of two novel expression systems for RPP in *P. pastoris* based on the new *PDF* promoter (P_{PDF}, a commercial variant of the *Hansenula*

polymorpha FMD promoter (Fischer *et al.*, 2019b; Vogl *et al.*, 2020), which drives strong transcription by simple methanol-free de-repression and can be also further induced with methanol) and *UPP* promoter (P_{UPP}, a constitutive commercial variant of the *Pichia* promoter P_{GCW14}, (Wang *et al.*, 2019)), have been thoroughly characterized and compared with $P_{GAP,}$ the most frequently used constitutive promoter, considered a reference standard for methanol-free expression systems. Expression strains for the lipase B from *Candida antarctica* (CalB) were constructed with the same parental strain and identical vectors, except for the promoter sequence driving *CALB* gene expression. To avoid an eventual gene dosage effect, for each selected representative transformant, the correct integration of only a single copy of the expression cassette in the *Pichia* genome was confirmed by qPCR (data not shown). To compare these expression systems, a set of chemostat cultivations designed to assess the effect of different μ values on the production kinetics was performed. Furthermore, *CALB* transcript levels were determined and compared to the expression levels for each condition tested in chemostat. Finally, as a first step of the scale-up and using the optimal μ conditions found in the chemostat mode, the same selected clones were cultivated in 15 L fed-batch processes to evaluate their performance in this operational mode.

This work was performed within the EU project IBISBA1.0 including the two IBISBA partners UAB and VTT. EU-IBISBA obtained a European Strategy Forum on Research Infrastructures (ESFRI) status in 2018, with the aim to accelerate biotechnology and synthetic biology activities in Europe through service infrastructure and know-how. Inter-laboratory studies are key in providing smooth project executions from production host design and construction to large pilot-scale production.

Results and discussion

Strain generation, screening and gene dosage

Isogenic clones were generated to compare the performance of the promoters P_{UPP}, P_{PDF} and P_{GAP} for the expression of CalB as a model recombinant protein. The parental strain, *P. pastoris* BSYBG11(*aox1-/Mut*^{S}), which is a BioGrammatics (Carlsbad, CA) *Pichia pastoris* BG11 strain, deposited at Bisy in Austria, was transformed with each of the expression vectors, only differing in the respective promoters. Considering the potential clonal variability usually observed in *Pichia* clone generation methods, care was taken to select a clone for each expression system with a single expression cassette integrated into the genome. Thus, low amounts of the linearized expression plasmid (cut with *Smi*I) were used (1 µg) for the *Pichia* transformation, which has been reported to be appropriate for single copy clone generation previously (Vogl *et al.*, 2014; Vogl, Sturmberger, Kickenweiz, Wasmayer, Schmid, A.-M. Hatzl, *et al.*, 2016). Subsequently, around 90 individual transformants were analyzed in a high-throughput screening based on deep well plate (DWP) system to develop a "landscape" of expression data for clone characterization according to Weis *et al.* (Weis *et al.*, 2004). Putative single-copy integration transformants for each of the different promoter constructs were picked from the majority of transformants which showed very similar lipase activity in the supernatant after cultivation and induction in 96-DWP and an initial screen measuring CalB activity of secreted reporter enzyme (data not shown). Clones secreting higher amounts of CalB were suspected to be associated with multicopy or random integration events; while lower activity observed in another clones might be related with detrimental effects exerted by ectopic integration (Schwarzhans *et al.*, 2016). Subsequently, a second round of DWP screening was performed with each clone cultivated at least in triplicates, to confirm the expression levels from the average clones for each expression system. From these single

copy integration candidates, numerous clones for each of the alternative promoter constructs were selected for gene dosage analysis by both qPCR and droplet digital PCR (ddPCR) (data not shown) to confirm that they were all clones that only integrated a single-copy of the respective expression vector in *Pichia*'s genome. Confirmed single copy clones for each expression system presenting an average CalB production were therefore selected to start the producer clone characterization to ensure production differences are only a result of the effect of each promoter's specific influence on CalB recombinant expression, and not due to a different gene dosage.

Physiological state comparison of the *P. pastoris* clones harboring different expression systems

Chemostat cultivations were performed with a selected CalB production clone for each of the three different expression systems (GAP-C, PDF-C and UPP-C). This comparison test was performed at three different dilution rates (D): 0.05 h^{-1}, 0.10 h^{-1} and 0.15 h^{-1}. This characterization allowed to determine the optimal conditions for CalB production in subsequent fed-batch (FB) cultivations.

As described in the previous screening section, P_{UPP} and P_{PDF} clones had significantly higher expression levels than those based on P_{GAP}. High levels of recombinant protein expression have been shown to cause a burden on the protein secretion machinery likely due to an overload of the processing capacity (Young *et al.*, 2011; Delic *et al.*, 2014). Therefore, an impact of the three expression systems on the physiological state was tested in chemostat cultivations by analyzing glycerol and O_2 consumption rates, and CO_2 production rates (Figure 6.1).

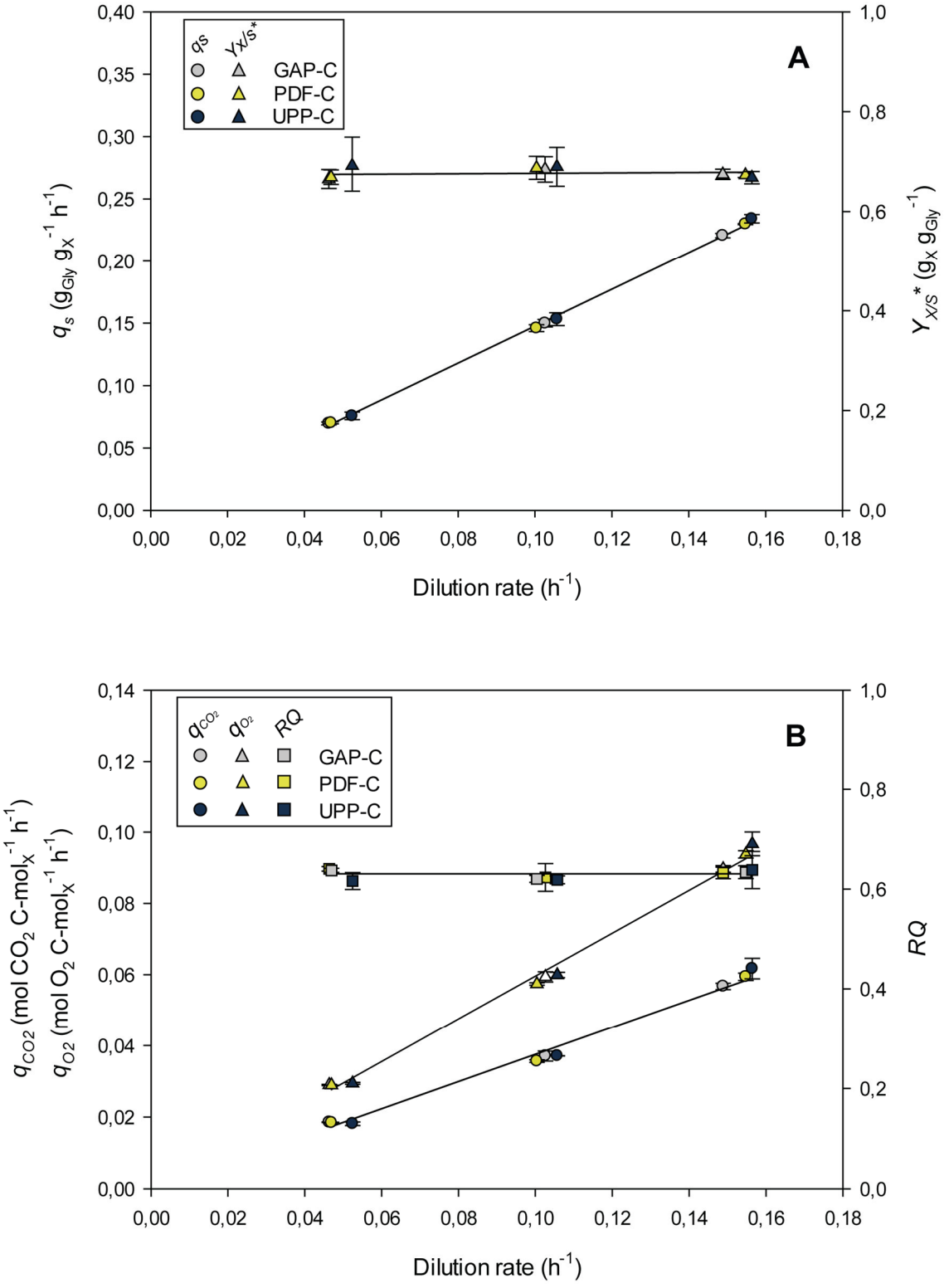


Figure 6.1. Physiological state indicators of *Pichia pastoris* CalB producer clones—GAP-C, PDF-C, UPP-C— in chemostat cultivations. **A**) Specific glycerol consumption rate (q_s), overall glycerol-to-biomass yield ($Y_{X/S}$*). **B**) Specific oxygen uptake rate (q_{O2}), specific carbon dioxide production rate (q_{CO2}) and respiratory quotient (RQ). Error bars represent the standard deviation of two biological replicates.

In this regard, no significant differences were evident across the *D* tested. The specific glycerol consumption rate (q_s) and overall biomass-to-substrate yield ($Y_{X/S}$*) were rather similar. As expected, q_s increases linearly over the *D*, whereas $Y_{X/S}$* values were constant (only slight differences could be observed at the highest D). All the clones presented similar specific CO_2 production rates (q_{CO_2}) and specific O_2 consumption rates (q_{O_2}) and followed standard linear trends. Consequently, similar respiratory quotient values were exhibited by all the studied clones (RQ, i.e. about 0.62, see figure 6.1B). Summarizing, based on the analysis at macrokinetic level, it can be stated that the higher CalB production provided by the new generation expression vectors employing the new promoters P_{PDF} or P_{UPP} did not alter any of the studied physiological parameters compared to the GAP-C with lower *CALB* expression levels.

Novel expression systems outperformed P_{GAP}-based CalB production

Compared to P_{GAP}, the use of both new promoters resulted in notably higher q_p values, between 4 and 9-fold higher at any *D* (Figure 6.2A). UPP-C also had q_p values significantly higher than PDF-C at the lowest and middle *D*. At the highest *D*, UPP-C was similar to PDF-C, with only a slightly higher q_p.

Different production kinetic profiles, q_p at different *D*, were obtained with the three expression systems compared (Figure 6.2A): UPP-C presented a bell-shape profile with a maximum at mid *D*, 0.10 h^{-1}. On the other hand, while for PDF-C was observed a clearly saturated profile, for the GAP-C a mainly linear pattern but including a slight saturation trend at higher *D* values was determined. This differs from other examples using P_{GAP} as constitutive expression systems in which, clearly, q_p increases linearly with *D* (Garcia-

Ortega *et al.*, 2016; Nieto-Taype *et al.*, 2019). Thus, these results indicate that production kinetics, in most cases, are protein dependent.

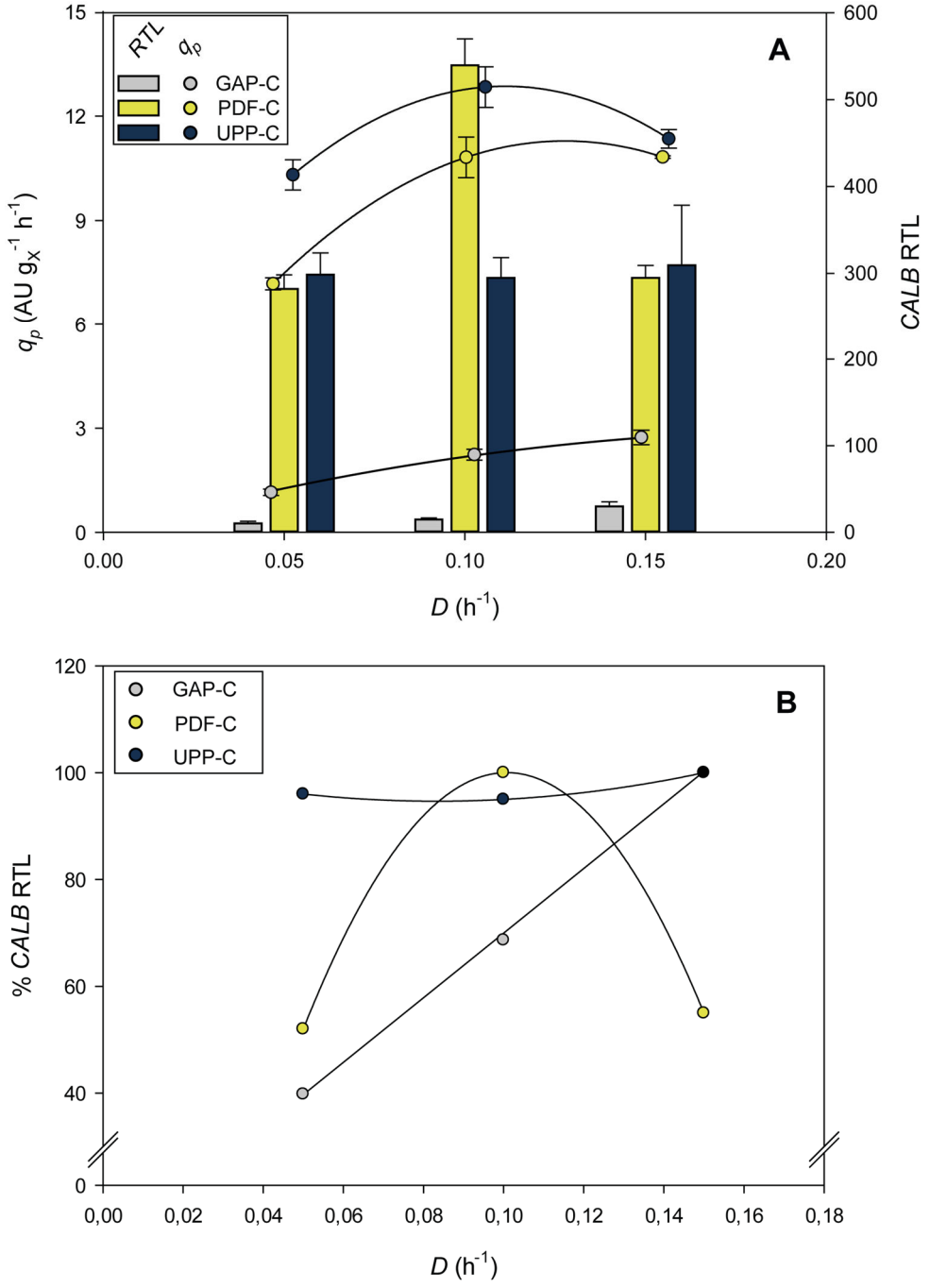


Figure 6.2. **A)** CalB production kinetics (q_P vs D) and *CALB* relative transcription levels determined in chemostat cultivations for the three expression systems studied. Transcript levels were normalized to the levels of the *MTH1* transcript, which was used as housekeeping gene for the analysis. Error bars represent the standard deviation of two biological replicates. **B)** Percentage of *CALB* relative transcription levels respect to the maximum observed for the corresponding expression system.

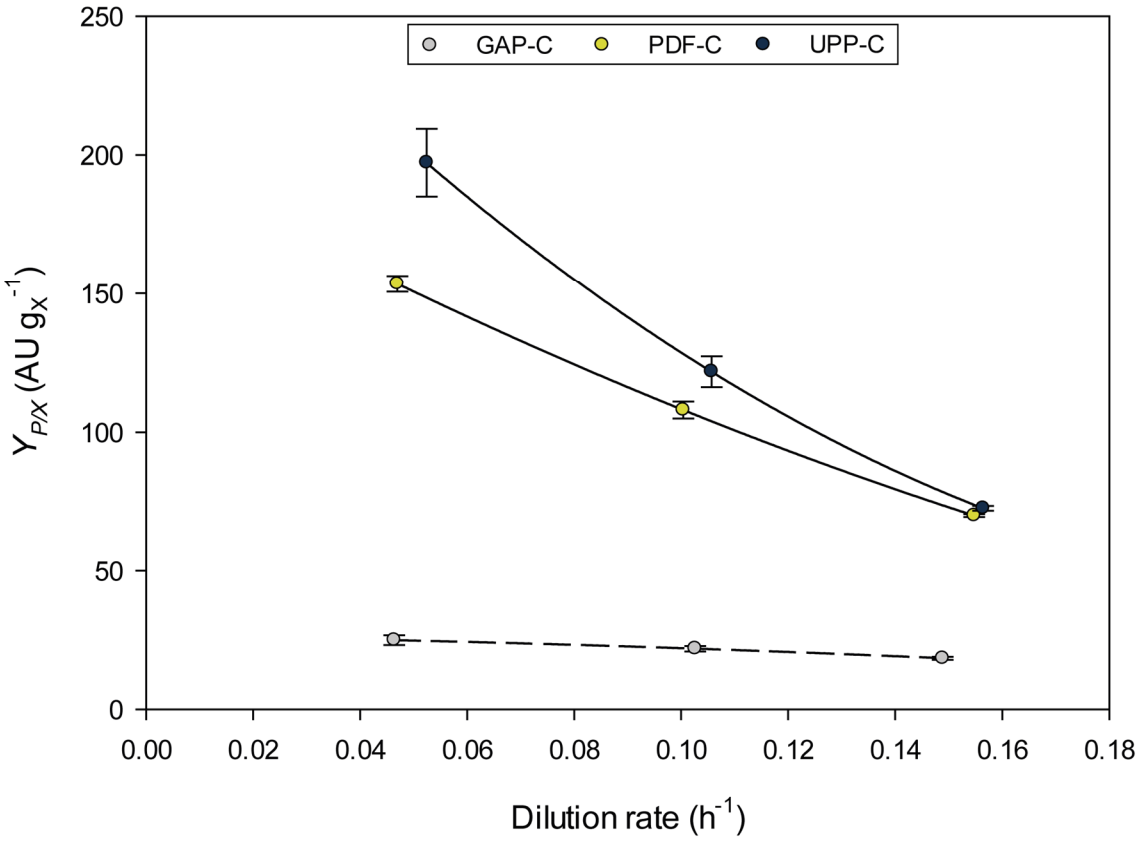


Figure 6.3. Comparison of the overall CalB product-to-biomass yield ($Y_{P/X}$*) in chemostat cultivations for the three expression systems tested. Error bars represent the standard deviation of two biological replicates.

Another important key parameter to be considered is the overall product to biomass yield ($Y_{P/X}$*); it determines the overall capacity of cells to produce recombinant protein under certain conditions. The P_{UPP} and P_{PDF} expression systems are similar to other systems where increasing *D* is detrimental for $Y_{P/X}$* (Garrigós-Martínez *et al.*, 2019; Nieto-Taype *et al.*, 2019), as shown in Figure 6.3. As observed with q_p, the biggest difference between UPP-C and PDF-C $Y_{P/X}$* values was at the lowest *D*, yet similar at higher *D*. Importantly, the highest $Y_{P/X}$* values for the UPP-C and PDF-C were notably higher than those obtained with the GAP-C (i.e. up to 8.9-fold higher with UPP-C at 0.05 h^{-1}) (Figure 6.3).

Based on the CalB production-related parameters q_p and $Y_{P/X}$*, both novel expression systems should be considered good candidate promoters to produce recombinant proteins.

Different conditions could be optimal depending on the objective: to reach higher product titer, or maximal productivity. If the objective is the highest protein titer, the lowest *D* should be selected (Garcia-Ortega *et al.*, 2013; Garrigós-Martínez *et al.*, 2019; Nieto-Taype *et al.*, 2019), especially in the case of UPP-C, where $Y_{P/X}$* reductions with *D* were more pronounced than with PDF-C. On the other hand, in order to maximize q_p, the optimal conditions for production should be at a *D* of 0.10 h^{-1}, observed most markedly with PDF-C, where *D* has a bigger impact (Figure 6.2A).

The new promoters enable increased tunability of recombinant protein expression processes in *Pichia pastoris*

mRNA levels are not always directly correlated with the level of recombinant protein production obtained (Landes *et al.*, 2016; Garrigós-Martínez *et al.*, 2019; Nieto-Taype *et al.*, 2019), especially when knowing about the physiological stress caused by high *CALB* transcript levels (Theron *et al.*, 2020). In the present work, variable *CALB* transcript levels were observed in all three expression systems compared: GAP-C, UPP-C, and PDF-C. As shown in Figure 6.2, a linear profile of *CALB* relative transcript levels (RTL) was observed across *D* for GAP-C, confirming the widely reported constitutive and growth-coupled regulation pattern of P_{GAP}. For this clone, even both RTL-*D* and q_p-*D* present rather similar profiles, a slight saturation trend of q_p can be observed at high *D* (Figure 6.2A), likely because the low production rates observed P_{GAP} are not expected provoke a big overload of the processing and secreting capacity. For UPP-C, only slight differences in *CALB* RTL were observed among the different *D* tested. Therefore, the regulation of *CALB* expression under P_{UPP} control should be considered growth independent. Strikingly, the *CALB* mRNA expression patterns did not correlate with the bell-shaped q_p profile described in the previous section (highest at a 0.10 *D*). For the PDF-C, RTL presents a bell-shape trend, while the q_p-*D* profile presents a saturation pattern. Therefore,

according the RTL results (Figure 6.2B), the P_{PDF}-based expression system exhibits a growth-rate dependent regulation, which thus can be considered a system with a promising tunable expression pattern.

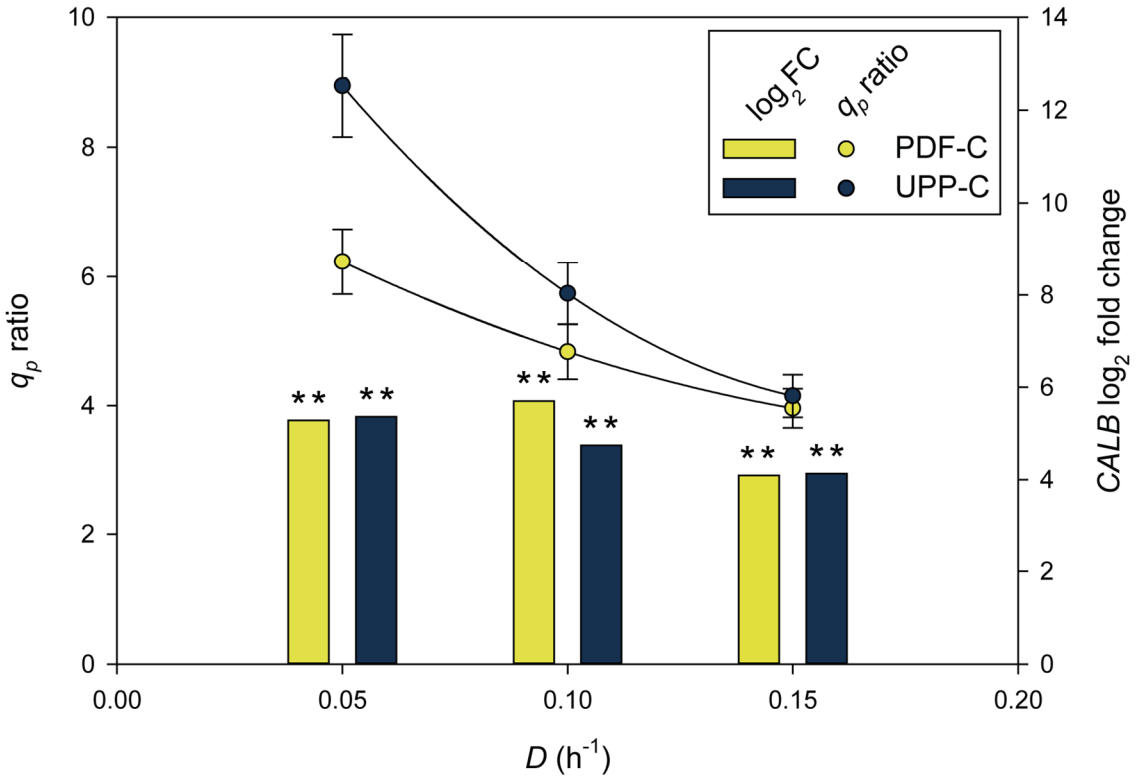


Figure 6.4. Effect of dilution rate on the q_p ratio and the *CALB* differential RTL calculated as $\log_2$ fold change relative to GAP-C values. GAP-C values were used as a control for ratio calculations. P-values (t-test) were calculated in order to determine the *CALB* expression significance between the different clones (* significance level $p \leq 0.05$; ** significance level $p \leq 0.01$).

As presented in Figure 6.4, the comparison of *CALB* expression regulated by the new promoters, relative to the P_{GAP}, illustrated an interesting contrasting behavior. The weaker, growth-coupled, P_{GAP}-based expression system performs better at higher μ, likely because the highest target transcription levels are obtained while the recombinant protein "burden" can still be sorted, or processed, properly in the ER (at least for the studied single copy integrations). On the other hand, with both novel promoters, *CALB* transcription rates and specially q_p ratios, followed a decrease over D, thus indicating that

high specific growth rates are detrimental for the more productive systems. Both UPP-C and PDF-C generated *CALB* transcripts, as well as secreted protein at significantly higher levels than GAP-C. The higher transcript levels could be overwhelming the secretory pathway and trigger the unfolded protein response (UPR). Therefore, UPP-C and PDF-C might be better at low and medium μ, when most of transcription can be converted into protein, as is demonstrated by the higher productivity rates. Consequently, the optimal conditions need to balance growth and protein production towards to reach the optimization criteria.

UPR influence on CalB production

In order to assess the potential endoplasmic reticulum (ER) stress derived from the excessive heterologous protein production, the expression analysis of key UPR genes was conducted. The reporters selected were two well-known ER-resident chaperones, *KAR2* and *ERO1,* and a gene product generally considered to be an UPR master regulator, *HAC1* (Zhu *et al.*, 2011; Ahmad *et al.*, 2014; Puxbaum *et al.*, 2015). The relative transcription levels of these three UPR related genes were measured in UPP-C and PDF-C and compared to the GAP-C, presented as a $\log_2$ fold change (Figure 6.5).

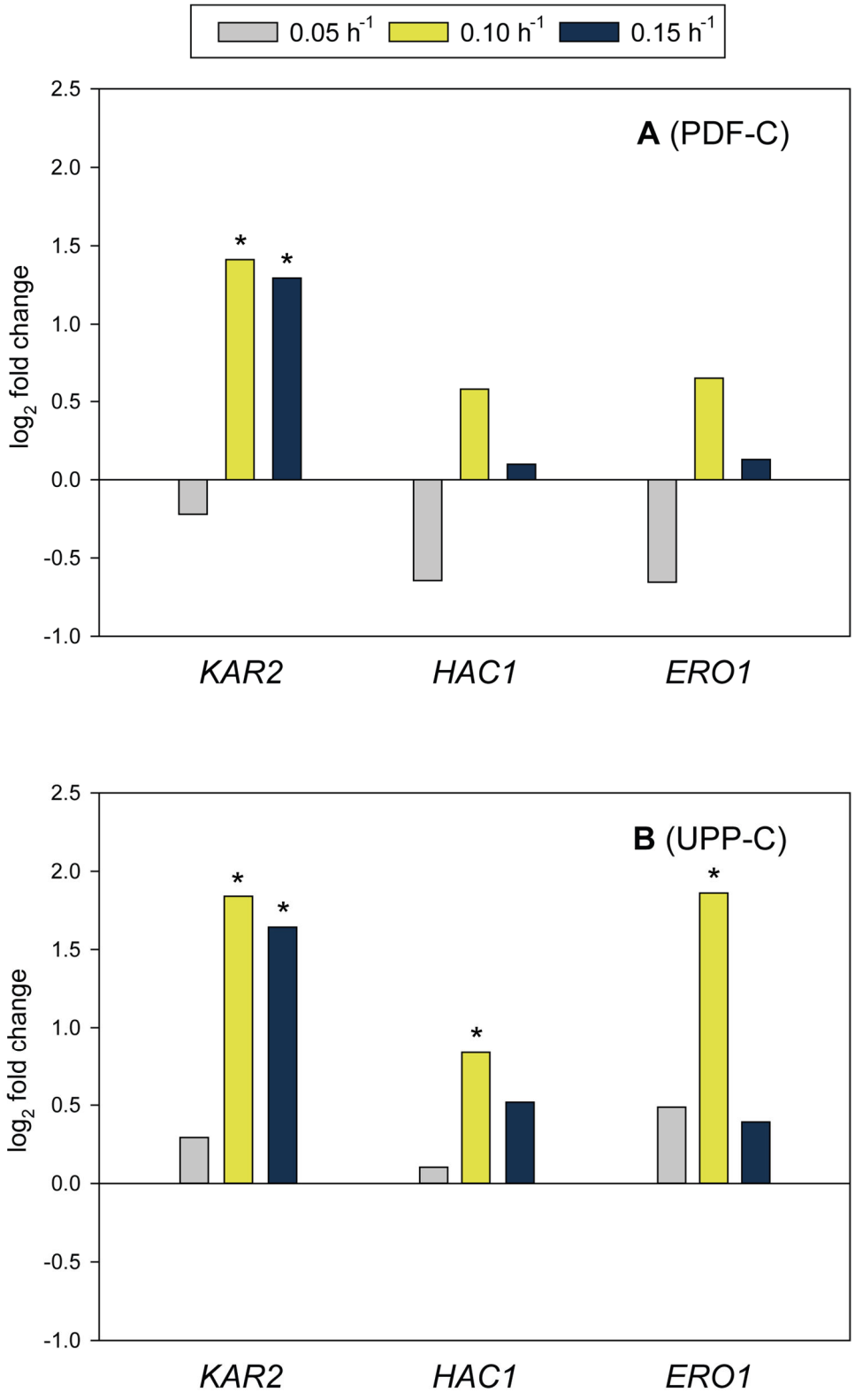


Figure 6.5. Transcriptional levels of three different UPR-related genes at the different dilution rates. A) PDF-C, and B) UPP-C (presented as a comparison of transcript levels with the levels from the control GAP-C, log2 scale). P-values (t-test) were calculated for all the genes and conditions in order to determine the gene expression significance between the different clones (* significance level $p \leq 0.05$).

Overall, the expression of all three of the UPR reporter genes presented a similar pattern across all the *D* rates tested. Their expression was the highest when the clones were cultivated at intermediate *D* of 0.10 h^{-1}, whereas only moderate expression increases were found at *D* 0.15 h^{-1}. These results are in line with other published work reporting that not only recombinant protein production, but also μ, having a significant impact on UPR induction (Rebnegger *et al.*, 2014; Raschmanová *et al.*, 2019). The cited studies are also similar to the data presented here, showing unremarkable UPR levels at lower μ, whereas μ increases led to UPR upregulation.

Strikingly, the growth-coupled expression of *CALB* by GAP-C was the only case in which the RTL presented a similar pattern to the q_p. Remarkably, nevertheless, the UPP-C q_p values mimic the UPR sensor gene expression profiles across the *D* tested, regardless of *CALB* transcription rates, which were rather similar for all the *D* rates. Therefore, UPR might have an influence in subsequent steps of CalB processing and secretion. The UPR impact on *CALB* expression is demonstrated by comparing q_p and *CALB* RTL values at *D* 0.10 h^{-1}. In the Figure 6.5, UPR-associated gene expression is higher for UPP-C (Figure 6.5B) than PDF-C (Figure 6.5A) and may explain why UPP-C q_p is higher than PDF-C q_p at this *D* despite UPP-C presenting 84% less *CALB* RTL. In this sense, it has been also described that the co-expression of protein disulfide isomerase, which is also upregulated at higher UPR, enhances active lipase production by *P. pastoris* (Sha, X.-W. Yu, *et al.*, 2013). Lastly, the comparison between two *D* conditions of PDF-C in continuous cultivations supports this hypothesis. In both *D* 0.05 h^{-1} and *D* 0.15 h^{-1}, the *CALB* RTL levels are rather similar. However, the UPR-related gene expression is growth coupled, enhanced at higher *D*. This could contribute to a 50% higher q_p at the highest *D* even though the target gene RTL levels are rather similar.

Together, these analyses indicate that q_p for CalB in *P. pastoris* is influenced by several factors: heterologous gene transcription rates, recombinant protein-associated UPR, and *D*-associated UPR. Comparing the different alternatives for methanol-free expression presented in this work, it could be observed that the new generation constructs, based on the *PDF* and *UPP* promoters, allowed to achieve *CALB* transcription levels of up to 8-fold higher than P_{GAP}-regulated expression, for all the *D* tested. However, at the high expression levels, a direct correlation between *CALB* RTL and q_p was not observed. In contrast, higher q_p values were usually observed at *D* conditions with enhanced expression of UPR-related genes, suggesting a relevant impact of UPR on CalB production.

Fed-batch cultures for further scalable bioprocess development

Through chemostat cultivations, a thorough characterization of the three expression systems studied was carried out, gathering information both at macrokinetic and transcriptional level. From these results, the optimal μ conditions for reaching the maximum CalB production were determined in order to maximize product yields and/or productivities. Ideally, the best μ values found in chemostat cultivations should be implemented to FB cultures, which is currently the most widely used scalable operational mode for industrial production of recombinant proteins. However, production kinetics may present relevant differences between the different operational modes (Garrigós-Martínez *et al.*, 2019; Nieto-Taype *et al.*, 2019). Therefore, fed-batch cultivations were also conducted to confirm CalB production kinetics determined in chemostat cultivations.

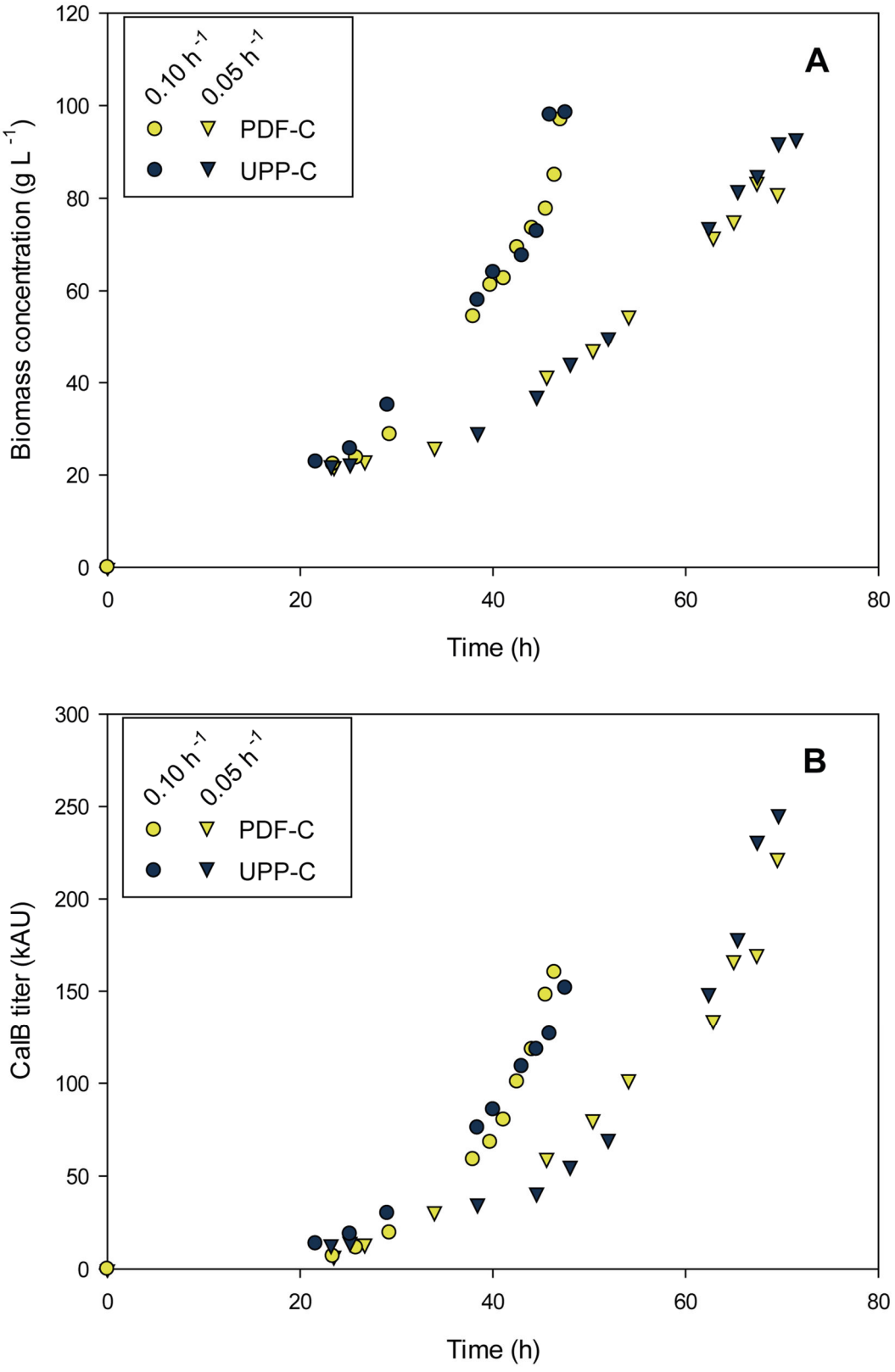


Figure 6.6. Fed-batch culture time profiles for Biomass (**A)** and CalB production (**B**) expressed as cell concentration (DCW) and total activity units, respectively.

Carbon-limited fed-batch cultivations, the culture strategy usually considered as most efficient with *P. pastoris* methanol free processes, were performed with the UPP-C and PDF-C to obtain biomass and CalB production profiles (Figure 6.6). Accordingly, based on the results obtained in chemostat cultures in which strong expression systems perform better at mid-low μ, the μ of 0.15 h^{-1} was discarded for further fed-batch implementation.

As expected, biomass production of all the cultures presented the targeted exponential profile, reaching a maximum between 90 and 100 g L^{-1} of dry cell weight (DCW), which is considered a standard endpoint for *Pichia* high-cell density fed-batch (Figure 6.6A). CalB production, expressed in activity units (kAU), also increased pseudo-exponentially over time (Figure 6.6B). Product titers obtained at the lowest μ tested (~0.05 h^{-1}) were substantially higher than with the intermediate μ (~0.10 h^{-1}), being 38% higher with PDF-C and 67% with UPP-C. Therefore, $Y_{P/X}$* values were also markedly higher at low μ cultures (Table 6.1). Comparing product-related parameters between chemostat and FB cultivations, PDF-C performed better in FB mode (24% higher, on average q_p, and 41% on average higher $Y_{P/X}$* at the equivalent μ, Table 6.1). On the other hand, UPP-C presented smaller differences of q_p and $Y_{P/X}$* values at low μ, on average below 10%. Strikingly, for the range μ 0.10 h^{-1}, UPP-C presented performance parameters significantly worse in FB cultures relative to the chemostat cultures (Table 6.1).

At the end of the batch phase, a 2-fold higher titer was obtained using UPP-C compared to PDF-C. That confirmed the relevant difference in the expression systems regulation in presence of carbon excess, a situation that only occurs during the batch phase. Thus, we concluded that the P_{UPP} presents a constitutive and strong expression in excess of C-source, conditions in which the P_{PDF}-regulated expression was repressed. Its expression can be derepressed by C-limiting conditions, usually at the end of the batch phase or during the fed-batch phase, thus being an interesting tunable expression system, which

allows simple uncoupling of biomass growth and transcription rates without altering the carbon source.

Table 6.1. Comparison of the main production parameters obtained in chemostat and fed-batch cultures with the producer clones at different specific growth rates (μ).

Clone	Operational mode	*Nominal* μ	*Experimental* μ	q_p	$Y_{P/X}$*
		h^{-1}	h^{-1}	AU g_x^{-1} h^{-1}	AU g_x^{-1}
GAP	Chemostat	0.05	0.046	1.16	24.9
		0.10	0.103	2.24	21.9
		0.15	0.149	2.74	18.4
PDF	Chemostat	0.05	0.047	7.20	153
		0.10	0.100	10.8	108
		0.15	0.155	10.8	70.1
	Fed-batch	0.05	0.042	9.20	219
		0.10	0.087	13.1	150
UPP	Chemostat	0.05	0.052	10.3	197
		0.10	0.106	12.9	122
		0.15	0.156	11.4	72.7
	Fed-batch	0.05	0.051	11.1	217
		0.10	0.084	9.95	102

Conclusions

In this study, a characterization of two promising methanol-free expression systems were conducted for RPP in *P. pastoris*. In terms of substrate and respiration-related parameters, all three expression systems showed similar behavior, suggesting that the potential differences in CalB production does not significantly alter the yeast homeostasis in chemostat cultivations. As main outcome, the CalB production kinetics with the two novel expression systems exhibited values significantly higher than the reference, GAP-C, up to 9-fold higher in terms of q_p. These important differences in product-related parameters were mostly attributed to the significantly higher *CALB* transcription levels. Interestingly, under carbon-limiting conditions, the P_{PDF}-based expression system showed a *D*-dependent tunable expression, while P_{UPP}-regulated expression was rather constant, regardless the growth rates tested. Furthermore, an UPR upregulation was noted specially for the UPP-C at D=0.1 h^{-1}. At this dilution rate, the three UPR sensor genes monitored were at their highest level. Notable is that the highest q_p was also at this *D* condition.

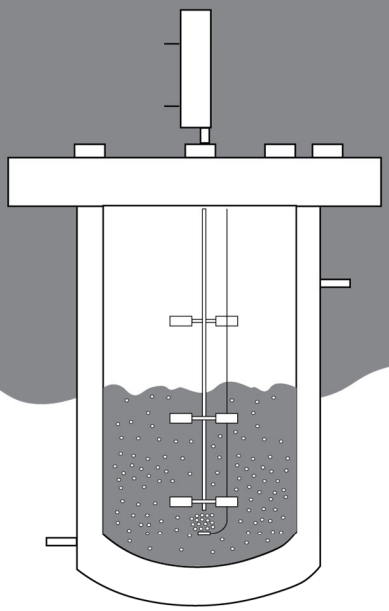
The chemostat results were used to design and implement a scalable production process for fed-batch cultivations. Both novel expression systems, based on P_{UPP}, $P_{PDF,}$ were also tested in fed-batch fermentations. The difference in regulation patterns was reproduced in a fed-batch mode as UPP-C had around 2-fold higher CalB production than PDF-C at the end of batch phase, illustrating a strong constitutive *CALB* expression under P_{UPP} regulation. On the other hand, highest *CALB* expression in the PDF-C was obtained under C-limiting conditions, in which the expression is derepressed, thus presenting an interesting tunable expression pattern. Concerning the production kinetics, UPP-C showed much better performance at low μ in fed-batch, as the q_p at this μ outperformed

the levels obtained at mid μ by 11%. On the other hand, PDF-C behavior was enhanced at mid μ as q_p was significantly higher than in low μ condition.

This work, testing the promoter design at larger scales, validated the scalability of small scale screenings and characterization in chemostats and, most importantly, that industrially relevant production conditions should be taken into consideration already when making the molecular design of the host. The access to complementary expertise and equipment at different scale enabled these Inter-laboratory studies and smooth project executions from production host design and construction to large pilot-scale production, thanks to the EU project IBISBA1.0 including the two IBISBA partners UAB and VTT.

RESULTS III

Production of human cytochrome P450 2C9 in *Pichia pastoris* whole cells: from strain and bioprocess engineering to preparative scale demonstration



7

Chapter submitted as a research article in Microbial Cell Factories Journal

Garrigós-Martínez, J., Weninger, A., Montesinos-Seguí, JL., Schmid, C., Valero, F., Rinnofner, C., Glieder, A., Garcia-Ortega, X.

Scalable production and application of *Pichia pastoris* whoce cell catalyst expressing human cytochrome P450 2C9

Results III content

Background

To avoid adverse drug reactions (ADRs) not only the drug as such, but also the impact of its metabolites on the organism are considered of great importance (Slaughter and Edwards, 1995). Cytochrome P450s (CYPs) are a super-family of heme-containing membrane-associated monooxygenases involved in the oxidation of monoterpenes, saturated fatty acids and alkanes, vitamins, steroids and eicosanoids (Guengerich, 2015; Bhattacharya and Yadav, 2017) and the clearance of drugs and xenobiotics in humans (Bernhardt, 2006; Schroer, Kittelmann, *et al.*, 2010). Among the human variants, next to CYP3A4 and CYP2D6, CYP2C9 is considered one of the most important drug oxidizing enzyme (Daly *et al.*, 2017). In particular, CYP2C9 oxidizes several drugs such as celecoxib, diclofenac, etodolac, ibuprofen, indomethacin, lornoxicam, mefenamic acid, suprofen, tenoxicam and also metabolizes some endogenous compounds, such as, for example, arachidonic acid, linoleic acid, and non-drug xenobiotics (Daly *et al.*, 2017).

Drugs and drug-metabolite standards are needed for toxicological, biological and drug-to-drug interaction studies. These kinds of studies are typically elucidated by analysis of tissue samples (e.g. human liver homogenate) or microsomal preparations thereof (Guengerich, 2015; Rinnofner *et al.*, 2019), because their synthesis is often limited as hydroxylation of non-activated carbon atoms is difficult in chemical synthesis (Geier *et al.*, 2012). Therefore, based on their versatile applications in drug metabolite synthesis, there is a great interest and high demand for human CYPs (hCYP) in pharmaceutical and chemical industry, especially when applicable on a preparative scale level (Schroer, Kittelmann, *et al.*, 2010; Urlacher and Girhard, 2019). However, biotransformations performed with purified enzymes are hardly feasible on a large scale, since CYPs usually show low efficiency and stability (Urlacher and Girhard, 2012, 2019). Moreover, for their

activity, CYPs require an electron transport system, which provides the electrons to the CYPs for oxygen activation and substrate oxidation (Nebert *et al.*, 2013). hCYPs rely on the presence of cytochrome P450 reductase (CPR), which is needed for electron transfer from the co-factor NAD(P)H.

During the last years, to fulfil industrial biotransformations that require high conversion rates, easy handling and low costs, the recombinant production of CYP enzymes has been intensively investigated (Urlacher and Girhard, 2012; Olaofe *et al.*, 2013; Quehl *et al.*, 2016; Forman *et al.*, 2018; Hausjell *et al.*, 2018; Worsch *et al.*, 2018; Rinnofner *et al.*, 2019). Specifically, successful recombinant expression of hCYP2C9 has been reported using insect cells (Rushmore *et al.*, 2000), *Escherichia coli* (Vail *et al.*, 2005) and *Schizosaccharomyces pombe* (Drăgan *et al.*, 2011). Thus, this enzyme was used for the preparation of milligram amounts of the 4'-hydroxy metabolite of diclofenac using whole cells or isolated membranes (Vail *et al.*, 2005). Showing the potential of yeast systems, Neunzig *et al.* first demonstrated the applicability of a whole cell based process for ibuprofen conversion using fission yeast for co-expression of human CYP2C9 and CPR (Neunzig *et al.*, 2012, 2013). Geier *et al.* also highlighted the methylotrophic yeast *Pichia pastoris,* recently re-classified as *Komagataella phaffii*, as the best cell factory for whole cell bioconversions using hCYP2D6 (Geier *et al.*, 2012).

P. pastoris is currently considered as an efficient alternative to bacterial and mammalian cell factories. Most of desirable membrane proteins (MPs) such as CYPs often come from eukaryotic organisms, thus bacterial expression systems usually fail to express them due to the lack capacity to perform post-translational modifications, the inefficient secretory pathways, which are needed to direct the protein to the membrane, as well as to the differences in lipid bilayer composition (Ramón and Marín, 2011; Byrne, 2015). On the other hand, mammalian expression systems are able to produce high quality proteins

(Byrne, 2015). However, their cell growth is significantly slower and more expensive compared to microbial cultures, which includes *P. pastoris*. Thus, on the basis of its high efficiency and cost-effectiveness, *P. pastoris* has become into a reference host system for the production of recombinant proteins, including membrane-anchored and soluble proteins—either intra or extracellular (Byrne, 2015). Works focused on the recombinant expression of MP's usually address issues related to protein degradation, ER folding problems and post-translational modifications in order to increase final product titers, since it may have a relevant impact on its production (Suades *et al.*; Claes *et al.*, 2016). However, to the date there has not been reported any study that assess the impact of operational strategies in the recombinant production of MPs targeted to be anchored in the cell membrane. So, as a relevant novelty of this work, it has been aimed to determine the interrelation between the specific production rate (q_p) and the specific growth rate (μ) of the culture. This relationship is called production kinetics, and it is often considered a key to design optimal bioprocess strategies for recombinant protein production (RPP) (Looser *et al.*, 2014; Rebnegger *et al.*, 2014; Çalık *et al.*, 2015; Garrigós-Martínez *et al.*, 2019; Nieto-Taype *et al.*, 2019).

The aim of this work is to produce highly active CYP2C9/CPR whole cell biocatalysts in *P. pastoris* by the design and implementation of a feasible and scalable production process. In this respect, an innovative strategy based on the dual gene expression by means of a bidirectional promoter system (BDS) has been implemented for the simultaneous expression of both *CYP2C9* and *CPR* genes (Vogl, Kickenweiz, *et al.*, 2018). Bioprocesses that requires the co-expression of multiples genes, i.e. *de novo* building of biosynthesis pathway, are considered challenges in metabolic engineering and synthetic biology, since components of the synthetic system need to be balanced to prevent potential bottlenecks. Furthermore, the total amounts of recombinant proteins

might have a relevant impact, since as an excessive recombinant protein burden might be detrimental (Galanie *et al.*, 2015; Tan and Prather, 2017). Consequently, in order to achieve the best results, several combinations of expression systems with offering different performances should be tested in order to find the best balance of components that enhances the overall system activity. In this regard, the BDS has shown as a simple and fast tool —half time respect to the monodirectional cloning methods— for versatile library construction. The resulting clone library included a variety of both regulatory patterns and expression strength, which therefore offers a high degree of freedom for bioprocess design and development due to its high degree of tunability. Moreover, further expression pattern diversity could be achieved by implementing promoter hybridisation, directed deletions/truncations or even fusions (Vogl, Kickenweiz, *et al.*, 2018) as shown in Figure 7.1. Interestingly, this work has evaluated the performance of up to eight different innovative promoters, which are considered promising alternatives to the widely used P_{GAP} and P_{AOX1}.

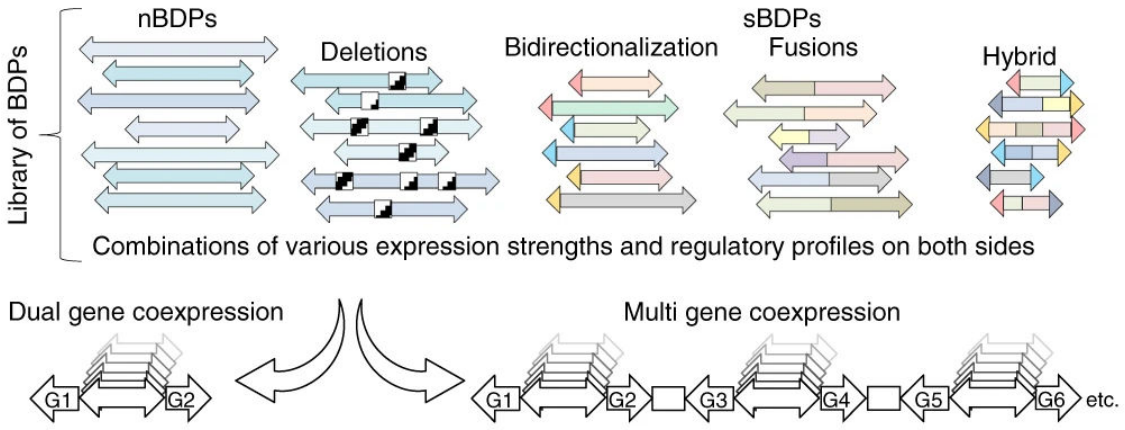


Figure 7.1. A library of diversely regulated natural and synthetic BDPs (nBDPs and sBDPs) covering a wide range of regulatory profiles facilitates optimization of dual gene co-expression and the assembly of multi-gene co-expression cassettes. Adapted from (Vogl, Kickenweiz, *et al.*, 2018).

Remarkably, one of the promoters tested for driving the expression of CYP/CPR is the P_{PDF}, which showed an outstanding performance in CalB production, as demonstrated in Chapter II of results. Therefore, an additional objective of this work is to ascertain if the notable outcomes obtained with this promoter for CalB production can be replicated in the production of an industrially interesting protein such as CYP2C9.

Once the best promoter combination was selected for the bidirectional expression system, then the bioprocess was rationally designed at bench-top bioreactor scale taking into consideration the impact of several factors. The marginal experience in producing membrane proteins in *P. pastoris* led to check out the importance of cultivation pH on CYP2C9/CPR production. Moreover, the experience gathered in works detailed in previous results sections concerning the effect of μ on soluble protein production suggested to assess the influence of this parameter in the production of membrane-anchored proteins. Furthermore, the effect of adding methanol as co-substrate was also investigated as the promoter regulation pattern used for driving the expression of CYP2C9 and CPR were described as derepressible and further inducible by methanol. At this point, it is important to remark that biocatalyst activity as whole cells requires the optimal balance between two different proteins. Consequently, both selection of the right promoter combination and the characterization of bioprocess conditions that lead to the highest biocatalyst activity becomes crucial. Finally, active whole cells obtained in bioreactor were successfully applied for a preparative scale biotransformation (0.5 L) of ibuprofen to 3-hydroxyibuprofen.

Results and Discussion

Strain generation

Bidirectional cloning offers a simple and quick solution to identify optimal promoter combinations in a single experiment, enabling multi-gene co-expression in which the expression of each gene can be optimally tuned towards to achieve the desired objective (Vogl, Kickenweiz, *et al.*, 2018). To fully exploit the potential of this bidirectional expression system, a library including up to 7 alternative promoters with diverse regulation patterns were used in different combinations to drive the expression of the full-length genes: *CYP2C9* (accession AL359672) including its hydrophobic N-terminal sequence and *CPR* (accession AAH34277.1). In this way, different expression levels and regulatory profiles were evaluated in order to achieve the best performance of the whole cell biocatalysts, in which is required that both domains are anchored to the membrane. The study included the strong methanol inducible promoters P_{DAS1} and P_{DAS2} (and their short variants $P_{DAS1\text{-}552}$ and $P_{DAS2\text{-}699}$), and the promoters P_{PDC} and P_{PDF}, two de-repressed promoters, which start expression upon glucose/glycerol depletion, and which are further inducible with methanol (MetOH). In earlier studies using GFP as a reporter, P_{DAS1} showed a strong methanol-inducible performance, reaching at least half of the level of the commonly used *AOX1* promoter, while P_{DAS2} even outperformed P_{AOX1} expression (Vogl, Hatzl, Gerstmann, Pitzer, Wagner, G. G. Thallinger, *et al.*, 2016). On the other hand, the expression of the P_{PDC} promoter (native catalase gene promoter) is repressed in the presence of glucose or glycerol and derepressed upon the depletion of these carbon sources. Its expression can further be induced with methanol and remarkably also with oleic acid (Fischer *et al.*, 2019a). Besides the P_{PDC}, an orthologous promoter, i.e. P_{PDF}, which exhibits a similar regulation profile, has been also tested. This kind of promoters

are highly interesting for the industry because allow to uncouple the recombinant protein expression and the cell growth.

In earlier studies, efficient CYP2D6 amounts detectable by carbon monoxide (CO) difference spectroscopy in *P. pastoris* whole cells were shown (Geier *et al.*, 2012). However, highest CYP levels determined by CO-spectroscopy could not be correlated to highest conversions levels when using whole cells. Working with whole cells, their conversion efficiencies are influenced by electron and substrate supply (Geier *et al.*, 2012), and it also affected by the complex interactions of multiple parameters (e.g. CYP/CPR ratio, membrane permeability, NADH production/regeneration or expression level of ROS degrading enzymes between among others) (Rußmayer *et al.*, 2015). Thus, this work is focused on to achieve the production of active whole cells as biocatalyst regardless the amount of CYP2C9 anchored to the cell membrane.

For these reasons, when carrying out the producer clones screening to identify the optimal promoter combination for the bidirectional expression system two alternative whole cell bioconversions based screening assays were performed (Weis *et al.*, 2004). One based on diclofenac conversion and the other one based on the fluorescence of the product 7-hydroxy-4-(trifluoromethyl)-2H-chromen-2-one (HFC). In figure 7.2, the products obtained for both substrates are presented, and in Figure 7.3, a comparison of the best clones of each construct is shown. The qualitative behaviour of both screening methods was quite similar, construct 2C9-PDC/PDF-CPR showed the best conversion (53%) of diclofenac (2mM). The same construct with inverted promoter orientation 2C9-*PDF*/*PDC*-CPR, also presented good substrate conversion reaching 39%, while all other constructs showed only 20% conversion of diclofenac or lower. Similar results were obtained using the alternative fluorometric MCF-based assay. As for the diclofenac screening, the P_{PDF}/P_{PDC} based expression system gave the best results, reaching a

maximum value of 28 RFU per OD_{600}. The fluorometric MCF-based assay is considered more suitable for HTS since online fluorescence signal can be easily monitored in 96-deep well plates with low cell density cultures. However, the use of this method to determine the biocatalyst activity of the subsequent cultivations carried out in bench-top fermenters was considered not accurate enough due to fluorescence shielding derived from the high cell densities reached. Consequently, to determine the activity of samples obtained from bench-top fermenters, only the diclofenac-based method was used.

From the results presented, the clone 2C9-PDC/PDF-CPR was selected for the further bench-top fermenter studies since it presented the best conversion rates for both methods.

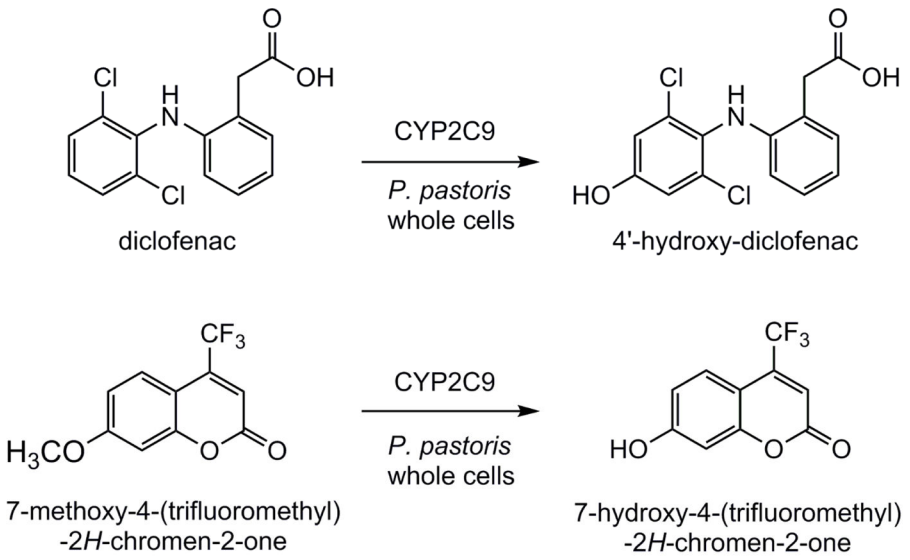


Figure 7.2. Diclofenac and 7-methoxy-4-(trifluoromethyl)-2H-chromen-2-one (MCF) oxidation reactions catalysed by *P. pastoris* CYP2C9/CPR whole cells.

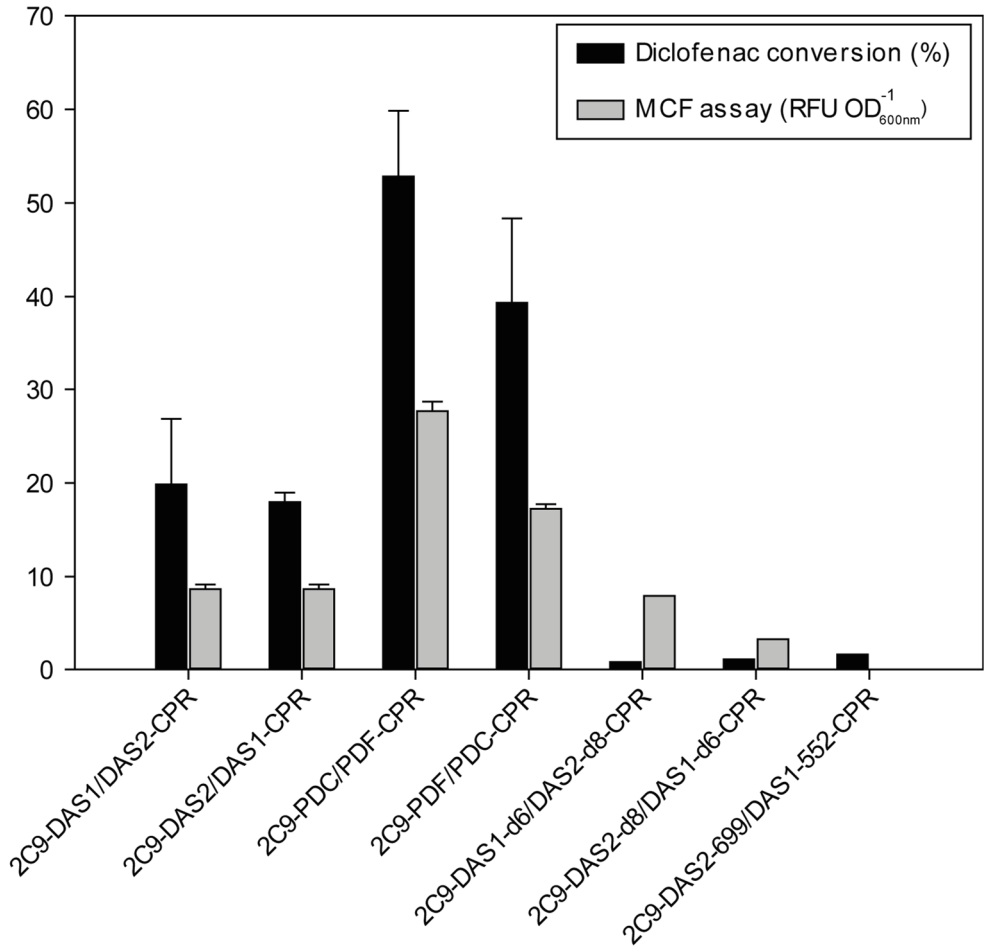


Figure 7.3. Screening of different bidirectional promoter variants for obtaining the most active CYP2C9/CPR whole cells.

Production of *P. pastoris* CYP2C9/CPR whole cell biocatalyst in bioreactor cultures

For the very first time, CYP2C9/CPR expressing cell factories were cultivated in bioreactor. Thus, it was considered interesting to compare CYP2C9/CPR expressing cells cultivated in either shake flask or bioreactor regarding their ability to convert diclofenac and the conversion per mg of DCW was quite similar, as shown in Figure 7.4.

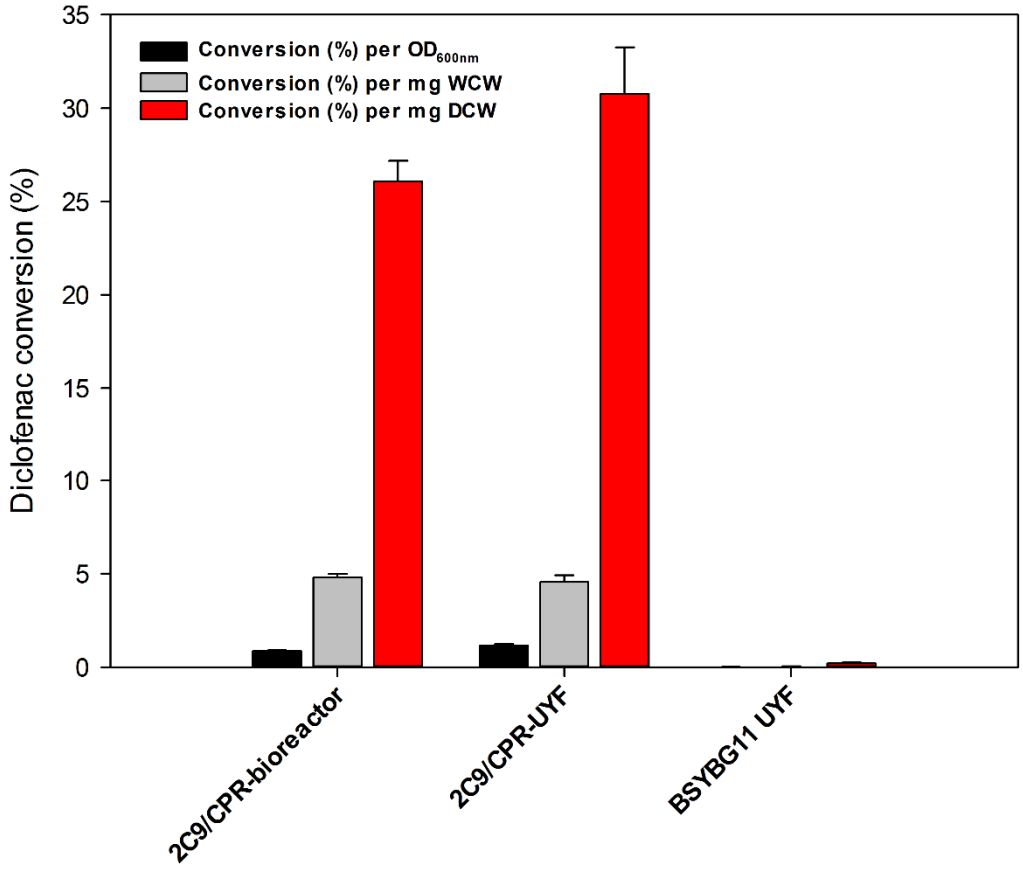


Figure 7.4. Whole cell hydroxylation of diclofenac by *P. pastoris* CYP2C9/CPR whole cells cultivated respectively in bioreactor and in shake flask (UYF = ultra-yield flask). The wild type strain BSYBG11 was used as a negative control.

Once it is assured that cells obtained from bench-top fermenters display similar diclofenac oxidation activities than cells from shake flask, a thorough study of CYP2C9/CPR whole cell production was carried out including chemostat cultivation and bioprocess optimization in fed-batch cultures.

Among the physical-chemical parameters, optimal pH is the most protein dependent. Thus, the effect of different pH values (pH 5-7) on the production of CYP2C9/CPR whole cells in chemostat cultivations at intermediate D=0.10 h^{-1} was assessed. Working in steady state conditions for a constant specific growth rate (D=μ) provides a robust and reliable comparison among different operational conditions for one variable while

keeping the others constant. Activity tests showed that the best condition was pH=5, obtaining a specific enzyme activity up to 3-fold higher than in the other conditions tested (Figure 7.5). Consequently, further fed-batch (FB) cultivations were always carried out at pH=5.

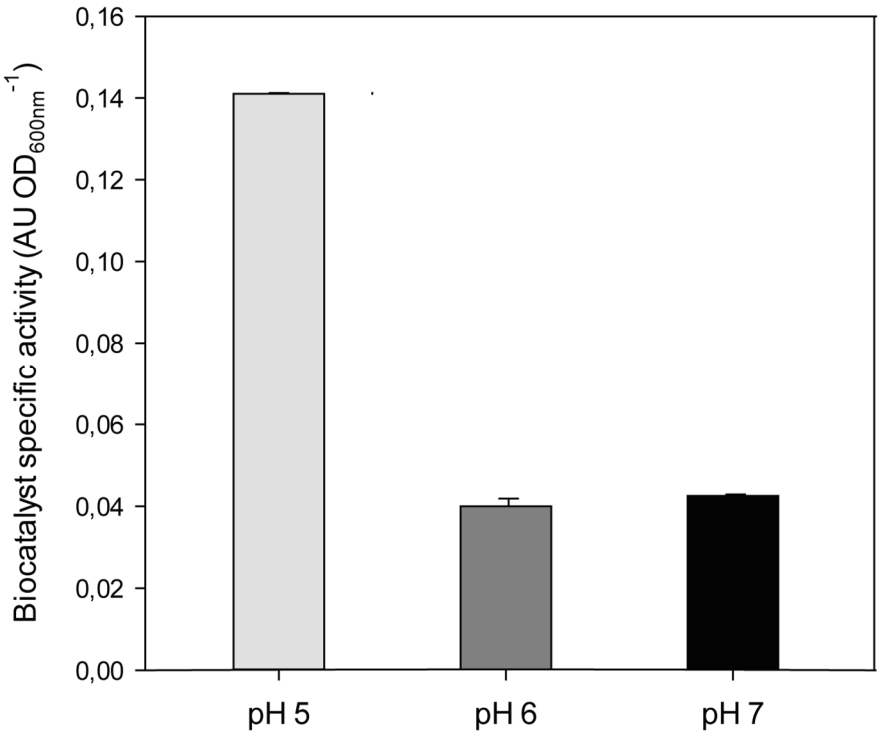


Figure 7.5. Optimal culture pH determination for the *P. pastoris* CYP2C9/CPR whole cells production. Chemostat cultivations were carried out at dilution rate 0.10 h^{-1}. Error bars represent the activity assay standard deviation determined by triplicates.

Biomass growth characterization in fed-batch cultures

Since the producer strain selected from the screening phase is based on 2C9-PDC/PDF-CPR, it allows to implement either methanol-induced and/or methanol-free bioprocess strategies. Therefore, as the current trend of *P. pastoris* bioprocesses is to avoid the use of methanol due to its operational associated drawbacks, a primary set of methanol-free FB cultures with glycerol as sole carbon source was performed. In these cultures, the

effect of the specific growth rate (μ) on biocatalyst production was evaluated by performing glycerol-limiting cultures at different nominal μ (FBs 1-3) ranging from 0.05 to 0.15 h^{-1}. Furthermore, the potential boosting effect of methanol on biocatalyst production was additionally assessed by growing the culture at μ=0.05 h^{-1} on glycerol-limiting conditions and adding methanol as co-substrate either by pulses (FB4) or keeping a constant methanol concentration of 3 g L^{-1} (FB5). The last strategy was selected since it was previously reported to be the best condition for *Rhizopus oryzae* lipase (Rol) production under the control of P_{AOX1} promoter (Barrigón *et al.*, 2013).

In co-substrate cultures designed for Mut^S phenotype strains, it is necessary to maintain a low μ—0.05 h^{-1}, otherwise methanol utilisation (MUT) genes might be repressed (Arnau *et al.*, 2011; Capone *et al.*, 2015). Consequently, methanol would not be consumed and its impact as promoter inducer might not be properly assessed.

Concerning biomass generation, the selected producer clone was able to grow as expected in all three methanol-free FB cultures (Figure 7.6A), so the recombinant expression did not have any significant limiting effect on the cell growth. Most of the cultures were grown up to reach 100 g L^{-1} of DCW. Nevertheless, the fermentation performed at 0.15 h^{-1} (FB1) had to be stopped at a biomass concentration around 80 g L^{-1}. At this point, the temperature control could no longer be able to keep constant the temperature at 25 ˚C due to the high amount of heat generated at high specific growth rate. Anyway, numerous samples were taken throughout the process, which allowed the determination of the production-related macrokinetic parameters needed for its comparison with the rest of fed-batches cultures.

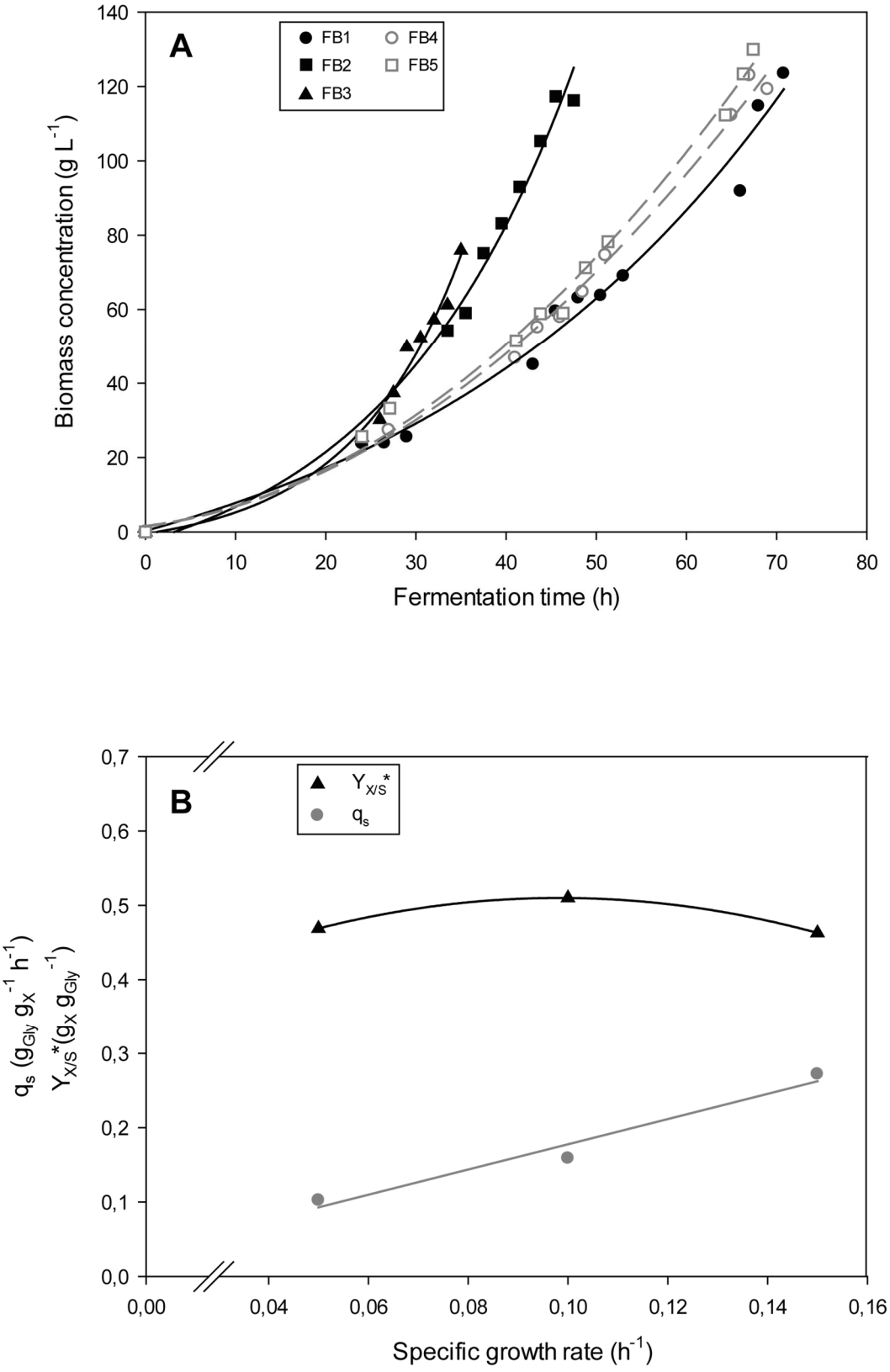


Figure 7.6. **A)** Biomass time-profile of the fed-batch cultures performed. **B)** Biomass related parameters, q_S and $Y_{X/S}^*$ for methanol-free experiments. Experiment labels: FB1: μ=0.05 h^{-1}; FB2: μ=0.10 h^{-1}; FB3: μ=0.15 h^{-1}; FB4: μ=0.05 h^{-1} + MetOH pulses; FB5: μ=0.05 h^{-1} + constant MetOH=3 g L^{-1}.

The use of methanol as a co-substrate did not modify the growth curves substantially, in comparison to the methanol-free processes, due to the low biomass-to-methanol yield, $Y_{X/MetOH}$*, and the slow specific methanol uptake rate, q_{MetOH}, of the *P. pastoris* Mut^S phenotype. Moreover, CYP2C9/CPR production seemingly did not alter cell physiology rates, as indicated by the specific glycerol uptake rate (q_{Gly}) and biomass-to-substrate yield ($Y_{X/Gly}$*) (Figure 7.6B). The obtained values were similar to those previously reported as standard for *P. pastoris* strains producing different recombinant proteins (Garcia-Ortega *et al.*, 2013; Looser *et al.*, 2017; Rajamanickam *et al.*, 2017).

Production of CYP2C9/CPR in *P. pastoris* whole cells

Regarding biocatalyst production and its optimization, production-time profiles were analysed for the five alternative FB strategies. From the results obtained (Figure 7.7), a significant difference in terms of biocatalyst production could be observed between the methanol-free fed-batches (FB 1-3) and the methanol-based ones (FB 4-5).

In methanol-free processes, a saturation profile is observed at the lowest specific growth rate tested —0.05^{-1}, 0.10 h^{-1}, whereas at the highest μ tested —0.15 h^{-1}— a growth associated production profile was observed. On the contrary, the methanol induced processes present an exponential trend up to 65 AU mL^{-1}, thus reaching biocatalyst activity relevantly higher than methanol-free fermentations. Specifically, when further inducing the system with methanol a 2-fold biocatalyst activity increase could be observed.

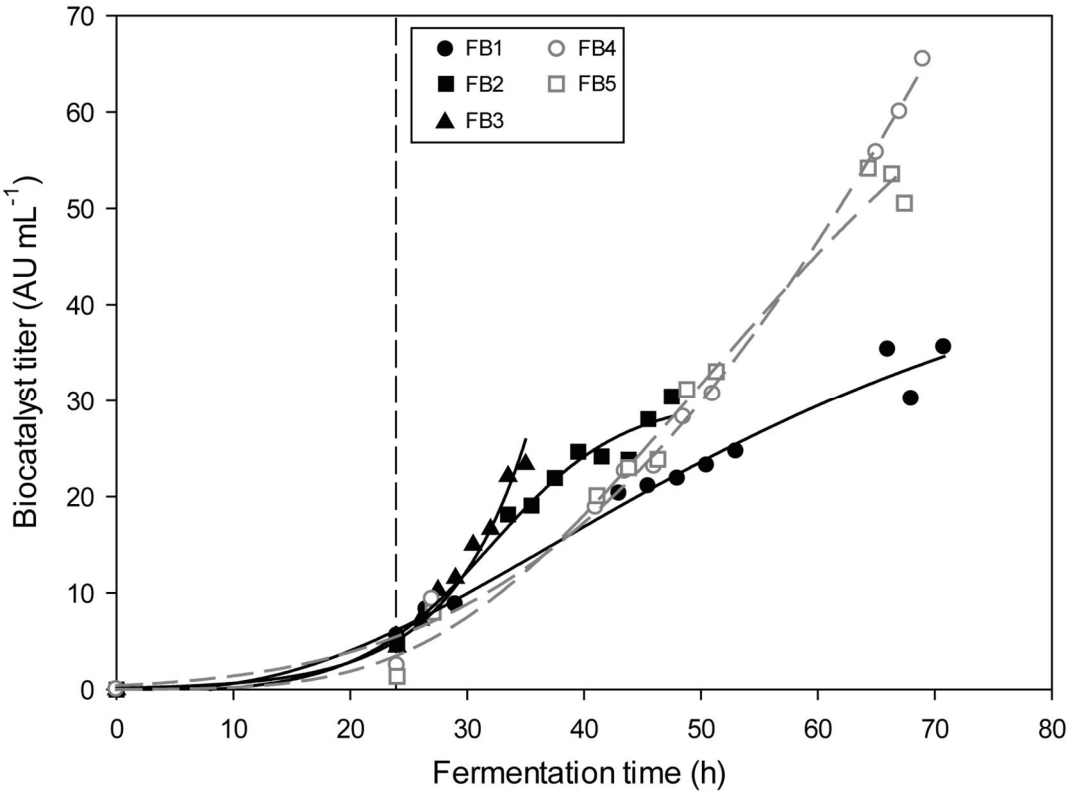


Figure 7.7. Biocatalyst titer time-profiles for fed-batch (FB) cultures 1-5. Experiment labels: FB1: μ=0.05h^{-1}; FB2: μ=0.10 h^{-1}; FB3: μ=0.15 h^{-1}; FB4: μ=0.05 h^{-1} + MetOH pulses; FB5: μ=0.05 h^{-1} + constant MetOH=3 g L^{-1}. Dashed line defines the feeding start.

As presented in Figure 7.7, a significant but low amount of active biocatalyst was already obtained during the batch phase in all five cultures. This fact is an effect of the de-repressed expression in non-limited glycerol conditions of the promoters used, P_{PDF}/P_{PDC}. However, later, on glycerol-limiting conditions the expression is boosted as expected. The similar starting point at the end of the batches carried out demonstrates also the high reproducibility between the different bioreactor runs.

It has been widely described that when expressing a heterologous protein under the regulation of several different promoters the q_p is importantly affected by μ, i.e., it can be frequently considered that production is coupled to cell growth (Jungo *et al.*, 2006; Garcia-Ortega *et al.*, 2013, 2016; Rebnegger *et al.*, 2014). As depicted in Figure 7.8, the q_p is clearly correlated with μ within the experimental range tested. Results point out a

slight q_p raise when increasing the specific growth rate from 0.05 h^{-1} (FB1) to 0.10 h^{-1} (FB2), but a bigger increase have been found when growing culture at 0.15 h^{-1} (FB3), reaching up to 2.5-fold q_p increase respect to the worse condition. In contrast, the μ effect on $Y_{P/X}$ values might be considered negligible as yield values are rather similar for all three fed-batches around, 230 AU g_{MetOH}^{-1} (FB 1-3).

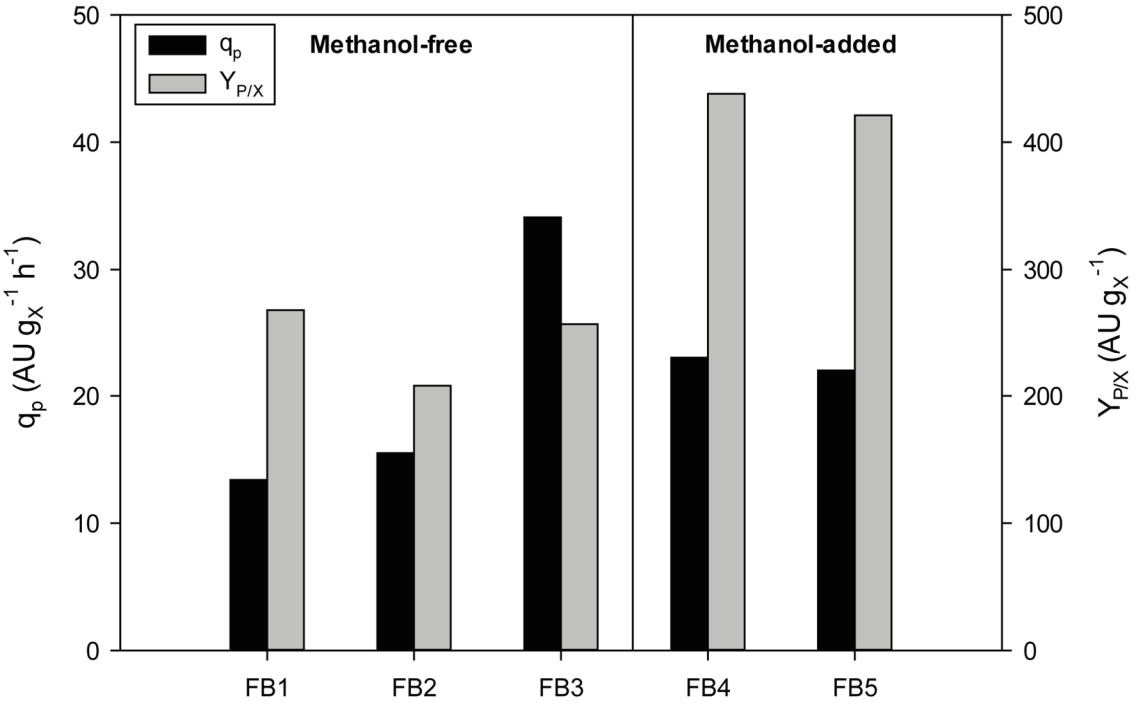


Figure 7.8. Product-related parameters: specific product generation rate (q_p) and product-to-biomass yield ($Y_{P/X}$) obtained from fed-batch cultivations. Experiment labels: FB1: μ=0.05h^{-1}; FB2: μ=0.10 h^{-1}; FB3: μ=0.15 h^{-1}; FB4: μ=0.05 h^{-1} + MetOH pulses; FB5: μ=0.05 h^{-1} + constant MetOH=3 g L^{-1}.

Relating to the effect of methanol utilization on biocatalyst production, bioreactor cultivations with the same μ (FB1, FB4 and FB5) but either adding or not methanol were compared. In all the cultures, the use of this co-substrate further induced the protein expression, significantly increasing the biocatalyst production rates. More specifically, it led to q_p improvements of up to 2.5-fold, and up to 3-fold of $Y_{P/X}$. On the one hand, this can be mainly explained because the promoters used in the producer strains are highly

inducible by methanol. However, it could also be attributed to other differences in carbon source dependent systems-level organization, such as for example cofactor regeneration, membrane permeability, or the expression levels of other proteins (Rußmayer *et al.*, 2015). As a summary, a comparison of the main process parameters is presented in Table 7.1. When comparing both methanol-based culture strategies implemented, strikingly, controlling the methanol concentration in the culture broth (FB5) only provides an improvement of rather 20% in active biocatalyst produced per gram of methanol fed ($Y_{P/MetOH}$) compared to the methanol-pulses based strategy (FB4). Moreover, other product-related parameters —q_p, $Y_{P/X}$— are only slightly lower when controlling the residual methanol concentration at a constant value. Consequently, since the results obtained are similar, the methanol pulses-based process should be considered preferable due to its technical simplicity comparing with a methanol closed-loop control cultivation strategy.

Table 7.1. Main process parameters obtained in the different FB cultures performed for the production of CYP2C9/CPR in *P. pastoris* whole cells.

μ	$Y_{X/S}$*	q_{Gly}	q_{MetOH}	*Titer*	q_p	$Y_{P/X}$*	$Y_{P/MetOH}$*
h^{-1}	$g_X\, g_S^{-1}$	$g_{Gly}\, g_X^{-1}\, h^{-1}$	$g_{MethOH}\, g_X^{-1}\, h^{-1}$	$AU\, mL^{-1}$	$AU\, g_X^{-1}\, h^{-1}$	$AU\, g_X^{-1}$	$AU\, g_{MetOH}^{-1}$
FB1	0.47	0.102	-	35.6	13.4	268	-
FB2	0.51	0.159	-	30.4	15.5	208	-
FB3	0.46	0.272	-	23.4	34.1	257	-
FB4	0.55	0.084	0.029	65.5	23.0	438	793
FB5	0.53	0.087	0.023	53.6	22.0	421	956

FB1: 0.05 h-1; **FB2**: 0.10 h-1; **FB3**: 0.15 h-1; **FB4**: 0.05 h-1 + MetOH pulses; **FB5**: 0.05 h-1 + constant [MetOH].

In order to maximize bioprocess efficiency, different optimization criteria based on diverse performance indexes can be considered; among them, biocatalyst titer, productivity and yield are the most often selected. In *P. pastoris* fermentation processes, the culture is often limited by the maximum amount of biomass that can be reached, about 100 g/L in DCW. Then, the optimization method should aim to maximise the production up to reaching this limiting cell concentration. To maximise the product titer (P) and yield ($Y_{P/X}$), the process should be prolonged in time, growing the cells not only at a low μ, but also adding methanol as co-substrate, preferably through pulses. In this way, up to 65 AU mL^{-1} could be achieved, this representing a 3-fold increase respect to the worst condition. On the contrary, if q_p is selected as the parameter to be optimized, the best operational strategy would be a carbon-limiting fed-batch operation, with glycerol as a sole carbon source, setting the specific growth rate at the highest value tested, μ=0.15 h^{-1} (FB3) being this a 3-fold increase respect to worst conditions, which was achieved in FB1.

However, from an industrial point of view, in order to select the optimal bioprocess strategy to produce active biocatalyst, the whole process including downstream should also be taken into account. Since this bioprocess approach uses whole cells, downstream processing can be considered quite simple and not costly. Therefore, to reach high biocatalyst titer, which is essential to be able to afford high cost separations, should not be considered as main priority. Thus, taking into consideration the different criteria discussed previously, the methanol-free C-limiting strategy based on glycerol at highest μ could be considered the best alternative that allows maximization of q_p. Despite the generally lower titer levels reached, avoiding the utilisation of methanol would allow to circumvent the important methanol-associated drawbacks, such as difficulties in storage, transportation and handling and further increase the process feasibility. In any case, a detailed economic analysis of the production process should be necessary in order to

clearly define the best alternative towards the biocatalyst production process development.

Preparative scale demonstration

The drug ibuprofen is mainly converted into 3-hydroxyibuprofen, 3-carboxyibuprofen, and 2-hydroxyibuprofen during Phase I metabolism in the human liver (Neunzig *et al.*, 2012). CYP2C8 and CYP2C9 are the main enzymes responsible for its stereoselective hydroxylation, of which CYP2C9 is the only known CYP to produce relevant amounts of 3-OH-ibuprofen as it has been shown for fission yeast by Neunzig *et al.* (Neunzig *et al.*, 2012). To evaluate the feasibility of a preparative scale production using *P. pastoris*, CYP2C9 expressing cells from bioreactor were applied in a whole cell biotransformation approach, i.e. 500 ml reaction volume and a substrate concentration of 2 mM.

As it can be observed in Figure 7.9B, within 16 h an almost linear increase in conversion to 25% of total product was observed, which further increased up to 40% within 86 hours of reaction. As expected from small scale results (data not shown), metabolite 2 (35%), which was identified as 3-OH-ibuprofen by NMR spectroscopy (data not shown), was mainly produced, while metabolite 1, was only found to an extent of 5%. No relevant amounts of 1-OH-ibu (metabolite 3) could be detected. This corresponds to a total of about 0.8 mM of products, 0.69 mM of 3-OH-ibu and 0.14 mM of 2-OH-ibu, respectively, and confirms the potential of recombinant CYP2C9 catalysed biotransformations using *P. pastoris* whole cells. The ratios between the two products (M2/M1) in terms of space-time yield and specific production rate were around 5, which was higher than previous published values (Neunzig *et al.*, 2012), confirming a high selectivity for the 3-OH-metabolite. A comparison of the main reaction parameters is presented in Table 7.2.

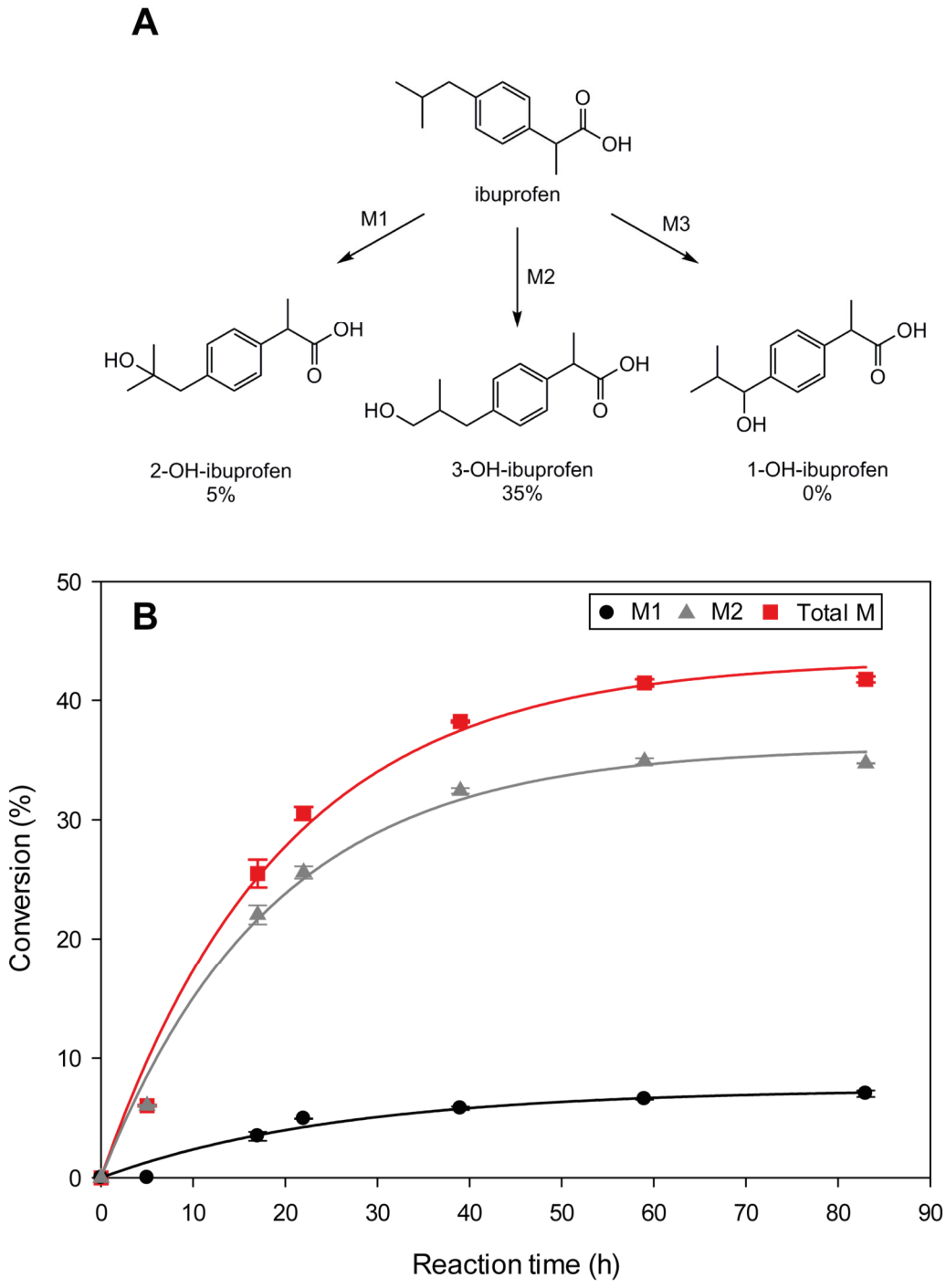


Figure 7.9. A) Ibuprofen oxidation reactions catalysed by human CYPs **B)** Preparative scale ibuprofen biooxidation by *P. pastoris* CYP2C9/CPR whole cells. Substrate concentration: 2 mM, reaction volume: 500 mL. Error bars represent the standard deviation of two analytical replicates.

Table 7.2. Main parameters obtained for the preparative scale conversion of ibuprofen.

Reaction time (h)	*Average substrate conversion (%)*	*Product formed (mM)*		*Space-time yield (µmol product L^{-1} d^{-1})*		*Specific prod. rate (µmol product g whole $cells^{-1}$ d^{-1})*		*Yield (µmol product g whole $cells^{-1}$)*	
		M1	*M2*	*M1*	*M2*	*M1*	*M2*	*M1*	*M2*
86	41.7	0.14	0.69	39	194	1.22	6.06	4.39	21.7

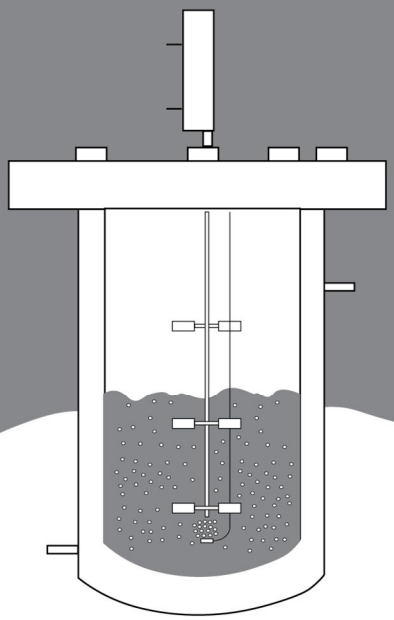
Conclusions

Recombinant production of CYPs has been intensively investigated and it is expected to continue attracting high attention by the pharmaceutical and chemical industry. In this work, a whole bioprocess for fine-tuned co-expression of CYP2C9 and CPR in *P. pastoris* has been successfully attained. To achieve this objective, an innovative fine-tunable protein expression based on bidirectional system of the co-expressed genes has been successfully implemented. Furthermore, by including several promising alternative promoters for *P. pastoris,* different combinations of innovative promoters have been evaluated in a high-throughput methodology in order to find the best expressing clones. Finally, for the first time, whole cell biocatalysts were obtained from bioreactor cultures, which therefore were tested for a biotransformation of interest based on the degradation of ibuprofen. The biocatalyst introduced in the preparative test could convert more than 40% of total substrate added to the reaction and, more importantly, it showed a high selectivity for the hydroxylated product in position three.

Furthermore, a deep study of the *P. pastoris* CYP2C9/CPR whole cells regarding the production kinetics has been made at bench top bioreactor scale. It has been demonstrated that the μ has a positive effect on q_p, supporting that the production is coupled with the cell growth. Therefore, selecting a high μ in the bioprocess is advantageous to maximize q_p on methanol-free strategies. However, if the objective is to maximise the $Y_{P/X}$ and, therefore, the biocatalyst titer, a process using mixed substrate based in the co-feeding of glycerol and methanol should be conducted. In these latter strategies μ is decreased to the lowest value tested (0.05 h^{-1}) to promote the methanol consumption. This co-substrate is then added either through pulses or by means of a closed loop control.

Altogether, it is expected that this work will contribute not only to the development of new bioprocesses and the respective catalysed biotransformations for recombinant human CYPs in *P. pastoris* but also to the production of other products of interest that could be expressed recombinantly with the novel promoters tested in this work. Finally, it can help to overcome current limitations in producing complex human enzymes and proteins anchored to the cell membrane.

CONCLUSIONS



8

Over the last two decades, the research on *Pichia pastoris* has increased exponentially, which has helped to put this yeast in the spotlight for hosting recombinant protein and metabolite production. Lately, a wealth of knowledge has been gathered about the methanol inducible P_{AOX1}, mainly regarding its strength and regulation pattern. An exhaustive study of the relationship between P_{AOX1} regulation, heterologous gene expression and recombinant protein production has been conducted using two Crl1 producing clones with different heterologous gene dosage (1 and 3 copies). From these experiments, it was concluded that μ has a relevant importance on *MIT1* expression, which has been reported as a crucial transcription factor for the regulation of the *AOX1* gene. As *CRL1* gene expression is driven by the P_{AOX1}, it was observed a clear correlation between *MIT1, AOX1*, *CRL1* transcript levels over μ in chemostat cultivations. Interestingly, a limitation in *MIT1* levels was noticed, as the *AOX1* expression was lower in the MCC than in the SCC. Regarding *CRL1* gene dosage influence, increasing *CRL1* copy number from 1 to 3 significantly boosted the *CRL1* gene expression and q_p —up to 2.2 and 2.4-fold, respectively. In addition, the clone's performance on Crl1 production was enhanced by cultivating them in fed-batch mode, where the product-related parameter values —q_p, $Y_{P/X}$— were more than twice greater than the ones showed in chemostat mode.

However, this classical expression system presents several drawbacks from an operational perspective. Consequently, a lot of efforts have been put in the investigation of methanol-free expression systems that could eventually replace P_{AOX1} for driving RPP bioprocesses. In this regard, an in-depth characterization of two novel methanol-free expression systems for RPP in *P. pastoris* has been conducted in chemostat cultivations, whose results were compared with the widely used P_{GAP}-based expression system as reference. The heterologous gene regulation of these new expression system depends on

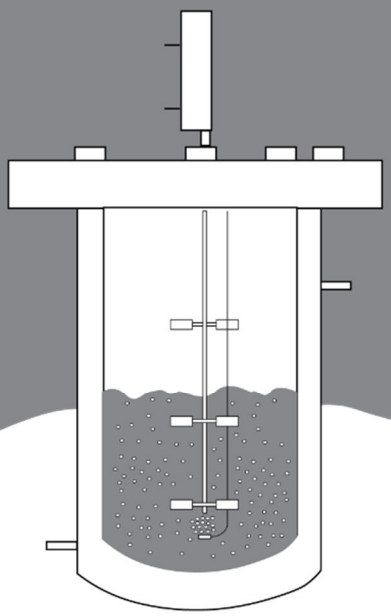
the P_{PDF} and $P_{UPP.}$ The main outcome of this study was the great performance on CalB production delivered by novel expression systems, as their q_p values significantly outperformed the ones obtained with P_{GAP}-based expression system up to 9-fold in chemostat cultivations. The outstanding CalB production outcomes obtained with the novel expression systems were mostly referred to a powerful gene expression driven by the new promoters used, in comparison to P_{GAP}. Concerning the regulation of the two novel promoters, noticeable differences were observed. The P_{UPP} showed a constitutive and growth independent regulation, since the *CALB* transcript levels remained stable regardless the μ of the cultivation, whereas the P_{PDF} exhibited a certain degree of tunability with μ. Interestingly, the transcriptional analysis of UPR sensors resulted in a striking correlation between UPR RTL and CalB q_p.

Its features led to apply the P_{PDF} in the production of a high industrially valuable catalyst such as the CYP2C9 in *P. pastoris* whole cells. This enzyme was co-expressed with its redox partner CPR. From a bidirectional promoter screening experiment, the best substrate oxidation activity was achieved by expressing the CYP2C9 and CPR under the control of the P_{PDF} and P_{CAT}, respectively. Then, the active biocatalyst production was rationally studied in fed-batch mode. Remarkably, it is the first time that a comprehensive bioprocess study is performed with a *P. pastoris* whole cell system. In methanol-free FB processes, it was observed a beneficial effect of raising μ on q_p values. On the other hand, even though the methanol-free bioprocesses provided the highest q_p values, the methanol addition allowed boosting both the active biocatalyst titer and the $Y_{P/X}$* —up to 2.8 and 2.1- fold higher, respectively. The whole cell biocatalyst obtained from the FB bioprocess conducted in bench top fermenters was directly applied to an ibuprofen proof of concept oxidation in order to test it in a reaction of industrial interest. In this regard, whole cell

biocatalyst delivered a promising performance in terms of ibuprofen conversion —40%— and selectivity —88% for the 3'OH product.

Overall, the integration of transcriptional and bioprocess analyses has been proved as successful for the extensive characterization of protein expression systems in *P. pastoris* and to maximize the RPP. This working platform has allowed to extend the knowledge about the widely studied P_{AOX1}-derived expression system as well as the novel P_{UPP} and P_{PDF}-based expression system. Furthermore, the P_{PDF} has been used for driving heterologous gene expression in the production bioprocess of an active biocatalyst for the oxidation of the broadly used ibuprofen.

BIBLIOGRAPHY



9

Ahmad, M., Hirz, M., Pichler, H., and Schwab, H. (2014) Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol 98: 5301–5317.

Ahn J, Jang MJ, Ang KS, Lee H, Choi ES, Lee DY. (2016). Codon optimization of *Saccharomyces cerevisiae* mating factor alpha prepro-leader to improve recombinant protein production in *Pichia pastoris*. Biotechnol. Lett. 38:2137–2143.

Ali S, Perez-Pardo MA, Aucamp JP, Craig A, Bracewell DG, Baganz F. (2012). Characterization and feasibility of a miniaturized stirred tank bioreactor to perform *E. coli* high cell density fed-batch fermentations. Biotechnol. Prog. 28:66–75.

Anasontzis GE, Salazar Penã M, Spadiut O, Brumer H, Olsson L. (2014). Effects of temperature and glycerol and methanol-feeding profiles on the production of recombinant galactose oxidase in *Pichia pastoris*. Biotechnol. Prog. 30:728–35.

BCC Publishing. Global Markets for Enzymes in Industrial Applications. (2018). BIO030K.

Strategic Plan for Regulatory Science (2011) FDA.

Arnau C, Casas C, Valero F. (2011). The effect of glycerol mixed substrate on the heterologous production of a *Rhizopus oryzae* lipase in *Pichia pastoris* system. Biochem. Eng. J. 57:30–37.

Ata Ö, Prielhofer R, Gasser B, Mattanovich D, Çalık P. (2017). Transcriptional engineering of the glyceraldehyde-3-phosphate dehydrogenase promoter for improved heterologous protein production in *Pichia pastoris*. Biotechnol. Bioeng. 114:2319–2327.

Ben Azoun S, Belhaj AE, Göngrich R, Gasser B, Kallel H. (2016). Molecular optimization of rabies virus glycoprotein expression in *Pichia pastoris*. Microb. Biotechnol. 9:355–368.

Ben Azoun S, Eya Belhaj A, Gasser B, Kallel H. Molecular optimization of rabies virus glycoprotein expression in *Pichia pastoris*. (2016) 9:355-368.

Baeshen MN, Al-Hejin AM, Bora RS, Ahmed MMM, Ramadan HAI, Saini KS, Baeshen NA, Redwan EM. 2015. Production of biopharmaceuticals in *E. Coli*: Current scenario and future perspectives. (2015) J. Microbiol. Biotechnol. Korean Society for Microbiolog and Biotechnology. 25:953-962.

Baghban R, Farajnia S, Rajabibazl M, Ghasemi Y, Mafi AA, Hoseinpoor R, Rahbarnia L, Aria M. (2019). Yeast Expression Systems: Overview and Recent Advances. Mol. Biotechnol. (2019). 61:365-384.

Bandaranayake AD, Almo SC. 2014. Recent advances in mammalian protein production. FEBS Lett. (2014). 588:253-260.

Baratti J, Couderc R. (1980). Oxidation of methanol by the yeast *Pichia pastoris*. Purification and properties of the alcohol oxidase+. Agric. Biol. Chem. 44:2279–2289.

Barcelos MCS, Lupki FB, Campolina GA, Nelson DL, Molina G. (2018). The colors of biotechnology: general overview and developments of white, green and blue areas. FEMS Microbiol. Lett. 365:21.

Bareither R, Bargh N, Oakeshott R, Watts K, Pollard D. (2013). Automated disposable small scale reactor for high throughput bioprocess development: A proof of concept study. Biotechnol. Bioeng. 110:3126–3138.

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. 17:161.

Barrigón JM, Montesinos JL, Valero F. (2013). Searching the best operational strategies for *Rhizopus oryzae* lipase production in *Pichia pastoris* Mut+ phenotype:

Methanol limited or methanol non-limited fed-batch cultures? Biochem. Eng. J. 75:47–54.

Barrigón JM, Valero F, Montesinos JL. (2015). A macrokinetic model-based comparative meta-analysis of recombinant protein production by *Pichia pastoris* under *AOX1* promoter. Biotechnol. Bioeng. 112:1132–1145.

Baumann K, Carnicer M, Dragosits M, Graf AB, Stadlmann J, Jouhten P, Maaheimo H, Gasser B, Albiol J, Mattanovich D, Ferrer P. (2010). A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. BMC Syst. Biol. 4:141.

Ben Azoun S, Ben Zakour M, Sghaier S, Kallel H. (2017). Expression of rabies virus glycoprotein in the methylotrophic yeast *Pichia pastoris*. Biotechnol. Appl. Biochem. 64:50–61.

Van Den Berg MA, Maruthachalam K (2016). Genetic transformation systems in fungi. Fungal Biology. 1.

Bernales S, Papa FR, Walter P. (2006). Intracellular Signaling by the Unfolded Protein Response. Annu. Rev. Cell Dev. Biol. 22:487–508.

Bernhardt R. (2006). Cytochromes P450 as versatile biocatalysts. J. Biotechnol. 124:128–145.

Bernhardt R, Urlacher VB. (2014). Cytochromes P450 as promising catalysts for biotechnological application: Chances and limitations. Appl. Microbiol. Biotechnol. 98:6185-6203.

Berrios J, Flores M-O, Díaz-Barrera A, Altamirano C, Martínez I, Cabrera Z. (2017). A comparative study of glycerol and sorbitol as co-substrates in methanol-induced cultures of *Pichia pastoris*: temperature effect and scale-up simulation. J. Ind. Microbiol. Biotechnol. 44:407–411.

Betts JI, Baganz F. (2006). Miniature bioreactors: Current practices and future opportunities. Microb. Cell Fact. 5: 21.

Bhattacharya SS, Yadav JS. (2017). Microbial P450 Enzymes in Bioremediation and Drug Discovery: Emerging Potentials and Challenges. Curr. Protein Pept. Sci. 19: 75-86.

Blazeck J, Alper H. (2010). Systems metabolic engineering: Genome-scale models and beyond. Biotechnol. J. 5:647-659.

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

Braren I, Greunke K, Umland O, Deckers S, Bredehorst R, Spillner E. (2007). Comparative expression of different antibody formats in mammalian cells and *Pichia pastoris*. Biotechnol. Appl. Biochem. 47:205-214.

Butz JA, Niebauer RT, Robinson AS. (2003). Co-expression of molecular chaperones does not improve the heterologous expression of mammalian G-protein coupled receptor expression in yeast. Biotechnol. Bioeng. 84:292–304.

Byrne B. (2015). *Pichia pastoris* as an expression host for membrane protein structural biology. Curr. Opin. Struct. Biol. 32:9–17.

Çalık P, Ata Ö, Güneş H, Massahi A, Boy E, Keskin A, Öztürk S, Zerze GH, Özdamar TH. (2015). Recombinant protein production in *Pichia pastoris* under glyceraldehyde-3-phosphate dehydrogenase promoter: From carbon source metabolism to bioreactor operation parameters. Biochem. Eng. J. 95:20–36.

Cámara E, Albiol J, Ferrer P. (2016). Droplet digital PCR-aided screening and characterization of *Pichia pastoris* multiple gene copy strains. Biotechnol. Bioeng. 113:1542–1551.

Cámara E, Landes N, Albiol J, Gasser B, Mattanovich D, Ferrer P. (2017). Increased dosage of *AOX1* promoter-regulated expression cassettes leads to transcription attenuation of the methanol metabolism in *Pichia pastoris*. Sci. Repor. 7:44302.

Cámara E, Monforte S, Albiol J, Ferrer P. (2019). Deregulation of methanol metabolism reverts transcriptional limitations of recombinant *Pichia pastoris* (*Komagataella spp*) with multiple expression cassettes under control of the *AOX1* promoter. Biotechnol. Bioeng. 116:1710-1720.

Canales C, Altamirano C, Berrios J. (2015). Effect of dilution rate and methanol-glycerol mixed feeding on heterologous *Rhizopus oryzae* lipase production with *Pichia pastoris* Mut+ phenotype in continuous culture. Biotechnol. Prog. 31:707–714.

Capone S, Horvat J, Herwig C, Spadiut O. (2015). Development of a mixed feed strategy for a recombinant *Pichia pastoris* strain producing with a de-repression promoter. Microb. Cell Fact. 14: 101.

Casas-Godoy L, Duquesne S, Bordes F, Sandoval G, Marty A. (2012). Lipases: An overview. Methods Mol. Biol. 861:3-30.

Çelik E, Çalik P. (2012). Production of recombinant proteins by yeast cells. Biotechnol. Adv. 30:1108-1118.

Cereghino GPL, Cereghino JL, Ilgen C, Cregg JM. (2002). Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. Curr Opin Biotech 13:329–332.

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

Chang S-W, Lee G-C, Shaw J-F. (2006). Codon Optimization of *Candida rugosa* lip1 Gene for Improving Expression in *Pichia pastoris* and Biochemical Characterization of the Purified Recombinant LIP1 Lipase. 54:815-822.

Chao R, Yuan Y, Zhao H. (2015). Recent advances in DNA assembly technologies. FEMS Yeast Res. 15:1-9.

Chemistry B, Biotechnology I, United T, Convention N, Diversity B, Ministry P, Education H. (2012). science Rainbow code of biotechnology:814–816.

Chen GQ. (2012). New challenges and opportunities for industrial biotechnology. Microb. Cell Fact. 11:111.

Chen GQ, Kazlauskas R. (2011). Chemical biotechnology in progress. Curr. Opin. Biotechnol. 22:747-748.

Cheng J, Maier KC, Avsec Ž, Rus P, Gagneur J. (2017). Cis-regulatory elements explain most of the mRNA stability variation across genes in yeast. RNA 23:1648–1659.

Chiruvolu V, Cregg JM, Meagher MM. (1997). Recombinant protein production in an alcohol oxidase-defective strain of *Pichia pastoris* in fedbatch fermentations. Enzyme Microb. Technol. 21:277–283.

Claes K, Vandewalle K, Laukens B, Laeremans T, Vosters O, Langer I, Parmentier M, Steyaert J, Callewaert N. (2016). Modular Integrated Secretory System Engineering in *Pichia pastoris* To Enhance G-Protein Coupled Receptor Expression. ACS Synth. Biol. 5:1070–1075.

Cos O, Ramón R, Montesinos J, Valero F. (2006). Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. Microb. Cell Fact. 5:17.

Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. (2005). Combined effect of the methanol utilization (Mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. J. Biotechnol. 116:321–335.

Cragg GML, Kingston DGI, Newman DJ. (2012). Anticancer agents from natural products. CRC Press. 2n edition.

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

Cregg JM, Tolstorukov I, Kusari A, Sunga J, Madden K, Chappell T. (2009). Chapter 13: Expression in the Yeast *Pichia pastoris*. Methods Enzymol. 463:169–189.

Curvers S, Brixius P, Klauser T, Thommes J, Weuster-Botz D, Takors R, Wandrey C. (2001). Human Chymotrypsinogen B Production with *Pichia pastoris* by Integrated Development of Fermentation and Downstream Processing. Part 1. Fermentation. Biotechnol. Prog. 17:495–502.

Dai J, Yoon SH, Sim HY, Yang YS, Oh TK, Kim JF, Hong JW. (2013). Charting microbial phenotypes in multiplex nanoliter batch bioreactors. Anal. Chem. 85:5892–5899.

Daly A, Rettie A, Fowler D, Miners J. (2017). Pharmacogenomics of CYP2C9: Functional and Clinical Considerations. J. Pers. Med. 8:1.

Daly R, Hearn MTW, Hearn MTW. (2005). Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. J. Mol. Recognit. J. Mol. Recognit 18:119–138.

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. 74:381–389.

Damasceno LM, Pla I, Chang HJ, Cohen L, Ritter G, Old LJ, Batt CA. (2004). An optimized fermentation process for high-level production of a single-chain Fv antibody fragment in *Pichia pastoris*. Protein Expr. Purif. 37:18–26.

Damasceno LM, Huang CJ, Batt CA. (2012) Protein secretion in *Pichia pastoris* and advances in protein production. Appl. Microbiol. Biotechnol. 93:31-39.

Delic M, Göngrich R, Mattanovich D, Gasser B. (2014). Engineering of protein folding and secretion - Strategies to overcome bottlenecks for efficient production of recombinant proteins. Antioxidants Redox Signal. 21:414-437.

Dietzsch C, Spadiut O, Herwig C. (2011). A fast approach to determine a fed batch feeding profile for recombinant *Pichia pastoris* strains. Microb. Cell Fact. 10:85.

Dong JJ, Fernández-Fueyo E, Hollmann F, Paul CE, Pesic M, Schmidt S, Wang Y, Younes S, Zhang W. (2018). Biocatalytic Oxidation Reactions: A Chemist's Perspective. Angew. Chemie. Int. Ed. Eng. 57:9238-9261.

Doran PM. 2013. Bioprocess Engineering Principles. 2n edition. Elsevier Ed.

Drăgan C-A, Peters FT, Bour P, Schwaninger AE, Schaan SM, Neunzig I, Widjaja M, Zapp J, Kraemer T, Maurer HH, Bureik M. (2011). Convenient gram-scale metabolite synthesis by engineered fission yeast strains expressing functional human P450 systems. Appl. Biochem. Biotechnol. 163:965–80.

Dumont J, Euwart D, Mei B, Estes S, Kshirsagar R. (2016). Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit. Rev. Biotechnol. 36:1110-1122.

Dyson MR. (2016). Fundamentals of expression in mammalian cells. Adv. Exp. Med. Biol. 896:217–224.

Erickson B, Nelson, Winters P. (2012). Perspective on opportunities in industrial biotechnology in renewable chemicals. Biotechnol. J. 7:176-185.

Esmonde-White KA, Cuellar M, Uerpmann C, Lenain B, Lewis IR. (2017). Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing. Anal. Bioanal. Chem. 409:637-649.

Fei L, Wang Y, Chen S. (2009). Improved glutathione production by gene expression in *Pichia pastoris*. Bioprocess Biosyst. Eng. 32:729–735.

Fidan O, Zhan J. (2015). Recent advances in engineering yeast for pharmaceutical protein production. RSC Adv. 5:86665–86674.

Filho DG, Silva AG, Guidini CZ. (2019). Lipases: sources, immobilization methods, and industrial applications. Appl. Microbiol. Biotechnol. 103:7399-7423.

Fischer JE, Glieder A. (2019). Current advances in engineering tools for *Pichia pastoris*. Curr. Opin. Biotechnol. 59:175–181.

Fischer JE, Hatzl AM, Weninger A, Schmid C, Glieder A. (2019). Methanol Independent Expression by < *Pichia pastoris* Employing De-repression Technologies. J. Vis. Exp:e58589.

Fischer R, Holland T, Sack M, Schillberg S, Stoger E, Twyman RM, Buyel JF. (2018). Glyco-Engineering of Plant-Based Expression Systems. Adv. Biochem. Eng. Biotechnol.

Forman V, Bjerg-Jensen N, Dyekjær JD, Møller BL, Pateraki I. (2018). Engineering of CYP76AH15 can improve activity and specificity towards forskolin biosynthesis in yeast. Microb. Cell Fact. 17:181.

Galanie S, Thodey K, Trenchard IJ, Interrante MF, Smolke CD. (2015). Complete biosynthesis of opioids in yeast. Science. 349:1095–1100.

Garcia-Ortega X, Adelantado N, Ferrer P, Montesinos JL, Valero F. (2016). A step forward to improve recombinant protein production in *Pichia pastoris*: From specific growth rate effect on protein secretion to carbon-starving conditions as advanced strategy. Process Biochem. 51:681–691.

García-Ortega X, Cámara E, Ferrer P, Albiol J, Montesinos-Seguí JL, Valero F. (2019). Rational development of bioprocess engineering strategies for recombinant protein production in *Pichia pastoris* (*Komagataella phaffii*) using the methanol-free GAP promoter. Where do we stand? N. Biotechnol. 53:24–34.

Garcia-Ortega X, Ferrer P, Montesinos JL, Valero F. (2013). Fed-batch operational strategies for recombinant Fab production with *Pichia pastoris* using the constitutive GAP promoter. Biochem. Eng. J. 79:172–181.

Garrigós-Martínez J, Nieto-Taype MA, Gasset-Franch A, Montesinos-Seguí JL, Garcia-Ortega X, Valero F. (2019). Specific growth rate governs *AOX1* gene expression, affecting the production kinetics of *Pichia pastoris* (*Komagataella phaffii*) P_{AOX1}-driven recombinant producer strains with different target gene dosage. Microb. Cell Fact. 18:187.

Gasser B, Maurer M, Gach J, Kunert R, Mattanovich D. (2006). Engineering of *Pichia pastoris* for improved production of antibody fragments. Biotechnol. Bioeng. 94:353–361.

Gebhardt G, Hortsch R, Kaufmann K, Arnold M, Weuster-Botz D. (2011). A new microfluidic concept for parallel operated milliliter-scale stirred tank bioreactors. Biotechnol. Prog. 27:684–690.

Geier M, Brandner C, Strohmeier GA, Hall M, Hartner FS, Glieder A. (2015). Engineering *Pichia pastoris* for improved NADH regeneration: A novel chassis strain for whole-cell catalysis. J. Org. Chem 11:1741–1748.

Geier M, Braun A, Emmerstorfer A, Pichler H, Glieder A. (2012). Production of human cytochrome P450 2D6 drug metabolites with recombinant microbes - a comparative study. Biotechnol. J. 7:1346–1358.

Gellissen G. (2005). Production of recombinant proteins: novel microbial and eukaryotic expression systems. Wiley-VCH.

Gleeson MA, Sudbery PE. (1988). The methylotrophic yeasts. Yeast 4:1–15.

Goodrick JC, Xu M, Finnegan R, Schilling BM, Schiavi S, Hoppe H, Wan NC. (2001). High-level expression and stabilization of recombinant human chitinase produced in a continuous constitutive *Pichia pastoris* expression system. Biotechnol. Bioeng. 74:492–497.

Graf A, Dragosits M, Gasser B, Mattanovich D. (2009). Yeast systems biotechnology for the production of heterologous proteins. FEMS Yeast Res. 9:335–348.

Gu L, Zhang J, Liu B, Du G, Chen J. (2014). High-Level Extracellular Production of Glucose Oxidase by Recombinant *Pichia pastoris* Using a Combined Strategy. Appl. Biochem. Biotechnol. 175:1429–1447.

Guengerich FP. (2015). Human Cytochrome P450 Enzymes. Cytochrome P450 - Struct. Mech. Biochem. 523–785.

Guerfal M, Ryckaert S, Jacobs PP, Ameloot P, Van Craenenbroeck K, Derycke R, Callewaert N. (2010). The *HAC1* gene from *Pichia pastoris*: Characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins. Microb. Cell Fact. 9:49.

Gupta SK, Dangi AK, Smita M, Dwivedi S, Shukla P. (2019). Effectual Bioprocess Development for Protein Production. Appl. Microbiol. Bioeng. 203–227.

Gupta SK, Shukla P. (2016). Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. Crit. Rev. Biotechnol. 36:1089:1098.

Gupta SK, Shukla P. (2017). Sophisticated cloning, fermentation, and purification technologies for an enhanced therapeutic protein production: A review. Front. Pharmacol. 8:419.

Gurramkonda C, Polez S, Skoko N, Adnan A, Gäbel T, Chugh D, Swaminathan S, Khanna N, Tisminetzky S, Rinas U. (2010). Application of simple fed-batch technique to high-level secretory production of insulin precursor using *Pichia pastoris* with subsequent purification and conversion to human insulin. Microb. Cell Fact. 9:31.

Hackel BJ, Huang D, Bubolz JC, Wang XX, Shusta E V. (2006). Production of soluble and active transferrin receptor-targeting single-chain antibody using *Saccharomyces cerevisiae*. Pharm. Res. 23:790–797.

Hang HF, Chen W, Guo MJ, Chu J, Zhuang YP, Zhang S. (2008). A simple unstructured model-based control for efficient expression of recombinant porcine insulin precursor by *Pichia pastoris*. Korean J. Chem. Eng. 25:1065–1069.

Hartner FS, Glieder A. (2006). Regulation of methanol utilisation pathway genes in yeasts. Microb. Cell Fact. 5:39.

Hartner FS, Ruth C, Langenegger D, Johnson SN, Hyka P, Lin-Cereghino GP, Lin-Cereghino J, Kovar K, Cregg JM, Glieder A. (2008). Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. Nucleic Acids Res. 36:e76.

Hasan F, Shah AA, Hameed A. (2006). Industrial applications of microbial lipases. Enzyme Microb. Technol. 39:235–251.

Hasslacher M, Schall M, Hayn M, Bona R, Rumbold K, Lückl J, Griengl H, Kohlwein SD, Schwab H. (1997). High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* in microbial hosts. Protein Expr. Purif. 11:61–71.

Hausjell J, Halbwirth H, Spadiut O. (2018). Recombinant production of eukaryotic cytochrome P450s in microbial cell factories. Biosci. Rep. 38:2.

Havlik D, Brandt U, Bohle K, Fleißner A. (2017). Establishment of *Neurospora crassa* as a host for heterologous protein production using a human antibody fragment as a model product. Microb. Cell Fact. 16:128.

Van der Heijden RTJM, Heijnen JJ, Hellinga C, Romein B, Luyben. (1994). Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. Biotechnol. Bioeng. 43:3-10.

Helenius A, Marquardt T, Braakman I. (1992). The endoplasmic reticulum as a protein-folding compartment. Trends Cell Biol. 2:227–231.

Hemmerich J, Adelantado N, Barrigón JM, Ponte X, Hörmann A, Ferrer P, Kensy F, Valero F. (2014). Comprehensive clone screening and evaluation of fed-batch

strategies in a microbioreactor and lab scale stirred tank bioreactor system: Application on *Pichia pastoris* producing *Rhizopus oryzae* lipase. Microb. Cell Fact. 13:36.

Hendershot L, Bole D, Kohler G, Kearney JF. (1987). Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein. J. Cell Biol. 104:761–767.

Hesketh AR, Castrillo JI, Sawyer T, Archer DB, Oliver SG. (2013). Investigating the physiological response of *Pichia* (*Komagataella*) *pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures. Appl. Microbiol. Biotechnol. 97:9747–9762.

Heyland J, Fu J, Blank LM, Schmid A. (2010). Quantitative physiology of *Pichia pastoris* during glucose-limited high-cell density fed-batch cultivation for recombinant protein production. Biotechnol. Bioeng. 107:357–368.

Hoffmeister D, Keller NP. (2007). Natural products of filamentous fungi: Enzymes, genes, and their regulation. Nat. Prod. Rep. 24:393-416.

Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D. (2004). Effects of Gene Dosage, Promoters, and Substrates on Unfolded Protein Stress of Recombinant *Pichia pastoris*. Biotechnol. Bioeng. 85:367–375.

Houde A, Kademi A, Leblanc D. (2004). Lipases and their industrial applications: An overview. Appl. Biochem. Biotechnol.118:155–170.

Hu XQ, Chu J, Zhang Z, Zhang SL, Zhuang YP, Wang YH, Guo MJ, Chen HX, Yuan ZY. (2008). Effects of different glycerol feeding strategies on S-adenosyl-l-methionine biosynthesis by P_{GAP}-driven *Pichia pastoris* overexpressing methionine adenosyltransferase. J. Biotechnol. 137:44–49.

Hu Y, Daoud W, Cheuk K, Lin C. 2016. Newly Developed Techniques on Polycondensation, Ring-Opening Polymerization and Polymer Modification: Focus on Poly(Lactic Acid). Materials (Basel). 9:133. http://www.mdpi.com/1996-1944/9/3/133.

Ikonomou L, Schneider YJ, Agathos SN. (2003). Insect cell culture for industrial production of recombinant proteins. Appl. Microbiol. Biotechnol. 62:1-20.

Inan M, Aryasomayajula D, Sinha J, Meagher MM. (2006). Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. Biotechnol. Bioeng. 93:771–778.

Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N. (2009). Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. Nat. Protoc. 4:58–70.

Jaeger KE, Eggert T. (2002). Lipases for biotechnology. Curr. Opin. Biotechnol. 13:390-397.

Jakočiūnas T, Jensen ED, Jensen MK, Keasling JD. (2018). Assembly and Multiplex Genome Integration of Metabolic Pathways in Yeast Using CasEMBLR. Methods Mol. Biol.1671:185–201.

Janakiraman V, Kwiatkowski C, Kshirsagar R, Ryll T, Huang YM. (2015). Application of high-throughput mini-bioreactor system for systematic scale-down modeling, process characterization, and control strategy development. Biotechnol. Prog. 31:1623–1632.

Jarvis DL, Fleming JGW, Kovacs GR, Summers MD, Guarino LA. (1990). Use of early raculovirus promoters for continuous expression and efficient processing of foreign gene products in stably transformed lepidopteran cells. Bio/Technology 8:950–955.

Jigami Y, Odani T. (1999). Mannosylphosphate transfer to yeast mannan. Biochim. Biophys. Acta - Gen. Subj. 1426:335-345.

Jin Z, Han SY, Zhang L, Zheng SP, Wang Y, Lin Y. (2013). Combined utilization of lipase-displaying *Pichia pastoris* whole-cell biocatalysts to improve biodiesel production in co-solvent media. Bioresour. Technol. 130:102–109.

Johnson MA, Waterham HR, Ksheminska GP, Fayura LR, Cereghino JL, Stasyk O V., Veenhuis M, Kulachkovsky AR, Sibirny AA, Cregg JM. (1999). Positive selection of novel peroxisome biogenesis-defective mutants of the yeast *Pichia pastoris*. Genetics 151:1379–1391.

Jordà J, De Jesus SS, Peltier S, Ferrer P, Albiol J. (2014). Metabolic flux analysis of recombinant *Pichia pastoris* growing on different glycerol/methanol mixtures by iterative fitting of NMR-derived 13C-labelling data from proteinogenic amino acids. N. Biotechnol. 31:120–132.

Jungo C, Rérat C, Marison IW, Von Stockar U. (2006). Quantitative characterization of the regulation of the synthesis of alcohol oxidase and of the expression of recombinant avidin in a *Pichia pastoris* Mut+ strain. Enzyme Microb. Technol. 39:936–944.

Juturu V, Wu JC. (2018). Heterologous Protein Expression in *Pichia pastoris*: Latest Research Progress and Applications. ChemBioChem 19:7–21.

Kamm B, Kamm M, Gruber PR, Kromus S. (2008). Biorefinery Systems - An Overview. Biorefineries-Industrial Process. Prod. Status Quo Futur. Dir. 1:1–40.

Kang HA, Kang W, Hong WK, Kim MW, Kim JY, Sohn JH, Choi ES, Choe KB, Rhee SK. (2001). Development of expression systems for the production of recombinant

human serum albumin using the *MOX* promoter in *Hansenula polymorpha* DL-1. Biotechnol. Bioeng. 76:175–185.

Karaoglan M, Karaoglan FE, Inan M. (2016). Comparison of *ADH3* promoter with commonly used promoters for recombinant protein production in *Pichia pastoris*. Protein Expr. Purif. 121:112–117.

Khan MA, Hassan N, Ahmad N, Khan MI, Zafar AU, Khan F, Husnain T. (2014). Studies to analyse the relationship between IFNα2b gene dosage and its expression, using a *Pichia pastoris* -based expression system. Yeast 31:13–28.

Kim SJ, Lee JA, Kim YH, Song BK. (2009). Optimization of the functional expression of *Coprinus cinereus* peroxidase in *Pichia pastoris* by varying the host and promoter. J. Microbiol. Biotechnol. 19:966–971.

Klein T, Schneider K, Heinzle E. (2013). A system of miniaturized stirred bioreactors for parallel continuous cultivation of yeast with online measurement of dissolved oxygen and off-gas. Biotechnol. Bioeng. 110:535–542.

Knittler MR, Dirks S, Haas IG. (1995). Molecular chaperones involved in protein degradation in the endoplasmic reticulum: Quantitative interaction of the heat shock cognate protein BiP with partially folded immunoglobulin light chains that are degraded in the endoplasmic reticulum. Proc. Natl. Acad. Sci. U. S. A. 92:1764–1768.

Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Ohi H, Tomomitsu K, Ohmura T. (2000). High-level expression of recombinant human serum albumin from the methylotrophic yeast *Pichia pastoris* with minimal protease production and activation. J. Biosci. Bioeng. 89:55–61.

Kolewe ME, Gaurav V, Roberts SC. (2008). Pharmaceutically active natural product synthesis and supply via plant cell culture technology. Mol. Pharm. 5:243-256.

Kost TA, Condreay JP, Jarvis DL. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. Nat. Biotechnol. 23:567-575.

Kottmeier K, Günther TJ, Weber J, Kurtz S, Ostermann K, Rödel G, Bley T. (2012). Constitutive expression of hydrophobin HFB1 from *Trichoderma reesei* in *Pichia pastoris* and its pre-purification by foam separation during cultivation. Eng. Life Sci. 12:162–170.

Krainer FW, Gerstmann MA, Darnhofer B, Birner-Gruenberger R, Glieder A. (2016). Biotechnological advances towards an enhanced peroxidase production in *Pichia pastoris*. J. Biotechnol. 233:181–189.

Krainer FW, Dietzsch C, Hajek T, Herwig C, Spadiut O, Glieder A. (2012). Recombinant protein expression in *Pichia pastoris* strains with an engineered methanol utilization pathway. Microb. Cell Fact. 11:22.

Kranthi BV, Kumar HRV, Rangarajan PN. (2010). Identification of Mxr1p-binding sites in the promoters of genes encoding dihydroxyacetone synthase and peroxin 8 of the methylotrophic yeast *Pichia pastoris*. Yeast 27:705–711.

Kranthi BV, Kumar R, Kumar NV, Rao DN, Rangarajan PN. (2009). Identification of key DNA elements involved in promoter recognition by Mxr1p, a master regulator of methanol utilization pathway in *Pichia pastoris*. Biochim. Biophys. Acta - Gene Regul. Mech. 1789:460–468.

Krause M, Ukkonen K, Haataja T, Ruottinen M, Glumoff T, Neubauer A, Neubauer P, Vasala A. (2010). A novel fed-batch based cultivation method provides high cell-density and improves yield of soluble recombinant proteins in shaken cultures. Microb. Cell Fact. 9:11.

Krivoruchko A, Nielsen J. (2015). Production of natural products through metabolic engineering of *Saccharomyces cerevisiae*. Curr. Opin. Biotechnol. 35:7:15.

Kumar D, Batra J, Komives C, Rathore AS. (2019). QbD based media development for the production of fab fragments in *E. Coli*. Bioengineering. 6.

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

Landes N, Gasser B, Vorauer-Uhl K, Lhota G, Mattanovich D, Maurer M. (2016). The vitamin-sensitive promoter P_{THI11} enables pre-defined autonomous induction of recombinant protein production in *Pichia pastoris*. Biotechnol. Bioeng. 113:2633–2643.

Laursen L. (2010). How green biotech turned white and blue. Nat. Biotechnol. 28:393-395.

Lee B, Yurimoto H, Sakai Y, Kato N. (2002). Physiological role of the glutathione-dependent formaldehyde dehydrogenase in the methylotrophic yeast *Candida boidinii*. Microbiology 148:2697–2704.

Lee CY, Lee SJ, Jung KH, Katoh S, Lee EK. (2003). High dissolved oxygen tension enhances heterologous protein expression by recombinant *Pichia pastoris*. Process Biochem. 38:1147–1154.

Lee JW, Kim HU, Choi S, Yi J, Lee SY. (2011). Microbial production of building block chemicals and polymers. Curr. Opin. Biotechnol. 22:758-767.

Li D, Tang Y, Lin J, Cai W. (2017). Methods for genetic transformation of filamentous fungi. Microb. Cell Fact. 16:168.

Li P, Anumanthan A, Gao XG, Ilangovan K, Suzara V V., Düzgüneş N, Renugopalakrishnan V. (2007). Expression of recombinant proteins in *Pichia pastoris*. Appl. Biochem. Biotechnol. 142:105-124.

Liang J, Yuan J. (2006). Oxygen transfer model in recombinant *Pichia pastoris* and its application in biomass estimation. Biotechnol. Lett. 29:27–35.

Liang S, Zou C, Lin Y, Zhang X, Ye Y. (2013). Identification and characterization of P GCW14: a novel, strong constitutive promoter of *Pichia pastoris*. Biotechnol Lett. 35:1865-1871.

Lin-Cereghino GP, Stark CM, Kim D, Chang J, Shaheen N, Poerwanto H, Agari K, Moua P, Low LK, Tran N, Huang AD, Nattestad M, Oshiro KT, Chang JW, Chavan A, Tsai JW, Lin-Cereghino J. (2013). The effect of α-mating factor secretion signal mutations on recombinant protein expression in *Pichia pastoris*. Gene 519:311–317.

Lin-Cereghino GP, Godfrey L, De La Cruz BJ, Johnson S, Khuongsathiene S, Tolstorukov I, Yan M, Lin-Cereghino J, Veenhuis M, Subramani S, Cregg JM. (2006). Mxr1p, a Key Regulator of the Methanol Utilization Pathway and Peroxisomal Genes in *Pichia pastoris*. Mol. Cell. Biol. 26:883–897.

Liu J, Han Q, Cheng Q, Chen Y, Wang R, Li X, Liu Y, Yan D. (2020). Efficient Expression of Human Lysozyme Through the Increased Gene Dosage and Co-expression of Transcription Factor Hac1p in *Pichia pastoris*. Curr. Microbiol.:1–9.

Liu M, Wang B, Wang F, Yang Z, Gao D, Zhang C, Ma L, Yu X. (2019a). Soluble expression of single-chain variable fragment (scFv) in *Escherichia coli* using superfolder green fluorescent protein as fusion partner. Appl. Microbiol. Biotechnol. 103:6071–6079.

Liu Y, Huang L, Fu Y, Zheng D, Ma J, Li Y, Xu Z, Lu F. (2019b). A novel process for phosphatidylserine production using a *Pichia pastoris* whole-cell biocatalyst with

overexpression of phospholipase D from *Streptomyces halstedii* in a purely aqueous system. Food Chem. 274:535–542.

Long Q, Liu X, Yang Y, Li L, Harvey L, McNeil B, Bai Z. (2014). The development and application of high throughput cultivation technology in bioprocess development. J. Biotechnol. 192:323–338.

Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, Fotiadis D, Kovar K. (2014). Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. Biotechnol. Adv. 33:1177–1193.

Looser V, Lüthy D, Straumann M, Hecht K, Melzoch K, Kovar K. (2017). Effects of glycerol supply and specific growth rate on methanol-free production of CALB by *P. pastoris*: functional characterisation of a novel promoter. Appl. Microbiol. Biotechnol. 101:3163-3176.

De Luca V, Salim V, Atsumi SM, Yu F. (2012). Mining the biodiversity of plants: A revolution in the making. Science. 336:1658-1661.

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. (2005). Heterologous protein production using the *Pichia pastoris* expression system. Yeast. 22:249-270.

Macedo Robert J, Garcia-Ortega X, Montesinos-Seguí JL, Guimaraes Freire DM, Valero F. (2019). Continuous operation, a realistic alternative to fed-batch fermentation for the production of recombinant lipase B from *Candida antarctica* under the constitutive promoter PGK in *Pichia pastoris*. Biochem. Eng. J. 147:39–47.

Magaña-Ortíz D, Fernández F, Loske AM, Gómez-Lim MA. (2018). Extracellular Expression in *Aspergillus niger* of an Antibody Fused to *Leishmania sp.* Antigens. Curr. Microbiol. 75:40–48.

Maksimenko OG, Deykin AV, Khodarovich YM, Georgiev PG. (2013). Use of Transgenic Animals in Biotechnology: Prospects and Problems. Acta Naturae 5:33–46.

Manahan M, Nelson M, Cacciatore JJ, Weng J, Xu S, Pollard J. (2019). Scale-down model qualification of ambr® 250 high-throughput mini-bioreactor system for two commercial-scale mAb processes. Biotechnol. Prog. 35:e2870.

Marsalek L, Puxbaum V, Buchetics M, Mattanovich D, Gasser B. (2019). Disruption of vacuolar protein sorting components of the HOPS complex leads to enhanced secretion of recombinant proteins in *Pichia pastoris*. Microb. Cell Fact. 18:1.

Marx H, Mecklenbräuker A, Gasser B, Sauer M, Mattanovich D. (2009). Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. FEMS Yeast Res. 9:1260–1270.

Massahi A, Çalık P. (2018). Naturally occurring novel promoters around pyruvate branch-point for recombinant protein production in *Pichia pastoris (Komagataella phaffii*): Pyruvate decarboxylase- and pyruvate kinase- promoters. Biochem. Eng. J. 138:111–120.

Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D. (2012). Recombinant protein production in yeasts. Methods Mol. Biol. 824:329–358.

Mattanovich D, Gasser B, Hohenblum H, Sauer M. (2004). Stress in recombinant protein producing yeasts. J. Biotechnol.113:121–135.

Maurer M, Kühleitner M, Gasser B, Mattanovich D. (2006). Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris*. Microb. Cell Fact. 5:37.

McChesney JD, Venkataraman SK, Henri JT. (2007). Plant natural products: Back to the future or into extinction? Phytochemistry 68:2015–2022.

Mellitzer A, Ruth C, Gustafsson C, Welch M, Birner-Grünberger R, Weis R, Purkarthofer T, Glieder A. (2014). Synergistic modular promoter and gene optimization to push cellulase secretion by *Pichia pastoris* beyond existing benchmarks. J. Biotechnol. 191:187–195.

Mellitzer A, Weis R, Glieder A, Flicker K. (2012). Expression of lignocellulolytic enzymes in *Pichia pastoris*. Microb. Cell Fact. 11:61.

Min CK, Lee JW, Chung KH, Park HW. (2010). Control of specific growth rate to enhance the production of a novel disintegrin, saxatilin, in recombinant *Pichia pastoris*. J. Biosci. Bioeng. 110:314–319.

Minning S, Serrano A, Ferrer P, Solá C, Schmid RD, Valero F. (2001). Optimization of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. J. Biotechnol. 86:59–70.

Moon TS, Yoon SH, Lanza AM, Roy-Mayhew JD, Jones Prather KL. (2009). Production of glucaric acid from a synthetic pathway in recombinant *Escherichia coli*. Appl. Environ. Microbiol. 75:589–595.

Moura RR, Melo LM, Freitas VJ de F. (2011). Production of recombinant proteins in milk of transgenic and non-transgenic goats. Brazilian Arch. Biol. Technol. 54:927–938.

Nakano A, Lee CY, Yoshida A, Matsumoto T, Shiomi N, Katoh S. (2006). Effects of methanol feeding methods on chimeric α-amylase expression in continuous culture of *Pichia pastoris*. J. Biosci. Bioeng. 101:227–231.

Nakata PA. (2017). Construction of pDUO: A bicistronic shuttle vector series for dual expression of recombinant proteins. Plasmid 89:16–21.

Nebert DW, Wikvall K, Miller WL. (2013). Human cytochromes P450 in health and disease. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368:20120431.

Neunzig I, Göhring A, Drăgan C-A, Zapp J, Peters FT, Maurer HH, Bureik M. (2012). Production and NMR analysis of the human ibuprofen metabolite 3-hydroxyibuprofen. J. Biotechnol. 157:417–20.

Neunzig I, Widjaja M, Peters FT, Maurer HH, Hehn A, Bourgaud F, Bureik M. (2013). Coexpression of CPR from Various Origins Enhances Biotransformation Activity of Human CYPs in *S. pombe*. Appl. Biochem. Biotechnol. 170:1751–1766.

Newman DJ, Cragg GM, Snader KM. (2003). Natural products as sources of new drugs over the period 1981-2002. J. Nat. Prod. 66:1022-1037.

Nieto-Taype MA, Garcia-Ortega X, Albiol J, Montesinos-Seguí JL, Valero F. (2020). Continuous Cultivation as a Tool Toward the Rational Bioprocess Development With *Pichia pastoris* Cell Factory. Front. Bioeng. Biotechnol. 8:632.

Nieto-Taype MA, Garrigós-Martínez J, Sánchez-Farrando M, Valero F, Garcia-Ortega X, Montesinos-Seguí JL. (2019). Rationale-based selection of optimal operating strategies and gene dosage impact on recombinant protein production in *Komagataella phaffii* (*Pichia pastoris*). Microb. Biotechnol.:1751–7915.13498.

Ogata K, Nishikawa H, Ohsugi M. (1969). A yeast capable of utilizing methanol. Agric. Biol. Chem. 33:1519–1520.

Olaofe OA, Fenner CJ, Gudiminchi RK, Smit MS, Harrison ST. (2013). The influence of microbial physiology on biocatalyst activity and efficiency in the terminal hydroxylation of n-octane using *Escherichia coli* expressing the alkane hydroxylase, CYP153A6. Microb. Cell Fact. 12:8.

Orman MA, Calik P, Ozdamar TH. (2009). The influence of carbon sources on recombinant-human- growth-hormone production by *Pichia pastoris* is dependent on phenotype: a comparison of MutS and Mut+ strains. Biotechnol. Appl. Biochem. 52:245–55.

Owczarek B, Gerszberg A, Hnatuszko-Konka K. (2019). A Brief Reminder of Systems of Production and Chromatography-Based Recovery of Recombinant Protein Biopharmaceuticals. Biomed. Res. Int. 4216060.

Pal Y, Khushoo A, Mukherjee KJ. (2006). Process optimization of constitutive human granulocyte-macrophage colony-stimulating factor (hGM-CSF) expression in *Pichia pastoris* fed-batch culture. Appl. Microbiol. Biotechnol. 69:650–657.

Pandey A, Benjamin S, Soccol CR, Nigam P, Krieger N, Soccol VT. (1999). The realm of microbial lipases in biotechnology. Biotechnol. Appl. Biochem. 29 :119–131.

Panula-Perälä J, Šiurkus J, Vasala A, Wilmanowski R, Casteleijn MG, Neubauer P. (2008). Enzyme controlled glucose auto-delivery for high cell density cultivations in microplates and shake flasks. Microb. Cell Fact. 7:31.

Park JH, Lee SY, Kim TY, Kim HU. (2008). Application of systems biology for bioprocess development. Trends Biotechnol. 8:404-412.

Park KY, Wi SJ. (2016). Potential of plants to produce recombinant protein products. J. Plant Biol. 59:559-568.

Peña DA, Gasser B, Zanghellini J, Steiger MG, Mattanovich D. (2018). Metabolic engineering of *Pichia pastoris*. Metab. Eng. 50:2–15.

Pollard DJ, Kirschner TF, Hunt GR, Tong IT, Stieber R, Salmon PM. (2007). Scale up of a viscous fungal fermentation: Application of scale-up criteria with regime analysis and operating boundary conditions. Biotechnol. Bioeng. 96:307–317.

Ponte X, Barrigón JM, Maurer M, Mattanovich D, Valero F, Montesinos-Seguí JL. (2018). Towards optimal substrate feeding for heterologous protein production in *Pichia pastoris* (*Komagataella spp*) fed-batch processes under P_{AOX1} control: a modeling aided approach. J. Chem. Technol. Biotechnol. 93:3208–3218.

Ponte X, Montesinos-Seguí JL, Valero F. (2016). Bioprocess efficiency in *Rhizopus oryzae* lipase production by *Pichia pastoris* under the control of P_{AOX1} is oxygen tension dependent. Process Biochem. 51:1954–1963.

Portela RMC, Vogl T, Ebner K, Oliveira R, Glieder A. (2018). *Pichia pastoris* Alcohol Oxidase 1 (*AOX1*) Core Promoter Engineering by High Resolution Systematic Mutagenesis. Biotechnol. J. 13:1700340.

Potgieter TI, Kersey SD, Mallem MR, Nylen AC, d'Anjou M. (2010). Antibody expression kinetics in glycoengineered *Pichia pastoris*. Biotechnol. Bioeng. 106:918–927.

Potvin G, Ahmad A, Zhang Z. (2012). Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review. Biochem. Eng. J. 64:91–105.

Prielhofer R, Reichinger M, Wagner N, Claes K, Kiziak C, Gasser B, Mattanovich D. (2018). Superior protein titers in half the fermentation time: Promoter and process engineering for the glucose-regulated *GTH1* promoter of *Pichia pastoris*. Biotechnol. Bioeng. 115:2479-2488.

Puxbaum V, Mattanovich D, Gasser B. (2015). Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 99:2925-2938.

Quehl P, Hollender J, Schüürmann J, Brossette T, Maas R, Jose J. (2016). Co-expression of active human cytochrome P450 1A2 and cytochrome P450 reductase on the cell surface of *Escherichia coli*. Microb. Cell Fact. 15:26.

Rajamanickam V, Metzger K, Schmid C, Spadiut O. (2017). A novel bi-directional promoter system allows tunable recombinant protein production in *Pichia pastoris*. Microb. Cell Fact. 16:152.

Ramón A, Marín M. (2011). Advances in the production of membrane proteins in *Pichia pastoris*. Biotechnol. J. 6:700–706.

Raschmanová H, Zamora I, Borčinová M, Meier P, Weninger A, Mächler D, Glieder A, Melzoch K, Knejzlík Z, Kovar K. (2019). Single-Cell Approach to Monitor the Unfolded Protein Response During Biotechnological Processes With *Pichia pastoris*. Front. in Microb. 10:335.

Raskin I, Ribnicky DM, Komarnytsky S, Ilic N, Poulev A, Borisjuk N, Brinker A, Moreno DA, Ripoll C, Yakoby N, O'Neal JM, Cornwell T, Pastor I, Fridlender B. (2002). Plants and human health in the twenty-first century. Trends Biotechnol. 20:522–31.

Rathore AS. (2009). Roadmap for implementation of quality by design (QbD) for biotechnology products. Trends Biotechnol. 27:546-553.

Rebnegger C, Graf AB, Valli M, Steiger MG, Gasser B, Maurer M, Mattanovich D. (2014). In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. Biotechnol. J. 9:511–525.

Reis P, Holmberg K, Watzke H, Leser ME, Miller R. (2009). Lipases at interfaces: A review. 147:237-250.

Resina D, Cos O, Ferrer P, Valero F. (2005). Developing high cell density fed-batch cultivation strategies for heterologous protein production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. Biotechnol. Bioeng. 91:760–767.

Resina D, Serrano A, Valero F, Ferrer P. (2004). Expression of a *Rhizopus oryzae* lipase in *Pichia pastoris* under control of the nitrogen source-regulated formaldehyde dehydrogenase promoter. J. Biotechnol. 109:103–113.

Rinnofner C, Kerschbaumer B, Weber H, Glieder A, Winkler M. (2019). Cytochrome P450 mediated hydroxylation of ibuprofen using *Pichia pastoris* as biocatalyst. Biocatal. Agric. Biotechnol. 17:525–528.

Robert JM, Lattari FS, Machado AC, de Castro AM, Almeida RV, Torres FAG, Valero F, Freire DMG. (2017). Production of recombinant lipase B from *Candida antarctica* in *Pichia pastoris* under control of the promoter PGK using crude glycerol from biodiesel production as carbon source. Biochem. Eng. J. 118:123–131.

Roggenkamp R, Janowicz Z, Stanikowski B, Hollenberg CP. (1984). Biosynthesis and regulation of the peroxisomal methanol oxidase from the methylotrophic yeast *Hansenula polymorpha*. Mol. Gen. Genet. 194:489–493.

Rosano GL, Morales ES, Ceccarelli EA. (2019). New tools for recombinant protein production in *Escherichia coli*: A 5-year update. Protein Sci. 8:1412-1422.

Rushmore TH, Reider PJ, Slaughter D, Assang C, Shou M. (2000). Bioreactor Systems in Drug Metabolism: Synthesis of Cytochrome P450-Generated Metabolites. Metab. Eng. 2:115–125.

Rußmayer H, Buchetics M, Gruber C, Valli M, Grillitsch K, Modarres G, Guerrasio R, Klavins K, Neubauer S, Drexler H, Steiger M, Troyer C, Al Chalabi A, Krebiehl G, Sonntag D, Zellnig G, Daum G, Graf AB, Altmann F, Koellensperger G, Hann S, Sauer

M, Mattanovich D, Gasser B. (2015). Systems-level organization of yeast methylotrophic lifestyle. BMC Biol. 13:80.

Ruth C, Zuellig T, Mellitzer A, Weis R, Looser V, Kovar K, Glieder A. (2010). Variable production windows for porcine trypsinogen employing synthetic inducible promoter variants in *Pichia pastoris*. Syst. Synth. Biol. 4:181–191.

Sagt CM, Müller WH, Boonstra J, Verkleij AJ, Verrips CT. (1998). Impaired secretion of a hydrophobic cutinase by *Saccharomyces cerevisiae* correlates with an increased association with immunoglobulin heavy-chain binding protein (BiP). Appl. Environ. Microbiol. 64:316–24.

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

Sakai Y, Nakagawa T, Shimase M, Kato N. (1998). Regulation and physiological role of the DAS1 gene, encoding dihydroxyacetone synthase, in the methylotrophic yeast *Candida boidinii*. J. Bacteriol. 180:5885–5890.

Sang YL, Lee DY, Tae YK. (2005). Systems biotechnology for strain improvement. Trends Biotechnol. 23:349-358.

Sasson A. (2005). Medical biotechnology: achievements, prospects and perceptions. United Nations University Press.

Schillberg S, Raven N, Spiegel H, Rasche S, Buntru M. (2019). Critical analysis of the commercial potential of plants for the production of recombinant proteins. Front. Plant Sci. 10:720.

Schroer K, Kittelmann M, Lütz S. (2010a). Recombinant human cytochrome P450 monooxygenases for drug metabolite synthesis. Biotechnol. Bioeng. 106:699–706.

Schroer K, Peter Luef K, Stefan Hartner F, Glieder A, Pscheidt B. (2010b). Engineering the *Pichia pastoris* methanol oxidation pathway for improved NADH regeneration during whole-cell biotransformation. Metab. Eng. 12:8–17.

Schwarzhans J-P, 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.

Sha C, Yu XW, Li F, Xu Y. (2013)a. Impact of gene dosage on the production of lipase from *Rhizopus chinensis* CCTCCM201021 in *Pichia pastoris*. Appl. Biochem. Biotechnol. 169:1160–1172.

Sha C, Yu X-W, Lin N-X, Zhang M, Xu Y. (2013b). Enhancement of lipase r27RCL production in *Pichia pastoris* by regulating gene dosage and co-expression with chaperone protein disulfide isomerase. Enzyme Microb. Technol. 53:438–443.

Shao Z, Zhao H, Zhao H. (2009). DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. Nucleic Acids Res. 37:e16.

Shen Q, Wu M, Wang H Bin, Naranmandura H, Chen SQ. (2012). The effect of gene copy number and co-expression of chaperone on production of albumin fusion proteins in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 96:763–772.

Shen S, Sulter G, Jeffries TW, Cregg JM. (1998). A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. Gene 216:93–102.

Shen W, Xue Y, Liu Y, Kong C, Wang X, Huang M, Cai M, Zhou X, Zhang Y, Zhou M. (2016). A novel methanol-free *Pichia pastoris* system for recombinant protein expression. Microb. Cell Fact. 15:178.

Shi L, Wang J, Wang X, Zhang Y, Song Z, Cai M, Zhou X. (2018). Transcriptome and metabolome analyses reveal global behaviour of a genetically engineered methanol-independent *Pichia pastoris* strain. Process Biochem. 76:46-54.

Shi L, Wang D, Chan W, Cheng L. (2007). Efficient expression and purification of human interferon alpha2b in the methylotrophic yeast, *Pichia pastoris*. Protein Expr. Purif. 54:220–226.

Shirvani R, Yazdanpanah S, Barshan-tashnizi M, Shahali M. (2019). A Novel Methanol-Free Platform for Extracellular Expression of rhGM-CSF in *Pichia pastoris*. Mol. Biotechnol. 61:521–527.

Shu M, Shen W, Yang S, Wang X, Wang F, Wang Y, Ma L. (2016). High-level expression and characterization of a novel serine protease in *Pichia pastoris* by multi-copy integration. Enzyme Microb. Technol. 92:56–66.

Shukla AK, Haase W, Reinhart C, Michel H. (2007). Heterologous expression and comparative characterization of the human neuromedin U subtype II receptor using the methylotrophic yeast *Pichia pastoris* and mammalian cells. Int. J. Biochem. Cell Biol. 39:931–942.

Siró I, Kápolna E, Kápolna B, Lugasi A. (2008). Functional food. Product development, marketing and consumer acceptance-A review. Appetite. 51_456-467.

Slaughter RL, Edwards DJ. (1995). Recent Advances: the Cytochrome P450 Enzymes. Ann. Pharmacother. 29:619–624.

Soley A, Fontova A, Gálvez J, Sarró E, Lecina M, Bragós R, Cairó JJ, Gòdia F. (2012). Development of a simple disposable six minibioreactor system for suspension mammalian cell culture. Process Biochem. 47:597–605.

Sørensen HP. (2010). Towards universal systems for recombinant gene expression. Microb. Cell Fact. 9:27.

Stewart MQ, Esposito RD, Gowani J, Goodman JM. (2001). Alcohol oxidase and dihydroxyacetone synthase, the abundant peroxisomal proteins of methylotrophic yeasts, assemble in different cellular compartments. J. Cell Sci. 114:2863–2868.

Sturmberger L, Chappell T, Geier M, Krainer F, Day KJ, Vide U, Trstenjak S, Schiefer A, Richardson T, Soriaga L, Darnhofer B, Birner-Gruenberger R, Glick BS, Tolstorukov I, Cregg J, Madden K, Glieder A. (2016). Refined *Pichia pastoris* reference genome sequence. J. Biotechnol. 235:121–131.

Suades A, Alcaraz A, Cruz E, Álvarez-Marimon E, Whitelegge JP, Manyosa J, Cladera J, Perálvarez-Marín A. (2019) Structural biology workflow for the expression and characterization of functional human sodium glucose transporter type 1 in *Pichia pastoris*. Sci. Repor. 9:1203.

Sudbery P. (2003). Hansenula polymorpha: Biology and Applications. Yeast 20:1307–1308.

Tan SZ, Prather KL. (2017). Dynamic pathway regulation: recent advances and methods of construction. Curr. Opin. Chem. Biol. 41:28-35.

Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff H V., Snoep JL. (2000). Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur. J. Biochem. 267:5313–5329.

Theron CW, Vandermies M, Telek S, Steels S, Fickers P. (2020). Comprehensive comparison of *Yarrowia lipolytica* and *Pichia pastoris* for production of *Candida antarctica* lipase B. Sci. Rep. 10:1–9.

Tolner B, Smith L, Begent RHJ, Chester KA. (2006). Production of recombinant protein in *Pichia pastoris* by fermentation. Nat. Protoc. 1:1006–1021.

Tomàs-Gamisans M, Ødum ASR, Workman M, Ferrer P, Albiol J. (2019). Glycerol metabolism of *Pichia pastoris* (*Komagataella spp*.) characterised by 13C-based metabolic flux analysis. N. Biotechnol. 50:52–59.

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

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. 15:3859-3876.

Urlacher VB, Girhard M. (2012). Cytochrome P450 monooxygenases: An update on perspectives for synthetic application. Trends Biotechnol. 30: 26-36.

Urlacher VB, Girhard M. (2019). Cytochrome P450 Monooxygenases in Biotechnology and Synthetic Biology. Trends Biotechnol. 37:882–897.

Vail RB, Homann MJ, Hanna I, Zaks A. (2005). Preparative synthesis of drug metabolites using human cytochrome P450s 3A4, 2C9 and 1A2 with NADPH-P450 reductase expressed in *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 32:67–74.

Valderrama-Rincon JD, Fisher AC, Merritt JH, Fan YY, Reading CA, Chhiba K, Heiss C, Azadi P, Aebi M, DeLisa MP. (2012). An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. Nat. Chem. Biol. 8:434–436.

Vanz AL, Nimtz M, Rinas U. (2014). Decrease of UPR- and ERAD-related proteins in *Pichia pastoris* during methanol-induced secretory insulin precursor production in controlled fed-batch cultures. Microb. Cell Fact. 13:23.

Veenhuis M, Van Dijken JP, Harder W. (1983). The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts. Adv. Microb. Physiol. 24:1–82.

Vicente T, Roldão A, Peixoto C, Carrondo MJT, Alves PM. (2011). Large-scale production and purification of VLP-based vaccines. J. Invertebr. Pathol. 107:42-48-

Vieira Gomes A, Souza Carmo T, Silva Carvalho L, Mendonça Bahia F, Parachin N. (2018). Comparison of Yeasts as Hosts for Recombinant Protein Production. Microorganisms 6:38.

Vogl T, Fischer JE, Hyden P, Wasmayer R, Sturmberger L, Glieder A. (2020). Orthologous promoters from related methylotrophic yeasts surpass expression of endogenous promoters of *Pichia pastoris*. AMB Express 10:38.

Vogl T, Glieder A. (2013). Regulation of *Pichia pastoris* promoters and its consequences for protein production. N. Biotechnol. 30:385-404.

Vogl T, Hartner FS, Glieder A. (2013). New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. Curr. Opin. Biotechnol. 24:1094–1101.

Vogl T, Kickenweiz T, Pitzer J, Sturmberger L, Weninger A, Biggs BW, Köhler E-M, Baumschlager A, Fischer JE, Hyden P, Wagner M, Baumann M, Borth N, Geier M, Ajikumar PK, Glieder A. (2018a). Engineered bidirectional promoters enable rapid multi-gene co-expression optimization. Nat. Commun. 9:3589.

Vogl T, Sturmberger L, Fauland PC, Hyden P, Fischer JE, Schmid C, Thallinger GG, Geier M, Glieder A. (2018b). Methanol independent induction in *Pichia pastoris* by

simple derepressed overexpression of single transcription factors. Biotechnol. Bioeng. 115:1037–1050.

Vogl T, Ruth C, Pitzer J, Kickenweiz T, Glieder A. (2014). Synthetic core promoters for *Pichia pastoris*. ACS Synth. Biol. 3:188–191.

Vogl T, Sturmberger L, Kickenweiz T, Wasmayer R, Schmid C, Hatzl AM, Gerstmann MA, Pitzer J, Wagner M, Thallinger GG, Geier M, Glieder A. (2016). A Toolbox of Diverse Promoters Related to Methanol Utilization: Functionally Verified Parts for Heterologous Pathway Expression in *Pichia pastoris*. ACS Synth. Biol. 5:172–186.

Wang J, Zhang T, Li Y, Li L, Wang Y, Yang B, Wang Y. (2019). High-level expression of *Thermomyces dupontii* thermo-alkaline lipase in *Pichia pastoris* under the control of different promoters. 3 Biotech. 9:33.

Wang J, Wang X, Shi L, Qi F, Zhang P, Zhang Y, Zhou X, Song Z, Cai M. (2017). Methanol-Independent Protein Expression by AOX1 Promoter with trans-Acting Elements Engineering and Glucose-Glycerol-Shift Induction in *Pichia pastoris*. Sci. Rep. 7:41850.

Wang J, Lin M, Xu M, Yang ST. (2016a). Anaerobic fermentation for production of carboxylic acids as bulk chemicals from renewable biomass. Adv. Biochem. Eng. Biotechnol. 156:323–363.

Wang X, Wang Q, Wang J, Bai P, Shi L, Shen W, Zhou M, Zhou X, Zhang Y, Cai M. (2016b). Mit1 Transcription Factor Mediates Methanol Signaling and Regulates the Alcohol Oxidase 1 (AOX1) Promoter in *Pichia pastoris*. J. Biol. Chem. 291:6245–61.

Waterham HR, Digan ME, Koutz PJ, Lair S V., Cregg JM. (1997). Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene. 186:37–44.

Weckbecker A, Gröger H, Hummel W. (2010). Regeneration of nicotinamide coenzymes: Principles and applications for the synthesis of chiral compounds. Adv. Biochem. Eng. Biotechnol. 120:195–242.

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. 5:179–89.

Weninger A, Fischer JE, Raschmanová H, Kniely C, Vogl T, Glieder A. (2018). Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers. J. Cell. Biochem. 119:3183–3198.

Weninger A, Hatzl A-M, Schmid C, Vogl T, Glieder A. (2016). Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. J. Biotechnol. 235:139–149.

Worsch A, Eggimann FK, Girhard M, Bühler CJ, Tieves F, Czaja R, Vogel A, Grumaz C, Sohn K, Lütz S, Kittelmann M, Urlacher VB. (2018). A novel cytochrome P450 mono-oxygenase from *Streptomyces platensis* resembles activities of human drug metabolizing P450s. Biotechnol. Bioeng. 115:2156–2166.

Wu D, Chu J, Hao Y-Y, Wang Y-H, Zhuang Y-P, Zhang S-L. (2011). High efficient production of recombinant human consensus interferon mutant in high cell density culture of *Pichia pastoris* using two phases methanol control. Process Biochem. 46:1663–1669.

Wu J, Kaufman RJ. (2006). From acute ER stress to physiological roles of the unfolded protein response. Cell Death Differ. 13:374-384.

Wu MH, Huang S Bin, Cui Z, Cui Z, Lee G Bin. (2008). A high throughput perfusion-based microbioreactor platform integrated with pneumatic micropumps for three-dimensional cell culture. Biomed. Microdevices 10:309–319.

Xiong AS, Yao QH, Peng RH, Zhang Z, Xu F, Liu JG, Han PL, Chen JM. (2006). High level expression of a synthetic gene encoding *Peniophora lycii* phytase in methylotrophic yeast *Pichia pastoris*. Appl. Microbiol. Biotechnol. 72:1039–1047.

Xu J, Towler M, Weathers PJ. (2016). Platforms for Plant-Based Protein Production. Bioprocess. Plant Vitr. Syst, 1–40.

Xu P, Clark C, Ryder T, Sparks C, Zhou J, Wang M, Russell R, Scott C. (2017). Characterization of TAP Ambr 250 disposable bioreactors, as a reliable scale-down model for biologics process development. Biotechnol. Prog. 33:478–489.

Yamawaki S, Matsumoto T, Ohnishi Y, Kumada Y, Shiomi N, Katsuda T, Lee EK, Katoh S. (2007). Production of Single-Chain Variable Fragment Antibody (scFv) in Fed-Batch and Continuous Culture of *Pichia pastoris* by Two Different Methanol Feeding Methods. J. Biosci. Bioeng. 104:403–407.

Yan J, Zheng X, Li S. (2014). A novel and robust recombinant *Pichia pastoris* yeast whole cell biocatalyst with intracellular overexpression of a *Thermomyces lanuginosus* lipase: Preparation, characterization and application in biodiesel production. Bioresour. Technol. 151:43–48.

Yang H, Zhai C, Yu X, Li Z, Tang W, Liu Y, Ma X, Zhong X, Li G, Wu D, Ma L. (2016a). High-level expression of Proteinase K from *Tritirachium album* Limber in *Pichia pastoris* using multi-copy expression strains. 122:38-44.

Yang J, Lu Z, Chen J, Chu P, Cheng Q, Liu J, Ming F, Huang C, Xiao A, Cai H, Zhang L. (2016b). Effect of cooperation of chaperones and gene dosage on the expression of porcine PGLYRP-1 in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 100:5453–5465.

Yang Z, Zhang Z. (2017). Codon-optimized expression and characterization of a pH stable fungal xylanase in *Pichia pastoris*. Process Biochem. 53:80–87.

Yang Z, Zhang Z. (2018). Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: A review. Biotechnol. Adv. 36:182–195.

Yao J, Weng Y, Dickey A, Wang KY. (2015). Plants as factories for human pharmaceuticals: Applications and challenges. Int. J. Mol. Sci. 16:28549-28565.

Yee CM, Zak AJ, Hill BD, Wen F. (2018). The Coming Age of Insect Cells for Manufacturing and Development of Protein Therapeutics. Ind. Eng. Chem. Res. 57:10061–10070.

Young CL, Yuraszeck T, Robinson AS. (2011). Decreased secretion and unfolded protein response upregulation. Methods Enzymol. 491:235–260.

Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, Woodcock J. (2014). Understanding pharmaceutical quality by design. AAPS Journal. 16:771-783.

Yu XW, Sun WH, Wang YZ, Xu Y. (2017). Identification of novel factors enhancing recombinant protein production in multi-copy *Komagataella phaffii* based on transcriptomic analysis of overexpression effects. Sci. Rep. 7:16249.

Zarei N, Vaziri B, Shokrgozar MA, Mahdian R, Fazel R, Khalaj V. (2014). High efficient expression of a functional humanized single-chain variable fragment (scFv) antibody against CD22 in *Pichia pastoris*. Appl. Microbiol. Biotechnol. 98:10023–10039.

Zhan C, Yang Y, Zhang Z, Li X, Liu X, Bai Z. (2017). Transcription factor Mxr1 promotes the expression of Aox1 by repressing glycerol transporter 1 in *Pichia pastoris*. FEMS Yeast Res. 17:4.

Zhang C, Ma Y, Miao H, Tang X, Xu B, Wu Q, Mu Y, Huang Z. (2020). Transcriptomic Analysis of *Pichia pastoris* (*Komagataella phaffii*) GS115 During Heterologous Protein Production Using a High-Cell-Density Fed-Batch Cultivation Strategy. Front. Microbiol. 11:463.

Zhang J. (2010) Mammalian Cell Culture for Biopharmaceutical Production. Manual of Industrial Microbiology and Biotechnology, Third Edition.

Zhang W, Zhao H, Xue C, Xiong X, Yao X, Li X, Chen H, Liu Z. (2006). Enhanced Secretion of Heterologous Proteins in *Pichia pastoris* Following Overexpression of *Saccharomyces cerevisiae* Chaperone Proteins. Biotechnol. Prog. 22:1090–1095.

Zhu T, Guo M, Tang Z, Zhang M, Zhuang Y, Chu J, Zhang S. (2009). Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast *Pichia pastoris*. J. Appl. Microbiol. 107:954–963.

Zhu T, Guo M, Zhuang Y, Chu J, Zhang S. (2011). Understanding the effect of foreign gene dosage on the physiology of *Pichia pastoris* by transcriptional analysis of key genes. Appl. Microbiol. Biotechnol. 89:1127-1135.

Zydney AL. (2015). Perspectives on integrated continuous bioprocessing - opportunities and challenges. Curr. Opin. Chem. Eng. 10:8-13.