Universitat Autònoma de Barcelona

Escola d'Enginyeria

Departament d'Enginyeria Química

Model-based design and development of operational strategies for *Rhizopus oryzae* lipase production in *Pichia pastoris* under the *AOX1* promoter

PhD Thesis

José Manuel Barrigón de San Marcos

Bellaterra, 18 de diciembre 2014

FRANCISCO VALERO BARRANCO y JOSÉ LUIS MONTESINOS SEGUÍ, catedrático y profesor titular del Departament d'Enginyeria Química de la Universitat Autònoma de Barcelona,

CERTIFICAMOS:

Que el Ingeniero Químico José Manuel Barrigón de San Marcos ha realizado, bajo nuestra dirección, en los laboratorios del Departament d'Enginyeria Química, el trabajo con el título "**Model-based design and implementation of operational strategies for *Rhizopus oryzae* lipase production in *Pichia pastoris* under the *AOX1* promoter**" que se presenta en esta memoria, la cual constituye su Tesis para optar al Grado de Doctor por la Universitat Autònoma de Barcelona.

Y para dejar constancia y que tenga los efectos que corresponda, se presenta ante la Escola d'Enginyeria de la Universitat Autònoma de Barcelona la mencionada Tesis, firmando esta certificación.

Bellaterra, 18 de diciembre de 2014

Dr. Francisco Valero Barranco Dr. José Luis Montesinos Seguí

Acknowledgements

Me gustaría agradecer a todos los que me habéis acompañado mi camino de doctorando hasta la culminación del gran objetivo de esta etapa: la tesis doctoral.

En primer lugar, a mis directores de tesis Paco y José Luis con los cuales he podido aprender hasta el último día, divagar y discutir ampliamente sobre los progresos y decepciones científicas y poner en común nuestra particular visión del mundo. A parte de discusiones científicas y nuestros e-mails para cubrirnos de gloria, también hemos tenido tiempo para los grandes "banquetes" de corderito y el famoso menú de bravas y chocos.

A mis compañer@s y colegas Xavi G, Xavi P, Núria, Elena, Marius, Albert, Margot, Carol, Marc, Joel, Marina, Kristin, Nelsy del grupo de lipasas, y por extensión a los aldolaseros Alfred, Martina, Dani y Marcel·la. Que aparte de compañeros de despacho, de videos, cervezas, squash, y fricadas varias siempre seremos "adictos" a *Pichia* y *Coli.*

I would like to thank Michael Maurer and Diethard Mattanovich the acceptance for a stay at University of Natural Resources and Life Sciences. Although, the experimental work developed during my stay has not been fruitful, the collaboration allowed me to improve my knowledge of *Pichia*'s physiology and, of course, to meet your teamwork that are a small family of brilliant researchers. Nils, Corinna, Frederik, Marizela, Justina, Markus, Fabian, Brigitte, Diethard and Michi, I will never forget you.

También os la dedico a vosotros mis amigos Javi G, Edu, Carles, Cris, Olaia, Belen, Mabel, Javi D los que hemos hecho la tesis B, no en contabilidad, sino en fiestas, anedotas, cotilleos, alegrías, peleas, aventuras, deporte, visitas intempestivas a los laboratorios, bodas, etc. Y que esta tesis aun no se ha acabado de escribir.

A mi familia por estar apoyándome en el camino que he elegido y porque nunca habéis cuestionado mi elección. Gracias por estar ahí.

Y finalmente quiero dedicar este trabajo a Raquel, el amor de mi vida. Creo que aunque no hayas hecho una tesis, te has doctorado en paciencia y empatía, porque sólo tú sabes lo que me ha costado concretar el trabajo que comencé ya hace más de 5 años. Gracias por aguantar esos largos fines de semana con la tesis, la tesis, la tesis, la cual es más larga que el campo de futbol de Oliver y Benji, Gracias por sobrellevar que haya tenido que reducir nuestro tiempo a cambio de poder estar un poquito más cerca de cerrar esta bonita y dura etapa.

MUCHAS GRACIAS A TODOS.

Table of contents

Abstract

Pichia pastoris is recognized as one of the most efficient cell factories for the production of recombinant proteins. More than 500 proteins have been expressed using this system. *P. pastoris* combines the ability of growing on minimal medium at very high cell densities with secreting the heterologous protein, aiding to their recovery.

In this work, the selected target protein has been is the recombinant *Rhizopus oryzae* lipase (ROL). The heterologous ROL production in *P. pastoris* P_{AOX1} (Mut^+) fed-batch cultures has been studied for bioprocess monitoring and control, kinetic modelling, and the design and development operational strategies.

The first step in the research has been the estimation of biomass, substrate and specific growth rate (μ) by means of two non-linear observers and a linear estimator. The aim of this study has been to compare the performance of the different algorithms in *P. pastoris* bioprocesses.

Heterologous protein production is closely related to μ. So, due to its high relevance in the bioprocess, μ has been estimated by on-line gas analyses or substrate concentration measurements. Biomass and substrate have been straightforwardly obtained solving their corresponding mass balances.

The most frequently used cultivation strategy to achieve high cell densities and high heterologous protein production levels with P_{AOX1}-(Mut^+)-based system is the fed-batch operation. The standard operational strategies are based on the control of the substrate concentration close to zero (limiting strategies) or keeping the concentration at a constant value (non-limiting strategies). Consequently, the effect of methanol non-limiting fed-batch (MNLFB) and methanol limited fed-batch (MLFB) operational

strategies on ROL production has been studied. These most commonly applied control strategies allow maintaining rather constant key specific rates: cell growth (μ), substrate uptake (q_s), and protein production (q_p) through the quasi-steady state hypothesis for substrate. Results have been analyzed in order to determine the most suitable operating conditions in terms of yields and productivities.

Furthermore, mean specific rates and state variables for various fed-batch cultures, under methanol limited and non-limited conditions were used for modeling. Hence, an unstructured macrokinetic model for heterologous ROL production by a *P. pastoris* P_{AOX1} based system has been developed. Then, a comparative meta-analysis of heterologous protein production of various target proteins by *P. pastoris* under *AOX1* promoter has been conducted and a general strategy for improving protein production from process kinetics was developed as a key to bioprocess optimization.

Additionally, the characterization of oxygen transfer capacity in different laboratory/pilot scale bioreactors has been studied. The oxygen transfer model was also developed and validated in heterologous ROL production by *P. pastoris* under *AOX1* promoter. Finally, the previous kinetic model for heterologous ROL production and oxygen transfer model have been applied to define alternative operational strategies based on oxygen limited fed-batch operations (OLFB) and have been compared to the standard strategies.

Resumen

Pichia pastoris es identificada como una de las factorías celulares más eficientes para la producción de proteínas recombinantes. Más de 500 proteínas han sido expresadas usando este sistema de expresión. *P. pastoris* combina la habilidad de crecer en medios mínimos a altas densidades celulares con la de secreción de proteína heteróloga, ayudando a su recuperación.

En este trabajo, la proteína modelo seleccionada es la lipasa recombinante de *Rhizopus oryzae* (ROL). La producción heteróloga de ROL en cultivos *fed-bach* de *P. pastoris* P_{AOX1} (Mut^+) ha sido estudiada desde el punto de vista de la monitorización y control del bioproceso, la modelización cinética y el diseño y desarrollo de estrategias operacionales.

El primer paso en la investigación ha sido la estimación de la concentración de biomasa, sustrato y velocidad específica de crecimiento (μ) mediante dos observadores no lineales y uno lineal. El objetivo de este estudio ha sido comparar las prestaciones de los diferentes algoritmos en bioproceso de *P. pastoris.*

La producción de proteína recombinante está estrechamente relacionada con la μ. Por lo tanto, debido a su elevada relevancia en el bioproceso, μ fue estimada utilizando el análisis de los gases de salida en línea o las medidas de la concentración de sustrato. La biomasa y el sustrato fueron obtenidos directamente de sus correspondientes balances de materia.

La estrategia de cultivo más frecuentemente utilizada para conseguir altas densidades celulares y elevada producción de proteína heteróloga con el sistema P_{AOX1} (Mut^+) es la operación *fed-batch*. Las estrategias operacionales estándar están basadas en el control de la concentración de sustrato próximas a cero (estrategias limitantes) o

manteniendo la concentración a un valor contante (estrategias no limitantes). En consecuencia se estudió el efecto de las estrategias operacionales *fed-batch* de metanol no limitante (MNLFB) y metanol limitante (MLFB) en la producción de ROL. Éstas son las estrategias de control más comunes que tienen como objetivo mantener contantes las velocidades específicas claves: crecimiento celular (μ), consumo de sustrato (q_s) y producción de proteína (q_p) a partir de la hipótesis de estado quasi-estacionario para el sustrato. Los resultados se analizaron con el objetivo de determinar la condición más apropiada en función de rendimientos y productividades.

Las velocidades específicas medias y las variables de estado para varios cultivos f*ed-batch*, bajo condiciones de metanol limitante y no limitante, se usaron para el desarrollo de un modelo macrocinético no estructurado para la producción heteróloga de la ROL por el sistema *P. pastoris* P_{AOX1}. Posteriormente, se ha realizado un meta-análisis comparativo sobre la producción de varias proteínas heterólogas modelo para *P. pastoris* bajo el promotor *AOX1* y se ha desarrollado una estrategia general para mejorar la producción de proteínas a partir de la cinética como clave para la optimización del proceso.

Adicionalmente se ha estudiado y caracterizado la capacidad de transferencia de oxigeno en diferentes biorreactores a escala laboratorio y piloto. El modelo de transferencia de oxigeno ha sido también desarrollado y validado en la producción de ROL en *P. pastoris* bajo el promotor *AOX1*. Finalmente, el modelo cinético previamente desarrollado para la producción heteróloga de ROL y el modelo de transferencia de oxigeno han sido aplicados para definir estrategias alternativas de operación basadas en operaciones con oxígeno limitante (OLFB) comparándose con las estrategias estándar.

Objectives

The main goal of the present thesis is a Model-based design and implementation of operational strategies for *Rhizopus oryzae* lipase production in *Pichia pastoris* phenotype Mut^+ under the *AOX1* promoter using methanol as sole carbon source. To achieve this general objective, the PhD thesis has been organized to achieve the following detailed objectives:

- To design and develop algorithms for on-line estimation of the state variables, biomass and substrate and the specific growth rate for the heterologous protein production in *P. pastoris* cultures.
- To study the impact of the residual methanol concentration in the recombinant *R. oryzae* lipase expression in fed-batch cultures by the *P. pastoris* P_{AOX1} (Mut^+) cell factory.
- To study the kinetics for the heterologous protein production in *P. pastoris* cultures and specifically, modeling the kinetics for the *R. oryzae* lipase expression by the *P. pastoris* P_{AOX1}(Mut+)-based system.
- To characterize the oxygen transfer capacity of some bioreactors at laboratory and pilot scale.
- To design and implement standard and alternative operational strategies to improve the heterologous *R. oryzae* lipase production in fed-batch cultures by the *P. pastoris* P_{AOX1} (Mut^+) cell factory.

Chapter 1:

Introduction

1. Introduction

1.1. Recombinant expression systems

The recombinant protein production in microorganisms has been developed in order to produce the suitable stable and functional protein in higher amount than by isolation from the native source (Porro et al., 2011). The first heterologous protein product using genetic recombination was the human insulin in 1977. Nowadays, the recombinant production of biopharmaceuticals is a global business, particularly the recombinant antibodies and vaccines (Bill et al., 2014). The biosynthesis of numerous valuable molecules such as antitumor, anticancer, antiviral, antiparasitic, antioxidant, immunological, enzymes, antibiotics and hormones are the main applications (Jeandet et al., 2013). Modern biotechnology are based on bacteria, lower eukaryotes (yeasts and fungi), invertebrates (insect cells and larvae), vertebrates (cells and transgenic animals), and plants (cells and transgenic plants) expression systems for heterologous protein expression (Demain and Vaishnav, 2009; Sreekrishna, 2010; Porro et al., 2011; Valero 2013; Bill et al., 2014).

Bacteria expression systems are easy to culture, to perform genetic modifications, to grow rapidly in minimal basalt media, achieving high product yield and easy to scale up. The disadvantages are the lack of post-translational modifications (glycosylation and disulfide bonds formation), problems with protein folding, which imply expensive downstream and purification difficulties. *Escherichia coli* is still the most popular host cell for producing recombinant proteins in both commercial and research applications. Recent new progresses in *E. coli* system have been developed for producing glycosylated proteins but glycosylation homogeneity has not achieved (Chen, 2012;

Corchero et al., 2013). Other bacteria expression system, such as *Lactoccocus lactis,* is a great choice to express membrane proteins. *Pseudomonas* systems have achieved product titers comparable to *E. coli* systems. Additionally, *Streptomyces, coryneform* bacteria, and halophilic bacteria have emerged in recent years as alternative bacterial expression systems (Demain and Vaishnav, 2009; Chen, 2012; Gopal and Kumar, 2013),

Yeast and fungi are eukaryotic organisms capable of growing at high growth rates similar to bacteria, can produce a high yield of secreted recombinant proteins reaching high productivities, assist on protein folding, carry out post-translational modifications similar to higher eukaryotic cells and allow scaling-up the bioprocess. The glycosylation patter differs between species and for example yeasts tend to hiperglycosilate (high level of mannose). Yeast and fungal homologous proteases may reduce the recombinant product yield. Some typically yeast and fungi expression systems are: *Saccharomyces cerevisiae, Pichia pastoris, Hansenula polymorpha* and *Aspergillus niger* (Porro et al., 2011, Çelik and Çalik, 2012; Mattanovich et al., 2012; Ward, 2012; Valero, 2013; Bill et al., 2014; Nevalienen and Peterson, 2014).

Insect cell lines are able to perform many of the post-translational modifications such as glycosylation, disulfide bond formation and phosphorylation of many complex proteins, the bioprocess is scalable and safe. However, this expression system grows slower than bacteria, yeast and fungi, needs more expensive media to produce recombinant protein production and may originate intracellular aggregates due to misfolded proteins. The baculovirus-insect cell expression system has been explored for the production of viral and parasitic antigens and, more recently, commercial

vaccines. The common host is *Spodoptera frugiperda* and baculovirus is the most typically vector (*Autographa californica*) (Cox, 2012; Kollewe and Vilcinskas, 2013).

Therapeutic protein production in transgenic plants is safe (low risk of animal pathogens contamination), low cost of upstream biomass process and flexible production scalability. Plants are considered as a potential bioreactor system for virus vectors. Although high levels of accumulation of protein in plant issues could be reached, low yield of recombinant soluble proteins may be recovered and it requires expensive downstream process, it represents up to 80% of overall production costs (Fischer et al., 2012). Proteolysis and plant-specific *N*-glycosylation pattern of glycosylated proteins are the main drawbacks. However, new strategies in *N*-glycosylation engineering for humanization the plant glycosylation pathway have been explored (Gomord et al., 2010; Kim et al., 2014). The representative transgenic expression system is *Nicotiana tabacum* (Mason et al., 1992).

Mammalian expression system is selected when the proteins require proper folding, human-like post-translational modifications. Mammalian cells, compared with the previous cell factories, grow slower, product titter is lower too, virus is the higher potential contamination focus agents when processes require exogenous animal proteins such as serum, cell line development is a very time-consuming process which takes at least 6 months, the cost of the culture media is expensive and also the investment for scale-up the process (Bil, 2014; Maccani et al., 2014). Various mammalian cell lines such as Chinese hamster ovary (CHO), mouse myeloma (NS0), baby hamster kidney (BHK), human embryonic kidney (HEK-293), and human embryonic retinoblast (PER.C6) cells have been established. Nevertheless, almost all

approved mammalian cell-derived biopharmaceutical proteins have been produced in CHO cells (Lai et al., 2013, Maccani et al., 2014).

To sum up, the expression system of recombinant proteins is selected depending on the target protein. Approximately 20% of biopharmaceutical proteins are produced in yeasts 30% in *E. coli* and 50% in mammalian cell-lines and hybridomas (Ferrer-Miralles et al., 2009; Mattanovich et al., 2012, Bill, 2014). The mammalian cell-lines are currently the workhorse of the biopharmaceutical industry because they are able to produce complex recombinant proteins with human-glycoforms, which are bioactive in humans. However, new advances have been developed in all expression systems to produce recombinant human proteins with humanized *N*-glycans (Bill, 2014; Maccani et al., 2014).

1.2. The methylotrophic yeasts

In the early 1970's, methylotrophic yeasts were studied for the production of yeast biomass for single cell protein application from methanol as the sole carbon and energy source by Phillips Petroleum Company (Potvin, et al., 2012). However, in 1980’s, the interest on this group of yeast was focused in their potential as recombinant production systems due to their ability to grow to high cell densities in a basal media and the strong high expression of methanol oxidizing enzymes induced by methanol (Yurimoto et al 2011; Mattanovich et al., 2012).

Nowadays, the methylotrophic yeasts used as a host for recombinant protein expression are: *Pichia pastoris* (also called *Komagataella pastoris*), *Hansenula polymorpha* (formerly *Pichia Angusta*), *Pichia methanolica*, *Candida boidinii* and

Pichia minuta (or *Ogataea minuta*). They are exploited as cellular factories in industrial and academic applications for recombinant production of a large number of proteins including enzymes, antibodies, cytokines, plasma proteins, and hormones (Hartner and Glieder, 2006; Mattanovich et al., 2012).

The methanol-utilizing pathway of all methylotrophic yeasts is analogous. A methanol metabolism schema from Yurimoto et al., (2011) is showed in Figure 1. Methanol (CH_3OH) is first oxidized by alcohol oxidase (AOX) to form formaldehyde (CH_2O) and hydrogen peroxide (H_2O_2), which are both highly toxic compounds. The hydrogen peroxide formed is removed by the action of a catalase (CTA) (Yurimoto et al., 2011). A portion of formaldehyde generated by AOX, leaves peroxisome to be oxidized to formate, and finally to CO_2, the rest can be assimilated into carbohydrates (Sola et al., 2007).

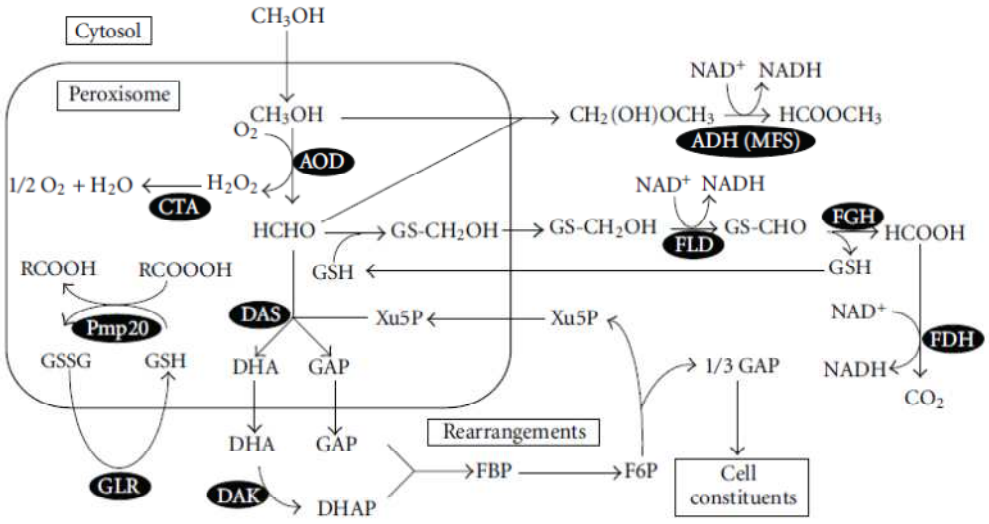

Figure 1.: Methanol metabolism schema in methylotrophic yeasts. Figure is obtained from Yurimoto et al 2011.

The formaldehyde oxidized to CO_2 by the cytosolic dissimilation pathway is to supply a source of energy (NADH) (Gao and Shi, 2013). Formaldehyde reduced form of glutathione (GSH) to generate S-hydroxymethyl glutathione (GS-CH_2OH). The dependent formaldehyde dehydrogenase (FLD) regenerates the NADH from NAD^+ catalyzing the *S*-formylglutathione (GS-CHO) formation from GS-CH_2OH. S-formylglutathione hydrolase (FGH) hydrolyses the GS-CHO in formate and glutathione. The final oxidation of formate to produce CO_2 and regenaretes the NADH is catalyzed by the formate dehydrogenase (FDH) (Yurimoto et al., 2011).

Formaldehyde assimilation pathway involves a dihydroxyacetone synthase (DAS), which catalyses the following reaction: formaldehyde is fixed to xylulose 5-phosphate (Xu5P) forming dihydroxyacetone (DHA) and glyceraldehyde 3-phosphate (GAP), which are used for the synthesis of cell constituents and the regeneration of Xu5P (Yurimoto et al., 2011). The enzymes (AOX, CTA and DAS) are peroxisomal (Solà et al., 2007). DHA and GAP are further assimilated within the cytosol. DHA is phosphorylated by dihydroxyacetone kinase (DHAK), and, subsequently, dihydroxyacetone phosphate (DHAP) and GAP form fructose 1,6-bisphosphate which is then utilized for regeneration of Xu5P and for biosynthesis of cell constituents (Yurimoto et al., 2011).

Under pure methanol induction, glycolysis route and the tricarboxylic acid (TCA) cycle flux are either blocked or weakened, and formaldehyde dissimilatory pathway is considered as the only route for energizing protein and cell synthesis (Gao and Shi, 2013)

1.3. *Pichia pastoris* as an expression system

In the 1970s, *P. pastoris* high density cultivations (higher than 100 g dry cell weight /L) were performed by Phillips Petroleum Company (Sreekrishna 2010). In the 1980s, efficient classical and molecular tools were developed for recombinant expression in *P. pastoris* in order to exploit pathway of methanol utilization for protein production (Cereghino and Cregg, 2000).

P. pastoris combines the ability of growing on minimal medium at very high cell densities with secreting the heterologous protein, aiding to their recovery (Cos et al., 2006b).

Nowadays, *P. pastoris* is recognized as one of the most efficient cell factories for the production of recombinant proteins (Macauley-Patrick et al., 2005; Potvin et al., 2012). More than 500 proteins have been expressed using this system (Cregg, 2014) and it also has been selected by several protein production platforms for structural genomics programs (Yokohama et al., 2003).

Also, it performs many of the higher eukaryotic post-traductional modifications such as processing of signal sequences (both pre and prepro type), folding, disulfide bond formation, certain types of lipid addition and glycosylation (O-and N-linked oligosaccharide structures) (Cereghino and Cregg, 2000; Gao and Shi, 2013).

Nevertheless, the most important characteristic of *P. pastoris* as host microorganism is the existence of a strong and tightly regulated promoter from the alcohol oxidase 1

gene, the alcohol oxidase 1 promoter (P_{AOX1}) (Cereghino and Cregg, 2000; Cos et al., 2006b; Potvin et al., 2012).

1.3.1. *AOX* Promoters

Currently, the P_{AOX1} is the most widely promoter used in the *Pichia* system for recombinant protein expression and it has associated the use of methanol as an inducer substrate (Cereghino and Cregg, 2000). There are two alcohol oxidase genes in *P. pastoris* that code for AOX enzyme, the alcohol oxidase 1 gene (*AOX1*), which is responsible for greater than 90% of the enzyme in the cell, and the alcohol oxidase 2 gene (*AOX2*) for less than 10%. There are three types of *P. pastoris* host phenotypes available that vary regarding their ability to utilize methanol. The wild type or methanol utilization plus phenotype (Mut^+), and those resulting from deletions in the *AOX1* gene (methanol utilisation slow Mut^S) or both *AOX* genes (methanol utilization minus Mut^-) (Cos et al., 2006b; Potvin et al., 2012).

High expression levels of recombinant Antithrombin III (Mochizuki et al., 2001), mini-proinsulin (País et al., 2003) and horseradish peroxidase isoenzyme C1A (Dietzsch et al., 2011) proteins using the P_{AOX2} or with a truncated version of the promoter have been reported. However, the expression levels using the P_{AOX1} have been higher than those reported for the P_{AOX2} (Cos et al., 2006b). Some examples of proteins expressed using P_{AOX1} are: human β_2-glycoprotein I Domain V (Katakura et al., 1998), heavy-chain fragment C of botulinum neurotoxin serotype A (Zhang et al., 2000), human chymotrypsinogen B (Curvers et al., 2002), hirudin variant 2 (Zhou and Zhang, 2002) and single-chain variable fragment antibody (Yamawaki et al., 2007).

1.3.2. *GAP* Promoter

In recent years, the glyceraldehyde-3-phosphate dehydrogenase constitutive promoter (P_{GAP}) has emerged as the most used alternative to P_{AOX1} (Maurer et al., 2006; Heyland et al., 2011). P_{GAP} is strong and constitutive *P. pastoris* promoter. *GAP* promoter allows the protein expression using glucose, glycerol and other carbon sources as a substrate. Several studies have reported that P_{GAP} is more efficient than P_{AOX1}, whereas others showed opposite results (Potvin et al., 2012). Operationally, the protein expression by P_{AOX1} is performed during an induction phase, where microorganisms grow at slower growth rate than during the non-induction phase. In contrast, the constitutive protein expression can be performed during all culture at the same growth rate by using P_{GAP}. When the protein expression does not impact negatively in the cell growth rate, the use of P_{GAP} for protein expression, instead of P_{AOX1}, may improve the bioprocess productivity (Volg and Glieder, 2013). Thus, it appears that expression levels achieved for a given protein using different promoters vary significantly based on properties of the expressed protein (Garcia-Ortega et al., 2013).

The use of the constitutive *GAP* promoter avoids the use of methanol in the fermentations. So, from a large-scale processing perspective, the P_{GAP} expression system may be advantageous because it eliminates the hazard and cost associated with the storage and delivery of large volumes of methanol and significantly decreases heat production and oxygen requirements of the processes (Heyland et al., 2010).

High expression levels of recombinant antibody Fab fragment (Maurer et al., 2006; Baumann et al., 2008; Garcia-Ortega et al., 2013), *Candida rugosa* lipase (Zhao et al., 2008) and phytase (Tang et al., 2010) proteins using the P_{GAP} have been reported.

1.3.3. *FLD1* Promoter

The *FLD1* gene encodes a formaldehyde dehydrogenase (FLD), an enzyme that plays an important role in the methanol catabolism as carbon source, as well as in the methylated amines metabolism as nitrogen source. The P_{FLD1} from *P. pastoris* was tightly and independently induced either by methanol as carbon source or methylamine as nitrogen source (Resina et al., 2004).

Moreover, the P_{FLD1} has shown similar transcriptional efficiency and tight regulation as P_{AOX1}. Nonetheless, the P_{FLD1} was used for methanol-independent expression of foreign genes in *P. Pastoris* using methylamine as sole nitrogen source and inductor (Cos et al 2005a, Resina et al., 2005).

Gelatin and green fluorescent protein under the control of P_{AOX1} and P_{FLD1} respectively, were simultaneously expressed in *P. Pastoris* (Duan et al., 2009). *Rhizopus oryzae* lipase was successfully expressed under P_{FLD1} (Resina et al., 2004, 2005).

1.3.4. Alternative promoters

1.3.4.1. *DAS* promoters

P. pastoris dihydroxyacetone synthase promoter (P_{DAS}) is a strong and inducible promoter from the methanol utilization pathway showing a similar regulatory pattern and expression level as P_{AOX1}. There are two highly similar *DAS* genes in *P. pastoris* (91 % similarity). The transcription of *DAS1* and *DAS2* was found to be equally high induced upon methanol induction (Vogl and Glieder, 2013).

Different approaches have been reported using the *DAS* promoters. Alternative expression to P_{AOX1} was performed by using the promoter sequence of *DAS2* for β-galactosidase expression (Takagi et al., 2012). On the other hand, horseradish peroxidase (HRP) and *Candida antarctica* lipase B (CalB) were coexpressed by the methanol induction of *DAS1* and *AOX2* promoters (Krainer et al., 2012).

1.3.4.2. *PGK1* promoter

Phosphoglycerate kinase (*PGK*) is a constitutively expressed housekeeping gene involved in glycolysis and gluconeogenesis. The *B. subtilis* α-amylase was expressed in *P. pastoris* by using the P_{PGK1}. Three different carbon sources were used to regulate the expression. mRNA levels were two times higher on glucose than glycerol. Similar α-amylase expression levels were observed in the P_{AOX1} and P_{PGK1} by methanol carbon source (de Almeida et al., 2005). However, the strength of P_{PGK1} was evaluated by the expression of three different proteins (β-Gal, HSA and GFP). P_{PGK1} showed a weak driven expression, only 10% of PGAP (Stadlmayr et al., 2010).

1.3.4.3. *TFL1* promoter

Translation elongation factor 1 alpha (TEF1) is a crucial component of the eukaryotic translation machinery and mediates the delivery of aminoacyl tRNAs to the ribosomes to sustain the elongation of the peptide chain. The P_{TEF1} is a strong constitutive promoter highly coupled to the growth rate, suitable for heterologous protein production (Volg and Glieder, 2013).

Lipase from *Bacillus stearothermophilus* L1 fused to a cellulose binding domain (CBD) from *Trichoderma harzianum* (CLLip), enhanced green fluorescent protein

(eGFP), β-galactosidase (β-Gal) and human serum albumin (HSA) were expressed under P_{TEF1} and P_{GAP}. P_{TEF1} driven expression was similar to P_{GAP} (Ahn et al., 2007; Stadlmayr et al., 2010).

1.3.4.4. *AOD* promoter

Mitochondrial alternative oxidases (AOX or AOD) are key enzymes for a shortcut to the standard respiratory pathway in plants, many fungi and yeasts. Alternative oxidases are involved in stress responses, programmed cell death and maintenance of the cellular redox balance (Volg and Glieder, 2013).

The fusion protein of the alternative oxidase and GFP (AOD-GFP) expression was tested under control of the AOD promoter and the GAP promoter. P_{AOD} is a strong promoter, not only depend to the methanol carbon source for induction, achieving similar expression levels as P_{GAP} on glucose (Kern et al., 2007).

1.3.4.5. *PHO89* promoter

P. pastoris PHO89 gene encodes the putative Na^+-coupled phosphate symporter (PHO89). The putative promoter of PHO89 showed a strong regulation by the phosphate concentration in the growth medium. The use of a phosphate-responsive promoter is an alternative to constitutive and inducible promoters to drive expression of a heterologous gene in *P. pastoris* (Volg and Glieder, 2013).

The lipase gene from *B. stearothermophilus* L1 fused to the CBD (CLLip) was expressed under $P_{PHO89.}$ P_{PHO89} driven expression appears showing similar expression levels as P_{GAP} and P_{TEF1} (Ahn et al., 2009).

1.3.4.6. *THI11* promoter

THI11 gene of *P. pastoris* codes the protein involved in synthesis of the thiamine precursor hydroxymethylpyrimidine. P_{THI11} is regulated by the availability of thiamine in the growth medium. The eGFP, β-Gal and HSA expression were evaluated under inducer P_{THI11} and compared to constitutive P_{GAP} in batch (shake flask) and fed-batch operational modes. In the absence of thiamine the relative expression levels of P_{THI11} are growth phase dependent in the range of 19–73% compared to P_{GAP} while the addition of thiamine represses P_{THI11} (Stadlmayr et al., 2010; Delic et al., 2013).

1.3.4.6. *GTH1* promoter

The high affinity glucose transporter gene (*GTH1*) is under the control of the innovative promoter P_{GTH1}. DNA microarray analysis was used to identify genes with both, strong repression on glycerol and high-level expression on glucose. Six novel promoters abbreviated (P_{G1}, P_{G3}, P_{G4}, P_{G6}, P_{G7} and P_{G8}) were applied to express the eGPF. The highest eGPF expression level in pre-cultures were achieved with the P_{GTH1} (alos named P_{G1}). The HSA expression levels under inducer P_{GTH1} was compared to constitutive $P_{GAP,}$ where was increased 2-fold. These six promoters provide a tool box for recombinant genes expression, where the P_{GTH1} shows the best performance (Prielhofer et al., 2013).

1.3.4.5. Weak alternative promoters

The YPT1 gene encodes a guanosine triphosphatase (GTPase) involved in secretion. The bacterial beta-glucuronidase (GUS) expression was performed with the constitutive *GAP* and *YPT1* promoters and the inducible P_{AOX1}. Although, a weak GUS expression was observed under the P_{YPT1} (>10 times weaker than P_{GAP}), the expression was driven on different carbon sources: glucose, methanol, or mannitol (Sears et al. 1998).

The PEX8 gene encodes a peroxisomal matrix protein that is essential for peroxisome biogenesis. The PEX8 can be expressed at low level on glucose and can be induced by methanol or oleate (Liu et al., 1995).

Other inducer promoter is the isocitrate lyase 1 (P_{ICL1}). This promoter was used to express a dextranase and shown to be regulated on the transcriptional level by ethanol and repressed by glucose in the exponential phase, but not in the stationary phase (Menendez et al., 2003; Valero et al., 2013).

P_{PEX8} and P_{YPT1} are weak promoters that were considered as alternatives to the classical P_{AOX1} and P_{GAP} (Cereghino and Cregg, 2000). The P_{ICL1} strength is unclear, no comparison to classical promoters have been performed yet. However, they have only rarely been used (Volg and Glieder, 2013).

1.4. Fermentation procedures in *PAOX1*-based systems

1.4.1 Operational conditions

Fermentation basal salt medium (BSM) is the medium recommended by Invitrogen fermentation guidelines to perform *P. pastoris* high cell density fermentation (Invitrogen, 2014). Although most researchers use the BSM medium (Cos et al., 2006b; Potvin et al., 2012), it may not be the optimum because its composition may be unbalanced, form precipitates and have too much ionic strength. Furthermore, the BSM medium provides a high concentration of salts (Sreekrishna, 2010).

Alternative media have been proposed to avoid the use of BSM: FM22 (Zhang et al., 2007), d'Anjou and Daugulis medium (2000), and Baumann et al., medium (2008). Baumann's and d'Anjou's media have a low concentration of salts. In contrast, the FM22 medium contains similar concentrations to BSM's with the exception of potassium salts (Sreekrishna, 2010). Furthermore, all defined media are supplemented by PTM1 trace salts. PTM1 is the supplementary trace salts stock solution proposed by Invitrogen (2014).

One of the most important points in a medium formulation is the nitrogen source (Sreekrishna, 2010). In BSM, FM21 and Baumann's media, nitrogen is added as ammonium hydroxide by means of the pH controller. In the D'Anjou's medium, all nitrogen is provided at the initial formulation. Nitrogen starvation starts around 50 $g \cdot L^{-1}$ of biomass in BSM media (Cos et al., 2006b). However, lack of nitrogen can affect the production at high biomass concentration, over 60 $g \cdot L^{-1}$ of dry cell weight (DCW) in BSM, FM22 and Baumann media. The lack of nitrogen is directly related to the increase in proteolytic activity, resulting in the degradation of extracellular

proteins (Yang et al., 2004). Alternatively, the use of organic nitrogen sources such as yeast extract, peptone, casamino acids or beef extract could help to improve the cell concentration and achieve higher expression levels than using organic nitrogen sources (Gao and Shi, 2013).

The optimal growth and production temperature in *P. pastoris* is 30°C, above 32°C protein expression stops and growth quickly decays (Cos et al., 2006b). However, some authors work at lower temperatures to improve the bioprocess performance (Curvers et al., 2002, Potgieter et al., 2010). Low temperature, such as 20ºC, could activate the AOX expression for cultivation on methanol, relieve cell lysis and protease secretion, and reduce extracellular proteolytic activity (Gao and Shi, 2013).

Although *P. pastoris* can grow at pH values ranging from 3 to 7, the most commonly pH used in different studies is between 5 and 6. Beyond this range, pH may affects protein stability, may provoke extracellular protease activity (Cos et al., 2006b; Potvin et al., 2012; Gao and Shi, 2013). Additionally, the use of BSM medium at pH greater than 5.0 may cause precipitation of medium salts (Cos et al., 2006b).

Another important parameter that affects the protein expression is the dissolved oxygen (O_2). The oxygen is used for methanol oxidation to formaldehyde and also in the formaldehyde dissimilatory pathway (See 1.2 section). The oxygen is provided to the culture by air flow supplying. The oxygen from the gas flow is driven to the liquid, which is available to cell uptake (Gao and Shi, 2013). Generally, the O_2 is kept at a certain level, > 20% of dissolved oxygen saturation in air (O_2^*), to ensure that *Pichia* grows without any oxygen limitation to metabolize glycerol or methanol (Invitrogen,

2014). In order to assist the oxygen transfer, agitation and aeration are controlled between 500-1500 rpm and 0.1-2 vvm respectively.

1.4.2. Fed-batch cultivation

The most applied cultivation strategy to achieve high cell densities and high heterologous protein production levels with P_{AOX1}-based systems is the fed-batch operation. Typically, such processes are divided into three phases: glycerol batch phase (GBP), transition phase (TP), and finally the methanol induction phase (MIP).

The objective of the GBP is the fast generation of biomass previous to the induction by methanol. Glucose is an alternative carbon source for GBP, but it is lightly utilized because has a repressing effect on the heterologous expression during MIP (Gao and Shi, 2013). The specific growth rate and biomass yield of *P. pastoris* growing on glycerol are higher than methanol. The maximum specific growth rate of wild type *P. pastoris* growing on glycerol (0.18-0.20 h^{-1}) (Cos et al., 2005a; Garcia-Ortega et al., 2013) is higher than growing on methanol (0.14 h^{-1}) and this methanol μ_{max} is, generally, lower when *Pichia* is producing a heterologous protein because of the negative effect that heterologous protein production has on the microorganism's growth (Cos et al., 2006b). GBP strategy is independent of the *Pichia* phenotype under P_{AOX1} and, in general, the initial concentration of glycerol used is about 40 g/L. This concentration is selected because a glycerol concentration over 40 g/L could inhibit growth (Invitrogen, 2014). The observed biomass to substrate yield ($Y_{X/S}$) (biomass expressed as dry cell weight) was about 0.5. At the end of GBP, the final biomass concentration is around 20 g/L. During GBP (which takes 18–24 h), the O_2 is controlled by adjusting the agitation rate up to 800 rpm for a typical 4-L bioreactor,

and around 450 rpm for 60-L reactor, and then manually adjusting the air (or oxygen-enriched-air) flow rate into bioreactor or by controlling at O_2 set-point about 20% of O_2^*. Once the GBP is finished, indicated by a spike in measured O_2, the TP starts (Zhang et al., 2007).

Although the MIP may be carried out after GBP (Invitrogen, 2014) the TP has been recommended to achieve both high cell density cultures and the derepression of the P_{AOX1}, under glycerol limiting conditions in presence of methanol, in order to adapt the cell metabolism to the induction of P_{AOX1} (Cos et al., 2006b).

Some authors selected different strategies for TP, constant (Cos et al., 2005b) or exponential feeding rates (Zhang et al., 2000) with glycerol as sole carbon source. However, during the TP, glycerol feeding rate is generally complemented with a methanol feeding rate to start the derepression and the induction of P_{AOX1}. The procedure consists of a glycerol feeding by decreasing the flow rate during 3–5 h and adding methanol with a pulse (Zhang et al., 2000) or with a low feed rate (Cos et al., 2005b). The final biomass levels reached at the end of the TP depends on the authors but generally is above 30 $g \cdot L^{-1}$. O_2 is usually maintained about 30% of O_2^*.

Finally, in the MIP, methanol is used as the carbon and inducer substrate. Although MIP may depend on the operational conditions (e.g. temperature, pH, and culture medium), phenotype and specific characteristics of the heterologous protein produced, the selection of a methanol feeding strategy is one of the most important factors to maximize heterologous protein production (Zhang et al., 2007).

1.4.3. Monitoring and control the bioprocesses

The real-time monitoring of the key fermentation variables (biomass, substrate, and products) is of major relevance to allow controlling and also improving the overall productivity of the bioprocesses. Furthermore, monitoring and controlling the biosystem guaranties the reproducibility for the final expected product quality and quantity (Kaiser et al., 2008).

However, some bioprocess variables are not measurable with standard devices (Veloso et al., 2009). For example, biomass concentration may not be measured online during the production to sufficient accuracy (Jenzsch et al. 2006). When the online measurement of any state variable or specific rate is not available, robust and reliable algorithms can be used for on-line estimation of non-measured key variables and parameters (Veloso et al., 2009; de Battista et al., 2011; Lyubenova et al., 2011).

Biomass growth rate and biomass concentration have been estimated by using different estimators based on online measurements like dissolved oxygen concentration and carbon dioxide concentration in the exhaust gases (Lubenova et al., 2003). An accurate control of the glycerol feeding rate has been developed by using adaptive controller dissolved oxygen measurement (Oliveria et al., 2005). On the other hand, the monitoring of different variables can support some control of the process, aiding on the adjustment of the control action, so, enhancing the bioprocess stability (Gao et al., 2013).

1.4.4. Operational strategies

Most common standard operational strategies are based on the control of the substrate concentration close to zero (limiting strategies). However, maintenance of the concentration at a constant value is also extensively described (non-limiting strategies) (Cos et al., 2006b, Jahic et al., 2006, Zhang et al., 2007). Since aerobic conditions are strictly essential for methanol assimilation, there are at least two possible limiting substrates for the process: oxygen and methanol.

1.4.4.1. Operational strategies based on oxygen

1.4.4.1.1. Dissolved oxygen control (DO-stat).

DO-stat is a control method that adjusts the methanol feeding depending on the O_2 in the medium. O_2 must be kept above a minimal level, generally set at 20% but excess of oxygen, high O_2 levels, is cytotoxic and significantly reduce cell viability (Potvin et al., 2012, Invitrogen, 2014).

Lim et al., (2003) developed DO-stat control that automatically handles the partial pressure of oxygen in the inlet air stream and the methanol feeding rate during induction. Alternatively, Yamawaki et al., (2007) controlled the O_2 by adjusting the methanol flow rate. Although DO-stat systems are simple to operate, they are not reliable. Methanol limitation or methanol excess may occur by the O_2 controller due to the complexity to tune correctly in systems that present time-varying consumption dynamics.

The weakest point of this kind of control is that it is quite difficult to obtain bioprocess reproducibility, because some key variables and specific rates, such as methanol concentration and specific growth rate (μ) are not constant (Cos et al., 2006b).

1.4.4.1.2. Oxygen limited fed-batch (OLFB),

OLFB is a control method that adjusts the methanol consumption by means of the oxygen limitation, consequently by the oxygen transfer rate in the *P. pastoris* P_{AOX1}-based systems (Khatri and Hoffmann, 2006). Although oxygen limitation should generally be avoided during the induction phase (Invitrogen, 2014), OLFB cultivations have been successfully implemented (Potvin et al., 2012). In OLFB, the residual methanol concentration is kept constant at about 1%, but O_2 concentration always drops to 0% due to oxygen limitation (Charoenrat, et al., 2005; Khatri and Hoffmann 2006, Berdichevsky et al., 2011).

The main advantage of OLFB is that this control strategy minimizes oxygen requirements and, therefore, may improve the economic feasibility of the process by reducing the cost on oxygen (Khatri and Hoffmann 2006). However, the methanol requirements were increased in these oxygen-limited conditions (Potvin et al., 2012).

OLFB cultivations have been emerging as promising alternatives to MLFB in the field of production of a monoclonal antibody with glycoengineered *Pichia pastoris*. Oxygen limitation reduces post-translational product modifications. The purity of product secreted (N-glycan composition, galactosylation, and fragmentation) may increase depending on the oxygen uptake rate (Potgieter et al., 2009; Berdichevsky et al., 2011).

Specific OLFB is the hypoxic fed-batch cultivation. Hypoxic conditions have been implemented to chemostat and fed-batch cultivations of *Pichia* expressing an antibody fragment (Fab) under the *GAP* promoter (Bauman et al., 2008). The oxygen concentrations in the inlet air represent following conditions: fully aerobic (20.97%), limited aerobic (10.91%), and hypoxic (8.39%, 5.87%). In both limited and hypoxic conditions the O_2 was 0%, whereas in fully aerobic conditions O_2 was about 45%. Limited aerobic conditions still allow fully aerobic glucose metabolism while no oxygen remains in the culture, whereas hypoxic conditions lead to partially fermentative metabolism. At the same time the biomass yield decreased and ethanol was produced, indicating a shift from oxidative to oxidofermentative condition. In chemostat and fed-batch cultivations, productivity and specific productivity was increased at hypoxic conditions (Bauman et al., 2008).

1.4.4.2. Operational strategies based on methanol

1.4.4.2.1. Methanol limited fed-batch (MLFB)

MLFB strategy is based in a control method that adjusts the methanol feeding rate based on mass balance equations to maintain a constant specific growth rate (μ) under methanol limiting conditions. So, no accumulation of methanol should be observed (Cos et al., 2005a). This method is the most widely reported for heterologous protein production by *P. pastoris,* although it can be referred to as specific growth rate control (Dabros et al., 2010), open loop control (Cos et al., 2006a) or pre-programmed exponential feeding rate addition (Cos et al., 2006b).

This feed-forward control allows considering a simple cell growth model and does not require on-line information about the system. It is assumed both that the residual

methanol concentration on the cultivation broth is close to zero and that a constant biomass/substrate yield is constant during the induction phase.

Although MLFB strategy can be easy to implement, they do not respond to any perturbations of the bioprocess. To avoid this problem the set point of μ is fixed far enough from the μ_{max} diminishing the productivity of the process. However, a feedback term can be introduced to minimize deviations of the methanol feeding rate and to compensate any perturbations (Dabros et al., 2010).

On the other hand, the substrate uptake rate (q_S), which is a rate generally μ-linearly dependent, is also proposed to pre-program the exponential methanol feeding rate addition (Zalai et al., 2012; Spadiut et al., 2014). Although some alternative MLFB feeding conditions were developed (stepwise q_S ramp or linear increasing q_S) as potential optimal feeding, the classical MLFB at constant specific rate (around the q_{Smax}) was the most productive.

The main advantage of MLFB is that keeping μ constant (or q_S which is μ-linear) improves process reproducibility and allows studying the μ effects on the heterologous protein production. The μ control is an effective strategy for bioprocess optimization because production is either directly or indirectly associated with cell growth (Potvin et al., 2012).

1.4.4.2.2. Methanol non-limited fed-batch (MNLFB)

MNLFB is a control strategy that consists in keeping constant the methanol concentration [MeOH]. An accurate methanol monitoring, for example by methanol

sensors based on liquid-gas equilibrium (Katakura et al., 1998; Ramón et al., 2004), and efficient control are required for robust and reproducible bioprocesses.

A feedback methanol control strategy has been used in numerous studies (Zhang et al 2002; Cos et al., 2006a; Pla et al. 2006). Some authors implemented the MNLFB strategy applying the simplest feed-back control, the on-off control, by different authors (Katakura et al., 1998; Zhang et al., 2000). However, this sort of control is only suitable for linear systems while heterologous protein production in *Pichia* is considered more complex and highly non-linear process (Potvin et al., 2012). This control strategy is inadequate for precise control of methanol concentration in a high methanol demanding P_{AOX1}-system.

Proportional–integral (PI) or proportional–integral–derivative (PID) and control heuristics are reported to be effective to maintain constant the methanol concentration (Potvin et al., 2012). Nevertheless, the optimal settings of the PID controller (gain K_C, the integral time constant τ_I and the derivative time constant τ_D) are hardly ascertained by trial and error tuning or other empirical methods. Some authors have developed a PID control Bode stabilization criterion to achieve the parameters associated to this kind of control, obtaining good results on methanol regulation in short time fermentations (Zhang et al., 2002).

A predictive control algorithm combined with a PI feedback controller was used to optimize the production of *Rhizopus oryzae* lipase in *P. pastoris* (Cos et al., 2006a). Different methanol concentrations directly affect cell growth, the cell viability and the heterologous protein production (Jahic et al., 2006). Thus, an accurate monitoring and

control allow determining stoichiometric and kinetic models that are particularly useful to optimize the protein production and the productivity (Schenk et al., 2007).

A comparison between DO-stat and MNLFB strategies in both fed-batch and continuous cultures for recombinant scFv expression was performed by Yamawaki et al., (2007). Higher specific productivity of scFv was obtained at [MeOH]=3.9 $g \cdot L^{-1}$. Nevertheless, this approach, an extensive analysis comparing operational modes described previously for bioprocess optimization and how affects the recombinant expression in *P. pastoris*, is rarely published. On the other hand, the high level of protein expression in several proteins using the MNLFB strategy, make it a good choice for process optimization.

1.4.4.3. Alternative operational strategies

1.4.4.3.1. Temperature limited fed-batch (TLFB)

TLFB is a control strategy that consists in keeping the O_2 constant, like DO-stat, through temperature limitation. During the induction stage, the temperature controller is programmed to use $O_{2,set\text{-}point}$=25%. When O_2 is below the set point the temperature decreases, and it increases when the O_2 is above the set point (Jahic et al., 2006).

In Surribas et al., (2007), the transition phase length was designed to reach a final biomass level about 54 $g \cdot L^{-1}$, when the oxygen demand could no longer be supported. Similar biomass concentration was achieved at the end of the glycerol phase in Jahic et al., (2003). Then, the induction phase started and methanol was fed to the culture as the sole carbon source. Additionally the methanol concentration was controlled via outlet gas analysis at 300 $mg \cdot L^{-1}$ (Jahic et al., 2003) and 3 $g \cdot L^{-1}$ (Surribas et al., 2007).

TLFB process was compared with a methanol limited fed-batch by Jahic et al. (2003). By reducing the temperature from 30ºC to 25ºC, the product amount was increased 2-fold. However, comparison between MNLFB and TLFB revealed that methanol control at high concentration (3 $g \cdot L^{-1}$) achieves a high ROL production and productivities (Surribas et al., 2007). Additionally, some adjustments to the standard process have been proposed to minimize proteolysis, including the manipulation of culture media and temperature (Surribas et al., 2007).

1.4.4.3.2. Mixed substrates

In general, all feeding strategies can be applied to bioprocess involving whichever phenotype (Mut^+ and Mut^S). However, multicarbon substrate feeding is usually carried out in order to optimize bioprocess using Mut^S phenotype (Cos et al., 2006b, Potvin et al., 2012; Valero et al., 2013). Mixed substrates strategy requires methanol, because it is the inductor, and a supplementary carbon source. The complementary substrate such as glycerol and sorbitol do not participate in methanol metabolism, so neither activate AOX nor induce the heterologous protein expression, but generates energy supply (NADH/ATP) via TCA cycle/oxidative phosporylation reaction (Gao and Shi, 2013).

Glycerol has been proposed to feed as supplementary carbon source in order to improve the productivity using Mut^+ phenotype. Theoretically, if *P. pastoris* grows faster utilizing glycerol than methanol, the productivity should increase. However, the high specific growth rates reduce the productivity of the bioprocess (d'Anjou and Dagoulis, 2001). On the other hand, the strong point of this strategy is that, when the glycerol is utilized as co-substrate, the system reduces the heat production and the oxygen consumption rates, both aids reduce process costs (Valero, 2013).

Although it has reported that glycerol represses the expression of AOX (Potvin et al., 2012), glycerol and methanol substrate feeding in fed-batch fermentation have been designed. This strategy is more frequently applied for Mut^S strains. For example, the novel pre-programmed exponential feeding rate based on q_S (q_S stepwise ramp for glycerol and constant q_S for methanol) was used to maximize the horseradish peroxidise (HRP) production (Zalai et al., 2012). However, glycerol also has been applied in cultures with Mut^+ phenotype (Zalai et al., 2012, Valero et al., 2013).

A pre-programmed exponential feeding rate with a different methanol-glycerol ratio was designed by Jungo et al. (2007a). Similar strategy, exponential glycerol feeding rate maintaining a residual methanol concentration between 1- 2 $g \cdot L^{-1}$ was proposed by d'Anjou and Dagoulis, (2001). The maximum specific growth rate of *P. pastoris,* utilizing glycerol as a single carbon source, is about 0.2 h^{-1}. However, the optimum of μ was 0.06 h^{-1} in Mut^+ phenotype, applying co-substrate feeding. Different ratios were proposed to maximize the q_P but similar production was found for sole methanol feeding or co-substrate feeding.

In contrast with glycerol, sorbitol is a non-repressing carbon source of P_{AOX1} up to40 $g \cdot L^{-1}$ is less critical than mixed feeds of glycerol and methanol because it does not affect the expression level of recombinant protein (Resina et al., 2004, 2005; Jungo et al., 2007b). The main disadvantage for sorbitol utilization is that μ_{max} is about 0.03 h^{-1} (Jungo et al., 2007b; Niu et al., 2013). This μ is too low in contrast to glycerol, but similar to μ of *P. pastoris* with Mut^S phenotype growing on methanol as sole carbon source.

Different operational strategies have been implemented using sorbitol as co-substrate (Jungo et al., 2007b; Çelik et al., 2009; Arnau et al., 2011). In case of Mut^+ strains, mixed substrates strategy with sorbitol could increase the production levels. However, the productivity may be adversely affected because fermentations may be longer than with other strategies. In the case of Mut^S, productivities were increased by sorbitol (Arnau et al 2010).

Arnau et al., (2010, 2011) designed an operational strategy using a Mut^S phenotype comparing both co-substrates sorbitol and glycerol in the production of *Rhizopus oryzae* lipase. To sum up, sorbitol presented better higher productivities than glycerol as co-substrate in the heterologous production of *Rhizopus oryzae* lipase. In addition, proteolytic activity was detected only when glycerol was used as co-substrate.

The carbon source cost per batch was a saving of approximately 70% when glucose was used instead of glycerol as the sole carbon source. Nevertheless, glucose has been rarely exploited as a supplementary carbon source substrate for *P. pastoris* Mut^+ phenotype strains cultivation (Paulová et al., 2012). The main reason is because glucose acts as a strong repressor of P_{AOX1} at the transcription level (Cos et al., 2006b). The repression with glucose is higher than glycerol. Complete diauxic growth was observed in batch with glucose-methanol and partially with glycerol-methanol (Inan and Meagher, 2001; Hang et al., 2009).

The glucose-methanol mixed substrates strategy has been explored in continuous cultures for Mut^+ phenotype (Paulová et al., 2012, Jordà et al., 2012). Different carbon contribution, (80-60%.) glucose and (20-40%) methanol was added in continuous

cultures for producing recombinant. High production of trypsinogen was performed at μ up to 0.15 h^{-1}. This strategy allowed increasing μ 1.3 times, to a value of $\mu_{max,met}$ (=0.12 h^{-1}) achieved with methanol as sole carbon source (Paulová et al., 2012).

In the case of Mut^S phenotype, a real-time parameter-based controlled glucose feeding strategy has been implemented for recombinant production of phytases in fed-bach. First, dissolved oxygen (O_2) was the controlled variable. O_2 was set at 20% until the ethanol-fermentative metabolism appeared at RQ<0.8. RQ was utilized as to re-adjust glucose feed rate by controlling RQ≥0.9. This glucose feeding addition strategy aided *Pichia* to assimilate methanol without glucose repression. The fed-batch reached a μ about 0.11 h^{-1} during the induction phase (Hang et al., 2009).

Glucose-methanol mixed substrate strategy allows producing protein at higher growth rates than other mixed substrate strategies. Methanol seems to play a key role as an auxiliary substrate to compensate for the increased energy demands derived from recombinant protein secretion and favouring metabolic adaptation to the new requirements (Jordà et al., 2012). Nevertheless, the main troubles to implement this strategy in fed-batch cultures are related to the metabolic glucose assimilation. A really good control system performance is required to avoid the P_{AOX1} repression by glucose, which provokes diauxic growth pattern or even cell death; and also avoids the ethanol sub-product formation (Hang et al., 2009).

1.5. Target protein: Lipase from *Rhizopus oryzae*

Rhizopus oryzae is a filamentous and lipolytic fungus which has been isolated from olive oil and seed industries. This microorganism only produces one form of extracellular lipase which has a high biotechnological potential as a catalyst for lipid modification, due to its high regiospecificity (Salah et al., 2006).

Lipases are enzymes that catalyze the hydrolysis of the ester bound of triacylglycerols. Lipases are found in all species of the animal kingdom as well as in plants and microorganisms such as yeast bacteria and fungi. The three-dimensional structure of many lipases of microbial is common, α/β hydrolase fold (Ollis et al., 1992). The region of highest conservation is the active site, which contains a 'classical' Ser-His-Asp catalytic triad, and residues involved in the oxyanion hole. In some cases the active site can be covered by an amphiphilic loop (lid or flap), which prevents access of the substrate (Salah et al., 2006).

Rhizopus oryzae lipase (ROL) acts only at the sn-1 and sn-3 locations because it belongs to a group of lipases that are active against esters of primary alcohols. (Guillén et al., 2011). The native structure of the enzyme is a 392 amino acids protein. The first 26 belong to a signal sequence (pre-region), the following 97 amino acids belong to the pro-region and the last 269 form the mature protein sequence. The deduced polypeptide sequence of the secreted form (ProROL) was found to be made up of 297 amino acids of which 269 belong to the mature protein sequence and the remaining 28 amino acids to the last part of the pro-region (Sayari et al., 2005). The molecular weight of this enzyme is 32 kDa and the isoelectric point is 6.85. It has four

potential sites of N-linked glycosylation and three disulphide bonds, between the amino acids 152 and 391, 163 and 166, and the 358 and 367 (Guillén et al., 2011).

ROL has been cloned and expressed in *Escherichia coli* (Di Lorenzo et al., 2005), *Saccharomyces cerevisiae* (Takahashi et al., 1998) and in the methylotrophic yeast *P. pastoris* (Cos et al 2005b). In *E. coli*, ProROL could be efficiently produced in high yield at high specific activity (166.0 U·mL) (Di Lorenzo et al., 2005). However, this expression system has the bottleneck that a cell disruption process is necessary, when compared with other expression systems in which the enzyme is secreted in the culture broth (Whang et al., 2013). In *S. cerevisiae*, ProROL could be secreted in the culture broth, but the activity of ProROL was very low (extracellular lipase activity reached only 2.9 U·mL at 120 h of cultivation in Yeast Extract Peptone Dextrose (YPD) medium (Takahashi,.et al., 1998). The mature sequence of ROL has been expressed in *P. pastoris* under the P_{AOX1} and secreted into the culture medium using the *S. cerevisiae* alpha-mating factor pre-prosequence obtaining up to 600 U·mL (Surribas et al., 2007, Arnau et al., 2010).

1.6. Background of recombinant ROL production by *Pichia pastoris.*

Heterologous production of *Rhizopus oryzae* lipase in *P. pastoris* has been widely studied by our research group. Different promoters, operational modes, strategies, an media have been explored with the aim of establishing specific protocols to improve the bioprocess reproducibility and also to increase the ROL production levels. A brief summary that includes all the information about ROL production is presented in Table 1.

The first approach was to test the ROL expression in batch cultures of *P. pastoris* under two different regulated promoters (P_{AOX1} and P_{FLD1}). These preliminary results showed a higher ROL production levels with P_{FLD1} than with P_{AOX1} Mut^+ system. Conversely, the higher productivity was observed under the P_{AOX1} Mut^+ system. Secondly, both promoters (P_{AOX1} and P_{FLD1}), including both phenotypes for P_{AOX1} Mut^+ and Mut^S were evaluated in fed-batch cultures. The Mut^S P_{AOX1}-system performed the higher ROL production. However, in terms of ROL productivity, the Mut^+ P_{AOX1}-system was the most productive. Although the strategies that were carried out for fed-batch were not optimal (e.g. heuristic control), the production was highly increased related to batch operation mode (Cos et al., 2005b).

In case of P_{FLD1}, a previous work was published (Resina et al., 2004), where, different inductors of the ROL expression under P_{FLD1} were tested. The inducers were the carbon or nitrogen source, or both. In the following approaches, the use of methanol was avoided for *FLD1* induction by nitrogen source. The strategy of substrate (sorbitol) limited fed-bacth was applied, being methylamine hydrochloride the nitrogen source and the inductor of ROL expression. The operational condition range for SLFB was: $\mu_{set\text{-}point}$ = 0.005, 0.01 and 0.02 h^{-1}; all lower than μ_{max} for sorbitol consumption. The ROL activity and productivity levels were increased related with $\mu_{set\text{-}point}$ =0.02 h^{-1} operational conditions (Resina et al., (2005).

In order to investigate how is the effect of *AOX* gen dosage for ROL production, a comparison of ROL expression between strains growing on methanol as single carbon source with different phenotype (Mut^+ and Mut^S) and different *ROL* gene copy number in fed-bach was carried out in Cos et al. (2005b). MNLFB strategy was applied for all

fed-batch cultures. The heuristic control of methanol in cultures with Mut^S strains showed better performance, lower fluctuations, than in the Mut^+ ones. Although the ROL concentration at the end of the fed-batch was higher for both multicopy strains (Mut^+ and Mut^S) compared to respective singlecopy, the productivity was better for both single copies (Mut^+ and Mut^S). Regarding the effect of phenotype, the specific productivity and the $Y_{P/X}$ were higher for Mut^s single copy strain than for Mut^+ single copy strain. On the contrary, the Mut^+ single copy strain was the most productive.

Although the use of the heuristic methanol control was a successful attempt for ROL production by *P. pastoris*, variations on the inducer concentration could affect the ROL expression (Cos et al., 2005b). A predictive control algorithm coupled with a PI feedback controller was implemented in the heterologous ROL production by *P. pastoris* P_{AOX1} (Mut^S phenotype (Cos et al., 2006a). The effect of methanol concentration, ranging from 0.5 to 1.75 $g \cdot L^{-1}$, on the heterologous protein production was evaluated in MNLFB cultures. Optimal ROL production was carried out at 1.0 $g \cdot L^{-1}$ of residual methanol. The implementation of the control increased more than double the ROL production and the productivity.

Alternative operational strategies were designed for ROL production using the *P. pastoris* P_{AOX1} (Mut^+) system (Surribas et al., 2007). Operational strategies that combined non-limiting substrate conditions, neither methanol nor oxygen, was compared. First, an MNLFB+DOstat strategy was implemented. The methanol not limiting conditions with methanol set-point of 3 $g \cdot L^{-1}$ were carried out until oxygen became limiting. Consecutively DO-stat control with 25 % of dissolved oxygen was exploited. The second strategy was quite similar, TLFB, methanol was kept constant

around to 3 $g \cdot L^{-1}$ and O_2 to 25%, but manipulating the temperature of the culture. The third one MNLFB+TLFB was tested in similar conditions. MNLFB+DOstat was shown as the best strategy, which allowed improving the ROL production and productivity (Table 1).

Optimization of *P. pastoris* P_{AOX1} (Mut^S) system using mixed substrate was also investigated (Arnau et al., 2010; 2011). Sorbitol or glycerol were the selected as co-substrates. SLFB was designed for co-substrate feeding and MNLFB for methanol. The operational range was $\mu_{set\text{-}point}$ = 0.05, 0.01, 0.02 h^{-1} for co-substrate and $[MetOH]_{set\text{-}point}$ = 0.5, 2.0, 4.0 $g \cdot L^{-1}$. The methanol optimal condition was fixed on 2.0 $g \cdot L^{-1}$. Optimal sorbitol feeding was set at μ=0.01 h^{-1} and for glycerol was at μ=0.01 h^{-1}. The mixed substrate strategies showed higher production and productivities than MNLFB for P_{AOX1} (Mut^S) system, being sorbitol-methanol the best one for that system.

Finally, a series of ^{13}C-tracer experiments were performed on aerobic chemostat cultivations with ROL producing strains growing at a dilution rate of 0.09 h−1 using a glucose:methanol 80:20 (w/w) mix as carbon source (Jordà et al., 2012; 2013). The research was focused on the investigation of the relationships between central metabolism and protein production, particularly for the case of multicarbon (methanol-glucose) source metabolism, by ^{13}C-based metabolic flux analysis studies. The glucose-methanol mixed substrates strategy showed a good performance to increase the productivity of ROL and reduce metabolic burden in *P. pastoris*.

Table 1: Comparative analysis between the different promoters, operational strategies, carbon and nitrogen source for ROL expression in *Pichia pastoris*.

Reference	Promoter	Operational			Carbon source	Nitrogen source	Lipolytic activity [U·mL^{-1}]	Productivity [U·L^{-1}·h^{-1}]
		Mode	Strategy	Conditions				
Cos et al., 2005a	P_{FLD1}	Batch	-	-	Sorbitol	CH_3NH_2· HCl	15	145
Cos et al., 2005a	P_{AOX1} (Mut^+)	Batch	-	-	MeOH	$(NH_4)_2SO_4$	6	201
Cos et al., 2005a	P_{AOX1} (Mut^S)	Fed-batch	MNLFB	Heuristic control [MeOH]≈1g·L^{-1}	MeOH	NH_4OH	205	2246
Cos et al., 2005a	P_{AOX1} (Mut^+)	Fed-batch	MNLFB	Heuristic control [MeOH]<5g·L^{-1}	MeOH	NH_4OH	150	2879
Cos et al., 2005a	P_{FLD1}	Fed-batch	SLFB	$\mu_{set\text{-}point}$= 0.015h^{-1}	Sorbitol	CH_3NH_2· HCl	170	2402
Resina et al., 2005	P_{FLD1}	Fed-batch	SLFB	$\mu_{set\text{-}point}$= 0.02h^{-1}	Sorbitol	CH_3NH_2· HCl	385	4379
Cos et al., 2005b	P_{AOX1} (Mut^S)	Fed-batch	MNLFB	Heuristic control [MeOH]≈1-2g·L^{-1}	MeOH	NH_4OH	205	2246
Cos et al., 2005b	P_{AOX1} (Mut^S)^M	Fed-batch	MNLFB	Heuristic control [MeOH]≈1-2g·L^{-1}	MeOH	NH_4OH	270	1500
Cos et al., 2005b	P_{AOX1} (Mut^+)	Fed-batch	MNLFB	Heuristic control [MeOH]≈1-2g·L^{-1}	MeOH	NH_4OH	150	3000
Cos et al., 2005b	P_{AOX1} (Mut^+)^M	Fed-batch	MNLFB	Heuristic control [MeOH]≈1-2g·L^{-1}	MeOH	NH_4OH	175	1894
Cos et al., 2006	P_{AOX1} (Mut^S)	Fed-batch	MNLFB	[MeOH] $_{set\text{-}point}$= 1 g·L^{-1}	MeOH	NH_4OH	490	4901
Surribas et al., 2007	P_{AOX1} (Mut^+)	Fed-batch	MNLFB+ DO-stat	[MeOH] $_{set\text{-}point}$= 3 g·L^{-1} DO-stat $_{set\text{-}point}$= 25%	MeOH	NH_4OH	644	8110
Surribas et al., 2007	P_{AOX1} (Mut^+)	Fed-batch	DO-TLFB	[MeOH] $_{set\text{-}point}$= 3 g·L^{-1} DO-stat $_{set\text{-}point}$= 25%	MeOH	NH_4OH	534	5200
Surribas et al., 2007	P_{AOX1} (Mut^+)	Fed-batch	MNLFB+ DO-TLFB	[MeOH] $_{set\text{-}point}$= 3 g·L^{-1} DO-stat $_{set\text{-}point}$= 25%	MeOH	NH_4OH	713	5980
Arnau et al., 2010	P_{AOX1} (Mut^S)	Fed-batch	Mixed substrates	$\mu_{set\text{-}point}$=0.01h^{-1} [MeOH] $_{set\text{-}point}$= 2 g·L^{-1}	Sorbitol MeOH	NH_4Cl	621	6624
Arnau et al., 2011	P_{AOX1} (Mut^S)	Fed-batch	Mixed substrates	$\mu_{set\text{-}point}$= 0.02h^{-1} [MeOH] $_{set\text{-}point}$= 2 g·L^{-1}	Glycerol MeOH	NH_4Cl	471	5416

Note: strains with multiple *ROL* gen copies were identify by (^{M}), in other case single copy was used.

1.7. Problems and bottlenecks on protein production

Heterologous protein production in any expression system, also in *P. pasto*ris, includes the following steps: genetic modification of microorganism and expression cassette; bioprocess control and protein expression; product purification and recovery (Minjie and Zhonping, 2013). So, enhancing the recombinant protein production or maximization bioprocess productivity using *P. pasto*ris as host can be achieved by means of determining the best operational conditions to produce the target recombinant product using the standard operational modes (see 1.5.4 section) in singular conditions (Minjie and Zhonping, 2013; Valero, 2013) or time-varing (Zhang et al. 2005; Maurer et al., 2006; Spadiut et al., 2014); improving the engineered strain to enhance the production, by solving some bottleneck that prevent the heterologous product production or by search any metabolic or physiological condition that aids the production (Buchetics et al., 2011; Vogl and Glieder, 2013, Rebnegger et al., 2014).

The potential bottlenecks of *P. pasto*ris expression system in the field of genetic and metabolic engineering field are: codon usage of the expressed gene (Valero, 2013), gene copy number (Gasser et al., 2013), efficient transcription using strong promoters/translation signals (Hohenblum et al., 2004), processing, folding in the ER and secretion (Damasceno et al., 2012) or intracellular proteolytic degradation (Pfeffer et al., 2012). Genetically, maximize these bottlenecks by increasing the rate of one step, can lead to rate-limitation of another one, which can then become a bottleneck in the expression causing the cellular stress response (Gasser et al., 2008, Damasceno et al., 2012). Protein folding and conformational stress imposes a burden on the central metabolism, occurring even at modest production levels (Jordà et al. 2012). Unfolding protein response (UPR) is induced by an increasing of growth rate. Genes of

translocation protein folding and glycosylation and cytolosolic chaperones are upregulated in high growth rate conditions (Rebnegger et al., 2014).

However, synthetic biology and metabolic engineering approaches allow designing tailor-made production strains (Volg and Glieder, 2013). The production of ROL lipase under *PFLD1* triggers the unfolding protein response (UPR) detected at transcriptional levels (Resina et al., 2007). Engineered strains were developed to overcome this problem, which co-express constitutively transcription factor *pHAC1* and/or present a deletion of the *GAS1* gene. In any case, the mutants present between 3-fold to 7-fold increase of the production (Resina et al., 2009). Alternatively, overproduction of heterologous proteins affects the primary metabolism of the producing cells. Enhancement of productivity was predicted by simulating the gene targets for deletion or overexpression in a genome scale metabolic model. Five out of 9 tested single gene modifications (overexpression or knockout) led to a significant improvement of recombinant protein production (Nocon et al., 2014). Moreover, cell cycle and protein secretion are interrelated. At high secretion rates cultures contain a large fraction of cells in the G2 and M phases of cell cycle. The constitutive overexpression of the cyclin gene *CLB2* allows increasing the fraction of cells in G2+M phase. This engineered strain improved the productivity of an antibody Fab fragment by 18% and the product titer by 53% (Buchetics et al., 2011). These three cases are an example that how the co-expression of proteins or the deletion of genes affect to bioprocess engineering (Valero, 2013).

One of the main characteristics of *P. pastoris* is the capability to perform many of the higher eukaryotic post-translational modifications such as protein folding, proteolytic

processing, disulfide bond and glycosylation and finally secret heterologous protein simplifying their recovery (Valero et al., 2013). However, glycosylation pattern of *P. pastoris* is quite different from mammalian cells. For example, monoclonal antibodies produced in mammalian cells carry three major N-glycans: $Gal_2Glc\text{-}Nac_2Man_3F$ (G_2F), $Gal_1GlcNac_2Man_3F$ (G_1F), and $GlcNac_2Man_3F$ (G_0F), and glycoproteins derived from fungal expression systems contain non-human N-glycans of the high mannose (Man) type, which are immunogenic in humans, being cleared rapidly from the human bloodstream (Choi et al., 2003; Ye et al., 2011). Nevertheless, glycoengineered *P. pastoris* expression system allows the production of human proteins with complex N-glycosylation modifications, mimicking the human N-glycosylation pathway (Potgieter et al., 2009, Berdichevsky et al., 2011).

Operational modes and operational conditions are usually evaluated to increase the production in *Pichia* cell factory. However, the substrate providing to the cells (specifically oxygen) may be the key to optimize the bioprocess (Potvin et al., 2012; Gao and Shi, 2013).

P. pastoris metabolises methanol by the oxidative pathway only in the presence of enough amounts of oxygen. Thus, the oxygen level in the culture and the methanol uptake rate are interrelated (Lim et al., 2003). High methanol supply may lead to sudden oxygen depletion. The oxygen limitation may negatively affect the expression of foreign genes (Cereghino and Cregg, 2000). So, for many years, avoiding oxygen limitation in MLFB and MNLFB cultures has been a priority (Khatri and Hoffmann 2006, Potvin et al., 2012, Invitrogen, 2014). Due to the bioreactor oxygen transfer capacity is unable to sustain the oxygen metabolic demand only flowing air, oxygen-

enriched gas mix is usually provided instead of air (Cos et al., 2006b, Jahic et al., 2006). The use of pure oxygen to perform oxygen-enriched gas mix increases the production costs and may cause difficulties in scale-up (Khatri and Hoffmann 2006, Potvin et al., 2012, Invitrogen, 2014). In order to determine if the system may affected by oxygen depletion, characterization of bioreactor oxygen transfer capacity is required (García-Ochoa et al., 2010). However, successful protein production has also been completed under oxygen limiting conditions (Charoenrat, et al., 2005; Khatri and Hoffmann 2006, Berdichevsky et al., 2011). Thus, OLFB cultivation strategy emerges as smart alternative to eliminate the drawbacks of the oxygen limitation.

1.8 References

1. Ahn J, Hong J,Lee H, Park M, Lee E, Kim C, Choi E, Jung J, Lee H. 2007. Translation elongation factor 1-α gene from *Pichia pastoris*: molecular cloning, sequence, and use of its promoter. Appl Microbiol Biotechnol. 74:601–608.
2. Ahn J, Hong J, Park M, Lee H, Lee E, Kim C, Lee J, Choi E, Jung J, Lee J. 2009. Phosphate-responsive promoter of a *Pichia pastoris* sodium phosphate symporter. Appl Environ Microbiol. 75:3528.
3. Arnau C, Ramon R, Casas C, Valero F. 2010. Optimization of the heterologous production of a *Rhizopus oryzae* lipase in *Pichia pastoris* system using mixed substrates on controlled fed-batch bioprocess. Enzyme Microb Techl. 46:494–500.
4. 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.
5. Baumann K, Maurer M, Dragosits M, Cos O, Ferrer P, Mattanovich D. 2008. Hypoxic fed-batch cultivation of *Pichia pastoris* increases specific and volumetric productivity of recombinant proteins. Biotechnol Bioeng. 100:177–183.
6. Berdichevsky M, d'Anjou M, Mallem MR, Shaikh SS, Potgieter TI. 2011. Improved production of monoclonal antibodies through oxygen-limited cultivation of glycoengineered yeast. J Biotechnol. 155:217–224.
7. Bill RM. 2014. Playing catch-up with *Escherichia coli*: using yeast to increase success rates in recombinant protein production experiments. Front Microbiol.5:1-5.
8. Buchetics M, Dragosits M, Maurer M, Rebnegger C, Porro D, Sauer M, Gasser B, Mattanovich D. 2011. Reverse engineering of protein secretion by uncoupling of cell cycle phases from growth. Biotechnol Bioeng. 108:2403-2412.
9. Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol Rev. 24:45–66.
10. Charoenrat T, Ketudat-Cairns M, Stendahl-Andersen H, Jahic M, Enfors SO. 2005. Oxygen-limited fed-batch process: an alternative control for *Pichia pastoris* recombinant protein processes. Bioprocess Biosyst Eng. 27:399–406.
11. Chen R. 2012. Bacterial expression systems for recombinant protein production: *E. coli* and beyond. Biotechnol Avd. 30:1102–1107.

12. Choi BK, Bobrowicz P, Davidson RC, Hamilton SR, Kung DH, Li H, Miele RG, Nett JH, Wildt S, Gerngross TU. 2003. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. Proc Natl Acad Sci USA. 100:5022–5027.
13. Corchero JL, Gasser B, Resina D, Smith W, Parrilli E, Vázquez F, Abasolo I, Giuliani M, Jäntti J, Ferrer P, Saloheimo M, Mattanovich D, Schwartz S Jr, Tution ML, Villaverde A. 2013. Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. Biotechnol Adv. 31:140-153.
14. Cos O, Resina D, Ferrer P, Montesinos J, Valero F. 2005a. Heterologous production of *Rhizopus oryzae* lipase in *Pichia pastoris* using the alcohol oxidase and formaldehyde dehydrogenase promoters in batch and fed-batch cultures. Biochem Eng J. 26:86–94.
15. Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005b. 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.
16. Cos O, Ramon R, Montesinos JL, Valero F. 2006a. A simple model-based control for *Pichia pastoris* allows a more efficient heterologous protein production bioprocess. Biotechnol Bioeng. 95:145–154.
17. Cos O, Ramón R, Montesinos JL, Valero F. 2006b. 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.
18. Cox MMJ. 2012. Recombinant protein vaccines produced in insect cells. Vaccine. 30:1759– 1766.
19. Cregg JM. Heterologous protein expressed in *Pichia pastoris*, available online: http://www.kgi.edu/documents/faculty/James_Cregg/heterologous_proteins_expressed_in_Pichia_pastoris.pdf (Accessed 06 December 2014).
20. Curvers S, Brixius P, Klauser T, Thömmes 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.

21. Curvers S, Klauser T, Wandrey C, Takors R. 2002. Recombinant protein production with *Pichia pastoris* in continuous fermentation - Kinetic analysis of growth and product formation. Eng Life Sci. 2:229–235.
22. Çelik E, Çalık P, Oliver SG. 2009. A structured kinetic model for recombinant protein production by Mut^+ strain of *Pichia pastoris*. Chem Eng Sci. 64:5028–5035.
23. Çelik E, Çalık P. 2012. Production of recombinant proteins by yeast cells. Biotechnol Adv. 30:1108–1118
24. d'Anjou MC, Daugulis a J. 2000. Mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast. Biotechnol Lett. 22: 341–346.
25. d'Anjou MC, Daugulis a J. 2001. A rational approach to improving productivity in recombinant *Pichia pastoris* fermentation. Biotechnol Bioeng. 72:1–11.
26. Dabros M, Schuler MM, Marison IW. 2010. Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. Bioprocess Biosyst Eng. 33:1109–1118.
27. Damasceno LM, Huang CJr, Batt CA. 2012. Protein secretion in *Pichia pastoris* and advances in protein production. Appl Microbiol Biotechnol. 93:31–39.
28. de Almeida JRM, de Moraes LMP, Torres FAG. 2005. Molecular characterization of the 3-phospoglycerate kinase gene (*PGK1*) from the methylotrophic yeast *Pichia pastoris*. Yeast. 22:725-737.
29. de Battista H, Picó J, Garelli F, Vignoni A. 2011. Specific growth rate estimation in (fed-)batch bioreactors using second-order sliding observers. J Process Control 21:1049–1055.
30. Delic M, Gasser B, Mattanovich D. 2013. Repressible promoters – A novel tool to generate conditional mutants in Pichia pastoris. Microb cell fact. 12:6.
31. Demain AL, Vaishnav P. 2009. Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv. 27: 297–306.
32. 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.
33. Di Lorenzo M, Hidalgo A, Haas M, Bornscheuer UT. 2005. Heterologous production of functional forms of *Rhizopus oryzae* lipase in *Escherichia coli*. Appl Environ Microbiol. 12:8974–8977.

34. Duan HM, Umar S, Hu YP, Chen JC. 2009. Both the *AOX1* promoter and the *FLD1* promoter work together in a *Pichia pastoris* expression vector. World J Microbiol Biotechnol. 25:1779–1783.
35. Ferrer-Miralles N, Domingo-Espín J, Corchero JL, Esther Vázquez E, Villaverde A.2009. Microbial factories for recombinant pharmaceuticals. Microb Cell Fact. 8:17.
36. Fischer R, Schillberg S, Hellwig S, Twyman RM, Drossard J. 2012. GMP issues for recombinant plant-derived pharmaceutical proteins. Biotechnol Adv.30:434–439
37. 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.
38. Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E, Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A. 2008. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. Microb Cell Fact 7:11
39. Gasser B, Prielhofer R, Marx H, Maurer M, Nocon J, Steiger M, Puxbaum V, Sauer M, Mattanovich D. 2013. *Pichia pastoris*: protein production host and model organism for biomedical research. Future Microbiol. 8:191–208.
40. Gao MJ, Shi Z. 2013. Process control and optimization for heterologous protein production by methylotrophic *Pichia pastoris*. Chinese J Chem Eng. 21:216-226.
41. Gao MJ, Zhan XB, Zheng ZY, Wu JR, Dong SJ, Li Z, Shi ZP, Lin CC. 2013. Enhancing pIFN-α production and process stability in fed-batch culture *of Pichia pastoris* by controlling the methanol concentration and monitoring the responses of OUR/DO levels. Appl Biochem Biotechnol. 171:1262–1275.
42. Gomord V, Fitchette AC, Menu-Bouaouiche, Saint-Jore-Dupas C, Plasson C, Michaud D, Faye L. 2010. Plant-specific glycosylation patterns in the context of therapeutic protein production. Plant Biotechnol J. 8:564–587
43. Gopal GJ, Kumar A. 2013. Strategies for the production of recombinant protein in *Escherichia coli*. Protein J. 32::419-425.

44. Guillén M, Benaiges MD, Valero F. 2011. Comparison of the biochemical properties of a recombinant lipase extract from *Rhizopus oryzae* expressed in *Pichia pastoris* with a native extract. Biochem Eng J. 54:117–123.
45. Hang H, Ye X, Guo M, Chu J, Zhuang Y, Zhang M, Zhang S. 2009. A simple fermentation strategy for high-level production of recombinant phytase by *Pichia pastoris* using glucose as the growth substrate. Enzyme Microb Technol. 44:185–188.
46. 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.
47. Heyland J, Fu J, Blank LM, Schmid A. 2011. Carbon metabolism limits recombinant protein production in *Pichia pastoris.* Biotechnol. Bioeng. 108:1942–1953.
48. 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.
49. Invitrogen. 2014. *Pichia* Fermentation Process Guidelines, available online: http://tools.lifetechnologies.com/content/sfs/manuals/pichiaferm_prot.pdf (Accessed 6 December 2014).
50. Jahic M, Wallberg F, Bollok M, Garcia P, Enfors SO. 2003. Temperature limited fed-batch technique for control of proteolysis in *Pichia pastoris* bioreactor cultures. Microb Cell Fact. 2:6.
51. Jahic M, Veide A, Charoenrat T, Teeri T, Enfors SO. 2006. Process technology for production and recovery of heterologous proteins with *Pichia pastoris*. Biotechnol Prog. 22:1465–1473.
52. Jeandet P, Vasserot Y, Chastang T, Courot E. 2013. Engineering microbial cells for the biosynthesis of natural compounds of pharmaceutical significance. Biomed Res Int. 27:780145.
53. Jenzsch M, Simutis R, Eisbrenner G, Stückrath I, Lübbert A. 2006. Estimation of biomass concentrations in fermentation processes for recombinant protein production. Bioprocess Biosyst Eng. 29:19–27.

54. Jordà J, Jouhten P, Cámara E, Maaheimo H, Albiol J, Ferrer P. 2012. Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose:methanol mixtures. Microb Cell Fact. 11:57.
55. Jordà J, Suarez C, Carnicer M, ten Pierick A, Heijnen JJ, van Gulik W, Ferrer P, Albiol J, Wahl A. 2013. Glucose-methanol co-utilization in *Pichia pastoris* studied by metabolomics and instationary ^{13}C flux analysis. BMC Syst Biol. 7:17.
56. Jungo C, Marison I, von Stockar U. 2007a. Mixed feeds of glycerol and methanol can improve the performance of *Pichia pastoris* cultures: A quantitative study based on concentration gradients in transient continuous cultures. J Biotechnol. 128:824–837.
57. Jungo C, Schenk J, Pasquier M, Marison IW, von Stockar U. 2007b. A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with *Pichia pastoris*. J Biotechnol. 131 57-66.
58. Kaiser C, Pototzki T, Ellert a., Luttmann R. 2008. Applications of PAT-process analytical technology in recombinant protein processes with *Escherichia coli*. Eng Life Sci. 8:132–138.
59. Katakura Y, Zhang W, Zhuang G, Omasa T, Kishimoto M, Goto Y, Suga K. 1998. Effect of Methanol Concentration on the production of human β2-Glycoprotein I domain V by a recombinant *Pichia pastoris*: a simple system for the control of methanol concentration using a semiconductor gas sensor. J Ferment Bioeng. 86:482–487.
60. Kern A, Hartner FS, Freigassner M, Spielhofer J, Rumpf C, Leitner L, Fröhlich KU, Glieder A. 2007. *Pichia pastoris* 'just in time' alternative respiration. Microbiology. 153, 1250–1260.
61. Khatri NK, Hoffmann F. 2006. Oxygen-limited control of methanol uptake for improved production of a single-chain antibody fragment with recombinant *Pichia pastoris*. Appl Microbiol Biotechnol. 72:492–498.
62. Kim HS, Jeon JH, Lee LJ, Ko K. 2014. N-glycosylation modification of plant-derived virus-like particles: an application in vaccines. Biomed Res Int. 12::485689
63. Kollewe C, Vilcinskas A. 2013. Production of recombinant proteins in insect cells. Am J Biochem Biotechnol. 9:255–271.

64. 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.
65. Lai T, Yang Y, Ng SK. 2013. Advances in Mammalian cell line development technologies for recombinant protein production. Pharmaceuticals. 6:579–603.
66. Lim HK, Choi SJ, Kim KY, Jung KH. 2003. Dissolved-oxygen-stat controlling two variables for methanol induction of rGuamerin in *Pichia pastoris* and its application to repeated fed-batch. Appl Microbiol Biotechnol. 62:342–348.
67. Liu H, Tan X, Russell KA, Veenhuis M, Cregg JM. 1995. *PER3*a, gene required for peroxisome biogenesis in *Pichia pastoris*, encodes a peroxisomal membrane protein involved in protein import. J Biol Chem. 270:10940–10951.
68. Lubenova V, Rocha I, Ferreira EC. 2003. Estimation of multiple biomass growth rates and biomass concentration in a class of bioprocesses. Bioprocess Biosyst Eng. 25:395–406.
69. Lyubenova V, Ignatova M, Salonen K, Kiviharju K, Eerikäinen T. 2011. Control of α-amylase production by *Bacillus subtilis*. Bioprocess Biosyst Eng. 34:367–374.
70. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast. 22:249–70.
71. Manson HS, Lam DM, Arntzen CJ. 1992. Expression of hepatitis B surface antigen in transgenic plants. Proc Natl Acad Sci U S A. 89:11745–11749.
72. Mattanovich D, Branduardi P, Porro L, Gasser B, Sauer M, Porro D. 2012. Recombinant protein production in yeasts. Methods Mol Biol. vol. 824:329-358
73. 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.
74. Menendez J, Valdes I, Cabrera N.2003. The ICL1 gene of Pichia pastoris, transcriptional regulation and use of its promoter. Yeast. 20:1097–1108.
75. Mochizuki S, Hamato N, Hirose M, Miyano K, Ohtani W, Kameyama S, Kuwae S, Tokuyama T, Ohi H. 2001. Expression and characterization of recombinant human antithrombin III in *Pichia pastoris*. Protein Expr Purif. 23:55–65.
76. Nevalainen H, Peterson R. 2014. Making recombinant proteins in filamentous fungi - are we expecting too much? Front Microbiol. 5:10

77. Niu H, Jost L, Pirlot N, Sassi H, Daukandt M, Rodriguez C, Fickers P. 2013. A quantitative study of methanol/sorbitol co-feeding process of a *Pichia pastoris* Mut^+/pAOX1-lacZ strain. Microb Cell Fact. 12:1–8.
78. Nocon J, Steiger MG, Pfeffer M, Sohn SB, Kim TY, Maurer M, Rußmayer H, Pflügl S, Ask M, Haberhauer-Troyer C, Ortmayr K, Hann S, Koellensperger G, Gasser B, Lee SY, Mattanovich D. 2014. Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. Metab Eng. 24:129–138
79. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington J, Silman I, Schrag J, Sussman JL, Verschueren KHG, Goldman A. 1992. The hydrolase fold. Protein Eng 5:197–211.
80. Oliveira R, Clemente JJ, Cunha AE, Carrondo MTJ. 2005. Adaptive dissolved oxygen control through the glycerol feeding in a recombinant *Pichia pastoris* cultivation in conditions of oxygen transfer limitation. J. Biotechnol. 116:35–50
81. País JM, Varas L, Valdés J, Cabello C, Rodríguez L, Mansur M. 2003. Modeling of mini-proinsulin production in *Pichia pastoris* using the AOX promoter. Biotechnol Lett. 25:251–255.
82. Paulová L, Hyka P, Branská B, Melzoch K, Kovar K. 2012. Use of a mixture of glucose and methanol as substrates for the production of recombinant trypsinogen in continuous cultures with *Pichia pastoris* Mut^+. J Biotechnol. 157:180–188.
83. Pfeffer M, Maurer M, Stadlmann J, Grass J, Delic M, Altmann F, Mattanovich D. 2012. Intracellular interactome of secreted antibody Fab fragment in *Pichia pastoris* reveals its routes of secretion and degradation. Appl. Microbiol. Biotechnol. 93:2503–2512.
84. Pla IA, Damasceno LM, Vannelli T, Ritter G, Batt C a, Shuler ML. 2006. Evaluation of Mut^+ and Mut^S *Pichia pastoris* phenotypes for high level extracellular scFv expression under feedback control of the methanol concentration, Biotechnol Prog. 22:881–888.
85. Porro D, Gasser B, Fossati T, Maurer M, Branduardi P, Sauer M, Mattanovich D. 2011. Production of recombinant proteins and metabolites in yeasts: when are these systems better than bacterial production systems? Appl Microbiol Biotechnol. 89:939-948.

86. Potgieter TI, Cukan M, Drummond JE, Houston-Cummings NR, Jiang Y, Li F, Lynaugh H, Mallem M, McKelvey TW, Mitchell T, Nylen A, Rittenhour A, Stadheim T a, Zha D, d'Anjou M. 2009. Production of monoclonal antibodies by glycoengineered *Pichia pastoris*. J Biotechnol. 139:318–325.
87. Potgieter TI, Kersey SD, Mallem MR, Nylen AC, d'Anjou M. 2010. Antibody expression kinetics in glycoengineered *Pichia pastoris*. Biotechnol Bioeng. 106:918–927.
88. 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.
89. Ramón R, Feliu JX, Cos O, Montesinos JL, Berthet FX, Valero F. 2004. Improving the monitoring of methanol concentration during high cell density fermentation of *Pichia pastoris*. Biotechnol Lett. 26:1447–1452.
90. 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.
91. 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.
92. 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.
93. Salah RB, Mosbah H, Fendri A, Gargouri A, Gargouri Y, Mejdoub H. 2006. Biochemical and molecular characterization of a lipase produced *by Rhizopus oryzae*. FEMS Microbiol Lett. 260:241–248.
94. Sayari A, Frikha F, Miled N, Mtibaa H, Ali YB, Verger R, Gargouri Y. 2005. N-terminal peptide of *Rhizopus oryzae* lipase is important for its catalytic properties. FEBS Lett. 579:976–982.
95. Schenk J, Marison IW, von Stockar U. 2007. A simple method to monitor and control methanol feeding of *Pichia pastoris* fermentations using mid-IR spectroscopy. J Biotechnol. 128:344–353.
96. Sears IB, Oconnor J, Rossanese OW, Glick BS. 1998. A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*. Yeast 14: 783–790.

97. Solà A, Jouhten P, Maaheimo H, Sánchez-Ferrando F, Szyperski T, Ferrer P. 2007. Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates. Microbiology. 153:281-290.

98. Spadiut O, Zalai D, Dietzsch C, Herwig C. 2014. Quantitative comparison of dynamic physiological feeding profiles for recombinant protein production with *Pichia pastoris.* Bioprocess Biosyst Eng. 37:1163-1172.

99. Sreekrishna K. 2010. *Pichia*, Optimization of Protein Expression. Ed. Michael C Flickinger. Encycl Ind Biotechnol.190:695–701.

100. Stadlmayr G, Mecklenbräuker A, Rothmüller M, Maurer M, Sauer M, Mattanovich D, Gasser B. 2010. Identification and characterisation of novel *Pichia pastoris* promoters for heterologous protein production. J Biotechnol. 150:519-529.

101. Surribas A, Stahn R, Montesinos JL, Enfors SO, Valero F, Jahic M. 2007. Production of a *Rhizopus oryzae* lipase from *Pichia pastoris* using alternative operational strategies. J Biotechnol. 130:291–299.

102. Takagi S, Tsutsumi N, Terui Y, Kong XY, inventors; Novozymes A/S, Bagsvaerd (DK), assignee. Method for methanol independent induction from methanol inducible promoters in *Pichia*. United States patent US 8,143,023. 2012 Jun 28.

103. Takahashi S, Ueda M, Atomi H, Beer HD, Bornscheuer UT, Schmid RD, Tanaka A. 1998. Extracellular production of active *Rhizopus oryzae* lipase by *Saccharomyces cerevisiae*. J Ferment Bioeng. 86:164–168.

104. Tang S, Reiche A, Potvin G, Zhang Z. 2010. Modeling of phytase production by cultivation of *Pichia pastoris* under the control of the *GAP* promoter. Int J Chem React Eng. 8:1–9.

105. Valero F. 2013. Bioprocess engineering of *Pichia pastoris*, an exciting host eukaryotic cell Expression system. Ed. Tomohisa Ogawa. Protein Eng - Technol Appl .3–32.

106. Veloso ACA, Rocha I, Ferreira EC. 2009. Monitoring of fed-batch *E. coli* fermentations with software sensors. Bioprocess Biosyst Eng. 32:381–388.

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

108. Wang JR, Li YY, Xu SD, Li P, Liu JS, Liu DN. 2013. High-level expression of pro-form lipase from *Rhizopus oryzae* in *Pichia pastoris* and its purification and characterization. Int J Mol Sci. 15:203-217.

109. Ward OP. 2012. Production of recombinant proteins by filamentous fungi. Biotechnol Adv. 30:1119–1139.

110. 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.

111. Yang J, Zhou X, Zhang Y. 2004. Improvement of recombinant hirudin production by controlling $NH4^+$ concentration in *Pichia pastoris* fermentation. Biotechnol Lett. 26:1013–1017.

112. Ye J, Ly J, Watts K, Hsu A, Walker A, McLaughlin K, Berdichevsky M, Prinz B, Sean Kersey D, d'Anjou M, Pollard D, Potgieter T. 2011. Optimization of a glycoengineered *Pichia pastoris* cultivation process for commercial antibody production. Biotechnol Prog. 27:1744–1750.

113. Yokohama S. 2003. Protein expression systems for structural genomics and proteomics. Curr Opin Chem Biol. 7:39-43.

114. Yurimoto H, Oku M, Sakai Y. 2011. Yeast methylotrophy: metabolism, gene regulation and peroxisome homeostasis. Int J Microbiol. 101298:1-8.

115. Zalai D, Dietzsch C, Herwig C, Spadiut O. 2012. A dynamic fed batch strategy for a *Pichia pastoris* mixed feed system to increase process understanding. Biotechnol Prog. 28: 878-886.

116. Zhang W, Bevins M a, Plantz B a, Smith L a, Meagher MM. 2000. Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. Biotechnol Bioeng. 70:1–8.

117. Zhang W, Smith LA, Plantz BA, Schlegel VL, Meagher MM. 2002. Design of methanol feed control in *Pichia pastoris* fermentations based upon a growth model. Biotechnol Prog. 18:1392–1399.

118. Zhang W, Sinha J, Smith LA, Inan M, Meagher MM. 2005. Maximization of production of secreted recombinant proteins in *Pichia pastoris* fed-batch fermentation. Biotechnol Prog. 21:386–393.

119. Zhang W, Inan M, Meagher MM. 2007. Rational design and optimization of fed-batch and continuous fermentations. Methods Mol Biol. 389:43–64.

120. Zhao W, Wang J, Deng R, Wang X. 2008. Scale-up fermentation of recombinant Candida rugosa lipase expressed in *Pichia pastoris* using the *GAP* promoter. J Ind Microbiol Biotechnol. 35:189–195.

121. Zhou X, Zhang Y. 2002. Decrease of proteolytic degradation of recombinant hirudin produced by *Pichia pastoris* by controlling the specific growth rate. Biotechnol Lett. 24:1449–1453.

Chapter 2:

State and specific growth rate estimation in heterologous protein production by *Pichia pastoris*.

Published as: Barrigón JM., Ramon R., Rocha I., Valero F., Ferreira EC., Montesinos JL. 2012. State and specific growth estimation in heterologous protein production by *Pichia pastoris*. Aiche J. 58:2966-2979.

---

2. State and specific growth rate estimation in heterologous protein production by *Pichia pastoris.*

2.0. Abstract

Estimation of biomass, substrate and specific growth rate (μ) by two non-linear observers (non-linear observer-based estimator - *NLOBE*, asymptotic observer with second-order dynamics tuning - *AO-SODE)* and a linear estimator (Recursive least-squares with variable forgetting factor - *RLS-VFF*) is presented. Heterologous protein production in *Pichia pastoris* P_{AOX1} (Mut^+) and P_{FLD1}-based systems is closely related to μ and has been addressed due to its high relevance in modern biotechnology and bioprocess engineering. μ was estimated by on-line gas analyses or substrate measurements, biomass and substrate considering yield coefficients and mass balances.

In simulation studies *NLOBE* showed high sensitivity to tuning and initialization variables. Validation experiments demonstrated *AO-SODE* performs better than the *RLS-VFF* for moderate to rapid changes of μ and model parameters being known. If low changes on μ are presented, for instance in substrate regulation, *RLS-VFF* comes up as the best option, because of its reduced requirements.

2.1. Introduction

In order to achieve high productivity, constant product quality, and also to allow optimization and control of biotechnological processes, real-time monitoring of the key fermentation variables (biomass, substrate, and products) is of major academic and industrial relevance. Some reviews and specific papers have addressed this topic, covering the wide range of techniques that have been applied (Junker and Wang, 2006; Surribas et al., 2006a, Becker et al., 2007; Amigo et al., 2008). Biomass is usually the central variable of the mathematical models used to describe microbial growth, where it is included as a state variable. Several analytical methods have been adapted to monitor cell density evolution during bioprocesses. However, a standard method for the on-line determination of biomass is not currently available, as each available technique has its own advantages and disadvantages (Kiviharju et al., 2008; Madrid and Felice, 2005).

The methylotrophic yeast *Pichia pastoris* has been widely reported as a suitable expression system for both basic research and industrial application (Macauley-Patrick et al., 2005). In recent years, more than 500 proteins have been expressed using this system (Cregg, 2014). The strong and tightly regulated alcohol oxidase 1 promoter (P_{AOX1}) is the most widely used *P. pastoris* promoter for recombinant expression, being induced by methanol (Lin Cerghino et al., 2001; Cregg et al., 1998). In order to avoid methanol utilization, *FLD1* promoter has been recently considered (Shen et al., 1998). The gene encodes a formaldehyde dehydrogenase (*FLD*), a key enzyme required for the catabolism of methanol and certain primary amines, such as methylamine used as nitrogen source in methylotrophic yeast (Harder and Veenhuis, 1989). With this promoter, methanol is

replaced by sorbitol as carbon source and methylamine is used as sole nitrogen source and inducer of protein production.

The specific growth rate is a critical parameter in the optimization of the heterologous protein production by *P. pastoris*. There is an optimum growth rate for optimal product formation, which is protein specific. Nevertheless, besides being difficult to estimate on-line, this parameter may vary during the fermentation process, since this is usually operated in batch and fed-batch modes. Moreover, there are variations in substrates and operational conditions applied within the different phases of the production process (Cos et al., 2006).

On-line estimation of the specific growth rate(s) is usually performed together with the determination of state variables such as biomass, since the direct estimation of the specific growth-rate is not possible. The determination of these variables can be achieved by the development of estimation algorithms or software sensors that require the existence of accurate mathematical models of the process (van Impe and Claes, 1999; James et al., 2000; Dochain, 2003; Sunström and Enfors, 2008; Kadlec et al., 2009; Jensch et al., 2006). Nonetheless, the development of a suitable mathematical description and its identification is difficult due to the complex interactions exhibited by the microorganism, as well as to the operating conditions and the state of the system.

In order to eliminate errors in estimation owing to inaccurate models, it is possible to design adaptive algorithms that estimate simultaneously the state and some parameters of the process, considered as time-varying parameters.

This type of estimation can be achieved by an adaptive extended Kalman filter (Valero et al., 1990; Arndt et al., 2005), a recursive prediction error method (Montesinos et al., 1995), an adaptive extended Luenberger observer (Loeblein and Perkins, 1999) and also by non-linear observers (Nadri et al., 2006), which are based on the accurate knowledge of both model structure and parameters. However, these observers present a problem of convergence over a wide range of operational conditions. In addition, guessing suitable initial values for different state variables and parameters is rather critical to obtain precise estimates.

Adaptive observers somewhat easier to implement have been designed (Bastin and Dochain, 1990). They estimate the state and the kinetics of the process, considered as time-varying parameters, through partial measurements of the state.

Observed-based estimators (*OBE*), whose fundamentals are based on the non-linear systems theory, have also been proposed. In this approach, reaction rates estimation is carried out from the measurement of state variables (Lubenova et al., 2003; Farza et al., 1999). The observer gain depends on measured state variables and also constraints, which are added for the calibration of the gain when the measurements used are relatively noisy. Furthermore, specific non-linear observer-based estimators (*NLOBE*) have been

developed (Farza et al., 1999). The main characteristic of those estimators lie in the ease with which they may be implemented and, specifically calibrated. Their tuning is reduced to the calibration of a simple tuning parameter.

Because of the complexity to tune *OBE´s*, some authors studied their stability, dynamics of convergence and suitable values for tuning parameters (Oliveira et al., 2002). Recently, an asymptotic observer (*AO*) for the estimation of state variables has been designed and its performance compared with that of a classical observer (Extended Kalman Observer, *EKO*) (Veloso et al,. 2009). The specific growth rates were obtained using an estimator based on the reformulation proposed by Dochain and Bastin, (1990; Pomerleau and Perrier, 1992). The *AO* allows reconstructing the missing state variables even if the process is not exponentially observable and the kinetics is unknown. Another advantage of this kind of observers is that there are no tuning parameters. However, it is assumed that the yield coefficients are known.

Additionally, estimation of parameters by the recursive least-squares (*RLS*) method has been proposed, especially for the specific growth rate and for the process state, thus achieving unbiased convergence and making possible the estimation of changing process parameters (Estler, 1995; Golobic et al., 1999). The main advantage of the *RLS* methods is that it does not require an accurate knowledge of the system because a linear model that only depends on the on-line measures is considered.

The modulating-functions method has also been used for the on-line identification of a microbial growth model (Ungarala and Co, 1998). Nevertheless, due to the large number of estimated coefficients and parameters, the method is of difficult application to experimental conditions.

In this chapter, the estimation of biomass, substrate and specific growth rate is presented and applied to the heterologous production of *Rhizopus oryzae* lipase (ROL) by *P. pastoris* in batch and fed-batch operational modes. Different algorithms and procedures are studied and discussed, always considering their applicability in terms of overall performance, taking into account aspects which often are not considered from an experimental point of view.

Two of the estimators studied belong to the *OBE* class (*NLOBE*, *AO*) and the other is based on the *RLS* method, all three selected owing to their main advantages described above. The aim of this work has been to compare performance of the different algorithms in *P. pastoris* bioprocesses. On-line exhaust gas analyses or substrate concentration measurements were used to estimate the specific growth rate. With the aim of simplicity and to reduce instabilities and the number of tuning parameters, biomass and substrate were straightforwardly obtained solving their corresponding mass balances.

This second chapter is based on Barrigón et al., (2102). The development and the implementation of the algorithm *NLOBE* estimator and RLS method have been obtained from Ramon, 2007. On the other hand, the *AO* estimator has been developed for the first

time to be appliedin the estimation of state and specific growth rate in the ROL heterologous production by *Pichia pastori*s. The chapter is structured as follows: firstly, process description and estimation algorithms are presented. Secondly, simulations for different alternatives are carried out with the *P. pastoris* P_{AOX1}-based system, Mut^+ phenotype, discussing the most suitable ones using single and global performance metrics. Finally, experimental validation for both *P. pastoris AOX1* (Mut^+) and *FLD* promoters is shown.

2.2. Materials and methods

2.2.1. Strains

The wild type *P. pastoris* X-33 strain (Mut^+), containing the vector pPICZαAROL was used for heterologous expression of ROL under the control of the *AOX1* promoter (Cos et al., 2005a). *P. pastoris* X-33 containing the vector pPICZFLDαROL was chosen for the expression of ROL controlled by the *FLD1* promoter (Cos et al., 2005a).

2.2.2. Cultivation set-up and operational conditions

Batch and fed-batch processes for the heterologous production by *P. pastoris* were studied. Cells were cultured in a 5 liter Braun Biostat ED fermenter (Braun Biotech, Melsungen, Germany). Fermentation conditions were: stirring rate 800 rpm; temperature 30 ºC; pH controlled at 5.5 adding NH_4OH 30 % (v/v) (batch) or 5 M KOH (fed-batch) for P_{AOX1} and 5 M KOH for P_{FLD1} (batch and fed-batch); dissolved oxygen controlled above 30 % with an air flow rate of 1.5-20 $L \cdot min^{-1}$.

For the *P. pastoris* P_{AOX1}-based systems, the cultivations were carried out in two phases: a first batch growth phase on glycerol (3.5 L volume with an initial glycerol concentration of 40 $g \cdot L^{-1}$) followed by a second phase (fed-batch), where first a mixture of glycerol plus methanol was fed to the culture and after that, only methanol was used as the sole carbon source and production of the recombinant protein took place (Cos et al., 2005a). Cultivations using *P. pastoris* P_{FLD1}-based systems proceeded as follows: after a batch phase on glycerol, a pulse of sorbitol and methylamine was added into the bioreactor. For the induction phase a 300 $g \cdot L^{-1}$ of sorbitol, 30 $g \cdot L^{-1}$ of methylamine chloride and 7.5 mL

of trace salts solution feeding medium was used in an exponential feeding rate strategy (Cos et al., 2005a).

Glycerol, methanol, sorbitol and methylamine were added in the transition and production phases by an automatic microburette MicroBU-2031 from Crison Instruments (Alella, Barcelona, Spain). On-line substrate measurements are possible by the implementation of automated manifolds as previously reported (Ramon et al., 2004; Horstkotte et al., 2008). Off-gas analyses were performed by infrared and paramagnetic detectors (Multor, Maihak, Hamburg, Germany). Carbon dioxide production rate (*CPR*) and oxygen uptake rate (*OUR*) were considered approximately equal to carbon dioxide transfer rate (*CTR*) and oxygen transfer rate (*OTR*) respectively, and thus, obtained through mass balances of CO_2 and O_2 measured in the exhaust gas.

For the estimation procedure, the integration step and the sampling time were fixed at 0.055 h (200 s). On-line gas measurements are a result of 20 averaged raw data points. So, the filtered signal was used as the input for the estimator.

A more detailed description of materials and methods can be found elsewhere (Cos et al., 2005a).

2.3. Theory and calculation

2.3.1. The process model

For the oxidative assimilation of substrate (glycerol, methanol or sorbitol), equations for growth (1); protein production (2) and maintenance (3) can be stated as:

$$k_1\,\mathrm{S} + k_2\,\mathrm{O}_2 \xrightarrow{r_X} X + k_3\,\mathrm{CO}_2 \qquad (1)$$

$$k_4\,\mathrm{S} + k_5\,\mathrm{O}_2 \xrightarrow{r_P} P + k_6\,\mathrm{CO}_2 \qquad (2)$$

$$\mathrm{S} + k_7\,\mathrm{O}_2 \xrightarrow{r_m} k_8\,\mathrm{CO}_2 \qquad (3)$$

where X, S, O_2, CO_2, and P represent biomass, substrate, dissolved oxygen, dissolved carbon dioxide and product respectively (in the sequel, the same symbols are used to represent component concentrations); r_x, r_p, r_m are the reaction rates; k_i are the yield (stoichiometric) coefficients.

The corresponding dynamic model can be represented as follows:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ O_2 \\ CO_2 \\ P \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -k_1 & -k_4 & -1 \\ -k_2 & -k_5 & -k_7 \\ k_3 & k_6 & k_8 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mu \\ q_p \\ q_m \end{bmatrix} \cdot X - D\begin{bmatrix} X \\ S \\ O_2 \\ CO_2 \\ P \end{bmatrix} + \begin{bmatrix} 0 \\ D \cdot S_0 \\ OTR \\ -CTR \\ 0 \end{bmatrix} \qquad (4)$$

where S_o is the substrate concentration in the feed, *OTR* is the oxygen transfer rate from gas to liquid phase and *CTR* is the carbon dioxide transfer rate from liquid to gas phase. μ, q_m and q_p are biomass, product and maintenance specific reaction rates respectively. $D = F_{in}/V$ is the dilution rate; F_{in} is the inlet volumetric flow rate, and V is the reactor volume. In case of the volume variation, the mass balance should be taken into account (see chapter 7). However, in this chapter it is approximated to $dV/dt = F_{in}$.

Neglecting product formation in mass balance equations and considering a global reaction scheme for growth and maintenance, then, equations (1-4) are transformed into equations (5-6), where overall yield coefficients ($k^*_i = Y^*_{c/x}$) are used.

$$k^*_1\,\mathrm{S} + k^*_2\,\mathrm{O_2} \xrightarrow{r_X} X + k^*_3\,\mathrm{CO_2} \qquad (5)$$

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ O_2 \\ CO_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -k_1^* \\ -k_2^* \\ k_3^* \end{bmatrix} \cdot \mu \cdot X - D\begin{bmatrix} X \\ S \\ O_2 \\ CO_2 \end{bmatrix} + \begin{bmatrix} 0 \\ D \cdot S_0 \\ OTR \\ -CTR \end{bmatrix} \qquad (6)$$

Kinetic parameters and yield coefficients used for *P. pastoris* process under P_{AOX1} (Mut^+) and P_{FLD1} for *ROL* production were obtained from previous experiments in batch and fed-batch cultivations (Cos et al., 2005a,b,c). The yield and maintenance coefficients in the kinetic models were identified as previously reported (Petkov and Davis, 1996; Surribas et al., 2006b) but adapted to fed-batch processes. Process model parameters and coefficients are listed in Table 1.

Table 1. Parameters and coefficients used for *P. pastoris* process under P_{AOX1} (Mut^+) and P_{FLD1} for *ROL* production.

	P_{AOX1} (Mut^+)		P_{FLD1}
Parameters and coefficients	**Batch**	**Fed-batch**	**Fed-batch**
Substrate	Glycerol	Methanol	Sorbitol
μ_{max} (h^{-1})	0.260	0.059	0.030
K_s ($g \cdot L^{-1}$)	0.20	0.22	---
$Y_{CO2/X}$ ($mol_{CO2} \cdot g_X^{-1}$)	$1.57 \cdot 10^{-2}$	$1.02 \cdot 10^{-1}$	$1.32 \cdot 10^{-2}$
$m_{CO2/X}$ ($mol_{CO2} \cdot g_X^{-1} \cdot h^{-1}$)	$2.6 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$	$6.5 \cdot 10^{-4}$
$Y_{O2/X}$ ($mol_{O2} \cdot g_X^{-1}$)	$2.4 \cdot 10^{-2}$	$1.69 \cdot 10^{-1}$	$2.80 \cdot 10^{-2}$
$m_{O2/X}$ ($mol_{O2} \cdot g_X^{-1} \cdot h^{-1}$)	$3.1 \cdot 10^{-4}$	$4.7 \cdot 10^{-4}$	$7.1 \cdot 10^{-4}$
$Y_{S/X}$ ($g_S \cdot g_X^{-1}$)	1.97	4.29	0.47
$m_{S/X}$ ($g_S \cdot g_X^{-1} \cdot h^{-1}$)	0.008	0.010	0.043
K_La (h^{-1})	---	360	---
$O_{2,sat}$ ($mol_{O2} \cdot L^{-1}$)	---	$6.6 \cdot 10^{-4}$	---

2.3.2. Estimation algorithms

The main objective of this work is to implement different estimators for the determination of state variables and specific growth rate using, if it is possible, single and easily available measurements from a bioprocess system. Equally, estimation algorithms can use either gas transfer rates or substrate concentration as on-line measurements to firstly estimate the specific growth rate without requiring a kinetic model for growth, after that, state variables are calculated. The estimators considered are two non-linear observers (*NLOBE*, *AO*) and another one based on the recursive least squares (*RLS*) method.

They are appropriate to systems where component kinetic rates (*CKR*) can be considered related to biomass concentration (X) and the specific growth rate (μ) through a component to biomass intrinsic yield ($Y_{c/x}$) and a maintenance coefficient ($m_{c/x}$) described by the so-called *Luedeking-Piret* relationship as follows:

$$CKR = Y_{c/x}\mu X + m_{c/x}X \approx Y^{*}_{c/x}\mu X \qquad (7)$$

So, the proposed methodology can be implemented where either the carbon dioxide production rate (*CPR*), the oxygen uptake rate (*OUR*), or the substrate uptake rate (*SUR*) can be determined on-line.

2.3.2.1. Non-linear observer based estimator (NLOBE)

The first estimator tested is based on a general design for a *non-linear observer* (Farza et al., 1999) represented by:

$$\frac{d\hat{z}}{dt} = F(s,\hat{z})\hat{z} + G(u,s,\hat{z}) - \Lambda^{-1}(s,\hat{z})S_{\theta}^{-1}C^{T}(C\hat{z} - y) \quad (8)$$

where $\hat{z}$ is the estimated state, and (*u, s*) and *y* are, respectively, the input and the output of the system.

The used estimator considers that only z_1 is available as on-line measurement. It is correlated to biomass and the specific growth rate by the *Luedeking-Piret* type model mentioned above. For instance, considering *CPR* as the component kinetic rate, equations (9) and (10) are taken into account.

$$CPR(t) = Y_{CO_2/X}\mu(t)X(t) + m_{CO_2/X}X(t) \quad (9)$$

$$z_1 = \text{CPR} \quad (10)$$

Thus, the structure of the estimator is developed as follows (Farza et al., 2000),

$$\frac{d\hat{z}_1}{dt} = \hat{z}_2\hat{z}_1 - D\hat{z}_1 - 3\delta(\hat{z}_1 - z_1) \quad (11)$$

$$\frac{d\hat{z}_2}{dt} = \hat{z}_3 - 3\frac{\delta^2}{z_1}(\hat{z}_1 - z_1) \quad (12)$$

$$\frac{d\hat{z}_3}{dt} = -\frac{\delta^3}{z_1}(\hat{z}_1 - z_1) \quad (13)$$

Therefore, the μ estimates can be obtained from equation (14):

$$\frac{d\hat{\mu}}{dt} = (\hat{z}_2 - \hat{\mu})\left(\hat{\mu} + \frac{m_{CO_2/X}}{Y_{CO_2/X}}\right) \quad (14)$$

In this approach, *a priori*, only a single tuning parameter δ is needed to calibrate the estimation procedure. However, suitable initial values for $\hat{z}_{10}$, $\hat{z}_{20}$ and $\hat{z}_{30}$ have to be provided in order to obtain satisfactory estimates.

2.3.2.2. Asymptotic observer (AO) with second-order dynamics tuning (SODE)

The second estimator tested was an *asymptotic observer* (*AO*) which is used for the estimation of state variables. The specific growth rate is obtained using an estimator based on the reformulation proposed by Dochain and Bastin, (1990) and Pomerleau and Perrier, (1992).

A general dynamic model for stirred tank bioreactors can be described following mass balance equations written in a matrix form as:

$$\frac{d\xi}{dt} = KH(\xi)\rho(\xi) - D\xi + F - Q(\xi) \qquad (15)$$

where ξ is the state vector (the set of *n* component concentrations), *K* is a (*n x m*) matrix of known yield coefficients, *D* the dilution rate, *F* the feed rate vector with *dim(F)=n* and *Q* the gaseous outflow rate vector which *dim(Q)=n*. *D*, *F* and *Q* are measured on-line. The reaction rates are defined as $\varphi(\xi) = H(\xi)\rho(\xi)$ to take advantage of any possible knowledge of the kinetic model, $H(\xi)$ being a (*m x r*) matrix of known functions of the state and $\rho(\xi)$ a vector of *r* unknown functions of state. The reaction rates can be estimated using the following the *AO* general structure (Oliveira et al., 2002):

$$\frac{d\hat{\xi}}{dt} = KH(\xi)\hat{\rho} - D\xi + F - Q - \Omega\left(\xi - \hat{\xi}\right) \qquad (16)$$

$$\frac{d\hat{\rho}}{dt} = \left[KH(\xi)\right]^T \Gamma\left(\xi - \hat{\xi}\right) \qquad (17)$$

As previously reported (Oliveira et al., 2002) the stability considerations for the related linear time varying perturbing system are developed for *AO* with *SODE*. The first step is the selection of a subset of *r* equations of the full state space model to use the kinetics estimator. For the specific growth rate estimation from the global reaction defined in equation (5), it is necessary to introduce a *"c"* one-dimensional sub-space related with the measured variable (*S*, O_2 or CO_2), its feeding rate or component transfer rate and its respective yield coefficient k^*_c.

Furthermore, the *SODE* algorithm usually uses a transformed state variable, defined as

$$\Psi = K_C^{-1} \cdot \xi_C = \frac{1}{\pm k_C^*} C \qquad (18)$$

Combining equations (16) and (17) with equation (18) the estimator is obtained:

$$\frac{d\hat{\psi}}{dt} = H\hat{\rho} - D\psi + K_C^{-1}(F_C - Q_C) - \Omega_C\left(\psi - \hat{\psi}\right) \qquad (19)$$

$$\frac{d\hat{\rho}}{dt} = H^T \Gamma_C\left(\psi - \hat{\psi}\right) \qquad (20)$$

Where the gain matrices (Ω_C, Γ_C) and their parameters are defined as follows:

$$\Gamma_C = diag\{-\gamma_i\} \qquad \Omega_C = diag\{-\omega_i\} \qquad (21)$$

$$\gamma = \frac{2 \cdot \zeta}{\tau} - \frac{1}{\hat{X}}\frac{d\hat{X}}{dt} \qquad \omega = \frac{1}{\hat{X} \cdot \tau^2} \qquad (22)$$

The estimation of biomass (*X*) is obtained from one measurement *C* (substrate, oxygen or carbon dioxide concentration and their respective feeding or gaseous outflow rates),

according to the choice of a sub-space *c*. Thus, the system model from equation (6) is simplified, according with the expression (23):

$$\frac{d}{dt}\begin{bmatrix} X \\ C \end{bmatrix} = \begin{bmatrix} 1 \\ \pm k^*_C \end{bmatrix} \cdot \mu \cdot X - D\begin{bmatrix} X \\ C \end{bmatrix} \pm \begin{bmatrix} 0 \\ F_C - Q_C \end{bmatrix} \qquad (23)$$

Finally, biomass estimation is calculated by an auxiliary variable (*Z*) that depends on *C* concentration, *CKR* measurements or the feeding rate, the overall yield coefficient and biomass[25].

$$\hat{Z} = \hat{X} + \frac{1}{\pm k^*_C} C \qquad (24)$$

$$\frac{d\hat{Z}}{dt} = -D\hat{Z} + \frac{1}{\pm k^*_C}\left(F_C - Q_C\right) \qquad (25)$$

$$\hat{X} = \hat{Z} - \frac{1}{\pm k^*_C} C \qquad (26)$$

As previously described (Oliveira et al., 2002; Veloso et al., 2009), when *D*, the dilution rate, is kept close to zero for long periods of time, the performance of the *AO* is expected to be not satisfactory, since the rate of convergence of the estimation fully depends on the values of that variable. Therefore, the *AO* estimation procedure was not applied in the batch phase and will be tested and validated only in fed-batch operation.

Furthermore, although the kinetics of the process may be considered unknown in these observers, the estimation of the state variables requires an accurate knowledge about the reaction scheme and stoichiometric coefficients. Consequently, uncertainties on these model parameters, besides noisy on-line measurements can generate a large bias in the

estimation procedure. For that reason, it is advisable, as implemented by some authors (Van Impe and Claes, 1999; Estler, 1995), to use filtered data with setting bounds on μ ($0.05\mu_{max} \leq \mu \leq \mu_{max}$) in order to avoid unrealistic estimated values for the specific growth rate and $Y^{*}_{c/x}$ in equation (7).

2.3.2.3. Recursive least-squares (RLS) with variable forgetting factor (VFF)

Estimators based on the r*ecursive least squares* (*RLS*) method consider a linear model for the system within the time interval where the identification procedure is performed. Some authors have shown the capacity of linear estimators to adequately estimate the fermentation process while not increasing the model structure complexity (Roux et al., 1966; Carrier and Stephanopoulos, 1998). In this way, some approaches have been developed to carry out the estimation from either measurement of oxygen (Estler, 1995) or carbon dioxide in the exhaust gas of the fermentor (Golobic et al., 1999). It has been applied to a type of processes where a sole substrate for growth and induction was added in a non-continuous mode. From equation (7), considering the *CPR* and its time-derivative, equation (27) can be obtained.

$$\frac{dCPR(t)}{dt} = \frac{d\mu(t)}{dt} X(t) Y_{CO2/X} + \frac{dX(t)}{dt} Y_{CO2/X}\, \mu(t) + m_{CO2/X} \frac{dX(t)}{dt} \quad (27)$$

Then, considering the mass balance for biomass

$$\frac{dCPR(t)}{dt} = \frac{d\mu(t)}{dt} X(t) Y_{CO2/X} + \big(\mu(t) - D(t)\big) \cdot CPR(t) \qquad (28)$$

where $\mu(t)$ is an unknown time-varying parameter, *CPR(t)* is an indirect on-line measurement, and biomass *X(t)* cannot be measured on-line.

Considering that μ varies slowly within the sampling interval, the term corresponding to the μ first-time derivative in equation (28) can be neglected. Discretization of the resulting equation is conducted using a first-order central Euler approximation,

$$\frac{CPR(t_{k+1}) - CPR(t_{k-1})}{2\Delta t} = \big(\mu(t_k) - D(t_k)\big) \cdot CPR(t_k) \quad (29)$$

Thereupon, a time varying parameter $\hat{\theta}_{(t_k)}$ can be defined, which will be recursively estimated with the *RLS* method.

$$CPR(t_{k+1}) = \hat{\theta} CPR(t_k) + CPR(t_{k-1}) \quad (30)$$

From the previous estimate, the specific growth rate is obtained:

$$\mu(t_k) = \frac{\hat{\theta}(t_k)}{2\Delta t} + D(t_k) \quad (31)$$

The proposed estimator uses the *RLS* method to estimate $\hat{\theta}_{(t_k)}$ according to the set of equations presented below:

$$\hat{\theta}(t_k) = \hat{\theta}(t_{k-1}) + K(t_k)\left(y(t_{k-1}) - \hat{y}(t_{k-1})\right) \quad (32)$$

$$\hat{y}(t_k) = \Psi^T(t_k)\,\hat{\theta}(t_{k-1}) \quad (33)$$

$$K(t_k) = \Psi^T(t_k)\,\hat{\theta}(t_{k-1}) \quad (34)$$

$$Q(t_k) = \frac{P(t_{k-1})}{\lambda + \Psi^T(t_k)\,P(t_{k-1})\Psi(t_k)} \quad (35)$$

$$Q(t_k) = \frac{1}{\lambda} + \left(P(t_{k-1}) - \frac{P(t_{k-1})\,\Psi(t_k)\Psi^T(t_k)\,P(t_{k-1})}{\lambda + \Psi^T(t_k)\,P(t_{k-1})\Psi(t_k)}\right) \quad (36)$$

where *P(t)* is the covariance matrix, *Q(t)* an auxiliary matrix, *y(t)* is the *CPR(t)* and Ψ the data vector.

In the heterologous protein production by *P. pastoris*, sometimes there are important changes in the system characteristics, such as substrates, concentrations and operating conditions, which may lead to less satisfactory estimation results. The nonlinear dynamics of the specific growth rate can be included in the estimator, considering a variable forgetting factor $\lambda(t)$ to conform the proposed estimator. The time-varying forgetting factor (*VFF*) prevents the constant reduction in the value of the covariance matrix during the dynamic process (Golobic et al., 1999). The calculation of the varying forgetting factor is based on model error and both the data vector and covariance matrix.

$$\lambda(t_k) = 1 - \frac{\left(y(t_{k-1}) - \hat{y}(t_{k-1})\right)^2}{\Sigma_0} + \left(1 - \frac{P(t_{k-1})\,\Psi(t_k)\Psi^T(t_k)\,P(t_{k-1})}{\lambda + \Psi^T(t_k)\,P(t_{k-1})\Psi(t_k)}\right) \quad (37)$$

Σ_0 is the only tuning parameter and it needs to be determined empirically, because it depends on process dynamics, sampling time and measurement noise.

Finally, it is necessary to impose biological restrictions in μ estimation in order to be always positive and bounded ($\leq \mu_{max}$).

2.3.3. Performance indicators

Single metrics *SSE* (sums of squared error), *RMSE* (root mean squared error), *MRE* (mean relative error), RMNS (root mean noise sensitivity), *ITAE* (integral time-weighted absolute error) and *RT* (rise time) were used for the evaluation of the "goodness-of-estimation" in the simulation phase. *SSE*, *RMSE* and *MRE* were calculated for μ, *X* and *S* estimations. *RMNS*, *ITAE* and *RT* were obtained for μ estimates.

The sum of squared error (*SSE*) can be defined as follows:

$$SEE = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \qquad (38)$$

where *n* is the number of data points, $\hat{y}_i$ is the i^{th} estimated value, y_i the corresponding i^{th} actual value from process model.

The root mean squared error (*RMSE*) is defined by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \qquad (39)$$

that is the average difference between estimated and actual target variables (specific growth rate, biomass and substrate), where $\hat{y}_i$ and y_i are the estimated and true value for the i^{th} data point; *n* is the total number of data points. Alternatively, calculation of the mean relative error (*MRE*) is also carried out to examine the "goodness-of-estimation".

$$MRE = \frac{1}{n} \sum_{i=1}^{n} \frac{|(\hat{y}_i - y_i)|}{y_i} \qquad (40)$$

The sum of squared noise sensitivity (*SSNS*) accounts for the continuous variation of μ estimations and it is defined as follows:

$$SSNS = \sum_{i=1}^{n-1} (\mu_{i+1} - \mu_i)^2 \qquad (41)$$

Root mean square noise sensitivity (*RMNS*) is then calculated as the average differences between consecutive estimated specific growth rate values μ_{i+1} and $\hat{\mu}_i$ from i= 1 to *n-1* data points.

$$RMNS = \sqrt{\frac{\sum_{i=1}^{n-1}(\mu_{i+1} - \mu_i)^2}{n-1}} \qquad (42)$$

In addition, overall performance indicators: root mean overall performance index (*RMOPI*) and combined mean relative error (*CMRE*), which consider some former single metrics, were used.

$$RMOPI = \sqrt{w_\mu RMSE_\mu{}^2 + w_x RMSE_x{}^2 + w_s RMSE_s{}^2 + w_{ns} RMNS^2} \qquad (43)$$

w_k weighting factors of the single metrics, selected to bring the terms within the same order of magnitude but with lower contribution for errors on *S*. $w_\mu = 0.40 \cdot 10^5$, $w_x = 0.19$, $w_s = 0.01$, $w_{ns} = 0.40 \cdot 10^5$.

$$CMRE = w_{c\mu} MRE_\mu + w_{cx} MRE_x \qquad (44)$$

w_k weighting factors of the mean relative error for μ and *X*, $w_{c\mu} = 0.50$, $w_{cx} = 0.50$.

Integral time-weighted absolute error (ITAE) weights errors which exist after long time much more heavily than those at the start of the estimation. Its evaluation is of interest in batch and fed-batch processes where the need to have rapidly reliable estimates is a crucial issue[15], because it can be taken as an indicator for speed of convergence.

$$ITAE = \int t\, e\, dt \qquad (45)$$

with *e* the absolute error between μ estimations and actual values $|(\bar{\mu} - \mu)|$, *t* time of the process. *ITAE* is computed over the simulated period using the trapezium rule to approximate the integral.

Moreover, rise time (*RT*) is defined at the time taken for the estimator response to first reach 95% of the change on specific growth rate. The lower *RT* is, the faster response the estimation system has.

Another point that has to be noted is the different effect which tuning parameters and estimator initialization values can be produced on the estimator's global performance. So, sensitivity of *RMOPI* and *CMRE* functions were examined through variations of their nominal values.

2.4. Results and discussion

2.4.1. Simulation results

The simulated system accounts for the cultivation of the Mut^+ phenotype of *P. pastoris* under P_{AOX1} for the *ROL* production. In order to study the performance of the different estimators, a process model was developed considering mass balances (X, S, O_2, and CO_2) and component kinetic rate expression for carbon dioxide, oxygen, and substrate (*CPR*, *OUR*, and *SUR)* as stated in equations (6,7).

NLOBE, two sorts of *AO-SODE* and *RLS-VFF* estimation algorithms were studied. Firstly, *NLOBE* and *RLS-VFF* estimation were carried out from CPR measures. Secondly, the estimation was made using substrate concentration (methanol) for the first of the *AO-SODE's*. Thirdly, the *AO-SODE* estimation performance, for the second *AO-SODE* estimator, was obtained from oxygen balance, in which the main variables are the dissolved oxygen concentration and the oxygen transfer rate *OTR.* Thus, it is necessary to include growth kinetics (μ(S)) and also the oxygen transfer rate model (OTR) for the *AO-SODE* simulations. The specific growth rate is approximated by the *Monod* equation. The oxygen transfer rate is computed as $OTR = K_La\ (O_{2,sat} - O_2)$ where K_La is the global mass transfer coefficient and $O_{2,sat}$ the saturation concentration of dissolved oxygen (Pérez et al., 2006).

For the simulation, and after a batch phase on glycerol, an induction phase in fed-batch mode was started by an exponential addition of methanol as substrate:

$$F_{in}(t) = \frac{\mu_{set}\left[X_{(0)}\, V_{(0)}\right]}{Y^{*}_{X/S}\; S_0} \exp\left[\mu_{set}\,[t - t_{(0)}]\right] \qquad (46)$$

Thus, the substrate feeding rate depends on the specific growth rate set-point, μ_{set} = 0.02 h^{-1}; the initial volume, $V_{(0)}$; the initial biomass concentration, $X_{(0)}$; the substrate concentration in the feed, S_0; and the biomass to substrate overall yield, $Y^{*}_{X/S} = Y^{*-1}_{S/X}$.

The *CTR* (*CPR)* and *OTR* (*OUR)* simulated data were corrupted by an additive Gaussian noise of mean zero and relative standard deviation of 2%, considered as a typical value after data filtering. Noise applied to substrate measurements was 6%.

Initial conditions for the simulations are given in Table 2. Initial estimates are taken as the same as the system model and are representative of the real process: ($\hat{X}_{(0)}$, $\hat{\mu}_{(0)}$, $\hat{S}_{(0)} = X_{(0)}, \mu_{(0)}, S_{(0)}$).

Table 2. Simulation conditions in batch and fed-batch *P. pastoris* process under *PAOX1* (Mut^{+}) for *ROL* production.

Process variables	Units	Batch	Fed-batch
Substrate (*S*)	[g·L^{-1}]	Glycerol	Methanol
Initial concentration ($S_{(0)}$)	[g·L^{-1}]	40	0
Inlet concentration (S_0)	[g·L^{-1}]	---	790
Process time (*t*)	[h]	14.7	60.3
Initial volume ($V_{(0)}$)	[L]	3.5	3.5
Initial Biomass ($X_{(0)}$)	[g·L^{-1}]	0.50	20
Specific growth rate set-point (μ_{set})	[h^{-1}]	---	0.02
Initial specific growth rate ($\mu_{(0)}$)	[h^{-1}]	0.0	0.0
Sampling time (*T*)	[h]	0.055	0.055

In general, tuning parameters and some initialization variables were determined empirically considering the convergence speed and noise sensitivity (Van Impe and Claes, 1999), which are tightly related to the deviation between estimations and real data. Weighted average of mean square deviations can be used as part of the objective function.

In this sense, the *RMOPI* indicator described before was used as cost criterion to be minimized. Depending on the sensitivity of this *RMOPI* on parameters and initialization values, nominal values were rounded to one or two significant digits. In some cases, restrictions related to stability and convergence properties were applied to get preliminary values. On the other hand, these values can sometimes be obtained straightforward attending to the particular structure of the estimator. Details related to each estimator are specified separately.

2.4.1.1. Tuning parameters and initialization variable

In the *NLOBE* estimator, initial guesses for δ, $\hat{z}_{10}$, $\hat{z}_{20}$ and $\hat{z}_{30}$ were selected according to the structure of the estimator: $\delta \approx 1/T$; $\hat{z}_{10} \approx CPR_{max}$; $\hat{z}_{20} < ((\mu_{max}/T)/\mu)$ for the maximal value of $d\mu/dt$; $\hat{z}_{30} = 10 \cdot \hat{z}_{20}$. Initial guesses and nominal values, obtained by *RMOPI* minimization, are listed in Table 3. Reinitialization of the estimator parameters is needed when the operational mode is switched from batch to fed-batch, when protein production is induced. Because of different system dynamics during batch and fed-batch

culture, the estimator is not able to overcome the change on feeding rate *(D>0)* if a similar estimation quality is required.

Two parameters must be tuned in the *AO-SODE* algorithm: ζ, the damping coefficient and τ, the time constant of a second order dynamic response of estimated to the true values. The damping coefficient, ζ, was preliminary fixed at 1 according to a common engineering rule of thumb (Bastin and Dochain, 1990). Tuning rules for discrete time implementation (Oliveira et al., 2002) were applied to determine preliminary value of τ according to $0< T < 2\zeta\tau$, thus setting τ to 0.05. Nominal values were finally obtained by minimization of the cost criterion function *RMOPI* and they are shown in Table 3 for the two *AO-SODE* methods tested.

The tuning of the *RLS-VFF* estimator starts with the assignation of first trials to the tuning parameter Σ_0 and the initial value θ_0.

Σ_0 has to be set in order to manage slow or moderate variations of μ by the variable forgetting factor. So, Σ_0 initial settings were fixed to be the same order of magnitude that square deviations in equation (37). For the initial value θ_0 , according to the inner structure of the *RLS* estimator a preliminary value $\theta_0 \approx 2T\mu$ can be assigned.

Initial and nominal values, obtained by *RMOPI* minimization, for Σ_0 and θ_0 can be found in Table 3. In contrast with *NLOBE* estimator, retuning of the estimator parameters was not necessary when the operational mode was changed. The *RLS-VFF* method was able

to cope with the different dynamics of the operational modes, batch and fed-batch process.

Initial value for the covariance matrix P_0 was set at 1, because simulations with high and low P_0 showed that it converges quickly, and weakly influence, to the estimation.

Table 3. Initial guesses, nominal values and acceptable ranges for tuning parameters and initialization estimator variables.

Algorithm	Parameters & estimator variables	Units	Initial guesses	Nominal value	Acceptable range
NLOBE	δ	$[h^{-1}]$	20	51	±1
	$\hat{z}_{10}$	$[mol \cdot L^{-1} \cdot h^{-1}]$	0.5	4.0	±0.3
	$\hat{z}_{20}$	$[h^{-1}]$	5.0	2.0	±0.1
	$\hat{z}_{30}$	$[h^{-2}]$	50	100	±3
AO-SODE (Methanol)	ζ	[-]	1.0	0.70	±0.40
	τ	$[h^{-1}]$	0.05	0.20	±0.11
AO-SODE	ζ	[-]	1.0	0.70	±0.40
	τ	$[h^{-1}]$	0.05	0.20	±0.11
RLS-VFF	Σ_0	$[h^{-2}]$	$5.0 \cdot 10^{-8}$	$1.0 \cdot 10^{-8}$	$\pm 0.9 \cdot 10^{-8}$
	θ_0	[-]	$2.0 \cdot 10^{-3}$	$1.5 \cdot 10^{-2}$	$\pm 0.5 \cdot 10^{-2}$

2.4.1.2. Estimators performance

The estimation results for *NLOBE,* both *AO-SODE* and *RLS-VFF* methods are illustrated in Figure 1, Figure 2 and Figure 3, respectively. The performance of the estimators at their nominal values for the tuning parameters and initialization variables are presented in Table 4. The goodness of estimation for *NLOBE* and *RLS-VFF* methods is nearly as satisfactory as for *AO-SODE*, which shows the higher values for the single metrics (*SSE*,

RMSE and *MRE*) for μ values. However, all of μ and *X* estimations were conducted properly with deviations lower than 1% (*MRE*) and with similar values of the overall performance indexes (*RMOPI*, *CMRE*). At this point, it is remarkable to emphasize that substrate estimations have a large mean relative error (*MRE*>1), due to substrate low values, which is confirmed by its *RMSE*. Terms in substrate balance corresponding to inlet and uptake rate are quite similar and so, computed substrate concentration is very sensitive to the estimation of μ.

The speed of convergence, indirectly measured by *ITAE* and *RT*, is also satisfactory for all estimators. Fast enough response to the change on the substrate feeding strategy is considered for rise time appeared to be shorter than 1 h, being the longest time for both *AO-SODE* estimators. Nonetheless, *ITAE* values are slightly better for *AO-SODE* estimators. In addition, all estimators showed that dynamics of convergence are time-varying and becoming faster when the bioprocess approaches the end.

Finally, the noise sensitivity of the estimators was quantified by the *RMNS* index and displayed in Table 4. The *NLOBE* estimator is provided for almost as sensitive as others and as a result the higher time-varying dynamics. However, *NLOBE* estimator gives the lowest estimation error *RMSE*. In contrast *AO-SODE* methods present the highest *RMSE* values with low *RMNS*. These facts are a consequence of the tuning procedure applied to obtain nominal values, which are a compromise between estimation error and noise sensitivity evaluated through the *RMOPI* index.

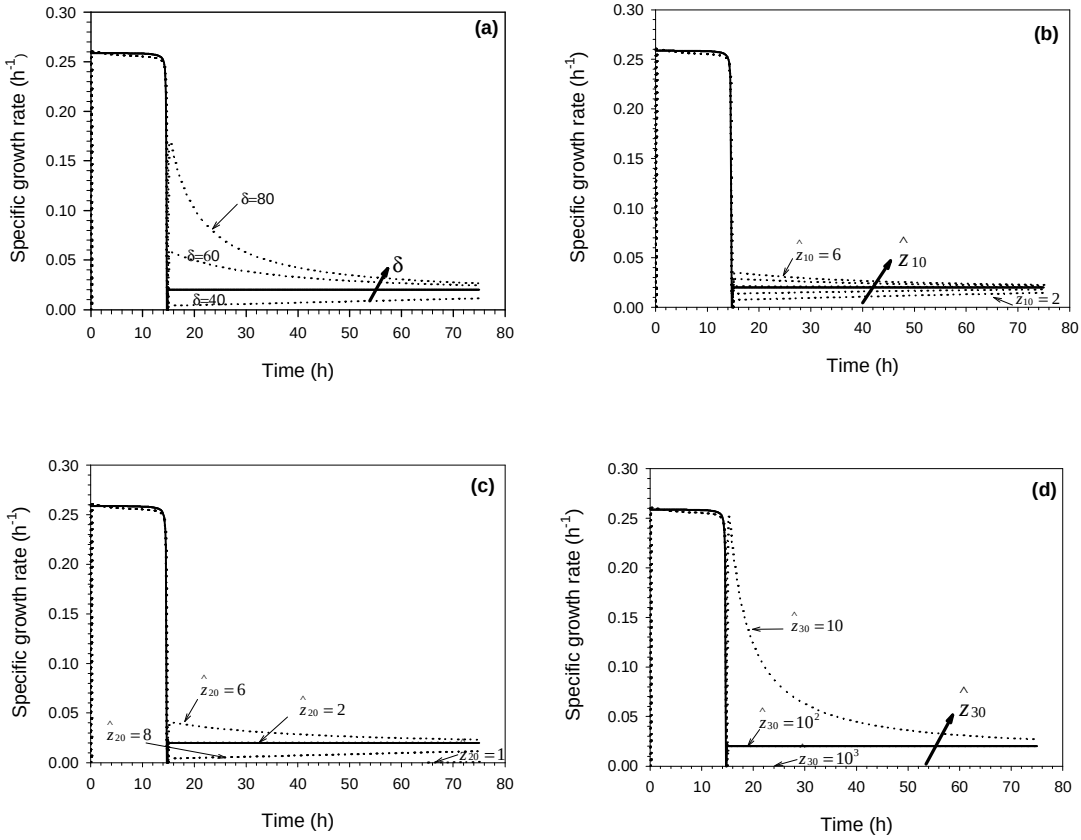

Figure 1. *NLOBE* estimator. Effect of tuning and initial parameters on μ estimation: **(a)** δ, **(b)** $\hat{z}_{10}$ **(c)** $\hat{z}_{20}$, **(d)** $\hat{z}_{30}$. S*olid line* stands for model data and *dotted lines* were stood for estimation from *CPR* simulated data.

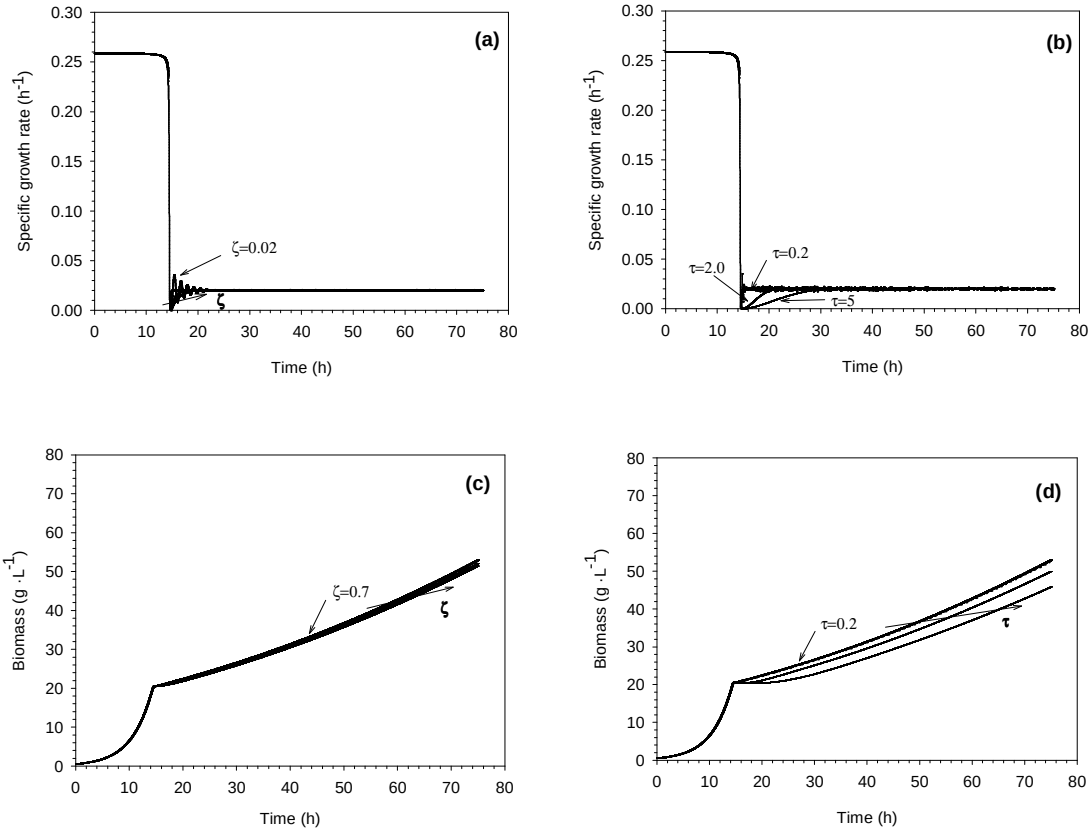

Figure 2. *AO-SODE* estimator. Effect of tuning parameters on μ estimation from substrate measurements: **(a)** ζ, **(b)** τ; and on reconstructed biomass: **(c)** ζ, **(d)** τ..

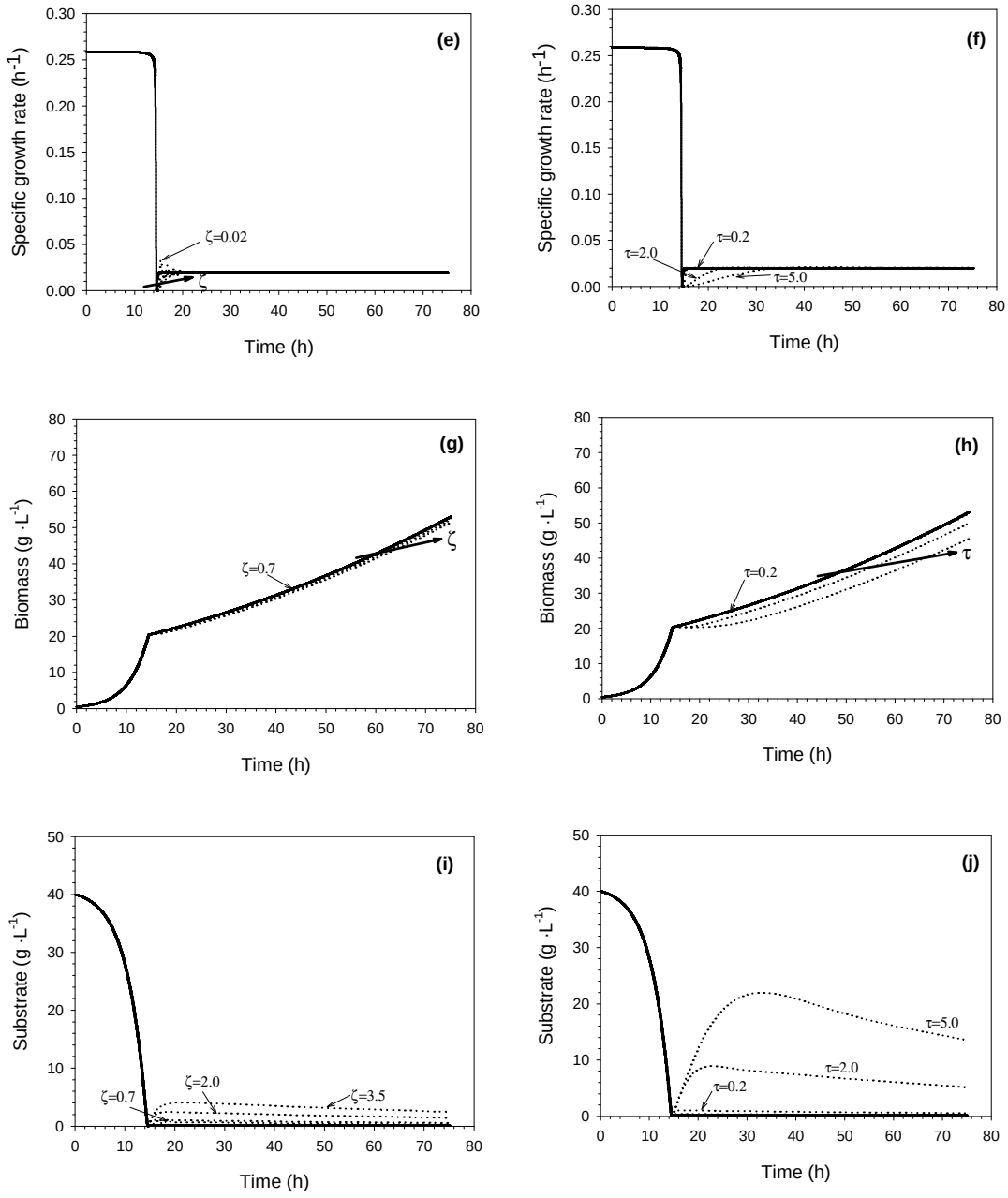

Figure 2. *AO-SODE* estimator. Effect of tuning parameters on μ estimation from pO_2 and *OTR*: **(e)** ζ, **(f)** τ; on reconstructed biomass: **(g)** ζ, **(h)** τ; and on reconstructed substrate: **(i)** ζ, **(j)** τ. *Solid line* stands for model data and *dotted lines* were stood for estimation from simulated data.

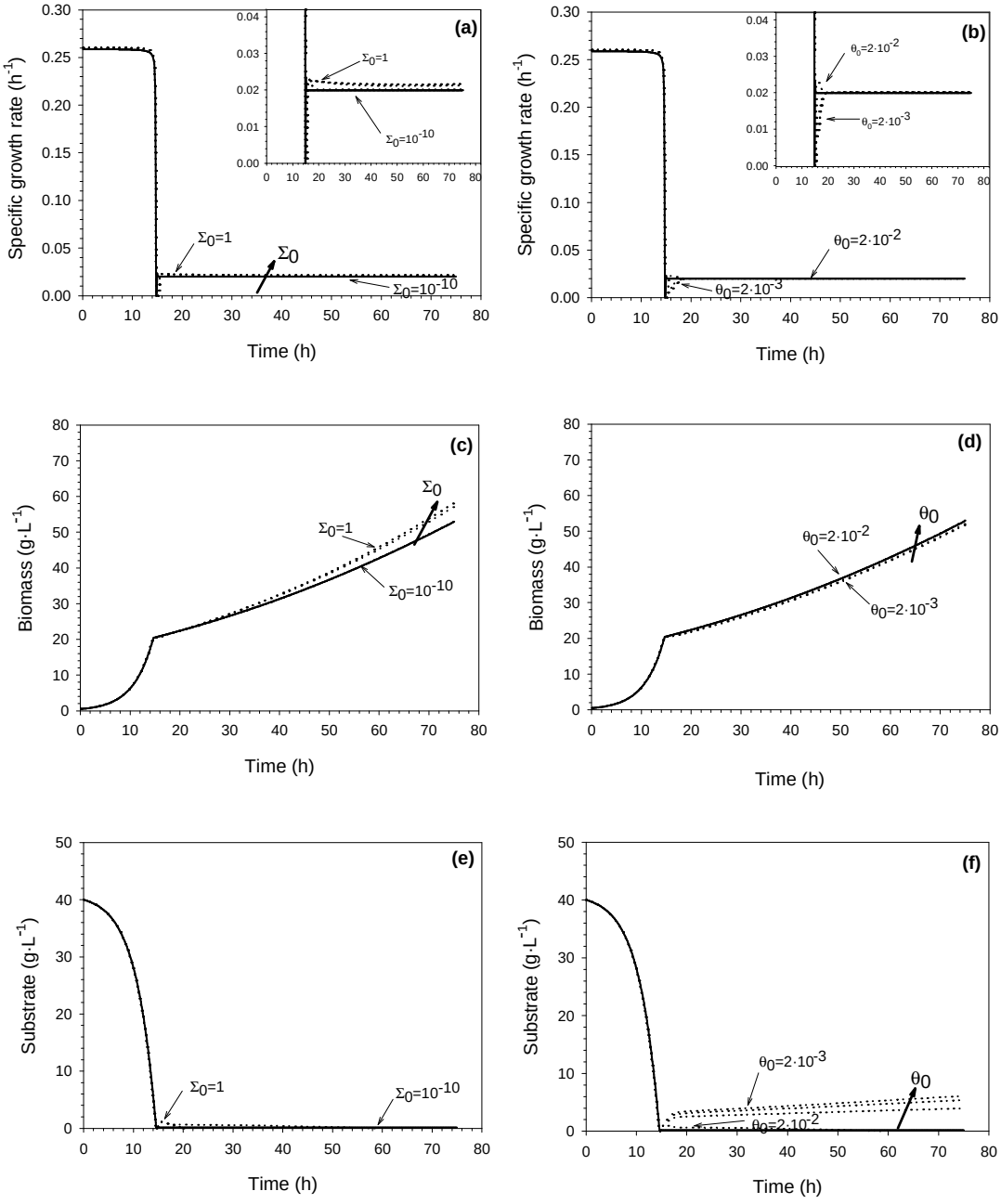

Figure 3. *RLS-VFF* estimator. Effect of tuning parameters on μ estimation: **(a)** Σ_0, **(b)** θ_0; on reconstructed biomass: **(c)** Σ_0, **(d)** θ_0; and on reconstructed substrate: **(e)** Σ_0, **(f)** θ_0. S*olid line* stands for model data and *dotted lines* were stood for estimation from *CPR* simulated data.

2.4.1.3. Sensitivity analysis

In order to make a thorough sensitivity analysis, firstly, *NLOBE's* simulations were carried out to study the effect of the tuning parameter δ and the initial value of some variables on the estimation results. The effect of δ, , $\hat{z}_{20}$ and $\hat{z}_{30}$ on the estimation of the specific growth rate from *CPR* measurements is visualized in Figure 1. Regarding the effects of the tuning parameter, high values of δ provoke a typical overshoot at the start-up of the fed-batch phase. Moreover, a high variation on overall performance is observed for low to moderate δ variations as displayed in Figure 1 and Figure 4. In spite of having only one tuning parameter, the extremely sensitivity of the *RMOPI* index to δ is the main drawback of the *NLOBE*. Equally, the results obtained for the initialization variables were presented in Figures 2.1 and 2.4. Results presented in Figure 4 shows more clearly the largest sensitive response. *RMOPI* is less sensitive to and $\hat{z}_{20}$ than $\hat{z}_{30}$, but the last one is no harder sensitive than δ. These results were somewhat expected because z_3 was defined as the z_2's time derivative which is supposed to be unknown and bounded. The dynamics of z_3 introduces some type of integral action which eliminates any static error when estimating z_2 (Farza et al., 1999).

Secondly, the effect of *AO-SODE* tuning parameters ζ and τ on the dynamics of convergence can be assessed from the plots in Figure 2. In Figure 2 (a-b) is illustrated that specific growth rate values are in agreement with typical second-order dynamic responses. It is shown that decreasing τ, the response becomes faster and decreasing ζ produces more oscillatory responses. With regard to stability conditions, the range allowed for the integration step T, set equal to the sampling time, is bounded and

conditioned by the choice of ζ and τ. These facts allow searching and selecting the tuning parameters with intuitive basis, simplifying the search for optimum (Oliveira et al., 1999).

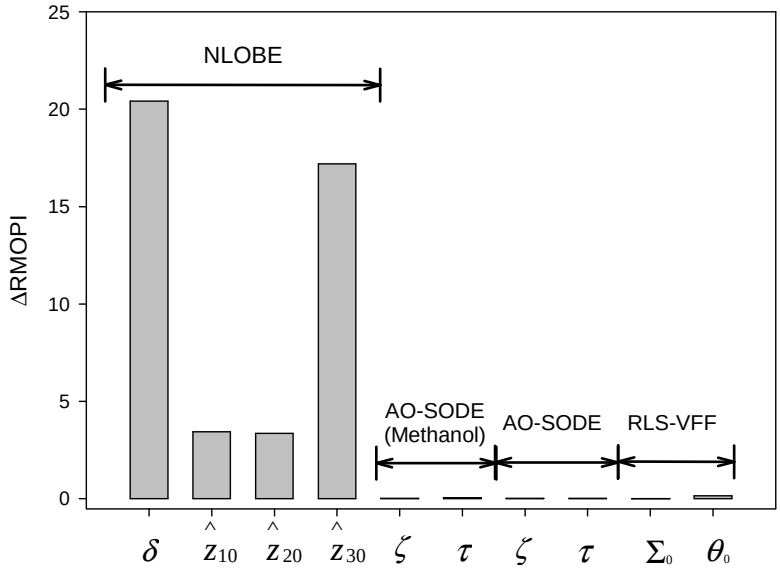

Figure 4. Absolute variation in the root mean overall performance index (*RMOPI*) of the estimation procedure upon variation in the estimator parameters by 20%.

As it is shown in Figure 2, biomass (X) estimations are quite sensitive to deviations of the specific growth rate (μ). When the estimation started, long time was taken to reach the specific growth rate for τ values rather far from nominal values. This initial X offset remained quite stable through the bioprocess, whenever μ estimation was well fitted to the simulated system response. Finally, in order to calculate the sensitivity to the tuning parameters ζ and τ, the *RMOPI* combined metrics was explored. In contrast to Figure 2, it can be observed in Figure 4 that the effect of ζ is not different than those of τ. This fact can be explained due to the tuning parameter τ in Figure 2 was tested in a wider range than ζ, with the aim to easily appreciate the overall estimation behavior.

Thirdly, in the second case of *AO-SODE*, Figure 2(e-j), based on dissolved O_2 and *OUR* data, the dynamics of convergence for tuning parameters ζ and τ were very similar to those illustrated for the methanol case. No significant differences on X and μ estimations were shown for all *AO-SODE* estimators tested (Methanol, O_2 and *OUR*, CO_2 and *CPR).* Nevertheless, *AO-SODE* estimators using gas measures also allowed computing substrate concentration. This estimation was performed applying the substrate balance, using X and μ estimation, as indicated in equations (18-26). Thereupon, it can be observed in Figure 2 (i-j), substrate estimations are more sensitive to deviations of the specific growth rate and biomass concentration than in their ζ and τ nominal values, as commented in the previous section. In addition, for the cases where the specific growth rate is estimated using gas measures (O_2 or CO_2 and *OTR* or *CTR*), the effect of the dissolved concentration in the *AO-SODE* algorithm is negligible compared with the gaseous transfer rates.

Finally, the study of *RLS-VFF* was focused on the effect of the initial value θ_0 and the tuning parameter Σ_0. Their effect on μ, X and S estimation, using *CPR* measurements, is presented in Figure 3 and Figure 4. As it can be observed, the performance of the estimation algorithm is more sensitive to θ_0 values than for Σ_0. In Figure 3(c-d), the tuning parameter Σ_0 for biomass calculations seems to have a stronger effect than initial value of θ_0 due to the wide range explored for Σ_0. From Figure 3, if a low Σ_0 value is introduced in the estimation of μ, it gives robustness to the overall procedure, providing a very satisfactory state estimation. These results clearly highlight the practical implementation of the proposed algorithm. With sub-optimal θ_0 values a good convergence speed of the estimator is not attained, as can be seen in Figure 3 at the

beginning of the fed-batch process, and consequently large deviations of substrate estimations are produced.

2.4.1.4. Selection of the estimator

Estimations showed similar good agreement with their actual values for the specific growth rate, biomass and substrate concentrations for all estimators tested. This goodness of estimation has been evaluated by overall performance indexes *RMOPI* and *CMRE* presented in Table 4. However, significant differences were observed concerning the range of application for tuning parameters and initialization variables, speed of convergence, number of tuning parameters needed and requirements for system model.

The acceptable ranges of tuning parameters were shown in Table 3. These ranges were calculated according to the sensitivity analysis on *CMRE* indicator to obtain variations always lower than 5%, except in the cases of ζ, τ and Σ_0 with 1.5% maximal *CMRE* variation. This is justified due to stability criteria used for *AO-SODE* which imposes setting bounds on tuning parameters (ζ, τ) and the far low sensitivity of *CMRE* to ζ, τ and Σ_0.

From results given in Table 3 and Figure 4, it can be concluded that tuning parameters and initial values for the *NLOBE* have stronger influence on the overall estimator performance than the *AO* and *RLS* methods. Furthermore, this influence is not so much different between *AO* and *RLS* methods, although the *RLS* showed more dependence on the θ_0 initial value.

Table 4. Single and combined performance metrics for *NLOBE*, *AO-SODE* and *RLS-VFF* estimators. μ [h^{-1}]; X [$g \cdot L^{-1}$]; S [$g \cdot L^{-1}$].

Algorithm	*Variable*	*SSE*	*RMSE*	*MRE*	*RMNS*	*RT*	*ITAE*	*RMOPI*	*CMRE*
NLOBE	μ	$1.2 \cdot 10^{-4}$	$3.4 \cdot 10^{-4}$	$5.0 \cdot 10^{-3}$	$6.5 \cdot 10^{-4}$	0.27	0.11	0.17	0.004
	X	24	0.15	$3.1 \cdot 10^{-3}$					
	S	370	0.59	>1					
AO-SODE (Methanol)	μ	$1.3 \cdot 10^{-3}$	$1.1 \cdot 10^{-3}$	$7.2 \cdot 10^{-3}$	$2.2 \cdot 10^{-4}$	0.50	0.06	0.24	0.006
	X	46	0.20	$5.4 \cdot 10^{-3}$					
	S	---	---	---					
AO-SODE	μ	$1.8 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$	$1.0 \cdot 10^{-2}$	$1.8 \cdot 10^{-4}$	0.72	0.12	0.27	0.006
	X	1.6	$3.8 \cdot 10^{-2}$	$1.2 \cdot 10^{-3}$					
	S	510	0.68	>1					
RLS-VFF	μ	$2.3 \cdot 10^{-4}$	$4.6 \cdot 10^{-4}$	$9.1 \cdot 10^{-3}$	$3.5 \cdot 10^{-5}$	0.10	0.18	0.11	0.006
	X	7.3	0.08	$2.3 \cdot 10^{-3}$					
	S	170	0.40	>1					

Speed of convergence indicated in Table 4 by the rise time (*RT*) was similar to all three estimators studied (< 1h), but presenting a rather sluggish initial response for the *AO-SODE*. It is not clearly reflected in the *ITAE* value because it weights initial errors less heavily than those which persist on time.

The number of critical initialization variables is greater in the *NLOBE* estimator, having some of them high influence on the *RMOPI* index (Figure 4). It has to be pointed out the effect of δ, , $\hat{z}_{20}$ and $\hat{z}_{30}$ which a variation of 2%, 8%, 5% and 3% respectively provides an 5% increase on *CMRE* as can be straightforwardly obtained from Table 3.

Observer based estimators (*NLOBE, AO*) are not recommended in a preliminary choice with poorly known kinetic parameters because of their fundamentals relies on an accurate knowledge of yields and maintenance equations. However, their stability, dynamics of convergence and parameter tuning can be improved using a variable gain-structure like

AO-SODE estimators presented. With this kind of variable gain observers, moderate to rapid changes in the specific growth rate can be anticipated properly. They allow higher identification frequency and reveal good adaptive behavior to process changes with high speed of convergence. As its main disadvantage, its application in the batch mode is not recommended due to expect poor performance and weak stability.

Making use of *RLS-VFF* methods allows decreasing the knowledge requirements about the system, not only in quantitative terms, but also in a qualitative or schematic system description. Among other advantages, it can be pointed out the use of linear equations instead of differential equations. Finally, a reduced number of tuning parameters gives an added value to this kind of estimators. The adaptive computation of the forgetting factor in the *RLS* method makes possible that moderate changes in the system, such as in substrates, state variable concentrations and operational conditions, can be processed to conduct a satisfactory tracking of the specific growth rate. This approach allows that, when the operational mode is switched from batch to fed-batch, reinitialization of estimator variables is not necessary.

On the whole, taking into account simulation studies *RLS-VFF* methods and *AO-SODE* estimators are the most promising ones. The main advantage for both is their tuning simplicity comparing to the *NLOBE* tested. Additionally, *RLS-VFF* does not require kinetic coefficients and *AO-SODE* presents good stability within a wide range of tuning parameter values. Thereupon, in the next section experimental validation is carried out for both estimators.

2.4.2. Experimental validation

The *RLS-VFF* method was applied to experimental data from both *P. pastoris* P_{AOX1} and P_{FLD1}-based systems. Experimental data for the specific growth rate were obtained from off-line biomass measurements using suitable smoothing spline functions (Cos et al., 2005b).

Firstly, the cultivation of the Mut^+ phenotype of *P. pastoris* under P_{AOX1} for the ROL production was studied. After batch cultivation on glycerol and transition phase, an induction phase in fed-batch mode was started by a pre-programmed exponential feeding of methanol.

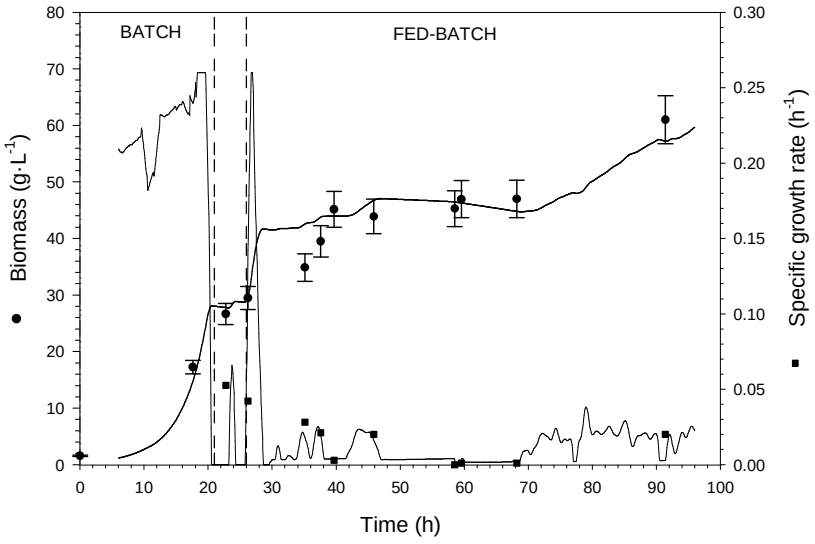

Figure 5. Specific growth rate and biomass estimation using *CPR* experimental data for the *P. pastoris* P_{AOX1}-based system (Mut^+) obtained with the *RLS-VFF* method. *Symbols* correspond to real values and *lines* stand for estimation.

As it can be seen in Figure 5, μ estimation and computed biomass estimation were carried out properly to the system response and a rapid convergence was obtained for the μ estimation during the batch phase. When the batch phase was completed, as found at 20 h, it was shown also unbiased μ and X estimations to substrate changes. This fast adjustment is again confirmed, as in previous section, during the fed-batch mode. Firstly, growth slowdown is observed, the growth almost was practically stopped mainly due to the regulation of the microorganism to the new conditions. After that, an increase in the growth rate is obtained. The estimator performance allows adapting to growth variations and, then, estimating the computed biomass successfully.

With the aim to validate the estimation algorithms in other systems with different promoters and process dynamics, experimental data from *P. pastoris* P_{FLD1}-based system was processed. In P_{FLD1}-system, a comparable experimental strategy than for the P_{AOX1}-system was used, but now utilizing sorbitol and methylamine instead of methanol and ammonium hydroxide as carbon and nitrogen sources, respectively. Variable substrate feeding was carried out, with a non-automatic substrate control procedure, attempting to keep the substrate concentration at 8 $g \cdot L^{-1}$.

With the goal to make a comparison among estimation methods, single metrics: *RMSE* and *MRE* were used to evaluate the goodness-of-estimation for the state variables, such as biomass and substrate, when they can be measured directly but they are not used as on-line measurement. A summary is presented in Table 5.

Table 5. Errors of the estimation algorithms for estimated state variables in the *P. pastoris PFLD1*-based system during the fed-batch mode.

Algorithm	Estimated variable	*RMSE*	*MRE*
RLS-VFF (CPR)	Biomass [$g \cdot L^{-1}$]	1.2	0.036
	Substrate [$g \cdot L^{-1}$]	1.3	0.16
AO-SODE (CO_2,CPR)	Biomass [$g \cdot L^{-1}$]	1.0	0.026
	Substrate [$g \cdot L^{-1}$]	1.5	0.19
AO-SODE (O_2,OUR)	Biomass [$g \cdot L^{-1}$]	0.9	0.025
	Substrate [$g \cdot L^{-1}$]	1.4	0.18
AO-SODE (Sorbitol)	Biomass [$g \cdot L^{-1}$]	2.0	0.051

Results for the *RLS-VFF* algorithm are shown in Figure 6 and Table 5. Once the specific growth rate was properly estimated, biomass and substrate can be calculated through the use of their corresponding mass balances. Biomass and substrate were estimated successfully with an estimation error $\leq 5\%$, although slight deviations were detected for substrate. From Figure 6, it was observed that the estimated substrate concentration and specific consumption rates are in good agreement with those obtained through substrate balancing with off-line specific growth rate data.

The estimation procedure showed suitable capacity of adjustment during the bioprocess, especially when switching from batch phase to fed-batch mode with different substrate types and changing concentration. Interestingly, the quality of the estimation was non-permanently affected by operational problems caused by aeration and foaming shift-up at 68 h.

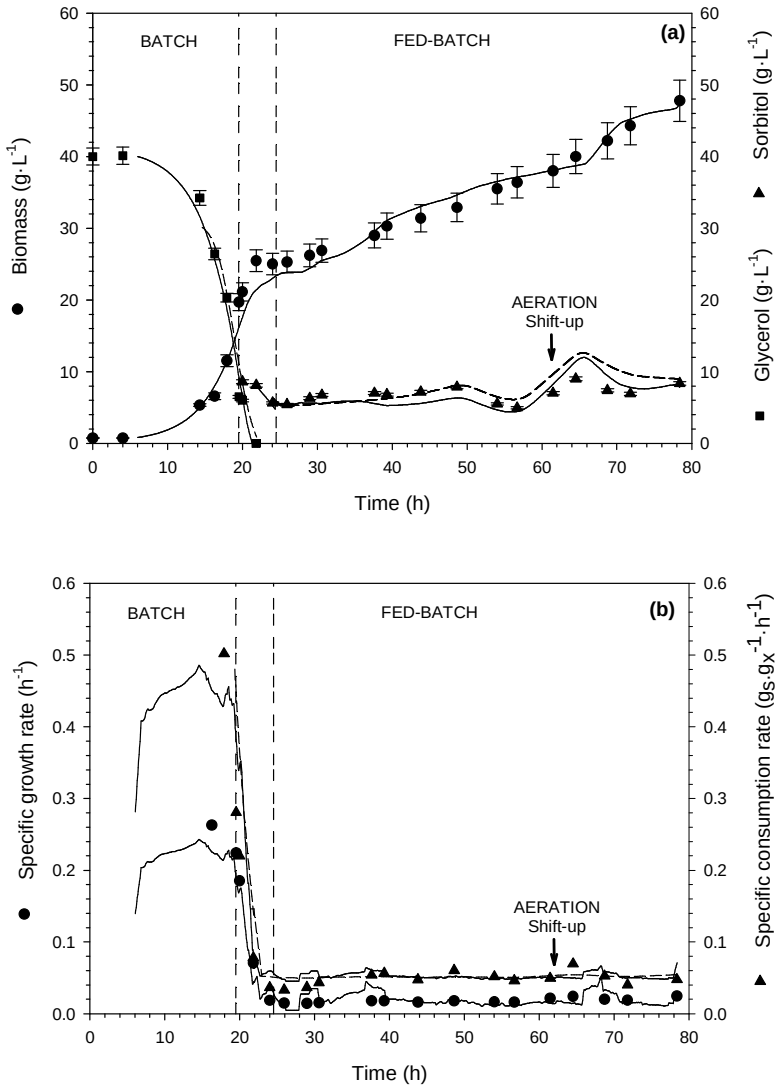

Figure 6. *RLS-VFF* validation from *CPR* experimental data for the *P. pastoris PFLD1*-based system. **(a)** Biomass and substrate estimation; **(b)** Specific rates estimation. *Symbols* correspond to process values, *solid lines* stand for estimation and *dashed lines* obtained through substrate balance with off-line specific growth rate data.

Furthermore, the performance of the *AO-SODE* observer was also tested in the *P. pastoris* P_{FLD1}-based system for validation and, finally, compared with the *RLS-VFF* estimator.

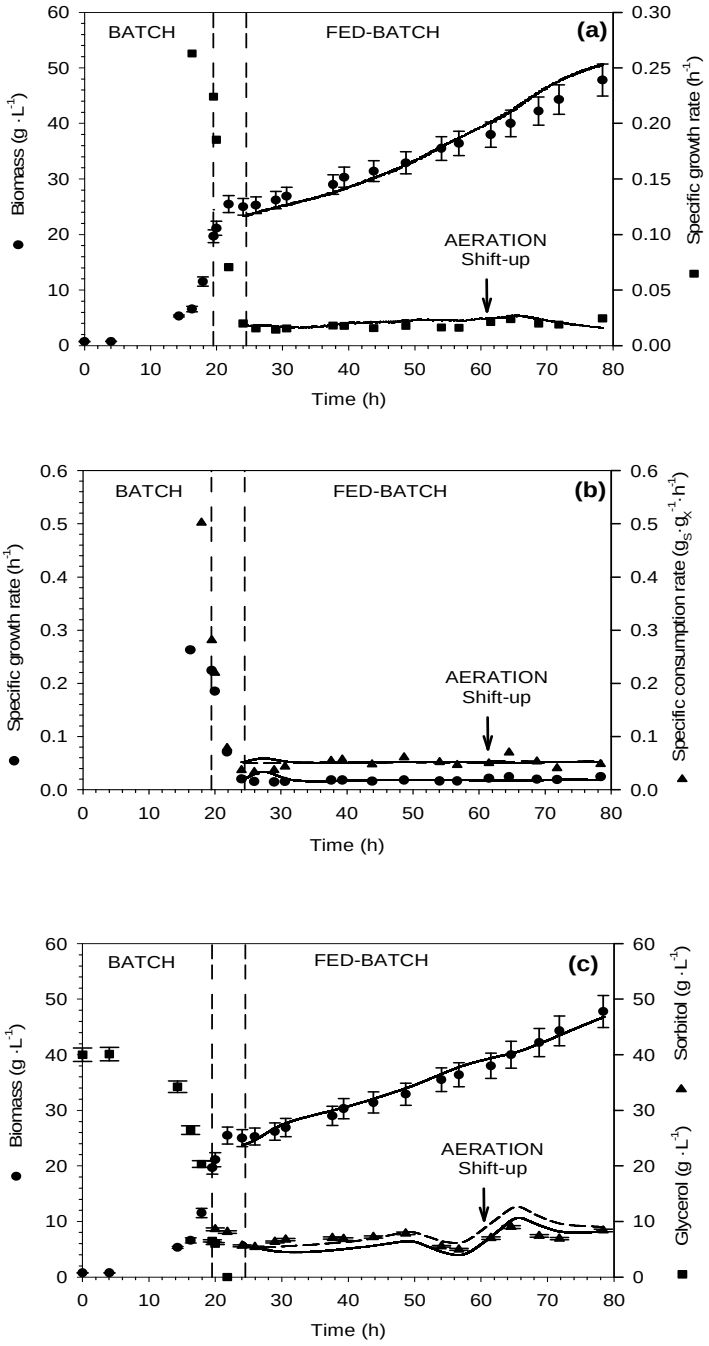

Figure 7**.** *AO-SODE* validation from experimental data for the *P. pastoris PFLD1*-based system. **(a)** Specific growth rate and biomass estimation using *S* measurements; **(b)** Specific growth rate estimation and specific consumption rate using O_2 and *OTR*; **(c)** Biomass estimation and reconstructed substrate using O_2 and *OTR*. *Symbols* correspond to real values, *solid lines* stand for estimation and *dashed lines* were obtained through substrate balance with off-line specific growth rate data.

Results from specific growth rate and biomass estimation using substrate or oxygen measurements are displayed in Figure 7 (a-c). Computed biomass estimation was achieved with an estimation error ≤ 5% although performance was better when gaseous measurements were used instead of substrate measurements as presented in Table 5. The low accuracy of the values of the overall yield coefficients used could explain the differences in biomass estimations (Veloso et al., 2009).

As previously, apart from X and μ estimations using gaseous measurements, substrate estimation was performed from the substrate balance and substrate uptake rate (SUR) estimation. Similarly, the substrate concentration was reconstructed employing off-line X and μ data with the aim of comparing the accuracy of computed estimates. Although slight deviations in the computed substrate were detected, reconstructed substrate has a similar precision than the estimated one.

Notably, when the *AO-SODE* method was applied to O_2 and *OUR* data or CO_2 and *CPR* measurements, similar results were observed. However, *AO-SODE* (O_2, *OUR*) was evaluated achieving the least deviation error (*RMSE* and *MRE*) of state variables, as detailed in Table 5.

Finally, all estimation methods could follow substrate evolution with less deviation error than obtained in simulation results, as presented in Table 4 and Table 5.

2.5. Conclusions

Simulation results obtained for the non-linear *OBE's* showed adequate global performance, but an optimal tuning cannot be easily derived. Despite of their stability, dynamics of convergence and parameter tuning could be improved using a variable gain-structure, it is necessary to know with high accuracy the yield and maintenance coefficients of the bioprocess model. However, these values can be different, depending on the carbon source and environmental conditions. So, when complex nutrients are consumed in a significant quantity, these parameters may also vary during the fermentation process. Thus, errors on model coefficients may produce inaccurate results in estimation of μ and, consequently, the error is propagated and amplified on X and S estimation. Hence, these observers are not recommended for processes with poorly known parameters in a preliminary selection. Additionally, *NLOBE* was established as the most tuning sensitive, being rather dependent on their tuning parameters and initial values.

The use of *RLS* methods to identify the specific growth rate allows diminishing requirements about the knowledge of the system, besides providing other important advantages. Among them, low mathematical complexity because it is only necessary to solve linear equations instead of differential equations like in *OBE's*, increase of identification frequency, high adaptation capacity to process changes, short response times and reduced number of both tuning and initialization variables. The implementation of the *RLS-VFF* method allows that the whole estimation system is fairly dependent on the tuning parameters applied. Furthermore, the variable forgetting factor allows that a

minimal excitation of the estimator is maintained through the bioprocess preventing a constant reduction of the covariance matrix. In the cases studied, this is a great advantage, since there is a continuous addition of substrate in a relatively long process with a low specific growth rate. Moreover, it includes phases where the carbon source has to be replaced and also operational perturbations often appear. This method was able to make a suitable identification of the specific growth rate both in batch and fed-batch processes, whenever slow variation of the specific growth rate is presented.

Asymptotic observers (*AO*), as well as recursive least square (*RLS*) methods can be applied to correctly estimate the specific growth rate. Concretely, *AO-SODE* performed better than the *RLS-VFF* when moderate to rapid changes of the specific growth rate appeared because model parameters were well known. On the other hand, when slow changes on the specific growth rate were presented in the bioprocess, for instance in a substrate control operation, *RLS-VFF* was come up as the best option, because of its reduced requirements. In addition, biomass and substrate were also satisfactorily predicted, solving their corresponding mass balances once the specific growth rate had been estimated.

To conclude, the method which was resulted especially efficient for the proposed system was the *AO-SODE (O_2, OUR)* and was evaluated achieving the least deviation error. The overall methodology presented in this paper can be used to improve the global performance of the process in terms of productivity, yields and reproducibility in the heterologous protein production by *P. pastoris* and for real-time monitoring of the key

fermentation variables. This will represent a significant contribution through a more efficient process of heterologous protein production by *P. pastoris*.

In a future approach, the chosen estimation procedure will allow the implementation of a "true" specific growth rate controller. This would represent an improvement regarding those based either on pre-programmed exponential feeding or an indirect μ-control, keeping constant the substrate concentration.

2.6. References

1. Amigo JM, Surribas A, Coello J, Montesinos JL, Maspoch S, Valero F. 2008. On-line parallel factor analysis. A step forward in the monitoring of bioprocess in real time. Chemom Intell Lab Syst. 92:44-52.
2. Arndt M, Kleist S, Miksch G, Friehs K, Flaschel E, Trierweiler J, Hitzmann BA. 2005. Feedforward-feedback substrate controller based on a Kalman filter for a fed-batch cultivation of *Escherichia coli* producing phytase**.** Comput Chem Eng**.** 29:1113-1120.
3. Bastin G, Dochain C. 1990. On-line estimation and adaptive control of bioreactors. Amsterdam: Elsevier science publishers, B. V.
4. Becker Th, Hitzmann B, Muffler K, Poertner R, Reardon KF, Stahl F, Ulber R. 2007. Future aspects of bioprocess monitoring. Adv Biochem Eng Biotechnol. 105:249-293.
5. Carrier JF, Stephanopoulos G. 1998. Wavelet-based modulation in control-relevant process identification. AIChE J. 44:341-360.
6. Cos O, Resina D, Ferrer P, Montesinos JL, Valero F. 2005a. Heterologous production of *Rhizopus oryzae* lipase in *Pichia pastoris* using the alcohol oxidase and formaldehyde dehydrogenase promoters in batch and fed-batch cultures. Biochem Eng J. 26: 86-94.
7. Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005b. 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.
8. Cos O. 2005c. Monitoritzacio i control del procés de producció de proteïnes heteròlogues en el llevat metilotròfic *Pichia pastoris*, PhD Thesis, Departament d'Enginyeria Química, ETSE, Universitat Autònoma de Barcelona.
9. Cos O, Ramon R, Montesinos JL, 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. (13!!!)
10. Cregg JM, Shen S, Johnson M, Waterham HR. 1998. Classical genetic manipulation. In: Higgins DR, Cregg JM. Methods in Molecular Biology. *Pichia* protocols. Totowa, NJ: Humana Press. 17-26.

11. Cregg JM. Heterologous protein expressed in *Pichia pastoris*, available online: http://www.kgi.edu/documents/faculty/James_Cregg/heterologous_proteins_expressed_in_Pichia_pastoris.pdf (Accessed 08 December 2014).
12. Dochain D. 2003. State and parameter estimation in chemical and biochemical processes: a tutorial. J Proc Control**.** 13:801-818.
13. Estler MU. 1995. Recursive on-line estimation of the specific growth rate form off-gas analysis for the adaptive control of fed-batch processes. Bioprocess Eng. 12:205-207.
14. Farza M, Hammouri H, Jallut C, Lieto, J. 1999. State observation of a nonlinear system: application to (bio)chemical processes. AIChE J. 45:93-106.
15. Farza M, Nadri M, Hammouri H. 2000. Nonlinear observation of specific growth rate in aerobic fermentation processes. Bioprocess Eng. 23:359-366
16. Golobic I, Gjerkes H, Malensek J. 1999. On-line estimation of the specific growth rate in the bacitracin fermentation process. AIChE J. 45:2550-2556.
17. Harder W, Veenhuis M. 1989. Metabolism of one-carbon compounds. In: Rose AH, Harrison JS. The Yeasts, vol. 3. London: Academic Press, 289-316.
18. Horstkotte B, Arnau C, Valero F, Elsholz O, Cerdà V. 2008. Monitoring of sorbitol in *Pichia pastoris* cultivation applying sequential injection analysis. Biochem Eng J. 42: 77-83.
19. James SC, Legge RL, Budman, H. 2000. On-line estimation in bioreactors: A review. Rev in Chem Eng. 16:311-340.
20. Jensch M, Simutis R, Eisbrenner G, Stückrath I, Lübbert A. 2006. Estimation of biomass concentrations in fermentation processes for recombinant protein production. Bioprocess Biosyst Eng. 29:19-27.
21. Junker BH, Wang HY. 2006. Bioprocess monitoring and computer control: Key roots of the current PAT initiative**.** Biotechnol Bioeng. 95:226-261.
22. Kadlec P, Gabrys B, Strandt S. 2009. Data-driven soft sensors in the process industry. Comput Chem Eng**.** 33:795-814.
23. Kiviharju K, Salonen K, Moilanen U, Eerokäinen T. 2008. Biomass measurement online: the performance of in situ measurements and software sensors. J Ind Microbiol Biotechnol. 35:657-665.

24. Lin Cereghino GP, Lin Cereghino J, Sunga AJ, Johnson MA, Lim M, Gleeson MAG, Cregg JM. 2001. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*. Gene. 263:159-169.
25. Loeblein C, Perkins JD. 1999. Structural design for on-line process optimization: I. Dynamic Economics of MPC. AIChE J. 45:1018-1029.
26. Lubenova V., Rocha I., Ferreira E.C. 2003. Estimation of multiple biomass growth rates and biomass concentration in a class of bioprocesses. Bioprocess Biosyst Eng. 25:395-406.
27. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast. 22:249-270.
28. Madrid RE, Felice CJ. 2005. Microbial biomass estimation. Crit Rev in Biotechnol. 25:97-112.
29. Montesinos JL, Lafuente J, Gordillo MA, Valero F, Solà C, Charbonnier S, Cheruy A. 1995. Structured modeling and state estimation in a fermentation process: lipase production by *Candida rugosa*. Biotechnol Bioeng. 48:573-584.
30. Nadri M, Trezzani I, Hammouri H, Dhurjati P, Longin P, Lieto J. 2006. Modeling and observer design for recombinant *Escherichia coli* strain. Bioprocess Biosyst Eng. 28:217-225.
31. Oliveira R, Ferreira EC, Feyo de Azevedo S. 2002. Stability, dynamics of convergence and tuning of observer-based kinetics estimators. J Process Control. 12:311-323.
32. Pérez J, Montesinos JL, Gòdia F. 2006. Gas-liquid mass transfer in an upflow cocurrent packed-bed biofilm reactor. Biochem Eng J. 188-196.
33. Petkov SB, Davis RA. 1996. On-line biomass estimation using a modified oxygen utilization rate. Bioprocess Eng. 15:43-45.
34. Pomerleau Y, Perrier M. 1992. Estimation of multiple specific growth rates: experimental validation. AIChE J. 38:1751-1760.
35. Ramon R, Feliu JX., Cos O, Montesinos JL, Berthet FX, Valero F. 2004. Improving the monitoring of methanol concentration during high cell density fermentation of *Pichia pastoris*. Biotechnol Lett. 26:1447-1452.

36. Ramon R 2007. Estratègies d'operació en el procés de producción de proteïnes heteròlogues en *Pichia pastoris*: Aplicació de tècniques de monitoritzacio i control, PhD Thesis, Departament d'Enginyeria Química, ETSE, Universitat Autònoma de Barcelona.
37. Roux G, Dahhou B, Queinnec I. 1996. Modelling and estimation aspects of adaptive predictive control in a fermentation process. Control Engineering Practice. 4: 55-66.
38. Shen S, Sulter G, Jeffries T, Cregg JM. 1998. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. Gene. 216:93-102.
39. Sunström H, Enfors SO. 2008. Software sensors for fermentation processes. Bioprocess Biosyst Eng. 31:145-152.
40. Surribas A, Geissler D, Gierse A, Scheper Th, Hitzmann B, Montesinos JL, Valero F. 2006a. State variables monitoring by *in*-situ mulit-wavelenght fluorescence spectroscopy in heterologous protein production by *Pichia pastoris.* J Biotechnol. 124:412-419.
41. Surribas A, Montesinos JL, Valero F. 2006b. Biomass estimation using fluorescence measurements in *Pichia pastoris* bioprocess. J Chem Technol Biotechnol. 81:23-28.
42. Ungarala S, Co TB. 1998. Model parameter tracking in microbial growth processes. AIChE J. 44:2129-2134.
43. Valero F, Lafuente J, Poch M, Solà C. 1990. Biomass estimation using on-line glucose monitoring by FIA: Application to *Candida rugosa* batch growth. Appl Biochem Biotechnol. 24-25:591-602.
44. Van Impe JF, Claes JE. 1999. On-line estimation of the specific growth rate based on viable biomass measurements: experimental validation. Bioprocess Eng. 21:389-395.
45. Veloso ACA, Rocha I, Ferreira EC. 2009. Monitoring of fed-batch *E. coli* fermentations with software sensors. Bioprocess Biosyst Eng. 32:381-388.

Chapter 3:

Searching the best operational strategies for *Rhizopus oryzae* lipase production in *Pichia pastoris* Mut$^+$ phenotype: Methanol Limited or Methanol Non-Limited Fed-Batch cultures?

Published as: 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.

---

3. Searching the best operational strategies for *Rhizopus oryzae* lipase production in *Pichia pastoris* Mut$^+$ phenotype: Methanol Limited or Methanol Non-Limited Fed-Batch cultures?

3.0. Abstract

Two different operational strategies, methanol limited (MLFB) and methanol non-limited (MNLFB) fed-batch cultures, have been compared for the production of the recombinant *Rhizopus oryzae* lipase (ROL) expressed in *Pichia pastoris* Mut$^+$ phenotype. Yields, productivities and specific production rate in all MLFB conditions were very low, obtaining the best results at the lower specific growth rates. Similar results were obtained for MNLFB strategy at methanol set-point concentrations up to 2 $g \cdot L^{-1}$. However, for methanol set-points higher than 2 $g \cdot L^{-1}$ a significant increase in the production was observed. Methanol set-point of 3 $g \cdot L^{-1}$ was the optimum to maximize product yield ($Y_{P/X}$), both volumetric and specific productivities and mean specific production rate ($q_{P,mean}$), although in terms of ROL production high values were obtained in the range of 3-10 $g \cdot L^{-1}$. An inhibitory effect of methanol on cell growth as well as slowness on ROL production during early induction phase was observed at concentration of 10 $g \cdot L^{-1}$.

3.1. Introduction

The methylotrophic yeast *Pichia pastoris* has become one of the most widely yeast systems for the production of heterologous proteins (Macauley-Patrick et al., 2005). The strong and tightly regulated promoter from the alcohol oxidase 1 gene P_{AOX1} still is, by far, the most common system used (Cereghino and Cregg, 2000; Sreekrishna, 2010; Potvin et al., 2012).

The most frequently used cultivation strategy to achieve high cell densities and high heterologous protein production levels with AOX promoter-based systems is the fed-batch operation. Typically, such processes are divided into three phases: Glycerol batch phase (GBP), transition phase (TP), and finally the methanol induction phase (MIP).

The objective of the GBP is the fast generation of biomass previous to the induction by methanol. The specific growth rate and biomass yield of *P. pastoris* growing on glycerol are higher than methanol. Once the GBP is finished, indicated by a spike in measured DO, the TP starts (Cos et al., 2006b). Although the MIP may be carried out after GBP (Invitogen, 2014), the TP has been recommended for Mut$^+$ phenotype (Zhang et al, 2007). The objective of the TP is to achieve both high cell density cultures and the derepression of the *AOX1* promoter, due to the non-excess of glycerol, in order to adapt the cell metabolism to the consumption of methanol by means of the induction of P_{AOX1}. The procedure consists on a glycerol feeding by decreasing the flow rate during 3-5 h and also adding methanol with a pulse (Zhang et al, 2000) or with a low feed rate (Cos et al, 2005b). Finally, in the MIP, methanol is used as the carbon and inducer substrate. The

selected methanol feeding strategy used in the MIP is one of the most important factors to maximize heterologous protein production (Zhang et al, 2000).

Standard operational strategies are based on the control of the substrate concentration close to zero (limiting strategies). However, maintenance of the concentration at a constant value is also extensively described (non-limiting strategies) (Zhang et al, 2000; Cos et al., 2006b; Jahic et al, 2006; Zhang et al, 2007). Since aerobic conditions are strictly essential for methanol assimilation, there are at least two possible limiting substrates for the process: oxygen and methanol. Thus, the most common methanol feeding strategies are: dissolved oxygen control (DO-stat), oxygen limited fed-batch (OLFB), methanol limited fed-batch (MLFB), methanol non-limited fed-batch (MNLFB), as well as temperature limited fed-batch (TLFB) (Cos et al., 2006b; Jahic et al, 2006; Potvin et al., 2012).

Lim et al., (2003) developed DO-stat control that automatically handles the partial pressure of oxygen in the inlet air stream and the methanol feeding rate during induction. Alternatively, Yamawaki and co-workers controlled the DO by adjusting the methanol flow rate (Yamawaki et al., 2007). The weakest point of this kind of control is that it is quite difficult to obtain bioprocess reproducibility, because some key variables and specific rates, such as methanol concentration and specific growth rate (μ) are not constant (Cos et al., 2006b).

Although oxygen limitation should generally be avoided during the induction phase, (Invitrogen, 2014), OLFB cultivations have been successfully implemented (Potvin et al., 2012). In OLFB, the residual methanol concentration is kept constant at about 1%, but DO concentration always drops to 0% due to oxygen limitation (Trentmann et al., 2004; Berdichevsky et al., 2011). The main advantage of OLFB is that this control strategy minimizes oxygen requirements and, therefore, may improve the economic feasibility of the process by reducing the cost on oxygen (Khatri et al., 2006). However, the methanol requirements were increased in these oxygen-limited conditions (Potvin et al., 2012).

MLFB is a control method that adjusts the methanol feeding rate based on mass balance equations to, theoretically, maintain a constant specific growth rate (μ) (Cos et al, 2005a). This method is the most widely reported for heterologous protein production by *P. pastoris*, although it can be referred as specific growth rate control (Dabros et al., 2010), open loop control (Cos et al, 2006a) or pre-programmed exponential feeding rate addition (Arnau et al., 2011). When a feed-forward structure control is selected, a simple cell growth model is considered and on-line information about the system is not strictly required. It is assumed that the residual methanol concentration on the cultivation broth is close to zero during the induction phase. In order to minimise deviations of the methanol feeding rate and compensate other perturbations, which can derive into methanol accumulation, a feedback term can be introduced (Sinha et al., 2003; Dabros et al., 2010).

The main advantage of MLFB is that keeping μ constant improves process reproducibility and allows for the study of the μ effects on the heterologous protein production.

The control strategy consisting in maintaining the methanol concentration constant is called MNLFB. Accurate methanol monitoring (Katakura et al., 1998) (Surribas et al., 2003; Amigo et al., 2008), and efficient control are required for robust and reproducible bioprocesses. A feedback methanol control strategy has been used in numerous studies (Pla et al., 2006; Surribas et al., 2007). Proportional-integral (PI) or proportional-integral-derivative (PID) and control heuristics are reported to be effective to maintain constant the methanol concentration (Potvin et al., 2012). A predictive control algorithm combined with a PI feedback controller was used to optimize the production of *Rhizopus oryzae* lipase in *P. pastoris* (Cos et al., 2006a). Different methanol concentrations directly affect cell growth, the viability and the heterologous protein production (Jahic et al., 2006). Thus, an accurate monitoring and control allow determining stoichiometric and kinetic models that are particularly useful to optimize the system (Schenk et al., 2007).

Some adjustments to the standard process have been proposed to minimize proteolysis, including the manipulation of culture media and temperature (Surribas et al., 2007) . The dissolved oxygen (DO) can be indirectly controlled by either temperature or the methanol feeding rate. TLFB process was compared with a methanol limited fed-batch by Jahic et al., (2003). By reducing the temperature from 30ºC to 25ºC, the product amount was increased 2-fold. Moreover, the final product purity was also increased.

In this work, the selected target protein was a recombinant *Rhizopus oryzae* lipase (ROL). Process development for ROL production has been widely studied from a bioprocess engineering point of view (Cos et al, 2005b; Cos et al, 2006a; Surribas et al., 2007; Arnau et al., 2010). Furthermore, ROL has been reported as biocatalyst with interesting applications in the field of structured lipids (Nunes et al., 2012; Tecelão et al., 2012) and flavours (Guillén et al., 2012). The effect of MNLFB and MLFB operational strategies on ROL production has been studied. These most commonly applied control strategies allow maintaining rather constant key specific rates: cell growth (μ), substrate uptake (q_s), and protein production (q_p) through the quasi-steady state hypothesis for substrate. Results were analysed in order to determine the most suitable operating conditions in terms of yields and productivities.

3.2. Materials and methods

3.2.1. Strains

The wild type *P. pastoris X-33* strain containing the vector pPICZαAROL was used for heterologous expression of *Rhizopus oryzae* lipase (ROL) under the control of the *AOX1* promoter (Cos et al, 2005a).

3.2.2. Inoculum preparation

Pre-inocula for bioreactor cultures were grown for 24 h in 1 L baffled shake flasks at 30 ºC, 150 rpm, in YPD medium containing 10 g yeast extract, 20 g peptone, 20 g glucose and 1 mL zeocin per litre of distilled water. Shake flasks contained 200 mL of YPD medium. The culture was centrifuged at 4500 x g, the harvested cells were re-suspended in bioreactor culture medium and used to inoculate a 5 L Biostat ED bioreactor (Braun Biotech, Melsungen, Germany).

3.2.3. Fed-batch cultivation set-up and operational conditions

The basal salt synthetic medium for fed-batch cultivations contained per litre of distilled water: H_3PO_4 (85%) 26.7 mL, $CaSO_4$ 0.93 g, K_2SO_4 18.2 g, $MgSO_4 \cdot 7H_2O$ 14.9 g, KOH 4.13 g, glycerol 40 g, 2 mL of biotin solution (200 $mg \cdot L^{-1}$), 5 mL of trace salts solution and 0.5 $mL \cdot L^{-1}$ of antifoam agent (A6426, Sigma-Aldrich Co., St. Louis, MO, USA). The trace salts solution contained per litre: $CuSO_4 \cdot 5H_2O$ 6.0 g, NaI 0.08 g, $MnSO_4 \cdot H_2O$ 3.0 g, $Na_2MoO_4 \cdot 2H_2O$ 0.2 g, H_3BO_3 0.02 g, $CoCl_2$ 0.5 g, $ZnCl_2$ 20.0 g, $FeSO_4 \cdot 7H_2O$ 65.0 g, biotin 0.3 g, H_2SO_4 concentrated 5 ml. The biotin and trace salts solutions were

sterilized separately by microfiltration (SLGV013SL 0.22 µm, Millipore Corporation, Billerica, MA, USA).

Cells were cultured in a 5 L Braun Biostat ED bioreactor (Braun Biotech, Melsungen, Germany) under the following cultivation conditions: initial volume 2 L, stirring rate 800 rpm, temperature 30 °C, pH controlled at 5.5 by adding NH_4OH 30% (v/v) during the batch phase, and KOH (5 M) during the transition and induction phase, dissolved oxygen above 7.0 % in pure oxygen saturation (≈30% air saturation), with a gas flow rate, air enriched in oxygen, between 0.5 and 7.5 $L \cdot min^{-1}$. The cultivation started with a 40 $g \cdot L^{-1}$ glycerol batch phase (GBP). Secondly, when glycerol was exhausted, detected by a sudden increase in DO concentration, a 5 h transition phase (TP) started (Cos et al, 2005b). Finally, the methanol induction phase (MIP) was carried out using methanol as sole carbon source and inducer substrate. The methanol addition rate was implemented with two different strategies in order to make either methanol limited fed-batch cultures (MLFB) or methanol non-limiting fed-batch cultures (MNLFB). Thus, the goal was to keep either the specific growth rate (µ) constant or to control the residual substrate concentration (S) that is directly related to µ.

Separated solutions of 500 mL of glycerol (50% (v/v)) and 2 L of pure methanol, complemented with 2 and 10 mL of trace salts solution respectively and also 0.15 and 0.6 g of biotin in each solution were added during the transition phase. In the induction phase only methanol solution was fed. It was made by an automatic Microburette 1S (Crison Instruments S.A., Alella, Barcelona, Spain). The microburette addition system was

chemically sterilized by using solutions of HCl 1 M, Ethanol 70%, NaOH 1 M and sterile water (Arnau et al., 2010).

To supply the nitrogen strain requirements, 40 g of ammonium chloride diluted on 200 mL of sterile water were added to the reactor, when approximately 30-35 $g \cdot L^{-1}$ of dry cell weight (DCW) were reached. All the fermentations were finished when biomass reached a value between 50-60 $g \cdot L^{-1}$ DCW to compare the experiments with a similar biomass concentration and always below maximal working volume.

3.2.3.1. Methanol limited fed-batch operation (MLFB)

A pre-programmed exponential substrate addition strategy was implemented for methanol limiting culture conditions during the induction phase. This feed-forward controller was based on fed-batch substrate balance for a quasi-steady state (Cos et al, 2005a) and it was programmed in order to maintain μ constant during this phase. The first step of this strategy is to initialize (t=0) the substrate feeding rate (F) by Eq.1, where $Y_{X/S}$ is the biomass to substrate yield [$g \cdot g^{-1}$], S_{feed} the methanol feeding concentration [$g \cdot L^{-1}$], $(XV)_0$ the initial overall biomass [g], S_0 residual methanol concentration [$g \cdot L^{-1}$], μ_{sp} the set-point value for μ [h^{-1}]. Initial values were: $Y_{X/S}$ =0.20 $g \cdot g^{-1}$; S_{feed} = 775 $g \cdot L^{-1}$; $(XV)_0$ = 50 g; $S_0 \approx 0.0$ $g \cdot L^{-1}$.

$$F_0 = \frac{(XV)_0 \cdot \mu_{sp}}{Y_{X/S} \cdot (S_{feed} - S_0)} \qquad (1)$$

After that, the feed flow rate can be updated according to the following equation:

$$F_{t+\Delta t} = F_t \cdot \exp(\mu_{sp} \cdot \Delta t) \quad (2)$$

3.2.3.2. Methanol non-limited fed-batch operation (MNLFB)

A predictive-PI control strategy was applied for methanol non-limiting culture conditions during the induction phase. The predictive-PI controller was adapted from the model-based controller previously described for Mut^S phenotype (Cos et al, 2006a) in order to improve the controller response for the Mut$^+$ phenotype, which presents higher specific methanol consumption rate. It was designed and implemented on NI-LabVIEW platform (LabVIEW 8.6 , National Instruments Corp, Austin, TX, USA).

The control law applied is described by the equation:

$$F_{t+\Delta t} = F_t - \frac{V}{(S_{feed} - S_t)} \cdot \frac{dS}{dt} + K_P \left(\varepsilon_t + \frac{1}{\tau_i} \int_o^t \varepsilon_t \, dt \right) \quad (3)$$

Firstly, from Eq (3) the first-time derivative of methanol concentration for each time interval was bounded taking into account biological constraints, i.e. maximum specific substrate consumption

$$0 \geq \frac{dS}{dt} \geq -\frac{1}{Y_{X/S}} \cdot \mu_{max} \cdot X_0 \cdot exp(\mu_{max} \cdot t) \quad (4)$$

The heuristic method was used to tune the controller parameters (K_P=400 μL·L·g^{-1}·min^{-1}; τ_i=6 min). The integral error term was computed for a moving-window of 30 minutes.

Finally, the feed flow rate updated for the next time interval ($F_{t+\Delta t}$) was also bounded to the feeding capacity (F_{max}).

$$0 \leq F_{t+\Delta t} \leq F_{max} \quad (5)$$

3.2.3.3. On-line methanol determination

Methanol concentration was on-line monitored using a sensor immersed in the culture broth (Raven Biotech, Vancouver, BC, Canada) (Arnau et al., 2011). Specifically, the sensing element that detects the methanol on that equipment was the Figaro TGS-822 (Figaro USA Inc., Glenview, USA) (Ramon et al., 2004).

3.2.4. Biomass analysis

Biomass concentration was quantified as DCW per litre of culture broth. Culture samples were collected by centrifugation. Pellets were washed and centrifuged twice in ddH_2O at 4500 x g during 3 min, dried at 105 ºC until constant weight. Relative standard deviation was about 5%.

3.2.5. Off-line glycerol and methanol determination

Methanol and glycerol were determined by HPLC as reported elsewhere (Arnau et al., 2011). Residual standard deviation (RSD) was estimated about 0.5%.

3.2.6. Lipolytic activity assay

Extracellular lipolytic activity was measured after removing cells from cultivation samples by centrifugation (10000 × g) during 1 min. Then, the lipolytic activity was followed spectrophotometrically in a Cary Varian 300 spectrophotometer (Varian Inc, Palo Alto, USA) at 30ºC in 400mM Tris-HCl + 10mM $CaCl_2$ buffer at pH 7.25 using the Roche lipase colorimetric kit (Roche kit 11821792, Mannheim, Germany) as previously described (Resina et al., 2004). The measurement was made at 580 nm and every analysis was carried out by triplicate. RSD was estimated as 5%.

3.3. Theory and calculation

3.3.1. Calculation of the state variables

From the total mass balance for fed-batch operation, volume variation can be obtained by the following equation:

$$\frac{dV}{dt} = \frac{\rho_{Feed}F - \rho_{H2O}F_{Evap} + \rho_{Base}F_{Base} - \rho_{Broth}F_O + M_{GAS}}{\rho_{Broth}} \quad (6)$$

where V is the volume of broth in the reactor [L], F the volumetric feeding rate [$L \cdot h^{-1}$], F_{Evap} the water evaporation rate [$L \cdot h^{-1}$], F_{Base} the base feeding rate [$L \cdot h^{-1}$], F_O the withdrawal rate [$L \cdot h^{-1}$], M_{GAS} net mass gas flow rate [$g \cdot h^{-1}$], ρ_{Feed} substrate feed density [$g \cdot L^{-1}$], ρ_{H2O} water density [$g \cdot L^{-1}$], ρ_{Base} base density [$g \cdot L^{-1}$], ρ_{Broth} mean broth density [$g \cdot L^{-1}$]. The net mass gas flow rate is calculated with the equation (7).

$$M_{GAS} = W_{O2}\, OUR - W_{CO2}\, CPR \quad (7)$$

where OUR is the oxygen uptake rate [$mol \cdot h^{-1}$], CPR carbon dioxide production rate [$mol \cdot h^{-1}$], W_{O2} oxygen molar mass [$g \cdot mol^{-1}$], W_{CO2} carbon dioxide molar mass [$g \cdot mol^{-1}$]. OUR and CPR were estimated as previously reported (Barrigón et al., 2012).

Biomass, substrates and product concentrations were determined as described earlier in section 2. Although biomass concentration (X) is referenced to the total volume (V), substrates (S_{Liq}): glycerol, methanol; and the lipolytic activity (P_{Liq}) were measured on the supernatant, so in the liquid phase, excluding the pellet volume.

Thus, substrates (SV) (on-line and off-line measurements) and total lipolytic activity (PV) were recalculated on the total phase by Eq (8) and Eq (9) (Borzani, 2004), where microorganism density is ρ = 1068 [g·L^{-1}] and the fraction of the dry matter in the biomass σ = 0.304 [g·g^{-1}].

$$(PV)_t = P_{Liq,t} V_t \left(1 - \frac{X_t}{\sigma\,\rho}\right) \quad (8)$$

$$(SV)_t = S_{Liq,t} V_t \left(1 - \frac{X_t}{\sigma\,\rho}\right) \quad (9)$$

3.3.2. Calculation of discrete specific rates

The estimation procedure to determine $\mu_{(t)}$, $q_{S(t)}$, and $q_{P(t)}$ was adapted from (Cos et al, 2005b). Firstly, global state variables ((XV), (SV) and (PV)) were estimated within the induction time by applying the smoothing tool (Matlab R2009a Curvefit Toolbox, The Mathworks Inc., Natik, USA) from off-line data. The first time-derivatives of the smoothed curves were also obtained. Finally, the $\mu_{(t)}$, $q_{S(t)}$, and $q_{P(t)}$ were calculated by using their corresponding mass balances, Eq (10), Eq (11) and Eq (12) within induction time. Uncertainties of discrete specific rates were calculated by error propagation.

$$\mu_{(t)} = \frac{1}{(XV)_t} \frac{d(XV)_t}{dt} \quad (10)$$

$$q_{S_{(t)}} = \frac{1}{(XV)_t} \left(F_t \cdot S_{feed} - \frac{d(SV)_t}{dt}\right) \quad (11)$$

$$q_{P_{(t)}} = \frac{1}{(XV)_t} \frac{d(PV)_t}{dt} \quad (12)$$

3.3.3. Calculation of mean specific rates

In order to estimate averaged specific induction rates different methods have been used by other authors, that is: arithmetic mean rates (Albaek et al., 2011), time-weighted average rates (Arnau et al., 2011) and linear regressions (Potgieter et al., 2010) .

For the two first methods, discrete specific rates have to be calculated for each off-line value which considers first-time derivatives of the global variables, Eq (10), Eq (11) and Eq (12). The derivatives calculation increases the estimation error.

In this work, in order to avoid the first-time derivatives calculation, a linear regression method was used to obtain the averaged specific rates. For this aim, the global off-line state variables data: (XV), (SV) and (PV); feed flow rate (F) and induction time (t) were utilized. The slope of each linear regression, Eq (13), Eq (14) and Eq (15), corresponds to the averaged specific rate. The standard error of the averaged values is obtained from the linear regression data.

$$\int_{(XV)_0}^{(XV)} d(XV) = \mu_{mean} \int_{t_0}^{t} (XV)\, dt \quad (13)$$

$$S_{feed} \int_{t_0}^{t} F\, dt - \int_{(SV)_0}^{(SV)} d(SV) = q_{S\,mean} \int_{t_0}^{t} (XV)\, dt \quad (14)$$

$$\int_{(PV)_0}^{(PV)} d(PV) = q_{P\,mean} \int_{t_0}^{t} (XV)\, dt \quad (15)$$

3.4. Results and discussion

3.4.1. Methanol limited fed-batch cultures (MLFB)

In a first set of experiments, several MLFB cultures were carried out with *P. pastoris* Mut$^+$ phenotype at different constant specific growth rates producing·heterologous *Rhizopus oryzae* lipase (ROL), aiming to test the effect of different specific growth rate under methanol limiting conditions on growth, yield and productivities. This strategy is quite simple to be implemented and it has been successfully applied in this cell factory for the production of other heterologous proteins, e.g. BoNT/A(Hc) (Zhang et al, 2000), r-oIFN-τ (Sinha et al., 2003) , α-Gal (Zhang et al., 2005) , YGLY4140 (Potgieter et al., 2010), etc.

The maximal specific growth rate of *P. pastoris* wild type strains growing on methanol as sole C-source has been previously reported to be about 0.09 h^{-1} (Zhang et al, 2000). However, this value is reduced up to 0.06 h^{-1} when *P. pastoris* is producing ROL (Cos et al, 2005a). Thus, three different specific growth rates were selected for this series of experiments: 0.015, 0.020 and 0.045 h^{-1}. The highest value was lower than the maximum reported growth rate in order to avoid possible methanol accumulation during the fed-batch. Time evolution of biomass, methanol and lipolytic activity are shown in Figure 1, whereas the evolution of μ, q_S, and q_P are presented in Figure 2. A summary of maximal lipolytic activity, yield, volumetric and specific productivities, as well as mean specific rates are presented in Table 1.

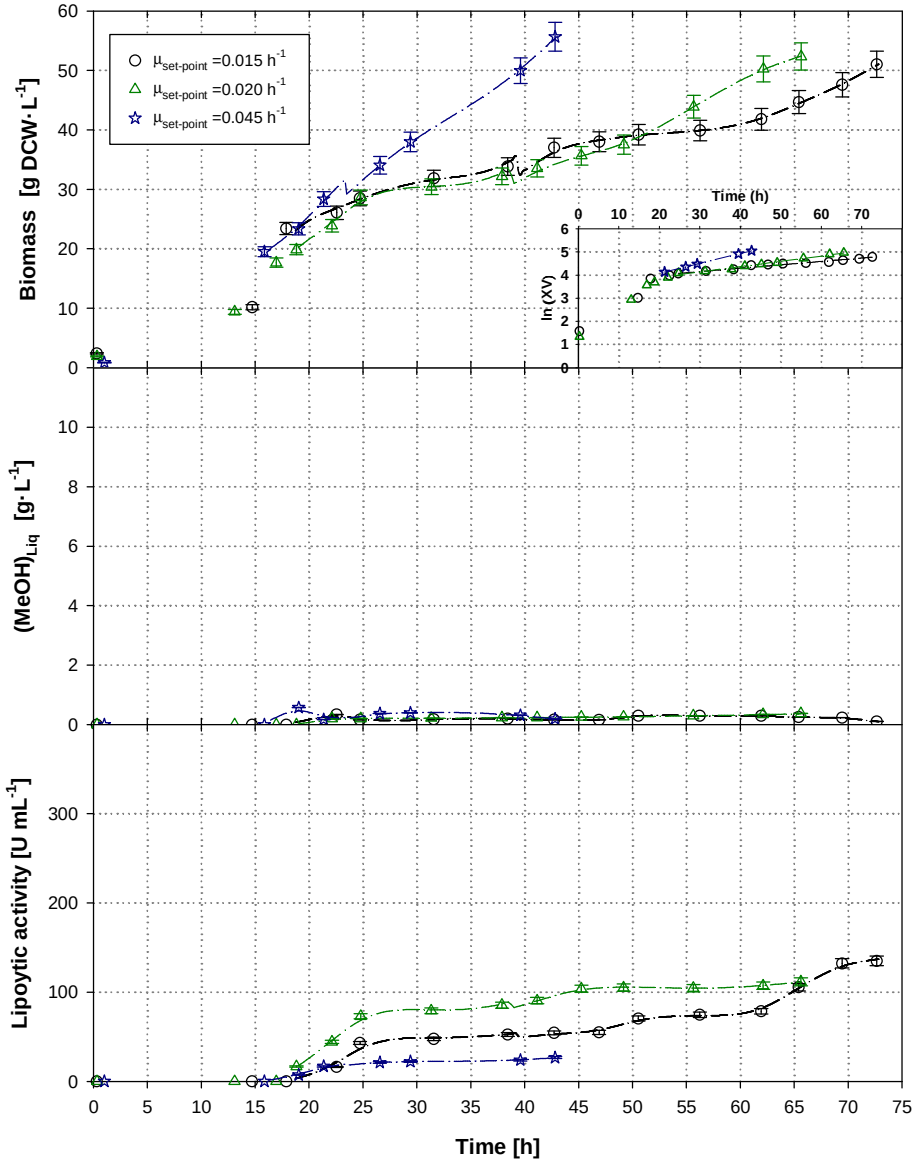

Figure 1. Time evolution of different state variables in fed-batch cultures for Mut$^+$ phenotype using the MLFB operational strategies: biomass, methanol and lipolytic activity. Error bars indicate standard deviation (SD).

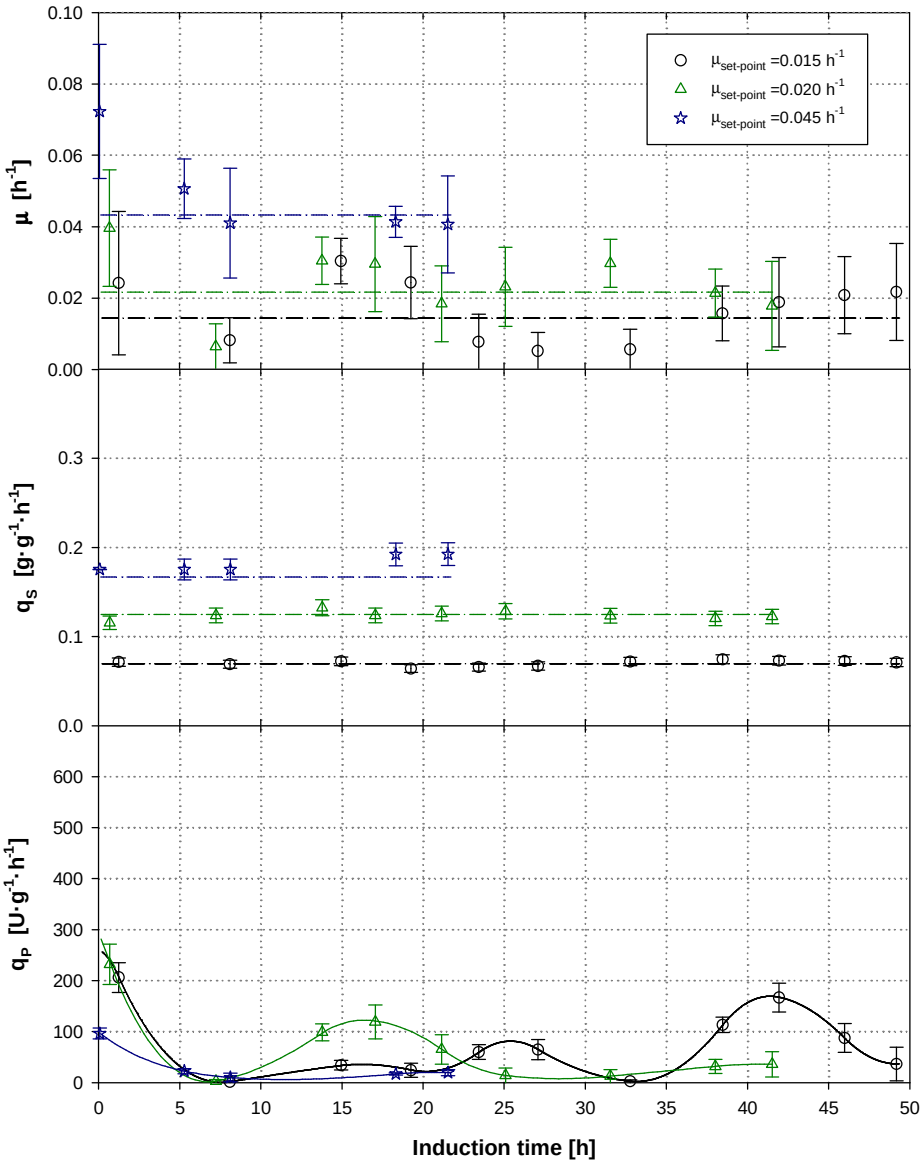

Figure 2. Specific rates in fed-batch cultures for Mut$^+$ phenotype using the MLFB operational strategies: μ, q_S and q_P. Dashed lines indicate the specific mean rate value. Uncertainties are represented by error bars.

In the three MLFB fermentations, the residual methanol concentration was negligible, and the μ_{mean} was conducted properly to the set-point with a relative error lower than 7% (Table 1). In addition, the μ_{mean} and q_{Smean} estimation corroborated that the $Y_{X/S}$ and

$[XV]_0$ values used in the pre-programmed exponential methanol feeding rate equation were suitable (Figure 2, Table 1).

Table 1. Comparison of process variables, yields, productivities and specific rates using different specific growth rate set-points during the induction phase in *Pichia pastoris* MLFB cultures expressing ROL. ± indicate standard error (SE)

μ_{sp}	[h^1]	0.015	0.02	0.045
Max. Lipolytic activity	[U mL^{-1}]	135	112	27
$Y_{P/X}$*	[U $\cdot g^{-1}$]	2644	2130	479
Volumetric Productivity*	[U L^{-1} h^{-1}]	1857	1700	623
Specific Productivity*	[U$\cdot g^{-1}\cdot h^{-1}$]	36	32	11
μ_{mean}	[h^{-1}]	0.014±0.001	0.022±0.001	0.043±0.002
$q_{S,mean}$	[g g^{-1} h^{-1}]	0.07±0.01	0.12±0.01	0.17±0.01
$q_{P,mean}$	[U$\cdot g^{-1}\cdot h^{-1}$]	46±4	46±3	18±2

(*) Estimated over all the fermentation time

During early phase of the induction period a slight adaptation time was observed, probably due to the growth rate reduction in relation to the transition phase. Thereafter, μ was rather constant throughout remaining induction time (Table 1).

For all three MLFB cultures, a slow but continuous product accumulation was observed during the induction phase, in contrast to the fast increase of lipolytic activity observed within the transition phase. Conversely, a sharp increase of the lipolytic activity was also observed towards the end of the fermentation at μ_{sp} = 0.015 h^{-1} (Figure 1).

Hence, $q_{p,max}$ values were achieved during the first induction hours due to the initial higher secretion rates observed in the TP, following the same behaviour as μ. Thereafter, q_p showed fluctuating profiles along the induction phase, reaching values of q_p close to zero in some periods (Figure 2). Overall, the MLFB approaches that yielded higher q_p were those with the lower μ set-points (Table 1). Similar results have been described in the production of r-oIFN-τ (Sinha et al., 2003), where the maximal production rate for MLFB cultures was obtained at a μ_{sp} below 0.025 h^{-1}.

3.4.2. Methanol non-limited fed-batch cultures (MNLFB)

Previous studies have determined that heuristic methanol control strategies cause important fluctuations of methanol concentration around the set-point. These fluctuations affect significantly the recombinant protein level and, consequently, the specific production rate (Cos et al, 2005b).

With the objective to determine the effect of methanol concentration on biomass and ROL production and their specific rates, a predictive-PI control, previously developed for a Mut^S phenotype (Cos et al, 2006a), was adapted to a Mut^+ phenotype. MNLFB fermentations were carried out at the operational range of methanol concentration of 1-10 $g \cdot L^{-1}$. Specifically, the selected MeOH set-points ($MeOH_{sp}$) were: 1.0, 2.0, 3.0, 5.0, and 10.0 $g \cdot L^{-1}$. The time course of biomass, methanol and lipolytic activity of these cultures is illustrated in Figure 3. The evolution of μ, q_S and q_P for the MIP is depicted in Figure 4. A summary of maximal lipolytic activity, yield, both volumetric and specific productivities, and mean specific rates are presented in Table 2.

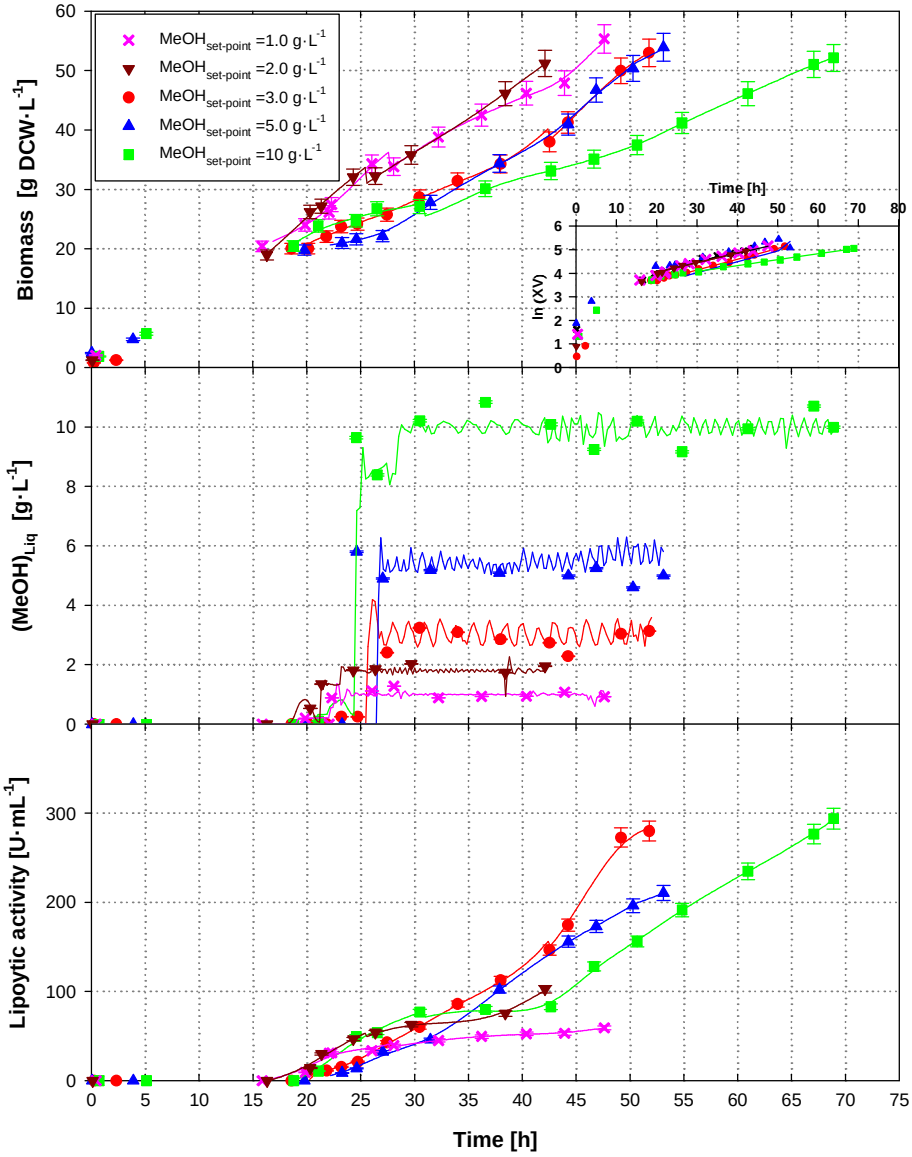

Figure 3. Time evolution of different state variables in fed-batch cultures for Mut$^+$ phenotype using the MNLFB operational strategies: biomass, methanol and lipolytic activity. Error bars indicate standard deviation (SD).

In all MNLFB cultures methanol concentration in the liquid phase was conducted properly to the $MeOH_{sp}$ (Figure 3), thereby proving the robustness of the enhanced control strategy.

Table 2. Comparison of process variables, yields, productivities and specific rates obtained using different methanol set-points during the induction phase in *Pichia pastoris* MNLFB cultures expressing ROL. ± indicate standard error (SE)

$MeOH_{sp}$	[g·L^{-1}]	1.0	2.0	3.0	5.0	10
Max. Lipolytic activity	[U·mL^{-1}]	59	103	280	210	294
$Y_{P/X}$*	[U ·g^{-1}]	1070	2004	5282	3905	5635
Volumetric Productivity*	[U l^{-1} h^{-1}]	1243	2437	5406	3964	4264
Specific Productivity*	[U·g^{-1}·h^{-1}]	22	48	102	74	82
μ_{mean}	[h^{-1}]	0.042±0.001	0.043±0.002	0.046±0.002	0.046±0.002	0.025±0.001
$q_{S,mean}$	[g g^{-1} h^{-1}]	0.23±0.01	0.19±0.01	0.20±0.01	0.20±0.01	0.14±0.01
$q_{P,mean}$	[U·g^{-1}·h^{-1}]	45±2	106±9	322±13	234±8	175±6

(*) Estimated over all the fermentation time

As observed in the MLFB fed-batch series, the initial values of μ in the earlier induction phase were high (Figure 4). The μ_{mean} of the first four MNLFB experiments were very similar, with a maximum value of 0.046 h^{-1} for a $MeOH_{sp}$ of 3 g·L^{-1} (Table 2). On the other hand, a drop of the specific growth rate of about 60% related to the maximum value was observed in the MNLFB culture performed at a methanol concentration of 1%. It is important to emphasize that the observed μ_{mean} is far from the μ_{max} =0.09 h^{-1} reported by Zhang et al. (Zhang et al, 2000). This low value may be related to the influence of high level expression on cell physiology, that is, ROL secretion appears to impose a metabolic burden on *P. pastoris* which results in a negative effect on cell growth (Surribas et al., 2007).

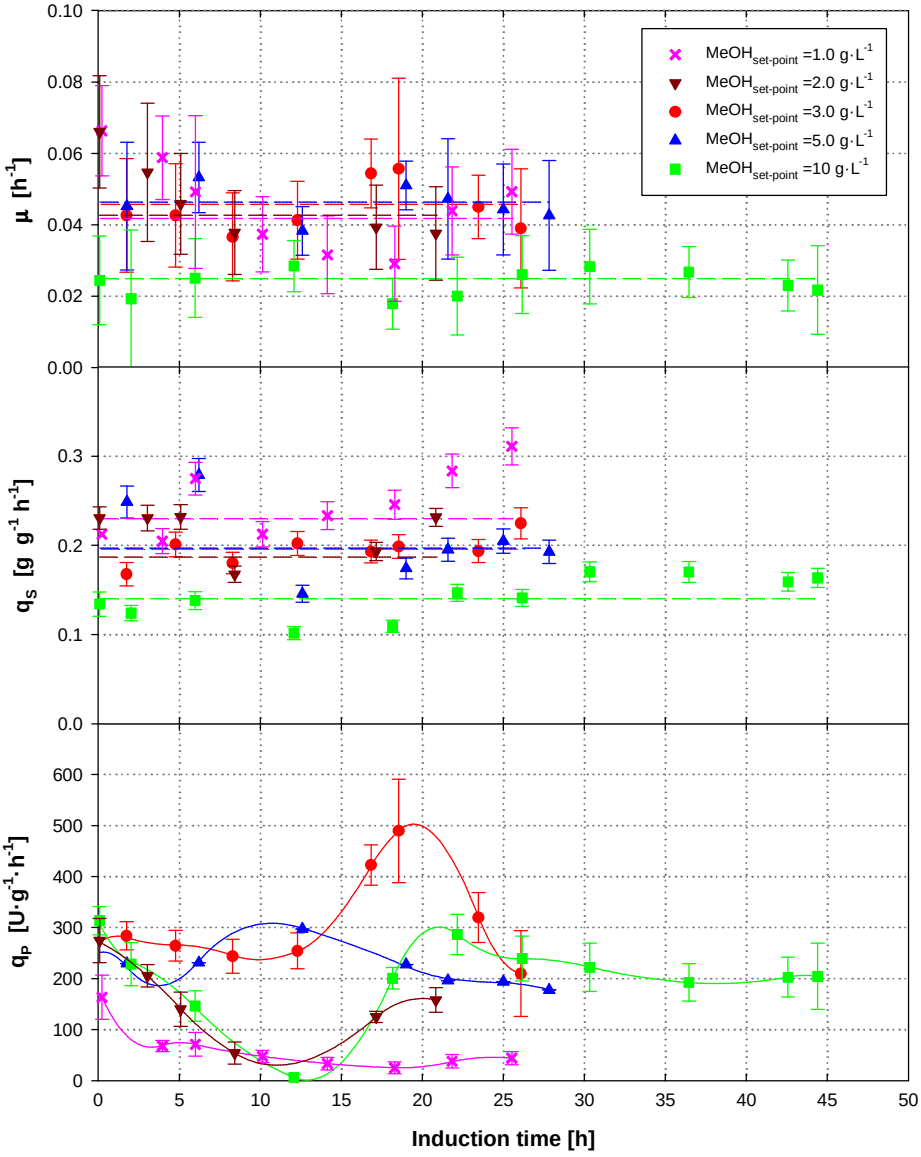

Figure 4. Specific rates in fed-batch cultures for Mut$^+$ phenotype using the MNLFB operational strategies operational feeding strategy of non-limiting methanol: μ, q_S and q_P. Dashed lines indicate the specific mean rate value. Uncertainties are represented by error bars.

Furthermore, a similar behaviour for q_s was obtained. Specifically, while q_S was about 0.2 g·g^{-1} h^{-1} in cultivations with residual methanol concentrations up to 5 g·L^{-1}, a

reduction to 0.14 $g \cdot g^{-1}\ h^{-1}$ was observed when the methanol was set at 10 $g \cdot L^{-1}$ (Table 2 and Figure 4).

Independently of the strategy applied, the relationship between μ and q_S was similar. Hence, μ_{mean} and q_{Smean} of MLFB cultures at μ_{sp}=0.045 h^{-1} were similar to those estimated for MNLFB cultures at $MeOH_{sp}$=2.0 and 3.0 $g \cdot L^{-1}$. Also, μ_{mean} and q_{Smean} of MLFB cultures at μ_{sp}=0.020 h^{-1} were similar to those calculated for MNLFB cultures at $MeOH_{sp}$=10 $g \cdot L^{-1}$. However, the feeding strategy applied affected q_p and ROL production. In this way, q_p and maximal lipolytic activity were much lower for MLFB.

In MNLFB cultures where the methanol concentration was lower than 2 $g \cdot L^{-1}$ (Figure 3, Table 2) the lipolytic activity during the induction time was similar to MLFB cultivations (Figure 1, Table 1). A slow down on lipolytic activity production rate at the beginning of the MIP (Figure 3) appeared. Consequently, a decrease on specific production rates was observed (Figure 4). On the other hand, in MNLFB cultures with $MeOH_{sp}$ of 3, 5 and 10 $g \cdot L^{-1}$ ROL level time profiles were clearly different to the previous fed-batch cultivation series. High ROL production and also specific production rates (> 200 $U \cdot g^{-1} \cdot l^{-1}$) were attained over the whole induction period (Table 2). Nevertheless, the behaviour of the MNLFB cultivation at 10 $g \cdot L^{-1}$ methanol set point in terms of ROL production and q_P was similar to that observed in MLFB cultures only during the first 20 hours of the MIP.

3.4.3. Comparison of strategies

In other to discuss about the best production strategy, Figures 5 and 6 are presented. In Figure 5, the variation of the total product in terms of (PV) against induction time is displayed, that is related to productivities which can be estimated from the slope of the curves. In Figure 6, variation of the final titer is also shown in order to estimate q_P from the slope of the curves. As observed in Figure 5, the highest total ROL production, about 800 kU, was reached in MNLFB cultures at $MeOH_{sp}$ 3.0, and 10 $g \cdot L^{-1}$. However, induction times in these cultures were different, being the strategy at $MeOH_{sp}$ 3.0 $g \cdot L^{-1}$ the one that maximizes the productivity. Finally, as observed in Figure 6, the highest specific production rate was also achieved with MNLFB culture at $MeOH_{sp}$ 3.0 $g \cdot L^{-1}$. Similar results are shown in Table 2 for global volumetric and specific productivity for the whole fermentation time. In addition, similar productivities were achieved for the cultivations with $MeOH_{sp}$ of 5 and 10 $g \cdot L^{-1}$ while the highest levels of final titers were for the cultivation at $MeOH_{sp}$=10 $g \cdot L^{-1}$ (Table 2).

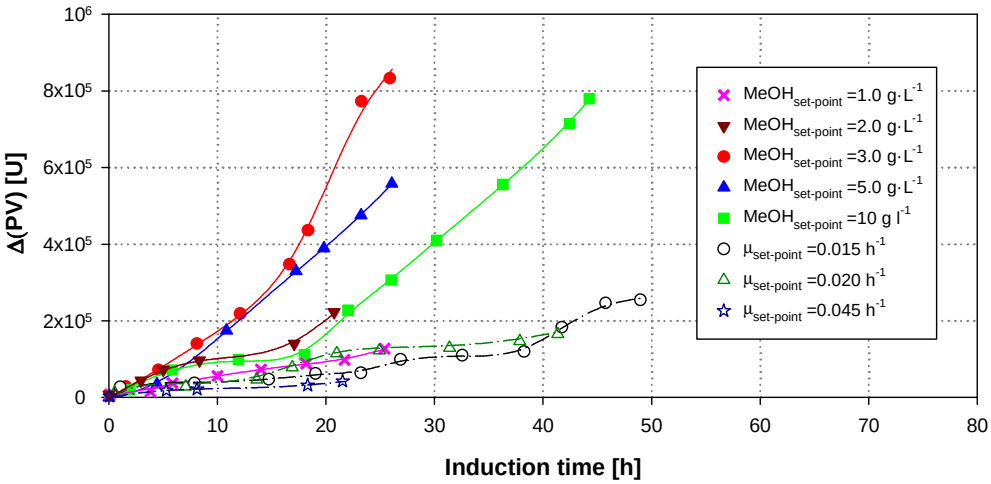

Figure 5. Variation of total amount of product through the induction time for the MLFB and MNLFB fermentations. Productivities can be estimated from the slope of the curves.

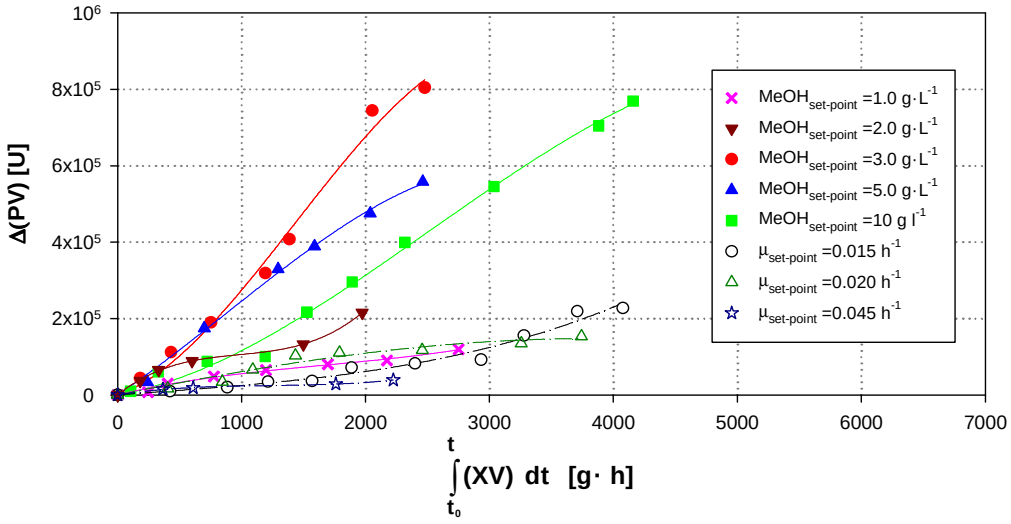

Figure 6. Variation of total lipolytic activity related to total biomass obtained for the fed-batch processes within the induction time. Specific production rates (q_P) can be estimated from the slope of the curves.

It is quite difficult to compare different operational strategies when the target protein is not the same, because the recombinant protein has influence in the specific rates. ROL has been the model protein of choice for several bioprocess engineering studies of *P. pastoris* performed by our research group, and, therefore, different operational strategies under different phenotypes and promoters have been implemented.

The results presented in this work are in good agreement with other previously reported studies on ROL production using Mut$^+$ strains. In a MNLFB culture operated under non-automatic control ($MeOH_{mean} \approx 2\ g \cdot L^{-1}$) showed analogous yield, maximal lipase activity, volumetric and specific productivity to those obtained at $MeOH_{sp} = 2\ g \cdot L^{-1}$ in this current work (Cos et al, 2005b). Conversely, some differences among them concerning specific rates were observed, probably due to distinct time dynamics related with the different

control approach applied. In addition, proteolysis observed under heuristic control strategy was not reproduced under MNLFB conditions.

Experiments made in different laboratories and bioreactors at the same methanol set-point (3 $g \cdot L^{-1}$) coupling MNLFB and MLFB strategies, where the latter were applied once oxygen limitation appeared, showed similar productivity (98 to 102 $U \cdot g^{-1} \cdot h^{-1}$), proving the reproducibility of the developed MNLFB strategy. Maximal lipolytic activity and $Y_{P/X}$ were higher because the bioprocess was longer. However, the mean specific rates presented in the present work were higher (i.e. $q_{P,mean}$ = 322 to 277 $U \cdot g^{-1} \cdot h^{-1}$) (Surribas et al., 2007).

MLFB strategies did not appear to be the best ones because the recombinant protein (ROL) production rate was drastically reduced after transition phase. This fact was also observed when ROL was produced under FLD promoter applying substrate limited fed-batch (SLFB) strategies (Resina et al., 2005). The implementation of non-limited substrate fed-batch strategies (SNLFB) overcame this problem. This production profile is rather similar to that observed using Mut$^+$ phenotype under the *AOX1* promoter. The behaviour of intracellular ROL activity could indicate an adaptation of cell's physiological state to ROL overexpression through all the cultivation time, leading to a down regulation of the ROL transcription levels (Resina et al., 2005), suggesting that a threshold level of intracellular ROL may exist, related to the triggering of the unfolded protein response (UPR) (Hohenblum et al., 2004).

3.5. Conclusions

MLFB operational strategy has been successfully applied in the production of recombinant proteins in the *P. pastoris* cell factory. However, independently of the phenotype and promoter used in ROL production, this strategy reached low production levels. A similar behaviour was observed in the MNLFB strategy at methanol concentration lower than 2 $g \cdot L^{-1}$. Although from an operational industrial point of view MLFB is an easy fed-batch strategy to be implemented, a low total lipolytic activity leads us to reject this approach.

MNLFB operation at constant methanol concentration higher than 2 $g \cdot L^{-1}$ is necessary to maximize ROL production. Proper monitoring and control of methanol concentration in MNLFB strategy revealed the importance to maintain methanol controlled during the bioprocess as key parameter to maximize protein production. It allowed not only increasing both yields and productivities, but also obtaining a more reproducible bioprocess and understanding the influence of methanol on ROL production.

Highest total ROL production, $Y_{P/X,}$ global volumetric and specific productivity, q_p and productivities were obtained for constant methanol concentration of 3 $g \cdot L^{-1}$. Lower methanol concentrations resulted in low product accumulation, whereas at the highest methanol concentration tested (10 $g \cdot L^{-1}$) methanol inhibition on growth was observed.

3.6. References

1. Amigo JM, Surribas A, Coello J, Montesinos JL, Maspoch S, Valero F. 2008. On-line parallel factor analysis. A step forward in the monitoring of bioprocess in real time. Chemom Intell Lab Syst. 92:44-52.
2. Albaek MO, Gernaey KV, Hansen MS, Stocks SM. 2011. Modeling enzyme production with *Aspergillus oryzae* in pilot scale vessels with different agitation, aeration, and agitator types. Biotechnol Bioeng 108:1828-1840.
3. Arnau C, Ramon R, Casas C, Valero F. 2010. Optimization of the heterologous production of a *Rhizopus oryzae* lipase in *Pichia pastoris* system using mixed substrates on controlled fed-batch bioprocess. Enzyme Microb Techl. 46:494–500.
4. 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.
5. Barrigón JM, Ramon R, Rocha I, Valero F, Ferreira EC, Montesinos JL. 2012. State and specific growth estimation in heterologous protein production by *Pichia pastoris*, AIChE J. 58:2967–2979.
6. Berdichevsky M, d'Anjou M, Mallem MR, Shaikh SS, Potgieter TI. 2011. Improved production of monoclonal antibodies through oxygen-limited cultivation of glycoengineered yeast. J Biotechnol. 155:217–224.
7. Borzani W. 2004. Calculation of fermentation parameters from the results of batch test taking account of the volume of biomass in the fermenting medium, Biotechnol Lett. 26:1953–1956.
8. Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol Rev. 24:45–66.
9. Cos O, Resina D, Ferrer P, Montesinos JL, Valero F. 2005a. Heterologous production of *Rhizopus oryzae* lipase in *Pichia pastoris* using the alcohol oxidase and formaldehyde dehydrogenase promoters in batch and fed-batch cultures. Biochem Eng J. 26: 86-94.

10. Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005b. 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.
11. Cos O, Ramon R, Montesinos JL, Valero F. 2006a. A simple model-based control for *Pichia pastoris* allows a more efficient heterologous protein production bioprocess. Biotechnol Bioeng. 95:145–154.
12. Cos O, Ramón R, Montesinos JL, Valero F. 2006b. 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.
13. Dabros M, Schuler MM, Marison IW. 2010. Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. Bioprocess Biosyst Eng. 33:1109–1118.
14. Guillén M, Benaiges MD, Valero F. 2011. Comparison of the biochemical properties of a recombinant lipase extract from *Rhizopus oryzae* expressed in *Pichia pastoris* with a native extract. Biochem Eng J. 54:117–123.
15. 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.
16. Invitrogen. 2014. *Pichia* Fermentation Process Guidelines, available online: http://tools.lifetechnologies.com/content/sfs/manuals/pichiaferm_prot.pdf (Accessed 06 December 2014).
17. Jahic M, Wallberg F, Bollok M, Garcia P, Enfors SO. 2003. Temperature limited fed-batch technique for control of proteolysis in *Pichia pastoris* bioreactor cultures. Microb Cell Fact. 2:6.
18. Jahic M, Veide A, Charoenrat T, Teeri T, Enfors SO. 2006. Process technology for production and recovery of heterologous proteins with *Pichia pastoris*. Biotechnol Prog. 22:1465–1473.
19. Jensch M, Simutis R, Eisbrenner G, Stückrath I, Lübbert A. 2006. Estimation of biomass concentrations in fermentation processes for recombinant protein production. Bioprocess Biosyst Eng. 29:19-27.

20. Katakura Y, Zhang W, Zhuang·g, Omasa T, Kishimoto M, Goto Y, Suga K. 1998. Effect of Methanol Concentration on the production of human β2-Glycoprotein I domain V by a recombinant *Pichia pastoris*: a simple system for the control of methanol concentration using a semiconductor gas sensor. J Ferment Bioeng. 86:482–487.
21. Khatri NK, Hoffmann F. 2006. Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant *Pichia pastoris*. Biotechnol Bioeng 93:871-879.
22. Lim HK, Choi SJ, Kim KY, Jung KH. 2003. Dissolved-oxygen-stat controlling two variables for methanol induction of rGuamerin in *Pichia pastoris* and its application to repeated fed-batch. Appl Microbiol Biotechnol. 62:342–348.
23. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast. 22:249-270.
24. Nunes PA, Pires-Cabral P, Guillén M, Valero F, Ferreira-Dias S. 2012. Batch operational stability of immobilized heterologous *Rhizopus oryzae* lipase during acidolysis of virgin olive oil with medium-chain fatty acids, Biochem Eng J. 67:265–268.
25. Pla IA, Damasceno LM, Vannelli T, Ritter G, Batt C a, Shuler ML. 2006. Evaluation of Mut^+ and Mut^S *Pichia pastoris* phenotypes for high level extracellular scFv expression under feedback control of the methanol concentration, Biotechnol Prog. 22:881–888.
26. Potgieter TI, Kersey SD, Mallem MR, Nylen AC, d'Anjou M. 2010. Antibody expression kinetics in glycoengineered *Pichia pastoris*. Biotechnol Bioeng. 106:918–927.
27. 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.
28. Ramon R, Feliu JX., Cos O, Montesinos JL, Berthet FX, Valero F. 2004. Improving the monitoring of methanol concentration during·high cell density fermentation of *Pichia pastoris*. Biotechnol Lett. 26:1447-1452.

29. 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.
30. 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.
31. Schenk J, Marison IW, von Stockar U. 2007. A simple method to monitor and control methanol feeding of *Pichia pastoris* fermentations using mid-IR spectroscopy. J Biotechnol. 128:344–353.
32. Sinha J, Plantz BA, Zhang W, Gouthro M, Schlegel V, Liu CP, Meagher MM. 2003. Improved production of recombinant ovine interferon-τ by Mu^{t+} strain of *Pichia pastoris* using an optimized methanol feed profile, Biotechnol Progr. 19 (2003) 794–802.
33. Sreekrishna K. 2010. *Pichia*, Optimization of Protein Expression. Ed. Michael C Flickinger. Encycl Ind Biotechnol.190:695–701.
34. Surribas A, Cos O, Montesinos JL, Valero F. 2003. On-line monitoring of the methanol concentration in *Pichia pastoris* cultures producing an heterologous lipase by sequential injection analysis. Biotechnol Lett. 25:1795–1800.
35. Surribas A, Stahn R, Montesinos JL, Enfors SO, Valero F, Jahic M. 2007. Production of a *Rhizopus oryzae* lipase from *Pichia pastoris* using alternative operational strategies. J Biotechnol. 130:291–299.
36. Tecelão C, Guillén M, Valero F, Ferreira-Dias S. 2012. Immobilized heterologous *Rhizopus oryzae* lipase: A feasible biocatalyst for the production of human milk fat substitutes. Biochem Eng J. 67:104–110
37. Trentmann O, Khatri NK, F. Hoffmann H. 2004. Reduced oxygen supply increases process stability and product yield with recombinant *Pichia pastoris*. Biotechnol Progr. 20:1766–1775.
38. 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.

39. Zhang W, Bevins M a, Plantz B a, Smith L a, Meagher MM. 2000. Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. Biotechnol Bioeng. 70:1–8.
40. Zhang W, Sinha J, Smith LA, Inan M, Meagher MM. 2005. Maximization of production of secreted recombinant proteins in *Pichia pastoris* fed-batch fermentation. Biotechnol Prog. 21:386–393.
41. Zhang W, Inan M, Meagher MM. 2007. Rational design and optimization of fed-batch and continuous fermentations. Methods Mol Biol. 389:43–64.

Chapter 4:

A macrokinetic model-based comparative meta-analysis of recombinant protein production by *Pichia pastoris* under *AOX1* promoter.

Published as: 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 (accepted).

---

4. A macrokinetic model-based comparative meta-analysis of recombinant protein production by *Pichia pastoris* under *AOX1* promoter.

4.0. Abstract

An unstructured macrokinetic model for heterologous production of *Rhizopus oryzae* lipase (ROL) by a *Pichia pastoris* P_{AOX1} based system was developed. Mean specific rates and state variables for various fed-batch cultures, under methanol limited and non-limited conditions were used for modeling. The most representative kinetic functions (viz., Monod, Haldane, Pirt and Luedeking-Piret) were used for the specific rates of cell growth (μ), substrate consumption (q_S) and product formation (q_P). The performance of two different process models was assessed via simulation tests, using an overall validation procedure. The best model describes the cellular growth by a non-monotonic substrate function, substrate uptake via Pirt's equation and product formation through a Luedeking-Piret equation. A comparative meta-analysis of heterologous protein production of various target proteins by *Pichia pastoris* under *AOX1* promoter was conducted and a general strategy for improving protein production from process kinetics was developed as a key to bioprocess optimization.

4.1. Introduction

Pichia pastoris is recognized as one of the most efficient cell factories for the production of recombinant proteins (Macauley-Patrick et al. 2005; Potvin et al. 2012). In fact, more than five hundred proteins have to date been expressed by this system (Cregg, 2014). Also, the system has been used in several protein production platforms for structural genomics programs (Yokohama et al. 2003), especially with alcohol oxidase 1 promoter (P_{AOX1}), which is currently the most widely used. In addition, the system has been used in combination with methanol as inducer substrate (Cereghino and Cregg, 2000). There are three types of *P. pastoris* host strains available regarding to their ability to utilize methanol: the wild-type or methanol utilization plus phenotype (Mut^+), and the strains resulting from deletions in the *AOX1* gene, methanol utilization slow (Mut^s), or both *AOX* genes, methanol utilization minus (Mut^-). Typically, *Pichia* bioprocesses under P_{AOX1} involve a glycerol batch phase (GBP), a transition phase (TP) and a methanol induction phase (MIP). The primary goal of GBP is to rapidly produce biomass prior to induction by methanol. A number of TP strategies use glycerol and methanol (Zhang et al. 2000; Cos et al. 2005a) to increase cell density under glycerol limiting conditions and the de-repression of P_{AOX1} to adapt cell metabolism to induction of the promoter by methanol. In MIP, methanol acts as both a carbon source and the inducer substrate. The particular methanol feeding strategy used in MIP is one of the most important factors with a view to maximizing production of heterologous proteins (Cos et al. 2006a; Cos et al. 2006b).

Heterologous protein expression systems can usually be optimized by using improved strains in combination with an optimal engineering strategy avoiding highly complex

media in order to reduce costs and facilitate downstream processes (Sreekrishna, 2010; Potvin et al. 2012). Modeling and control tools are especially useful for *Pichia* bioprocess engineering. A thorough knowledge of the process kinetics is essential to understand the behavior of this microorganism in terms of growth, uptake and production rates. When the cell population is assumed to be homogeneous, the "average cell" approximation to construct an appropriate non-segregated model, whether structured or unstructured (Bailey, 1998; Villadsen et al. 2011) can be considered. If the aim is to understand some crucial aspect of a bioprocess, then the model should include physiological and metabolic considerations (e.g., cell diversity and morphology, substrate and product transport through cells, protein synthesis, and distribution of substrates and products between phases) (Montesinos et al. 1995; Nielsen, 1996; Duboc and Stockar, 2000; Pérez et al. 2005). Accounting for complex interactions between the extracellular environment and intracellular enzymes and metabolites entails using complex models. Genome-scale stoichiometric models currently provide the best approximation to the metabolic capabilities of cells (Feist et al. 2007; Sohn et al. 2010). Combinations of metabolic and mechanistic models can lead to a more comprehensive knowledge of cellular organization (Carneiro et al. 2013). However, unstructured models provide quite acceptable results in many situations such as when the balanced growth condition is fulfilled (Heijnen et al. 1979; Chae et al. 2000; Albaek et al. 2011) or with estimation, control and optimization applications of little mathematical complexity for biotechnological processes (Barrigón et al. 2012; Ren and Yuan, 2005; Zhang et al. 2005; Maurer et al. 2006). Thus, using a bioreactor and a macrokinetic model accounting for macro reactions in combination can be an effective compromise between simplicity and

the comprehensive process description typically provided by highly complex alternatives such as cybernetic (Ramakrishna, 1996), genetically structured (Lee and Bailey, 1984) and metabolic models (Heyland et al. 2011; Soons et al. 2011; Jordà et al. 2012). In any case, models should be kept as simple as possible and modeling strategies only be made more elaborate when and as needed (Villadsen et al. 2011; Levenspiel, 2002).

Despite the large number of recombinant proteins reported for the *Pichia* cell factory, growth, consumption and production kinetics in this yeast have scarcely been modeled to date (Zhang et al. 2000; Ren et al. 2003; Çelik et al. 2009). The selected target protein in this work was a recombinant *Rhizopus oryzae* lipase (ROL). Process development for ROL production has been widely studied from a bioprocess engineering point of view under different phenotypes, promoters, co-expressing helpers proteins and strains with permeabilized membrane (Cos et al. 2005b; Cos et al. 2006a; Cos et al. 2006b; Resina et al. 2007; Marx et al. 2006; Surribas et al. 2007; Resina et al. 2009; Arnau et al. 2011). Recently, ROL production in *Pichia pastoris* Mut^+ phenotype under two different operational strategies, methanol limited (MLFB) and methanol non-limited fed-batch (MNLFB) cultures, have been intensively analyzed (Barrigón et al. 2013). These most commonly applied control strategies allow maintaining rather constant the key specific rates for cell growth (μ) and substrate uptake (q_S) through the quasi-steady state hypothesis for the substrate while seeking to maximize target protein production.

In this work, a macrokinetic model for fed-batch processes is developed and then, a comparative meta-analysis of heterologous protein production by *Pichia pastoris* under *AOX1* promoter, considering different target proteins, is presented.

4.2. Materials and methods

4.2.1. Strains

The wild type *P. pastoris X-33* strain containing the pPICZαAROL vector was used for heterologous expression of *R. oryzae* lipase (ROL) under control of the *AOX1* promoter (Cos et al. 2005b). Pre-inocula for bioreactor cultures were grown as described elsewhere (Barrigón et al. 2013).

4.2.2. Bioreactor cultivation

Pre-inocula for bioreactor cultures were grown for 24 h in 1 L baffled shake flasks at 30 ºC, 150 rpm, in YPD medium containing 10 g yeast extract, 20 g peptone, 20 g glucose and 1 mL zeocin per litre of distilled water. Shake flasks contained 200 mL of YPD medium. The culture was centrifuged at 4500 x g, the harvested cells were re-suspended in bioreactor culture medium and used to inoculate a 5 L Biostat ED bioreactor (Braun Biotech, Melsungen, Germany).

Cultivation tests were conducted in a 5L Braun Biostat ED bioreactor (Braun Biotech, Melsungen, Germany). Media and cultivation conditions have been previously described in detail (Barrigón et al. 2013). The fermentation process involved the three above-described phases: GBP, TP and MIP (Cos et al. 2005a). The MIP phase was conducted in two different ways, namely: MLFB and MNLFB as described elsewhere (Barrigón et al. 2013). A total of three MLFB cultures (0.015, 0.020 and 0.045 h^{-1}) and five MNLFB cultures (1.0, 3.0, 5.0, 8.0 and 10.0 $g \cdot L^{-1}$) were performed.

4.2.3. Analyses

Biomass concentration was quantified as dry cell weight (DCW) per litre of culture broth (Barrigón et al. 2013). Methanol and glycerol were determined by HPLC as described elsewhere (Arnau et al. 2011). Extracellular lipolytic activity was monitored spectrophotometrically (Resina et al. 2004). The methanol concentration was monitored online by using a sensor from Raven Biotech (Vancouver, BC, Canada) immersed in the culture broths (Arnau et al. 2011).

4.3. Theory and calculation

4.3.1. Mass balance and stoichiometric equations

The oxidative uptake of substrates (glycerol or methanol) to form biomass and products can be described by a single overall reaction, a so-called Black Box model, which is a simplification of all the biochemical reactions involved:

$$k_1 S + k_2 O_2 \rightarrow X + k_3 P + k_4 CO_2 \quad (1)$$

where *S* denotes one single limiting substrate (glycerol or methanol as the carbon and energy source), O_2 oxygen, *X* biomass, *P* product and CO_2 carbon dioxide; and k^*_i values are stoichiometric coefficients that can also be called overall i-biomass yields ($Y^*_{i/X}$).

For an ideal stirred tank-reactor, considering conversion rates of biomass formation, substrate uptake and product formation allows the following mass balance equations for a fed-batch cultivation process to be formulated:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ P \\ O_2 \\ CO_2 \end{bmatrix} = \begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ P \\ O_2 \\ CO_2 \end{bmatrix} + \begin{bmatrix} 0 \\ DS_0 \\ 0 \\ OTR \\ -CTR \end{bmatrix} \quad (2)$$

where μ is the specific growth rate [h^{-1}], q_S the specific substrate uptake rate [$g \cdot g^{-1} \cdot h^{-1}$], q_P the specific production rate [$U \cdot g^{-1} \cdot h^{-1}$], q_{O2} the specific oxygen uptake rate [$mol \cdot g^{-1} \cdot h^{-1}$], q_{CO2} the specific carbon dioxide production rate [$mol \cdot g^{-1} \cdot h^{-1}$], *D* the dilution rate (F/V), *F* the substrate feeding rate [$L \cdot h^{-1}$], *V* the volume of broth in the reactor [L], S_0 the substrate feeding concentration [$g \cdot L^{-1}$], *OTR* the oxygen transfer rate [$mol \cdot L^{-1} \cdot h^{-1}$] and *CTR* the carbon dioxide transfer rate [$mol \cdot L^{-1} \cdot h^{-1}$]. Substrate and product concentrations were referred to the whole medium, including biomass volume (Barrigón et al. 2013).

Some assumptions are made in Eq (2), among the most important, stoichiometric (yield) coefficients remain constant, all cellular components are pooled into one single representative biomass with concentration X, biomass composition stays constant and morphology does not change, all other rates than substrate uptake can be derived from total mass balances and rate expression for the limiting substrate (Villadsen, 2011). For more details see the chapter 7.

4.3.2. Kinetic models

The most widely used model for cell growth kinetics is a monotonically increasing function (viz., the well-known Monod kinetic model). It is an unstructured, nonsegregated model and assumes the growth to be independent of the biomass composition as any other unstructured model. The non-monotonically increasing function for cell growth called the "Haldane model" has also been used with yeasts to include inhibition by substrate (e.g., for *P. pastoris* growing on methanol as only carbon source) (Kobayashi et al. 2000). These two unstructured models assume that there is only one limiting substrate. Modeling of the q_S is commonly based on Pirt's maintenance energy model, which assumes a linear dependence between q_S and μ (Villadsen et al. 2011). The substrate consumption for maintenance is considered independently of the growth process. A large number of kinetic models have been used to describe the formation of heterologous proteins. Frequently, product formation is related to biomass production or substrate uptake in order to reduce the number of variables to be measured (Liu et al. 2011). Probably, the best known model for production kinetics is the linear relation between q_P and μ proposed by Luedeking-Piret. In this work, substrate monotonically

increasing function (Eq. 3), non-monotonically increasing function (Eq. 4) and linear law (Eq. 5) have been used:

$$q_i = \frac{q_{max,i}\, S}{K_{s,i} + S} \quad (3)$$

$$q_i = \frac{q_{max,i}\, S}{K_{s,i} + S + \frac{S^2}{K_{I,i}}} \quad (4)$$

$$q_{i2} = Y_{i2/i1} q_{i1} + m_{i2} \quad (5)$$

where q_i is the i-specific rate, $q_{max,i}$ the maximum value of the i-specific rate, $K_{S,i}$ the substrate monotonic or non-monotonic increasing model constant, $K_{I,i}$ the substrate inhibition non-monotonic increasing model constant, $Y_{i2/i1}$ the yield coefficient of component i_2 to i_1 and $m_{i,2}$ the maintenance coefficient of component i_2. With the use of linear laws the overall yield coefficients ($Y^*_{i2/i1}$) are not constant and depend on the *i1*-specific rate. Discrete specific rates (μ_t, $q_{S,t}$ and $q_{P,t}$) for each off-line data [$(XV)_t$, $(SV)_t$ and $(PV)_t$] were determined by using the respective mass balance equations. Averaged specific rates during the induction phase in all fermentations were estimated by linear regression (Barrigón et al. 2013).

4.3.3. Performance indicators

To evaluate "goodness-of-estimation", mean relative error (*MRE*) was calculated as a single metrics for *XV*, *SV* and *PV* (Eq. 6) and also an overall performance indicator (*OPI*), which is formed by the *MRE* of each simulated state variable (Eq. 7).

$$MRE = \frac{1}{n}\sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{y_i} \quad (6)$$

$$OPI = \frac{1}{N}\sum_{j=1}^{N} w_X\, MRE_{X,j} + w_S\, MRE_{S,j} + w_P\, MRE_{P,j} \quad (7)$$

where n is the number of data points for an individual experiment, $\widehat{y}_i$ the i^{th} estimated value, y_i the corresponding i^{th} actual value for the bioprocess and N the number of experiments, and w_k denote the weighting factors for the single metrics $w_X = 0.5$, $w_S = 0.1$ and $w_P = 0.4$.

In order to solve dynamic equations the MATLAB's ordinary differential equation solver function ode45s (Matlab R2009a, The MathworksInc., Natik, USA) was used.

4.4. Results and discussion

The kinetic model for ROL production by *P. pastoris* Mut$^+$ phenotype was experimentally established from the results for MLFB and MNLFB cultures. Eight fed-batch cultures with different set-points on μ (0.015, 0.020 and 0.045 h^{-1}) and S (1.0, 3.0, 5.0, 8.0 and 10.0 $g \cdot L^{-1}$) were used. These control strategies allow maintaining rather constant the key specific rates for cell growth (μ) and substrate uptake (q_S) through the quasi-steady state hypothesis for the substrate. Range of set-points, were selected in order to more easily observe the effect of the inducer–substrate on the specific rates. In Figure 1 the two standard strategies are shown.

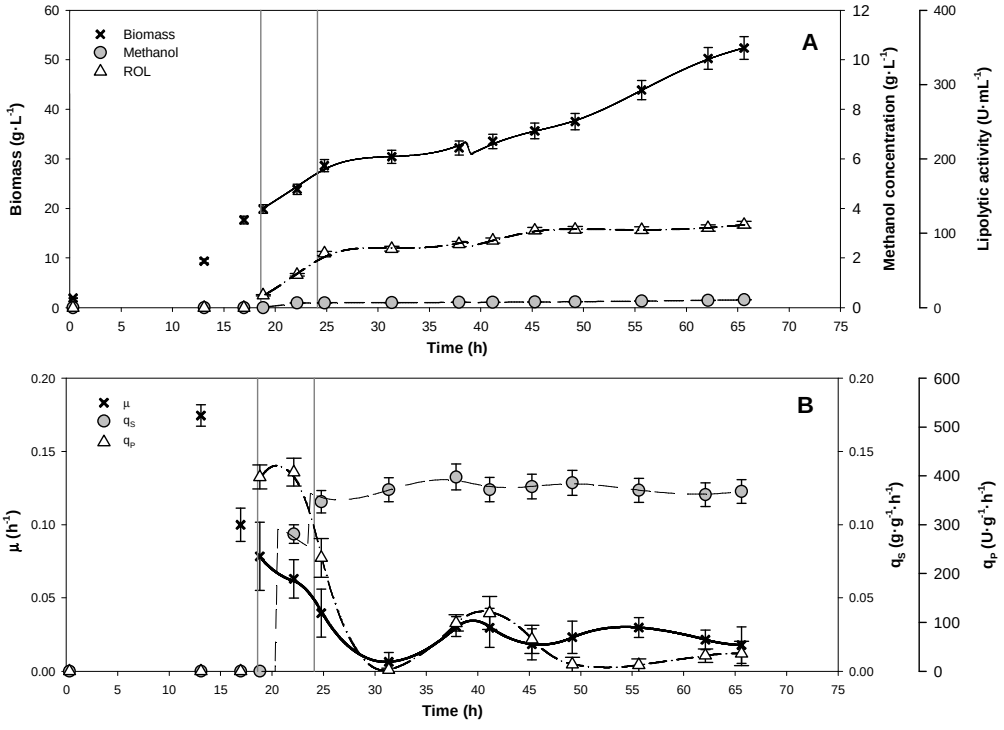

Figure 1. Time course of state variables and specific rates. MLFB-$\mu_{set\text{-}point}$ = 0.02 h^{-1} culture: (**A**) biomass, methanol and lipolytic activity; (**B**) μ, q_S and q_P. MNLFB-$MeOH_{set\text{-}point}$ = 10.0 $g \cdot L^{-1}$ culture; (**C**) biomass, methanol and lipolytic activity; (**D**) μ, q_S and q_P. Error bars represent standard deviations (SD).

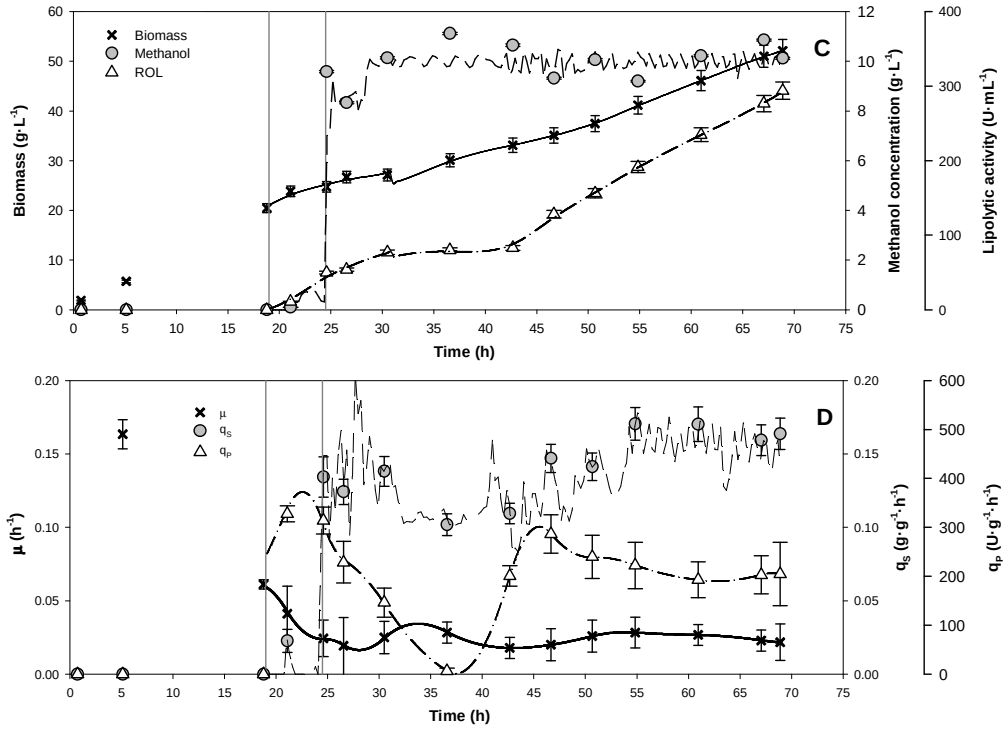

Figure 1. Continued

4.4.1. Kinetic model for ROL production

Figure 2 shows the monotonically increasing and non-monotonically increasing substrate models used (Eq. 3 and 4). Kinetic parameters were obtained by least squares minimization. Mean specific rates exhibited a monotonic increase to a maximum value not corresponding to the highest concentration of methanol; this was followed by a decrease in the specific rates with increasing concentration of substrate, which suggested that a monotonic model might not be the best choice. Parameter values, coefficients of variation (CV) and determination (R^2), and p-values of the analysis of variance are presented in Table I. The critical substrate concentration ($S_{crit,i}$), that is, the concentration leading to the maximum specific rate, was also calculated for a non-monotonic kinetic

model. As can be seen from Table 1, the resulting statistics confirmed that a non-monotonic model afforded better fitting than a monotonic function. The maxima for each specific rate function occurred at different methanol concentrations (viz., about 2 $g \cdot L^{-1}$ for $S_{crit,X}$ and $S_{crit,S}$, and approximately $3g \cdot L^{-1}$ for $S_{crit,P}$).

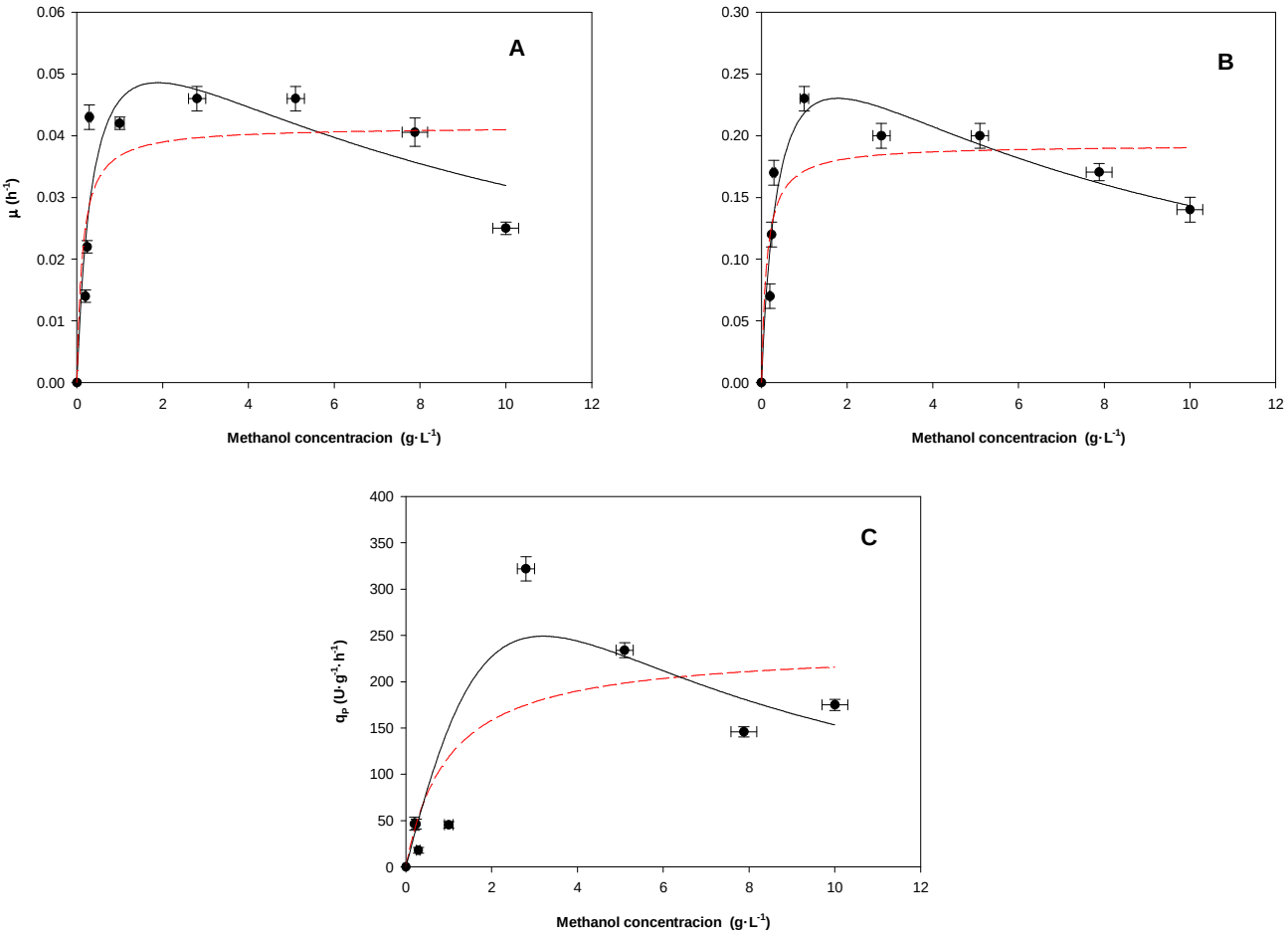

Figure 2. Relationship between specific rates (μ, q_S, q_P) and the substrate (MeOH) for ROL production by *P. pastoris* X-33 strain with Mut^+ phenotype. (**A)** Specific growth rate *vs* residual methanol concentration. (**B**) Specific substrate uptake rate (q_S) *vs* residual methanol concentration. (**C**) Specific product formation rate (q_P) *vs* residual methanol concentration. Dots represent experimental mean specific rates, the dashed line the substrate monotonically increasing model and the solid line the substrate non-monotonically increasing model. Error bars indicate standard error (SE).

Table 1. Macrokinetic models for ROL production by *P. pastoris* X-33 strain with Mut^+ phenotype.

Summary of model parameters and correlation factors between specific rates (μ, q_S and q_P) and the substrate (MeOH) concentration.

	***S* non-monotonically increasing function**					***S* monotonically increasing function**				
	Parameter	**Units**	**Value**	**CV %**	**Statistic**	**Parameter**	**Units**	**Value**	**CV %**	**Statistic**
q_X (μ) [h^{-1}]	μ_{max}	h^{-1}	0.069	33.3	R^2 (0.81)	μ_{max}	h^{-1}	0.042	12.8	R^2 (0.67)
	$K_{S,X}$	$g \cdot L^{-1}$	0.40	69.1						
	$K_{I,X}$	$g \cdot L^{-1}$	8.85	84.6	p (0.006)	$K_{S,X}$	$g \cdot L^{-1}$	0.13	76.1	p (0.007)
	$S_{crit,X}$	$g \cdot L^{-1}$	1.9	-						
q_S [$g \cdot g^{-1} \cdot h^{-1}$]	$q_{max,S}$	$g \cdot g^{-1} \cdot h^{-1}$	0.34	21.6	R^2 (0.92)	$q_{max,S}$	$g \cdot g^{-1} \cdot h^{-1}$	0.20	12.2	R^2 (0.76)
	$K_{S,S}$	$g \cdot L^{-1}$	0.42	43.3						
	$K_{I,S}$	$g \cdot L^{-1}$	7.57	60.4	p (0.001)	$K_{S,S}$	$g \cdot L^{-1}$	0.14	66.3	p (0.003)
	$S_{crit,S}$	$g \cdot L^{-1}$	1.7	-						
q_P [$U \cdot g^{-1} \cdot h^{-1}$]	$q_{max,P}$	$U \cdot g^{-1} \cdot h^{-1}$	1844	> 100	R^2 (0.80)	$q_{max,P}$	$U \cdot g^{-1} \cdot h^{-1}$	237	26.5	R^2 (0.64)
	$K_{S,P}$	$g \cdot L^{-1}$	10.2	> 100						
	$K_{I,P}$	$g \cdot L^{-1}$	1.0	> 100	p (0.008)	$K_{S,P}$	$g \cdot L^{-1}$	1.0	> 100	p (0.010)
	$S_{crit,P}$	$g \cdot L^{-1}$	3.2	-						

The study was completed by calculating correlations between specific rates (q_S and q_P *vs* μ), using linear functions in Pirt's model for q_S and in the Luedeking-Piret model for q_P (Eq. 5). As can be seen from Figure 3A, the specific uptake and growth rate were linearly related. Pirt's maintenance energy model proved as good as the non-monotonically increasing substrate model for q_S; in fact, both explained changes in q_S with a similar confidence level (Tables1 and 2).

Table 2. Linear kinetic models for ROL production by *P. pastoris* X-33 strain with Mut$^+$ phenotype. Summary of model parameters and correlation factors between specific rate for substrate uptake or product formation (q_S, q_P) and specific growth rate (μ).

	$\mathbf{q_S}$ $[g\cdot g^{-1}\cdot h^{-1}]$ (i = s)			$\mathbf{q_P}$ $[U\cdot g^{-1}\cdot h^{-1}]$ (i = p)		
Parameter	**Units**	**Value**	**CV %**	**Units**	**Value**	**CV %**
$Y_{i,X}$	$g\cdot g^{-1}$	4.21	11.6	$U\cdot g^{-1}$	4567*	17.3
m_i	$g\cdot g^{-1}\cdot h^{-1}$	0.0142	> 100	$U\cdot g^{-1}\cdot h^{-1}$	39.3	20.7
Statistic		**R^2**	**p value**		**R^2**	**p value**
		0.92	< 0.001		0.92	0.001

*Only applies when $S > S_{crit,X}$; zero when $S \leq S_{crit,X}$

Figure 3B illustrates the correlation between a Luedeking-Piret model for the specific ROL production rate and specific growth rate. However, two different submodels (q_P-constant and μ-linear) were required to accurately explain the variation of q_P with μ (Table2). The Luedeking-Piret function may be linear enough for a limited range of μ values in some heterologous recombinant protein production systems (de Hollander, 1993). The limit condition for switching between functions was stated for the critical substrate value ($S_{crit,X}$). In fed-batch cultures carried out below $S_{crit,X}$ (1.9 $g\cdot L^{-1}$), q_P is virtually independent of μ, so the intrinsic product to biomass yield ($Y_{P/X}$) ratio is

assumed to be zero. Thus, q_P for rHSA production was previously found to remain essentially constant with *P. pastoris* Mut$^+$ phenotype at low *S* (Ohya et al. 2005); above $S_{crit,X}$, however, q_P was linearly dependent on μ.

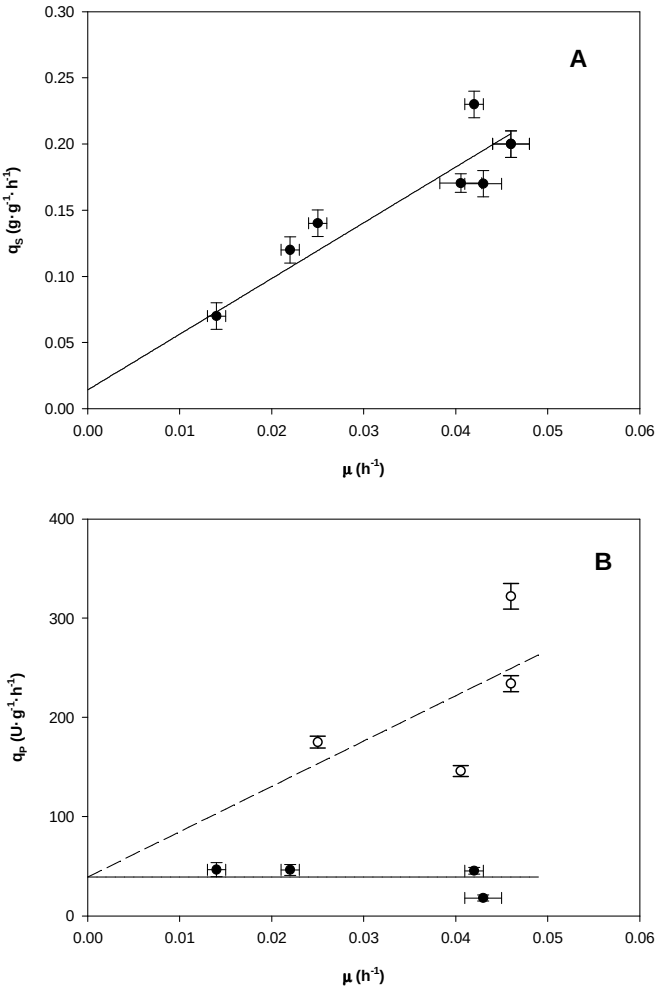

Figure 3. Linear relationships between specific rates of ROL production by *P. pastoris* X-33 strain with Mut$^+$ phenotype. (**A**) Specific substrate uptake rate (q_S) *vs* specific growth rate (μ). Dots represent experimental mean values. (**B**) Specific product formation rate (q_P) vs specific growth rate (μ). Black dots and white dots represent experimental mean specific rates for MeOH $\leq S_{crit,X}$ and MeOH > $S_{crit,X}$ respectively. Lines indicate the best fit to the Pirt model **(A)** or Luedeking-Piret model (solid line with MeOH $\leq S_{crit,X}$ and dashed line with MeOH > $S_{crit,X}$) **(B)**. Error bars corresponds to standard error (SE).

4.4.2. Model validation

Two alternative macrokinetic models were tested, namely: one assuming non-monotonically increasing kinetics for specific growth, uptake and production (Model A, Figure 2); and the other assuming non-monotonically increasing kinetics for μ (Model B, Figure 2A) and μ-linear functions for q_S (Figure 3A) and q_P (Figure 3B).

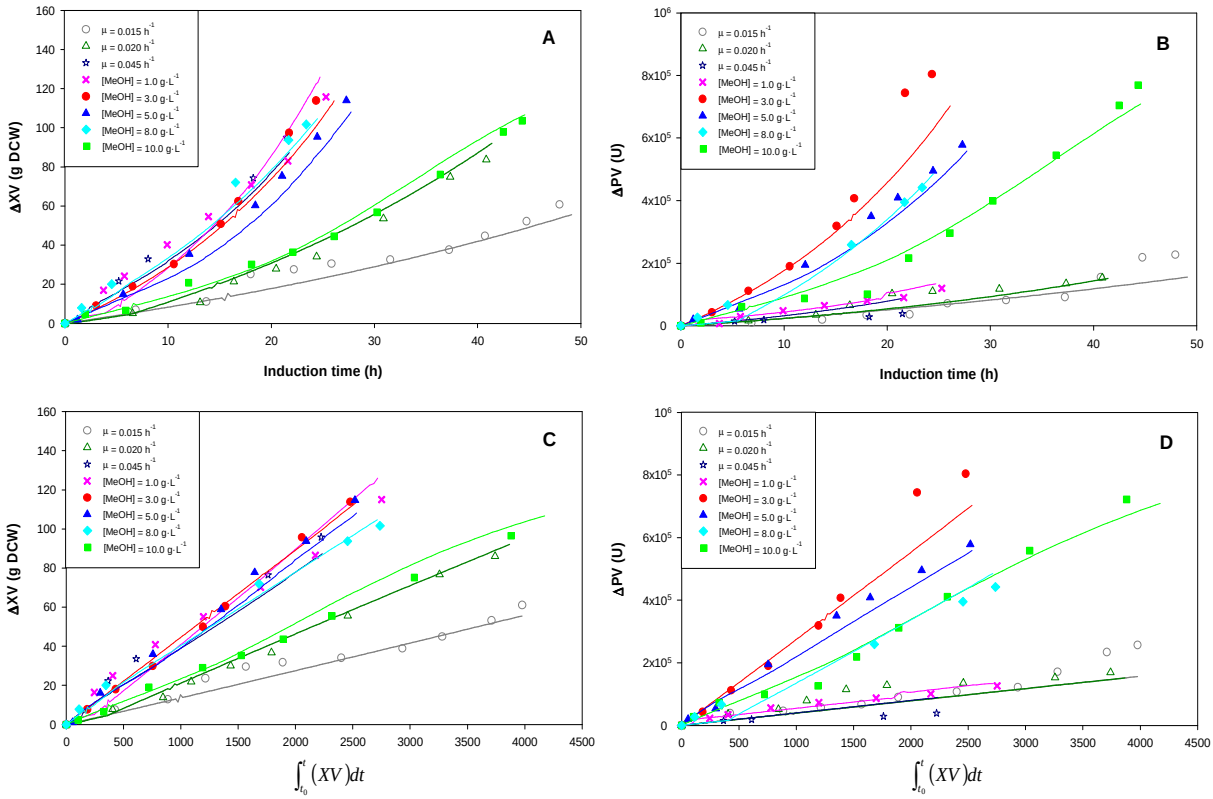

Figure 4. Model validation by simulation of global state variables (*XV* and *PV*) over the induction time. Time course of **(A)** total biomass and **(B)** total lipolytic activity. **(C)** Total biomass and **(D)** total lipolytic activity variations on cumulative total biomass. Set-points are indicated by μ values for MLFB cultures and methanol concentrations for MNLFB cultures.

Models A and B provided similarly accurate estimated biomass profiles, with a mean relative deviation of 5% (Table .3, substrate estimates not shown). Substrate simulations were subject to large errors with both models, especially in MLFB cultures (*MRE*

>100%) owing to the low *S*. Model B provided more accurate product profiles, with a mean relative deviation of 12%. Overall performance was clearly worse with Model A (*OPI*=17%) than with Model B (*OPI*=12%). Although Model A fitted mean specific rates quite well (Figure 2 and Table I), the simulated results for global state variables exposed a higher than expected deviation. Generally, prediction errors were lower (MRE_{min}) for biomass in MLFB cultures and higher (MRE_{max}) for product in MNLFB, thus according to the observed poorer fitting of the kinetic models at high substrate concentration. In Figure 4 the simulated and experimental values of the global state variables (*XV* and *PV*) as obtained with Model B are shown. Overall biomass and ROL productivities in the induction phase were estimated from the mean slope of the curve in Figure 4A and 4B, respectively. So, biomass productivity ranged from 1.1 to 5.1 g·DCW h^{-1} and ROL productivity from $3.2 \cdot 10^3$ to $2.7 \cdot 10^4$ U·h^{-1}. Mean specific rates for μ and q_p were also calculated from the slope of experimental curves in Figure 4C and 4D, now according to Eq. (8) and Eq. (9), respectively.

$$\int_{(XV)_0}^{(XV)} d(XV) = \mu_{mean} \int_{t_0}^{t} (XV)\, dt \quad (8)$$

$$\int_{(PV)_0}^{(PV)} d(PV) = q_{P\,mean} \int_{t_0}^{t} (XV)\, dt \quad (9)$$

Specific rates varied over the range 0.015-0.045 h^{-1} for μ and 40-280 U·g^{-1}·h^{-1} for q_p.

Table 3. Model validation of global state variables (*XV* and *PV*) for time course (*t*) and cumulative biomass ($\int_{t_0}^{t}(XV)dt$). MRE_{mean}, MRE_{min} and MRE_{max} are, respectively, the mean, lower and higher *MRE* obtained for all the experimental set.

Model	Statistic	XV vs t	XV vs $\int_{t_0}^{t}(XV)dt$	PV vs t	PV vs $\int_{t_0}^{t}(XV)dt$
Model A S **non-monotonically increasing functions for** μ, q_S, **and** q_P	MRE_{mean}	0.05	0.05	0.24	0.24
	MRE_{min}	0.03	0.03	0.07	0.07
	MRE_{max}	0.09	0.09	0.55	0.66
Model B S **non-monotonically increasing function for** μ μ**-linear function for** q_S **and** q_P	MRE_{mean}	0.05	0.05	0.12	0.11
	MRE_{min}	0.03	0.03	0.06	0.07
	MRE_{max}	0.07	0.07	0.20	0.17

In conclusion, although the estimated values provided by Models A and B were consistent with their experimental counterparts, the latter model performed somewhat better. In fact, Model B was highly accurate for its reduced mathematical complexity and is thus a promising choice for estimation, control and optimization in future tests.

4.4.3. Comparative analysis

Comparing the productivity of target proteins requires considering bioprocess engineering aspects including culture medium and physical variables such as temperature and pH, oxygen supply and pO_2 level. Also, operational strategies such us fed-batch operation at variable methanol feeding rates with or without methanol control, and the use of mixed substrates, can significantly affect specific rates, and hence production levels and the specific productivity of a process. Besides, genetic and physiological aspects affect protein expression levels for a given target protein (Brocca et al. 1998; Hohenblum et al. 2004; Boettner et al. 2007; Resina et al. 2007).

Table 4. Summary of kinetic models and process characteristics for various heterologous proteins produced by *P. pastoris* under *AOX* promoter.

Reference	*Strain*	*Protein*	***μ-model***			***q_s-model***		***q_p-model***		***Specific Productivity ($mg \cdot g^{-1} \cdot h^{-1}$)***	*Medium*
			Type	**Critical values**	**Operational range**	**Type**	**Parameters**	**Type**	**Operational range**		
This work	X-33 Mut$^+$	ROL	*S* non-monotonically increasing	$S_{crit} = 1.93$ $\mu_{crit} = 0.049$	–	μ-linear	$Y_{s/x} = 4.21$ $m_s = 0.0142$	q_p-constant μ-linear	$S < S_{crit} \approx 2.0$ $S > S_{crit} \approx 2.0$	0.05	BSM
Zhang et al., 2000	GS115 Mut$^+$	BoNT/ A(Hc)	*S* non-monotonically increasing	$S_{crit} = 3.65$ $\mu_{crit} = 0.08$	–	μ-linear	$\underline{Y_{s/x} = 3.53}$ $\underline{m_s = 0.0298}$	*S* non-monotonically increasing	$S_{crit} = 2.1$	<u>0.14</u>	BSM
Zhang et al., 2005	GS115 Mut$^+$	Coffee bean α-Gal	–	–	–	μ-linear	$\underline{Y_{s/x} = 3.05}$ $\underline{m_s = 0.016}$	μ non-monotonically increasing	$\mu_{crit} \approx 0.027$	<u>0.13</u>	BSM
Sinha et al., 2003	GS115 Mut$^+$	rOvIFN-τ	–	---	–	μ-linear	$\underline{Y_{s/x} = 3.15}$ $\underline{m_s = 0.0231}$	μ non-monotonically increasing	$\mu_{crit} \approx 0.025$	<u>0.03</u>	FM22
Zhang et al., 2004	X-33 Mut$^+$	rOvIFN-τ	*S* non-monotonically increasing	(Zhang et al., 2000)	–	μ-linear	$\underline{Y_{s/x} = 3.17}$ $\underline{m_s = 0.0126}$ (*)	μ non-monotonically increasing	$0.01 < \mu < 0.09$	0.03	FM22
Schenk et al., 2007&2008	GS115 Mut$^+$	recGAvi	*S* non-monotonically increasing	$1 < S_{crit} < 6$ $\mu_{crit} = 0.139$	–	–	$\underline{Y^*_{s/x} = 2.85}$	μ monotonically increasing	$\mu < 0.14$	0.01	BSM
Jacobs et al., 2010	GlycoSwitch -Man5 Mut$^+$	GM-CSF	–	$S_{crit} = 2.0$ $\mu_{crit} = 0.063$	$0.015 < \mu < 0.063$	μ-linear	$\underline{Y^*_{s/x} = 3.53}$	μ-linear decreasing q_p-constant	$0.015 \leq \mu \leq 0.033$ $0.033 \leq \mu \leq 0.063$ (*)	<u>0.16</u>	FM22
Cunha et al., 2004	GS115 Mut$^+$	scFv	–	–	–	–	–	q_S-linear decreasing q_p-constant	$\underline{0.011 \leq q_S \leq 0.026}$ $\underline{0.026 \leq q_S \leq 0.055}$	<u>< 0.01</u>	BMGY
Khatri and Hoffmann, 2006	GS115 BA11	scFv	–	–	$3 \leq S \leq 30$	μ-linear	$\underline{Y_{s/x} = 4.17}$ $\underline{m_s = 0.042}$ (*)	–	–	<u>0.04</u>	BSM
Yamawaki et al., 2007	GS115 Mut$^+$	scFv	–	–	$0.007 \leq \mu \leq 0.054$ $S_{set\text{-}point} = 3.9$	μ-linear	$Y_{s/x} \approx 3.33$ $m_s \approx 0.013$ (*)	μ monotonically increasing	$\mu_{crit} \approx 0.02$ (*)	0.04	BSM
Potgieter et al., 2010	YGLY4140	IgG1	–	–	$\mu < 0.03$	μ-linear	$\underline{Y_{s/x} = 3.09}$ $\underline{m_s = 0.0168}$	μ non-monotonically increasing	$\mu_{crit} \approx 0.0115$	<u>0.08</u>	BMGY

Reference	*Strain*	*Protein*	*μ-model*			*q_s-model*		*q_p-model*		*Specific Productivity ($mg \cdot g^{-1} \cdot h^{-1}$)*	*Medium*
			Type	**Critical values**	**Operational range**	**Type**	**Parameters**	**Type**	**Operational range**		
Katakura et al., 1998	GS115 Mut$^+$	β_2GPId V	*S*-linear decreasing	$\mu_{crit} = 0.12$ (*)	$1.5 \leq S \leq 31$	*S*-linear decreasing	–	*S*-linear increasing	$1.5 \leq S \leq 31$ (*)	0.05	BSM
Kobayashi et al., 2000	HA2 Mut$^+$ (*)	rHSA	*S* non-monotonically increasing	$S_{crit} = 3.05$ $\mu_{crit} = 0.154$	–	μ-linear	$Y_{s/x} = 2.57$ $m_s = 0.0226$	μ-linear decreasing q_p-constant	$0.002 \leq \mu \leq 0.015$ $0.015 \leq \mu \leq 0.025$	0.17	1.5 · PBM
Ohya et al., 2005	HA2 Mut$^+$ (*)	rHSA	–	–	$0.001 \leq \mu \leq 0.060$	μ-linear	$Y_{s/x} = 2.82$ $m_s = 0.0209$	two-μ-linear increasing	$0.001 \leq \mu \leq 0.015$ $0.015 \leq \mu \leq 0.060$	0.21	PBM
Curvers et al., 2002	GS115 Mut$^+$	hCTRB	*S* monotonically increasing	$\mu_{max} = 0.084$	$S < 4.0$	–	–	Luedeking-Piret	$0.03 < \mu < 0.09$	0.03	BSM
Zhou and Zhang, 2002	GS115 Mut$^+$	rHV2	*S* non-monotonically increasing	$S_{crit} = 3.09$ $\mu_{crit} = 0.046$	–	–	$\underline{Y^*_{s/x} = 2.0}$ (at $\mu = 0.02$) (*)	*S* non-monotonically increasing	$S_{crit} = 0.50$	0.07	BSM
Pais et al., 2003	GS115 Mut^S	MPI	–	–	–	μ-linear	$\underline{Y_{s/x} = 3.27}$ $\underline{m_s = 0.026}$	μ-linear	–	0.01	D'Anjou & Daugulis, 2000
Dietzsch et al., 2011	KM71 Mut^S	HRP	–	–	–	μ-linear	$Y_{s/x} = 2.05$ $m_s = 0.016$ (*)	q_p-constant q_S-linear decreasing	$0.016 \leq q_S \leq 0.048$ $0.048 \leq q_S \leq 0.080$(*)	Data not available	BSM
Jahic et al., 2002	SMD1168 Mut$^+$	CBM-CALB	*qs*-linear	–	$0.005 \leq \mu \leq 0.18$	*S* monotonically increasing μ-linear	$Y_{s/x} = 2.78$ $m_s = 0.013$	–	–	0.06	BSM

Notes:

- $Y_{s/x}$, intrinsic substrate to biomass yield; $Y_{s/x}^*$, overall substrate to biomass yield
- S, $S_{crit} = [g \cdot L^{-1}]$; μ, $\mu_{crit} = [h^{-1}]$; $Y_{s/x}$, $Y^*_{s/x} = [g \cdot g^{-1}]$; $m_s = [g \cdot g^{-1} \cdot h^{-1}]$
- The temperature was 30 ºC in all cases, except in Curvers et al.(2002) and Ohya et al.(2005), where it was 25ºC; and 24 ºC in Potgieter et al.(2010)
- Biomass is expressed in g dry cell weight (DCW). Underlined values were calculated by using the conversion factor 4.2 WCW g ≈ 1 g DCW from Potgieter et al.(2010)
- (*) Operational ranges, critical values and parameter models were calculated from numerical or graphical data reported in each reference
- Specific productivities were calculated from data reported in each reference and are expressed in mg extracellular target protein per gram DCW per hour fermentation. For BoNT/A(Hc) mg of intracellular protein per gram DCW per hour fermentation.
- BSM, BMGY, FM22 are described in Cregg (2007) and PBM is described in Kobayashi et al. (2000).

Many studies have been performed on modeling the production of recombinant protein by *P. pastoris* and some of them used to optimize fermentation conditions to let maximal yields and productivities. A summary of kinetic models for different target proteins produced by *P. pastoris* Mut^+ or Mut^S strain with methanol as sole carbon and energy source is presented in Table 4. It includes critical values for specific rates, operational ranges, yields and maintenance parameters, specific productivities yielded and culture medium used. Because the overall performance of the protein production process is dictated by optimal selection of many variables including genetic and bioprocess engineering factors, it is rather difficult to compare the production results of Table IV since yields, specific rates and productivities depend on appropriate selection of such factors. However, available data allow one to establish common strategies and select typical ranges for some operational process parameters with a view to efficiently exploiting the *Pichia* cell factory.

4.4.3.1. Specific growth rate

The most widely used kinetic model for μ in *Pichia* is the well-known Monod equation, in culture broths typically falling in the substrate range from 0 to 4-5 $g \cdot L^{-1}$ (Curvers et al. 2002; Barrigón et al. 2012). However, the system is also accurately described by a non-monotonically increasing pattern (Table IV), usually selected for the high-methanol operating region, which is cytotoxic and inhibitory of microbial growth (Potvin et al. 2012). Other kinetic models typically developed from continuous culture setups have been used to identify relationships between growth rates and other kinetic rates (Ohya et al. 2005; Potgieter et al. 2010). The μ of the wild type strain of *Pichia* (μ_{max} = 0.16 h^{-1}) is

rarely reached during production of a recombinant protein (Potvin et al. 2012). Examining the actual maximum observed μ values (μ_{crit}) for different proteins (Table IV) reveals that only heterologous production of β_2GPIdV (Katakura et al. 1998), rHSA (Kobayashi et al. 2000) and recGAvi (Schenk et al. 2007) was close to μ_{max} for the wild type (0.12-0.15 h^{-1}). However, production of rOvINF-r (Zhang et al. 2004), BoNT/A(Hc) (Zhang et al. 2000), scFv (Kathri and Hoffmann, 2006), rHV2 (Zhou and Zhang, 2002), hCTRB (Curvers et al. 2002) and ROL affected cell growth significantly and reduced μ_{crit} to 0.05-0.08 h^{-1}. Although these values may have been underestimated owing to differences in media and operational conditions (e.g., dissolved oxygen level, temperature), production of a recombinant protein by *Pichia* clearly has an adverse effect on its growth kinetics. Besides, a knowledge of strain construction and details about bioprocess implementation does not suffice to estimate production levels for a given target protein—not even qualitatively.

4.4.3.2. Specific substrate uptake rate

Practically all bioprocesses conformed to a Pirt's maintenance energy model (μ-linear). The experimental intrinsic yield ($Y_{S/X}$) ranged from 2.0 to 4.2 $g \cdot g^{-1}$ (mean, 3.2±0.6 $g \cdot g^{-1}$) and the maintenance coefficient (m_s) from 0.013 to 0.042 $g \cdot g^{-1} \cdot h^{-1}$(mean, 0.020±0.008 $g \cdot g^{-1} \cdot h^{-1}$). Similarly to the μ, it is rather difficult to compare reported results of q_S owing to differences in genetic approach and operational strategy. Also, some results are unrealistic or even inconsistent with the C-balance owing to potential loss of methanol through the off-gas stream, generation of by-products and the presence of large errors in biomass determinations. It should be noted that conversion factors from absorbance and

WCW to DCW can affect yield determinations. Since WCW to DCW conversion factors are rarely reported, the yields of Table IV were calculated by using the conversion factor value reported by Potgieter et al. (2010). The yield coefficients of Zhou and Zhang (2002), and Dietzsch et al. (2011), were considerably lower than others. In fact, $Y^*_{S/X}$ values in the region of 2 g g^{-1} are usually related to glycerol or glucose uptake. On the other hand, carbon from methanol can be dissimilated in CO_2 to a ratio of 80 % (Jordà et al. 2012), so $Y^*_{S/X}$ can be as high as ~5 g g^{-1}as a result. Although the $Y_{S/X}$ values obtained by using a Pirt's model in this work are among the highest reported, the maintenance coefficient was lower to the mean. The particular operational conditions used (viz., air enriched with pure oxygen, pulse additions of nitrogen source, specific operational strategies) may have modulated metabolic regulation (e.g., by boosting CO_2 production and reducing biomass production accordingly in our case). By contrast, the results of Katakura et al. (1998) suggest that q_S decreases linearly with the substrate concentration; however, if Pirt's model is applied, q_S is linearly dependent on μ. Jahic et al. (2002) assumed μ not to be the key specific rate and used a monotonically increasing substrate model for q_S instead.

4.4.3.3. Specific substrate formation rate

The model for q_P differs significantly depending on the particular target protein. In Table 4, q_P is described as a function of the specific rate (μ or q_S), S, X or combinations thereof, using increasing, decreasing or independent profiles. The production rate of most proteins is related to cell growth; as a result, the q_P is related to the μ. Hensing et al. (1995) developed the growth-coupled Luedeking-Piret production model, a μ-linear model for

homologous protein production. Other, partially growth-dependent, production patterns are μ monotonically and μ non-monotonically increasing models. The specific formation rates of recGAvi (Schenk et al. 2008) and scFv (Yamawaki et al. 2007) have been represented with monotonic functions, and so have those of coffee bean α-Gal(Zhang et al. 2005), rOvIFN-τ (Shinha et al. 2003) and IgG1 (Potgieter et al. 2010) with μ non-monotonically increasing models. These production profiles are to some extent associated to changes in metabolic regulation rates that are growth-rate dependent. Thus, protein production may be partially explained by cell growth or substrate uptake, but also by other rates potentially associated to production bottlenecks. Some authors have characterized q_P with a stepwise model assuming linearity between rates (μ-linear or q_S-linear models) and/or constant protein production (q_P-constant model) that is formulated via inequalities. When a biosystem reaches the limit condition (viz., an inducer concentration, μ or q_S value), its metabolism changes significantly and so does its protein production pattern as a result. The specific production rates of ROL, GM-CSF, Jacobs et al. (2010), scFv, Cunha et al. (2004), HRP, Dietzsch et al. (2011) and rHSA, Ohya et al. (2005) have been described with stepwise models. A variety of production profiles were thus obtained, which can be ascribed to a stronger dependence on intracellular inducer levels, metabolism regulation rates or saturation of secretion pathways (Potvin et al. 2012). On the other hand, cell growth-independent protein production has been reported for rHV2 (Zhou and Zhang, 2002), BoNT/A(Hc) (Zhang et al. 2000), β_2GPIdV (Katakura et al. 1998) and scFv (Khatri and Hoffman, 2006), and q_P related to the substrate/inducer concentration. This relationship is quite consistent with the fact that production is related to the uptake of methanol, which is the limiting substrate as well as the inducer.

4.4.3.4. Productivity

Table 4 lists the specific productivities [in mg target protein·g $DCW^{-1}\cdot h^{-1}$] calculated from reported data for comparison with heterologous protein production by *P. pastoris* under *AOX1* promoter. Although direct comparison is rather difficult because many of the bioprocesses were not optimized, typical ranges for specific productivities can be identified. The mean specific productivity was 0.07±0.06 $mg\cdot g^{-1}\cdot h^{-1}$. Lower values ($\leq$0.01 $mg\cdot g^{-1}\cdot h^{-1}$) were obtained for mini proinsulin (País et al. 2003), recombinant glycosilated avidin (recGAvi) (Schenk et al. 2008) and antibody fragment (scFV) production (Cunha et al. 2004) without optimization of the operational conditions. However, scFV productivity was greatly increased (up to 0.04 $mg\cdot g^{-1}\cdot h^{-1}$) under optimal conditions, which suggests that optimizing the bioprocess can substantially improve its outcome (Kathri and Hoffmann, 2006; Yamawaki et al. 2007). In some cases, authors focused on improving product purity for efficient recovery. Moderate specific productivities (0.03-0.13 $mg\cdot g^{-1}\cdot h^{-1}$) were previously obtained for rOvIFN-τ (Sinha et al. 2004; Zhang et al. 2004, β_2GPIdV (Katakura et al. 1998), hCTRB (Curvers et al. 2002), scFV(Yamawaki et al. 2007), ROL (this work), CMB-CALB (Jahic et al. 2002), rHV2 (Zhou and Zhang, 2002), IgG1 (Potgieter et al. 2010) and coffee bean α-Gal (Zhang et al. 2005). The fact that different recombinant proteins including hydrolytic enzymes, antibody fragments, interferon and polypeptides are expressed at moderate productivities under P_{AOX1} confirms the high potential of *P. pastoris* as a protein expression host. High specific productivities (0.14-0.16 $mg\cdot g^{-1}\cdot h^{-1}$) have been reported for intracellular BoNT/A(Hc) (Zhang et al. 2000) and GM-CSF (Jacobs et al. 2010) under optimal bioprocess conditions. The highest specific productivity reported so far is that for human

serum albumin (rHSA) used as a reference protein for *Pichia* Expression Kit (Invitrogen, 2014). The first optimization policy applied to this system was intended to maximize its productivity and involved designing μ-profiles and having the system follow an optimum substrate feeding rate trajectory leading to a specific productivity of 0.18 $mg \cdot g^{-1} \cdot h^{-1}$ (Kobayashi et al. 2000). Subsequently, using an improved kinetic model and new constraints raised rHSA productivity to 0.21 $mg \cdot g^{-1} \cdot h^{-1}$(Ohya et al. 2005).

4.4.3.5. Bioprocess optimization

How a fermentation process is optimized depends on the particular criterion used (e.g., maximizing yield or productivity) and the kinetics of the process, the relationship between product formation and biomass growth, mainly. Therefore, each cell factory producing a specific recombinant protein in a given fermentation mode has a kinetic condition, singular specific rate or range thereof that leads to optimum bioprocess production. Optimizing a fed-batch bioprocess often involves identifying the particular feeding strategy that will maximize yield or productivity. This optimal control problem can be solved by applying Pontryagin's Maximum Principle (Claes et al., 1999). In some cases, maximal yield and productivity can also be obtained by operating at a highly constant specific rate (e.g., $\mu_{max,}$ μ_{min} or $q_{P,max}$), using an increasing substrate feeding rate. Conversely, depending on the particular kinetics of product formation, the highest productivity or yield can be obtained by using an optimal trajectory for the controlled variable (typically, the substrate or specific growth rate) (Maurer et al. 2006). Using optimal profiles to optimize a process is quite advisable when correlation between cell growth and protein production is non-linear, but rather monotonically or non-

monotonically increasing (Claes et al. 1999; de Hollander, 1993). Only a few of the recombinant proteins produced by *P. pastoris* under *AOX1* promoter listed in Table IV were obtained by using an optimal profile. For example, some kinetic models were used to optimize production of α-Gal and rHSA.α-Gal production was optimized by maximizing either productivity or yield, using μ as the control variable (Zhang et al. 2005). Optimal μ trajectories were also calculated and used to optimize total rHSA production (Kobayashi et al. 2000; Ohya et al. 2005). The influence of some operational strategies (μ-constant, *S*-constant) on ROL production by *P. pastoris* under *AOX1* promoter has been the subject of systematic study (Barrigón et al. 2013). However, defining optimal profiles for this system was outside the scope of this work and will be left for future work.

4.5. Conclusions

Heterologous production of *Rhizopus oryzae* lipase (ROL) by a *P. pastoris* P_{AOX1}-based system can be accurately described with a macrokinetic model. Two different models (A and B) were assessed for overall performance and mathematical simplicity here. Model A uses non-monotonic substrate functions for growth, substrate uptake and production, whereas model B uses a non-monotonic substrate function for growth, Pirt's model for substrate uptake and Luedeking-Piret equation for protein production. Although the two models described specific rates with great accuracy, Model B performed better in an overall validation for the entire operational range. Models could be further tested for application to process control and optimization in future work. A comparative analysis of different target protein production by *P. pastoris* under *AOX1* promoter revealed the following: protein production had a direct effect on growth, substrate uptake was proportional to growth and production differed significantly between proteins. The results can be used to estimate a range of specific rates for growth and substrate uptake, and also to relate specific rates. This may be the starting point for developing a generic strategy to improve protein production from process kinetics as a key to bioprocess optimization. In this way, near-optimal production could be achieved without the need to define an optimal trajectory with its associated additional operational and mathematical complexity. In general, the typical non-monotonic nature of growth kinetics, linear q_S-μ behavior and proportionally decreasing relationship between q_P and μ, suggest that yields and productivities can be maximized at a specific growth rate where the specific product formation rate is not maximal. Usually, maximum yields and productivities will be obtained at moderate S (2-4 $g \cdot L^{-1}$) and moderate to high μ (0.02-0.06 h^{-1}). MLFB

operational strategy (μ-controlled) has been successfully applied in the production of recombinant proteins in the *P. pastoris* cell factory, but generally reaching low production levels associated with low *S*. Although from an operational industrial point of view MLFB is an easy fed-batch strategy to be implemented with an open-loop control structure, MNLFB operation at rather constant *S* higher than 2 $g \cdot L^{-1}$ is required to maximize protein production. For higher *S* than 6 $g \cdot L^{-1}$ significant inhibition on growth has been often observed.

4.6. References

1. Albaek MO, Gernaey KV, Hansen MS, Stocks SM. 2011. Modeling enzyme production with *Aspergillus oryzae* in pilot scale vessels with different agitation, aeration, and agitator types. Biotechnol Bioeng 108:1828-1840.
2. 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.
3. Bailey, JE. 1998. Mathematical modeling and analysis in Biochemical Engineering: Past accomplishments and future opportunities. Biotechnol Prog 14:8-20.
4. Barrigón JM, Ramon R, Rocha I, Valero F, Ferreira EC, Montesinos JL. 2012. State and specific growth estimation in heterologous protein production by *Pichia pastoris*. AIChE J 58:2967-2979.
5. 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.
6. Boettner M, Steffens C, von Mering C, Bork P, Stahl U, Lang C. 2007. Sequence-based factor influencing the expression of heterologous genes in the yeast *Pichias pastoris* - A comparative view on 79 human genes. J Biotechnol 130:1-10.
7. Brocca S, Schmidt-Dannert C, Lotti M, Alberghina L. 1998. Design, total synthesis and functional overexpression of the *Candida rugosa* lip1 gene coding for a major industrial lipase. Protein Sci 7:1415-1422.
8. Carneiro S, Ferreira EC, Rocha I. 2013. Metabolic responses to recombinant bioprocesses in *Escherichia coli*. J Biotechnol 164:396-408.
9. Çelik E, Çalik P, Oliver SG. 2009. A structured kinetic model for recombinant protein production by Mut^+ strain of *Pichia pastoris*. Chem Eng Sci 64:5028-5035.
10. Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. Fems Microbiol Rev 24:45-66.

11. Chae HJ, Delisa MP, Cha HJ, Weigand WA, Rao G, Bentley WE. 2000. Framework for online optimization of recombinant protein expression in high-cell-density *Escherichia coli* cultures using GFP-fusion monitoring. Biotechnol Bioeng 69:275-85.
12. Claes JE, Geeraerd AH, Van Impe JF. 1999. Heuristic feed rate profiles for optimal yield and productivity of fed-batch processes. Chem Eng Commun 172:189-216.
13. Cos O, Resina D, Ferrer P, Montesinos JL, Valero F. 2005a. Heterologous production of *Rhizopus oryzae* lipase in *Pichia pastoris* using the alcohol oxidase and formaldehyde dehydrogenase promoters in batch and fed-batch cultures. Biochem Eng J 26:86-94.
14. Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005b. 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.
15. Cos O, Ramon R, Montesinos JL, Valero F. 2006a. A simple model-based control for *Pichia pastoris* allows a more efficient heterologous protein production bioprocess. Biotechnol Bioeng 95:145-154.
16. Cos O, Ramon R, Montesinos JL, Valero F. 2006b. 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.
17. Cregg JM. 2007. Methods in Molecular Biology: *Pichia* Protocols 2nd ed. Totowa. New Jersey: Humana Press.
18. Cregg JM. Heterologous protein expressed in *Pichia pastoris*, available online: http://www.kgi.edu/documents/faculty/James_Cregg/heterologous_proteins_expressed_in_Pichia_pastoris.pdf (Accessed 09 December 2014)
19. Cunha AE, Clemente JJ, Gomes R, Pinto F, Thomaz M, Miranda, Pinto R, Moosmayer D, Donner P. 2004. Methanol induction optimization for scFv antibody fragment production in *Pichia pastoris*. Biotechnol Bioeng 86:458-467.
20. Curvers S, Linnemann J, Klauser T, Wandrey C, Takors R. 2002. Recombinant protein production with *Pichia pastoris* in continuous fermentation - kinetic analysis of growth and product formation. Eng Life Sci 8:229-235.

21. de Hollander JA. 1993. Kinetics of microbial product formation and its consequences for the optimization of fermentation processes. Antonie van Leeuwenhoek 63:375-381.
22. d'Anjou MC, Daugulis AJ. 2000. Mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast. Biotechnol Lett 22:341-346.
23. Dietzsch C, Spadiut O, Herwig C. 2011. A dynamic method based on the specific substrate uptake rate to set up a feeding strategy for *Pichia pastoris*. Microb Cell Fact 10:14.
24. Duboc P, von Stockar U. 2000. Modeling of oscillating cultivations of *Saccharomyces cerevisiae*: Identification of population structure and expansion kinetics based on on-line measurements. Chem Eng Sci 55:149-160.
25. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO. 2007. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Mol Syst Biol 3:121.
26. Heijnen JJ, Roels JA, Stouthamer AH. 1979. Application of balancing methods in modeling the penicillin fermentation. Biotechnol Bioeng 21:2175-2201.
27. Hensing MC, Rouwenhorst RJ, Heijnen JJ, Van Dijken JP, Pronk JT. 1995. Physiological and technological aspects of large-scale heterologous-protein production with yeasts. Antonie van Leeuwenhoek 67:261-279.
28. Heyland J, Fu J, Blank LM, Schmid A. 2011. Carbon metabolism limits recombinant protein production in *Pichia pastoris*. Biotechnol Bioeng 108:1942–1953.
29. 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.
30. Invitrogen. *Pichia* Expression Kit. cat. no. K1710-01 Version A.0, available online: http://tools.lifetechnologies.com/content/sfs/manuals/pich_man.pdf (Accessed 09 December 2014).
31. Jacobs PP, Inan M, Festjens N, Haustraete J, Van Hecke AV, Contreras R, Meagher MM, Callewaert N. 2010. Fed-batch fermentation of GM-CSF-producing

glycoengineered *Pichia pastoris* under controlled specific growth rate. Microb Cell Fact. 9:93.

32. Jahic M, Rotticci-Mulder J, Martinelle M, Hult K, Enfors S-O. 2002. Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. Bioprocess Biosyst Eng 24:385-393.
33. Jenzsch M, Simutis R, Luebbert A. 2006. Generic model control of the specific growth rate in recombinant *Escherichia coli* cultivations. J Biotechnol 122:483-493.
34. Jordà J, Jouhten P, Cámara E, Maaheimo H, Albiol J, Ferrer P. 2012. Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose: methanol mixtures. Microb Cell Fact 11:57.
35. Katakura Y, Zhang W, Zhuang G, Omasa T, Kishimoto M, Goto Y, Suga K-I. 1998. Effect of methanol concentration on the production of human β2-Glycoprotein I domain V by a recombinant *Pichia pastoris*: A simple system for the control of methanol concentration using a semiconductor gas sensor. J Ferment Bioeng 86:482-487.
36. Khatri NK, Hoffmann F. 2006. Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant *Pichia pastoris*. Biotechnol Bioeng 93:871-879.
37. Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Tomomitsu K. 2000. High level secretion of recombinant human serum albumin by fed-batch fermentation of the methylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy. J Biosci Bioeng 90:280-288.
38. Liu CH, Chen CY, Wang YW, Chang JS. 2011. Fermentation strategies for the production of lipase by an indigenous isolate *Burkholderia sp*. C20. Biochem Eng J 58-59:96-102.
39. Lee SB, Bailey JE. 1984. Genetically structured models for lac promoter-operator function in *Escherichia coli* chromosome and in multicopy plasmids: lac operator function. Biotechnol Bioeng 26:1372-1382.
40. Levenspiel O. 2002. Modeling in chemical engineering. Chem Eng Sci 75:4691-4696.
41. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* system. Yeast 22:249-270.

42. Marx H, Sauer M, Resina D, Vai M, Porro D, Valero F, Ferrer P, Mattanovich D. 2006. Cloning, disruption and protein secretory phenotype of the GAS1 homologue of *Pichia pastoris*. Fems Microbiol Lett 64:40-47.
43. 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.
44. Montesinos JL, Lafuente J, Gordillo MA, Valero F, Solà C, Charbonnier S, Cheruy A. 1995. Structured modeling and state estimation in a fermentation process: lipase production by *Candida rugosa*. Biotechnol Bioeng 48:573-584.
45. Nielsen J. 1996. Modelling the morphology of filamentous microorganisms. Trends Biotechnol 14, 438-443.
46. Ohya T, Ohyama M, Kobayashi K. 2005. Optimization of human serum albumin production in methylotrophic yeast *Pichia pastoris* by repeated fed-batch fermentation. Biotechnol Bioeng 90:876-887.
47. País JM, Varas L, Valdés J, Cabello C, Rodríguez L, Mansur M. 2003. Modeling of mini-proinsulin production in *Pichia pastoris* using the AOX promoter. Biotechnol Lett 25:251-255.
48. Pérez J, Poughon L, Dussap C, Montesinos JL, Gòdia F. 2005. Dynamics and steady state operation of a nitrifying fixed bed biofilm reactor: mathematical model based description. Process Biochem 40:2359-2369.
49. Potgieter TI, Kersey SD, Mallem MR, Nylen AC, d'Anjou M. 2010. Antibody expression kinetics in glycoengineered *Pichia pastoris*. Biotechnol Bioeng 106:918-927.
50. 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.
51. Ramakrishna R, Ramakrishna D, Konopka A. 1996. Cybernetic modeling of mixed substitutable substrate environments. Preferential and simultaneous utilization. Biotechnol Bioeng 52:144-154.
52. Ren HT, Yuan JQ, Bellgardt KH. 2003. Macrokinetic model for methylotrophic *Pichia pastoris* based on stoichiometric balance. J Biotechnol 106:53-68.

53. Ren H, Yuan J. 2005. Model-based specific growth rate control for *Pichia pastoris* to improve recombinant protein production. J Chem Technol Biotechnol 80:1268-272.
54. 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.
55. Resina D, Bollok M, Khatri NK, Valero F, Neubauer P, Ferrer P. 2007. Transcriptional response of *P. pastoris* in fed-batch cultivations to *Rhizopus oryzae* lipase production reveals UPR induction. Microb Cell Fact. 16:6-21.
56. Resina D, Maurer M, Cos O, Arnau C, Carnicer M, Marx H, Gasser B, Valero F, Mattanovich D, Ferrer P. 2009. Engineering of bottlenecks in *Rhizopus oryzae* lipase production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. New Biotechnol 25:6.
57. Schenk J, Balazs K, Jungo C, Urfer J, Wegmann C, Zocchi A, Marison IW, von Stockar U. 2008. Influence of specific productivity and glycosylation of a recombinant avidin produced by a *Pichia pastoris* Mut^+ strain. Biotechnol Bioeng 99:368-377.
58. Schenk J, Marison IW, von Stockar U. 2007. A simple method to monitor and control methanol feeding of *Pichia pastoris* fermentations using mid-IR spectroscopy. J Biotechnol 128:344-353.
59. Sinha J, Plantz BA, Zhang W, Gouthro M, Schlege Vl, Liu CP, Meagher MM. 2003. Improved production of recombinant ovine interferon-tau by Mut^+ strain of *Pichia pastoris* using an optimized methanol feed profile. Biotechnol Prog 19:794-802.
60. Sohn SB, Graf AB, Kim TY, Gasser B, Maurer M, Ferrer P, Mattanovich D, Lee SY. 2010. Genome-scale metabolic model of methylotrophic yeast *Pichia pastoris* and its use for *in silico* analysis of heterologous protein production. J Biotechnol 5:705-715.
61. Soons ZITA, Ferreira EC, Rocha I. 2011. Identification of minimal metabolic pathway models consistent with phenotypic data. J Process Control 21:1483-1492.
62. Sreekrishna K. 2010. *Pichia*, Optimization of protein expression. In: Flickinger MC, editor. Encyclopedia of Industrial Biotechnology: Bioprocess, Bioseparation, and Cell Technology. New York: John Wiley & Sons Inc. p 1-16.

63. Surribas A, Stahn R, Montesinos JL, Enfors SO, Valero F, Jahic M. 2007. Production of a *Rhizopus oryzae* lipase from *Pichia pastoris* using alternative operational strategies. J Biotechnol 130:291-299.
64. Villadsen J, Nielsen J, Lidén G. 2011. Bioreaction Engineering Principles 3rd ed. Growth kinetics of cell cultures. New York: Springer Science+Business Media. p 271-357.
65. Yamawaki S, Matsumoto Y, 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.
66. Yokohama S. 2003. Protein expression systems for structural genomics and proteomics. Curr Opin Chem Biol 7:39-43.
67. Zhang W, Bevins MA, Plantz BA, Smith LA, Meagher MM. 2000. Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. Biotechnol Bioeng 70:1-8.
68. Zhang W, Liu CP, Inan M, Meagher MM. 2004. Optimization of cell density and dilution rate in *Pichia pastoris* continuous fermentations for production of recombinant proteins. J Ind Microbiol Biotechnol 31:330-334.
69. Zhang W, Sinha J, Smith LA, Inan M, Meagher MM. 2005. Maximization of production of secreted recombinant proteins in *Pichia pastoris* fed-batch fermentation. Biotechnol Prog 21:386-393.
70. Zhou XS, Zhang YX. 2002. Decrease of proteolytic degradation of recombinant hirudin produced by *Pichia pastoris* by controlling the specific growth rate. Biotechnol Lett 24:1449-1453.

Chapter 5:

Design of alternative operational strategies in *Pichia pastoris* Mut^{+} cultures through an oxygen transfer model.

5. Design of alternative operational strategies in *Pichia pastoris* Mut$^+$ cultures through an oxygen transfer model.

5.0. Abstract

The characterization of oxygen transfer capacity in different laboratory/pilot scale bioreactors was carried out. Response time of the electrode was measured and taken into account in k_La measurements. A k_La comparison was performed between the bioreactors in different operational conditions (stirring 300-1000 rprm and airflow 0.25-3vvm). The most representative empirical correlation for k_La estimation was explored to define the oxygen transfer model. The mass transfer model was tested in a methanol limited fed-batch culture (MLFB) showing a good prediction capability during the fed-batch phase, where the k_La estimation was about 15 % deviation. After satisfactorily testing the model in a MLFB culture, some oxygen limited fed-batch (OLFB) strategies were simulated in order to explore the oxygen limiting conditions as alternative operational strategies.

5.1. Introduction

In aerobic bioprocesses, oxygen must be continuously provided to the broth by a gas phase. Oxygen is transferred from gas phase to biomass, through gas-liquid interphase, becoming soluble in the media and finally going through liquid-solid interphase (Nagata, 1975; Bailey and Ollis, 1986; García-Ochoa and Gómez, 2009). In this process, many oxygen transport resistances are involved, but the film resistance around bubble controls the overall transfer rate (García-Ochoa and Gómez, 2009). The gas-liquid mass transfer is usually modelled by the double film theory, the equation can be written:

$$K_G(p_G - p_G^*) = K_L\,(O_2^* - O_2) \quad (1)$$

$$\frac{1}{K_L} = \frac{1}{H_{O2}k_g} + \frac{1}{k_L} \quad (2)$$

where, K_i are overall mass transfer coefficients and k_i the local mass transfer coefficients [$m \cdot s^{-1}$], p the oxygen partial pressure [Pa], O_2 the oxygen concentration in the liquid phase [$mol \cdot m^{-3}$], L subscripts indicate the liquid phase and G the gas phase, * superscript the critical value, H_{O2} is the Henry's Law constant for oxygen solubility in water [$m^3 \cdot Pa \cdot mol^{-1}$]. K_L coefficient is approximated to k_L because the oxygen solubility is low ($H_{O2} >> k_G, k_L$). In water and also in culture the liquid resistance is the highest. Due to difficulties to measure k_L coefficient, volumetric mass transfer coefficient (k_La) [s^{-1}] is commonly studied instead of k_L. The specific interfacial area (a) is gas-liquid interfacial area of the bubbles per unit of liquid volume [$m^2 \cdot m^{-3}$]. (a) value depends on bioreactor geometry, system flow dynamics and liquid properties (Ochoa et al., 2010; Villadsen et al., 2011; Doran et al., 2013).

In order to prevent or determine any oxygen limitation during the bioprocess, it is essential to perform a study of the oxygen transfer capacity of the bioreactor. The oxygen transfer rate (OTR) of a bioreaction system [$mol \cdot L^{-1} \cdot s^{-1}$] is defined by the following equation:

$$OTR = k_L a\,(O_2^* - O_2) \quad (3)$$

Where (O_2^* - O_2) is driving force, which corresponds to difference between the oxygen solubility (O_2^*) and the dissolved oxygen concentration in the liquid (O_2) [$mol \cdot L^{-1}$]. Oxygen solubility depends on the culture media, ionic strength, viscosity, temperature, pressure, and the chemical reactions (Weisenberger and Schumple, 1996; Gros et al., 1999; Pérez et al., 2006; Liang and Yuan, 2007; García-Ochoa and Gómez, et al., 2010). OTR may be increased by taking advantage of the driving force (O_2^* - O_2), e.g. O_2^* is usually higher in water than saline or viscous media; or by studying k_La of the bioreactor and applying the optimal configuration that allows maximizing it.

Studies about the k_La determination in aerobic bioreactors are not a new approach. Differences between transfer capacities of different bioreactor geometries and operational conditions have been studied (Van't Riet 1979, Bailey y Ollis, 1986; Blanch and Clark, 1996; Fujisová et al., 2007; García-Ochoa and Gómez, 2009, Karimi et al., 2013). k_La can be measured experimentally by chemical or physical methods; or estimated by using empirical correlations. k_La determination by chemical methods are based on measuring the chemical reaction, where the reaction rate is faster than mass transfer. Thus, mass transfer controls the process. The most popular chemical methods are the sodium sulphite oxidation (Cooper et al., 1944; Fyferling et al., 2008; Pinelli et al. 2010) and the carbon

dioxide adsorption (Danckwerts, 1970; Bioshnoi and Rochelle, 2000). k_La value can be estimated by measuring Na_2SO_3 or CO_2 consumption rate respectively. However, both methods have the same problem: ion presence (SO_3^{-2} and OH^-). These ions affect into bubble size, making higher the k_La value than utilizing other methods (Van't Riet, 1979; García-Ochoa and Gómez, 2009).

Nowadays, physical methods are the most used to k_La determination. The dynamic method is based on measuring the dissolved oxygen concentration by polarographic oxygen probe during the oxygen desorption or absorption processes (Tribe et al., 1995, Galaction et al., 2004; Pérez et al., 2006; García-Ochoa and Gòmez, 2009; Patel and Thibault, 2009). The method consists in measuring the evolution of the dissolved oxygen during a period of time until the oxygen concentration reaches the steady state.

$$\ln\left(\frac{O_2^* - O_{2,2}}{O_2^* - O_{2,1}}\right) = -k_L a \quad (t_2 - t_1) \quad (4)$$

Where t_1 is the initial time, t_2 the final time [s], $O_{2,i}$ the dissolved concentration at t_i. In case of desorption method, after oxygen saturation in the liquid is reached, the oxygen is removed from the liquid phase, e.g. by means of nitrogen supply, until the oxygen concentration drops down to zero. In contrast, absorption method consists, after desorption process, in air supplying until oxygen saturation in the liquid is reached. However, in order to determine accurately the k_La parameter by the dynamic method, it may be necessary to take into account the response time of the probe (τ) [s]. The time required for the probe to reach 63.2% of its final value, when it is exposed to a step change in concentration, is τ. It is highly recommended to work with probes with

$\tau << k_L a^{-1}$ (Tribe et al., 1995; Doran et al., 2013). The dynamic of the electrode can be neglected only if time characteristic for the oxygen transport ($k_L a^{-1}$) is 10 times higher. The τ of the fast oxygen electrochemical sensors is about 5 s (García-Ochoa and Gòmez, 2009). However, commercially available steam sterilizable electrodes which are used for measurements in bioreactor systems have large response time constants 10-100 s (Tribe et al., 1995; Doran et al., 2013).

$k_L a$ may be correlated with dimensionless numbers of Sherwood (Sh), Reynols (Re), Schmidt (Sc) and with other variables and an extended list of operational conditions, geometries and reactors (García-Ochoa and Gómez, 2009; Villadsen et al., 2011). However, the most utilized correlation of $k_L a$ for Newtonian flows was defined by Van't Riet, (1979) and it is showed in Eq. 5 (Cooper et al., 1944, Bailey and Ollis, 1986; Martín et al., 2008; Fyferling et al., 2008; García-Ochoa and Gómez, 2009, Doran et al., 2013; Karimi et al., 2013).

$$k_L a = k_1 \left(\frac{P_g}{V}\right)^{\alpha} v_S^{\beta} \quad (5)$$

Where Pg is the effective power [W], which covers both the shaft power and the power provided by the isothermal expansion of the gas, V is the liquid volume [m^3 or L], v_S the superficial gas rate [$m \cdot s^{-1}$], k_1 the correlation constant, α and β are the exponents. $k_L a$ is proportional to the effective power per volume of liquid in aeration conditions and to superficial gas rate for Newtonian flows (Eq. 5).

Some authors have replaced Pg/V variable by stirring rate (N) [rpm] (Marques et al., 2009; Pinelli et al., 2010) or have introduced viscosity (μ_v) for non-Newtonian [Pa·s], see

in Eq. 6 (García-Ochoa and Gómez, 1998; Gogate et al., 2000; Badino et al., 2001; Puthli et al., 2005; Fujisova et al. 2007; Albeak et al., 2011).

$$k_L a = k_1 \left(\frac{P_g}{V}\right)^{\alpha} v_S^{\beta} \ \mu_v^{\gamma} \quad (6)$$

The global mass transfer capacity of the system can be modified by biochemical factors, such as tensoactive agents (Bailey and Ollis, 1986), ionic strength of the liquid (Van't Riet 1979; Fujisova et al. 2007; Fyferling et al. 2008; García-Ochoa and Gómez, 2009), oxygen diffusivity in the liquid and biomass (Fyferling et al. 2008; García-Ochoa and Gómez, 2009 and 2010). For example, surfactant substances may provoke an additional resistance. On the other hand, biomass may cause an acceleration effect on the global mass transfer due to oxygen taken up. These effects that correct the global mass transfer capacity are studied by enhancement factor (E) (Ju and Sundarajan, 1992; Galaction et al. 2004; Fyferling et al. 2008; García-Ochoa and Gómez, 2009; Çalik et al., 2010).

$k_L a$ and OTR can be also measured during the bioprocess. Oxygen mass balance in the bulk liquid phase is described by the following equations:

$$\frac{dO_2}{dt} = OTR - OUR \quad (7)$$

$$OUR = q_{O_2} \cdot X \quad (8)$$

Where the OUR is the oxygen uptake rate [$mol \cdot L^{-1} \cdot s^{-1}$ or $mol \cdot L^{-1} \cdot h^{-1}$], qO_2 the specific oxygen uptake rate [$mol \cdot g^{-1} \cdot s^{-1}$ or $mol \cdot g^{-1} \cdot h^{-1}$], and X the biomass concentration [$g \cdot L^{-1}$]. The evolution of dissolved oxygen concentration (O_2) in the broth depends essentially on OTR, from the gas to the liquid phase and to biomass film, and on OUR by the respiring

microorganism. OTR can be indirectly estimated by measuring the OUR directly by gas phase analysis or by dynamic method. OUR can be measured by the difference between the oxygen concentration of the air inflow and air outflow (Belluci and Hamaker, 2011; Doran et al., 2013). Supposing steady or pseudo-steady state hypothesis, this means: state variables, rates and operational conditions are not varying, OTR is equal to OUR. If O_2^* and O_2 are known, k_La can be calculated. The dynamic "gas out-gas in" method consists of measuring the evolution of the dissolved oxygen during the bioprocess by a method similar to adsorption, described previously. However, oxygen consumption by the respiratory activity of microorganisms has to be taken into account for k_La determination. When the gas supply is turned off to the bioreactor, desorption step is performed by microorganisms and OUR is measured. When gas supply is turned on, O_2 increases to the stationary concentration. The dynamic evolution of oxygen is used to determine the k_La (Tribe et al., 1995, Pérez et al., 2006; Hanson et al., 2009; Patel and Thibault, 2009; García-Ochoa et al., 2010; Belluci and Hamaker, 2011; Kirk and Szita, 2012; Doran et al., 2013).

OTR and k_La determination or estimation are essential for bioreactor design and process scale-up. Additionally, OTR may be used to control the dissolved oxygen and indirectly to estimate or control the microbial concentration (Jenzsch et al., 2004; Oliveira et al., 2004, 2005; Potgieter et al., 2010; Valero et al., 2013).

The methylotrophic yeast *Pichia pastoris* is one of the most efficient cell factories for the production of recombinant proteins. *P. pastoris* is able to grow to high cell densities on

minimal medium, potential for high expression levels, perform posttranslational modifications and secrete the recombinant proteins (Macauley-Patrick et al., 2005; Potvin et al., 2012). However, the most important characteristic of *P. pastoris* is the existence of a strong and tightly regulated promoter from the alcohol oxidase 1 gene, the alcohol oxidase 1 promoter (P_{AOX1}) (Cereghino and Cregg, 2000; Cos et al., 2006; Sreekrishna, 2010; Potvin et al., 2012). The fed-batch is the most used operational mode for Pichia's high-cell-density fermentation, which also allows achieving high productivities. The utilized substrates are usually glucose, glycerol, sorbitol and methanol. O_2 is usually kept above the 20% of dissolved oxygen saturation in air (O_2^*) (Cregg, 2007, Potvin et al., 2012; Valero, 2013, Invitrogen 2014). P_{AOX1} have been used in the *Pichia* system for high level expression of heterologous proteins using methanol as an inducer substrate (Cereghino et al., 2000). When methanol is utilized as a sole carbon source, high oxygen requirements are necessary to metabolize it. Especially in the first step of the oxidative methanol assimilation, which methanol is oxidated by the alcohol oxidase enzyme (AOX) to form formaldehyde and hydrogen peroxide (Cos et al., 2005, Cregg, 2007; Zhang et al., 2007; Yano et al 2009, Yurimoto et al., 2011, Valero et al., 2013). Utilizing methanol, q_{O2} may be up to 5 times higher than utilizing glycerol (Barrigón et al., 2012, citas). So, oxygen provide to *P. pastoris* may be a critical limitation for protein production, especially when using P_{AOX1} Mut^+ phenotype at high-cell-density (Wu and Fu, 2012). In order to avoid oxygen limitation, enhancing oxygen transfer studies have been performed. Enriching air flow by mixing air with pure oxygen or pressurizing the process allows an increase in the oxygen partial pressure in the gas mixture (Yang and Wang, 1992). Nevertheless, these operational conditions are only recommended when all safer

and cheaper strategies were exhausted, such as k_La optimization or increase the culture oxygen solubility (Zhang et al., 2007). Finally, when oxygen limitation cannot be overcome, oxygen limited fed-batch operational strategy (OLFB) is proposed.

OLFB is a control method that adjusts the methanol consumption by means of the oxygen transfer rate, often provoking oxygen taken up limitations to the microorganisms (Khatri and Hoffmann, 2006b). Although oxygen limitation should generally be avoided during the induction phase (Invitrogen, 2014), OLFB cultivations strategy have been successfully implemented in *P. pastoris* cultivations through different approaches (Potvin et al., 2012). MNLFB strategy was converted into OLFB, when the maximum OTR of the reactor was reached, and dissolved oxygen concentration declined to zero (Khatri and Hoffmann 2006a). Exponential and constant airflow rates were carried out to control the specific methanol uptake rate during production at the high methanol concentration (3% v/v) by Khatri and Hoffmann, (2006b). In Potgieter et al., (2010), the production of IgG1 in MLFB, MLFB+OLFB and OLFB cultures was explored by glycoengineered *Pichia pastoris* strain. The maximum of OTR was set at 150 mmol L^{-1} h^{-1} in OLFB cultures. No oxygen restriction was conducted in MLFB. A comparative analysis between MLFB ($\mu_{setpoint}$=0.008 h^{-1}) and OLFB ($OTR_{setpoint}$= 30 mmol$\cdot L^{-1}\cdot h^{-1}$) cultures was presented in Berdichevsky et al., (2011).

OLFB and MLFB were compared in the previous works in terms of product titers and productivities. In case of the production of IgG1, similar product titer was obtained for three operational modes, but OLFB achieved higher volumetric productivity (Potgieter et

al., 2010). Higher concentration of scFv (>3.5 $g \cdot L^{-1}$) were obtained by OLFB strategy (Khatri and Hoffmann, 2006b). Productivities of mAb were higher for MLFB cultures because the limiting nutrient (methanol) allowed working with higher μ's (Berdichevsky et al., 2011). However, the total titer of mAb was higher in OLFB because in OLFB was not affected by mAb proteolysis degradation, which was measured in MLFB cultures. The total β-glucosidase protein in the medium was lower in the OLFB than MLFB (Charoenrat et al., 2005). Nevertheless, in terms of product purity the OLFB presented better performance. N-glycans composition was 10 % higher in OLFB than in MLFB cultures. The fraction of antibody containing complex N-glycans was approximately 10% higher in the oxygen-limited process (Berdichevsky et al., 2011). The activity and specific activity of β-glucosidase was higher in OLFB bioprocess (Charoenrat et al., 2005).

In this work, the oxygen transfer capacity of different laboratory scale bioreactors has been characterized. The oxygen transfer model has been validated in heterologous *Rhizopus oryzae* lipase (ROL) protein production by *Pichia pastoris* under *AOX1* promoter in methanol limited fed-batch culture (MLFB) and therefore used to define alternative operational strategies based on oxygen limited fed-batch operations (OLFB).

5.2. Materials and methods

5.2.1. Strain and inoculum preparation

The wild type *P. pastoris X-33* strain containing the pPICZαAROL vector was used for heterologous expression of *R. oryzae* lipase (ROL) under control of the P_{AOX1} (Cos et al. 2005b). Pre-inocula for bioreactor cultures were grown in 1 L baffled shake flasks containing 200 mL of YPD medium (10 $g \cdot L^{-1}$ yeast extract, 20 $g \cdot L^{-1}$ peptone, 20 $g \cdot L^{-1}$ glucose and 1 $mL \cdot L^{-1}$ zeocin solution, 100 $mg \cdot mL^{-1}$) at 30 ºC with shaking at 150 rpm for 24 h. The cultures were centrifuged at 4500 ×g, and harvested cells resuspended in bioreactor culture medium and used to inoculate the bioreactor.

5.2.2. Equipment

Five different laboratory/pilot plant stirred-tank bioreactors: 4L *Biostat* B, 5L *Biostat* ED and 50L *Biostat* UD (B. Braun, Sartorius Biotech, Melsungen, Germany), 3L and 7L *Biobundle* (Applikon Biotechnology, Schiedam, Netherlands) were used for k_La determination in water at 30ºC. Stirring and aeration was set between 300-1000 rpm and 0.25-3vvm (L air· L liquid$^{-1} \cdot$ min^{-1}), respectively. Temperature and stirring was controlled by means of Data Control Unit (DCU). Aeration was controlled by thermal mass flow controller Brooks 5851-E and Brooks 5850-E (Brooks Instruments, Hatfield, USA).The bioreactors configuration is presented in Table and Figure 1. Temperature, pO_2, pH and methanol probes were integrated in each bioreactor setup. Dissolved oxygen was measured by InPro 6800 (Mettler-Toledo Process Analytica, Wobur, USA), Oxiferm XL (Hamilton Bonaduz AG, Bonaduz, Switzerland) or Applisens Low Drift **(**Applikon

Biotechnology, Schiedam, Netherlands) polarographic sensors, depending on the bioreactor system (Table 1).

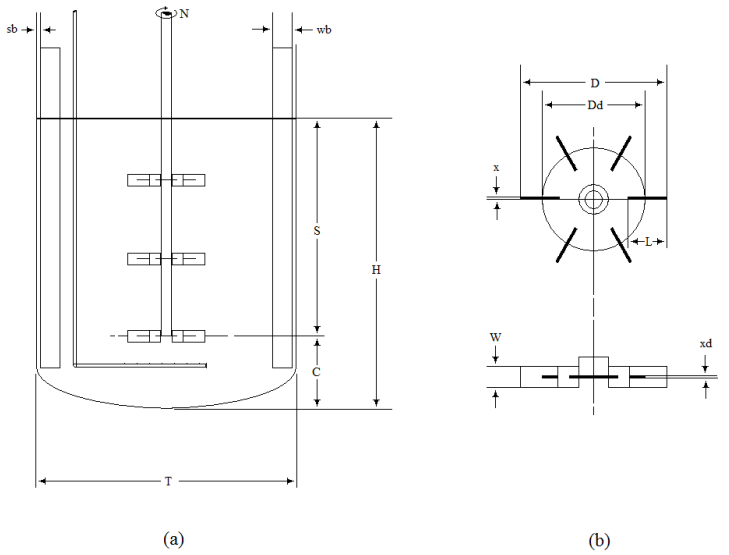

Figure 1. Bioreactor scheme.

Table 1. Bioreactor geometric parameters.

Variable	Acronym	Units	Biostat B	Biostat ED	Biostat UD	Biobundle 3L	Biobundle 7L
Tank diameter	T	m	0.16	0.14	0.315	0.125	0.16
Baffles number	nb	-	4	*	4	3	3
Baffler width	Wb	m	0.012	-	0.023	0.010	0.016
Rushton impellers	Ni	-	3	3	3	2	2
Impeller diameter	D	m	0.063	0.07	0.125	0.045	0.060
Liquid height	H	m	0.215	0.3	0.685	0.08	0.235
Liquid volume	V	L	4	5	5	1	5
Total height	H_T	m	0.276	0.392	0.964	0.245	0.341
Blade length	L	m	0.018	0.020	0.040	0.012	0.015
Blade width	W	m	0.013	0.02	0.025	0.012	0.012
Blade thickness	x	m	0.0015	0.002	0.003	0.0015	0.0014
Disk diameter	Dd	m	0.0345	0.045	0.075	0.030	0.040
Disk thickness	xd	m	0.0015	0.002	0.003	0.0015	0.003
Bottom to 1st impeller height	C	m	0.040	0.070	0.150	0.023	0.075
Ratio H/T	H/T	-	1.3	2.1	2.2	0.6	1.5
Ratio H_T/T	H_T/T	-	1.7	2.8	3.1	2.0	2.2
Ratio D/T	D/T	-	0.4	0.50	0.40	0.4	0.4

5.2.3. Bioreactor cultivation

The cultivation was conducted in a 3L *Biobundle* bioreactor. Invitrogen's fermentation basal salts medium (BSM) [26.7 $mL \cdot L^{-1}$ H_3PO_4 (85%), 0.93 $g \cdot L^{-1}$ $CaSO_4$, 18.2 $g \cdot L^{-1}$ K_2SO_4, 14.9 $g \cdot L^{-1}$ $MgSO_4 \cdot 7H_2O$, 4.13 $g \cdot L^{-1}$ KOH and 40 $g \cdot L^{-1}$ glycerol], 5 $mL \cdot L^{-1}$ PMT1 salts solution (6.0 $g \cdot L^{-1}$ $CuSO_4 \cdot 5H_2O$, 0.08 $g \cdot L^{-1}$ NaI, 3.0 $g \cdot L^{-1}$ $MnSO_4 \cdot H_2O$, 0.2 $g \cdot L^{-1}$ $Na_2MoO_4 \cdot 2H_2O$, 0.02 $g \cdot L^{-1}$ H_3BO_3,0.5 $g \cdot L^{-1}$ $CoCl_2$, 20.0 $g \cdot L^{-1}$ $ZnCl_2$, 65.0 $g \cdot L^{-1}$ $FeSO_4 \cdot 7H_2O$, 0.3 $g \cdot L^{-1}$ biotin and 5 $mL \cdot L^{-1}$ concentrated H_2SO_4), 2 $mL \cdot L^{-1}$ biotin solution (200 mg L^{-1}) and 0.5 $mL \cdot L^{-1}$ antifoam agent A6426 (Sigma–Aldrich, St. Louis, MO, USA)] were used to prepare 1 L of batch medium. The biotin and trace salts solutions were sterilized separately by passage through SLGV013SL filters of 0.22 μm pore size (Millipore Corporation, Billerica, MA, USA).

The cultivation conditions were as follows: temperature 30 °C, pH 5.5 (adjusted by adding 30% (v/v) NH_4OH during the batch phase, and 5 M KOH during the transition and induction phases), dissolved oxygen above 7.0 % in pure oxygen (≈ 30 % air saturation), oxygen-enriched air flow rate 0.5-5 $L \cdot min^{-1}$ and stirring rate 800-1000 rpm

5.2.4. Operational strategies

The fermentation process was divided in three phases: glycerol batch phase (GBP), transition phase (TP) and methanol induction phase (MIP). The batch phase started with inoculation of the bioreactor and ended when glycerol was depleted as signalled by an abrupt increase in O_2. GBP was followed by TP, which lasted 5 h, adding glycerol and methanol co-feeding with decreasing and constant feeding rates, respectively (Cos et al.

2005a). Methanol was used as the carbon source and inducer substrate in the methanol induction phase (MIP), which was conducted in by using a pre-programmed exponentially increasing feeding rate for methanol limited fed-batch culture (MLFB) as described in Barrigón et al. (2013). The MLFB culture was set at $\mu_{\text{set-point}}$ =0.020 h^{-1}.

5.2.5. Off-line analysis

Biomass concentration was quantified as dry cell weight (DCW) per litre of culture broth. Pellets and supernatants were separated by centrifugation at 10000 ×g for 1 min, washed, centrifuged twice in ddH_2O at 4500 ×g for 3 min and dried to constant weight at 105 ºC. The relative standard deviation for DCW was about 5%. The supernatant was used for substrate and product determinations. Methanol and glycerol were determined by HPLC as described elsewhere (Arnau et al. 2011). The residual standard deviation (RSD) was estimated about 0.5%. Extracellular lipolytic activity was monitored spectrophotometrically in 400 mM Tris-HCl + 10mM $CaCl_2$ buffer at pH 7.25 at 30 ºC, using the Roche lipase colorimetric kit (Ref. 11821792, Roche, Mannheim, Germany) as described elsewhere (Resina et al. 2004) on a Cary Varian 300 spectrophotometer (Varian Inc., Palo Alto, CA, USA). Measurements were made at 580 nm for 7 min in triplicate. RSD was estimated as 5%.

5.2.6. Online methanol determination

The methanol concentration was monitored online by using a sensor from Raven Biotech (Vancouver, BC, Canada) immersed in the culture broths (Arnau et al. 2011).

5.3. Theory and calculation

5.3.1. Experimental k_La determination

The k_La was experimentally determined by applying the dynamic method without biomass. Dissolved oxygen concentration was measured during the absorption step by sterilizable polarographic oxygen probe. The response of dissolved oxygen electrodes is usually assumed to follow first-order kinetics. The method to determine τ was detailed in Perez et al., (2006). The evolution of dissolved oxygen considering the response time is described by the following equation:

$$O_2 = O_2^* + \frac{O_2^* - O_2}{1 - \tau\, k_L a}\left(\tau\, k_L a\ \ exp\left(-\frac{t}{\tau}\right) - exp(-\tau\, k_L a)\right) \quad (9)$$

kLa determination was performed by fitting the measured curve to optimal solution of Eq. (9), which is calculated by the method of minimal least squares (SigmaPlot 11.0, Regression Wizard, Systat Software, Inc, San Jose, USA).

5.3.2. Van't Riet's correlation

The Van't Riet's correlation (Eq. 9) was used to describe the dependence of k_La on operational conditions. Superficial gas rate was calculated by the following equation:

$$v_S = \frac{Q_G}{\frac{\pi}{4}T^2} \quad (10)$$

Where, Q_G is the gas flow rate and T the internal bioreactor diameter. In Eq. (5), P_g is related with energy supply to liquid from stirring, and gas superficial rate is related with aeration. So, P_g can be calculated from the effective power in non-aeration conditions (P_0). A procedure to calculate P_0 is the Eq. (11), which requires to measure following

variables: Power number (N_p), Reynolds number (Re), liquid density (ρ), the kind and number of impellers (Blanch and Clark, 1996; Jenzsch et al., 2004; Amaral et al., 2007).

$$P_0 = N_P \cdot \rho \cdot N^3 \cdot T^3 \quad (11)$$

However, the P_0 correction for aeration systems proposed by Hughmark, (1980) is one of the most used equations in bioreactors that fulfil D/T is between 0.33-0.54 (Blanch and Clark 1996; Gogate et al., 2000; Amaral et al., 2007, Gill et al., 2007; Fyferling *et al.* 2008). In this work, Hughmark's equation (Eq. 12) was utilized to calculate the effective power.

$$\frac{P_g}{P_0} = 0.10\left(\frac{Q_G}{NV}\right)^{-0.20}\left(\frac{N^2T^4}{gw_bV^{2/3}}\right)^{-0.25} \quad (12)$$

5.3.3. Mass balance and stoichiometric equations

The reaction involved in the oxidative uptake of substrates (glycerol or methanol) to form biomass and products is described by equation 13:

$$k_1S + k_2O_2 \rightarrow X + k_3P + k_4CO_2 \quad (13)$$

where S denotes substrate (glycerol or methanol), O_2 oxygen, X biomass, P product and CO_2 carbon dioxide; and k values are stoichiometric coeficients. The mass balance equations for a fed-batch cultivation process, considering conversion rates of biomass formation, substrate uptake and product formation, can be formulated by equation 14:

$$\frac{d}{dt}\begin{bmatrix} XV \\ SV \\ PV \\ O_2V \\ CO_2V \end{bmatrix} = \begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} XV + \begin{bmatrix} 0 \\ F \cdot S_0 \\ 0 \\ OTR \cdot V \\ -CTR \cdot V \end{bmatrix} \quad (14)$$

where μ is the specific growth rate [h^{-1}], q_S the specific substrate uptake rate [$g \cdot g^{-1} \cdot h^{-1}$], q_P the specific production rate [$U \cdot g^{-1} \cdot h^{-1}$], q_{O2} the specific oxygen uptake rate [$mol \cdot g^{-1} \cdot h^{-1}$], q_{CO2} the specific carbon dioxide production rate [$mol \cdot g^{-1} \cdot h^{-1}$], F the substrate feeding rate [$L \cdot h^{-1}$], V the volume of broth in the reactor [L], S_0 the substrate feeding concentration [$g \cdot L^{-1}$], OTR the oxygen transfer rate [$mol \cdot L^{-1} \cdot h^{-1}$] and CTR the carbon dioxide transfer rate [$mol \cdot L^{-1} \cdot h^{-1}$].

The volume change due to substrate feeding, base and antifoam addition, sampling, water evaporation and net mass gas flow rate was obtained from the total mass balance. Substrate and product concentrations were referred to the whole medium, including biomass volume (Barrigón et al. 2013).

5.3.4. Kinetic model for ROL production

Heterologous production of *Rhizopus oryzae* lipase (ROL) by a *P. pastoris* P_{AOX1}-based system has been described previously (Barrigon et al., 2012 and in the 4th chapter). μ was modeled by substrate non-monotonically increasing function (Eq. 15), where the substrate is the methanol.

$$\mu = \frac{\mu_{max}\, S}{K_{s,X} + S + \frac{S^2}{K_{I,X}}} \quad (15)$$

The linear law, so-called Luedeking–Piret relationship, was used to model the q_S (Eq. 16), q_{O2} (Eq. 17) and q_P (Eq. 18 and 19).

$$q_S = Y_{S/X}\mu + m_S \approx \; Y^*_{S/X}\; \mu \quad (16)$$

$$q_{O2} = Y_{O2/X}\mu + m_{O2} \approx \; Y^*_{O2/X}\; \mu \quad (17)$$

$$q_P = m_P \qquad S < 2\ g \cdot L^{-1} \quad (18)$$

$$q_P = Y_{P/X}\mu + m_P \approx \ Y^*_{P/X}\,\mu \qquad S \geq 2\ g \cdot L^{-1} \quad (19)$$

where q_i is the i-specific rate, $q_{max,i}$ the maximum value of the i-specific rate, $K_{S,i}$ the i-substrate non-monotonic increasing model constant, $K_{I,i}$ the i-substrate inhibition non-monotonic increasing model constant, $Y_{i2/i1}$ the individual yield of component i_2 to i_1, the $m_{i,2}$ the maintenance coefficient of component i_2 and the $Y_{i2/i1}^*$ the overall yield coefficient yield of component i_2 to i_1.

The stoichiometric, yield and maintenance coefficients for ROL production used for bioprocess modelling were identified in Barrigón et al (2012) and updated in the 4th Chapter. In case of the overall oxygen-biomass yield coefficient ($Y_{O2/X}^*$) the electron balance were applied to estimated it from the overall substrate-biomass yield coefficient ($Y_{S/X}^*$) (Xie et al., 2013). Kinetic and stoichiometric coefficients are listed in Table 2. Initial conditions for simulation are shown in Table 3.

Table 2. Kinetic and stoichiometric coefficients used in *P. pastoris* P_{AOX1} (Mut$^+$) fermentation for ROL production.

Coefficients	Units	Batch	Fed-batch
Substrate	-	Glycerol	Methanol
μ_{max}	$[h^{-1}]$	0.260	0.069
$K_{s,X}$	$[g \cdot L^{-1}]$	0.20	0.40
$K_{I,X}$	$[g \cdot L^{-1}]$	-	8.85
$Y_{S/X}$	$[g_S \cdot g_X^{-1}]$	1.97	4.21
$m_{S/X}$	$[g_S \cdot g_X^{-1} \cdot h^{-1}]$	0.008	0.0142
$Y_{O2/X}^*$	[mol $O_2 \cdot g_X^{-1}$]	$2.6 \cdot 10^{-2}$	$1.9 \cdot 10^{-1}$
$Y_{P/X}$	$[U \cdot g_X^{-1}]$	-	4567
m_P	$[U \cdot L^{-1} \cdot h^{-1}]$	-	39.3

Table 3. Initial conditions used for fed-batch simulations.

Variable	Units	Batch / Glycerol	Fed-batch / Methanol
S	[g ·L^{-1}]	40	3.0
X	[g ·L^{-1}]	0.5	28
V	[L]	1.0	1.0

5.3.5. Calculation of specific rates

Discrete and mean specific rates in the induction phase were calculated in all fermentations; also, discrete specific rates (μ_t, $q_{S,t}$ and $q_{P,t}$) for each off-line data [$(XV)_t$, $(SV)_t$ and $(PV)_t$] were determined by using the respective mass balance equations (Barrigón et al. 2013).

5.3.6. Statistics indexes

With the aim to evaluate "goodness-of-estimation" of the k_La estimation, *MRE* was calculated as a single metric (Eq. 20).

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i} \quad (20)$$

where *n* is the number of data points for an individual experiment, $\hat{y}_i$ the i^{th} estimated value, y_i the corresponding i^{th} actual value for the bioprocess. Each individual experimental value of the mass transfer coefficient under process conditions is calculated by the oxygen mass balance (Eq. 21).:

$$k_L a_{exp} = \frac{q_{O2}X + \frac{dO_2}{dt}}{O_2^* - O_2} \quad (21)$$

Determination of oxygen solubility (O_2^*) in culture medium was calculated using Henry's law, using the methodology proposed by Gross et al., (1999) for Henry's constant estimation.

5.3.7. Control rules

In case of methanol strategies, the growth rate is directly controlled by the μ-feed-forward control in MLFB cultures, and MeOH control is programmed in MNLFB cultures. The control laws of both strategies are described in Barrigón et al., (2013). In the OLFB operation mode, the growth rate is limited by the OTR, as follows:

$$\mu = \frac{k_L a(O_2^* - O_2)}{Y_{O2/X}^* \cdot X} \quad (22)$$

5.4. Results and discussion

5.4.1. Response time

The characteristic response time was experimentally established from dynamic response of the probes. The response time constant is defined as the time that the probe takes to reach 63.2% of its final value when it is exposed to a step change in dissolved oxygen concentration (Van't Riet 1979; Gourich et al., 2008; García-Ochoa and Gomez, 2009; Doran, 2013). The Low drift - AppliSens, InPro 6800 - Mettler Toledo, and Oxyferm XL - Hamiltorn dissolved oxygen probes were used to τ determination. The Figure 2 shows the Low Drift - AppliSens electrode response to the concentration step, from the ratio O_2/O_2^*=0 to1. The aim of measuring τ is to conclude if the response of the electrodes are enough fast to neglect their response time in k_La determination. The oxygen mass transfer coefficient determination becomes inaccurate when probes manifest large τ (Van't Riet 1979; Gourich et al., 2008; García-Ochoa and Gomez, 2009; Doran, 2013). The characteristic response time of each probe is reported in Table 4.

Table 4. Sterilizable polagraphic oxygen probe response times

Nº	DO Sensor model	Length / Diameter [mm]	τ ± sd [s]	Bioreactor
1	Low drift - AppliSens	340 / 12	15.0 ± 0.5	*Biobundle 3L*
2	Low drift - AppliSens	430 / 12	12.0 ± 0.2	*Biobundle 7L*
3	InPro 6800 - Mettler Toledo	320 / 12	15.0 ± 0.3	*Biostat B 4L*
4	InPro 6800 - Mettler Toledo	80 / 25	16.4 ± 0.3	*Biostat ED 5L*
5	Oxyferm XL - Hamilton	80 / 25	14.1 ± 0.4	*Biostat UD 50L*

The tested probes presented a fast response (τ=12-15s) with differences between responses of 20%. The electrodes responded fast for steam-sterilizable electrodes because

the response times of commercially steam-sterilizable electrodes are usually in the range of 10 to 100 s (Gourich et al., 2008; García-Ochoa and Gomez, 2009; Doran, 2013). The measured τ values are in agreement with the values predicted by the manufacturer: 8.5-13 s (Low Drift), <15s (Oxyferm XL), <22.5s (InPro 6800) (AppliSens, 2014; Hamilton, 2014; Mettler Toledo, 2014). However, care must be taken into account the time response from the manufacturer probe specifications; because the manufactures usually report the time to arrive at 90-99% of the signal. In Bellucci and Hamaker, (2011), τ of polarographic probe Oxyferm XL was calculated from the oxygen evolution, provoked by step decrease in tank pressure. The dissolved oxygen drop was of 15 % for τ measurements, obtaining not enough accurate results τ=30±12 s. The τ result was not in agreement with the time response constant by the manufacturer. The response time that the probe takes to arrive at 98% of the signal is 30-60 s (Hamilton, 2014).

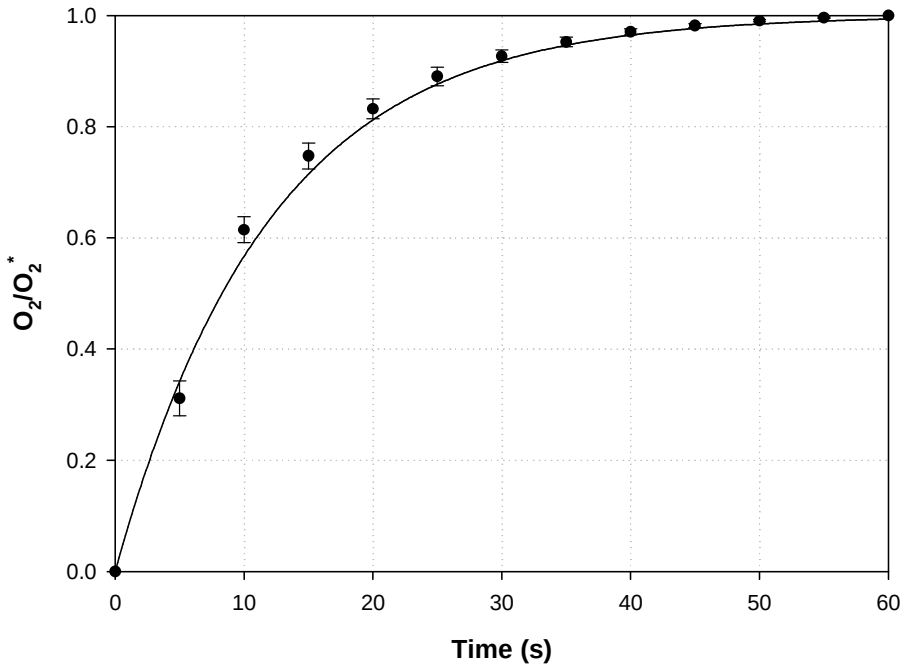

Figure 2. Dynamic response of the polarographic oxygen probe nº 2 (Low Drift - Applisens).

Additionally, there are faster non-autoclavable electrodes with response times of 2 to 3 s that respond more rapidly (García-Ochoa and Gomez, 2009; Doran, 2013). Nevertheless, non-autoclavable electrodes are not useless for *Pichia* cultivations.

A criterion for neglecting τ in k_La estimation, for reasonably accurate k_La values, is when k_La^{-1} is higher than 10τ. So, with our set of probes, τ can be neglected for $k_La < 0.0067\ s^{-1}$. In other case, the first-order response is included for k_La determination.

5.4.2. Comparison between the bioreactors' oxygen transfer capacity (k_La).

In order to characterize the transfer capacity of the bioreactors used for *Pichia* cultivations, k_La was measured in 4L *Biostat* B, 5L *Biostat* ED, 50L *Biostat* UD, 3L *Biobundle* and 7L *Biobundle* bioreactors under the same operational conditions: 300, 500, 700 and 1000 rpm; and aeration 0.25, 0.5, 0.75 and 1.0 vvm. The geometry and bioreactors dimensions are defined in the Table 1. The dynamic method was applied for k_La measurement. The probe time response was applied in k_La calculations (Eq. 9). The k_La results obtained for the 5 bioreactors are illustrated in the Figure 3. Results that correspond to intermediate aeration (0.5 and 0.75 vvm) are not shown.

The bioreactors showed similar mass transfer behaviour under the tested operational conditions (Figure 3). k_La could be modified by aeration and stirring. In general, k_La was increased by increasing the gas flow or the stirring rate. Stirring increases the fluid turbulence and it may affect the interfacial bubble area by bubble breakage. Aeration is

related to superficial gas rate, which is directly connected to interfacial area (Blanch and Clark, 1996; García-Ochoa and Gómez, 2010; Villadsen et al., 2011).

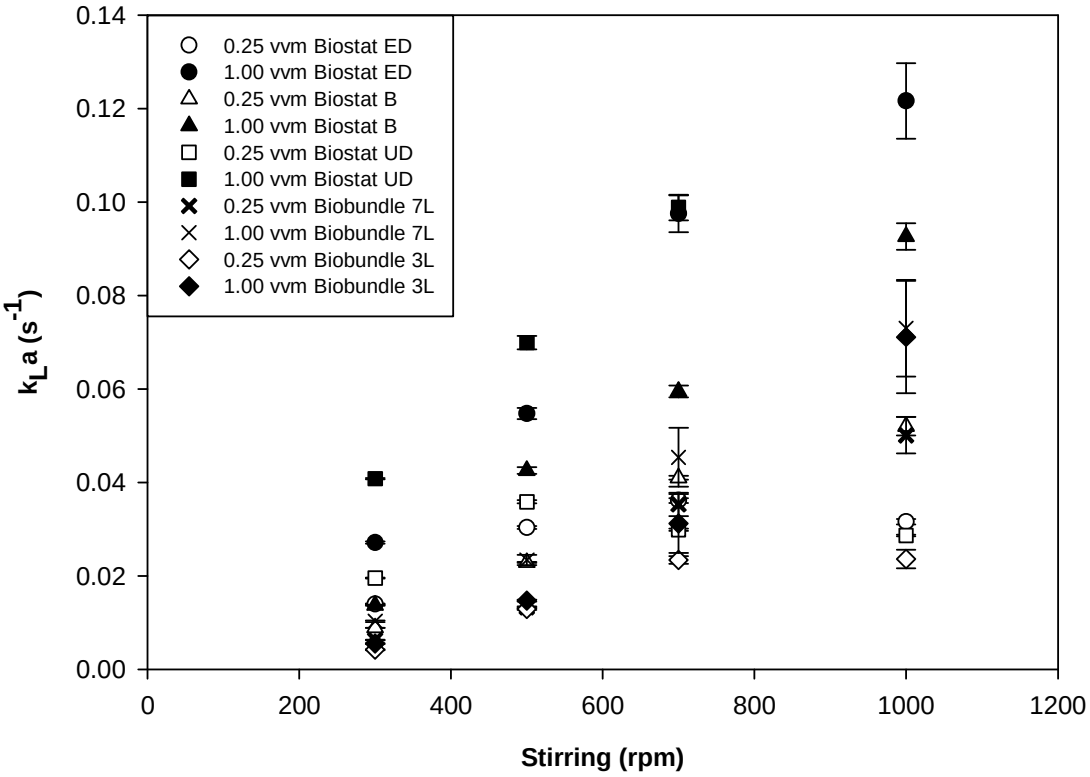

Figure 3. Comparison between k_La values in different operational conditions (bar error correspond to standard error).

Figure 3 shows that k_La depends linearly on stirring at high flow rates. However, linearity between k_La and stirring is not shown in the full stirring operational range, specifically at low aeration conditions. Biostat's (B, ED, UD) bioreactors presented better mass transfer performance than Biobundle's, ≥ 2 times higher. The best mass transfer capacity was shown by Biostat ED ($k_La = 0.12\ s^{-1}$; 1000rpm and 1.0 vvm). The poorest mass transfer capacity was measured in 3L Biobundle's ($k_La < 0.01\ s^{-1}$, 300 rpm and 0.25 vvm).

Bioreactor geometry, impeller distribution, liquid and gas turbulence are the basis for a good mass transfer capacity (Fujisová et al., 2007; Martín et al., 2008; Xia et al., 2009; Villadsen et al., 2011). Differences between geometry design ($H_T/T \approx 3.0$ in ED/ UD vs. $H_T/T \approx 2.0$ in Biobundle's), number of dipped impellers (Ni=3 in ED/ UD vs. Ni=2.0 in Biobundle's) and number of bafflers (nb=4 in ED/ UD vs. nb=3 in Biobundle's) may be the reason for large differences in mass transfer capacity. If Biobundle's impeller and baffler setup were similar to Biostats, the k_La could increase. Volumetric mass transfer coefficient was measured in a pilot-plant bioreactor ($H_T/T \approx 2.6$) with a single-, double- and triple-impeller setup in Moucha et al., (2009). In this case, the triple-impeller configurations gave the best mass transfer performance. Oxygen mass transfer characteristics for various twin and single-impeller systems were also investigated in Karimi et al., (2013). k_La enhancement was found depending on the impeller configuration. So, increasing only the impeller number could be improved k_La in Biobunddle's system.

5.4.3. Mass transfer correlation.

The empirical correlation proposed by Van't Riet, (1979) has been studied for k_La prediction during fermentations. k_La can be estimated as a function of the superficial gas velocity, and the specific power. Equation (5) was fitted to the k_La data by least squares regression. Pg and v_s was calculated under a wide operational range (300-1000 rpm; 0.25-3.0 vvm) in Biostat B, Biostat ED, 3L Biobundle and 7L Biobundle bioreactors. The constant and exponents (k_1, α, and β) of the correlation are presented in Table 5.

Table 5. Coefficients and parameters for k_La Van't Riet's correlation.

Bioreactor	k_1	α	β
Biobundle 7L	$5.5\ 10^{-3}$	0.61	0.42
Biostat B	$1.6\ 10^{-2*}$	0.37*	0.31*
Biostat ED	$8.3\ 10^{-2}$	0.42	0.70
Biobundle 3L	$5.3\ 10^{-2}$	0.70	0.83

*Data obtained from Ferrer, 2007.

The exponents obtained for each biosystem are in concordance with the literature. Typical values for exponents are ranged in $0.3 \leq \alpha \leq 0.7$ and $0.4 \leq \beta \leq 1$ (García-Ochoa and Gómez, 2009; Villadsen et al., 2011).

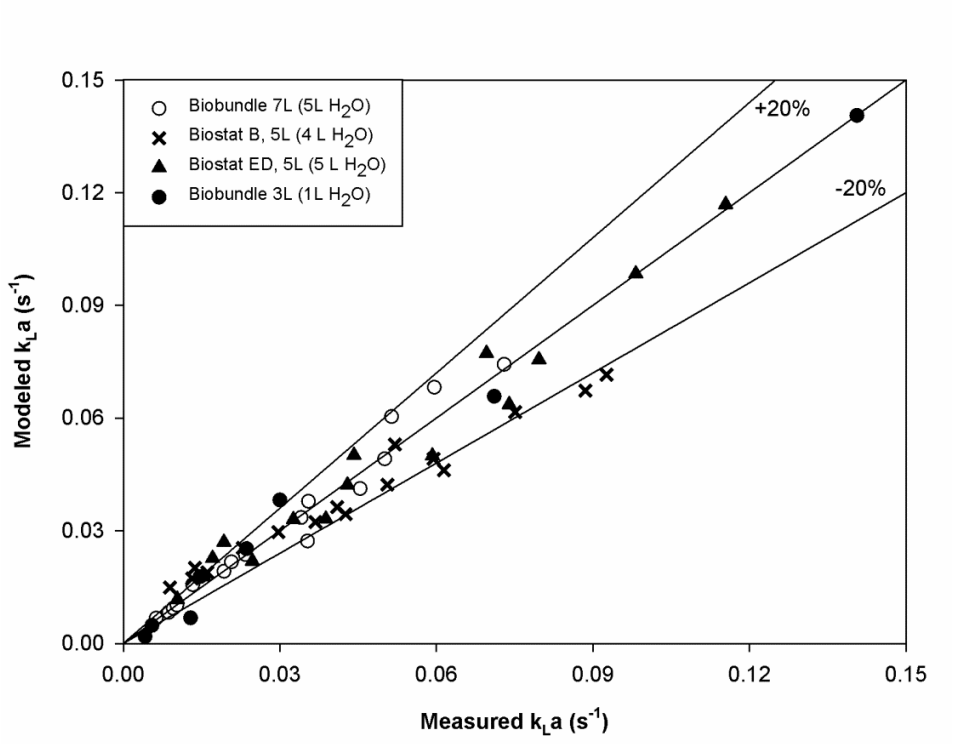

Figura 4. k_La estimation precision in Biobundle 3L, Biobundle 7L, Biostat B and Biostat ED bioreactors.

Figure 4 shows a linear plot of the experimental k_La values versus the modelled values. Precision of prediction is about ±20% of the measurement. The range of k_La measurements and prediction (0.01 - 0.14 s^{-1}) are in agreement with other systems (Marques et al., 2009; Moucha et al., 2009; Albaek et al., 2011). Nonetheless, other authors reported lower k_La values in their systems (Liang and Yuan, 2007; Bellucci and Hamaker, 2011). Low k_La values may be related to more restrictive operational conditions.

Furthermore, some authors have included more variables in their mass transfer model, such as the viscosity (μ_v) or the enhancement factor (E), in order to improve the model precision (Çalik et al., 2010, Albaek et al., 2011). In this work, it is considered that the proposed k_La correlations are accurate enough to be applied in *Pichia*'s bioprocess modelling. In any case, models should be kept as simple as possible and only be made more elaborate when it is required (Villadsen et al., 2011; Levenspiel, 2002).

5.4.4 Mass transfer model validation

MLFB culture was carried out with *P. pastoris* Mut$^+$ phenotype at $\mu_{set\text{-}point}$=0.020 h^{-1} producing heterologous ROL, in order to validate the mass transfer model. Figure 5 shows the evolution of the fermentation process during the three phases (GBP, TP and MIP). Biomass, substrate concentration (glycerol and methanol), dissolved oxygen concentration (%), lipase activity was displayed in Figure 5a. Methanol and dissolved oxygen concentrations were measured online. Biomass, glycerol and product activity was

measured off-line from the sampling and extrapolated to whole fermentation by splines (Barrigon et al., 2013).

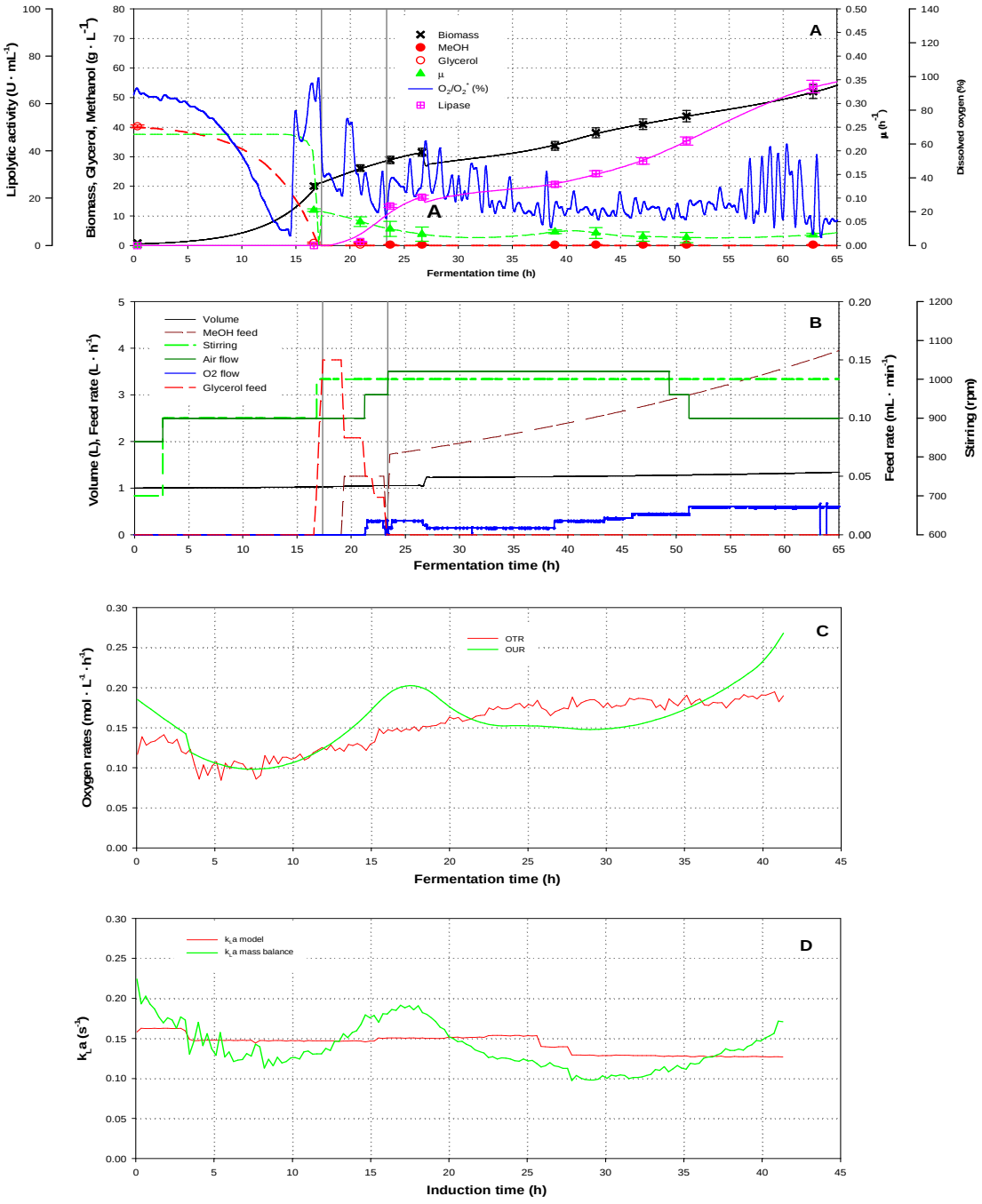

Figure 5. **a**. Evolution of state variables (X, S, P, O_2) and specific rate (μ) during MLFB culture. **b**. Operational conditions (stirring and flow rates), volume and feeding rates. **c**. Oxygen rates evolution during MIP. **d**. k_La prediction and estimation during MIP.

Operational conditions: stirring, flow rates, volume and feeding are presented in figure 5b. During the batch phase, biomass grew exponentially at μmax, the stirring and the air flow were kept approximately constant, dissolved oxygen decreased exponentially. During the transition phase, the oxygen consumption was increased by the methanol feed. Air enriched in oxygen was supplied. Finally, methanol was fed as the sole carbon source and also the inducer during the MIP. This production phase was carried out by using MLFB strategy, by a pre-programmed exponentially increasing methanol feeding rate at $\mu_{set\text{-}point}$=0.020 h^{-1}. The stirring and air flow were increased in order to maintain the dissolved oxygen concentration up to 20% of oxygen saturation with air (O_2^*). Growth rate was kept fairly constant during the induction. Biomass productivity was 1.1 g·DCW h^{-1}, volumetric productivity $1.1 \cdot 10^3$ $U \cdot L^{-1} \cdot h^{-1}$ and q_p=35 $U \cdot g^{-1} \cdot h^{-1}$. These results are in agreement with the MLFB presented in Barrigon et al., (2013)

The oxygen rates, OUR and OTR were compared during the MIP in Figure 5c. OTR was calculated from dissolved oxygen measurement and by applying the Van't Riet's correlation for k_La estimation. The mixture of air enriched in oxygen was taken into account to calculate the O_2* (Gross et al., 1999). Although k_La is typically reported in literature, OTR was studied because OUR may become the limiting rate (Bellucci and Hamaker, 2011). OUR and OTR presented an evolution quite similar with low deviations between them during the induction phase. However, large deviations between them were found during GBP and TP. The OTR maximal value was approximately 0.2 mol $L^{-1} \cdot h^{-1}$. In the literature, OTR_{max} about 0.05 mol $g^{-1} \cdot h^{-1}$ in a batch culture can be found (Bellucci

and Hamaker, 2011) and 0.16 - 0.18 mol·L^{-1}·h^{-1} in MLFB fedbatch cultivations of *P. pastoris* (Jenzsch et al., 2004, Liang and Yuan, 2007).

Finally, the mass transfer model was tested in figure 5d. Comparison between k_La experimentally determined, by the mass balance Eq (20), with the modelled k_La, by the Vant Riet's correlation (Eq 5) was plotted during the induction phase.

Table 6. k_La determination during the induction phase.

Parameter	Units	Value
k_La from model	s^{-1}	0.143±0.023
k_La oxygen mass balance	s^{-1}	0.139±0.056
MRE	%	15.4

The k_La predicted by the model, provided similarly accurate to the experimentally determinate k_La, with a mean relative deviation of 15% during the induction phase (Table 6). Large errors for k_La estimation were obtained during the batch and transition phase (data not shown). The estimated k_La by oxygen balance during GBP and TP may not be accurate. Although, dissolved oxygen was not accumulated and OTR was not required to achieve OTR_{max} value, OUR was 1.8-times lower than OTR. So, the oxygen consumption model for glycerol substrate was not as precise as required. The oxygen model of glycerol was extrapolated from glycerol limited fed-batch (Barrigon et al 2012), so at high growth rate it may not be precise enough. However, the k_La model presented a good fitting when the system was operated at OTR_{max} in MLFB, showing only few deviations between model prediction and mass balance estimation.

5.4.5. Alternative strategies based on oxygen transfer model

In this section, alternative operational modes (OLFB strategies) based on the oxygen transfer model of the Biobundle 3L bioreactor have been simulated for the production of the recombinant *Rhizopus oryzae* lipase (ROL) expressed in *P. pastoris* Mut$^+$ phenotype expression system under oxygen restrictive conditions but without methanol limitation during the MIP. Two OTR set-points were selected for OLFB simulations: $OTR_{set\text{-}point}$ = 0.135 and 0.232 $mol \cdot L^{-1} \cdot h^{-1}$, at $MeOH_{set\text{-}point}$ = 3.0 $g \cdot L^{-1}$. The set-points values were selected because the maximal oxygen transfer rate of the system is 0.135 $mol \cdot L^{-1} \cdot h^{-1}$ by air supplying, without enriching the air with pure oxygen; the OUR value of *Pichi*a's expressing ROL at the beginning of MIP is 0.232 $mol \cdot L^{-1} \cdot h^{-1}$ in non-limited conditions; and MeOH=3.0 $g \cdot L^{-1}$ is the most productive inductor concentration condition for this strain. The OLFB strategies have been compared to the most representative strategies based on methanol as the limiting substrate and tested in Barrigon et al., (2013): MLFB-$\mu_{set\text{-}point}$ = 0.015 h^{-1} and MLFB-$MeOH_{set\text{-}point}$ = 3.0 $g \cdot L^{-1}$

The simulations of total biomass (XV), the total product (PV), μ and q_P time course are illustrated for all strategies in Figure 6. Additionally, a summary about the predicted maximal lipolytic activity, yield, mean specific rates and the volumetric productivities for MLFB, MNLFB and OLFB strategies are presented in Table 7.

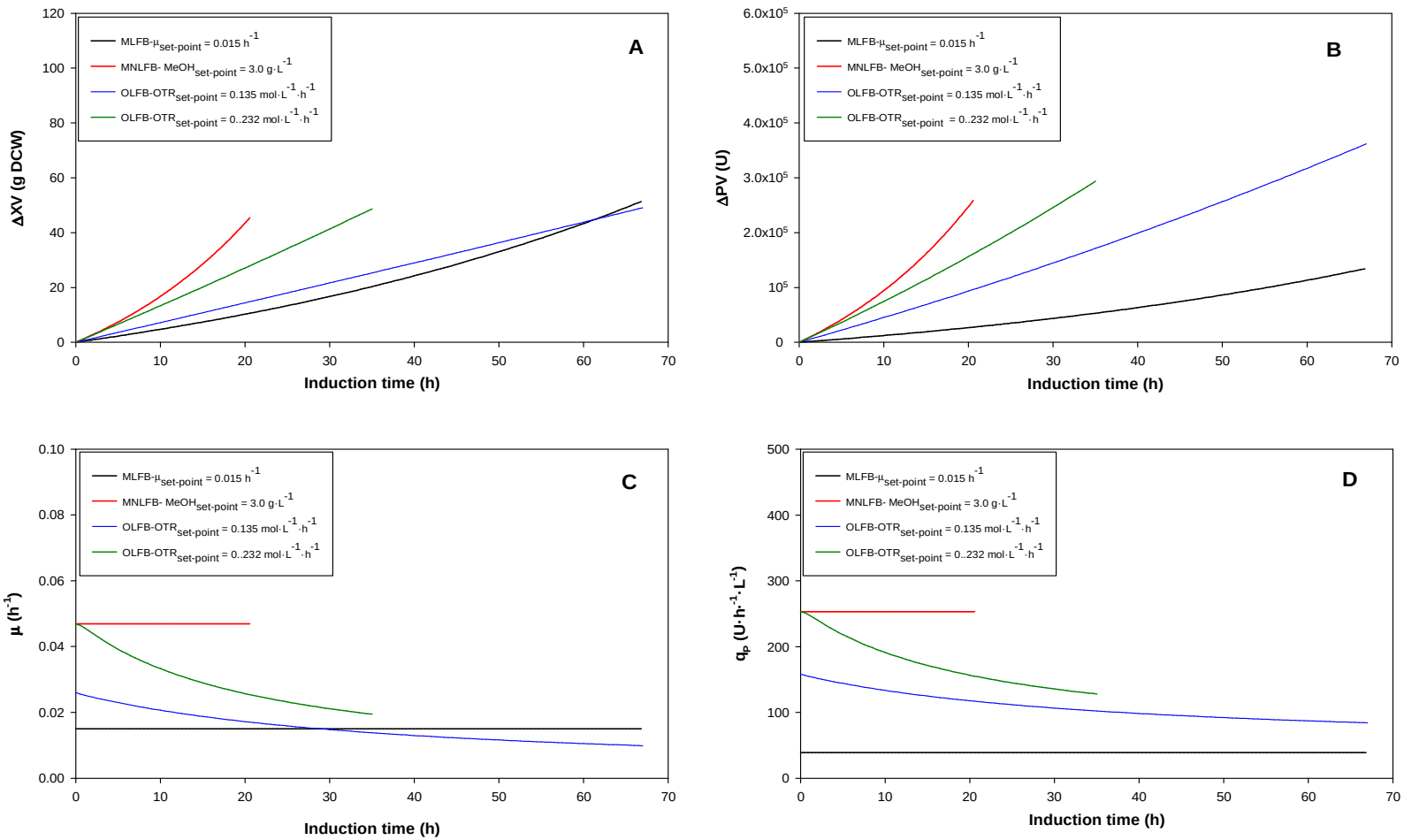

Figure 6. Simulation of *Pichia pastoris* bioprocess model: state variables (*XV* and *PV*) over the induction time. Time course of **(a)** total biomass and **(b)** total lipolytic activity. **(c)** specific growth rate evolution and **(d)** specific production rate evolution. Set-points are indicated by μ values for MLFB mode, methanol concentrations for MNLFB mode and maximal OTR for OLFB modes.

In Figure 6, different evolution patterns of XV, PV, μ and q_P are predicted during the MIP for the selected operational modes. For example, in MLFB and MNLFB, where the methanol is the limiting substrate, μ and q_P are constant. In contrast, decreasing profiles of the specific rates are shown by the OLFBs. Moreover, different production levels of XV and PV are observed. The dissimilarities between XV and PV can be explained by specific rates profiles.

Table 7. Comparison of process variables, yields, productivities and specific rates using different operational strategies during the induction phase in *Pichia pastoris* cultures expressing ROL.

Strategy	**Set-point**	**Fermentation time [h]**	**PV[b] [U]**	**$Y_{P/X}$[b] [U · g_X^{-1}]**	**μ[a] [h^{-1}]**	**q_P[a] [U·g^{-1}·h^{-1}]**	**Volumetric productivity[b] [U·L^{-1}·h^{-1}]**
MNLFB	MeOH = 3.0 [g·L^{-1}]	46.1	$2.8 \cdot 10_5$	3610	0.047	253.4	5255
OLFB	OTR =0.232 [mol·L^{-1}·h^{-1}]	60.8	$3.1 \cdot 10_5$	4090	0.029	172.8	4556
OLFB	OTR =0.135 [mol·L^{-1}·h^{-1}]	91.8	$3.8 \cdot 10_5$	4954	0.015	109.0	3599
MLFB	μ = 0.015 [h^{-1}]	91.6	$1.5 \cdot 10_5$	1942	0.015	39.3	769

[a] Estimated during the MIP.

[b] Estimated during all fermentation.

In Figure 6a, the biomass formation time course is plotted for all strategies. The stop condition in the simulated fermentations was that the biomass concentration (X) reached 60 g·L^{-1}, which correspond to 75-80 g of total biomass (XV). The biomass production is related to the growth rate. MNLFB strategy is the shortest strategy with the highest μ_{mean} (Table 7). On the other hand, the MLFB and OLFB-OTR_{max}=0.135 mol·L^{-1}·h^{-1} operational modes are shown as the longest strategies, 92 h to achieve the stop condition,

with the lowest μ_{mean} values (0.015 h^{-1}). OLFB-OTR_{max}=0.232 $mol \cdot L^{-1} \cdot h^{-1}$ reveals a μ_{mean}=0.029 h^{-1}. The MNLFB is about 2-times faster than MLFB and the slow OLFB, and about 1.3-times than the fast OLFB, refereed to all fermentation time (Table 7). Although a constant μ is characterized in MLFB and MNLFB cultures, OLFB strategies describe μ-decreasing profiles, where the μ_{max} is shown at MIP initial time (Figure 6c). The μ-decreasing profiles are explained by Eq. 22, because OTR keeps constant during all MIP and the biomass is in continuous proliferation.

A prediction of product evolution is presented for all strategies in Figure 6b. The model predicts that the ROL production level with the fastest strategy (MNLFB) is about $3.8 \cdot 10^5$ U. In contrast to biomass profiles, the OLFBs strategies predict higher product titers than MNLFB and MLFB. Slow OLFB increases the total production 1.4-fold related to MNLFB and the fast OLFB 1.1-fold to MNLFB. On the other hand, the ranking of operational strategies in terms of q_P is shown in Figure 6d, where MNNLB is the best strategy ($q_{P,mean}$= 253.4 $U \cdot g^{-1} \cdot h^{-1}$), MLFB reveals a low production rate ($q_{P,mean}$= 39.3 $U \cdot g^{-1} \cdot h^{-1}$) and OLFBs intermediate q_P values ($q_{P,mean}$= 172.8 and 109.0 $U \cdot g^{-1} \cdot h^{-1}$). The MNLFB and OLFBs strategies were designed with residual methanol concentration above 2 $g \cdot L^{-1}$, so, μ-lineal dependence is assumed by q_P (Eq. 19). So, OLFB strategies describe q_P-decreasing profiles due to the linear relation to the μ-decreasing profiles. Additionally, q_P linear dependence to μ is the reason that the OLFB reaches high product titers than MNLFB. The maximal of global product yield ($Y_{P/X}$*) is obtained at low μ, whenever the Lueddeking-Piret relationship can be considered (Eq. 19). Thus, in MLFB,

when the production is assumed that is not coupled to growth (Eq 18), a basal specific production rate and also a low production level is obtained.

In terms of volumetric productivity, the MNLFB strategy is shown as the most productive (Table 7). The shortest fermentation time and similar ROL production level as OLFB are the main reasons for the highest volumetric productivity. On the contrary, the longer fermentation time and the lower production levels of MLFB results in a low volumetric productivity. The productivities of OLFB are in concordance to the specific rates. In this way, the higher the OTR, the higher productivity is obtained.

5.5. Conclusions

k_La was measured in Biostat B, Biostat ED, Biostat UD, Biobundle 3L and Biobundle 7L bioreactors with a polarographic probes by the dynamic method. The tested probes presented fast responses (τ=12-16s) for steam-sterilizable electrodes. Bioreactors oxygen transfer capacity was measured, and kla modelled by a Van't Riet's correlation type within the range 0.01-0.14 s^{-1}. Biostat's bioreactors showed a better oxygen transfer capacity than Biobundle's.

The Biobundle 3L bioreactor was used in a *Pichia*'s fermentation to validate the k_La correlation. Error of k_La estimates (MRE) was about 15% during the induction phase. So, the k_La model was considered accurate enough to use for exploitation.

Finally, different operational modes were simulated to predict the levels of ROL production in *P. pastoris* fermentation. The oxygen transfer model was used to define OLFB strategies. The OLFB cultures show good bioprocess performance and represent a great alternative to the carbon source limiting strategies. The longest OLFB strategy shows better results than MFLB, higher product yield and higher volumetric productivity. Although the higher OTR condition for OLFB operating mode enhances the bioprocess yields respect to the other OLFB and MLFB, the simulation predicts that MNLFB operational mode presents the best bioprocess performance. The OLFB simulations predict a production and productivity enhancement in the OLFB cultures at higher OTR. However, a high increase of OTR may provoke a change of the limiting substrate from oxygen to methanol, converting the OLFB in a MNLFB strategy.

5.6. References

1. Albaek MO, Gernaey KV, Hansen MS, Stocks SM. 2011. Modeling enzyme production with *Aspergillus oryzae* in pilot scale vessels with different agitation, aeration, and agitator types. Biotechnol Bioeng 108:1828-1840.
2. Amaral PFF, Freire MG, Rocha-Leao MHM, Marrucho IM, Coutinho JAP,Coelho MAZ. 2008. Optimization of oxygen mass transfer in a multiphase bioreactor with Perfluorodecalin as a second liquid phase. Biotechnol Bioeng 99: 588-598.
3. AppliSens. Dissolved oxygen sensors. Technical data sheet STS11, available online: http://ravesolution.com/fusiontekco/images/mp/APPK_Bioreactors.pdf (Accessed 08 October 2014).
4. 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.
5. Badino Jr AC, Facciotti MCR, Schmidell W. 2001. Volumetric oxygen transfer coefficients (k_La) in batch cultivations involving non-Newtonian broths. Biochem Eng J 8:111–119
6. Bailey JE, Ollis DF. 1986. Biochemical Engineering Fundamentals. McGraw-Hill, New York.
7. Barrigón JM, Ramon R, Rocha I, Valero F, Ferreira EC, Montesinos JL. 2012. State and specific growth estimation in heterologous protein production by *Pichia pastoris*. AIChE J 58:2967-2979.
8. 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.
9. Bellucci JJ, Hamaker KH. 2011. Evaluation of oxygen transfer rates in stirred-tank bioreactors for clinical manufacturing. Biotechnol Prog 27:368-376
10. Berdichevsky M, d'Anjou M, Mallem MR, Shaikh SS, Potgieter TI. 2011. Improved production of monoclonal antibodies through oxygen-limited cultivation of glycoengineered yeast. J Biotechnol 155:217–224.

11. Blanch HW, Clark DS. 1996. Biochemical engineering. New York: Marcel Dekker. 702 p.
12. Çelik E, Çalik P, Oliver SG. 2009. A structured kinetic model for recombinant protein production by Mut$^+$ strain of *Pichia pastoris*. Chem Eng Sci 64:5028-5035.
13. Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. Fems Microbiol Rev 24:45-66.
14. Cooper CM, Fernstorm GA Miller SA. 1944. Performance of agitated gas-liquid contactors. Ind Eng Chem 36:504-509.
15. Cos O, Resina D, Ferrer P, Montesinos JL, Valero F. 2005a. Heterologous production of *Rhizopus oryzae* lipase in *Pichia pastoris* using the alcohol oxidase and formaldehyde dehydrogenase promoters in batch and fed-batch cultures. Biochem Eng J 26:86-94.
16. Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005b. 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.
17. Cos O, Ramon R, Montesinos JL, 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.
18. Çalık P, İnankur B, Soyaslan ES, Şahin M, Taşpınar H, Açık E, Bayraktar E. 2010. Fermentation and oxygen transfer characteristics in recombinant human growth hormone production by *Pichia pastoris* in sorbitol batch and methanol fed-batch operation. J Chem Technol Biotechnol 85: 226–233
19. Cregg JM. 2007. Methods in Molecular Biology: *Pichia* Protocols 2nd ed. Totowa. New Jersey: Humana Press.
20. Danckwerts PV. 1970. Gas-liquid reaction. New York: McGraw-Hill 239-250.
21. Doran PM. 2013. Bioprocess Engineering Principles, 2nd ed. San Diego: Academic Press 379-444.
22. Ferrer J. 2008. "Estudio de la transferencia de oxígeno en biorreactores utilizados para la producción de proteínas heterólogas en *Pichia pastoris*". *Msc thesis*, Universitat Autònoma de Barcelona, Bellaterra.

23. Fyferling M, Uribelarrea JL, Goma U, Molina-Jouve C. 2008. Oxygen transfer in intensive microbial culture. Bioprocess Biosyst Eng 31:595–604.
24. Fujasová M, Linek V, Maucha T. 2007. Mass transfer correlations for multiple-impeller gas–liquid contactors. Analysis of the effect of axial dispersion in gas and liquid phases on "local" $k_L a$ values measured by the dynamic pressure method in individual stages of the vessel. Chem Eng Sci 62:1650–1669.
25. Galaction AI, Cascaval D, Oniscu C, Turnea M. 2004. Enhancement of oxygen mass transfer in stirred bioreactors using oxygen-vectors. 1. Simulated fermentation broths. Bioprocess Biosyst Eng 26: 231–238.
26. Garcia-Ochoa F, Gomez E. 1998. Mass transfer coefficient in stirred tank reactors for xanthan solutions. Biochem Eng J 1:1–10.
27. Garcia-Ochoa F, Gomez E. 2009. Bioreactor scale-up and oxygen transfer rate in microbial processes: An overview. Biotechnol Adv 27:153–176.
28. Garcia-Ochoa F, Gomez E, Santos VE, Merchuk JC. 2010. Oxygen uptake rate in microbial processes: An overview. Biochem Eng J 49:289–307.
29. Gogate PR, Beenackers AACM, Pandit AB. 2000. Multiple-impeller systems with a special emphasis on bioreactors: a critical review. Biochem Eng J 6:109-144.
30. Gros JB, Dussap CG, Catte M. 1999. Estimation of O_2 and CO_2 solubility in microbial culture media. Biotechnol Prog 15:923-927.
31. Gourich B, Vial Ch, El Azher N, Belhaj Soulami M, Ziyad M. 2008. Influence of hydrodynamics and probe response on oxygen mass transfer measurements in a high aspect ratio bubble column reactior: Effect of the coalescence behaviour of the liquid phase. Biochem Eng J 39:1-14.
32. Hanson MA, Brorson KA, Moreira AR, Rao G. 2009. Comparisons of optically monitored small-scale stirred tank vessels to optically controlled disposable bag bioreactors. Microbial Cell Factories 8:44.
33. Hamilton. Process sensors. pH - Oxygen - Conductivity, available online: http://www.hamiltoncompany.com/downloads/Manual_OxyFermO2_105x297.pdf (Accessed 01 December 2014).
34. Hughmark GA. 1980. Power requirements and interfacial area in gas–liquid turbine agitated systems. Ind Eng Chem Proc Des Dev 19:638–641.

35. Invitrogen. *Pichia* Expression Kit. cat. no. K1710-01 Version A.0, available online: http://tools.lifetechnologies.com/content/sfs/manuals/pich_man.pdf (Accessed 10 December 2014).
36. Jenzsch M, Lange M, Bar J, Rahfeld JU, Lübbert A. 2004. Bioreactor retrofitting to avoid aeration with oxygen in *Pichia pastoris* cultivation processes for recombinant protein production. Chem Eng Res Des 82:1144–1152.
37. Jenzsch M, Simutis R, Lübbert A. 2006. Optimization and Control of Industrial Microbial Cultivation Processes. Eng Life Sci 6:117–124
38. Ju LK, Sundarajan A. 1992. Model analysis of biological oxygen transfer enhancement in surface-aerated bioreactors. Biotechnol Bioeng 40:1343-1352.
39. Karimi A, Golbabaei F, Mehrnia MR, Neghab M, Mohammad K, Nikpey A, Pourmand MR. 2013. Oxygen mass transfer in a stirred tank bioreactor using different impeller configurations for environmental purposes. Iranian J Environ Health Sci Eng 10:6.
40. Khatri NK, Hoffmann F. 2006b. Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant *Pichia pastoris*. Biotechnol Bioeng 93:871-879.
41. Khatri NK, Hoffmann F. 2006a. Oxygen-limited control of methanol uptake for improved production of a single-chain antibody fragment with recombinant *Pichia pastoris*. Appl Microbiol Biotechnol 72:492–48.
42. Kirk TV, Szita N. 2012. Oxygen transfer characteristics of miniaturized bioreactor systems. Biotechnol Bioeng 110:1005-1019
43. Liang J, Yuan J. 2007 Oxygen transfer model in recombinant Pichia pastoris and its application in biomass estimation. Biotechnol Lett 29:27–35
44. Levenspiel O. 2002. Modeling in chemical engineering. Chem Eng Sci 75:4691-4696.
45. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* system. Yeast 22:249-270.
46. Marques DAV, Torres BR, Porto ALF, Pessoa-Júnior A, Converti A. 2009. Comparison of oxygen mass transfer coefficient in simple and extractive fermentation systems Biochem Eng J 47:122-126.

47. Martín M, Montes FJ, Galán. 2008. Bubbling process in stirred tank reactors II: Agitator effect on the mass transfer rates. Chem Eng Sci 63:3223-3234.
48. Mettler Toledo. InPro6800/ InPro6900. The new O2 sensor family, available online: http://us.mt.com/us/en/home/supportive_content/product_documentation/technical_specificiations/TD_O2_sensors_InPro6XXX_series/_jcr_content/download/file/file.res/TD_DO_InPro69x0_6800_6050_EN_52200954_July2012.pdf (Accessed 01 December 2014)
49. Moucha T, Linek V, Erokhin K, Rejl JF, Fujasová M. 2009. Improved power and mass transfer correlations for design and scale-up of multi-impeller gas–liquid contactors. Chem Eng Sci 64:598-604
50. Nagata, S. 1975. Mixing Principles and Applications. John Wiley and Sons, New York. p126
51. Nielsen J. 1996. Modelling the morphology of filamentous microorganisms. Trends Biotechnol 14:438-443.
52. Patel N, Thibault J. 2009. Enhanced in situ dynamic method for measuring K_La in fermentation media. Biochem Eng J 47:48–54
53. Oliveira R, Simutis R, Feyo de Azevedo S. 2004. Design of a stable adaptive controller for driving aerobic fermentation processes near maximum oxygen transfer capacity. 14:617-626
54. Oliveira R, Clemente JJ, Cunha AE, Carrondoa MJT. 2005. Adaptive dissolved oxygen control through the glycerol feeding in a recombinant *Pichia pastoris* cultivation in conditions of oxygen transfer limitation. J Biotechnol 116:35-50
55. Pérez J, Montesinos JL, Gòdia F. 2006. Gas–liquid mass transfer in an up-flow cocurrent packed-bed biofilm reactor. Biochem Eng J 31:188-196.
56. Potgieter TI, Kersey SD, Mallem MR, Nylen AC, d'Anjou M. 2010. Antibody expression kinetics in glycoengineered *Pichia pastoris*. Biotechnol Bioeng 106:918-927.
57. Pinelli D, Liu Z, Magelli F. 2010. Analysis of KLa Measurement Methods in Stirred Vessels: The Role of Experimental Techniques and Fluid Dynamic Models. International Journal of Chemical Reactor Engineering 8:A115

58. 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.
59. Puthli MS, Rathod VK, Pandit AB. 2005. Gas–liquid mass transfer studies with triple impeller system on a laboratory scale bioreactor. Biochem Eng J 23:25–30.
60. Sreekrishna K. 2010. *Pichia*, Optimization of protein expression. In: Flickinger MC, editor. Encyclopedia of Industrial Biotechnology: Bioprocess, Bioseparation, and Cell Technology. New York: John Wiley & Sons Inc. p 1-16.
61. Tribe LA, Briens CL, Margaritis A. 1995. Determination of the volumetric mass transfer coefficient (k_La) Using the dynamic "gas oOut-gas in" method: analysis of errors caused by dissolved oxygen probes Biotecnol Bioneng 46:388-392
62. Valero F. 2013. Bioprocess engineering of *Pichia pastoris*, an exciting host eukaryotic cell expression system. Protein engineering - technology and application. Ed Dr. Tomohisa Ogawa
63. Van't Riet. 1979. Review of measuring methods and results in nonviscous gas-liquid mass transfer in stirred vessels. Ind Eng Chem Process Des Dev 18:357-364.
64. Villadsen J, Nielsen J, Lidén G. 2011. Bioreaction Engineering Principles 3rd ed. New York: Springer Science+Business Media.
65. Weisenberger S, Schumple A. 1996. Estimation of gas solubilities in salt solutions at temperatures from 273 K to 363 K. AIChE J 42:298-300
66. Wu JM, Fu WC. 2012. Intracellular co-expression of Vitreoscilla hemoglobin enhances cell performance and β-galactosidase production in *Pichia pastoris*. J Biosci Bioeng 113:332–337
67. Xia JY, Zhang YH, Zhang SL, Cheng N, Yin P, Zhuang YP, Chu J. 2009. Fluid dynamics investigation of variant impeller combinations by simulation and fermentation experiment. Biochem Eng J 43:252–260.
68. Xie J, Yang R, Zhou Q, Du P, Gan R, Ye Q. 2013. Efficiencies of growth and angiostatin expression in cultures of Pichia pastoris fed with mixed carbon sources. Chem Biochem Eng Q 27:235–244
69. Yano T, Takigami E, Yurimoto H, Sakai Y. 2009. Yap1-regulated glutathione redox system curtails accumulation of formaldehyde and reactive oxygen species in methanol metabolism of *Pichia pastoris*. Eukaryotic Cell 8:540–549.

70. Yurimoto H, Oku M, Sakai Y. 2011. Yeast methylotrophy: metabolism, gene regulation and peroxisome homeostasis. Int J Microbiol Res 101298:8
71. Zhang W, Inan M, Meagher MM . 2007. Rational design and optimization of fed-batch and continuous fermentations. Methods Mol Biol 389:43-64.

Chapter 6:

Conclusions

6. Conclusions

The heterologous ROL production in *P. pastoris* P_{AOX1} (Mut^+) fed-batch cultures has been studied for bioprocess monitoring and control, kinetic modelling, design and development of operational strategies.

First, the estimation of the state variables, biomass (X) and substrate (S), and the specific growth rate (μ) have been evaluated for heterologous protein production in *P. pastoris* cultures by two non-linear observers (*NLOBE* and *AO-SODE)* and a linear estimator (*RLS-VFF*).

Simulation results obtained for the *NLOBE's* show adequate global performance, but it has been established as the most tuning sensitive method, being rather dependent on their tuning parameters and initial values. Consequently, errors on model coefficients may produce inaccurate results in estimation of μ, and amplified on X and S estimation.

The μ has been correctly estimated by applying the asymptotic observers (*AO*), as well as recursive least square (*RLS*) methods. *AO-SODE* estimator shows a better estimation performance than the *RLS-VFF* methods, concretely, when rapid and moderate changes of μ appear because model parameters are well known. On the other hand, when slow changes on the specific growth rate are presented in the bioprocess, *RLS-VFF* comes up as the best option, due to its reduced requirements. In addition, biomass and substrate are also satisfactorily predicted with both methods.

The most efficient method to estimate the μ and the state variable of system is the *AO-SODE (O_2, OUR)*. The presented estimation methodology can be used for real-time monitoring of the key fermentation variables.

The MLFB and MNLFB operational strategies have been successfully applied for ROL production in the *P. pastoris* cell factory. MLFB is one of the easiest fed-batch strategies to be implemented from an operational industrial point of view. However, low ROL total lipolytic activity obtained leads us to reject this approach. In the MNFB, it has been observed that the methanol concentration is as key parameter to maximize protein production. MNLFB operation at constant methanol concentration higher than 2 $g \cdot L^{-1}$ is necessary to maximize ROL production

Highest total ROL production, $Y_{P/X,}$ global volumetric and specific productivity, q_p and productivities has been obtained at methanol concentration set-point of 3 $g \cdot L^{-1}$ in MNLFB culture. Lower methanol concentrations, up to 2 $g \cdot L^{-1}$, results in low product accumulation, whereas at higher methanol concentration tested (10 $g \cdot L^{-1}$) than 6 $g \cdot L^{-1}$ significant inhibition on growth is observed.

Heterologous ROL production by a *P. pastoris* P_{AOX1}-based system can be accurately described with a macrokinetic model. Two different models has been evaluated. Although the two models describe specific rates with great accuracy, the model that uses a non-monotonic substrate function for growth, Pirt's model for substrate uptake and Luedeking-Piret equation for protein production performs better in an overall validation for the entire operational range.

A comparative analysis of different target protein production by *P. pastoris* under *AOX1* promoter reveales: protein production has a direct effect on growth and the typical non-monotonic nature of growth kinetics is usually shown; a linear q_S-μ behaviour; and production differed significantly between proteins but often a proportionally decreasing relationship between q_P and μ is observed. So, these results imply that yields and productivities can be maximized at μ where the q_P is not maximal. Usually, maximum yields and productivities will be obtained at methanol concentration range 2-4 $g \cdot L^{-1}$ and μ about 0.02-0.06 h^{-1}.

k_La has been measured in Biostat B, Biostat ED, Biostat UD, Biobundle 3L and Biobundle 7L bioreactors with a polarographic probes by the dynamic method. Moreover, k_La has been modelled by a van Riet's correlation and validated to apply in *Pichia*'s fermentation.

The oxygen transfer model has been used to define OLFB strategies. The MLFB, MNLFB and OLFB operational modes has been simulated to predict the levels of ROL production in *P. pastoris* fermentation. The OLFB cultures show good bioprocess performance and represent a great alternative to the carbon source limiting strategies. Although the OLFB operating mode enhances the bioprocess yields respect to MLFB, the simulation predicts that MNLFB operational mode presents the best bioprocess performance.

Chapter 7:

Annex

7. Annex

7.1. Mass balance assumptions

The bioprocess described in all chapters is based on the reaction involved in the oxidative uptake of substrates (glycerol, methanol or sorbitol) to form biomass and products in the liquid phase. Equations for cell growth (1); protein production (2) and maintenance (3) can be stated as:

$$k_1 S + k_2 O_2 \xrightarrow{r_X} X + k_3 CO_2 \quad (1)$$

$$k_4 S + k_5 O_2 \xrightarrow{r_P} k_3 P + k_6 CO_2 \quad (2)$$

$$S + k_7 O_2 \xrightarrow{r_m} k_8 CO_2 \quad (3)$$

where *X, S, O_2, CO_2,* and *P* represent biomass, substrate, dissolved oxygen, dissolved carbon dioxide and product respectively (in the sequel, the same symbols are used to represent component concentrations); r_x, r_p, r_m are the reaction rates; k_i are the yield (stoichiometric) coefficients.

The volume variation can be obtained by the total mass balance for an ideal stirred tank reactor in fed-batch operation, as follows:

$$\frac{dV}{dt} = \frac{\rho_{Feed} F - \rho_{H2O} F_{Evap} + \rho_{Base} F_{Base} - \rho_{Broth} F_O + M_{GAS}}{\rho_{Broth}} \quad (4)$$

where V is the volume of broth in the reactor [L], F the volumetric feeding rate [$L \cdot h^{-1}$], F_{Evap} the water evaporation rate [$L \cdot h^{-1}$], F_{Base} the base feeding rate [$L \cdot h^{-1}$], F_O the withdrawal rate [$L \cdot h^{-1}$], M_{GAS} net mass gas flow rate [$g \cdot h^{-1}$], ρ_{Feed} substrate feed density

[$g \cdot L^{-1}$], ρ_{H2O} water density [$g \cdot L^{-1}$], ρ_{Base} base density [$g \cdot L^{-1}$], ρ_{Broth} mean broth density [$g \cdot L^{-1}$]. The net mass gas flow rate is calculated with the equation (5):

$$M_{GAS} = W_{O2}\, OUR - W_{CO2}\, CPR \quad (5)$$

where OUR is the oxygen uptake rate [$mol \cdot h^{-1}$], CPR carbon dioxide production rate [$mol \cdot h^{-1}$], W_{O2} oxygen molar mass [$g \cdot mol^{-1}$], W_{CO2} carbon dioxide molar mass [$g \cdot mol^{-1}$].

The corresponding dynamic model can be represented as follows:

$$\frac{d}{dt}\begin{bmatrix} XV \\ SV \\ PV \\ O_2V \\ CO_2V \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -k_1 & -k_4 & -1 \\ 0 & 1 & 0 \\ -k_2 & -k_5 & -k_7 \\ k_3 & k_6 & -k_8 \end{bmatrix} \begin{bmatrix} \mu \\ q_P \\ q_m \end{bmatrix} XV + \begin{bmatrix} 0 \\ F \cdot S_0 \\ 0 \\ OTR \cdot V \\ -CTR \cdot V \end{bmatrix} \quad (6)$$

where S_o is the substrate concentration in the feed, *OTR* is the oxygen transfer rate from gas to liquid phase and *CTR* is the carbon dioxide transfer rate from liquid to gas phase.

The first assumption is to consider a single overall reaction, a so-called Black Box model, for the oxidative uptake of substrates to form biomass and products as it is illustrated by the following equation:

$$k_1^* S + k_2^* O_2 \rightarrow X + k_3^* P + k_4^* CO_2 \quad (7)$$

where *S* denotes one single limiting substrate (glycerol or methanol) as the carbon and energy source), O_2 oxygen, *X* biomass, *P* product and CO_2 carbon dioxide; and k^*_i values are stoichiometric coefficients that can also be called overall i-biomass yields ($Y^*_{i/X}$).

So, the mass balance equations for an ideal stirred tank reactor in a fed-batch cultivation process, considering conversion rates of biomass formation, substrate and oxygen uptake, product and carbon dioxide formation, can be formulated:

$$\frac{d}{dt}\begin{bmatrix} XV \\ SV \\ PV \\ O_2V \\ CO_2V \end{bmatrix} = \begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} XV + \begin{bmatrix} 0 \\ F \cdot S_0 \\ 0 \\ OTR \cdot V \\ -CTR \cdot V \end{bmatrix} \quad (8)$$

where μ is the specific growth rate [h^{-1}], q_S the specific substrate uptake rate [$g \cdot g^{-1} \cdot h^{-1}$], q_P the specific production rate [$U \cdot g^{-1} \cdot h^{-1}$], q_{O2} the specific oxygen uptake rate [$mol \cdot g^{-1} \cdot h^{-1}$], q_{CO2} the specific carbon dioxide production rate [$mol \cdot g^{-1} \cdot h^{-1}$], F the substrate feeding rate [$L \cdot h^{-1}$], V the volume of broth in the reactor [L], S_0 the substrate feeding concentration [$g \cdot L^{-1}$], OTR the oxygen transfer rate [$mol \cdot L^{-1} \cdot h^{-1}$] and CTR the carbon dioxide transfer rate [$mol \cdot L^{-1} \cdot h^{-1}$].

The Eq. 4 and Eq. 8 can be used to simulate the time evolution of the state variables. However, the withdrawal of components (X, S, P, O_2 and CO_2) should be taken into account for the reconstruction of the state from an experimental data set and to validate the model (Eq. 9). In other case, neglecting the withdrawal of the components would be an approximation of the model.

$$\frac{d}{dt}\begin{bmatrix} XV \\ SV \\ PV \\ O_2V \\ CO_2V \end{bmatrix} = \begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} XV + \begin{bmatrix} -F_O \cdot X \\ F \cdot S_0 - F_O \cdot S \\ -F_O \cdot P \\ OTR \cdot V - F_O \cdot O_2 \\ -CTR \cdot V - F_O \cdot CO_2 \end{bmatrix} \quad (9)$$

Furthermore, the mass balance equations can be defined also in individual variables (or concentration), instead of using global variables (Eq. 8), as follows:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ P \\ O_2 \\ CO_2 \end{bmatrix} = \begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} X - \begin{bmatrix} X \\ S \\ P \\ O_2 \\ CO_2 \end{bmatrix} \frac{dV}{V \cdot dt} + \begin{bmatrix} 0 \\ (F/V) \cdot S_0 \\ 0 \\ OTR \\ -CTR \end{bmatrix} \quad (10)$$

Finally, in case the dV/dt term can be approximated to F, equation (10) will take the form of equation (11) including the dilution term D= F/V.

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ P \\ O_2 \\ CO_2 \end{bmatrix} = \begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ P \\ O_2 \\ CO_2 \end{bmatrix} + \begin{bmatrix} 0 \\ D \cdot S_0 \\ 0 \\ OTR \\ -CTR \end{bmatrix} \quad (11)$$

The mass balance equations used in Chapter 2, Chapter 3 and Chapter 4 are based on equation (11) and in Chapter 5 on equation (8).