UAB
Universitat Autònoma de Barcelona

Physiological control of Pichia pastoris for the production of recombinant proteins: A novel Respiratory Quotient control based on Artificial Intelligence and scale-up of the bioprocess

Arnau Gasset Franch

Physiological control of *Pichia pastoris* for the production of recombinant proteins:

A novel Respiratory Quotient control based on Artificial Intelligence and scale-up of the bioprocess

Memòria per optar al grau de Doctor per la Universiat Autònoma de Barcelona, sota la direcció dels doctors José Luis Montesinos Seguí i Francisco Valero Barranco.

Arnau Gasset Franch

Bellaterra, 2023

Universitat Autònoma de Barcelona. Escola d'Enginyeria. Departament d'Enginyeria Química, Biològica i Ambiental

Title: Physiological control of *Pichia pastoris* for the production of recombinant proteins: A novel Respiratory Quotient control based on Artificial Intelligence and scale-up of the bioprocess

Author: Arnau Gasset Franch

Supervisors: José Luis Montesinos Seguí i Francisco Valero Barranco

This work was funded by the Project PID 2019-104666GB-100 of the Spanish Ministry of Science and Innovation and the Project “Continued Process Verification (CPV) of the Future” funded by both the Parental Drug Association (PDA) and the Product Quality Research Institute (PQRI).

This work was also funded by the MCIN /AEI /10.13039/501100011033 / FEDER, UE. (Project PID2022-136936OB-I00) Spain.

The group is member of the 2021 SGR 00143 of the Department of Research and Universities of Generalitat de Catalunya. Arnau Gasset acknowledges the award of a scholarship FI-DGR2019 of the Generalitat de Catalunya.

El Dr. Francisco Valero Barranco i el Dr. José Luis Montesinos, Professor Catedràtic i Professor Titular, respectivament, membres del grup de recerca en Enginyeria de Bioprocessos i Biocatàlisi Aplicada del Departament d'Enginyeria Química, Biològica i Ambiental de la Universitat Autònoma de Barcelona

Certifiquem:

Que el biotecnòleg Arnau Gasset Franch ha dut a terme sota la nostra direcció al Departament d'Enginyeria Química, Biològica i Ambiental de la Universitat Autònoma de Barcelona i amb la nostra direcció el treball titulat "**Physiological control of *Pichia pastoris* for the production of recombinant proteins: A novel Respiratory Quotient control based on Artificial Intelligence and scale-up of the bioprocess**". El treball es presenta en aquesta memòria, la qual constitueix la seva Tesi per optar al grau de Doctor per la Universitat Autònoma de Barcelona dins del programa de Doctorat en Biotecnologia.

I per tal que se'n prengui coneixement i consti als efectes oportuns, signem la present declaració.

Dr. Francisco Valero Barranco
Co-director

Dr. José Luis Montesinos Seguí
Co-director

Arnau Gasset Franch
Autor

Bellaterra, novembre de 2023

Agraïments

En primer lloc voldria dedicar unes paraules d'agraïment a la meva família més propera: la meva germana i en especial els meus pares, que són qui m'han portat fins aquí. Els tres sempre han estat presents per escoltar les meves xapes sobre biologia cel·lular, genètica i en els últims anys, també sobre llevats i proteïnes recombinants. També han estat els meus pares qui sempre m'han insistit en seguir estudiant quan jo em plantejava si anar a la universitat o bé fer carrera de bomber, la qual segur que m'hagués donat menys maldecaps i més temps lliure del que he disposat aquest últims anys (i aquest últim any en particular). Que diguin el que vulguin, que això de fer una tesi és molt pitjor que presentar-se a unes oposicions, i fins i tot diria que encara té menys garanties d'èxit...

Seguidament, també vull tenir unes paraules per els tutors/directors de tesi, el Paco i el José Luis. Per mi han estat una mica com els meus pares dins del món acadèmic, a qui he demanat i de qui he escoltat consell, a qui de vegades he hagut de creure en contra de la meva opinió i de qui sempre m'he pogut refiar quan ho he necessitat. A més, sempre he tingut clar a qui li havia de preguntar cada cosa, ja que com a codirectors han format un molt bon tàndem, cadascun amb la seva expertesa particular: el Paco que coneix a tothom i que sempre sap a qui o a on s'ha d'anar a buscar una informació en particular i que amb un correu seu pot solucionar tots els tràmits administratius que jo tardaria una setmana a resoldre; i el José Luis, que amb 5 minuts i un full de paper et fa entendre aquelles equacions que quan les mires per primer cop sembla que t'hagis equivocat de llibre, o que amb una tarda i una plantilla d'excel et fa quadrar les dades de tots els experiments de la tesi. Per extensió també el Xavi, que tot i que no ha estat oficialment el meu director, també ha contribuït molt en l'execució d'aquest treball, i que és, de moment, la persona que m'ha aconseguit més feines i que per tant ha contribuït més en el meu desenvolupament professional. Per als tres, moltes gràcies.

En tercer lloc, els meus companys de feina també es mereixen unes molt importants paraules d'agraïment: Albert(s), Núria, Guille, Javi, Manito, Miquel, Juanjo, Mario, Sílvia, Kike, Lluc, Esther, Jessi, Laia, etc. No només els membres del grup de recerca, sinó també altres companys del departament, amb qui la relació ha anat més enllà de l'estrictament professional. Si bé és cert que la universitat pública catalana en general, i el nostre departament en particular, no destaca per l'abundància de recursos humans i materials, també s'ha de remarcar el bon ambient de treball, el companyerisme i el bon rotllo que hi ha. Per totes les tardes fent cerveses a la figuera, per tots els dimecres de bàsquet, i per totes les calçotades, barbacoes i "corderitos" als que he tingut el plaer de participar (i fins i tot organitzar), només en puc donar gràcies, ja que han fet que aquest pas pel departament hagi estat molt més agradable. Fins i tot, gràcies per

a totes les persones del departament i de la universitat (administratives, tècniques, secretaries, etc.) que sempre han estat allí per donar un cop de mà: Rosi, Manuel, Pili, Rosa, Laura… i per totes aquelles altres persones que sense arribar-les a conèixer personalment, m'han ajudat a entendre la figura del funcionari, cosa que els meus pares (funcionaris els dos, un bomber i l'altra mestra) no m'ho han pogut transmetre amb tanta facilitat i claredat. A tots ells, també moltes gràcies, ja que puc dir que marxo d'aquí tenint força més paciència que amb la que vaig arribar.

Finalment, i no per això menys important (jo diria que fins i tot més), vull dedicar unes sentides paraules d'agraïment per una colla de gent que s'ajunta els dimarts i els dijous al migdia per enfilar-se els uns sobre els altres. Sí, es tracta dels Ganàpies de la UAB, que m'han acompanyat al llarg de la meva carrera universitària i de qui m'hauré de despedir d'aquí poc temps, no sense una gran tristesa però veient com la colla creix, evoluciona i es renova any rere any. Potser 10 anys ja comencen a ser massa per formar part d'una colla castellera universitària... A tots ells, des del primer fins a l'últim, moltes gràcies. En especial, però, a aquells i aquelles que han acabat formant el que per mi ha estat la meva segona (o tercera, o quarta) família: Edu, Estel, Marti(s), Paula, Pau(s), Joan(s), Maria(s), Judit, Oriol, Adrià, Clara, Lídia, Arnau, Mar, Clàudia. Dani… i no acabaria mai. En realitat, quan vaig decidir fer un doctorat a la UAB, va ser en gran part per poder seguir compartint moments amb tots ells, moments que espero poder seguir compartint.

I ja, per acabar, donar les gràcies a la Universitat Autònoma de Barcelona com a institució per haver-me permès viure la meva vida universitària en un lloc tant idíl·lic pel que fa a l'entorn natural. I també per haver-me ensenyat que, com més gran és una institució, més difícil és de governar i més desigualtats acaben apareixent entre les diferents "castes" que s'hi troben. I que la política s'ha de fer des de la base, tant en una universitat com en un país.

Resum

Aquesta tesi constitueix una continuació d'un estudi iniciat als anys 2007–2008, on l'objectiu principal era analitzar com afectava la limitació d'oxigen en un procés de producció de proteïna recombinant (PPR) amb el llevat metilotròfic *Pichia pastoris*. Concretament, la proteïna recombinant expressada era un fragment d'anticòs (Fab), que estava sota la regulació del promotor constitutiu GAP (P_{GAP}), un promotor àmpliament utilitzat per a l'expressió heteròloga. En les primeres etapes d'aquest estudi es va demostrar que una limitació de l'aportació d'oxigen en cultius de *P. pastoris* en quimiòstat, utilitzant glucosa com a font de carboni, provocava un augment dels fluxos metabòlics de la glicòlisi degut a la sobreexpressió d'alguns gens glicolítics. Un d'aquests gens, el *THD3*, és el gen que de forma natural està regulat pel P_{GAP} en *P. pastoris*. Per tant, en condicions de limitació d'oxigen (o hipòxia), un gen recombinant sota la regulació del P_{GAP} també s'hauria de trobar sobreexpressat. En aquests estudis previs també es va observar que sota condicions hipòxiques, *P. pastoris* desenvolupa un metabolisme respiro-fermentatiu, produint sub-productes de fermentació com són l'etanol, i també en menor mesura arabitol i succinat.

En una primera etapa de la present tesi, es va dur a terme una caracterització de l'efecte de la hipòxia sobre la producció recombinant de la lipasa 1 de *Candida rugosa* (Crl1) mitjançant cultius en continu amb dos clons diferents de *P. Pastoris*, un que disposava d'una còpia del gen *CRL1* (*single-copy clone* o SCC) i un altre que en disposava de 5 còpies (*multicopy clone* o MCC). Es van provar diferents graus de limitació d'oxigen variant la concentració d'oxigen en el gas d'entrada, definit aquelles condicions que es consideraven òptimes en termes de producció de Crl1 i que no comprometien significativament el creixement del llevat. La implementació de cultius en continu en estat estacionari (quimiòstats) va permetre optimitzar l'estudi de diverses condicions arribant a diferents estats estacionaris. Les velocitats específiques de producció (q_P) van ser fins a 2 vegades més grans en hipòxia que en normòxia (condicions completament aeròbiques), tot i que aquest increment de producció de Crl1 va resultar més important per al SCC que per al MCC. Addicionalment, es va definir el Quocient Respiratori (*RQ*) com a variable per a controlar el nivell de limitació d'oxigen, ja que aquesta és una variable que depèn de l'estat fisiològic de la biomassa, és a dir, del tipus de metabolisme que aquesta duu a terme (respiratiu, fermentatiu o respiro-fermentatiu). A més, aquesta variable és independent de l'equipament i del sistema utilitzat, i per tant, es pot utilitzar com a variable indicadora a l'hora d'implementar el mateix grau d'hipòxia en un altre equip o mode d'operació.

Posteriorment, un cop definides les condicions de limitació d'oxigen òptimes, es van replicar aquestes condicions en fed-batch. Al tractar-se d'un cultiu dinàmic en què la concentració de biomassa augmenta amb el temps, el control sobre el nivell de limitació d'oxigen també havia de ser dinàmic. Els resultats obtinguts van ser similars als obtinguts en quimiòstat, duplicant la q_P del SCC en condicions d'hipòxia i obtenint, però, un increment menor en el cas del MCC. A més, els resultats van ser corroborats gràcies a un estudi de transcriptòmica dels gens relacionats amb la producció de Crl1 mitjançant qPCR.

Per finalitzar aquesta primera etapa de la tesi, es va dur a terme dues fermentacions a escala pilot, utilitzant el *RQ* com a criteri de canvi d'escala per tal d'aplicar les mateixes condicions d'hipòxia, obtenint una producció de Crl1 lleugerament més baixa que a escala laboratori però validant l'estratègia de limitació d'oxigen i la utilització del *RQ* com a paràmetre de control transferible entre diferents sistemes experimentals.

Un cop demostrat i caracteritzat l'efecte de la hipòxia sobre la PPR, la segona etapa de la tesi va consistir en millorar el controlador de *RQ* en fed-batch, que inicialment estava basat en la modificació manual i intermitent de la velocitat d'agitació per tal d'augmentar o disminuir la transferència d'oxigen al medi de cultiu (control manual-heurístic). En aquest segon capítol, es van establir dues estratègies de control de *RQ* sofisticades, les quals modificaven la velocitat d'agitació de forma automàtica quan era necessari. La primera estratègia, anomenada "controlador de lògica Booleana", va ser dissenyada, programada i desenvolupada a través del software *Eve*, el qual va permetre la implementació de *soft sensors* per tal de fer un càlcul precís i acurat del *RQ* i actuar sobre la velocitat d'agitació de forma automàtica. D'altra banda, la segona estratègia de control va consistir en un controlador proporcional adaptatiu, en què el valor del paràmetre proporcional del controlador s'anava actualitzant a cada minut de la fermentació. Aquesta actualització estava basant en un model d'intel·ligència artificial (IA), i va ser dissenyada conjuntament en col·laboració amb membres de l'empresa AIZON, que van ser qui en van programar el codi. Aquesta estratègia ha esdevingut una prova de concepte per demostrar com es pot utilitzar la IA per a dur a terme el control automàtic de bioprocessos, un camp on la indústria és encara reticent d'aplicar canvis.

Finalment, com a últim capítol de la tesi, l'estratègia hipòxica es va provar amb dues soques de *P. Pastoris* més, expressant dues proteïnes diferents: la lipasa de *Rhizopus oryzae* (proRol) i la lipasa B de *Candida antarctica* (CalB), dues lipases amb interès industrial. Les millores de producció obtingudes en aplicar condicions hipòxiques van ser equiparables a les obtingudes amb la producció de Crl1, validant així l'estratègia hipòxica per a la millora de producció de proteïnes recombinants.

Summary

This thesis represents a continuation of a study started in the years 2007-2008, in which the main goal was to analyze the impact of oxygen limitation on recombinant protein production (RPP) with the yeast *Pichia pastoris.* Specifically, the recombinant protein expressed was an antibody fragment (Fab), under the regulation of the constitutive GAP promoter or $P_{GAP,}$ a widely used promoter for RPP. In the initial stages of this study, it was demonstrated that a limitation of the oxygen supply in *P. pastoris* cultures operated in chemostat mode, and using glucose as a carbon source, led to an increase in metabolic flux through glycolysis due to the overexpression of some glycolytic genes. One of these genes, *THD3*, is the gene that is naturally regulated by the P_{GAP} in *P. pastoris*. Therefore, under oxygen-limiting conditions (or hypoxia), a recombinant gene under the regulation of P_{GAP} should be also overexpressed. In these previous studies it was also demonstrated that under hypoxic conditions, *P. pastoris* exhibits a respiro-fermentative metabolism, producing fermentation by-products such as ethanol, and also to a lesser extent arabitol and succinate.

In the first stage of the present thesis, a characterization of the effect of hypoxia on the recombinant production of *Candida rugosa* lipase 1 (Crl1) with *P. pastoris* was carried out using continuous cultures. Two different clones of *P. pastoris* were tested, one harboring one copy of the *CRL1* gene (single-copy clone or SCC) and another harboring 5 copies of it (multicopy clone). Different degrees of oxygen limitation were tested by varying the oxygen concentration in the inlet gas of the fermenter, and the optimal conditions for Crl1 production, which did not compromise severely the yeast growth, were defined. Continuous cultures in steady state (chemostats) were used to test different conditions reaching different steady states, resulting in specific production rates (q_P) up to two-fold higher under hypoxia compared to normoxia (completely aerobic conditions), although this increase in Crl1 production was higher for the SCC than for the MCC. Additionally, the Respiratory Quotient (*RQ*) was defined as a control variable to regulate the oxygen limitation degree, as it depends on the physiological state or the type of metabolism that cells conduct. Moreover, this variable is independent of the equipment and system used so, it can be used as a reporting variable to implement the same hypoxic degree in another equipment or operational mode.

Subsequently, once these optimal hypoxic conditions had been defined, they were replicated in fed-batch cultures. Since they are dynamic cultures where biomass concentration increases over time, the control of the process to maintain a constant level of oxygen limitation should also be dynamic. The results obtained in fed-batch cultures were similar to those in chemostat cultures, doubling the q_P

of SCC under hypoxic conditions, but obtaining a lower increase in the case of MCC. In addition, these results were further confirmed through transcriptomic analysis of certain genes related to Crl1 production using qPCR.

To conclude this first stage of the thesis, two pilot-scale fermentations were conducted, using *RQ* as a scale-up criterion to apply the same hypoxic conditions. Although Crl1 production was slightly lower at the pilot scale than at the lab scale, it validated the oxygen limitation strategy and the use of *RQ* as a transferable control parameter between different experimental setups.

After having demonstrated and characterized the hypoxic effect on bioprocess efficiency, the second stage of the thesis consisted of the refinement of the *RQ* controller in fed-batch, which was initially based on a manual and intermittent modification of agitation rate to adjust oxygen transfer into the culture broth (manual-heuristic control). Thus, two improved *RQ* control strategies were established in this second chapter, both of which consisted of adjusting the agitation rate automatically as the process required. The first strategy, named “Boolean-logic controller” was designed, programmed, and developed using a specific fermenter control software called *Eve*. It allowed the implementation of soft sensors to make a precise and accurate *RQ* calculation and act over agitation rate in an automated way. In contrast, the second control strategy consisted of an adaptive proportional controller, in which the value of the proportional parameter was updated every minute. This parameter value adaptation was based on an artificial intelligence (AI) model, and it was designed in collaboration with members of the company AIZON, who programmed the code. This strategy served as a proof of concept to demonstrate how AI can be used for the automatic control of bioprocesses, an area where the industry is still reluctant to implement changes.

Finally, as in the last chapter of the thesis, the hypoxic strategy was tested with two additional strains of *P. pastoris*, expressing two different proteins: *Rhizopus oryzae* lipase (proRol) and *Candida antarctica* lipase B (CalB), two lipases with industrial interest. The improvements in production observed under hypoxic conditions were comparable to those obtained with Crl1 production, thus validating the hypoxic strategy for the enhancement of recombinant protein production.

Abbreviations and Acronyms

AI: Artificial Intelligence
AI-APC: AI-aided Adaptive-Proportional Control
AU: Activity Units
BIRE: Big Reactor
BLC: Boolean Logic Control
CalB: *Candida antarctica* lipase B
CFD: Computational Fluid Dynamics
CHO: Chinese Hamster Ovary
CPV: Continuous Process Verification
Crl1: *Candida rugosa* lipase 1
DCW: Dry Cell Weight
ddPCR: digital droplet PCR
DO: Dissolved Oxygen
EMA: European Medicines Agency
Eng4Bio: Bioprocess Engineering and Applied Biocatalysis Group
FC_{GD}: Gene Dosage Fold-Change
FC_{OL}: Oxygen Limitation Fold-Change
FDA: U.S. Food and Drug Administration
GRAS: Generally Recognized As Safe
HPLC: High-Performance Liquid Chromatography
IoT: Internet of Things
kAU: kiloActivity Units
MCC: Multi-Copy Clone
MHC: Manual-Heuristic Control
MRE: Mean Relative Error
MW: Molecular Weight
nlpm: normal liter per minute
OD: Optical Density
PAT: Process Analytical Technologies
PD: Process Design
PDA: Parental Drug Association
PES: Polyethersulfone
PID: Proportional-Integrative-Derivative
p-NP: para-nitrophenol
p-NPB: para-nitrophenyl butyrate
PQ: Process Qualification
PTFE: Polytetrafluoroethylene
PTM: Post-Translational Modifications
PTM_1: *Pichia*'s Trace Metals
PV: Process Validation
qPCR: quantitative real-time PCR
RCT: Research Corporation Technologies
RMSD: Root Mean Square Deviation
RoI: *Rhyzopus oryzae* lipase
rpm: revolutions per minute
RPP: Recombinant Protein Production
RQ: Respiratory Quotient
RSD: Relative Standard Deviation
RTL: Relative Transcription Level
SCADA: Supervisory Control And Data Acquisition
SCC: Single-Copy Clone
SCP: Single Cell Protein
SD: Standard Deviation
SPC: Statistical Process Control
TCA: Tricarboxylic Acid Cycle
TF: Transcription Factors
TFBSs: Transcription Factors Binding Sites
UPR: Unfolded Protein Response
vvm: gas volume/liquid volume per minute

Table of contents

1. CONTEXT

The research work included in this PhD thesis has been performed in the Bioprocess Engineering and Applied Biocatalysis Research Group (Eng4Bio), located at the Department of Chemical, Biological, and Environmental Engineering of the Engineering School from the Universitat Autònoma de Barcelona.

The Eng4Bio research group is focused on the development of bioprocesses mainly for the production of recombinant proteins, but recently also to produce metabolites in different cell factories. Thus, the works carried out in the group often consist of an initial step of strain engineering to obtain the clones for recombinant production, mainly with the cell factory *Pichia pastoris*. Secondly, screenings of several clones to select the best candidates are performed, followed by a deep characterization of these producer clones to set the basis for bioprocess optimization. Finally, some of the recombinant proteins are used for final biocatalysis applications, whereas in other cases these proteins are delivered to third parties.

Thus, the Eng4Bio research group has a library of clones expressing different proteins with industrial interest. Most of them are lipases since one of the most performed biocatalysis applications in the group is the use of lipases for racemic solutions, biofuel production, and the synthesis of natural flavors, among others. However, this clone library is relatively limited, and not all the combinations of strain + expression cassette (including promoter, secretion signal, etc.) + gene of interest are available.

In addition, the research group has gained extensive and increasing know-how in the development of fermentation-based bioprocess and biocatalysis, and proof of that is the numerous scientific papers published yearly. Some of the last relevant works performed in the group are:

- 3-hydroxypropionic acid production with engineered *P. pastoris* (Fina et al., 2021, 2022).
- Implementation of a novel secretory signal based on co-translational translocation for *P. pastoris* (Barrero et al., 2018; Barrero, 2020).

- Clone construction, characterization, and bioprocess optimization for the production of different lipases with *P. pastoris* (Garrigós-Martínez et al., 2019; Nieto-Taype et al., 2020a; López-Fernández et al., 2021a).
- Synthesis of second and third-generation biodiesel and natural flavors with *Rhizopus oryzae* lipase (Rol) produced by *P. pastoris* (López-Fernández et al., 2021b; López-Fernández et al., 2022).
- Study of new promoters for recombinant protein production with *P. pastoris* (Garrigós-Martínez et al., 2021; Bernat-Camps et al., 2023).

Special interest must be placed on a previous work started in 2008, where oxygen-limiting conditions (or hypoxia) were found to increase recombinant protein production with *P. pastoris* (Baumann et al., 2008). During the following years, several studies thoroughly evaluated the effect of these oxygen-limiting conditions on cell growth and recombinant protein production, at a transcriptomic, proteomic, metabolomic, and lipidomic level (Carnicer et al., 2009, 2012; Adelantado et al., 2017; Baumann et al., 2010). Finally, in 2017, a more practical operational point of view was applied to this hypoxic effect, setting the basis for the present work (Ponte et al., 2016; Garcia-Ortega et al., 2017). After that, in 2018, this thesis was begun with the primary aim of studying the production of recombinant proteins under hypoxic conditions. Logically, this study would not have been possible without the participation of all the researchers who participated in these studies, so all of them deserve special recognition.

On the other hand, some of the work conducted in this thesis has been done in the framework of the Parental Drug Association (PDA) project “CPV of the Future”. This project aims to set a basis for the implementation of Artificial Intelligence (AI) as a valid analytical technique for Continuous Process Verification (CPV) in fermentation operations, mainly in the pharma sector, which is still reluctant to implement these advancements due to slow and expensive regulatory approvals. As can be found on the PDA website, “*This PDA initiative comprises a set of experimental studies that will focus on improving production of the recombined protein generated by the fermentation of the Pichia pastoris microorganism under different conditions of oxygen supply. Specific attention will be paid to hypoxia settings. Using a bioreactor of five-liter volume, fermentations*

will be performed that produce batches controlled through automatic and manual operations" (Manzano, 2020).

This section of the thesis has represented a significant challenge for the research group, as there was limited knowledge regarding the integration of signals from different devices, utilization of local networks, cloud connectivity for continuous data transmission, and, in general, the implementation of a digital twin of the fermenter. Moreover, there was not extensive expertise in programming within the research group, as the work of the Eng4Bio group mainly focuses on strain and bioprocess engineering. In this regard, the collaboration with the AIZON team has been crucial, as they provided all the necessary knowledge regarding equipment connectivity and data transmission on the one hand, and on the other hand, their extensive programming experience has been pivotal for the development of this work.

2. INTRODUCTION

2.1 General aspects of Biotechnology

Biotechnology is a multidisciplinary field in which mainly biological and life sciences knowledge is imperative, but where engineering approach, rational thinking, and data management skills are also essential. Furthermore, in the last years, computational and statistics abilities have emerged as powerful tools to improve biotechnology processes. To synthesize, biotechnology is the combination of biology and technology to leverage biological systems and processes to develop innovative products and services (Lynd et al., 1999). Biotechnology started with beer fermentation by the Sumerian and Egyptian cultures, more from an empirical approach rather than from a clear understanding of the bioprocess, and in the next years it will be evolved to high-tech production bioprocesses managed by artificial intelligence (AI) algorithms for the production of high added value pharmaceutical drugs (Manzano et al., 2021).

Another definition of biotechnology is the utilization of any living being to obtain a product or service. Therefore, from this point of view, agriculture and livestock could also be considered as biotechnology. However, this concept is not very accepted, and when it comes to agriculture and livestock, biotechnology and all the techniques that it encompasses are considered a way to improve traditional farming (European Commission, 2017).

The fields of biotechnology applications are very wide, ranging from industrial applications to medical research and new pharmaceutical products. All these subcategories are often divided into different “colors” based on the application and the complexity of the products involved. The "colors" of biotechnology include white biotechnology (industrial processes, fine chemical products, enzymes, biofuels), red biotechnology (health and medical research, including pharma industry), green biotechnology (agriculture), blue biotechnology (sea resources), and grey biotechnology (environmental bioengineering), among several others (DaSilva, 2004), as shown in **Figure 1**.

In recent years, industrial (or “white”) biotechnology has gained significant attention due to its potential to provide sustainable and eco-friendly solutions for industrial processes while reducing the environmental impact of traditional chemical manufacturing methods (Soetaert & Vandamme, 2006; Fröhling &

Hiete, 2020). The production of valuable products such as biofuels, biopolymers, and proteins using microorganisms is an essential aspect of white biotechnology (Koutinas et al., 2014; Heux et al., 2015).

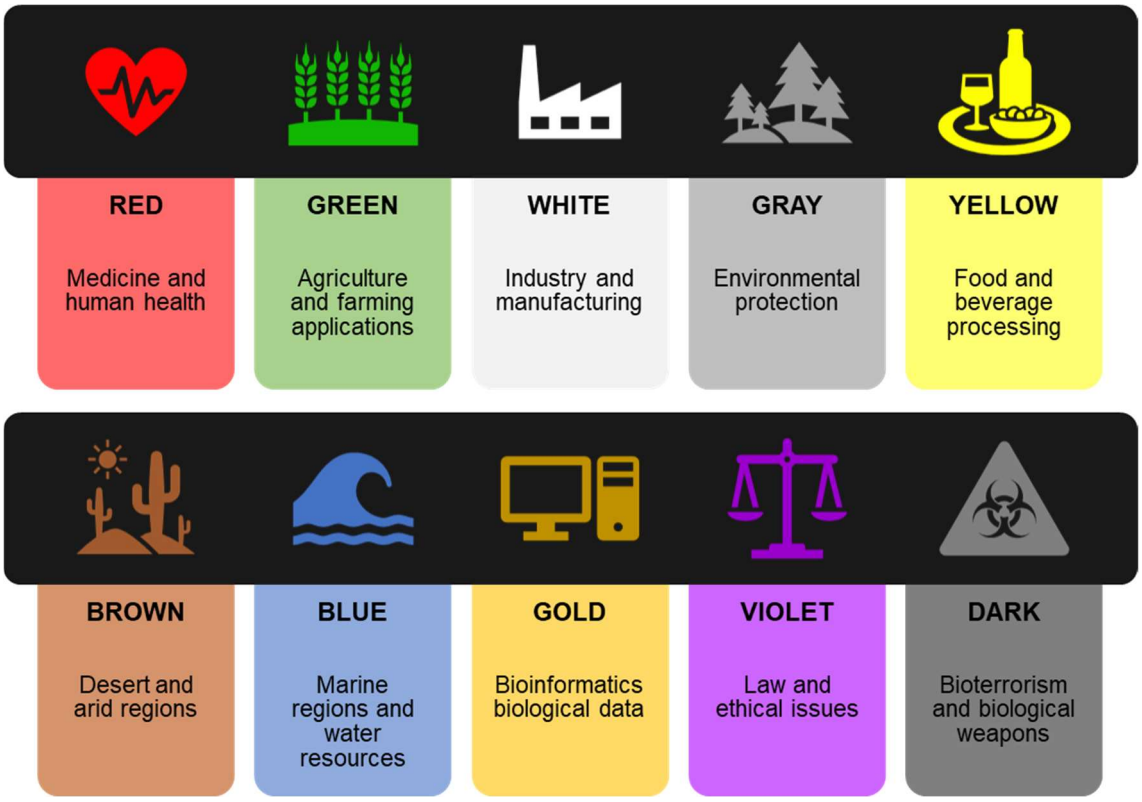


Figure 1. Graphical summary of the colors of biotechnology. Adapted from literature (Kafarski, 2012).

2.2 Recombinant proteins

Proteins are complex biomolecules that play critical roles in virtually every cellular process, including biocatalysis, regulation, signaling, transport, and cellular scaffolding. Since they are efficient biomolecules, cells do not produce enormous amounts of them, so normally the extraction from their native source results in insufficient yields and poor productivities. As a result, recombinant DNA technologies have been developed for the production of recombinant proteins, which are synthesized in a host cell frequently of a different species than their origin, often referred to as “cell factories”. Recombinant protein production (RPP) processes have become a reliable source of biocatalysts for several industrial sectors such as food, detergents, biofuels, textiles, polymers, paper, and pulp (Rigoldi et al., 2018), as well as biopharmaceuticals for human and veterinary medicine, including therapeutics and diagnostics biomolecules (Gupta & Shukla, 2017).

The production of recombinant proteins involves identifying the genes encoding the desired protein, cloning them into a microbial host, and optimizing the microorganism growth to find the best producing conditions.

In addition, the recombinant enzymes can be improved for specific applications, with a process that is called “protein engineering”. Rational design and directed evolution are protein engineering techniques that have been used to develop novel recombinant proteins with improved or new properties (Porter et al., 2016; Bilal et al., 2018). Rational design involves site-specific alterations of selected residues in a protein to cause predicted changes in function, while directed evolution involves repeated cycles of generating a library of protein variants and selecting the variants with desired properties. For example, a microbial acetyltransferase for modification of glyphosate has been improved 7,000-fold in enzyme activity and 5-fold in thermostability through directed evolution (Siehl et al., 2005). Additionally, in the last years, AI algorithms have been also applied to directed protein evolution to improve enzyme activity and thermostability (Yang et al., 2019).

2.2.1 Lipases

An industrially relevant class of proteins are lipases, which are a family of enzymes that catalyze the hydrolysis of the ester bonds of triglycerides, releasing free fatty acids and glycerol, as sketched in **Figure 2**. They are considered promising catalysts for many applications, such as in the food, detergents, pulp, and paper or biofuels industries, among many others (Mehta et al., 2020).

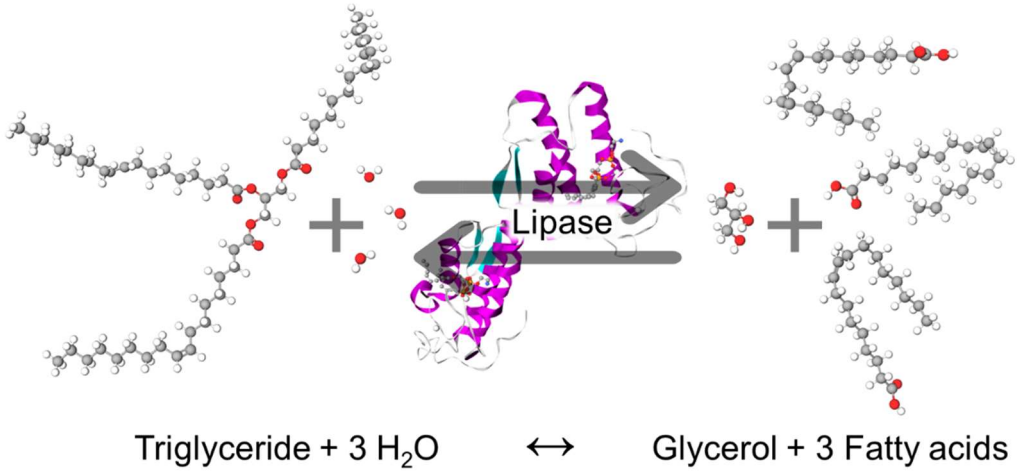


Figure 2. Schematization of the enzymatic reaction catalyzed by lipases. In this example, triolein is hydrolyzed into glycerol and three free molecules of oleic acid by the enzymatic action of a lipase.

One of these lipases that has been widely used by flavor, oil, and pharma industries is the *Candida rugosa* lipase (Crl) (Ken Ugo et al., 2017). Specifically, seven isoenzymes of Crl (Crl1-7) have been described, with Crl1 being the most common in commercialized lipase preparations (Ferrer et al., 2001). It can be produced recombinantly in cell factory hosts or purified from native *C. rugosa* extracts (Vanleeuw et al., 2019).

On the other hand, *Rhizopus oryzae* lipase (Rol) is a 1,3-regiospecific enzyme and highly enantioselective, distinguished in the biochemical synthesis of flavors, biofuels, and racemic resolutions in the pharma industry, to name some of its uses. It is obtained both from extraction and purification from its native source or produced *via* recombinant synthesis (López-Fernández et al., 2020).

Finally, another interesting lipase from an industrial point of view is *Candida antarctica* lipase B (CalB), which is one of the most widely studied and patented lipases. Among other applications, it has been largely used in detergent preparations due to its inherent ability to catalyze reactions at low temperatures, given the fact that its natural host is adapted to cold environments (Shivaji & Prasad, 2009; Rabbani et al., 2015).

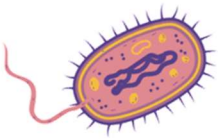
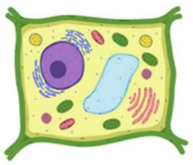
All these three lipases have been recombinantly produced and extensively studied in the Eng4Bio research group using *P. pastoris* as a cell factory (Barrigón et al., 2013; Ponte et al., 2016; Garrigós-Martínez et al., 2019, 2021; Nieto-Taype et al., 2020a; Gasset et al., 2022; Bernat-Camps et al., 2023).

2.3 Expression hosts

The production of complex proteins in large quantities while maintaining the native structure and biological activity presents significant challenges. The choice of a suitable host for recombinant protein expression is crucial and is one of the most critical steps in bioprocess design and development. Ideally, a universal expression system would enable the expression of all possible recombinant genes in a fast, cheap, and proper manner concerning yield, folding, and biological activity (Yin et al., 2007; Sørensen, 2010). However, the limitations of current systems make the selection of the expression system an empirical determination based on production parameters such as cost, yield, production timescale, scale-up capacity, downstream processes, and product properties.

Thus, different applications have different requirements for both quantity and quality, and the characteristics of the target protein must be taken into account, as displayed in **Table 1**. Glycosylation represents the most complex and widespread post-translational modification (PTM) and selecting the wrong expression host can result in the protein being misfolded, poorly expressed, lacking the necessary PTMs, or containing inappropriate modifications (Kim et al., 2015; Gupta & Shukla, 2017).

Table 1. Pros, cons, and common applications of the most used expression systems for recombinant protein production.

Expression System	Advantages	Disadvantages	Common Applications
Bacteria 	• Low-cost growth medium • Rapid growth and expression • Easy scale-up	• Lack of post-translational modifications (PTM) • Formation of inclusion bodies • Endotoxins	• Simple proteins with low MW (bacterial proteins, without PTM) • Industrial enzymes • Small peptides
Yeasts 	• Low-cost growth medium • Rapid growth and expression • High cell density • Some PTM	• Improper glycosylation compared to human pattern • Thick cell wall difficult cell disruption	• Low MW proteins • Industrial enzymes • Non-parenteral drugs • Diagnostic proteins
Plant cells 	• Low-cost growth • Rapid growth • Complex PTM • No pathogenic contamination risks • No endotoxins	• Low capacity of glycosylation • Lack of regulatory approval • Complex downstream purification	• Non-parenteral drugs • Diagnostic proteins • Veterinary vaccines
Mammalian cells 	• No endotoxins • Complex PTM with proper glycosylation patterns • More similar to human cells	• Very high-cost growth medium • Slow growth and low cell density • Human pathogen contamination risks	• Parenteral drugs (human enzymes, recombinant antibodies and antibody fragments, etc.) • Vaccines

Bacteria are an excellent expression system for proteins that do not require glycosylation or extensive PTMs, allowing for fast and inexpensive production processes. *Escherichia coli* has been the pre-eminent host cell for producing recombinant proteins for both commercial and research purposes, enabling a recombinant protein fraction of up to 50% of total biomass weight. There are many protocols available for high cell density cultivation with *E. coli*. However, producing eukaryotic proteins often results in low yields and inclusion body formation due to the faster rates of protein synthesis and folding in prokaryotes (Baeshen et al., 2015).

At the other extreme of cellular complexity, mammalian cell lines can overcome most of the limitations of producing recombinant eukaryotic proteins due to their capacity to perform human PTMs and to correctly fold and assemble complex human proteins. Chinese hamster ovary (CHO) cells are the preferred cell type

for the production of monoclonal antibodies and some other recombinant proteins due to their authenticity of glycosylation patterns (Ritacco et al., 2018). As a very recent example, the Covid-19 vaccine commercialized by HIPRA (Amer, Spain) is produced with CHO (Barreiro et al., 2023). However, the specific yields obtained with mammalian cell lines are often low, and their use is often non-prioritized due to their very low growth rates, extremely expensive medium and serum costs, and potential for viral product contamination (Yao & Asayama, 2017).

Recently, transgenic animals have also been developed as recombinant protein production systems, with high-quality proteins and high production yields at low cost. However, the production times are very long, and significant safety concerns such as transmission of infectious diseases, and adverse allergenic, immunogenic, and autoimmune responses of the host exist. Furthermore, the regulations are usually very prohibitive, especially in the European Union (Kleter & Kok, 2010; Hudson, 2017; Nishu et al., 2020).

On the other hand, transgenic plants offer an interesting possibility for producing recombinant proteins due to their lower production costs, timescale, safety risks, and ease of storage and distribution. Since there is a more permissive regulation, some bioprocess using genetically modified plants have been approved both in the USA and in the EU (Ma et al., 2015; Schillberg & Spiegel, 2022). Actually, Spain was in 2019 the only country in the EU growing genetically modified plants on a large scale (Karky & Perry, 2019).

Among all these aforementioned cell factories, yeasts, single-celled eukaryotic fungal organisms, offer a halfway house between bacteria and mammalian cells as an expression system, presenting important advantages for protein production processes. Yeasts are simple, genetically, and physiologically well-characterized, able to grow fast and up to very high cell densities in cheap mediums, generating high yields of recombinant proteins at low cost, which incorporate PTMs such as disulfide bonds and glycosylation (Gomes et al., 2018; Baghban et al., 2019).

In conclusion, the selection of the optimal expression system for recombinant protein production should be based on the specific requirements of the product

quality and the production parameters such as yield, cost, scale-up capacity, and downstream processes. While each expression system has its advantages and disadvantages, yeast and bacterial systems are currently the most widely used for commercial and research purposes due to their cost-effectiveness, scalability, and ease of production. However, mammalian cell lines, transgenic animals, and plants offer unique advantages to produce specific proteins, which may make them more attractive for certain applications.

2.3.1 Yeasts as expression hosts

Yeasts have been used for millennia in traditional and modern biotechnology to produce food, wine, beer, and a wide range of biochemical products. In modern biotechnology, yeasts are considered a versatile and efficient expression system for recombinant protein production due to their eukaryotic nature. As yeast cells are complex and compartmentalized, they have a great capacity for protein processing, including folding and assembling in the endoplasmic reticulum, introducing post-translational modifications (PTMs) such as glycosylation in the Golgi apparatus, and secretion to the extracellular medium through the vesicle system, among other abilities (Ahmad et al., 2014). In addition, since yeasts are single-celled microorganisms, they can grow fast and up to high cell densities in a chemically defined low-cost medium, especially compared with the medium needed for animal cell culture. Finally, they have a thick cell wall which gives them a high endurance and shear stress resistance.

Among the yeasts used for recombinant protein production, *Saccharomyces cerevisiae* is the most common and well-known yeast system. It is a GRAS (Generally Recognized As Safe) organism and has been used for producing therapeutic proteins since the approval of recombinant insulin by Novo Nordisk in 1986 (Ladisch' & Kohlmann, 1992). *S. cerevisiae* has been used to produce dozens of other therapeutic proteins, including human growth hormone, hepatitis B vaccine, and human serum albumin (Wang et al., 2017). However, one of the major drawbacks of using *S. cerevisiae* for producing biopharmaceuticals is the

difference between its glycosylation pattern and that of humans, which may limit therapeutic use.

Pichia pastoris is another yeast with GRAS status that has gained increasing attention as an efficient and versatile system for the expression of heterologous proteins (Love et al., 2018). *P. pastoris* is a methylotrophic yeast that can grow on methanol as a sole carbon source (Ogata et al., 1969). Due to this advantageous feature, it was initially commercially exploited to produce single-cell protein (SCP) for animal feed in the 70's by Phillips Petroleum Company (Bartlesville, OK, USA) using methanol, a by-product of petroleum distillation, as a substrate. Later in the early '80s, *P. pastoris* began to be studied as a heterologous gene expression system, and finally in the '90s Phillips sold its patent position with the *P. pastoris* expression system to RCT (Tucson, AZ, USA) and released it to academic laboratories for research purposes (Higgins & Cregg, 1998). Since then, it has been used to produce a wide range of recombinant proteins, including human insulin, interleukins, and monoclonal antibodies, among others, and is considered a promising platform to produce biopharmaceuticals (Valero, 2013). Already in the 2000's, the publication of the sequenced genome set the stage for *P. pastoris* to become an even more important cell factory for recombinant proteins (De Schutter et al., 2009; Sturmberger et al., 2016). In addition, during this decade, *P. pastoris* was taxonomically reclassified as *Komagataella phaffii* (Kurtzman, 2005, 2009). Even so, much of the scientific community has continued to call it "*Pichia pastoris*" (Mattanovich et al., 2017; Gasser & Mattanovich, 2018).

One of the advantages of *P. pastoris* over *S. cerevisiae* is the availability of strong promoters that regulate the expression of the desired protein, allowing high production yields. The main and most studied *P. pastoris* promoter is the alcohol oxidase 1 promoter or P_{AOX1}, which actually regulates the expression of the first gene related to methanol assimilation and was consequently the first to be studied (Cregg et al., 1989). Additionally, in *S. cerevisiae*, the use of vectors that are replicated episomally is still dominant, whereas in *P. pastoris* the recombinant gene of interest is integrated into the genome, which is a method that is significantly more stable (Baghban et al., 2019).

Another main advantage of *P. pastoris* is that it also has a secretion pathway that is more similar to higher eukaryotes than *S. cerevisiae*, with less elaborate hyperglycosylation, which makes it more similar to mammalian cells (Darby et al., 2012). Thus, it can produce high levels of secreted recombinant proteins, up to 80% of the total secreted protein (Li et al., 2007). Additionally, *P. pastoris* has been successfully glycoengineered, replacing the native pathway with the humanized N-glycosylation pathway to generate a set of strains producing proteins with human glycosylation patterns (Beck et al., 2010; Love et al., 2018).

However, one of the main disadvantages of P_{AOX1}-based processes is the use of methanol as a carbon source, which poses an important safety risk from an industrial point of view. Methanol-based expression systems using *P. pastoris* have been successful in obtaining high product titers in a cheap strategy, but they present significant drawbacks for large-scale bioprocesses associated with the storage and delivery of large quantities of methanol, as well as important increases in heat production and oxygen requirements (Yang & Zhang, 2018).

In the last two decades, the use of the glyceraldehyde-3-phosphate dehydrogenase promoter (GAP promoter or P_{GAP}) for constitutive expression of recombinant proteins in *P. pastoris* has been also used as an alternative methanol-free expression system (Zhang et al., 2009; Waterham et al., 1997). P_{GAP} regulates the levels of protein production, leading to strong and constitutive gene expression when the yeast grows on glucose, glycerol, or other carbon sources as a substrate, including methanol, and it is a promising alternative for industrial recombinant protein production using *P. pastoris*, as it avoids the drawbacks associated with methanol (García-Ortega et al., 2019; Gasset et al., 2022).

In conclusion, yeasts are versatile and efficient expression systems for recombinant protein production processes, and *P. pastoris* is a very established cell factory due to its efficient secretory pathway, strong promoters, and its ability to produce high levels of recombinant proteins. The engineering tools and bioprocess engineering know-how developed in *P. pastoris*, illustrated in **Figure 3**, are proof of the extensive use of this yeast.

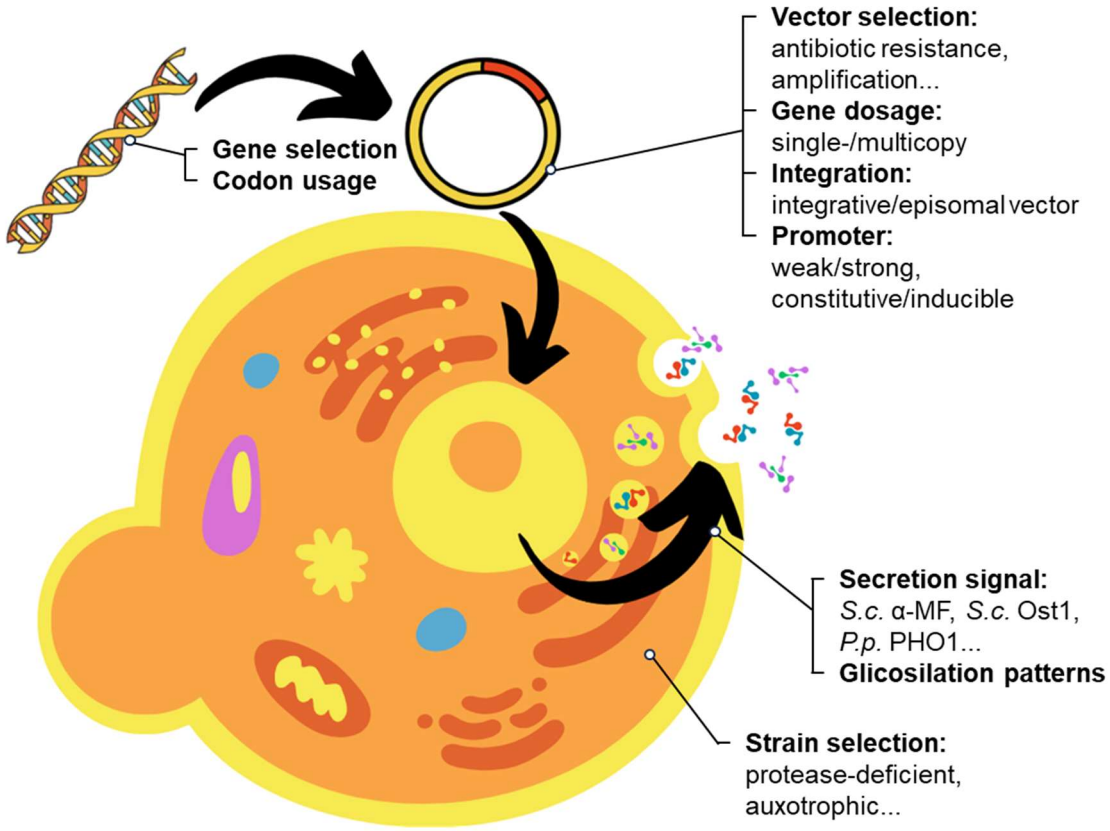


Figure 3. An overview of the know-how of recombinant gene expression with *P. pastoris*: from the DNA sequence optimization to strain selection.

2.4 Bioprocess engineering as a powerful tool to optimize RPP

Bioprocessing has been established as a more sustainable alternative to classic chemical industrial processes and for the synthesis of new products such as metabolites, proteins, or even whole cells. The optimization of a productive bioprocess for recombinant protein production involves a combination of strain engineering and bioprocess engineering, as presented in **Figure 4**. There are many different ways in which strain engineering can be applied. Some well-established examples are promoter engineering, codon usage optimization, co-expression of transcription factors or molecular chaperones, increasing gene dosage, etc. Nonetheless, the focus of this thesis is on bioprocess engineering, so it is not necessary to go into more detail since numerous examples can be found in the literature (Westbrook et al., 2019; De et al., 2021; Bustos et al., 2022). However, it must be said that strain engineering can easily improve productivity by more than one order of magnitude (Abad et al., 2011; Kroukamp et al., 2018).

Bioprocess engineering involves selecting the best bioreactor design and operational strategy, designing, and optimizing the process parameters, selecting appropriate culture media, and establishing the best conditions for the cells for optimal growth and production. In bioprocess optimization, selecting the correct operational mode is a key factor. Although there is a growing trend to move to continuous cultivation, batch or fed-batch are still the most established modes for production at an industrial scale, while continuous chemostat is often used for strain characterization (Looser et al., 2015; Kumar et al., 2020; Nieto-Taype et al., 2020b). In fed-batch processes, the feeding strategy is crucial to achieving high levels of production, since the expression of recombinant proteins is often coupled with cell growth, although in some cases this is not fulfilled (Bernat-Camps et al., 2023). Increasing the specific growth rate is often beneficial for protein production, and an optimal value can be found from which bioprocess efficiency no longer increases (Garrigós-Martínez et al., 2019; Brignoli et al., 2020). Also involved with the feeding strategy, the selection of the appropriate substrate is critical to achieve a productive bioprocess (García-Ortega et al., 2019).

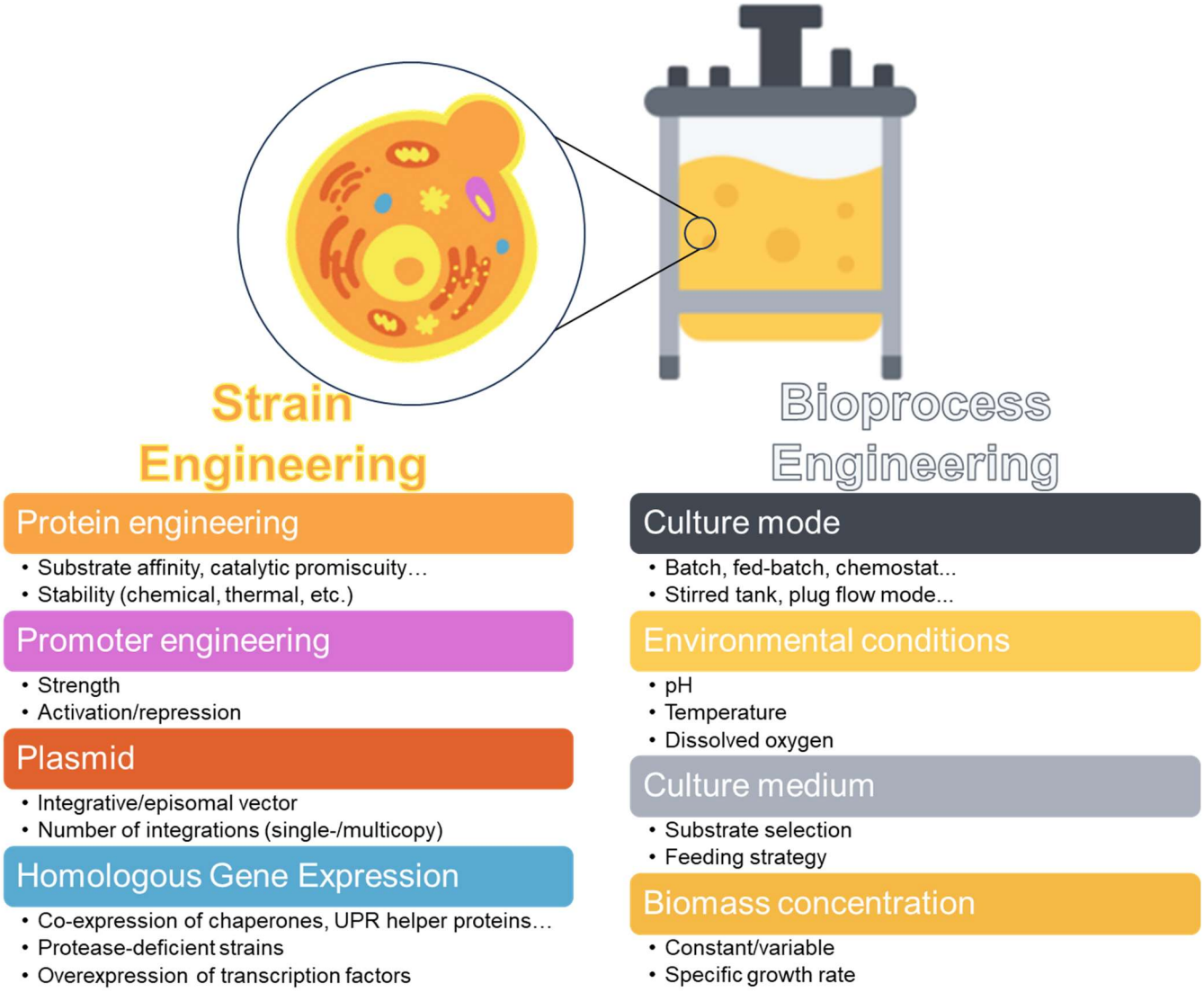


Figure 4. General considerations for recombinant protein production with *P. pastoris* for both micro- (strain optimization) and macro- (bioprocess engineering) scales. The genetic construction of the strain has a great influence on the design of the productive bioprocess.

On the other hand, maintaining the desired conditions along the bioprocesses is challenging, especially in dynamic cultures such as fed-batch. Therefore, control is a critical aspect of bioprocess engineering. Industrial control technology applied to bioindustry aims to monitor and regulate bioprocess performance to ensure the desired product quality and quantity, and it is crucial for ensuring the

reproducibility, scalability, and robustness of industrial bioprocesses. Bioprocess control involves the use of various tools and techniques, including statistical process control, process analytical technologies, and advanced process control systems, to monitor and optimize bioprocesses (Cos et al., 2006; Rathore et al., 2021).

Innovatively, the utilization of various cell stress methodologies has been found to exert a noteworthy and favorable impact on RPP bioprocesses. These well-established strategies, when appropriately employed, exhibit the potential to improve the overall bioprocess efficiency. For instance, one effective approach involves the implementation of substrate-starving periods during the bioprocess. By subjecting the cells to controlled periods of substrate deprivation, metabolic shifts are induced, prompting the cells to adapt and enhance their productivity (Garcia-Ortega et al., 2016; Çalık et al., 2018).

Similarly, another beneficial technique involves the manipulation of oxygen availability within the bioprocess environment. Oxygen limitation can be strategically induced, forcing cells to adjust their metabolic pathways to cope with reduced oxygen levels. This was demonstrated in previous studies, which presented compelling evidence regarding the impact of oxygen limitation on *P. pastoris* cultures grown on glucose as the sole carbon source. This limitation in oxygen availability or hypoxia was shown to trigger a notable increase in metabolic fluxes along the glycolytic and ethanol generation pathways, accompanied by a simultaneous reduction in TCA fluxes, indicating a shift from respirative to respiro-fermentative metabolism (Baumann et al., 2010). Moreover, these studies did not only report an increase in metabolic fluxes but also an upregulation of specific genes, both at the transcription and secretion levels. It is noteworthy that *TDH3* (glyceraldehyde-3-phosphate dehydrogenase) was one of the overexpressed genes. Remarkably, the GAP promoter or P_{GAP} is the natural promoter of the *TDH3* gene.

Subsequent studies demonstrated that under hypoxic conditions, the cells have a higher pool of NADH, which leads to an increase in the NADPH/$NADP^+$ ratio (Carnicer et al., 2012). In turn, this also could lead to higher recombinant protein production (Tomàs-Gamisans et al., 2020).

Additionally, other effects of oxygen limitation on *P. pastoris* were also described, which include lipidome, metabolome, and elemental composition, among others (Carnicer et al., 2009, 2012; Adelantado et al., 2017; Zahrl et al., 2017).

Succeeding research further investigated the effects of varying the degree of oxygen limitation on the production of an antibody fragment regulated by P_{GAP}. Astonishingly, these studies demonstrated that the application of distinct levels of oxygen limitation had a favorable influence on the specific production rate (q_P) of the antibody fragment, leading to an impressive up to 3-fold increase when compared to non-limiting conditions (Garcia-Ortega et al., 2017). As a result of these findings, it was proposed the adoption of a physiological and system-independent parameter as a transferable operating criterion to implement desired oxygen-limiting conditions in diverse equipment and operational modes. For this purpose, two parameters were suggested: the specific ethanol production rate (q_{EtOH}) and the respiratory quotient (RQ).

2.5 Towards Industry 4.0

Recently, the bioprocess engineering field has seen a rapid increase in the use of artificial intelligence (AI) algorithms to improve the efficiency, accuracy, and cost-effectiveness of bioprocesses. The incorporation of machine learning, deep learning, and other AI techniques facilitates the design, control, and monitoring of bioprocesses, resulting in higher yields, improved product quality, and reduced waste (Singh & Singhal, 2022). This trend is part of the larger movement towards Industry 4.0, which seeks to integrate advanced technologies such as AI, the Internet of Things (IoT), and cloud computing into industrial processes, leading to more flexible, responsive, and intelligent manufacturing systems. Therefore, the integration of Industry 4.0 into bioprocess engineering has presented new opportunities for optimizing and controlling bioprocesses, particularly in the context of process validation (PV) in drug manufacturing.

In 2011, the US Food and Drug Administration (FDA) published the "Guidance for Industry. Process Validation: General Principles and Practices" establishing some guidelines for the manufacture of drugs and biological products for humans and animals, including active pharmaceutical ingredients (FDA, 2011). Shortly after, in 2014 the European Medicines Agency (EMA) also joined in, publishing the "Guideline on process validation for the manufacture of biotechnology-derived active substances and data to be provided in the regulatory submission" (EMA, 2014). In both cases, they define PV as a systematic collection and analysis of data throughout the product and process lifecycle to ensure a consistent and high-quality product. It involves three stages: Process Design (PD), Process Qualification (PQ), and Continued Process Verification (CPV), as shown in **Figure 5**.

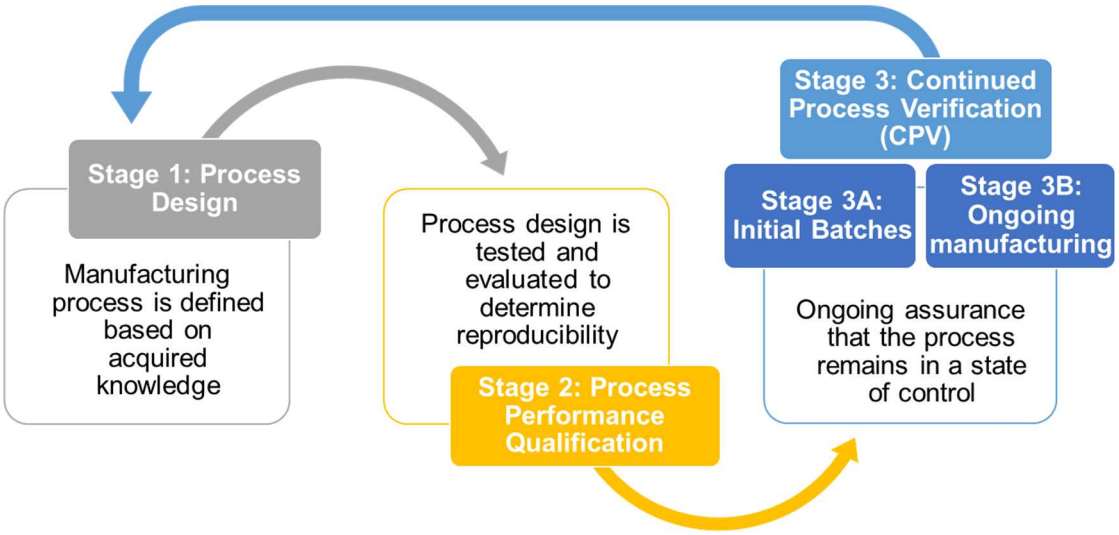


Figure 5. Stages of the Process Validation. Adapted from the "Guidance for Industry. Process Validation: General Principles and Practices" (FDA, 2011).

CPV, the third stage of PV, plays a crucial role in maintaining control over drug manufacturing by utilizing real-time data acquisition and direct interaction with production elements. However, the production environment is subject to uncertainty due to various factors like equipment stress, seasonality, and raw material suppliers. To address this, Stage 3 is divided into two phases: 3A, a preliminary production phase, and 3B, based on experimented iterations with successful batch releases (Ondracka et al., 2022).

In this sense, Statistical Process Control (SPC) is a valuable tool for monitoring drug manufacturing, enabling data-driven models to establish a mathematical mechanism, and therefore allowing to make decisions and interact with the system (Keller, 2011). Regulatory bodies encourage the use of multivariable statistical approaches and Process Analytical Technologies (PAT) to ensure the right batch production, but these mechanisms are not widely deployed (Read et al., 2010; FDA, 2019). Thus, in the last years, the tendency to bet on an experimental approach has begun, using AI as a multivariate statistical control to manage biotech processes (Duong-Trung et al., 2023). A digital twin, a virtualized

version of the physical system, is used for emulation, and AI algorithms are applied as valid analytical methods in pharmaceutical contexts (Bao et al., 2019). AI's ability to control complex processes with multiple variables, detect anomalies, and automate tasks adds value to CPV by keeping the process under control and reducing human-related issues during production (Manzano 2021).

In conclusion, the integration of Industry 4.0 and AI into bioprocess engineering has the potential to enhance process validation in drug manufacturing, ensuring consistent product quality and operational efficiency.

3. OBJECTIVES

As summarized in this thesis, all the work conducted during the doctorate was aimed at achieving two major objectives. The first main objective had a high chance of being achieved since it was focused on characterizing the effect of oxygen limitation (or hypoxia) on the production of a recombinant protein, *C. rugosa* lipase 1 (Crl1), under the regulation of the P_{GAP} with *P. pastoris*. This effect had already been characterized before with another protein (Fab) in a chemostat system, so the purpose of this study was to go a step further and test this effect in a dynamic system, more representative of an industrial process. This main goal was divided into five secondary objectives:

1. To characterize the effect of several degrees of hypoxia on the production of Crl1 under the regulation of P_{GAP} in chemostat mode and to define the optimal conditions regarding oxygen limitation for Crl1 production.

2. To evaluate the effect of different *CRL1* gene dosage on the production and expression of Crl1 and to test the synergistic effect with oxygen limitation.

3. To reproduce in a fed-batch system the optimal hypoxic conditions previously defined in chemostat operation, maintaining the same level of oxygen limitation throughout the whole process.

4. To analyze the effect of oxygen limitation on the transcription of several genes related to substrate consumption (glycolytic genes) and Crl1 production (the recombinant *CRL1* gene and other genes related to protein folding and secretion pathways) with the aim of identifying possible bottlenecks in transcription, translation, and/or secretion pathways.

5. To evaluate the reproducibility and transferability of the hypoxic production process to a higher scale process, validating the strategy from an industrial perspective.

After achieving the first key objective, the second one was to develop and test a control strategy to accurately maintain the required operating conditions regarding oxygen limitation during an entire fed-batch process. With this aim, several specific objectives were defined:

6. To define which variable/s must be controlled and which parameters must be manipulated to maintain a constant hypoxic level in a dynamic culture system.

7. To develop an automatic control loop based on some on-line measurements of the physiological state of the biomass, integrating all the available measurements coming from the bioreactor and external devices.

8. To test the control strategy with different *Pichia* strains expressing different recombinant proteins to validate both the control strategy and the effect of hypoxia.

Finally, with the advent of the “CPV of the Future” project and with collaboration with the AIZON company, an icing-on-the-cake final goal was defined:

9. To connect the *SCADA* system containing all the bioprocess measurements with a digital twin of the fermenter stored on the cloud.

10. To apply Artificial Intelligence (AI) algorithms to develop an innovative control strategy for the hypoxic process and contribute to setting the foundations for Industry 4.0.

4. METHODOLOGY

4.1 Strains and clones

Four clones of *P. pastoris* expressing different recombinant lipases were used in this work: two of them expressing *C. rugosa* lipase 1 (Crl1), another expressing the 28 aminoacids of prosequence jointly with the mature sequence of *R. oryzae* lipase (proRol) and the last one expressing the *C. antarctica* lipase B (CalB). Further details about the prosequence of proRol can be found in the literature (López-Fernández et al., 2019). These clones were kindly provided by Dr. Miguel-Ángel Nieto Taype (Crl1 and CalB clones) and Dr. Josu López Fernández (proRol clone).

Although the clones' construction was not part of this work, a brief explanation of the clone construction, screening, and gene copy number determination is given in the following sections. However, only information and details regarding Crl1-producing clones have been included as an example of clone construction.

More details concerning proRol- (López-Fernández et al., 2021a; López-Fernández, 2022) and CalB-producing clones (Nieto-Taype, 2020; Garrigós-Martínez et al., 2021) can be found in previous works conducted in the research group. However, it is noteworthy that the CalB-producing clone was built on the BG11 strain, while the remaining clones are derived from an X-33 strain.

4.1.1 Clones expressing *C. rugosa* lipase 1 (Crl1)

4.1.1.1 Crl1 clone transformation

Two recombinant clones of *P. pastoris* harboring one and five copies of the *CRL1* gene under the regulation of the constitutive *GAP* promoter (P_{GAP}) were selected among a battery of clones. The clones are named from now on as follows: Single-Copy Clone (SCC) and Multi-Copy Clone (MCC).

These clones were obtained by transforming a wild-type *P. pastoris* X-33 strain with different concentrations of the plasmid *PGAPαA* from Invitrogen (Carlsbad, CA, USA) containing the codon-optimized sequence of *CRL1* gene from GeneScript (Piscataway, NJ, USA) inserted through the restriction-ligation

method as described previously (Cos et al., 2005a; Cámara et al., 2016). By using the α-mating factor of *S. cerevisiae* the recombinant protein is secreted into the culture broth. Both clones had already been used in a previous study, thus the clone construction was already described in the literature (Nieto-Taype et al., 2020a).

4.1.1.2 Crl1 clone screening

For the selection of these clones, the protocol described in the manual *Pichia Expression Kit* from Invitrogen (Carlsbad, CA, USA) was followed.

Firstly, a pre-culture of 10 mL with YPG (2% w/v peptone, 1% w/v yeast extract, 2% w/v glycerol, pH 7.0) and 100 mg · L^{-1} of zeocin was inoculated for each clone and incubated 24 h at 25 ºC and 150 rpm of agitation. This pre-culture was used to inoculate a 40 mL culture in 0.25 L shake-flasks with BMG 1% medium (phosphate buffer 0.1 M pH 6.0, 1.34% YNB, 0.4 mg · L^{-1} biotin and 1% glycerol). The inoculum volume was adjusted to achieve an initial OD_{600} ≈ 0.1 in the shake-flasks. After approximately 16 h of incubation under the conditions noted above, each clone was inoculated in a separate shake-flask with an initial OD_{600} ≈ 0.1 with 40 mL BMD 1% medium (phosphate buffer 0.1 M pH 6.0, 1.34% YNB, 0.4 mg · L^{-1} biotin and 1% (w/v) glucose) for Crl1 expression. These cultures were kept under the same incubation conditions for approximately 48 h before Crl1 lipase activity determination.

4.1.1.3 *CRL1* copy number determination

Droplet digital PCR (ddPCR) was used to determine the number of the expression cassettes integrated. The protocol was adapted from a previous work (Cámara et al., 2016).

Firstly, genomic DNA was isolated using the commercially available DNA purification kit *QX200 ddPCR EvaGreen Supermix* (BioRad, Hercules, CA, USA). Then, to assure an accurate quantification, the genomic DNA was digested with

the restriction enzymes HindIII and EcoRI (NewEnglandBiolabs, Ipswich, MA, USA). *Droplet Generator* from BioRad (Hercules, CA, USA) was used to generate droplets containing PCR reaction mixture. Droplets were transferred to a 96-well plate and incubated at 95 ºC for 10 min for the activation of hot-start polymerase, followed by 40 cycles of two-step denaturalization (94 ºC, 30 s) and annealing/extension (60 ºC, 1 min). Droplet detection was done using the *QX200 Droplet Digital PCR System* and the software *QuantaSoft* from BioRad (Hercules, CA, USA).

More details can be found in the literature (Garrigós-Martínez et al., 2019; Nieto-Taype et al., 2020a).

4.2 Bioreactor operation

4.2.1 Inoculum preparation

4.2.1.1 Working cell bank

1.5 mL cryovials at OD_{600} ≈ 60 were prepared for each clone utilized in this work. Firstly, biomass from an agar plate was inoculated to a 1 L Erlenmeyer flask containing 150 mL of sterile YPG medium (1% yeast extract, 2% peptone, and 1% glycerol). When biomass was supposed to be at the exponential growth phase (OD_{600} ≈ 8 - 12) it was sterilely collected and centrifuged at 6000 xg and mixed with a volume of fresh YPG necessary to obtain an OD_{600} ≈ 90. Then, for each 1 mL of this concentrated biomass suspension, 0.5 mL of sterile glycerol/water 50% w/w was added, reaching a final OD_{600} ≈ 60 with a final glycerol concentration of 15% w/w, approximately. Finally, 1.5 mL aliquots were taken and introduced to cryovials and stored at -80 ºC for future utilization.

4.2.1.2 Pre-inoculum

For inoculum preparation, 1 to 3 cryovials prepared following the instructions of the previous section were mixed with 150 to 450 mL of sterile YPG medium, depending on the batch working volume in 1 to 3 sterile 1 L Erlenmeyer flasks (maximum 150 mL of YPG per flask). To assure the aseptic cell growth of the desired clones, 100 mg · L^{-1} zeocin were added to the inoculum medium.

Flasks were placed on a *Multitron Shaker* (INFORS HT, Bottmingen, Switzerland) and incubated at 25 ºC and 120 rpm for approximately 16 h until bioreactor inoculation.

4.2.2 Chemostat cultivations

Biological duplicates of chemostat cultivations were performed in a 2 L *Biostat B* fermenter (Braun Biotech, Melsungen, Germany) with a working volume of 1 L. Two clones were cultivated in chemostat mode, namely SCC and MCC. As detailed below, seven different conditions regarding oxygen supply were tested in duplicate for each clone.

4.2.2.1 Initial batch phase and bioreactor preparation

An initial batch phase was conducted before the chemostat operation. The batch medium contained per liter of distilled water: citric acid 1.8 g, glycerol 40 g, $(NH_4)_2HPO_4$ 12.6 g, $MgSO_4 \cdot 7H_2O$ 0.5 g, KCl 0.9 g, $CaCl_2 \cdot 2H_2O$ 0.02 g, antifoam *Glanapon 2000kz* (Bussetti & Co. GmbH, Vienna, Austria) 0.05 ml, biotin 0.4 mg, and 4.6 ml PTM_1 (trace salts solution). The total volume was 1L. The PTM_1 trace salts solution contained per liter: $CuSO_4 \cdot 5H_2O$ 6.0 g, NaI 0.08 g, $MnSO_4 \cdot H_2O$ 3.36 g, $Na_2MoO_4 \cdot 2H_2O$ 0.2 g, H_3BO_3 0.02 g, $CoCl_2$ 0.82 g, $ZnCl_2$ 20.0 g, $FeSO_4 \cdot 7H_2O$ 65.0 g, and 5.0 mL H_2SO_4 (98%).

Firstly, the culture medium without trace salts and biotin was introduced to the fermenter and the pH sensor was calibrated with pH standard solutions before placing it in the fermenter. The pH sensor used in chemostat cultivations was an *InPro 3030/225* from Mettler Toledo (Columbus, OH, USA). After that, a sterilization process (121 ºC, 30 min) was conducted with an *Autester Dry 80* autoclave (JP Selecta, Abrera, Spain). Then, biotin and trace salts solutions were sterilized by microfiltration with nylon filters and regenerated cellulose filters, respectively, both hydrophilic and with 0.2 µm of nominal pore size, and introduced through the septum. The commercial references of these filters are "*DMSO-Safe Acrodisc™17124311*" from Pall (Port Washington, NY, USA) and "*Syringe filter Minisart RC 17764-ACK*", from Sartorius (Goettingen, Germany), respectively.

After that, the dissolved oxygen (DO) sensor was calibrated. The DO sensor for the chemostat cultivations was a polarographic *InPro 6820/20/220* from Mettler

Toledo (Columbus, OH, USA). Afterward, the inoculum was introduced sterilely to the fermenter, leading to the beginning of the batch. The initial OD_{600} was between 1 and 2, leading to different batch durations depending on the fermentation schedule, following **Equation 1**.

$$X_t = X_0 \cdot e^{\mu \cdot t} \quad \textbf{Eq. 1}$$

Where X_0 and X_t are the biomass concentration at the beginning and the end of the batch, respectively (both in OD_{600} units or $g \cdot L^{-1}$), μ is the specific growth rate (h^{-1}) and t is time (h). X_t and μ were empirically defined for each clone.

Then, the necessary volume of pre-inoculum was defined with **Equation 2**,

$$V_1 \cdot OD_1 = V_2 \cdot OD_2 \quad \textbf{Eq. 2}$$

where V_1 is the batch volume (L) and OD_1 is the initial OD_{600} (X_0) defined in **Equation 1** (in OD_{600} units). OD_2 is the OD_{600} in the pre-inoculum shake flask (in OD_{600} units) and V_2 is the necessary volume of pre-inoculum (L). This volume was then sterilely centrifuged, resuspended with sterile water, and sterilely introduced to the bioreactor.

To reduce significatively lag phase at the beginning of the batch, it is noteworthy that the biomass used for the inoculum must be at exponential growth phase, so the OD_{600} in the pre-inoculum flasks should be $OD_{600} \approx 8$ - 12. Otherwise, an unexpectedly longer batch duration can be observed.

Temperature was kept at 25 ºC using the temperature control unit of the fermenter. Concretely, an external *Alpha RA8* cooling device or chiller (Lauda, Lauda-Königshofen, Germany) was used to supply cold water at 4 ºC to the control system due to intense heat generation within the cultivations. DO was maintained above 30% of air saturation by automatically modifying the agitation rate, using the fermenter's own DO cascade controller. No pure oxygen was needed to maintain the DO set-point neither the batch nor the chemostat operations. pH was controlled at 6.0, far from the pI (isoelectric point) = 4.5 of Crl1 (Benjamin & Pandey, 1998), using NH_4OH 15% (v/v). Additionally, NH_4OH also served as a nitrogen source. Evaporation losses were minimized using a gas

condenser with cooling water at 4 ºC. Moreover, a 250 mL desiccation column filled with silica gel was installed in the outgas stream to reduce its humidity and minimize water condensation in the exhaust gas filter. The filters for inlet gas and exhaust gas used in all fermenters (including chemostat and fed-batch fermentations) were made of polytetrafluoroethylene (PTFE) and hydrophobic, with a 0.2 µm of nominal pore size.

4.2.2.2 Chemostat operation

Right after the batch end, marked by a sudden rise of DO, the dilution rate (*D*) was set at 0.10 h^{-1} by controlling the out flowrate to 0.10 $L \cdot h^{-1}$. It is important to start the chemostat operation right after the end of the batch to have the biomass as close as possible to the exponential growth and avoid long dead times. Otherwise, long transitory states, fluctuations, and even bioreactor washout could be observed. To assure that the desired *D* was achieved, the outflow was measured every 2-3 residence times (τ), and the feeding pump power was adjusted depending on this measurement of the outflow, as detailed in previous works (Garcia-Ortega et al., 2016). The peristaltic pump used for these cultures was an *ISM935C* from Ismatec (Glattbrugg, Switzerland), and it was calibrated before each chemostat run. The chemostat feeding medium contained per liter of distilled water: citric acid 0.9 g, glucose 50 g, $(NH_4)_2HPO_4$ 4.35 g, $MgSO_4 \cdot 7H_2O$ 0.65 g, KCl 1.7 g, $CaCl_2 \cdot 2H_2O$ 0.01 g, antifoam *Glanapon 2000kz* (Bussetti & Co. GmbH, Vienna, Austria) 0.2 ml, biotin 1.2 mg, and 15 ml PTM_1 trace salts solution. Temperature, pH, and agitation were kept constant at 25 ºC, 6.0, and 700 rpm, respectively.

For the chemostat medium preparation, a 10 L or 20 L glass bottle with 2-3 L of water and all the necessary amount of antifoam were sterilized via autoclave (121 ºC, 30 min), since antifoam may block the filter used for medium sterilization. The rest of the medium components were dissolved in distilled water and then introduced to the 10 L or 20 L bottle through a polyethersulfone (PES) filter which consisted of a prefilter membrane of 0.45 µm of nominal pore size and an endfilter membrane of 0.2 µm nominal pore size. The commercial reference for this filter

is "*Sartopore 2 Midicap 5445307H8--OO—A*" from Sartorius (Goettingen, Germany).

Seven different conditions regarding oxygen availability were tested with the Crl1-producing Single- and Multicopy Clones (SCC and MCC) to evaluate the effect of oxygen limitation (hypoxia) and to compare with normal aerobic conditions (normoxia). To reduce the oxygen supply without altering the mixing conditions of the process, different mixtures of air and N_2 were implemented using two external *EL-FLOW* mass flow meters/controllers (Bronkhorst, Ruurlo, Netherlands), maintaining a constant aeration flowrate of 0.8 nlpm (normal liter per minute) (0.8 vvm). The flowrates of air and N_2, as well as the nominal corresponding oxygen proportion in the inlet gas, are displayed in **Table 2**. To assure an accurate respirometric parameter calculation, O_2 and CO_2 composition from inlet gas were analyzed in every steady state, and constantly from exhaust gas. For each condition tested, samples were taken and analyzed at 3, 4, and 5 residence times to check the achievement of steady state conditions. From these samples, cell density, lipolytic activity, and substrate and by-product quantification via HPLC were analyzed.

Table 2. Different air and nitrogen flowrates to achieve the desired oxygen concentrations in the inlet gas for chemostat cultivations.

Oxygen in the inlet gas (%)	**21**	**14**	**12**	**11**	**10**	**9**	**8**
Air flowrate (nlpm)	0.800	0.540	0.460	0.424	0.385	0.344	0.311
N_2 flowrate (nlpm)	0	0.260	0.340	0.376	0.415	0.456	0.489

As carried out in batch phase, temperature, and pH were kept at 25ºC and 6.0, respectively. DO was not controlled during chemostat operation since the effect of the oxygen availability on the Crl1 production process had to be evaluated and consequently diverse DO values were attained for the different conditions tested.

A picture and a schematic drawing of the experimental set-up for chemostat cultivations are displayed in **Figure 6**.

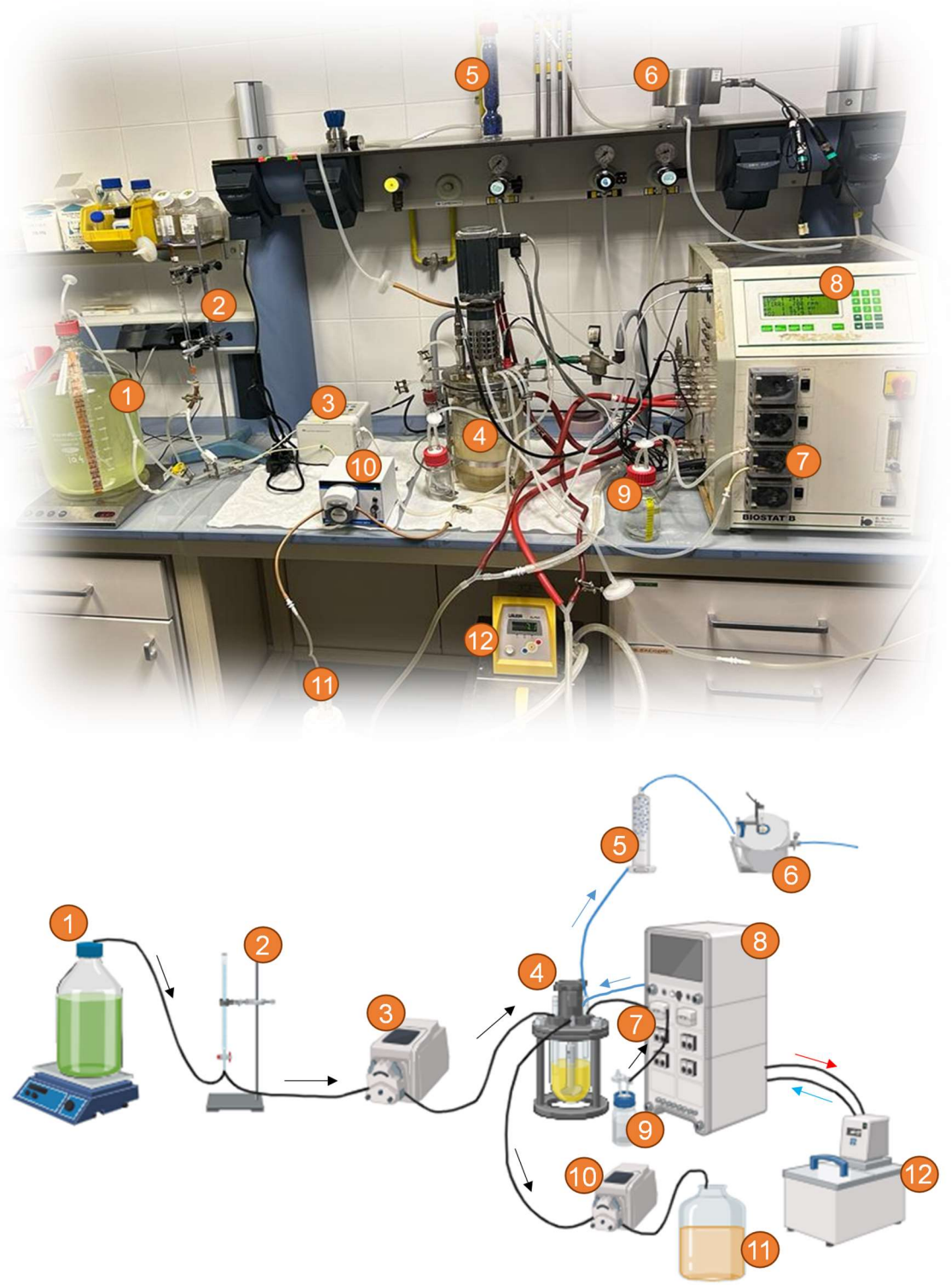


Figure 6. Picture and scheme of the experimental set-up for chemostat fermentations. **1**: substrate feeding tank on top of a magnetic stirrer; **2**: pipette for flowrate measurement; **3**: substrate addition peristaltic pump; **4**: bioreactor vessel; **5**: desiccation column; **6**: gas analyzer; **7**: bioreactor base pump; **8**: bioreactor control unit; **9**: base bottle; **10**: outflow peristaltic pump; **11**: outflow tank; **12**: chiller

Created with BioRender.com

4.2.3 Fed-batch cultivations

A total of 20 fed-batch cultures were performed with 4 different *P. pastoris* clones were included in this thesis. To provide a clear picture of all fed-batch fermentations and to make them easier to identify, in **Table 3** a classification of all of them is presented. For each fermentation, a codification related to the clone and O_2 condition has been assigned. “1” or “2” at the end of the code means they are replicates of the same clone and conditions. HIRQ and GASC stand for “High *RQ*” and “Gas-Controlled”, corresponding to a fermentation with a higher *RQ* set-point and a fermentation where *RQ* was controlled through gas mixing, respectively. In addition, they are subclassified by the *RQ* control strategy as explained in **Chapter 6**: Manual-Heuristic Control (MHC), Boolean-Logic Control (BLC), and AI-aided Adaptive-Proportional Control (AI-APC). In addition, as explained in more detail in **Sections 5.4.2** and **7.4**, two additional fermentations carried out in the Eng4Bio research group have been included for comparative purposes, although they are not part of this work.

Unless otherwise noted, all these fermentations have been carried out in two equal 5L *Biostat B* fermenters (Sartorius Stedim, Goettingen, Germany). NX50L and HPX50L fermentations were carried out in a 50 L fermenter prototype from ZETA GmbH (Lieboch, Austria), as part of a research stay in Bisy (Wünschendorf, Austria). Bisy is a research company specialized in promoter and strain engineering, mainly with *P. pastoris*. It has actively collaborated with the Eng4Bio research group, where this thesis has been done.

Table 3. Summary of the fed-batch cultivations included in this work.

** Fermentations included in this document that are not part of this thesis.

Codification	O_2 condition	Scale / Volume	Clone / Protein	*RQ* control strategy	*μ* (h^{-1})	Section
SCC-NX1	Normoxia	Lab / 5L	SCC	-	0.10	5.2
SCC-NX2	Normoxia	Lab / 5L	SCC		0.10	5.2
SCC-HPX1	Hypoxia	Lab / 5L	SCC	Manual-Heuristic	0.10	5.2, 6.3
SCC-HPX2	Hypoxia	Lab / 5L	SCC		0.10	5.2, 6.3
MCC-NX1	Normoxia	Lab / 5L	MCC	-	0.10	5.2
MCC-NX2	Normoxia	Lab / 5L	MCC		0.10	5.2
MCC-HPX1	Hypoxia	Lab / 5L	MCC	Manual-Heuristic	0.10	5.2
MCC-HPX2	Hypoxia	Lab / 5L	MCC		0.10	5.2
NX50L	Normoxia	Pilot / 50L	SCC	-	0.065	5.4
HPX50L	Hypoxia	Pilot / 50L	SCC	Manual-Heuristic	0.065	5.4
HPX-BLC1	Hypoxia	Lab / 5L	SCC	Boolean-Logic	0.10	6.4
HPX-BLC2	Hypoxia	Lab / 5L	SCC		0.10	6.4
HPX-APC1	Hypoxia	Lab / 5L	SCC	AI-aided Adaptive-Proportional	0.10	6.5
HPX-APC2	Hypoxia	Lab / 5L	SCC		0.10	6.5
proRoI-NX	Normoxia	Lab / 5L	proRoI	-	0.065	7.2
proRoI-HPX	Hypoxia	Lab / 5L	proRoI	Boolean-Logic	0.065	7.2
CalB-NX	Normoxia	Lab / 5L	CalB	-	0.065	7.3
CalB-HPX	Hypoxia	Lab / 5L	CalB	Boolean-Logic	0.065	7.3
HPX-HIRQ	Hypoxia	Lab / 5L	SCC	Manual-Heuristic	0.10	Annex
HPX-GASC	Hypoxia	Lab / 5L	SCC	Manual-Heuristic	0.10	Annex
** 065-NX	Normoxia	Lab / 5L	SCC	-	0.065	5.4, 7.4
** 065-HPX	Hypoxia	Lab / 5L	SCC	Boolean-Logic	0.065	5.4, 7.4

4.2.3.1 Initial batch phase and bioreactor preparation

A batch phase was carried out before fed-batch operation, following the same procedure described in **Section 4.2.2.1** with two exceptions in volume and cooling system, since different equipment was used, especially in the case of pilot-scale fermentations. In addition, different pH and DO sensors were utilized in each fermenter used.

Lab-scale batch phase and bioreactor preparation

For the 18 lab-scale cultivations performed in two *Biostat B* fermenters (Sartorius Stedim, Goettingen, Germany), they had an operational volume of ≈ 5 L, so the initial batch volume was 2 L. Additionally, the chiller used for these fermentations was a *Frigomix FX-1000* (Sartorius Stedim, Goettingen, Germany). Apart from that, the pH and DO sensors were an *EasyFerm Plus PHI VP 325* and a *VisiFerm DO 325*, respectively, both from Hamilton (Reno, NV, USA). On the other hand, for the 5 L fermenter set-up, a bigger desiccation column of 500 mL was used.

Pilot-scale batch phase and bioreactor preparation

The 2 pilot-scale fermentations were performed in a 50 L fermenter prototype from ZETA GmbH (Lieboch, Austria). From this point on, it will be referred to as "*BIRE*", which stands for "*Big Reactor"* as per the instructions of Bisy's staff.

On the one hand, the initial batch volume was 30 L, with the same culture medium composition described in **Section 4.2.2.1**. On the other hand, since the fermenter was designed for an industrial environment, it had its own temperature control unit, which was used to cool down and heat the cooling liquid (glycol water) for temperature control. In addition, the fermenter was connected to a boiler for a constant steam supply. This enabled to perform the sterilization *in-situ* following the standard procedure in industrial bioprocessing. Additionally, the pH and DO sensors used in these fermentations were *InPro 3100/120/Pt1000* and *InPro 6800/12/120*, respectively, both from Mettler Toledo (Columbus, OH, USA).

Contrarily to lab-scale fermentations, the *BIRE* was not equipped with a gas condenser to minimize evaporation losses. It is interesting to note that steam from the boiler was used to heat the outgas in order to prevent any water condensation that could block the output filter. In addition, no desiccation column was installed in the exhaust gas line of this fermenter. Thus, evaporation losses were expected to be higher than in the lab-scale.

4.2.3.2 Fed-batch operation

After the batch phase, a fed-batch phase was conducted with an exponential addition of glucose according to the different nominal specific growth rates selected. Some of the cultures were conducted under hypoxic conditions, as detailed below. However, to compare with non-limiting oxygen conditions, some of the cultivations were carried out in normoxic conditions, and they were considered as control experiments to assess the improvement in bioprocess efficiency when applying oxygen limitation.

After the end of the batch, indicated by a sudden rise of DO, the culture was kept around 1 h without substrate addition before starting the feeding phase to let biomass consume all potential by-products generated during the batch phase. Then, the addition pump was activated with an exponential pre-programmed feeding profile. For lab-scale fermentations, a piston pump *Burette 1S* (Crison, Alella, Spain) was used. This piston pump was named the “microburette”. On the other hand, for the pilot-scale fermentations, a peristaltic pump was used for the substrate addition, which was calibrated before the first run. As can be observed in **Table 3**, the nominal specific growth rate (μ) was set at 0.10 h^{-1} in some fermentations and at 0.065 h^{-1} in others by adjusting the feeding flowrate following **Equation 3**. To calculate the initial flowrate (F_0), **Equation 4** was used.

$$F_t = F_0 \cdot e^{\mu \cdot t} \qquad \textbf{Eq. 3}$$

$$F_0 = {X_0 \cdot V_0 \cdot \mu}\Big/{(S_0 \cdot Y_{X/S})} \qquad \textbf{Eq. 4}$$

Where F_t and F_0 are the feeding flowrates at t = t and t = 0, respectively ($L \cdot h^{-1}$), X_0 and V_0 are the biomass concentration (both in OD_{600} units or $g \cdot L^{-1}$) and the culture broth volume (L) at the beginning of the fed-batch phase, μ is the selected nominal specific growth rate (h^{-1}), t is the feeding phase time (h), S_0 is the substrate concentration in the feeding tank ($g \cdot L^{-1}$) and $Y_{X/S}$ is the overall biomass-to-substrate yield ($g_{DCW} \cdot g_S^{-1}$).

As performed in chemostat cultivations, temperature was kept at 25 ºC using the cooling jacket of the fermenter, and pH was maintained using NH_4OH 15% (v/v), with a set-point of 6.0 when working with SCC and MCC clones expressing Crl1 and a set-point of 5.0 with the other clones expressing proRol and CalB.

The feeding medium contained per liter of distilled water: glucose 400 g (unless otherwise specified), $MgSO_4 \cdot 7H_2O$ 6.45 g, KCl 10 g, $CaCl_2 \cdot 2H_2O$ 0.35 g, antifoam *Glanapon 2000kz* 0.2 ml, biotin 1.2 mg, and 15 ml PTM1 trace salts solution.

A rearrangement of **Equations 3** and **4** shows that biomass evolution depends on three factors: F_t, S_0, and $Y_{X/S}$, as represented in **Equation 5**. $Y_{X/S}$ is a biological parameter and highly dependent on oxygen availability, so F_t and S_0 are the only parameters that can be modified to achieve the desired biomass production.

$$X \cdot V \cdot \mu = F_t \cdot S_0 \cdot Y_{X/S} \qquad \textbf{Eq. 5}$$

Therefore, if the biomass production is required to be the same in normoxic and hypoxic conditions, where $Y_{X/S}$ differs notably, different $F_t \cdot S_0$ must be applied.

Thus, apart from the obvious exception of the pilot-scale fermentations, which had a higher volume and so, higher substrate flowrates, it is important to note that the initial flowrate (F_0) and the flowrate profile (F_t) were modified in some normoxic-hypoxic pair fermentations, whereas in other cases the glucose (and salts) concentration in the feeding tank (S_0) was the modified parameter between the normoxic and the corresponding hypoxic fermentation. This is explained in more detail in each subsection where different flowrate profiles or alternatively different glucose concentrations were implemented.

Regarding oxygen availability, in all normoxic fed-batches DO was maintained above 30% of air saturation, as done during the batch phase. This was achieved through the modification of the agitation rate during most of the fed-batch phases and additionally by mixing air with pure oxygen in the inlet gas during the last hours, when agitation was at maximum, except for one case (the pilot-scale normoxic fermentation) where fermenter overpressure was applied to increase oxygen solubility. A total of 7 fed-batches (SCC-NX1, SCC-NX2, MCC-NX1, MCC-NX2, and NX50L, proRol-NX, and CalB-NX) were done in normoxic conditions, as previously shown in **Table 3.**

On the other hand, 13 hypoxic fed-batches (SCC-HPX1, SCC-HPX2, MCC-HPX1, MCC-HPX2, HPX50L, HPX-BLC1, HPX-BLC2, HPX-APC1, HPX-APC2, proRol-HPX, CalB-HPX, HPX-HIRQ, and HPX-GASC) were included in this work, but not all of them were conducted following the same strategy regarding oxygen availability, *RQ* control, and agitation rate modification. Thus, the description of the methodology is classified based on the control strategy implemented during the feeding phase. Further details regarding the development of the different control strategies are included in **Chapter 6**.

Manual *RQ* control

The manual *RQ* control was applied in 4 lab-scale hypoxic fermentations and in the pilot-scale hypoxic fermentation (SCC-HPX1, SCC-HPX2, MCC-HPX1, MCC-HPX2, and HPX50L). Basically, it consisted of a manually controlled closed-loop system with the *RQ* and agitation rate as the controlled variable and agitation rate as the modified variable.

Thus, the target degree of oxygen limitation was reached by maintaining the *RQ* at the corresponding set-point value. To do so, *RQ* was calculated on-line through the "*CER/OUR/RQ*" Plugin from *BlueVis* software (BlueSens, Herten, Germany). Then, the agitation was modified manually following heuristic rules: to increase *RQ*, agitation was reduced to decrease the oxygen transfer rate to the culture broth and, therefore, decrease the oxygen uptake rate. To reduce *RQ*, the opposite action was performed. More details about the frequency, intensity, and optimization of the control actions are provided in **Chapter 6**.

In the case of lab-scale hypoxic fermentations with manual *RQ* control, the feeding medium composition was identical to the feeding composition of the associated normoxic fermentations (glucose concentration = 400 g · L^{-1}). Thereby, in order to achieve the same biomass production and specific growth rate as in normoxic conditions, the initial flowrate (F_0) and the flowrate profile (F_t) were higher than in the normoxic fermentations, since $Y_{X/S}$ was higher in normoxia than in hypoxia, as discussed in **Chapter 5**. This led to different biomass concentration and volume profiles but with comparable total biomass production, as discussed throughout the thesis.

On the other hand, two fermentations were carried out in a 50 L *BIRE* fermenter (HPX50L and NX50L) to test the performance of the hypoxic strategy with the manually controlled closed-loop system in pilot scale, and the results obtained are included in **Section 5.4**. These fermentations were intended to represent a change of scale from the lab-scale cultivations, so all parameters except for the volume were planned to be the same. Thus, the *RQ* control strategy for the hypoxic fermentation was the same as the one applied in the lab-scale, consisting of a manual modification of the agitation rate, and in the normoxic fermentation, the DO was kept at > 30%, as mentioned previously. The fed-bath phase of the pilot-scale fermentations was mainly conducted in the same manner as for the lab-scale cultures with only slight differences.

Firstly, as previously pointed out, substrate addition was achieved using an internal peristaltic pump of the fermenter equipment, which was calibrated only at the beginning of the normoxic fermentation since it was the first one to be done.

Secondly, glucose concentration in the feeding tank was increased in the hypoxic fermentation. Contrary to the lab-scale fermentations, where different feeding flowrate profiles were applied for the normoxic and hypoxic fermentations due to the distinct $Y_{X/S}$, in this case, the feeding flowrate profiles were the same and the S_0 was the modified variable. Thus, with the aim of reaching equal biomass concentration and volume profiles in both the normoxic and the hypoxic pilot-scale fermentations, S_0 = 440 g · L^{-1} in the hypoxic fermentation.

Finally, the μ selected was lower than in the lab-scale fermentations (μ = 0.065 h^{-1}) to avoid the need to enrich the inlet gas with pure oxygen, especially in the normoxic fermentation, as done in the 5 L scale. However, this was not enough to maintain the complete aerobic conditions at the last stages of the normoxic pilot-scale fermentation, and the internal pressure of the fermenter had to be slightly increased to increase oxygen solubility. This is discussed in **Section 5.4**.

Boolean logic *RQ* control

Four hypoxic fed-batches were carried out with Boolean logic *RQ* control mode. The procedure was the same as described for the "manual *RQ* controlled" fed-batches but with two important differences. On the one hand, glucose concentration in the feeding tank was increased (S_0 = 440 g · L^{-1}) for the same reason as in the pilot-scale hypoxic fermentation.

On the other hand, the *Eve* software (INFORS HT, Bottmingen, Switzerland) was used for these fermentations as a "Supervisory Control And Data Acquisition" or "*SCADA"* system (Ivarsson, 2017; Brunner et al., 2021). The software allowed the development of several soft sensors used for the *RQ* control, as further detailed in **Section 6.4**. So, in this case, the "*CER/OUR/RQ*" Plugin from BlueVis software was not necessary. Additionally, the automation of the control actions through *Eve* software also led to an increase in their frequency.

As discussed later, humidity both in the inlet and exhaust gas has a significant impact on *RQ* calculation. Thus, for these automated control fermentations, a 500 mL desiccation column filled with silica gel was installed also in the inlet gas stream before the inlet gas filter. Apart from that, in terms of methodology, no other variations concerning the manual control strategy were implemented in this experimental set-up.

AI-aided Adaptive-P *RQ* control

Finally, two fed-batches were conducted with SCC and with the Artificial Intelligence-aided Adaptive-Proportional *RQ* controller. However, apart from the *RQ* control strategy, the fermenter and peripherals set-up were the same as described for the Boolean logic *RQ*-controlled fermentations with no exceptions.

A picture and a schematic drawing of the experimental set-up for fed-batch cultivations are displayed in **Figure 7.**

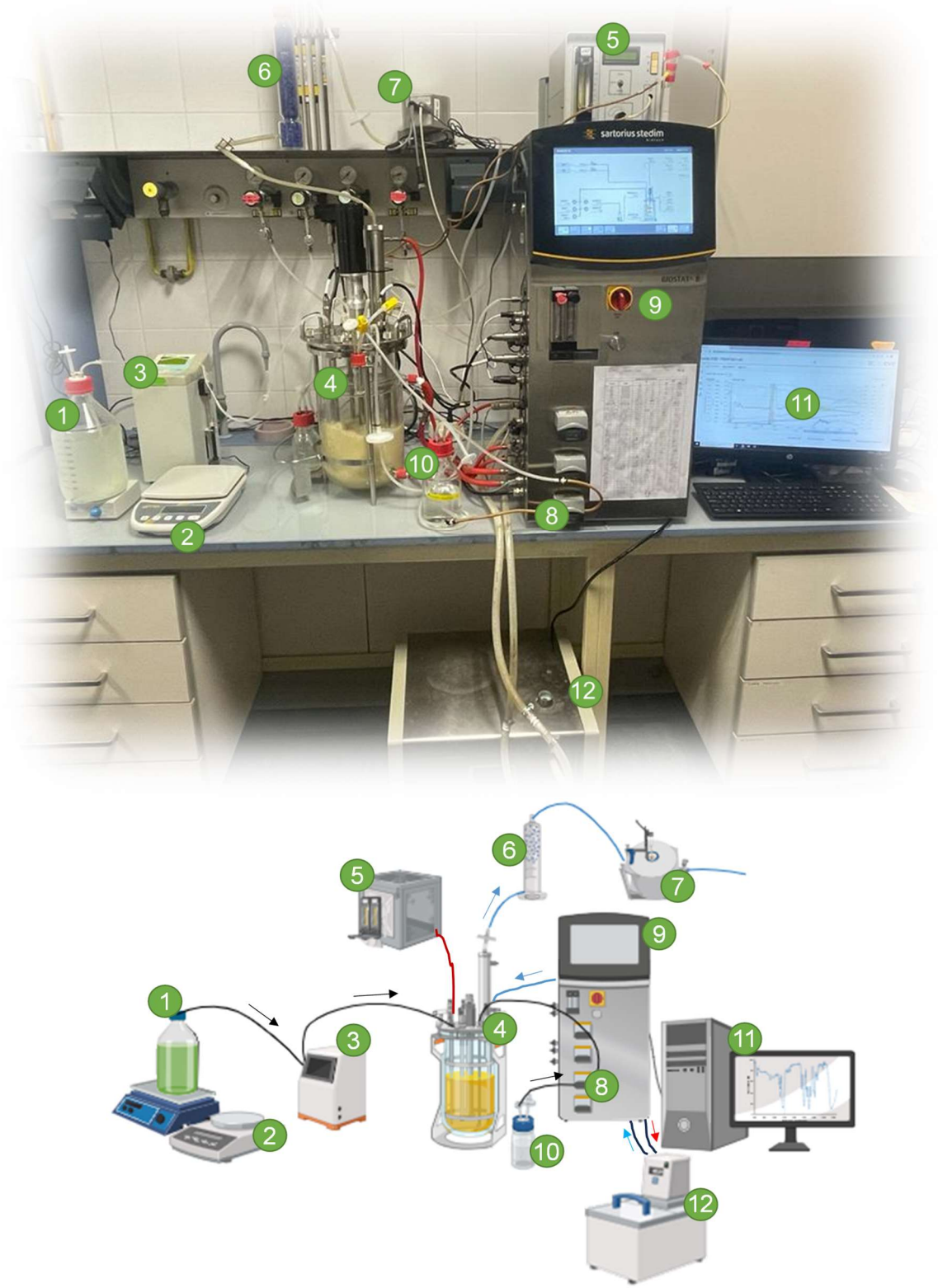


Figure 7. Picture and scheme of the experimental set-up for fed-batch fermentations. **1**: substrate feeding bottle on top of a magnetic stirrer; **2**: scale for flowrate verification; **3**: substrate addition piston pump (microburette); **4**: bioreactor vessel; **5**: ethanol sensor; **6**: desiccation column; **7**: gas analyzer; **8**: bioreactor base pump; **9**: bioreactor control unit; **10**: base bottle; **11**: computer for monitoring and control; **12**: chiller

Created with BioRender.com

4.3 Analytical methods

4.3.1 Biomass quantification by Optical Density (OD_{600})

To be able to make decisions during the process based on the biomass measurement, the optical density (OD) was measured at λ = 600 nm with a *DR 3900* spectrophotometer from Hach Lange (Düsseldorf, Germany). Samples were diluted until the OD_{600} was within the range from 0.2 to 0.8 to ensure a linear correlation between OD_{600} and biomass concentration. Dilutions were made with 0.9% NaCl dissolved in distilled water to avoid osmotic cell lysis.

4.3.2 Biomass quantification by Dry Cell Weight (DCW)

In addition to OD_{600} measurement, in order to have a precise and reliable biomass concentration value, as well as to be able to calculate the different process parameters in mass units, the value of the dry cell weight (DCW) was measured in quadruplicate for each fermentation sample.

Firstly, pre-labeled filters (0.7 µm) were dried for 24 h in a stove at 105 ºC and then placed in a desiccator for 2 h to cool down, before weighting with a precision scale *ME204* from Mettler (Columbus, OH, USA). Worth mentioning, "*Glass Fiber Prefilters*" (APFF04700, Millipore, Billerica, MA, USA) were used for DCW analysis.

After taking a bioreactor sample during a fermentation, these weighted filters were pre-moistened with 10 mL of 0.9% NaCl dissolved with distilled water, using a *Kitasato* flask connected to a vacuum pump and equipped with a magnetic filter funnel on the top. After that, 2 mL of fresh sample were filtered with the aid of the vacuum pressure. 2 mL of 0.9% NaCl were used to clean the micropipette's tip and added to the filtration cup. Finally, 10 mL of 0.9% NaCl were poured into the filtration cup to wash the filter and drag all the salts precipitated in the filter to avoid overweighting.

Again, the filters were dried for 24 h on a stove at 105 ºC and then placed in a desiccator for 2 h to cool down. Then the filters were weighed with a precision

scale to obtain the dry cell weight. It is important to ensure that after 24 h the weight of the filter does not decrease further, as this would indicate that it has not dried completely. If that were the case, it would be necessary to leave it in the oven for more hours until it has a constant weight.

When biomass in the bioreactor exceeds OD_{600} > 200 (≈ 60 $g_{DCW} \cdot L^{-1}$), only 1 mL of sample was necessary to obtain an accurate DCW determination. The Relative Standard Deviation (RSD) was always below 5%.

4.3.3 Biomass composition

With the aim of determining carbon and electron balances, biomass elemental composition was analyzed in terms of C, H, O, N, and S content.

Firstly, 40 mL samples from SCC producing Crl1 from a chemostat culture at an intermediate hypoxic level (*RQ* ≈ 1.4) were taken in a 50 mL falcon tube and centrifuged at 4 ºC and 4500 xg for 5 min. Then, the supernatant was discarded, and cells were resuspended with 35 mL of Tris-HCl 20 mM, pH 7.5, and centrifuged again with the same conditions. This washing step was repeated a total of three times. Then, the falcon tube containing the pellet was introduced overnight at the -80 ºC freezer. After that, the falcon tube was introduced to a freeze-dryer container and lyophilized for 24 h with a *VirTis Sentry* freeze dryer (SP Scientific, Warminster, PA, USA).

Lyophilized samples were manually ground and hermetically sealed until analysis. Elemental composition determination was done in triplicate with an elemental analyzer *EA CE 1108* for C, H, N, and S determination and *EA Flash 2000* for O determination, both from Thermo Fisher Scientific (Waltham, MA, USA).

Finally, to define the biomass' chemical formula, the percentages of each element were divided by its molecular weight and normalized by the C content. The molecular weight (MW) of biomass expressed in C-mol (grams of biomass per mole of carbon of the biomass) can be calculated using **Equation 6**:

$$MW\ C-mol = {}^{1}\Big/_{\frac{{}^{\%C}/_{100}}{12}} \qquad \textbf{Eq. 6}$$

Where *MW C-mol* is the molecular weight (MW) of biomass expressed in C-mol ($g \cdot C\text{-}mol^{-1}$) and *%C* is the mass percentage of carbon in the biomass ($g_{Carbon} \cdot g_{DCW}^{-1}$).

The elemental composition, chemical formula, and molecular weight in C-mol used within this work are displayed in **Table 4**.

Table 4. Chemical composition of the biomass considered in this work.

% C	% H	% O	% N	% S	Chemical formula	MW in C-mol
43.63 ± 0.17	6.66 ± 0.38	34.38 ± 0.41	7.50 ± 0.05	0.13 ± 0.02	$CH_{1.831}O_{0.591}N_{0.147}S_{0.001}$	27.50

4.3.4 Carbon source and by-products quantification

Glycerol, glucose, as well as potential fermentation by-products (mainly ethanol, arabitol, and succinate) were detected via HPLC. An *Ultimate 3000* system (Dionex, Sunnyvale, CA, USA) equipped with an *ICSep Coregel 87H3* ion exchange column (Transgenomic, Omaha, NE, USA) was used for the analyses. The detection of each component was performed with a *2410 Refractive Index Detector* from Waters Corporation (Milford, MA, USA). The column and software used for this purpose are further described elsewhere (Jordà et al., 2014). RSD was below 1%.

4.3.5 Ethanol concentration monitorization

In addition to the analysis of the ethanol concentration of the samples through HPLC, a *Methanol Sensor MRK002* from Raven Biotech (Vancouver, Canada) was used in all fermentations performed in the 5 L *Biostat B* fermenters to monitor

the ethanol production during the hypoxic feeding phases. Although commercialized as a methanol sensor and usually used for this purpose, this sensor is actually a volatile compounds sensor, and it is described as capable of measuring all volatile compounds that can diffuse through the probe membrane (Barrigon et al., 2015; Ponte et al., 2016; López-Fernández et al., 2019). Thus, it can be also used for ethanol measurement (Gasset et al., 2022; Ondracka et al., 2022).

After each fermentation, a correlation between the ethanol probe signal in millivolts and the real ethanol concentration measured by HPLC was experimentally determined following **Equation 7**.

$$C_{EtOH} = a \cdot (e^{(b \cdot (V - V_0)} - 1) \qquad \textbf{Eq. 7}$$

Where V is the signal of the probe (mV), V_0 is the basal signal of the probe (≈ 800 – 1000 mV), a (g · L^{-1}) and b (mV^{-1}) are the fitting coefficients of the correlation and C_{EtOH} is ethanol concentration (g · L^{-1}).

4.3.6 Inlet- and outgas analysis

A *BlueInOne FERM* gas analyzer (BlueSens, Herten, Germany) was used in both chemostat and fed-batch cultures. CO_2 and O_2 molar fractions and absolute humidity were monitored and recorded online from exhaust gas and intermittently measured from inlet gas. The gas analyzer was re-calibrated every steady-state in chemostat cultures and every fed-batch cultivation to assure an accurate measurement. For the gas analyzer calibration, several gas mixes with different and known CO_2 and O_2 concentrations were passed through the analyzer, and a linear correlation between CO_2 / O_2 and the sensor electronic signal (in millivolts) was established.

The data obtained were used to calculate the respirometric parameters: oxygen uptake rate (*OUR*), carbon dioxide evolution rate (*CER*), their corresponding specific rates (q_{O2} and q_{CO2}), and respiratory quotient (*RQ*). RSD was below 5%.

4.3.7 Lipolytic activity determination

Since in all the fermentations performed in this work, the recombinant proteins produced were lipases, an enzymatic assay based on colorimetric changes was used for product quantification.

The substrate of this enzymatic assay was para-nitrophenyl butyrate (p-NPB) (2635 – 84 – 9, Sigma Aldrich, Burlington, MA, USA), an ester that can be hydrolyzed by lipases to para-nitrophenol (p-NP) and butyrate. p-NP has a maximum of absorbance at 317 nm. However, at assay conditions, p-NP is spontaneously deprotonated to para-nitrophenolate (pK_a = 7.15 at 25ºC), which has a yellow coloration due to a maximum of absorbance at 405 nm. The isosbestic point, defined as the specific wavelength at which the total absorbance of the sample does not change during a chemical reaction such as a protonation/deprotonation, for para-nitrophenol/para-nitrophenolate, is at 348 nm (Biggs, 1954). One activity unit (AU) was defined as the amount of enzyme needed to obtain 1 µmol of product per min under assay conditions. The reaction scheme is shown in **Figure 8**.

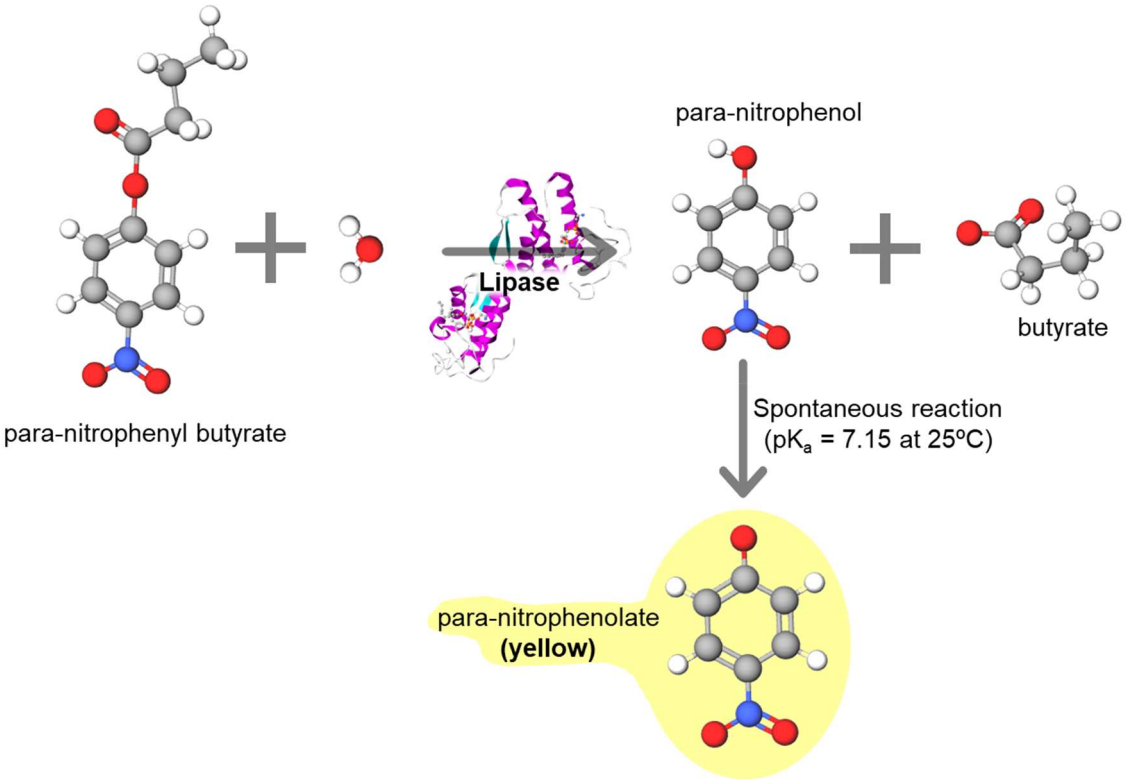


Figure 8. Scheme of the enzymatic hydrolysis of p-NPB catalyzed by a lipase, used in this work to determine the amount of lipase produced in each fermentation.

The assay conditions and duration were slightly different for each lipase, and so was the molar attenuation coefficient (ε). More details are given in the next sections. This enzymatic assay was adapted from literature, and variations of it (concerning temperature and pH) have been used in several laboratories (Chang et al., 2006).

4.3.7.1 Crl1 activity

For Crl1 enzymatic activity determination, the reaction buffer contained 1 mM p-NPB, 50 mM phosphate buffer (pH 7.00), and 4% (v/v) acetone. To prepare this reaction buffer, firstly 10 µL of pure p-NPB were dissolved in 2.5 mL of pure acetone. Subsequently, 0.4 mL of this mix was added to 9.6 mL of 50 mM phosphate buffer (pH 7.00) to obtain the reaction buffer. This way, p-NPB is 1 mM.

Then, after centrifuging the samples taken from the bioreactor for 90 s at 13400 rpm with a *MiniSpin Plus* (Eppendorf, Hamburg, Germany) and diluting the resulting supernatant with 50 mM phosphate buffer (pH 7.00) to enter in the selected range of the absorbance slope, 20 µL of properly diluted supernatant were mixed with 980 µL of reaction buffer, which was pre-heated at 30 ºC by using a thermoblock. The absorbance at λ = 348 nm was measured by using a *Specord 200 Plus* spectrophotometer from Analytic Jena (Jena, Germany) for 120 s, with temperature controlled at 30 ºC.

Absorbance slope was considered correct if it was within the range of 0.15 – 0.35 Absorbance Units · min^{-1}. If a lower or a higher slope was obtained, a different dilution had to be applied.

At these conditions (pH 7.00, temperature = 30ºC, phosphate buffer 50 mM, and a wavelength of λ = 348 nm) the molar attenuation coefficient (ε) was 5.47 $mM^{-1} \cdot cm^{-1}$. RSD was less than 5%.

4.3.7.2 proRol activity

As already mentioned, for proRol activity determination the substrate and product of the enzymatic reaction were the same, as well as the equipment used. However, in this case, the reaction buffer contained 1 mM p-NPB, 50 mM phosphate buffer (pH 7.25), and 4% (v/v) acetone. To summarize, the preparation was the same as for the Crl1 assay but with a slightly higher pH, since a proper comparison with other proRol-producing bioprocesses was intended (López-Fernández et al., 2019, 2020).

In the following step, 500 µL of sample supernatant diluted with 50 mM phosphate buffer (pH 7.25), was mixed with 800 µL of reaction buffer pre-heated to 30°C, and the absorbance at λ = 348 nm was measured during 420 s at a temperature of 30 ºC. Again, the absorbance slope was considered correct if it was within the range of 0.15 – 0.35 Absorbance Units · min^{-1}. With this assay conditions (pH 7.25, temperature = 30ºC, phosphate buffer 50 mM, and a wavelength of λ = 348), nm, and buffer composition ε = 5.37 mM^{-1}· cm^{-1}. RSD was less than 5%.

4.3.7.3 CalB activity

For the CalB enzymatic assay, the procedure was very similar to that for the previous lipases except for the buffer composition and the detection wavelength. In this case, for the preparation of the reaction buffer, 10 µL of p-NPB were dissolved in 109 µL of acetone. Then, 100 µL of this dissolved substrate were added to 9.9 mL of 300 mM Tris-HCl buffer (pH 7.00), giving a final concentration of 5.25 mM of p-NPB and 1% (v/v) acetone.

For the assay, 100 µL of supernatant from the centrifuged samples were diluted with 300 mM Tris-HCl buffer (pH 7.00) and then mixed with 900 µL of reaction buffer previously heated to 30 ºC. The absorbance was measured during 120 s at λ = 405 nm and with a temperature of 30 ºC. The absorbance slope was considered correct if it was within the range of 0.15 – 0.35 Absorbance Units · min^{-1}. With this assay conditions (pH 7.00, temperature = 30ºC, Tris-HCl buffer

300 mM and a wavelength of λ = 405), ε = 9.06 $mM^{-1}\cdot cm^{-1}$. RSD was less than 5%.

Importantly, in this case, a different wavelength was used for the enzymatic product detection. This implies that only the absorbance of the deprotonated form of p-NP (para-nitrophenolate) was measured. Thus, to have reliable and comparable results, it was imperative to adjust the buffer pH to 7.00. Otherwise, the pH of the buffer could have affected the activity measurement.

4.3.7.4 Determination of the molar attenuation coefficient (ε) of para-nitrophenol/para-nitrophenolate in different conditions.

For the determination of ε, six 1:2 serial dilutions of p-NP were prepared, starting from 1.265 mM and decreasing to 0.020 mM. The absorbance was measured for all dilutions, and a linear correlation between absorbance and p-NP concentration was established, discarding those points where linearity was not maintained. This was done in duplicate for each assay condition. The slope of this regression equals ε in $mM^{-1}\cdot cm^{-1}$ units.

4.3.8 Transcriptional analysis

Transcriptional analyses were carried out with samples taken from chemostats and for lab-scale SCC and MCC fed-batch fermentations.

4.3.8.1 RNA extraction

Samples of 1 mL were taken from both chemostat and fed-batch cultures, mixed with 0.5 mL of a 95% ethanol:5% phenol mixture, and centrifuged at maximum speed for 2 min at 4 ºC. Samples were then stored at -80 ºC until RNA extraction.

For RNA extraction, samples were first mixed with 300 µL of glass beads (425–600 µm, Sigma-Aldrich, St. Louis, MO, USA) and vortexed with a *TissueLyser* (Qiagen, Hilden, Germany). RNA extraction was then carried out with the *SV Total RNA Isolation System* kit (Promega, Madison, WI, USA) following the manufacturer's instructions.

4.3.8.2 cDNA synthesis and qPCR

cDNA synthesis was performed as reported previously (Garrigós-Martínez et al., 2019; Nieto-Taype et al., 2020a).

For the qPCR, a set of primers was designed to amplify the target cDNA, which are displayed in **Table 5**. The set of genes included: *CRL1* (the recombinant gene), *TDH3* (the gene encoding glyceraldehyde-3-phosphate dehydrogenase, which is the native gene expressed under the regulation of P_{GAP}), and the *PGK1* gene (which encodes phosphoglycerate kinase, another glycolytic enzyme). Additional genes involved in the unfolded protein response (UPR), such as *KAR2* and *HAC1*, were also studied (Raschmanová et al., 2019; Guerfal et al., 2010). To ensure maximum accuracy, reaction mixes were prepared by the EpMotion robot (Eppendorf, Hamburg, Germany), and the SYBR™ Select Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) was used for qPCR amplification. The qPCR was performed on the QuantStudio 12K Flex Real-Time equipment

from Thermo Fisher Scientific (Waltham, MA, USA) with an annealing/extension temperature of 57.4°C, following the same procedure as reported previously (Garrigós-Martínez et al., 2019; Nieto-Taype et al., 2020), with the only exception that in this case the β-actin gene (*ACT1*) was selected as the house-keeping gene, as done in previous transcriptional analyses under hypoxic conditions (Adelantado et al., 2017).

Table 5. Pairs of primers used for the qPCR analysis.

Gene	Primers' sequence
CRL1	CCTGAGGGTACTTACGAAG CCAGGTGGTCTAACAACG
TDH3	GGTGAGGTTTCTGCCAGC GTGGACTCAATGACGTAGTC
PGK1	AGAACGGTGGAACTGTCATCGT AAAGAAGCTCCTCCTCCAGTGG
HAC1	CATTACAGCAGGCTCCATC GTCAACTGATATGTGCCAAC
KAR2	GATGAAGTCGGGTCGTGTAC TCTTAGCAGCATCACCCAACC
ACT1	CACCACACCTTCTACAAC AGAAGGCTGGAACGTTG

4.4 Calculations

4.4.1 Mass balances, stoichiometric equations, and kinetic parameters

A single overall reaction defined on a C-molar basis could be used to describe the oxidative and oxidoreductive processes in which the culture growth consists of. This so-called “Black Box” model is a simplification of all the biochemical reactions involved and it is described by **Equation 8**.

$$S + Y_{O_2/S}O_2 \rightarrow Y_{X/S}X + Y_{CO_2/S}CO_2 + Y_{P/S}P \qquad \textbf{Eq. 8}$$

where *S* represents the single limiting substrate as the carbon and energy source; O_2, oxygen; *X*, biomass; CO_2, carbon dioxide, and *P* denotes the potential products or by-products. $Y_{i/s}$ are the stoichiometric coefficients, but they can also be referred to as overall "*i*" component-substrate yields. More generally, overall yields ($Y_{i/j}$) are defined as positive ratios between rates (r_i) or specific rates (q_i), as reflected in **Equation 9**. Accordingly, specific rates (q_i) are conversion rates related to the biomass concentration and can be defined for each component of the system as shown in **Equation 10**.

$$Y_{i/j} = \frac{r_i}{r_j} = \frac{q_i}{q_j} \qquad \textbf{Eq. 9}$$

$$q_i = \frac{r_i}{X} \qquad \textbf{Eq. 10}$$

The units of these parameters can be defined as: $Y_{i/j}$ ($g_i \cdot g_j^{-1}$), r_i ($g_i \cdot L^{-1} \cdot h^{-1}$), r_i ($g_i \cdot g_{DCW}^{-1} \cdot h^{-1}$), and *X*, which is the biomass concentration ($g \cdot L^{-1}$).

Thus, mass balance equations can be formulated for an ideal stirred tank-reactor, considering conversion rates of biomass formation, substrate uptake, oxygen consumption, carbon dioxide generation, and product or by-products formation, for both continuous operations at steady-state and fed-batch operation. These equations are necessary for the estimation of specific rates and overall yields, and thereby essential for the understanding of the behavior of the bioprocess under different operating conditions. Furthermore, they can be used to compare

different bioprocess efficiencies and to optimize the bioprocess variables to achieve the desired outcome. Therefore, a precise and accurate determination of the concentration of the main components is essential to obtain reliable results.

4.4.2 Correction of substrate and product concentrations

The concentration of biomass, substrate, product, and by-products was determined using the methods described in **Section 4.3**. Although the biomass concentration (*X*) was related to the total volume (V_{broth}), the substrates (S_{Liq}), products, and by-products (P_{Liq}) were measured on the supernatant after centrifuging the sample, without taking the volume of biomass into account. To overcome this issue, the substrate concentration (*S*), lipase titer (*P*), and by-products concentration (*EtOH*, *Ara*, *Suc*) were recalculated on the total volume using **Equation 11**,

$$C = C' \cdot \left(1 - \frac{X}{\sigma \cdot \rho}\right) \qquad \textbf{Eq. 11}$$

where *C* is the concentration or activity of any compound (substrate, lipase, by-products, etc. in $g \cdot L^{-1}$ or $AU \cdot L^{-1}$), *X* is the biomass concentration ($g \cdot L^{-1}$), σ is the fraction of dry matter in the biomass ($g \cdot g^{-1}$) and ρ is the yeast density ($g \cdot L^{-1}$), which had been experimentally determined in the literature as $\sigma = 0.304\ g \cdot g^{-1}$ and $\rho = 1068\ g \cdot L^{-1}$, respectively (Barrigón et al., 2013; Garcia-Ortega et al., 2013).

4.4.3 Determination of kinetic parameters in chemostat cultures

In a continuous cultivation operated at steady state, the time derivative is eliminated, as well as derivatives with respect to the volume. This is because in a steady state all variables remain constant, and it makes the determination of the kinetic parameter quite easier.

Equation 12 describes mass balance equations for chemostat operation,

$$\begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} \cdot XV = \begin{bmatrix} F_{out} \cdot X \\ -F \cdot S_0 + F_{out} \cdot S \\ F_{out} \cdot P \\ OUR \cdot V \\ CER \cdot V \end{bmatrix} \qquad \textbf{Eq. 12}$$

where μ is the specific growth rate (h^{-1}); q_S, specific substrate uptake rate ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); q_P, specific production rate ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); q_{O2}, specific oxygen uptake rate ($mol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); q_{CO2}, specific carbon dioxide production rate ($mol_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); F, substrate feeding rate ($L \cdot h^{-1}$); F_{out}, outlet flow rate ($L \cdot h^{-1}$); V, volume of broth in the reactor (L); S_0, substrate feeding concentration ($g \cdot L^{-1}$); OUR, oxygen uptake rate ($mol_{O2} \cdot L^{-1} \cdot h^{-1}$); CER, carbon dioxide evolution rate ($mol_{CO2} \cdot L^{-1} \cdot h^{-1}$). For the specific case of ethanol, an additional term was included to correctly estimate its corresponding specific rate taking into account stripping losses.

In chemostat, F_{out} can be obtained by the total mass balance for an ideal stirred tank reactor in continuous operation at steady state, as reflected by **Equation 13**,

$$F_{out} = \frac{\rho_{Feed} \cdot F - \rho_{H_2O} \cdot F_{Evap} + \rho_{Base} \cdot F_{Base} - \rho_{Bro} \cdot F_o + M_{GAS}}{\rho_{Broth}}$$

Eq. 13

where ρ_{Feed} denotes the substrate feeding density ($g \cdot L^{-1}$); ρ_{H2O}, water density ($g \cdot L^{-1}$); F_{Evap} is the water evaporation rate ($L \cdot h^{-1}$); ρ_{Base}, NH_4OH 15% (v/v) density ($g \cdot L^{-1}$); F_{Base}, base addition rate ($L \cdot h^{-1}$); ρ_{Broth}, mean broth density ($g \cdot L^{-1}$); F_o, exit flowrate ($L \cdot h^{-1}$); and finally, M_{GAS} is the net mass gas flowrate ($g \cdot h^{-1}$). The latter, in turn, is defined by **Equation 14**,

$$M_{GAS} = -(W_{O_2} \cdot OUR \cdot V + W_{CO_2} \cdot CER \cdot V) \qquad \textbf{Eq. 14}$$

where W_{O2} is the oxygen molecular weight ($32\ g \cdot mol^{-1}$) and W_{CO2} is the carbon dioxide molecular weight ($44\ g \cdot mol^{-1}$).

4.4.4 Determination of kinetic parameters in fed-batch cultures

In a fed-batch culture, all variables change over time, including volume. The calculation of key process parameters is therefore not trivial, being necessary to make an estimation of the volume throughout fed-batch fermentation as described below.

Equation 15 describes mass balance equations for fed-batch operation,

$$d\begin{bmatrix} XV \\ SV \\ PV \\ O_2V \\ CO_2V \end{bmatrix} \Big/ dt = \begin{bmatrix} \mu \\ q_S \\ q_P \\ q_{O_2} \\ q_{CO_2} \end{bmatrix} \cdot XV + \begin{bmatrix} -F_o \cdot X \\ F \cdot S_0 - F_o \cdot S \\ -F_o \cdot P \\ OTR \cdot V - F_o \cdot O_2 \\ -CTR \cdot V - F_o \cdot CO_2 \end{bmatrix} \qquad \textbf{Eq. 15}$$

where V is volume of broth in the fermenter (L); μ, specific growth rate (h^{-1}); q_S, specific substrate uptake rate ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); q_P, specific production rate ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); q_{O2}, specific oxygen uptake rate ($mol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); q_{CO2}, specific carbon dioxide production rate ($mol_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); F_o, outlet flowrate ($L \cdot h^{-1}$), which includes sampling; F, substrate feeding rate ($L \cdot h^{-1}$); S_0, substrate feeding concentration ($g \cdot L^{-1}$); S, substrate concentration in the broth ($g \cdot L^{-1}$); *OUR* oxygen uptake rate ($mol_{O2} \cdot L^{-1} \cdot h^{-1}$); and *CER* carbon dioxide evolution rate ($mol_{CO2} \cdot L^{-1} \cdot h^{-1}$), being *OUR* and *CER* considered equal to *OTR* and *CTR,* respectively (oxygen and carbon dioxide transfer rates). Again, for the case of ethanol, a stripping term was included. Although included in **Equation 15**, the "exit" term of the mass balances was considered negligible, since it only included very small sample volumes.

In order to estimate the culture volume, mass balance equations can be applied for the total mass of broth in the reactor, taking into account several elements such as the feeding rate of substrate, addition of ammonium hydroxide used in the pH control loop, volume of the samples withdrawn from the bioreactor, water evaporation losses, and gaseous exchange term. The expression used for this approximation is shown in **Equation 16**,

$$\frac{dV}{dt} = \frac{\rho_{Feed} \cdot F - \rho_{H_2O} \cdot F_{Evap} + \rho_{Base} \cdot F_{Base} - \rho_{Brot} \cdot F_S + M_{GAS}}{\rho_{Broth}}$$

Eq. 16

where V is the volume of broth in the bioreactor (L); ρ_{Feed}, feeding medium density (g · L^{-1}); F, flowrate of feeding medium (L · h^{-1}); ρ_{Base}, NH_4OH 15% (v/v) density (g · L^{-1}); F_{Base} base addition rate (L · h^{-1}); ρ_{Broth}, broth density (g · L^{-1}); F_S is the sampling withdrawal (L · h^{-1}) and M_{GAS} is the net mass gas flowrate (g · h^{-1}), which is defined by **Equation 14** from the previous **Section 4.4.3**.

Final broth densities and volumes were determined gravimetrically at the end of the fermentations, obtaining slight deviations between estimated and measured volumes. Specific correction factors for the volume profiles were therefore applied to all the fermentations.

Then, since some variables may be time-varying, a distinction between discrete values and mean values comes into necessity in the calculation of key process parameters.

Discrete values can be calculated at each time point considering the variation between two consecutive samples and they are useful to analyze whether the fed-batch process has reached a pseudo steady-state or not. On the other hand, mean values are mean specific rates of the overall feeding phase, and they are of great interest when comparing different processes or strategies.

Determination of discrete values of specific rates

The estimation procedure for determining the discrete rates or time-dependent specific growth rate ($\mu(t)$), specific substrate consumption rate ($q_S(t)$), specific product (or by-product) formation rate ($q_P(t)$), specific O_2 consumption rate $q_{O2}(t)$, and specific CO_2 production rate $q_{CO2}(t)$ was adapted from a previous study(Cos, Serrano, et al., 2005). Initially, the global state variables (biomass (*XV*), substrate (*SV*), and product (*PV*)) were estimated using off-line data within the feeding phase. The smoothing tool *Matlab R2015a Curvefit Toolbox* (The Mathworks Inc.,

Natick, MA, USA) was applied to obtain the smoothed curves of the global state variables, and their first time-derivatives were calculated. In parallel, on-line data of O_2 and CO_2 concentration was used to calculate *OUR* and *CER*, as explained in more detail in **Chapter 6**. Finally, the specific rates of $\mu(t)$, $q_S(t)$, and $q_P(t)$ were obtained by applying their corresponding mass balances, as described by **Equations 17**, **18**, and **19**, respectively. On the other hand, $q_{O2}(t)$, and $q_{CO2}(t)$ were calculated by dividing *OUR* and *CER* (calculated on-line) by the biomass concentration measured from off-line samples, as shown in **Equations 20** and **21**. The uncertainties in the discrete specific rates can be calculated by error propagation.

$$\mu_{(t)} = \frac{1}{(XV)_t} \cdot \frac{d(XV)_t}{dt} \qquad \textbf{Eq. 17}$$

$$q_{S\,(t)} = -\frac{1}{(XV)_t} \cdot \left(F_t \cdot S_0 - \frac{d(SV)_t}{dt}\right) \qquad \textbf{Eq. 18}$$

$$q_{P\,(t)} = \frac{1}{(PV)_t} \cdot \frac{d(PV)_t}{dt} \qquad \textbf{Eq. 19}$$

$$q_{O_2\,(t)} = \frac{OUR_t}{X_t} \qquad \textbf{Eq. 20}$$

$$q_{CO_2\,(t)} = \frac{CER_t}{X_t} \qquad \textbf{Eq. 21}$$

The units of all variables included in these equations are the same as in **Equation 15**, in this case expressed as discrete values.

Determination of mean values of specific rates

Several methods have been proposed in the literature for estimating the mean specific induction rates, including arithmetic mean rates, time-weighted average rates, and linear regressions (Potgieter et al., 2010; Albaek et al., 2011; Garcia-Ortega et al., 2013). However, the first two methods require the calculation of discrete specific rates for each off-line value, which can lead to increased

estimation errors due to the need for calculating first time-derivatives of the global variables. In this work, a linear regression method was employed to estimate the mean specific rates without the need for calculating first-time derivatives.

The method utilized off-line data of some of the global state variables (biomass, substrate, and product concentrations), while for other variables on-line data was available and therefore used for the calculations (feed flow rate and O_2 and CO_2 concentrations in the exhaust gas). **Equations 22**, **23**, **24**, **25,** and **26** show the linear regression equations applied for mean specific rates calculation. It is noteworthy that in this study several potential by-products were considered as product.

$$\int_{(XV)_0}^{(XV)} d(XV) = \mu_{mean} \cdot \int_{t_0}^{t} (XV)dt \qquad \textbf{Eq. 22}$$

$$S_0 \cdot \int_{t_0}^{t} F\ dt - \int_{(SV)_0}^{(SV)} d(SV) = -q_{S_{mean}} \cdot \int_{t_0}^{t} (XV)dt \qquad \textbf{Eq. 23}$$

$$\int_{(PV)_0}^{(PV)} d(PV) = q_{P_{mean}} \cdot \int_{t_0}^{t} (XV)dt \qquad \textbf{Eq. 24}$$

$$\int_{t_0}^{t} OUR \cdot V\ dt = q_{O_2 mean} \cdot \int_{t_0}^{t} (XV)dt \qquad \textbf{Eq. 25}$$

$$\int_{t_0}^{t} CER \cdot V\ dt = q_{CO_2 mean} \cdot \int_{t_0}^{t} (XV)dt \qquad \textbf{Eq. 26}$$

The units of all variables included in these equations are the same as in **Equation 15**, in this case expressed as mean values.

4.4.5 Calculation of bioprocess yields and productivities

Yields are critical parameters in bioprocess engineering as they represent the conversion efficiency of one variable with respect to another. For example, the overall biomass-to-substrate yield is defined as the amount of biomass produced per unit of substrate consumed.

However, these coefficients are not constant throughout the entire fermentation process, as they rapidly change according to the energy needs of the microorganism. Additionally, there is a close relationship between the overall biomass-to-substrate yield and the substrate maintenance coefficient (Pirt, 1982).

They can be calculated over a specific interval of the fermentation process, resulting in partial overall yields, or alternatively, over the entire fermentation process, resulting in total overall yields.

As mentioned in **Section 4.4.1**, overall yields can be estimated as quotients between ratios (r_i) or specific ratios (q_i), as shown in **Equations 9** and **10**. However, if overall yields are to be calculated as mean values over a fed-batch culture, a similar approach to that used to calculate specific rates can be employed based on measurements of the state variables *X*, *S* and *P*, as well as *V*, *F*, and S_0 as functions of time, as shown in **Equations 27** and **28** for the overall biomass-to-substrate yield ($Y_{X/S}$) and the overall product-to-biomass yield ($Y_{P/X}$). On the contrary, for chemostat cultivations, overall yields can easily be obtained by dividing rates (r_i) or specific rates (q_i).

$$\int_{(XV)_0}^{(XV)} d(XV) = Y_{X/S} \cdot \left(S_0 \cdot \int_{t_0}^{t} F \, dt - \int_{(SV)_0}^{(SV)} d(SV) \right) \qquad \textbf{Eq. 27}$$

$$\int_{(PV)_0}^{(PV)} d(PV) = Y_{P/X} \cdot \int_{(XV)_0}^{(XV)} d(XV) \qquad \textbf{Eq. 28}$$

Where X and S are biomass and substrate concentrations ($g \cdot L^{-1}$) in the culture broth, P is the product titer ($AU \cdot L^{-1}$), S_0 is the substrate concentration in the feeding tank ($g \cdot L^{-1}$), F is the feeding flowrate ($L \cdot h^{-1}$), V is broth volume (L), $Y_{X/S}$ is the overall biomass-to-substrate yield ($g_{DCW} \cdot g_S^{-1}$) and $Y_{P/X}$ is the overall product-to-biomass yield ($AU \cdot g_{DCW}^{-1}$).

On the other hand, productivity is defined as the ratio of the product obtained to the factors required to produce it during a specific time interval. Specifically, in bioprocess engineering, volumetric productivity or Space-time-yield (Q_V or STY) is defined as the amount of product formed per unit of volume and time, while specific productivity (Q_e) is calculated as the amount of product formed per unit of cell mass and time, as shown in **Equations 29** and **30**.

$$Q_V = \frac{\Delta(PV)}{V_f \cdot \Delta t} \qquad \textbf{Eq. 29}$$

$$Q_e = \frac{\Delta(PV)}{(XV)_f \cdot \Delta t} \qquad \textbf{Eq. 30}$$

Where Q_v is the volumetric productivity or Space-time-yield ($AU \cdot L^{-1} \cdot h^{-1}$), Q_v is the specific productivity ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$), X and P are biomass concentration ($g \cdot L^{-1}$) and product titer ($AU \cdot L^{-1}$), V and V_f volume and final volume, respectively (L) and t is time (h).

Similar to yields, productivities can be calculated over a specific time interval or the entire culture, with the latter providing more valuable information.

In a chemostat culture, the specific productivity Q_e and the q_P are the same, since a steady state for X, P, and V is reached and therefore they do not vary over time. In this case, the residence time (D^{-1}) is considered as Δt in **Equation 30**.

Consequently, from the application of product mass balances in a chemostat, the following relation shown by **Equation 31** can be derived:

$$q_P = \frac{D \cdot (P - P_0)}{X} = Q_e = \frac{D \cdot \Delta P}{X} \qquad \textbf{Eq. 31}$$

Where q_P is the specific production rate (AU · g_{DCW}^{-1} · h^{-1}), D is the dilution rate (h^{-1}), P and P_0 are the product titers in the broth and in the feeding flowrate, the latter being normally negligible (AU · L^{-1}), Q_e is the specific productivity (AU · g_{DCW}^{-1} · h^{-1}), and X is the biomass (g · L^{-1}).

However, in batch or fed-batch modes, the specific productivity (Q_e) is a ratio between variables calculated at the endpoint, contrarily to the specific product formation rate (q_P), which is calculated within the process evolution, although both have the same units (AU · g_{DCW}^{-1} · h^{-1}).

From the equation that allows the calculation of q_{Pmean} (**Eq. 24**), another relationship including q_{Pmean} and $(XV)_{mean}$ can be derived as shown in **Equation 32**. Note that all variables shown in the following equations and their corresponding units have already been described in previous equations, using the same abbreviations.

$$\Delta(PV) = q_{P_{mean.}} \cdot (XV)_{mean} \cdot \Delta t \qquad \textbf{Eq. 32}$$

So, in combination with **Equation 30**, the ratio between $q_{Pamean.}$ and Q_e can be easily obtained, as defined by **Equation 33**:

$$\frac{q_{P_{mean}}}{Q_e} = \frac{(XV)_f}{(XV)_{mean}} \qquad \textbf{Eq. 33}$$

A detailed explanation can be stated from the definition of $(XV)_{mean}$ itself. In its calculation, a time mean is done following **Equation 34**:

$$(XV)_{mean} = \frac{1}{\Delta t} \int_{t_0}^{t} (XV) dt \qquad \textbf{Eq. 34}$$

Thus, for a process where μ is rather constant, the time evolution of the total biomass $(XV)_t$ follows an exponential trend, as explained by **Equation 1** in **Section 4.2.2.1** for biomass concentration, but it can also be applied to total biomass (XV), as shown in **Equation 35**:

$$XV = X_0 V_0 \cdot e^{\mu \cdot (t - t_0)} \qquad \textbf{Eq. 35}$$

Equation 35 can be integrated with $t_0 = 0$ to give a new valuable expression for $(XV)_{mean}$, demonstrated by **Equation 36**.

$$(XV)_{mean} = \frac{(XV)_f - X_0 V_0}{\mu \cdot \Delta t} \approx \frac{(XV)_f}{\mu \cdot \Delta t} \qquad \textbf{Eq. 36}$$

In turn, it can be combined with **Equation 33**, and considering an exponential cell growth, **Equation 37** can be obtained, a worthy expression to compare $q_{Pmean.}$ and Q_e values easily.

$$\frac{q_{P_{mean}}}{Q_e} \approx \mu \cdot \Delta t = \ln\left(\frac{(XV)_f}{X_0 V_0}\right) \qquad \textbf{Eq. 37}$$

Where Δt here refers to the total cultivation time (h) considered for $q_{Pmean.}$ and Q_e calculations.

Therefore, it can be concluded that the ratio between $q_{Pmean.}$ and Q_e depends on the total biomass reached in the process related to its initial value.

Therefore, in those cases when q_P and Q_e are not both calculated in the same way, such is a batch or a fed-batch cultivation, and their values vary depending on how and how much total biomass is increased during the process, is quite common to obtain up to 2-3 fold higher q_P than Q_e. In the present work, this ratio had values always around 2.

4.4.6 Consistency check and data reconciliation

The accurate determination of specific rates and yields is essential in bioprocess engineering to evaluate the performance of a system. However, the calculation of these parameters can be influenced by random errors, drifts, and gross errors, which can affect their reliability. Mean values or moving average methods are typically used to reduce random noise. Elemental balances are applicable constraints that can remove measurement errors with little prior knowledge (van der Heijden et al., 1994a). Specifically, carbon and redox balances are often used as constraints to evaluate the consistency of the measurements.

To ensure the consistency of the measurements, standard statistical tests were used, considering carbon and redox balances as constraints (Wang & Stephanopoulos, 1983; Stephanopoulos et al., 1998). The measurement of the 7 key specific rates in the black−box process model was performed, considering the lipase production negligible in carbon and redox balances. The specific rates included in the balances were: biomass specific growth rate (μ), glucose uptake (q_S), oxygen uptake (q_{O2}), carbon dioxide production (q_{CO2}), and ethanol, arabitol, and succinate production (q_{EtOH}, q_{Ara}, q_{Suc}). The system was therefore overdetermined, with the degree of redundancy equal to the number of constraints, allowing for the detection of gross errors as well as unidentified metabolites and improving the accuracy of measured conversion rates using data reconciliation methods (van der Heijden et al., 1994b).

The statistical test output for the presence of gross errors or neglected components was the *h* value, which was determined by the sum of the weighted squares of the residuals ε, as shown in **Equation 38**,

$$h = \varepsilon^T \cdot P^{-1} \cdot \varepsilon \qquad \textbf{Eq. 38}$$

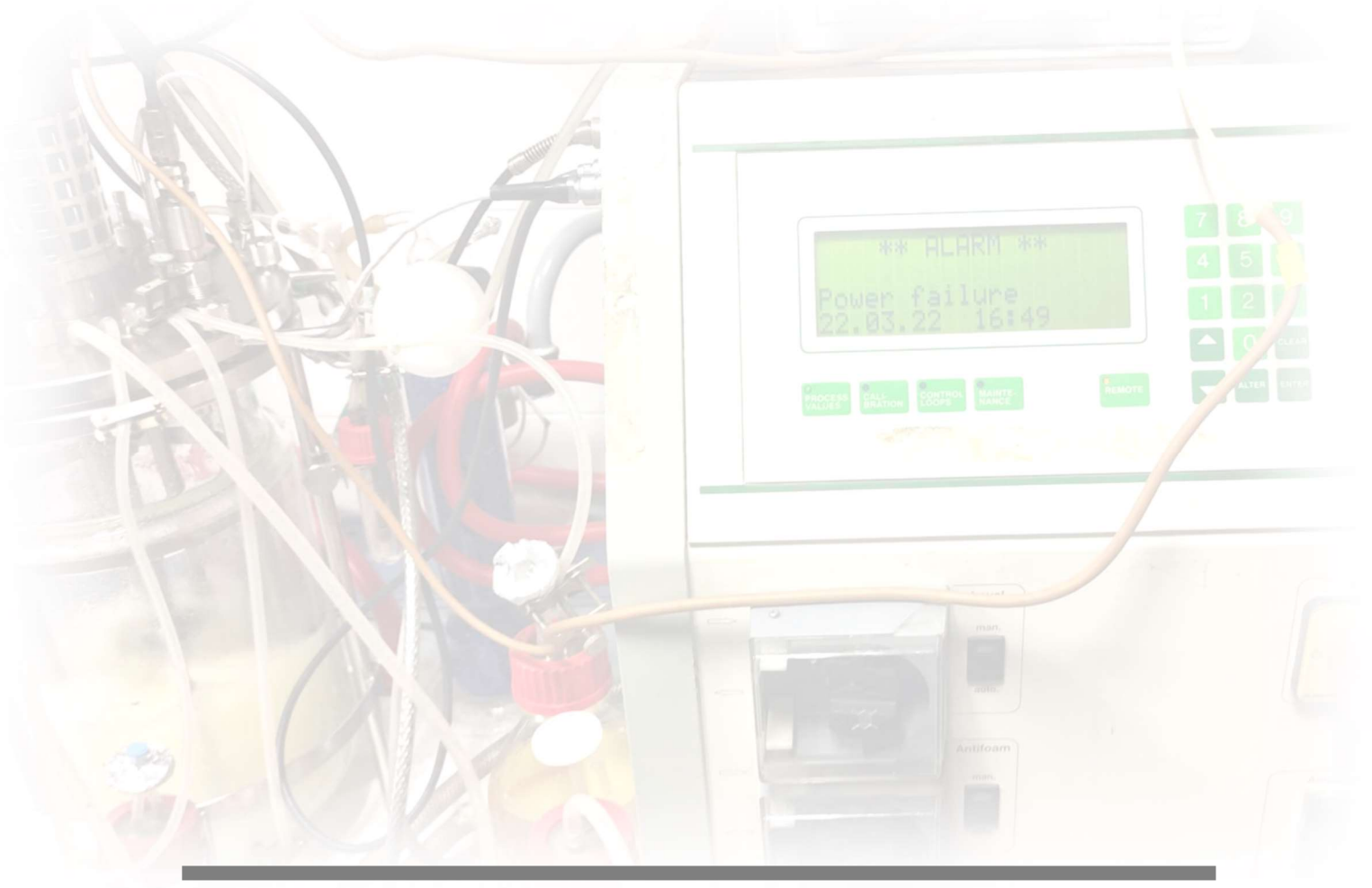
where ε corresponds to the vector of error attributed to each specific rate, ε^T is the transposed vector, and *P* is the covariance matrix. If the *h* value exceeded the threshold value, it was concluded that significant errors were present in the measurements or that any compound had not been considered in the black box process model. The threshold value, in turn, depends on the confidence level,

defined as 95% in this work, and the degree of redundancy according to the χ2 distribution. The variances of all specific rate measurements were estimated by the variability between replicates and/or measurement errors.

In this study, the χ2-test showed that the measurements mostly satisfied the stoichiometric model, with both C-balance and e-balance. It is worth mentioning that in the case of fed-batch cultivations, this accomplishment was lower than in chemostat cultures, especially in those fed-batches conducted under hypoxic conditions. Therefore, a reconciliation data procedure was applied to all data sets for all specific rates, considering the *h* values of each specific rate and the mandatory accomplishment of both carbon and redox balances (Wang & Stephanopoulos, 1983; van der Heijden et al., 1994b; Verheijen, 2009).

5. RESULTS I

Establishing a hypoxic strategy for Crl1 production



An adapted version of Sections 5.1, 5.2, and 5.3 has been published as a research article in *Frontiers in Bioengineering and Biotechnology*:

Gasset, A., Garcia-Ortega, X., Garrigós-Martínez, J., Valero, F., & Montesinos-Seguí, J. L. (2022). Innovative Bioprocess Strategies Combining Physiological Control and Strain Engineering of *Pichia pastoris* to Improve Recombinant Protein Production. *Frontiers in Bioengineering and Biotechnology, 10*(January), 1–12 https://doi.org/10.3389/fbioe.2022.818434

An adapted version of Section 5.4 has been submitted to be published as a research article in *Microbial Biotechnology:*

Sales, A., Gasset, A., Requena-Moreno, G., Valero, F., Montesinos-Seguí, J. L., Garcia-Ortega, X. (2023). Synergic kinetic and physiological stress control to improve bioprocess efficiency of *Komagataella phaffii* recombinant protein production bioprocesses. *Microbial Biotechnology.*

5.1 Characterization of hypoxia and gene dosage effects on recombinant Crl1 production and cell physiology using chemostat cultures

In the first approach, to evaluate the combined effect of gene dosage and oxygen limitation on *P. pastoris* physiology and Crl1 production, a set of chemostat cultures was performed using SCC and MCC, applying a wide range of different oxygen supply conditions.

Chemostat cultures are often used for strain characterization since they allow testing a wide range of environmental conditions such as pH or temperature, analyzing substrate utilization or studying the production kinetics, for example, with relatively little effort and eliminating dead times in cleaning and sterilization procedures as well as in mounting and dismounting steps compared with batch and fed-batch cultures (Hoskisson & Hobbs, 2005; Nieto-Taype et al., 2020b).

As detailed in the methodology section, to achieve these different oxygenation conditions air was mixed with nitrogen in different proportions, so the proportion of oxygen in the inlet gas was the only manipulated variable among all the chemostat conditions tested. The oxygen composition in the inlet gas ranged from 21%, which corresponded to normoxic conditions, to 8%, corresponding to severe hypoxia, passing through intermediate conditions (14%, 12%, 11%, 10%, and 9%). In previous studies with these two clones under normoxic conditions, the Crl1 production rate (q_P) showed an increasing trend with respect to μ with a saturation tendency at μ > 0.10 h^{-1}, especially for the MCC (Nieto-Taype et al., 2020a). Therefore, a dilution rate of $D = \mu = 0.10$ h^{-1} was selected for testing all the conditions. Since an important reduction of μ_{max} was expected when applying oxygen limiting conditions, and also considering that $\mu_{max} \approx 0.19$ h^{-1} when growing on glucose, half of μ_{max} was applied to be conservative and avoid bioreactor washout (Mattanovich et al., 2009, 2017). Other operational parameters are described in the methodology section.

The values of biomass (measured as Dry Cell Weight, DCW), ethanol concentration, Crl1 titer, and dissolved oxygen (DO) at different steady states

with distinct oxygen proportions are plotted in **Figure 9A** for SCC and **Figure 9B** for MCC.

This primary data indicates that within the range from 21% to 14% of oxygen in the inlet gas, DO was still above 0%, although it reaches relatively low values with the 14% condition. Thus, these two conditions can be considered normoxia or fully aerobic conditions since no ethanol was detected either. Biomass concentrations of SCC and MCC were constant within this range of oxygen supply. The same behavior was observed with Crl1 titers, despite being 2- to 3-fold higher in MCC than in SCC.

Below 12% of oxygen in the inlet gas, DO = 0%, and fermentation by-products were detected in the broth as explained below, so the conditions became hypoxic. Clear increases in Crl1 titers were observed for both clones, with similar profiles but higher values for MCC as observed in normoxia. Regarding biomass and ethanol production, similar profiles were observed for both clones: as the oxygen supply is reduced, biomass generation decreases linearly and ethanol production increases proportionally, since metabolism is redirected (Carnicer et al., 2009; Baumann et al., 2010; Adelantado et al., 2017; Garcia-Ortega et al., 2017).

Although a preliminary comparison between SCC, MCC, hypoxic, and normoxic conditions can be done with this primary data, a deeper analysis of specific rates must be done to establish the optimal oxygen limiting condition among all conditions tested. Otherwise, false conclusions could be reached, thereby correlating the maximum in Crl1 titer with the maximum in productivity or q_P, because biomass concentration also depends on oxygen-limitation as clearly shown in **Figure 9**. With this aim, in the next section, a proper comparison of q_S, $Y_{X/S}$, q_{EtOH}, and q_P between clones and also between oxygenation levels is carried out.

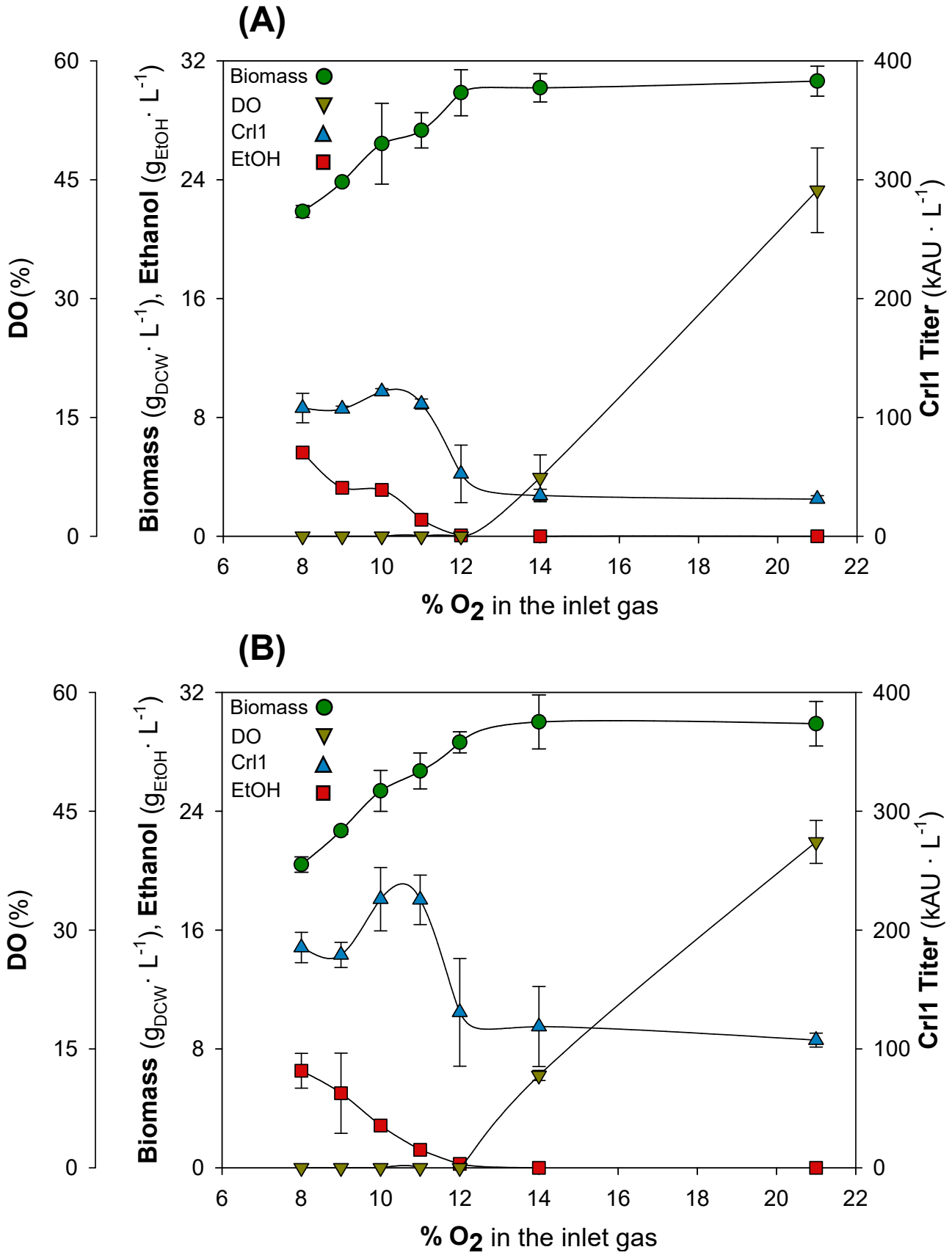


Figure 9. Primary data of the main key process parameters for the chemostat cultivations grown at different oxygen-limiting conditions. Biomass concentration (●, g · L^{-1}); DO, dissolved oxygen (▼, %); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}). **(A)** SCC, **(B)** MCC. Error bars represent the SD of biological replicates.

5.1.1 Increased gene dosage and oxygen limitation application for the enhancement of Crl1 production

The variation of different key parameters (q_S, $Y_{X/S}$, q_{EtOH}, q_{O2}, q_{CO2}, RQ, q_P, and $Y_{P/X}$) at different oxygenation levels can be observed in **Figure 10** and **Figure 11** for SCC and MCC, respectively. The numeric values of these parameters are also shown in **Tables 5** and **6,** together with other relevant parameters such as biomass concentration, Crl1 titer, and ethanol concentration.

Generally, the behavior of both clones is shown to be rather similar, except for the production parameters q_P and $Y_{P/X}$, which are between 2-fold and 4-fold higher for MCC, depending on the oxygenation condition. Thus, it can be stated that having a higher gene dosage does not significantly affect the yeast's physiology and its capacity to grow at μ = 0.10 h^{-1}. This fact has already been observed in some previous studies with the same promoter. Different clones expressing a recombinant protein under the regulation of the P_{GAP} with different gene dosage showed rather similar physiologic parameter values (Marx et al., 2009; Aw & Polizzi, 2013; Nieto-Taype et al., 2020a), whereas in other studies, clones with different gene dosage also under the regulation of the P_{GAP} did show different behavior regarding substrate consumption and growth capacities (Hohenblum et al., 2004; Aw & Polizzi, 2013; Zheng et al., 2014). Thus, it seems that it may depend on the target protein. Alternatively, in other studies, clones with different copy number of a recombinant expression cassette with *AOX1* promoter (P_{AOX1}), another widely used *Pichia*'s promoter, did show different physiologic behavior concerning substrate consumption (Zhu et al., 2009a, 2009b; Cámara et al., 2017; Garrigós-Martínez et al., 2019). The authors first hypothesized and later provided evidence that the cause might be the limited amount of *AOX1*-related transcription factors, as discussed thoroughly in **Section 5.3**. In addition, the results show that the increases in q_P and $Y_{P/X}$ between both clones are not proportional to the increase in gene dosage, since MCC has between 2- and 3-fold higher q_P values and between 2- and 4-fold higher $Y_{P/X}$ values than SCC while the gene copy number is 5 versus 1. Similar effects were also reported previously (Macauley-Patrick et al., 2005; Cámara et al., 2016), indicating that it may be protein-dependent.

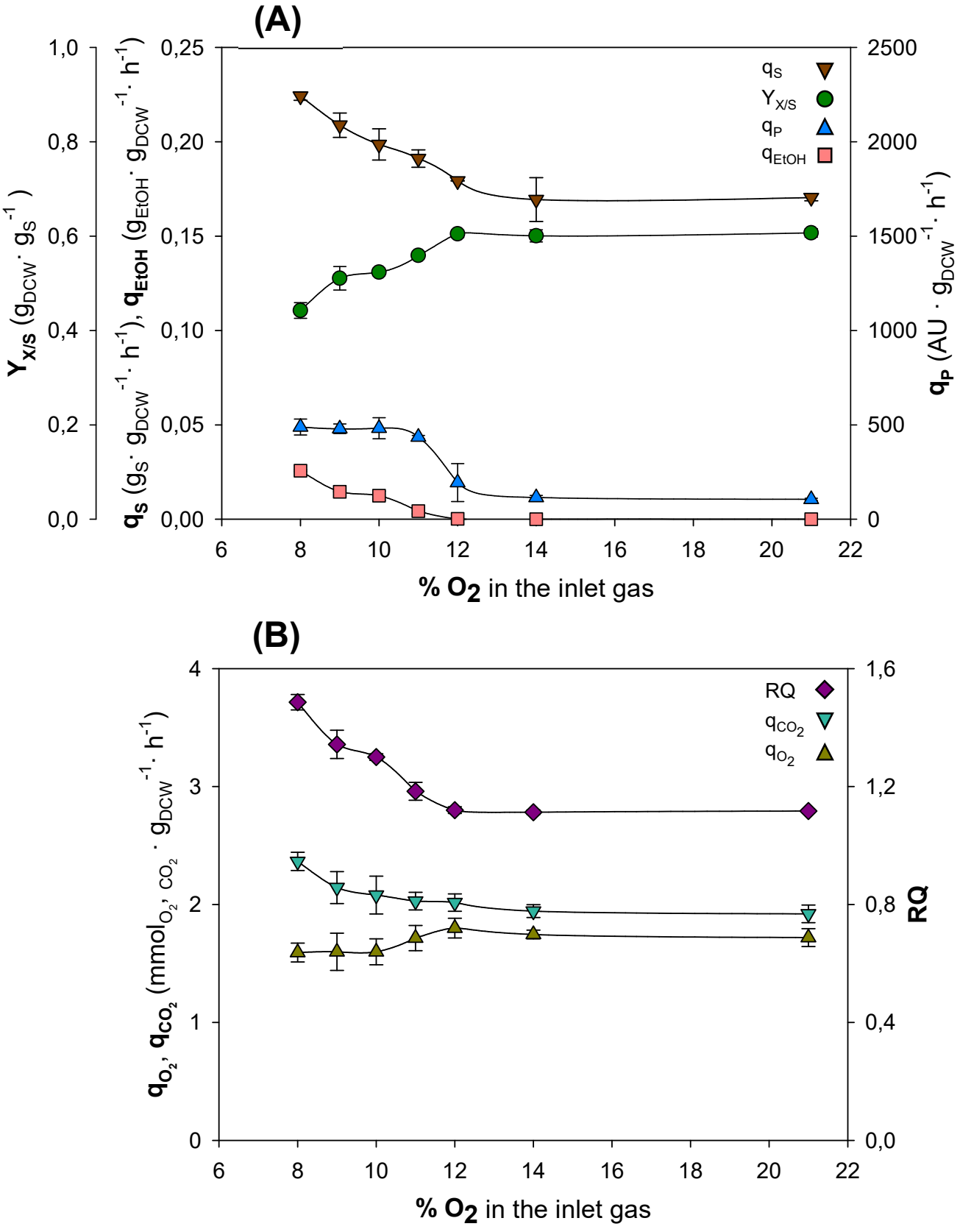


Figure 10. Main physiological and production parameters for chemostat cultures at different molar fractions of oxygen in the inlet gas with SCC. **(A)** Specific substrate consumption rate, q_S (▼, $g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); overall biomass-to-substrate yield, $Y_{X/S}$ (●, $g_{DCW} \cdot g_S^{-1}$); specific ethanol production rate, q_{EtOH} (■, $g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); and specific Crl1 production rate, q_P (▲, $AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); **(B)** Respiratory quotient, *RQ* (◆); specific oxygen consumption rate, q_{O2} (▲, $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); and specific carbon dioxide production rate, q_{CO2} (▼, $mmol_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$). Error bars represent the SD of biological replicates.

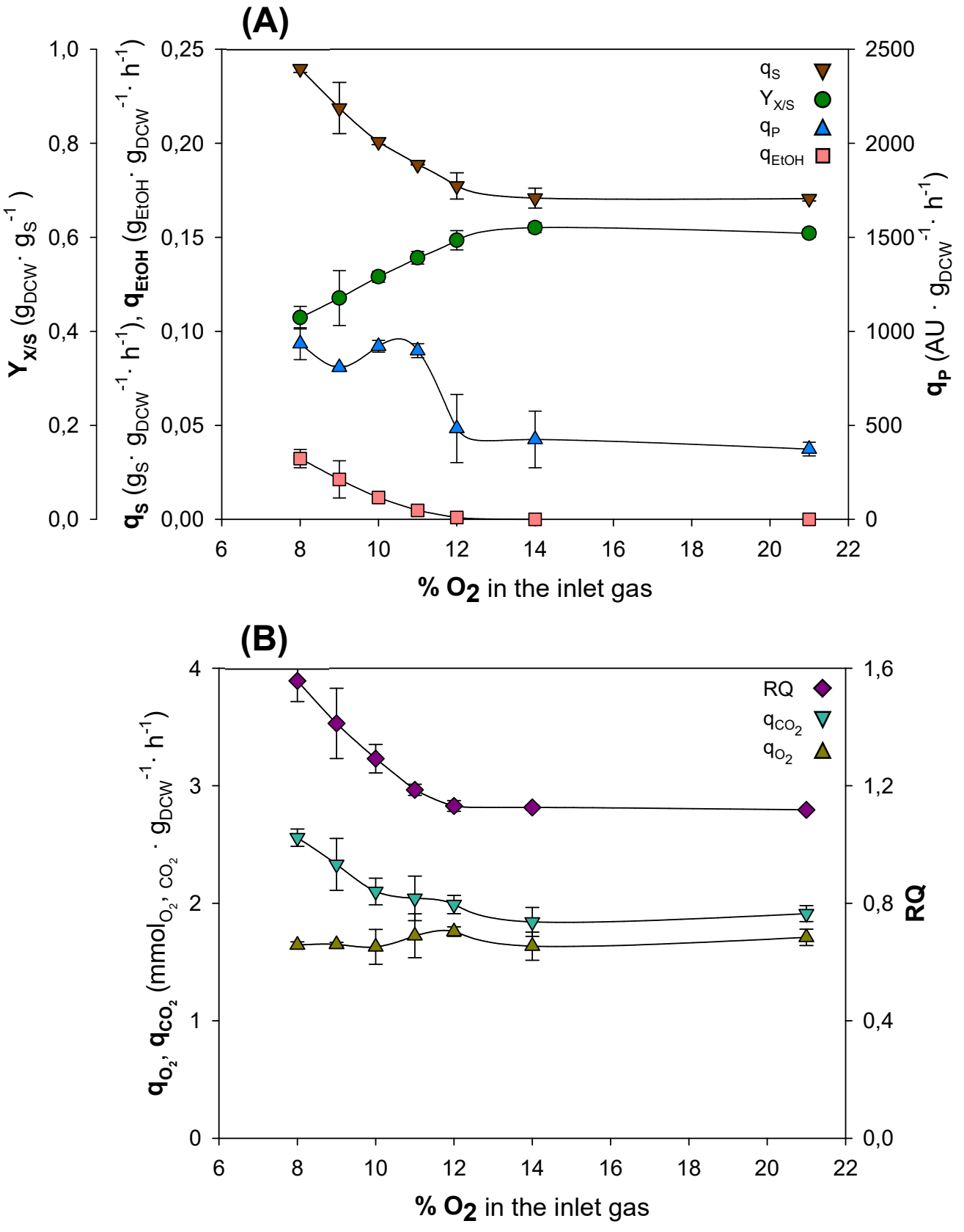


Figure 11. Main physiological and production parameters for chemostat cultures at different molar fractions of oxygen in the inlet gas with MCC. **(A)** Specific substrate consumption rate, q_S (▼, $g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); overall biomass-to-substrate yield, $Y_{X/S}$ (●, $g_{DCW} \cdot g_S^{-1}$); specific ethanol production rate, q_{EtOH} (■, $g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); and specific Crl1 production rate, q_P (▲, $AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); **(B)** Respiratory quotient, *RQ* (◆); specific oxygen consumption rate, q_{O2} (▲, $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); and specific carbon dioxide production rate, q_{CO2} (▼, $mmol_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$). Error bars represent the SD of biological replicates.

Regarding the effect of oxygen limitation, no physiological changes are observed within the range of 21% to 14% of oxygen in the inlet gas, which corresponds to a fully aerobic or oxygen non-limiting condition (normoxia). As of 12% and lower, oxygen availability becomes a crucial factor, implying that cells cannot oxidize the glucose to carbon dioxide completely, and so less energy is obtained from glucose. This causes a decrease in overall biomass-to-substrate yield ($Y_{X/S}$) and subsequently, a lower production of biomass. This fact is confirmed by the absence of glucose in the culture medium. Therefore, the specific substrate consumption rate (q_S) increases, since the same amount of glucose is introduced to the system and consumed by a lesser amount of biomass.

At this point, with around 12% of oxygen in the inlet gas, the cells cannot maintain a completely oxidative metabolism, so there is a shift from respirative to respiro-fermentative metabolism (Baumann et al., 2010; Garcia-Ortega et al., 2017), confirmed by the accumulation of fermentative by-products such as ethanol, arabitol, and succinic acid, already detected in comparable glucose-based hypoxic chemostat and fed-batch cultures (Baumann et al., 2008, 2010; Carnicer et al., 2009; Adelantado et al., 2017). It is described in the literature that increased glycolysis fluxes cause a metabolic imbalance, especially with NADH/NAD^+ cofactor, so cell metabolism re-adapts by generating these by-products (Heyland et al., 2011; Ata et al., 2018; Tomàs-Gamisans et al., 2020). The specific ethanol production rate (q_{EtOH} = 0.021 ± 0.006 $g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$) is higher but comparable to that observed in similar hypoxic chemostat cultures (q_{EtOH} = 0.014 ± 0.002 $g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$ (Adelantado et al., 2017); q_{EtOH} = 0.017 ± 0.001 $g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$ (Tomàs-Gamisans et al., 2020) with the same $D = \mu$ = 0.10 h^{-1}, and the same hypoxic level (RQ = 1.4 ± 0.1) expressing a Fab antibody. It has been described that high ethanol concentrations can result in yeast growth inhibition. However, the maximum concentrations reached in these chemostat fermentations (~7.5 $g \cdot L^{-1}$) did not attain these inhibitory levels, which were hypothesized above 30 $g \cdot L^{-1}$ (Ergün et al., 2019; Wehbe et al., 2020).

As the oxygen molar fraction in the inlet gas decreases, hypoxia increases, but no glucose accumulation is observed until the oxygen molar fraction in the inlet gas is reduced below 8%, where bioreactor washout could be observed (data not

shown). This indicates that, under such severe hypoxic conditions, critical dilution (D_C) is lower than 0.10 h^{-1} (D_C < 0.10 h^{-1}) because μ_{max} decreases substantially.

Regarding Crl1 production, around 5-fold and 2.5-fold increases in q_P are observed in hypoxic conditions for SCC and MCC, respectively, compared with normoxic conditions. On the other hand, the increases in $Y_{P/X}$ in hypoxic conditions with respect to normoxic conditions are around 4-fold and 2-fold for SCC and MCC, respectively. Strikingly, it is worth mentioning that both q_P and $Y_{P/X}$ reach rather similar values under all normoxic conditions (21% and 14% of oxygen in the inlet gas) and under all hypoxic conditions (11%, 10%, 9%, and 8% of oxygen in the inlet gas), regardless of the particular oxygen proportion value in the inlet gas, appearing to be only two main different states (normoxia and hypoxia) concerning Crl1 production.

The highest variability between replicates, in terms of both q_P and $Y_{P/X}$, is observed in the transition between normoxic and hypoxic conditions (around 12% of oxygen molar fraction in the inlet gas) as previously described (Garcia-Ortega et al., 2017). This fact could indicate, accordingly, that is around this point where the transition from respirative to respiro-fermentative metabolism is being attained.

Table 5. Values of key process parameters obtained in chemostat fermentations with SCC. Biomass concentration ($g_{DCW} \cdot L^{-1}$); specific substrate consumption rate, q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); biomass-to-substrate yield, $Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$); ethanol concentration ($g_{EtOH} \cdot L^{-1}$); specific ethanol production rate, q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); specific oxygen consumption rate, q_{O2} ($mmols_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); specific carbon dioxide evolution rate, q_{CO2} ($mmols_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); respiratory quotient, *RQ*; Crl1 Titer ($kAU \cdot L^{-1}$); specific Crl1 production rate, q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); and product-to-biomass yield, $Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$). ± indicate standard deviation (SD) of the biological replicates.

O_2 molar fraction in the inlet gas	**21**	**14**	**12**	**11**	**10**	**9**	**8**
Biomass ($g_{DCW} \cdot L^{-1}$)	30.6 ± 1.0	30.2 ± 1.0	29.9 ± 1.5	27.3 ± 1.2	26.4 ± 2.7	23.8 ± 0.1	21.9 ± 0.4
q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.17 ± 0.00	0.17 ± 0.01	0.18 ± 0.01	0.19 ± 0.01	0.20 ± 0.01	0.21 ± 0.01	0.22 ± 0.01
$Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$)	0.61 ± 0.01	0.60 ± 0.01	0.61 ± 0.01	0.56 ± 0.01	0.52 ± 0.01	0.51 ± 0.03	0.44 ± 0.02
Ethanol ($g_{EtOH} \cdot L^{-1}$)	n.d.	n.d.	0.06 ± 0.09	1.12 ± 0.40	3.13 ± 0.23	3.27 ± 0.25	5.64 ± 0.07
q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	n.d.	n.d.	0.001 ± 0.001	0.004 ± 0.001	0.013 ± 0.003	0.015 ± 0.001	0.026 ± 0.001
q_{O2} ($mmols_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	1.72 ± 0.08	1.75 ± 0.04	1.80 ± 0.08	1.71 ± 0.11	1.60 ± 0.11	1.60 ± 0.16	1.59 ± 0.08
q_{CO2} ($mmols_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	1.92 ± 0.07	1.94 ± 0.05	2.02 ± 0.07	2.03 ± 0.07	2.08 ± 0.16	2.14 ± 0.14	2.37 ± 0.08
RQ	1.12 ± 0.01	1.11 ± 0.01	1.12 ± 0.01	1.18 ± 0.03	1.30 ± 0.01	1.34 ± 0.05	1.49 ± 0.03
Crl1 Titer ($kAU \cdot L^{-1}$)	31 ± 3	34 ± 5	53 ± 24	111 ± 4	122 ± 2	107 ± 2	108 ± 12
q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	106 ± 6	115 ± 11	194 ± 100	435 ± 9	482 ± 56	480 ± 25	489 ± 42
$Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$)	1.02 ± 0.06	1.14 ± 0.21	1.79 ± 0.90	4.07 ± 0.02	4.63 ± 0.38	4.50 ± 0.12	4.94 ± 0.66

Table 6. Values of key process parameters obtained in chemostat fermentations with MCC. Biomass concentration ($g_{DCW} \cdot L^{-1}$); specific substrate consumption rate, q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); biomass-to-substrate yield, $Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$); ethanol concentration ($g_{EtOH} \cdot L^{-1}$); specific ethanol production rate, q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); specific oxygen consumption rate, q_{O2} ($mmols_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); specific carbon dioxide evolution rate, q_{CO2} ($mmols_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); respiratory quotient, *RQ*; Crl1 Titer ($kAU \cdot L^{-1}$); specific Crl1 production rate, q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); and product-to-biomass yield, $Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$). ± indicate standard deviation (SD) of the biological replicates.

O_2 molar fraction in the inlet gas	**21**	**14**	**12**	**11**	**10**	**9**	**8**
Biomass ($g_{DCW} \cdot L^{-1}$)	29.9 ± 1.5	30.0 ± 1.8	28.6 ± 0.7	26.7 ± 1.2	25.4 ± 1.4	22.7 ± 0.2	20.4 ± 0.5
q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.17 ± 0.01	0.17 ± 0.01	0.18 ± 0.01	0.19 ± 0.01	0.20 ± 0.00	0.22 ± 0.01	0.24 ± 0.01
$Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$)	0.61 ± 0.01	0.62 ± 0.01	0.59 ± 0.02	0.56 ± 0.01	0.52 ± 0.01	0.47 ± 0.06	0.43 ± 0.02
Ethanol ($g_{EtOH} \cdot L^{-1}$)	n.d.	n.d.	0.26 ± 0.37	1.21 ± 0.07	2.84 ± 0.20	5.02 ± 2.70	6.53 ± 1.17
q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	n.d.	n.d.	0.001 ± 0.001	0.005 ± 0.001	0.012 ± 0.002	0.021 ± 0.009	0.032 ± 0.005
q_{O2} ($mmols_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	1.71 ± 0.07	1.63 ± 0.12	1.76 ± 0.04	1.72 ± 0.19	1.63 ± 0.15	1.65 ± 0.02	1.64 ± 0.03
q_{CO2} ($mmols_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	1.92 ± 0.07	1.84 ± 0.12	1.99 ± 0.08	2.04 ± 0.19	2.10 ± 0.11	2.33 ± 0.22	2.56 ± 0.07
RQ	1.12 ± 0.01	1.13 ± 0.01	1.13 ± 0.02	1.19 ± 0.02	1.29 ± 0.05	1.41 ± 0.12	1.55 ± 0.07
Crl1 Titer ($kAU \cdot L^{-1}$)	107 ± 6	119 ± 34	131 ± 45	225 ± 21	226 ± 27	179 ± 11	185 ± 13
q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	374 ± 36	425 ± 150	483 ± 181	897 ± 37	920 ± 31	808 ± 6	934 ± 85
$Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$)	3.60 ± 0.38	4.00 ± 1.36	4.59 ± 1.70	8.43 ± 0.40	8.89 ± 0.57	7.90 ± 0.55	9.08 ± 0.40

5.1.2 *RQ* dependence on hypoxia degree

In **Figures 10B** and **11B**, the values of respirometric parameters (specific oxygen consumption rate, q_{O2}; specific carbon dioxide evolution rate, q_{CO2}; and *RQ*) for SCC and MCC, respectively, can be observed. The value of these parameters under normoxic conditions (q_{O2} = 1.72 ± 0.08 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; *RQ* = 1.12 ± 0.01) is within the range of those reported in previous glucose-based chemostat cultures (q_{O2} = 1.80 ± 0.01 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; *RQ* = 1.15 ± 0.05) (Nieto-Taype et al., 2020a); (q_{O2} = 1.65 ± 0.25 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; *RQ* = 1.05 ± 0.05) (Tomàs-Gamisans et al., 2020); or glucose/methanol co-feeding fed-batch fermentations (q_{O2} = 1.75 ± 0.25 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; *RQ* = 1.03 ± 0.07) (Zavec et al., 2020). As observed in the previous section, both clones performed very similarly, regardless of the gene dosage.

Regarding the oxygenation level, from 21% to 14% of oxygen in the inlet gas, the respirometric behavior of both clones is constant, corresponding to the values of normoxic conditions mentioned above. When the oxygen molar fraction is reduced below 12%, the biomass has less oxygen available, so q_{O2} decreases. This fact, combined with a slight increase in q_{CO2}, has a direct impact on *RQ* values, which present a linear increase from 1.1 under normoxic conditions to 1.6 with the most severe hypoxic conditions, indicating a direct relationship between *RQ* and the degree of oxygen limitation (from 12% to 8% of oxygen in the inlet gas). The values of q_{O2}, q_{CO2} and *RQ* under the most severe hypoxic conditions (q_{O2} = 1.59 ± 0.08 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; q_{CO2} = 2.37 ± 0.08 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; *RQ* = 1.49 ± 0.03) is comparable to those reported in previous glucose-based hypoxic chemostat cultures (q_{O2} = 1.4 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; q_{CO2} = 2.0 $mmol_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$; *RQ* = 1.48) (Tomàs-Gamisans et al., 2020).

5.1.3 Respiratory Quotient (*RQ*) as a transferable operational parameter to apply oxygen limitation

As stated in a preliminary study (Garcia-Ortega et al., 2017), the oxygen proportion in the inlet gas is an easy-to-manipulate operational parameter that becomes useful for the application of different levels of oxygen limitation, but it cannot be directly correlated with the degree of hypoxia and the effect that it generates over *P. pastoris* metabolism, since the effect of different molar fractions of oxygen in the inlet gas is specific for each culture system. The reason for this is that the oxygen transfer rate (*OTR*) to the culture broth does not rely only on the oxygen molar fraction in the inlet gas, but also on other operational conditions such as the gas flow rate, the agitation rate, the number, size, type, and position of stirrer impellers and even the pressure and temperature of the system, among others, which affect K_La and oxygen solubility (Garcia-Ochoa & Gomez, 2009; Garcia-Ochoa et al., 2010; Doran 2013).

On the other hand, *RQ* is a parameter that is also related to oxygen consumption and carbon dioxide production, giving very valuable information about the degree of oxygen limitation affecting biomass and so, the type of metabolism that it is carrying out (respirative or respiro-fermentative). At the same time, *RQ* is a parameter that does not depend on the configuration of the culture system but solely and exclusively on the physiological state of the cells. Thus, *RQ* can be used as a transferable operating criterion to apply the same hypoxic level to a different system.

Considering the results shown in **Figures 10A** and **11A**, q_{EtOH} could be also considered a reporting parameter of the hypoxic conditions, since it also presents a correlation with respect to the degree of oxygen limitation. However, its on-line determination would represent additional complexity since the on-line determination of biomass and ethanol concentrations would be necessary. In contrast, *RQ* determination is expected to be less complex, being a parameter commonly calculated on-line through O_2 and CO_2 analysis from the outgas stream.

In **Figure 12**, the relationship between q_P, q_{EtOH}, and *RQ* for SCC and MCC is presented. In both cases, two sets of grouped dots with low q_P (triangles) and zero q_{EtOH} (squares) corresponding to an *RQ* ≈ 1.1 are observed, which corresponds to normoxic conditions. When *RQ* reaches values around 1.2, a sudden rise in q_P can be observed, which becomes saturated when *RQ* > 1.3. From this point and above, q_P remains rather constant regardless of the *RQ* value. As stated before, in terms of production there seem to be only two main conditions: normoxia (*RQ* < 1.2) and hypoxia (*RQ* > 1.2), and a transition zone between them (*RQ* ≈ 1.2). On the contrary, q_{EtOH} presents a clear correlation with *RQ*, being zero in normoxic conditions and increasing linearly when *RQ* increases. Moreover, the same values of q_{EtOH} are observed for both clones.

Considering both applied strategies, namely gene dosage increase and oxygen limitation, an improvement of about one order of magnitude for mean q_P values is achieved comparing SCC in normoxia (q_P = 106 ± 6 CrI1 UA · g_{DCW}^{-1} · h^{-1}) with MCC in hypoxia (q_P = 934 ± 85 CrI1 UA · g_{DCW}^{-1} · h^{-1}). A similar increase of 8.9-fold in $Y_{P/X}$ is also obtained.

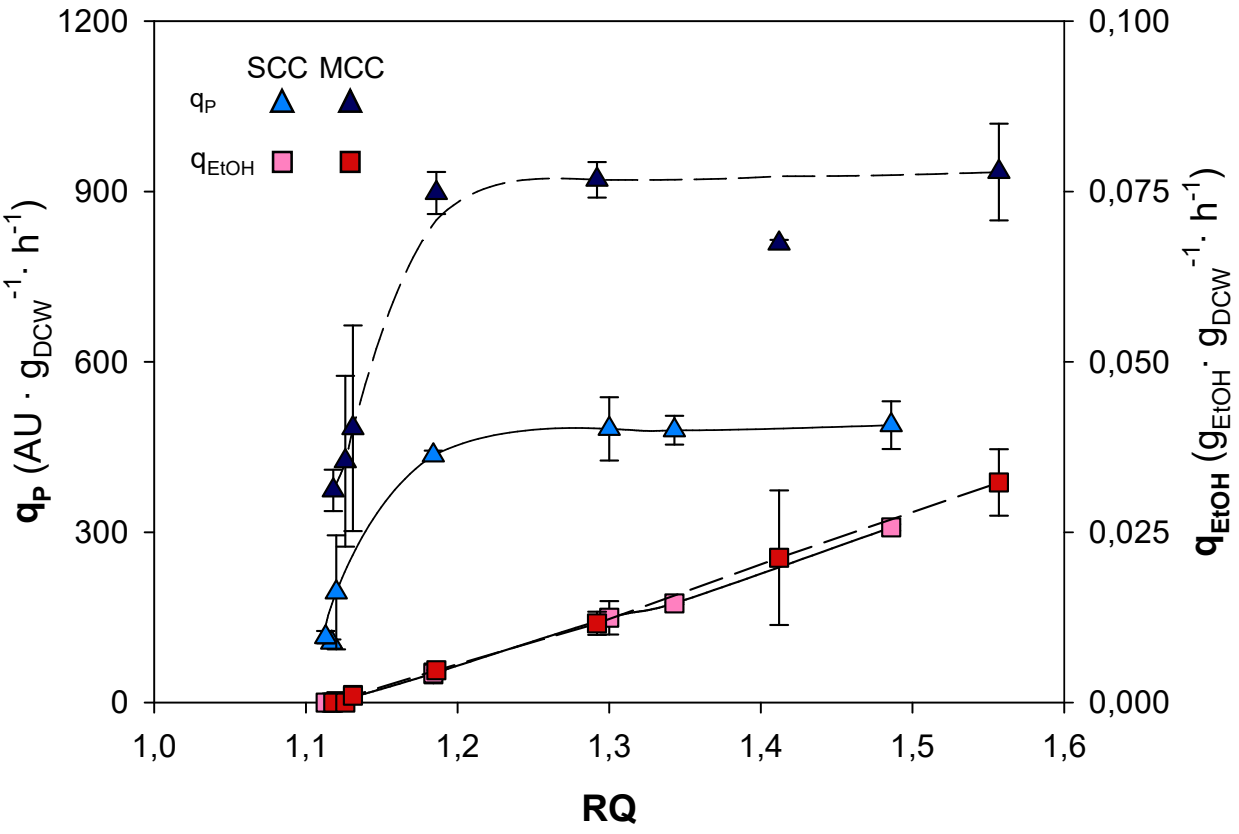


Figure 12. Relationships of specific Crl1 production rate, q_P (▲, AU · g_{DCW}^{-1} · h^{-1}) and specific ethanol production rate, q_{EtOH} (■, g_{EtOH} · g_{DCW}^{-1} · h^{-1}) with respect to respiratory quotient (*RQ*) with SCC (light symbols) and MCC (dark symbols) observed in chemostat cultures. Error bars represent the SD of biological replicates.

5.2 Fed-batch culture strategy to improve bioprocess efficiency

5.2.1 Agitation rate as manipulated variable to achieve target *RQ*

Once the combined effect of gene dosage and oxygen limitation had been characterized, the next step was to transfer the bioprocess to a fed-batch mode, which is more representative of an industrial production process (Kumar et al., 2020; Nieto-Taype et al., 2020b; De Brabander et al., 2023).

A set of eight fed-batch cultures was performed with a feeding strategy consisting of a pre-programmed exponential profile with μ = 0.10 h^{-1}, mimicking the carbon-limited growth applied in chemostat. With this strategy, a pseudo-steady state is reached during the feeding phase, so a calculation of a mean value for each physiological key parameter and a comparison with chemostat results can be conducted (Cos et al., 2005b; Garcia-Ortega et al., 2013). Four of these fed-batches were done with SCC and four with MCC, applying two different conditions: normoxia (dissolved oxygen higher than 30%, *RQ* ≈ 1.1) and hypoxia (*RQ* set-point = 1.4), and duplicates for each condition were carried out to ensure the reliability and robustness of the results.

Applying a constant level of oxygen limitation in a dynamic culture poses a challenge since the control action over *RQ* will not be constant. If *RQ* is to be maintained at a certain value and biomass grows exponentially, oxygen supply to the culture should also be increased exponentially.

As observed in chemostat cultures, the variation of the oxygen fraction in the inlet gas is an easy-to-modify parameter by means of mixing nitrogen with air. However, at an industrial scale the use of pure gases, either nitrogen or oxygen, involves associated costs that should be avoided if possible. Instead, agitation is also an easy-to-modify parameter and is directly related to oxygen supply, since it notably affects the volumetric oxygen transfer coefficient (k_La).

Thus, after having checked that the modification of both the oxygen molar fraction in the inlet gas and the stirring rate can be successfully used to maintain *RQ* within a particular range in a dynamic culture such as fed-batch (data included in Annex I), it was decided to implement an *RQ* control based on agitation rate

variation. This entails a new strategy, different from that implemented in chemostat cultures, which allows a possible saving in energy and pure gases, and, *a priori*, it seems easier to implement in an industrial process, whenever proper mixing can be guaranteed and no heterogeneities in the bioreactor are produced. Actually, agitation rate is a parameter commonly used to modify the oxygen transfer rate to the culture broth, although not to maintain hypoxic conditions but to keep the dissolved oxygen (DO) above a certain value (Garcia-Ochoa & Gomez, 2009; Garcia-Ochoa et al., 2010; Doran 2013).

The entire strategy applied in this work aims to be more robust and reproducible than that reported in previous studies, which was based on a start-and-stop substrate feeding that only targeted the avoidance of high ethanol accumulation, considered to be inhibitory, without taking into account other parameters than can affect q_P and productivity, such as μ and degree of oxygen limitation (Baumann et al., 2008).

As seen from the chemostat results in **Figure 12**, maximum q_P values were reached at RQ > 1.2. Even so, too high RQ values might cause an important ethanol production and a significant $Y_{X/S}$ decrease. Consequently, in fed-batch cultures, ethanol is accumulated in the culture broth and can reach inhibitory concentrations (Ergün et al., 2019; Potvin et al., 2016; Wehbe et al., 2020), which does not occur in chemostat since the medium is constantly renewed. On the other hand, a significant $Y_{X/S}$ decrease entails a remarkably lower biomass production, since in a carbon-limited fed-batch, the open-loop feeding profile is designed considering a constant and known $Y_{X/S}$. If $Y_{X/S}$ were to decrease during the feeding phase due to a highly fermentative metabolism, so would μ. This, in turn, would cause a decrease in the specific production rate and productivity of the bioprocess and a potential substrate accumulation if μ_{max} decreases below 0.10 h^{-1}, which was observed in those chemostat conditions with less than 8% of oxygen in the inlet gas. This effect differs from that observed in chemostat, where μ is governed by the dilution rate ($\mu = D$), whenever $\mu = D < D_C \sim \mu_{max}$.

A preliminary exploratory fed-batch, with RQ controlled above 1.8, was performed to check this possible ethanol inhibition effect. Ethanol accumulation achieved inhibitory concentrations (> 30 g/L) and caused a progressive halt in biomass

growth, as well as a $Y_{X/S}$ reduction of more than 50% (included in the Annex). This ethanol inhibitory effect was not observed within the *RQ* range of 1.2 to 1.6. Consequently, an *RQ* set-point of 1.4 was selected. Thus, possible minor deviations of *RQ* due to the non-automatization of the *RQ* control are expected to have a smaller effect on q_P, since specific productivity is quite constant within this *RQ* range, as observed in **Figure 12**.

In this first approach of *RQ* control operation, the actuation over agitation rate was performed manually approximately every hour, following heuristic rules: to increase *RQ*, agitation was reduced in order to decrease oxygen transfer rate to the culture broth and therefore decrease the oxygen uptake rate, since in an oxygen-limited fed-batch one can consider $OUR \approx OTR$. To reduce *RQ*, the opposite action was conducted.

On the other hand, as mentioned in **Section 4.2.3.2**, in the normoxic fermentations DO was kept above 30% of air saturation by modifying the agitation rate and also mixing air with pure oxygen in the inlet gas steam, especially at the last stages of the fermentation when the oxygen demand was higher.

5.2.2 Implementation of an optimal hypoxic fed-batch strategy for Crl1 production based on previous chemostat characterization

The time evolutions of biomass (dry cell weight), ethanol concentration, and Crl1 titer throughout the feeding phase, both for SCC and MCC and under normoxic and hypoxic conditions are shown in **Figure 13**. Accordingly, in **Table 7**, the final values of biomass and ethanol concentrations, Crl1 titers, and volumes, together with the mean values of their related key process parameters (μ, q_S, $Y_{X/S}$, q_{EtOH}, q_{O2}, q_{CO2}, RQ, q_P, $Y_{P/X}$) are presented.

The feeding phase lasted for 20 hours for all fermentations, reaching a final biomass concentration of around 110 $g \cdot L^{-1}$ in the normoxic fermentations and around 90 $g \cdot L^{-1}$ in the hypoxic ones. As observed in chemostat cultures, both clones (SCC and MCC) exhibit similar behavior, except for Crl1 production, which increases when applying hypoxia. As explained in more detail in the methodology section, the feeding flowrate profile was slightly higher in the hypoxic fermentations to counteract the expected $Y_{X/S}$ reduction. However, results could seem to indicate that this difference was not enough, since considerably lower biomass concentrations were obtained in hypoxia. Nevertheless, the difference in the total volume due to this differential feeding profile must be considered, since lower biomass concentrations combined with higher volumes result in quite comparable total biomass profiles. As presented in **Figure 14**, the total biomass productions are very similar for both clones and conditions, despite being slightly smaller in hypoxia. This leads to the conclusion that, in order to offset the differences in $Y_{X/S}$ between normoxic and hypoxic conditions, another variable should be modified instead of the volume of feeding addition, since although the total biomass production is comparable for both conditions, different concentration profiles can lead to misinterpretations. For instance, varying the substrate concentration of the feeding tank in the same proportion that $Y_{X/S}$ decreases between normoxia and hypoxia should be enough to achieve *a priori* identical biomass profiles with identical fermentation volumes. This will be discussed in **Section 5.4**.

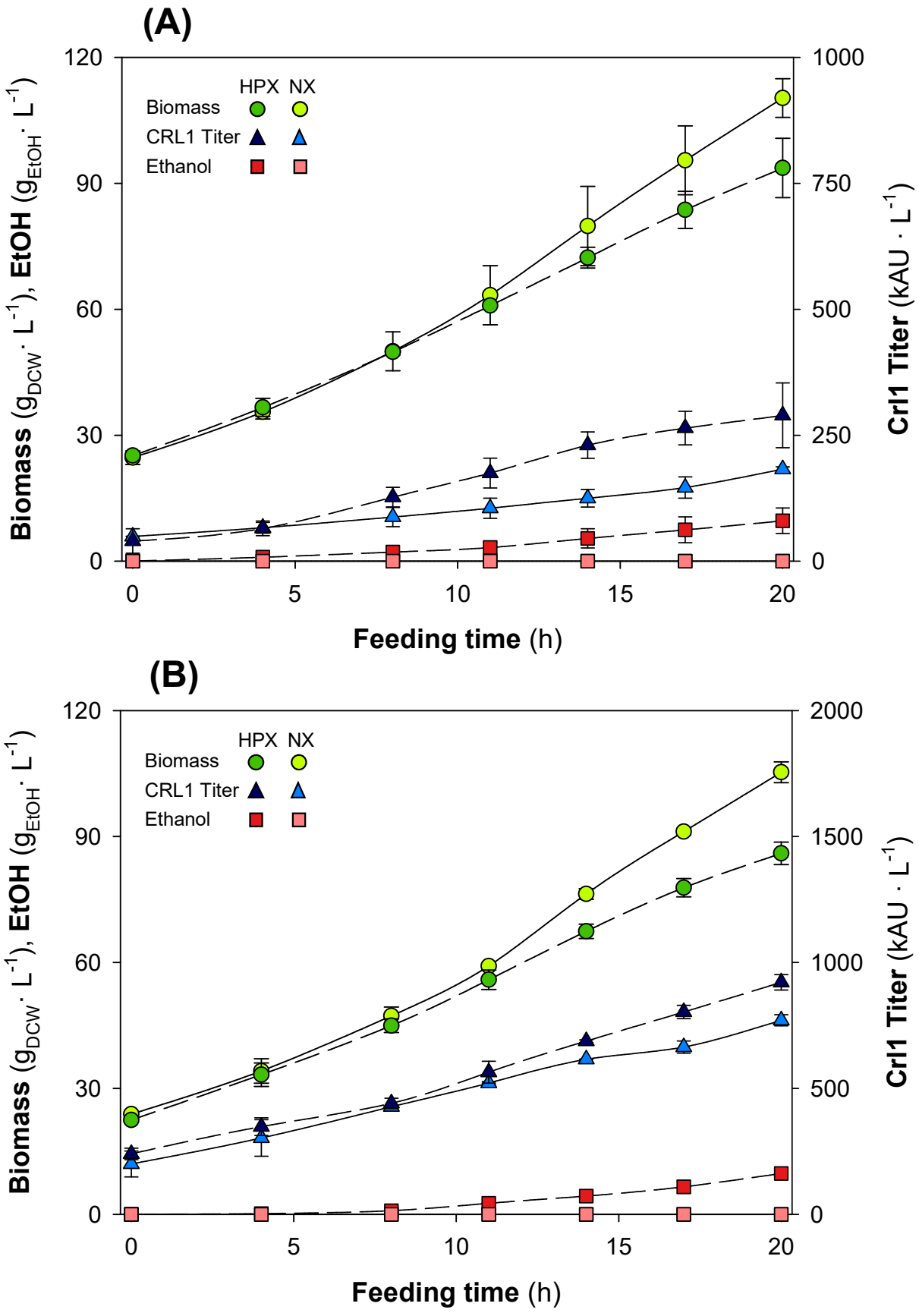


Figure 13. Primary data of the main key process parameters for the fed-batch cultivations grown at hypoxic (dark symbols) and normoxic (dark symbols) conditions with SCC and MCC. Biomass concentration (●, g · L^{-1}); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}). **(A)** SCC (SCC-NX1, SCC-NX2, SCC-HPX1, and SCC-HPX2) **(B)** MCC (MCC-NX1, MCC-NX2, MCC-HPX1, and MCC-HPX-2). Error bars represent the SD of biological replicates.

The values shown in **Table 7** corroborate these results. In hypoxic conditions, $Y_{X/S}$ decreases by about 20% on average, and q_S increases accordingly compared with normoxic conditions.

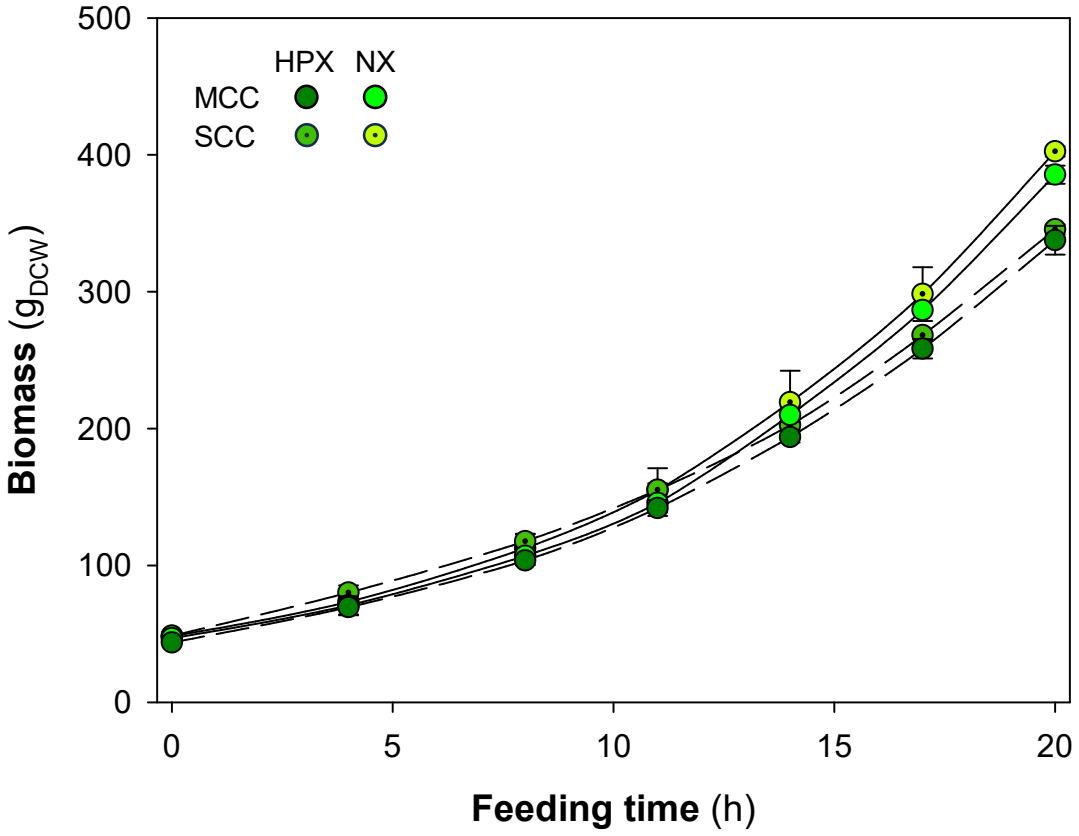


Figure 14. Total biomass generation with SCC (⊙) and MCC (●) obtained under hypoxic (dark symbols) and normoxic (light symbols) conditions (SCC-NX1, SCC-NX2, SCC-HPX1, SCC-HPX2, MCC-NX1, MCC-NX2, MCC-HPX1, and MCC-HPX-2). Error bars represent the SD of biological replicates.

HPLC analysis results showed an ethanol accumulation of up to 10 $g \cdot L^{-1}$ in hypoxic conditions, although no growth inhibition was observed. Furthermore, arabitol (up to 7 $g \cdot L^{-1}$) and succinic acid (up to 1 $g \cdot L^{-1}$) were also detected in the culture broth in hypoxic fermentations. The concentration of these by-products was higher than those observed in chemostats due to the accumulation effect, typical of this operational mode. Finally, no glucose accumulation was detected neither in normoxia nor in hypoxia.

Table 7. Value of key process parameters obtained in fed-batch fermentations with SCC and MCC. Final biomass concentration ($g_{DCW} \cdot L^{-1}$); specific growth rate, μ (h^{-1}); specific substrate consumption rate, q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); biomass-to-substrate yield, $Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$); final ethanol concentration ($g_{EtOH} \cdot L^{-1}$); specific ethanol production rate, q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); specific oxygen consumption rate, q_{O2} ($mmols_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); specific carbon dioxide evolution rate, q_{CO2} ($mmols_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$); respiratory quotient, *RQ*; Crl1 Titer ($kAU \cdot L^{-1}$); specific Crl1 production rate, q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); product-to-biomass yield, $Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$); and final volume (L). ± indicate standard deviation (SD) of the biological replicates.

O_2 supply condition	SCC		MCC	
	Normoxia	Hypoxia	Normoxia	Hypoxia
Biomass ($g_{DCW} \cdot L^{-1}$)	110 ± 5	94 ± 7	105 ± 2	86 ± 3
μ (h^{-1})	0.107 ± 0.003	0.101 ± 0.006	0.106 ± 0.006	0.105 ± 0.010
q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.18 ± 0.01	0.21 ± 0.01	0.18 ± 0.01	0.22 ± 0.01
$Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$)	0.59 ± 0.00	0.47 ± 0.04	0.58 ± 0.01	0.50 ± 0.01
Ethanol ($g_{EtOH} \cdot L^{-1}$)	n.d.	9.64 ± 3.04	n.d.	9.73 ± 1.00
q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	n.d.	0.03 ± 0.01	n.d.	0.03 ± 0.01
q_{O2} ($mmols_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	1.93 ± 0.07	1.67 ± 0.04	2.02 ± 0.04	1.78 ± 0.24
q_{CO2} ($mmols_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	2.13 ± 0.08	2.30 ± 0.10	2.23 ± 0.05	2.42 ± 0.19
RQ	1.11 ± 0.00	1.38 ± 0.09	1.10 ± 0.00	1.37 ± 0.07
Crl1 Titer ($kAU \cdot L^{-1}$)	182 ± 5	290 ± 64	770 ± 22	921 ± 30
q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	164 ± 8	329 ± 27	696 ± 7	1029 ± 45
$Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$)	1.55 ± 0.02	3.51 ± 0.12	6.61 ± 0.31	10.49 ± 0.20
Volume (L)	3.65 ± 0.12	3.72 ± 0.24	3.66 ± 0.02	3.92 ± 0.00

With regards to Crl1 production, Crl1 titers are clearly higher than in chemostat. As expected, MCC production is higher than SCC and, in parallel, in hypoxic conditions Crl1 production increases compared to normoxic state. However, due to differential volume profiles in hypoxic and normoxic fermentations, a proper comparison of Cr1 production should be made in terms of total production, q_P, and $Y_{P/X}$.

Crl1 total production is presented in **Figure 15**, where Crl1 titers (total kAU - kiloActivity Units) are plotted. Regarding oxygen limitation, SCC has 70% higher Crl1 production under hypoxic rather than normoxic conditions, whereas the increase for MCC is only 20%. On the other hand, with regard to gene dosage, MCC presents 4-fold higher Crl1 production than SCC in normoxia, whereas only 3-fold higher in hypoxia. Accordingly, considering the results from **Table 7**, q_P increases considerably with oxygen limitation. As observed in chemostat, the q_P increase when applying hypoxia is higher for SCC (2-fold) than for MCC (1.5-fold). In terms of $Y_{P/X}$, a similar effect can be observed. Whereas in normoxic conditions the $Y_{P/X}$ is higher in fed-batch than in chemostat, the increase of $Y_{P/X}$ when applying hypoxia is lower in fed-batch, so in these conditions, higher $Y_{P/X}$ values are observed in chemostat. In general, both q_P and $Y_{P/X}$ follow similar trends.

This represents an overall 6.3-fold increase of q_P and a 6.7-fold increase of $Y_{P/X}$, considering both increased gene dosage and oxygen limitation effects, lower than that achieved in chemostat cultures but still very important, considering the save in the use of pure gases and power consumption by a lower agitation whenever possible.

From the results obtained in chemostat cultures, fed-batch data seem to indicate that, in this operational mode the application of oxygen limitation is less effective than in chemostat, since higher Crl1 titers, q_P, and $Y_{P/X}$ values were expected when implementing hypoxic conditions to fed-batch cultures, especially with MCC. This could be explained by the combination of two effects: firstly, oxygen limitation causes metabolic stress, forcing cell metabolism towards the fermentative branch, generating a lower quantity of energetic molecules such as ATP (Carnicer et al., 2009; Baumann et al., 2010; Tomàs-Gamisans et al., 2020);

secondly, cell aging caused by the time-effect related to fed-batch bioprocesses, which has also been described to affect RPP (Curvers et al., 2001; Cos et al., 2006), probably amplifies the metabolic stress, which does not occur in chemostats, were cells are constantly being renewed.

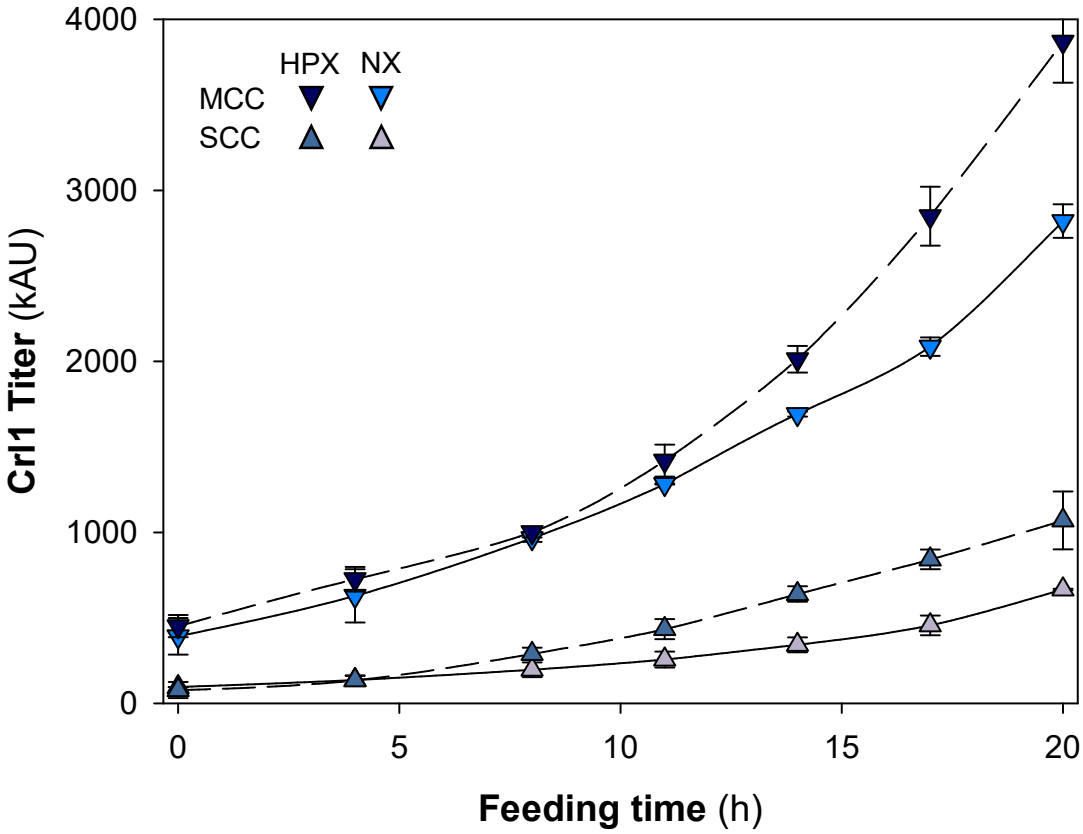


Figure 15. Total Crl1 production with SCC (▲) and MCC (▼) obtained under hypoxic (dark symbols) and normoxic (light symbols) conditions (SCC-NX1, SCC-NX2, SCC-HPX1, SCC-HPX2, MCC-NX1, MCC-NX2, MCC-HPX1, and MCC-HPX-2). Error bars represent the SD of biological replicates.

5.2.3 Strong dependence between ethanol production, *RQ,* and hypoxia degree

In **Figure 16** the off-line calculation of *RQ* and the off-line measurements of ethanol concentration for SCC (**16A**) and MCC (**16B**) are plotted. In order to further investigate the relationship between *RQ* and ethanol production in fed-batch mode as well as to assess the variability between duplicates, each replica is plotted separately: HPX1 and HPX2 corresponding to replicas 1 and 2 with hypoxic conditions; NX1 and NX2 for replicas 1 and 2 with normoxic conditions.

As explained in more detail in **Section 6.3**, it can be stated that *RQ* has been satisfactorily controlled within the desired range, being within the range from 1.0 to 1.2 in normoxic fermentations and between 1.2 and 1.8 in hypoxic ones, by means of an intermittent manual modification of agitation rate. However, high peaks in the *RQ* profile are detected, corresponding to those manual control actions increasing or reducing the agitation rate, especially when applying oxygen limitation. As detailed in **Section 6.3**, in these hypoxic fermentations the on-line calculation of *RQ* did not take into account neither the humidity concentration of the inlet and exhaust gases nor the updated calibration values of the gas analyzer. Likewise, this data was collected and incorporated *a posteriori* to the *RQ* calculated on-line to give the "off-line" *RQ* profile shown in **Figure 16**. For these reasons, mainly due to the inability to incorporate the gas analyzer's calibration data in the *RQ* calculation during the feeding phase, the final *RQ* values are not so constant and may differ considerably between replicas.

Very interestingly, a strong correlation between *RQ* values and ethanol production rate (shown by the slope of ethanol concentration profiles) can be established. The fermentations with higher final ethanol concentration (SCC-HPX2 and MCC-HPX2) are those with higher *RQ* values during the feeding phase. Furthermore, in SCC-HPX1 a sudden halt in ethanol production is observed at t = 5 h, corresponding to a decrease in the respiratory quotient from $RQ \approx 1.5$ to $RQ \approx 1.2$. Besides, in SCC-HPX2, ethanol is not detected until the third sample, when *RQ* rapidly increases from 1.2 to 1.7. These facts confirm the great influence of the oxygen limitation degree on the metabolism of *P. pastoris* growing on glucose as the sole carbon source, as observed in **Section 5.1.2**.

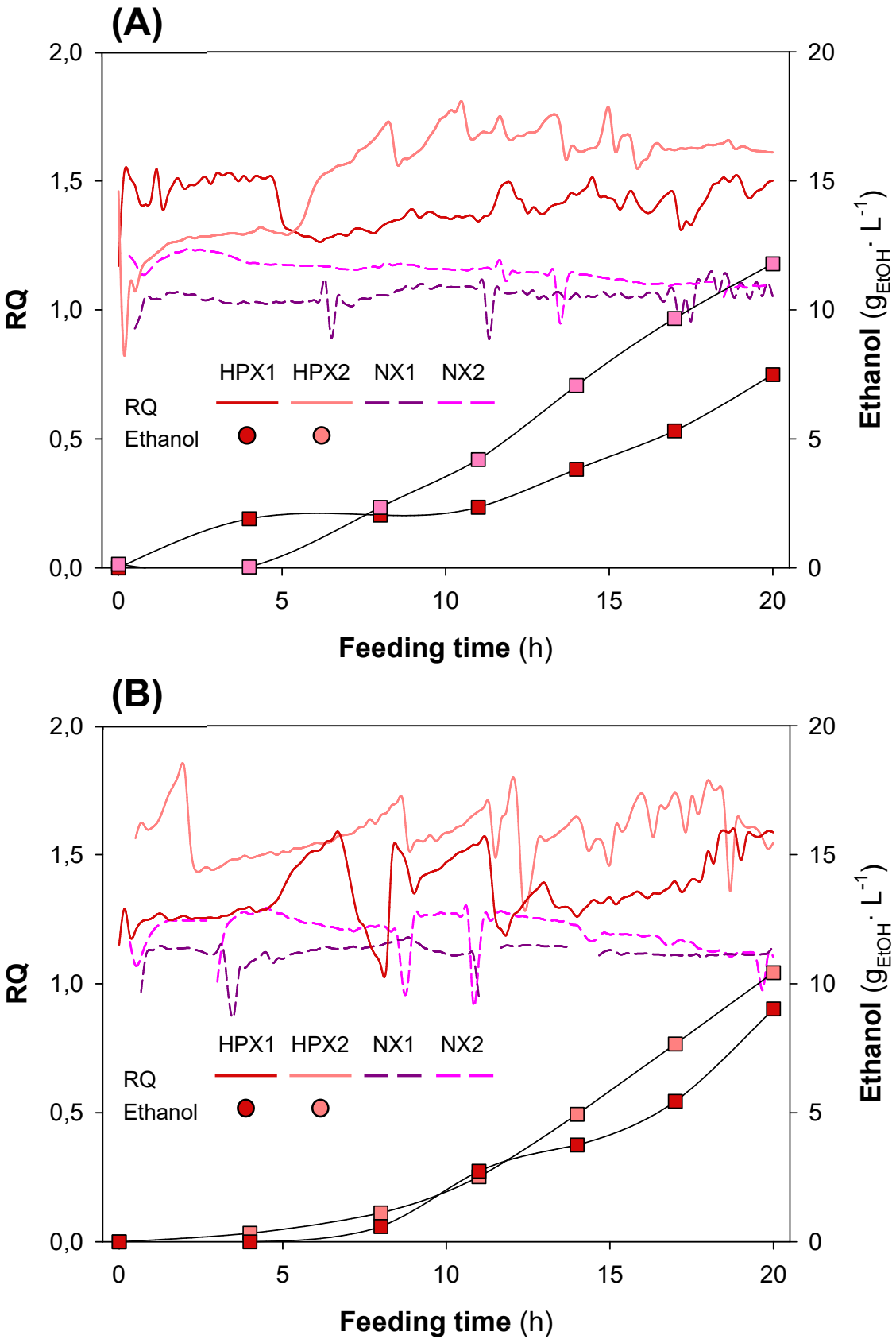


Figure 16. Off-line calculation of *RQ* and ethanol concentration for the fed-batch cultivations grown at hypoxic (red) and normoxic (purple) conditions with SCC and MCC. Off-line *RQ* (lines); Ethanol concentration (■, $g \cdot L^{-1}$). **(A)** SCC (SCC-NX1, SCC-NX2, SCC-HPX1, and SCC-HPX2), **(B)** MCC (MCC-NX1, MCC-NX2, MCC-HPX1, and MCC-HPX-2).

In concordance with these results, as seen in **Table 7**, qO_2 decreases by about 15% under hypoxic conditions, while qCO_2 has rather the same value regardless of the oxygenation conditions. Apart from the Crl1 production already discussed, no significant differences were observed between both clones.

Anticipating the discussion of **Section 6.3**, from these results it can be concluded that this manual-heuristic *RQ* control strategy, based on manual actions on agitation rate, is sufficient to maintain the *RQ* within the range of 1.2 – 1.6, but is definitely unsatisfactory if a certain *RQ* set-point has to be maintained with minor deviations. Furthermore, it ends up being a high labor- and time-consuming strategy which is not the perfect scenario for its application to a realistic industrial process. Besides, more reproducibility in terms of *RQ* control should be achieved before its final robust and reliable industrial application.

As already said, it is noteworthy the relatively high variability between replicates, not only for *RQ* values in hypoxic fermentations but also for Crl1 production, which has more variability in hypoxic duplicates than in normoxic ones, since error bars from hypoxic Crl1 profiles shown in **Figure 15** are clearly higher than those of normoxic fermentations. This indicates that the reproducibility of the hypoxic conditions, which seems *a priori* more difficult to implement than in normoxic conditions, can be further improved. This will be assessed in **Chapter 6**.

5.3 Hypoxia and gene dosage effects on transcriptional patterns

It has been thoroughly described that harboring several copies of a recombinant gene can affect the physiology and fitness of *P. pastoris* (Zhu et al., 2009; Jordà et al., 2012; Cámara et al., 2017). Besides, it is well-known that *P. pastoris* regulates its response to different carbon sources at the transcriptional level (Prielhofer et al., 2015). This can be extended to oxygen availability since it has been also demonstrated that oxygen limitation significantly alters transcriptional patterns (Baumann et al., 2010). Thus, transcriptional analysis of key genes was performed with the aim of identifying potential correlations between gene regulation and protein production under hypoxic conditions.

The genes selected were the recombinant *CRL1* gene, which allows the expression of the target product of the bioprocess, the glycolytic genes *TDH3* and *PGK1* that encode two crucial enzymes for central carbon metabolism (Glyceraldehyde-3-phosphate dehydrogenase and Phosphoglycerate Kinase 1, respectively), and the Unfolded Protein Response (UPR)-related genes *HAC1* and *KAR2*, described as a UPR-related transcription factor and a UPR-related chaperone, respectively (Guerfal et al., 2010; Raschmanová et al., 2019).

The transcription levels of the genes of interest have been analyzed in accordance with the protocol described in **Section 4.3.8**. The results of the relative transcription levels (RTL) for *CRL1, TDH3,* and *PGK1*, combined with the mean q_P values, are shown via bar graphs in **Figure 17**, and the numerical values are displayed in **Table 8**.

In order to facilitate a comprehensive comparison of both chemostat and fed-batch results, the formers are grouped into normoxic ($1.1 < RQ < 1.2$) and hypoxic results ($RQ > 1.2$) as the RTL of all the genes plotted presented no significant differences between any condition for each of the two groups (normoxic and hypoxic). As previously pinpointed, the same pattern was observed with q_P, so the values of this parameter are also grouped into normoxic and hypoxic values.

No clear trends could be observed for *HAC1* and *KAR2* RTLs with respect to neither *RQ* nor *CRL1* gene dosage, so they are excluded from the graphs and the discussion. Previous studies, in which the same two Crl1-producer clones were studied, showed that these genes remained unregulated under different specific growth rate (μ) conditions (Nieto-Taype et al., 2020a).

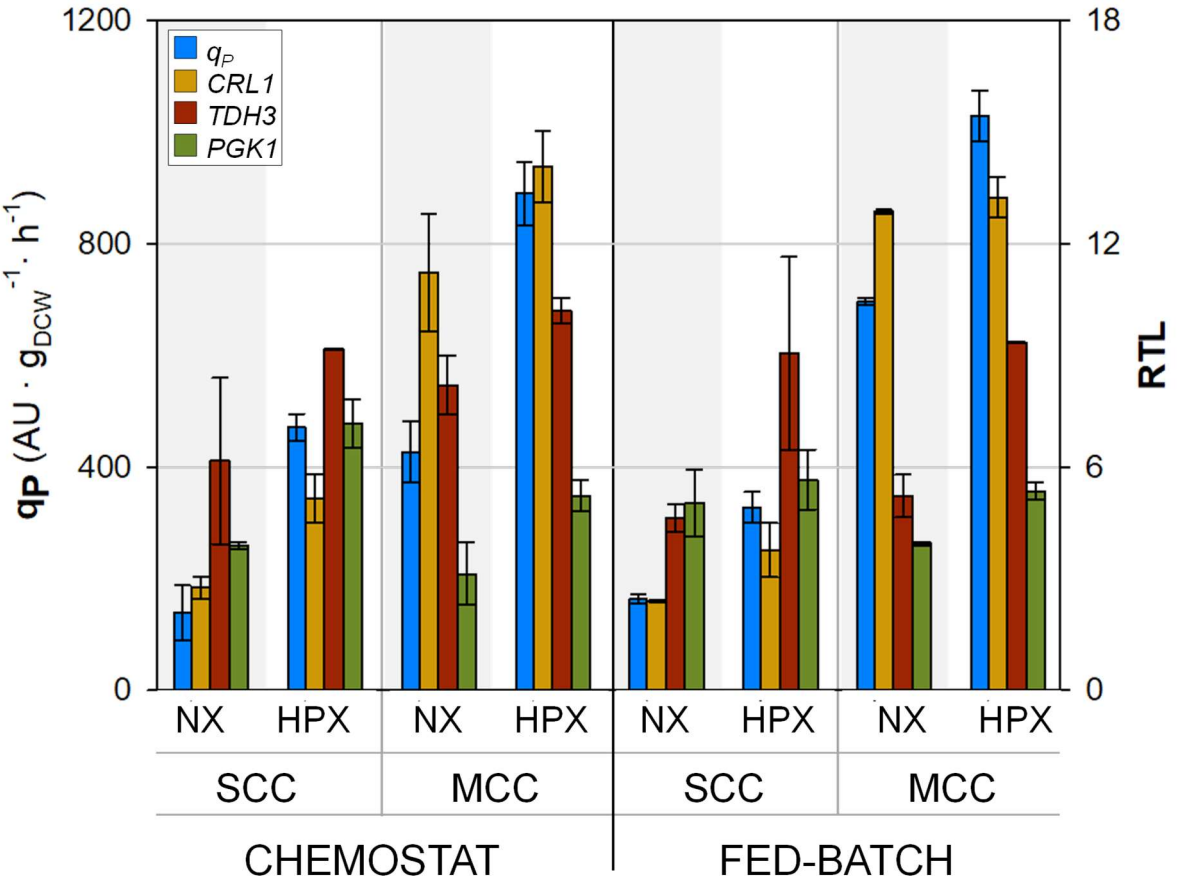


Figure 17. Comparison of q_P and relative transcript levels (RTL) of key genes related to Crl1 production and glucose metabolism under all conditions and operational modes tested. Error bars represent the SD of all samples belonging to the same group of culture conditions (hypoxic and normoxic).

As regards the glycolytic genes selected, *TDH3* and *PGK1*, as expected, there are no significant differences with respect to *CRL1* gene dosage in terms of transcript levels. Interestingly, on the other hand, the transcription levels of the glycolytic genes analyzed are boosted under oxygen limitation conditions. Accordingly, it could be stated that oxygen-limiting conditions have a significant effect on the upregulation of glycolysis-related gene expression, as previously reported (Baumann et al., 2010; Tomàs-Gamisans et al., 2020). Additionally, in previous studies using the same clones but analyzing the effect of the specific growth rate, a direct correlation between *TDH3* RTL and μ was also observed. However, this effect was not that evident with *PGK1* RTL (Nieto-Taype et al., 2020a).

When comparing q_P and *CRL1* RTL, a similar trend could be observed: they increase in MCC with respect to SCC. In addition, oxygen-limiting conditions also increase q_P and *CRL1* RTL. Interestingly, it is clearly seen that the largest increase is due to the higher gene copy number rather than the effect of oxygen limitation.

It is also important to note that no main differences could be observed between chemostat and fed-batch transcriptional results, although the effects of gene dosage and oxygen limitation seem to have more impact on chemostat than on fed-batch cultures, except for the *TDH3* gene, where the differences between the normoxic and hypoxic state are clearer in fed-batch cultures.

To provide further insights regarding the transcriptional results, two new parameters were defined: gene dosage fold-change (FC_{GD}) which corresponds to the RTL quotient or fold-change between MCC and SCC of the aforementioned genes (and q_P quotient or fold-change); and oxygen limitation fold-change (FC_{OL}), which corresponds to the RTL (and q_P) quotient or fold-change between hypoxic and normoxic conditions. The former facilitates the comprehension of the effect of gene dosage, while the latter highlights the effect of oxygen limitation. The values of these ratios are plotted in **Figures 18** and **19**. The values of these parameters are discussed in **Sections 5.3.1** and **5.3.2**.

Table 8. Value of Relative Transcript Levels (RTL) of key genes related to Crl1 production (*CRL1*), glucose metabolism (*TDH3* and *PGK1*), and UPR (*HAC1* and *KAR2*) under all conditions and operational modes tested. ± indicate standard deviation (SD) of the biological replicates.

O_2 supply condition		Chemostat		Fed-batch	
		Normoxia	Hypoxia	Normoxia	Hypoxia
SCC	***CRL1* RTL**	2.75 ± 0.31	5.17 ± 0.66	2.39 ± 0.02	3.78 ± 0.73
	***TDH3* RTL**	6.17 ± 2.23	9.18 ± 0.02	4.64 ± 0.38	9.06 ± 2.59
	***PGK1* RTL**	3.90 ± 0.09	7.18 ± 0.65	5.04 ± 0.89	5.67 ± 0.8
	***HAC1* RTL**	0.30 ± 0.11	0.26 ± 0.02	0.34 ± 0.13	0.54 ± 0.08
	***KAR2* RTL**	0.91 ± 0.34	0.69 ± 0.01	0.59 ± 0.17	0.89 ± 0.01
MCC	***CRL1* RTL**	11.23 ± 1.57	14.07 ± 0.96	12.87 ± 0.05	13.25 ± 0.55
	***TDH3* RTL**	8.21 ± 0.79	10.21 ± 0.34	5.24 ± 0.57	9.35 ± 0.02
	***PGK1* RTL**	3.14 ± 0.85	5.24 ± 0.42	3.95 ± 0.06	5.36 ± 0.24
	***HAC1* RTL**	0.22 ± 0.04	0.22 ± 0.05	1.21 ± 0.08	0.24 ± 0.02
	***KAR2* RTL**	0.91 ± 0.02	1.48 ± 0.30	1.30 ± 0.11	0.80 ± 0.19

5.3.1 FC_{GD} – Gene dosage fold-change

Concerning gene dosage (FC_{GD}), as expected, there is no effect on glycolytic reporting genes (*TDH3*, *PGK1*), since the values of *TDH3* FC_{GD} and *PGK1* FC_{GD} are around 1 in all cases, as can be observed in **Figure 18**. However, these results identified a possible bottleneck regarding Crl1 production in MCC, considering that the *CRL1* FC_{GD} ratio is systematically superior to the q_P FC_{GD}, irrespective of the operational mode or oxygenation conditions. This fact suggests a possible limitation in Crl1 processing, including translation, folding, and/or secretion, as already described in the literature for RPP processes (Ahmad et al., 2014; Puxbaum et al., 2015). Specifically, under normoxic conditions, the *CRL1* FC_{GD} values are around 5, in consonance with the increased number of *CRL1* gene copies, whereas q_P FC_{GD} values are between 3 – 4. A similar effect at high μ with this Crl1 producer MCC was also reported and thoroughly discussed (Nieto-Taype et al., 2020a).

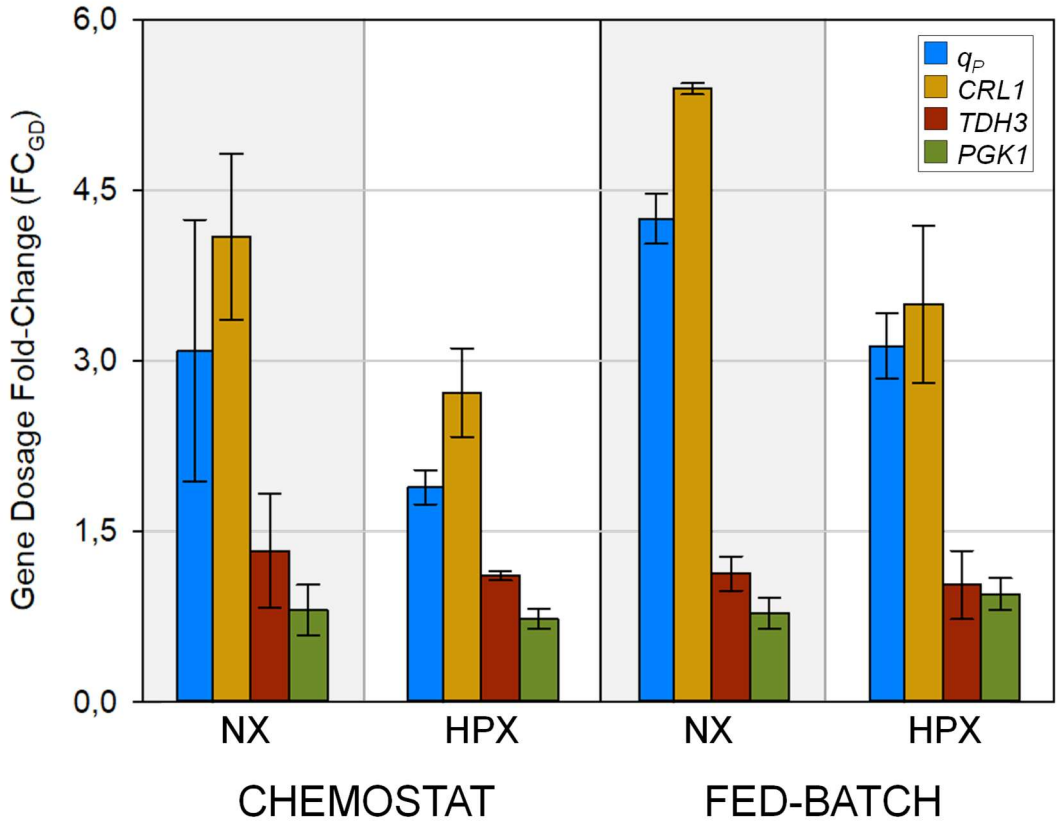


Figure 18. Gene dosage fold-change (FC_{GD}) or ratio between MCC and SCC for q_P and relative transcript levels (RTL) of key genes under all conditions and operational modes tested.

to ensure the genetic stability of the clones and to remark that there were no losses of *CRL1* expression cassettes, as reported in the literature for different recombinant multicopy strains (Zhu et al., 2009; Aw & Polizzi, 2013).

Very interestingly, the values of the *CRL1* FC_{GD} and q_P FC_{GD} in chemostat with $D = \mu = 0.10\ h^{-1}$ and under normoxic conditions (*CRL1* FC_{GD} = 4.09 ± 0.73; q_P FC_{GD} = 3.09 ± 1.15) are very similar to those obtained in the previous work (*CRL1* FC_{GD} = 4.60 ± 0.60; q_P FC_{GD} = 3.40 ± 0.40), which demonstrates the robustness and accuracy of both the chemostat cultures and the transcriptional analysis (Nieto-Taype et al., 2020a).

5.3.2 FC_{OL} – Oxygen limitation fold-change

Concerning the oxygen limitation effect (FC_{OL}), **Figure 19** shows that glycolytic genes (*TDH3* and *PGK1*) are about 1.5-fold overexpressed under hypoxic conditions, with small deviations. This aligns with the generally observed q_S increase of about 1.25-fold. Furthermore, in previous works where hypoxic conditions were implemented, an important increase in *TDH3* transcription was also observed (Tomàs-Gamisans et al., 2020). On the other hand, regarding Crl1 production, and reiterating the above results, hypoxic conditions have less impact on MCC than on SCC, both at the transcription (*CRL1* RTL) and protein secretion (q_P) levels. A plausible explanation for this effect is given in the next **Section 5.3.3**. Strikingly, the positive effect of oxygen limitation over Crl1 production is quite more pronounced in chemostat operational mode, but only in terms of q_P. This supports the hypothesis that cell stress caused by hypoxic conditions combined with the cell aging effect, characteristic of fed-batch operation, leads to smaller Crl1 production with the latter operational mode. Besides, since *CRL1* FC_{OL} values are similar in both operational modes but q_P FC_{OL} is higher in chemostat, this limitation in Crl1 production probably occurs in a step somewhere after transcription, namely folding, secretion, etc.

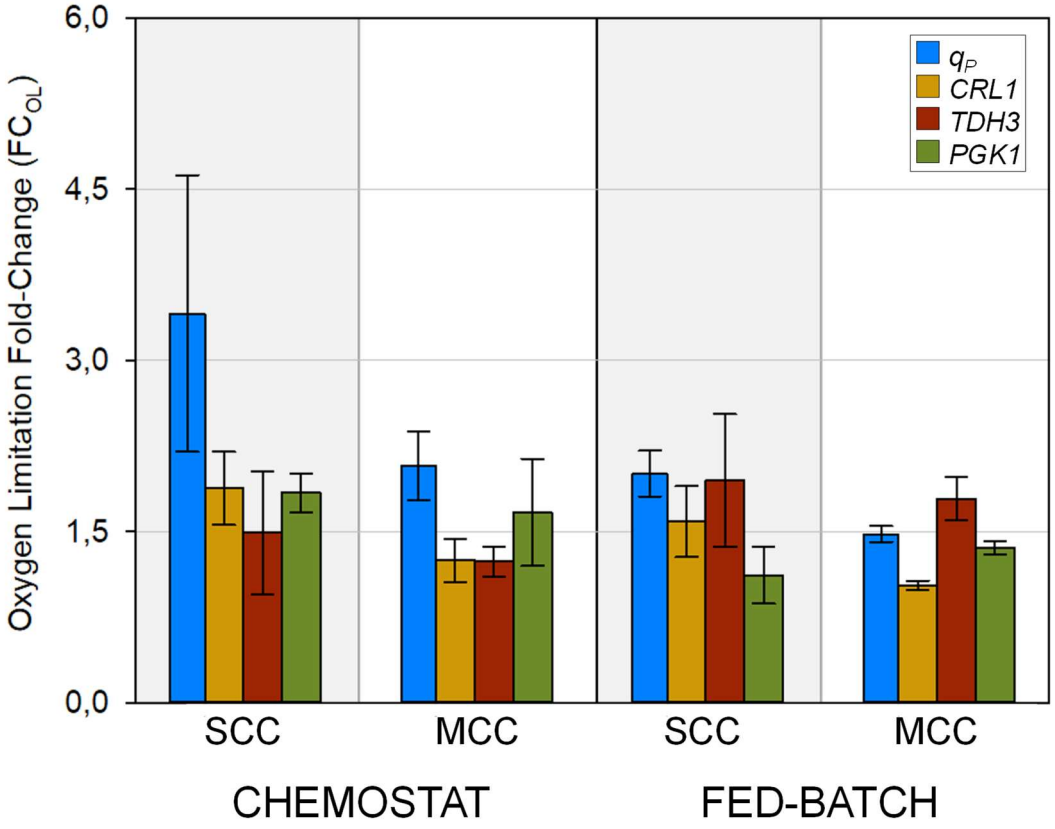


Figure 19. Oxygen limitation fold-change (FC_{OL}) or ratio between hypoxic and normoxic conditions for q_P and relative transcript levels (RTL) of key genes for both clones and operational modes tested.

5.3.3 Possible bottleneck in transcription due to a limited pool of transcription factors

From the results shown in **Figures 17**, **18**, and **19**, it can be stated that both parameters studied in this work enhance the *CRL1* transcription rate. However, the gene dosage effect is reduced under hypoxic conditions, where both *CRL1* FC_{GD} and q_P FC_{GD} are lower. This could indicate that, when combining both the heterologous gene copy number increase and hypoxic conditions, the transcriptional machinery is working closer to its upper rate limit, apart from the potential protein processing limitation mentioned previously.

Although there is not much information about the *cis*- and *trans*-acting elements involved in P_{GAP} regulation (Vogl & Glieder, 2013; Ata et al., 2017), a plausible explanation for such effect could be a finite availability of *TDH3*-related transcription factors (TF), such as Gal4-*like* family, found to be crucial in the central carbon metabolism regulation (Ata et al., 2017, 2018; Kalender & Çalık, 2020). Specifically, it was found that Cra1, a TF from the Gal4-*like* family, plays an important role in the shift from respirative to respiro-fermentative metabolism through the interaction with P_{GAP} and other glycolytic promoters, being able to activate the so-called Crabtree effect, increasing the metabolic fluxes through glycolysis and ethanol generation pathway and reducing the tricarboxylic acid cycle (TCA) activity (Ata et al., 2017, 2018; Prielhofer et al., 2018). This effect was also described for *S. cerevisiae* with *GAL4* (the homolog of *CRA1* in *S. cerevisiae*) (Frick & Wittmann, 2005; Vemuri et al., 2007). Thus, dramatically increasing the number of transcription factor binding sites (TFBSs), which occurs in MCC, with a limited pool of these transcription factors, could cause an attenuation of the transcription of all P_{GAP} regulated genes, as previously described for clones with multiple copies of a recombinant gene under the regulation of P_{AOX1} (Cámara et al., 2017; Garrigós-Martínez et al., 2019).

Indeed, this would align with the results from a previous study: a clone in which *CRA1* was overexpressed and the Cra1-binding site (within the sequence of P_{GAP}) was duplicated showed about 1.5-fold higher recombinant production than a clone in which only the Cra1-binding site was duplicated, indicating a possible scarcity of this TF when its TFBSs are duplicated (Ata et al., 2017). Thus, the

same effect might be happening in this case with MCC, which has 5 copies of P_{GAP} and so, many more Cra1-binding sites than SCC. However, in this work, no significant attenuation of neither *TDH3* nor *PGK1* transcription was observed, and q_S had very similar values regardless of the gene dosage.

To sum up, the combination of oxygen limitation and increased gene dosage has a synergic but not summatory effect at the transcriptional level. Accordingly, the increase in Crl1 production achieved by the combination of these two strategies is larger than that obtained using these two strategies separately. Moreover, the insights described shed light on the implications of the P_{GAP} regulation over RPP and bioprocess and strain engineering.

5.4 Scale up of the Crl1 production process

5.4.1 Using RQ as a transferable operating parameter to scale up

Scaling up a fermentation bioprocess is a complex engineering procedure that involves several critical factors, from biological to physical and chemical (Hewitt & Nienow, 2007; Takors, 2012). Traditionally, different scaling up strategies have been described in the literature, using distinct criteria to be able to reproduce the results obtained in the lab to an industrial scale. Some of these strategies are based on engineering principles, often related to the mixing properties or the oxygen transfer capacity. For example, keeping a constant k_La or applying a constant volumetric power consumption (P/V) are the most common strategies in industrial production applications to scale up, although geometry is also used as a scale up criterion (Marques et al., 2010; Neubauer & Junne, 2016). Other strategies are based on avoiding excessive shear stress to damage the cells, whenever it can be a critical factor. For instance, constant impeller tip speed is also used as a rule for scale up purposes (Dubey et al., 2008; Marques et al., 2010; Neubauer & Junne, 2016). More recently, scale up strategies based on the use of metabolic mathematical models combined with computational fluid dynamics (CFD) have been successfully applied (Xia et al., 2015; Wang et al., 2020).

However, a common problem in a scaling up procedure is the relatively low representativeness of the measures on which the control of these bioprocesses is based, since it is quite usual to observe chemical heterogeneities in a large-scale fermenter regarding substrate availability, for example, or temporal pH gradients due to relatively high mixing times, frequent in large-scale fermenter (Hewitt & Nienow, 2007; Takors, 2012; Neubauer & Junne, 2016; Ask & Stocks, 2022). In addition, the great height of industrial fermenters, compared to their laboratory counterparts, generates higher hydrodynamic pressures at the bottom of the fermenters than at the top. This fact, combined with the aeration at the underside of the fermenter, generates huge oxygen gradients in the culture broth (Onyeaka et al., 2003; Takors, 2012; Neubauer & Junne, 2016). All these heterogeneities lead cells to be exposed to an alternating environment regarding oxygen and substrate availability, pH, and even temperature (Xia et al., 2015).

On the other hand, exhaust gas analysis is a very reliable measurement, and it is not affected by possible heterogeneities or concentration gradients since mixing in the gas phase is much more complete than in the liquid phase due to the fluid characteristics. Therefore, using a scale up strategy based on the exhaust gas analysis instead of another invasive measurement seems *a priori* a good starting point for a scaling up process (Garcia-Ortega et al., 2017).

With this approach, a series of two fermentations, one under normoxic conditions and another under hypoxic conditions were carried out in a pilot-scale 50 L fermenter, trying to mimic the same conditions applied in the fed-batches from **Section 5.2** but with a starting batch volume of 30 L and a final volume of around 50 L at the end of the fed-batch, giving an important scaling factor of 1:15. However, due to operational limitations associated to the non-availability of pure oxygen to enrich the inlet gas in the pilot plant, implementing a μ = 0.10 h^{-1} was not possible, since in the last stages of the fed-batch, especially for the fermentation under normoxic conditions, some pure oxygen would have been necessary. Thus, a lower μ was selected to avoid having operational limitations. In a preliminary study whose results have not been included in this work, it was established for SCC that 0.12 h^{-1} < μ_{max} < 0.15 h^{-1} under hypoxic conditions with RQ = 1.4. Thus, a medium μ = 0.065 h^{-1} was selected for these scaling up fermentations, which is about half of the μ_{max} and allows to avoid the use of pure oxygen. As in both the normoxic and the hypoxic fermentations, the same μ = 0.065 h^{-1} was implemented, a proper comparison between both conditions can be carried out. Besides, since the improvement in Crl1 production and productivity is less important with MCC, as observed in **Sections 5.1** and **5.2**, SCC was selected for this scale up implementation. Additionally, bringing biomass from cell bank storage to the final industrial fermenter requires several additional steps compared to a lab-scale fermentation, which can affect the genetic stability of multicopy strains and force them to lose expression cassettes, thus giving more obvious reasons to test only the SCC (Takors, 2012).

As explained in **Section 4.2.3**, these fermentations were carried out in the facilities of Bisy GmbH (Wünschendorf, Austria) as part of an international research stay. In addition, these fermentations were also used by ZETA's team to gain further understanding of the behavior of the *BIRE* fermenter, which is actually a scale-down of one of their large-scale fermenters. Among other factors, the relationship between agitation and k_La was especially studied with the aim of developing a DO cascade control.

The values of biomass dry cell weight, Crl1 titer, and ethanol concentration during the feeding phase of normoxic and hypoxic fermentations are plotted in **Figure 20**.

In contrast to the previous fed-batch cultures, which were grown at μ = 0.10 h^{-1} and lasted for 20 h, in this case, the feeding phase was longer and ended at t = 30 h. In the normoxic fermentation, the final biomass dry cell weight was around 110 g · L^{-1}, considerably higher than for the hypoxia, which reached almost 90 g · L^{-1}. As explained below in more detail, the biomass profile in the hypoxic fermentation had some deviation from the expected exponential profile during the first 10 h. This deviation can be detected more clearly in **Figure 23** in the next **Section 5.4.2**.

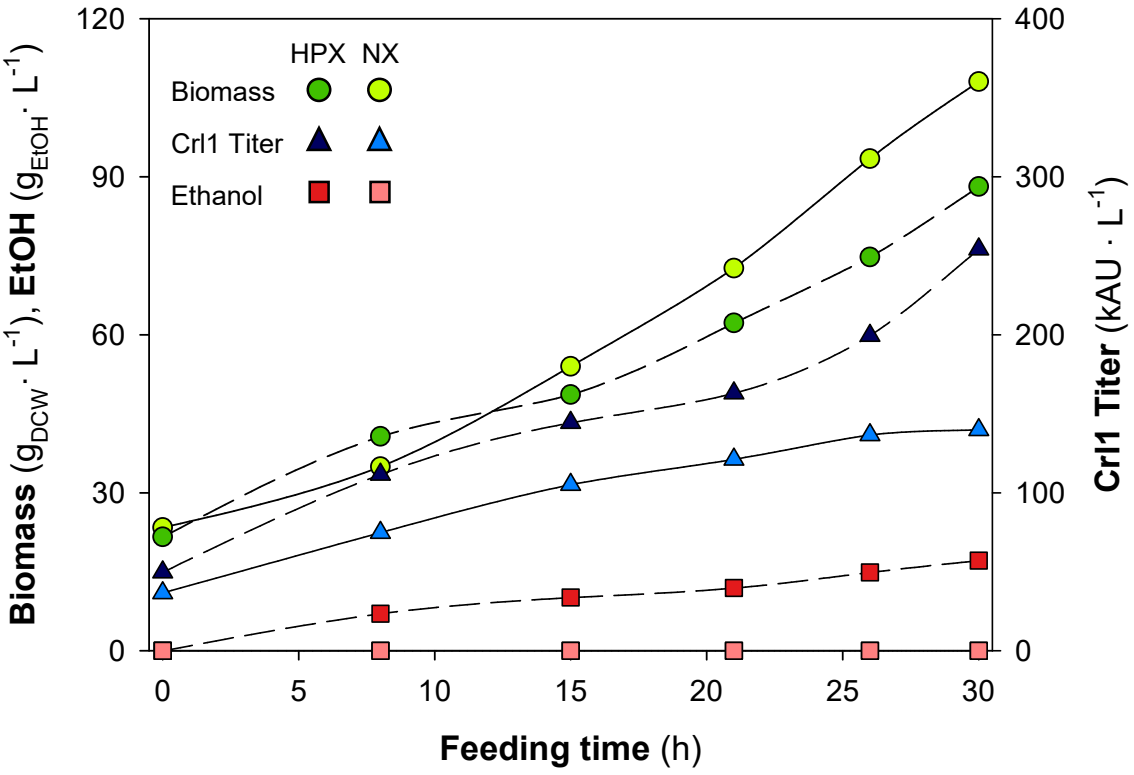


Figure 20. Primary data of the main key process parameters for the pilot-scale fed-batch cultivations grown at hypoxic (dark symbols) and normoxic (dark symbols) conditions (HPX50L, NX50L). Biomass concentration (●, g · L^{-1}); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}).

Contrary to the previous lab-scale fermentation, in this case, the total volume of substrate added to the fermenter was 20 L for both fermentations, since together with the 30 L batch medium, it was the maximum limit of the fermenter's working capacity (50 L). Then, to rectify the reduced biomass generation in hypoxic conditions due to the decrease in biomass-to-substrate yield ($Y_{X/S}$), the glucose concentration of the feeding was notably higher in the feeding tank of the hypoxic fermentation (glucose concentration = 455 g · L^{-1}). Compared to the normoxic fermentation (glucose concentration = 410 g · L^{-1}), it accounted for 10% more substrate addition. However, the expected yields estimated from lab-scale fermentations were $Y_{X/S}$ = 0.59 $g_{DCW} \cdot g_S^{-1}$ and $Y_{X/S}$ = 0.47 $g_{DCW} \cdot g_S^{-1}$, respectively, giving a $Y_{X/S}$ decrease of about 20% in hypoxic conditions. Still, it was decided to be conservative and not to reduce the substrate concentration even more in the normoxic fermentation, since it was the first occasion when this procedure of

varying substrate concentration was tested. As a result, this correction of the substrate concentration and thus of the substrate addition was not sufficient to counteract the reduction of $Y_{X/S}$, since the final biomass concentration differs greatly between the two fermentations. Indeed, as observed in **Table 9** from **section 5.4.2**, the final reduction of $Y_{X/S}$ when implementing hypoxia was around 25%, slightly higher than the expected 20% reduction and clearly higher than the 10% reduction of substrate concentration applied in this pair of fermentations.

Crl1 production increased through the feeding phase, reaching a final Crl1 enzymatic activity of 140 kAU · L^{-1} in normoxic conditions and 254 kAU · L^{-1} in hypoxic conditions. During the hypoxic fermentation, higher Crl1 production than the achieved was expected. However, the same operational issue that caused the growth deviation described above also caused a low level of Crl1 production during the first half of fermentation. It is discussed at the end of this section.

Regarding *RQ* control through manual agitation rate modification, **Figure 21** shows the time evolution of *RQ* together with the agitation speed for both strategies. As previously mentioned, in industrial scale dissolved oxygen (DO) gradients are likely to be present in the culture broth (Onyeaka et al., 2003; Takors, 2012; Neubauer & Junne, 2016). Thus, since the DO probe was installed at the bottom of the fermenter, it was decided to keep the agitation constant at 900 rpm in the normoxic fermentation to assure complete aerobic conditions, maintaining a DO > 30% for all the fed-batch. Furthermore, this was done also to avoid a having variable agitation rate that could lead to operational issues related to foam formation, which is known to impact more in large-scale fermenters (Vardar-Sukan, 1998). It is noteworthy that 200 mbar of overpressure was applied to the fermenter in the last 2 hours of the feeding phase to maintain the DO at the set-point since oxygen transfer increases with overpressure following the Henry-Dalton's law (Garcia-Ochoa & Gomez, 2009; Takors, 2012). On the other hand, in the hypoxic fermentation, the DO was just above zero and no overpressure was found necessary.

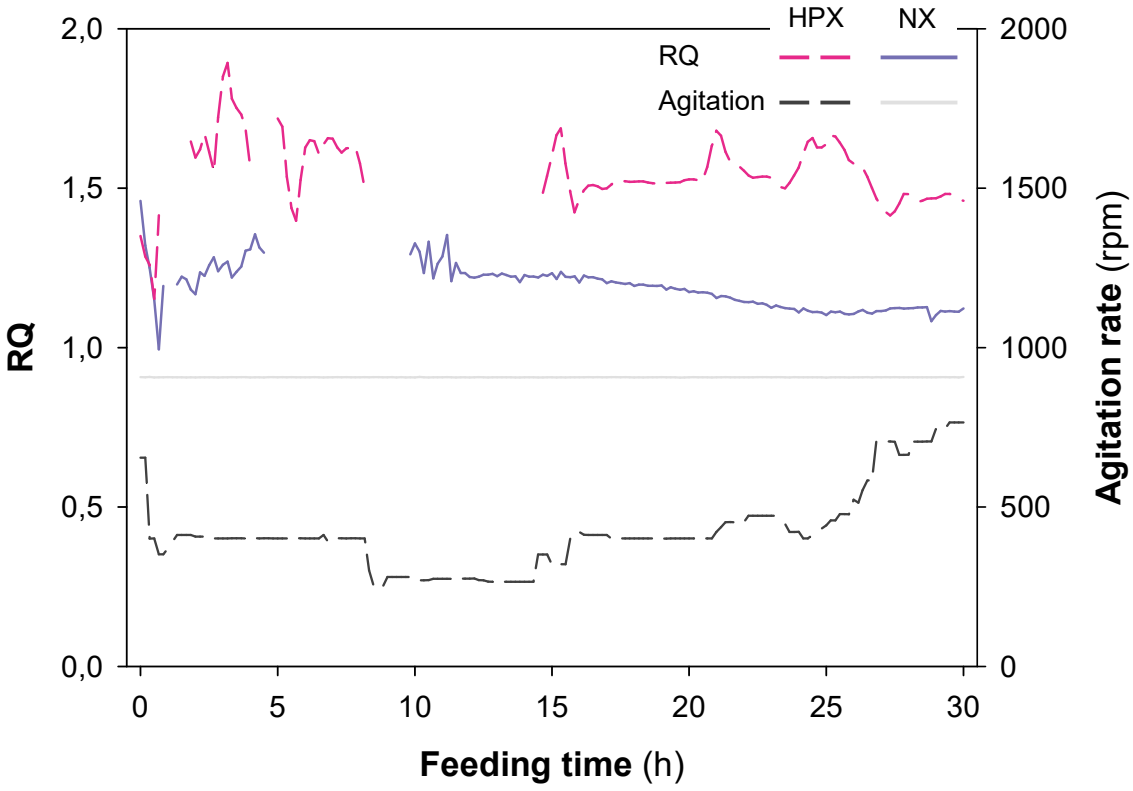


Figure 21. Off-line calculation of *RQ* and on-line data of agitation rate for the pilot-scale fed-batch cultivations grown at hypoxic (discontinuous lines) and normoxic (continuous lines) conditions (HPX50L and NX50L). Off-line *RQ* (upper lines); Agitation rate (lower lines, rpm).

*A gap in the *RQ* measurements indicates an unexpected stop of the data logging software during the night hours.

RQ during the normoxic feeding phase was rather constant around 1.1 – 1.25 with some oscillations in the first hours. DO sensor values above 30% during all the normoxic fermentation confirmed that there was no oxygen limitation, although these *RQ* values are slightly higher than those observed in lab-scale normoxic fermentations. Inasmuch as these fermentations were not performed with the same fermentation equipment, external measuring devices, and software as all the other lab-scale fermentations carried out in this thesis, the calculation of *RQ* was not so accurate. Basically, the gas analyzer could not be calibrated with different gas mixtures with known O_2 and CO_2 concentrations as done in the

lab-scale fermentations, as explained in **Section 4.3.6**, and it was only calibrated with air. This could have led to a small overestimation of CO_2 concentration during the first fermentation (normoxic) which, in turn, caused an overestimated *CER* values and thus, a higher *RQ* profile than it actually was. This could explain why *RQ* > 1.2 when DO > 30%, since in lab-scale normoxic fermentations *RQ* had been always between 1.0 and 1.1.

Bearing this in mind, in the subsequent fermentation, which was the hypoxic one, *RQ* was controlled with an *RQ* set-point = 1.5, higher than 1.4 applied in lab-scale. However, this manual control at a large-scale was not sufficiently accurate, and *RQ* oscillated between 1.5 and 1.8 approximately, with small peaks associated with manual control actions over agitation rate, as can be observed in **Figure 21.**

This uncertainty in the *RQ* measurement and control during the hypoxic fermentation caused more severe hypoxic conditions than those applied in lab-scale, thus causing an increase of the fermentative metabolism confirmed by a higher ethanol accumulation than those observed with the 5 L benchtop fermenter. Specifically, the final ethanol concentration reached in this fermentation was above 17 $g \cdot L^{-1}$, significantly superior to the 10 $g \cdot L^{-1}$ obtained in the lab-scale. Additionally, up to 1.3 and 1.2 $g \cdot L^{-1}$ of arabitol and succinic acid, respectively, were detected in the culture broth at the end of the feeding phase.

As advanced previously, in the hypoxic fermentation, due to a malfunction of the feeding pump during the first 8 hours of the feeding phase, the substrate addition was higher than desired. Therefore, the biomass generation was also higher, as can be detected in the hypoxic biomass profile (t = 8 h) in **Figure 20**. After realizing this operational issue, the feeding profile was slightly recalculated to reach a total substrate addition of 20 L at t = 30 h, thus leading to a reduced substrate addition from t = 8 h to t = 15 h.

Nonetheless, this deviation from the predefined exponential feeding profile had some effect on the growth rate, thus causing a $\mu >> 0.065\ h^{-1}$ during the first 8 hours and a $\mu << 0.065\ h^{-1}$ from this point to t = 15 h. During these periods, μ can be estimated as $\mu \approx 0.10\ h^{-1}$ and $\mu \approx 0.05\ h^{-1}$, respectively. From this point on, μ

was controlled appropriately. Although the final mean μ calculated for all the feeding phase was μ = 0.062 h^{-1}, really close to the μ set-point, this deviation caused a gradual halt in Crl1 production. **Figure 22** shows the discrete values of μ and q_P, calculated at each sample time point as explained in the methodology section. The deviation of μ from the set-point and its relationship with a clear reduction of q_P when μ was decreased can be clearly observed in the graph. This is consistent with previous results obtained at lab-scale with this clone, in which a clear correlation between μ and q_P was observed (Nieto-Taype et al., 2020a). Meaningfully, if μ was lower than the predefined value, so was q_P. In addition, these values seem to indicate that this negative effect on recombinant production was more extended in time than the effect on growth.

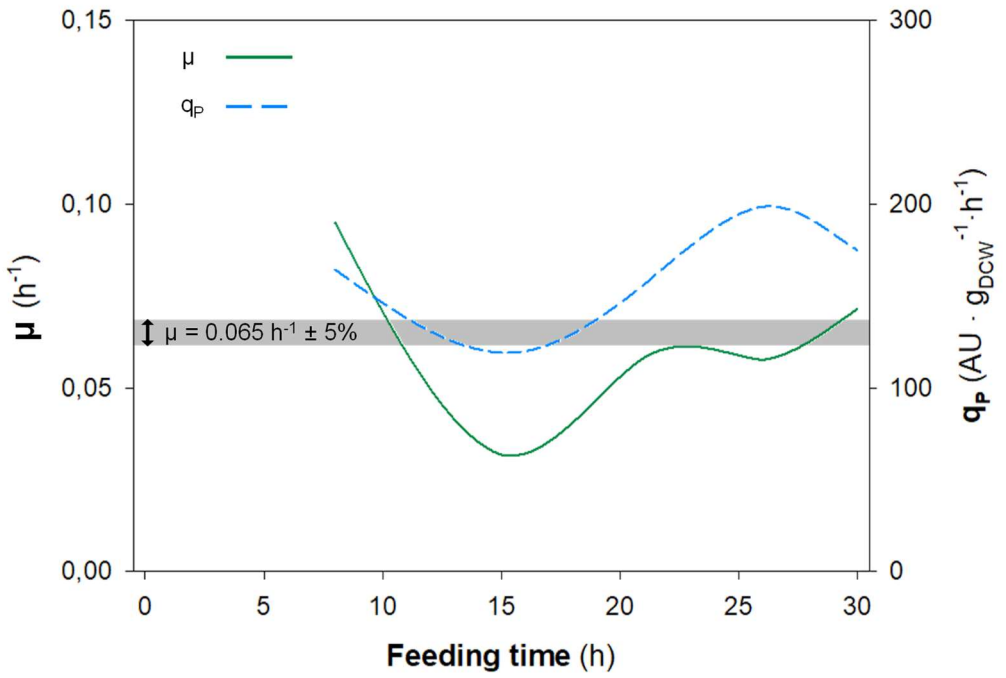


Figure 22. Off-line calculation of discrete values of specific growth rate, μ (green continuous line, h^{-1}) and specific production rate, q_P (blue discontinuous line, AU · g_{DCW}^{-1} · h^{-1}) for the hypoxic pilot-scale fermentation (HPX50L). A gray horizontal band indicates the $\mu_{set\text{-}point}$ = 0.065 h^{-1} with a ± 5% deviation.

5.4.2 Performance comparison between pilot and lab-scales

The values of different key process parameters (μ, q_S, $Y_{X/S}$, q_{EtOH}, RQ, q_P, $Y_{P/X}$, and final product $P \cdot V$) are displayed in **Table 9**. It is worth noting that the q_P value of the pilot-scale hypoxic fermentation was calculated considering only the last three sample points, considering all the explained in the previous section and shown in **Figures 20** and **22**. To provide a proper comparison, two lab-scale fermentations, one under normoxic conditions and one under hypoxic conditions, grown at μ = 0.065 h^{-1} (065-NX and 065-HPX), have also been included in **Table 9**. These fermentations are part of a set of cultivations conducted with SCC and applying different specific growth rates, to evaluate the production kinetics, and they were performed by the PhD candidates Albert Sales and Guillermo Requena, who gently provided the data (Sales et al., 2023).

Table 9. Value of key process parameters obtained in fed-batch fermentations with SCC at lab- and pilot-scales. Specific growth rate, μ (h^{-1}); specific substrate consumption rate, q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); biomass-to-substrate yield, $Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$); specific ethanol production rate, q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); respiratory quotient, RQ; specific Crl1 production rate, q_P (AU $\cdot g_{DCW}^{-1} \cdot h^{-1}$); product-to-biomass yield, $Y_{P/X}$ (kAU $\cdot g_{DCW}^{-1}$); and final total product amount, expressed as multiple of 10^6 (AU).

O_2 supply condition	Pilot-Scale (50 L)		Lab-Scale (5 L)	
	Normoxia	Hypoxia	Normoxia	Hypoxia
μ (h^{-1})	0.066	0.062	0.059	0.064
q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.11	0.14	0.13	0.15
$Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$)	0.58	0.43	0.48	0.42
q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	n.d.	0.027	n.d.	0.027
RQ	1.10	1.45	1.05	1.43
q_P (AU $\cdot g_{DCW}^{-1} \cdot h^{-1}$)	84	198	128	267
$Y_{P/X}$ (kAU $\cdot g_{DCW}^{-1}$)	1.40	2.91	2.17	4.18
$P \cdot V$ ($\cdot 10^6$ AU)	6.99	12.7	0.53	1.33

A general reduction of around 25% in q_P was observed comparing lab- and pilot-scale fermentations, which is commonly found in industrial scaling up (Lara et al., 2006; Xia et al., 2015). Accordingly, comparing the $Y_{P/X}$ values achieved at pilot-scale with values obtained at the lab-scale, an approximate 30% reduction is observed both in hypoxic but also in normoxic conditions. This indicates that, although a sub-optimal performance could be considered for the pilot-scale hypoxic fermentation, the reduction of Crl1 production could be associated more with the scale-up process rather than the capability to reproduce normoxic and hypoxic conditions.

Figure 23 shows the total amount of biomass and ethanol produced in pilot- and lab-scales. To evaluate the reproducibility of the results, the ratio between y-axes (left: lab-scale; right: pilot-scale) has been defined at 15, the same as the scale up factor. Thus, it can be observed that biomass and ethanol concentrations are comparable in both scales. However, slightly higher ethanol concentration is observed in the large-scale fermentation, caused by highly hypoxic conditions due to the improper *RQ* measurement and control already mentioned.

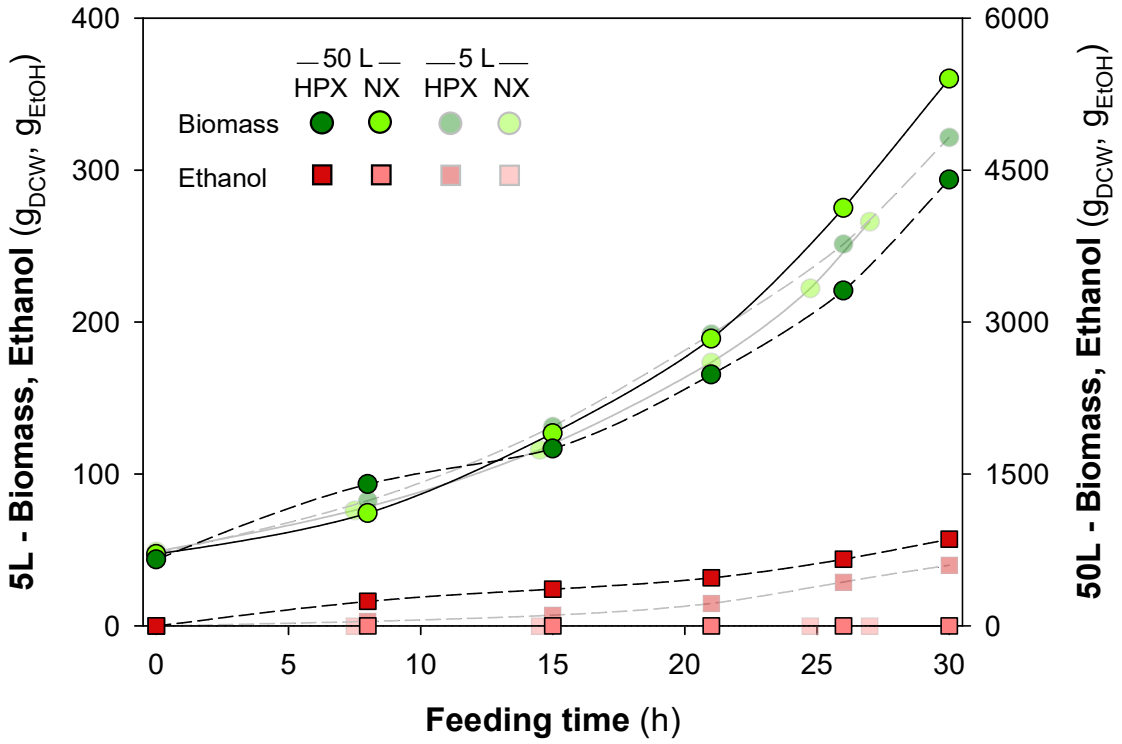


Figure 23. Total biomass and ethanol generation in the 50 L pilot-scale (opaque symbols) and 5L lab-scale (clear symbols) fermentations grown in hypoxic (dark symbols) and normoxic (light symbols) conditions (HPX50L, NX50L, 065-HPX, and 065-NX). Total biomass amount (●, g); Total ethanol amount (■, g). For a fair comparison, the relationship between y-axes is the same as the scale-up factor (15x).

Regarding Crl1 production, **Figure 24** shows the total Crl1 titer in lab-scale and pilot-scale fermentations, also with a ratio of 15 between the y-axis scales. Whereas in normoxic conditions the Crl1 production was quite comparable between lab- and pilot-scales, a proportionally lower Crl1 production was observed in the 50 L hypoxic fermentation. As already stated, the slowdown in Crl1 production observed in this fermentation should be ascribed to the halt in biomass growth that caused the transient lower *μ*.

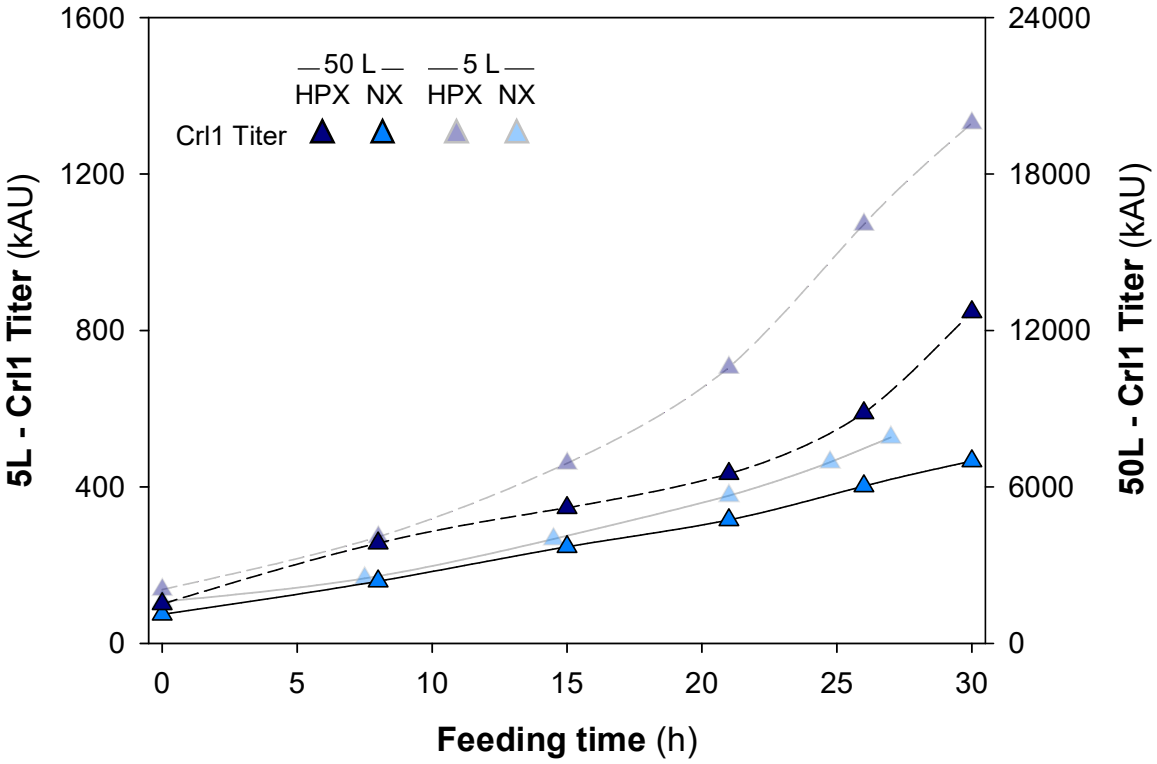


Figure 24. Total Crl1 production in the 50 L pilot-scale (opaque symbols) and 5L lab-scale (clear symbols) fermentations grown in hypoxic (dark symbols) and normoxic (light symbols) conditions (HPX50L, NX50L, 065-HPX, and 065-NX). Total Crl1 Titer (▲, kAU). For a fair comparison, the relationship between y-axes is the same as the scale-up factor (15x).

From these results, it can be concluded that the Crl1 production process has been partially successfully scaled up by a factor of 1:15, slightly higher than the standard 1:10 generally employed in industry (Garcia-Ochoa & Gomez, 2009). Although not achieving the same levels of Crl1 production, biomass grew as expected and no gross differences in μ, q_S, and $Y_{X/S}$ were observed, indicating an equivalent biomass performance. Despite the mentioned operational constraints, a 2.3-fold increase in q_P was achieved between hypoxic and normoxic conditions, quite similar to the 2-fold q_P improvement obtained in lab-scale.

More importantly, as stated previously, this scale up procedure is based on a non-invasive measurement of the exhaust gas, less vulnerable to the presence of substrate heterogeneities or oxygen gradients inside the fermenter, characteristic of large-scale fermenters and easier to apply to an already implemented production process, giving an innovative scale up criterion to an industrial field where no universal criteria are defined for scaling up purposes (Hewitt & Nienow, 2007; Marques et al., 2010; Takors, 2012; Xia et al., 2015; De Brabander, 2023). Nonetheless, proper calibration of the gas analyzer is highly necessary to apply this strategy, since the control of the bioprocess depends on it, as these results show.

6. RESULTS II

Implementing control strategies for hypoxic fed-batch fermentation: from manual control to the application of Artificial Intelligence (AI) algorithms

An article containing an adapted and summarized version of Chapter 6 is being prepared for submission to *Frontiers in Bioengineering and Biotechnology*:

Gasset, A., Van Wijngaarden, J., Mirabent, F., Sales, A., Garcia-Ortega, X., Montesinos-Seguí, J. L., Manzano, T., Valero, F. (2023/2024). Continued Process Verification 4.0 application in upstream: Adaptiveness implementation managed by AI in hypoxic bioprocess of *Pichia pastoris*. *Frontiers in Bioengineering and Biotechnology*.

The information included in Chapter 6 is complemented by a research paper that has been published in *PDA Journal of Pharmaceutical Science and Technology*:

Ondracka, A., Gasset, A., Garcia-Ortega, X., Hubmayr, D., van Wijngaarden, J., Montesinos-Seguí, J. L., Valero, F., Manzano, T. (2023). CPV of the Future: AI-powered continued process verification for bioreactor processes. *PDA Journal of Pharmaceutical Science and Technology*, 77(3), 146-165. https://doi.org/10.5731/pdajpst.2021.012665

6.1 Defining the desirable characteristics of a hypoxic control strategy

As a conclusion from **Chapter 5**, a clear increase in RPP under P_{GAP} regulation is observed in hypoxic conditions. As confirmed by transcriptional analyses, this increase can be attributed to the overexpression of glycolytic genes, due to the aforementioned shift from respirative to respiro-fermentative metabolism. Among them, *TDH3* is the glycolytic gene that has P_{GAP} as its natural promoter, so recombinant genes under P_{GAP} regulation are also overexpressed (Gasset et al., 2022).

As also stated in the previous section, this increase in RPP is quite constant regardless of the oxygen limitation level, being only two different states in terms of Crl1 productivity or q_P: normoxia (RQ < 1.2) and hypoxia (RQ > 1.2). Nevertheless, the oxygen limitation level does have a clear impact on the metabolism of the yeast, since q_{EtOH} and q_S increase linearly with RQ, and $Y_{X/S}$ decrease also proportionally to RQ, as observed in **Section 5.1**.

A very precise control of the RQ is therefore not essential in order to achieve an improvement in Crl1 production, but it is necessary to ensure that the whole bioprocess can be reproducible, since oscillations in RQ or deviations from the set-point may cause different metabolic states (distinct glucose utilization) and so, different $Y_{X/S}$, $q_{EtOH,}$ and biomass generation. Although q_P may remain the same for the whole range of hypoxic conditions, Crl1 production will differ from one fermentation to another if biomass generation is not equivalent. From an industrial perspective, reproducibility is crucial when implementing a productive bioprocess (Galvanauskas et al., 2019; Simutis & Lübbert, 2015; Veloso & Ferreira, 2017). It is therefore necessary to develop a precise RQ control strategy if the project is to be scaled up to an industrial level. As mentioned in **Section 2.5**, both EMA and FDA specify that reproducibility is an essential aspect of the validation process of the manufacturing of biotechnological or biomedical products (FDA, 2011; EMA, 2014). Furthermore, the selection of a suitable set-point of RQ is equally essential to assure the desired results. The selection of the appropriate RQ set-point is addressed in **Section 5.2.1**, but it can be summarized in the following two points: first, hypoxic conditions can be assured if RQ > 1.2,

so a higher set-point should be selected. Secondly, the higher the *RQ*, the greater the ethanol production and the lower the $Y_{X/S}$, both factors affecting the efficiency of the bioprocess. As shown in the Annex, a set-point for *RQ* greater than 1.6 resulted in an excessive amount of ethanol being produced, which can result in growth-inhibitory (Ergün et al., 2019; Wehbe et al., 2020). Additionally, such a high *RQ* led to a very low $Y_{X/S}$, resulting in less biomass production than was desired. Taking into account these considerations, *RQ* should be maintained within the range of 1.2 to 1.6, with 1.4 being the optimal value.

Besides reproducibility, automation is also an imperative step in the implementation of industrial bioprocesses, especially in those with a high degree of complexity, as can be a fed-batch with controlled hypoxic conditions. Automation is crucial not only to guarantee the quality of the product but also to improve the reliability and the economics of the process and to be able to deal with process failures appropriately (Alford, 2006; Stanke & Hitzmann, 2013; Rathore et al., 2021). Aside from improving product quality and safety, automation also enhances human security by reducing the number of manual control activities required by plant operators, which can be potentially dangerous or jeopardize the whole bioprocess (Clementschitsch & Bayer, 2006; Stanke & Hitzmann, 2013; Daniyan et al., 2014; Luo et al., 2021). It should be noted, however, that despite both the EMA and FDA encourage biotech and biopharma industries to automate their bioprocesses as a method for continuous process verification, these industries still seem unenthusiastic about automation applying novel control strategies, such as those based on AI, as this entails re-validation steps (FDA, 2011; EMA, 2014; Embury & Clayton, 2017).

Considering all the above, when establishing a precise, reproducible, and automated control strategy, another thing to keep in mind is the requirements of the bioprocess in terms of its control. To put it another way, it is important to be aware of how the process evolves and when softer or more severe control actions will be needed.

Regarding the present bioprocess, during the initial stages of the feeding phase, when biomass concentration is low, the system dynamics are relatively slow. Contrarily, at the end of the fed-batch phase, biomass concentration and

substrate addition rate are much higher than at the beginning, causing much faster dynamics. Regardless of the duration of the whole process, the fed-batch phase can therefore be divided into three parts, each lasting one third of the entire feeding phase duration. **Figure 25** shows the total biomass generation and Crl1 production through the feeding phase of the couple of hypoxic fermentations performed with SCC and presented in **Section 5.2.2**. According to biomass profiles, the increase of biomass during the first third of the fed-batch phase constitutes less than 15% - 20% of the overall increase. During the intermediate third of the fed-batch phase, it can be estimated that the biomass increase is approximately 25% - 30% of the total biomass increase, and during the final third of the fed-batch phase, the biomass increase is approximately 50% - 60% of the overall biomass increase. The same effect can also be observed in Crl1 titers shown in **Figure 25**, and the same pattern can be inferred from substrate addition whenever it follows an exponential profile. In short, because of the exponential growth of the biomass, there are many more changes in the culture conditions during the last third of the feeding phase rather than during the first two thirds of the fed-batch (total volume, biomass amount, oxygen and heat removal requirements, etc.) (Funke et al., 2010; Garcia-Ortega et al., 2013).

From the control perspective, this implies that the control actions, whether manual or automated, during the first third of the fed-batch phase should be few and soft. Ideally, control actions should be performed more often throughout the intermediate third, and almost constantly during the last third of the fed-batch phase, with more marked actions, being critical at the last hours when biomass concentration reaches 100 $g \cdot L^{-1}$ (Brignoli et al., 2020). Hence, the control strategy should take this time-effect variation into consideration.

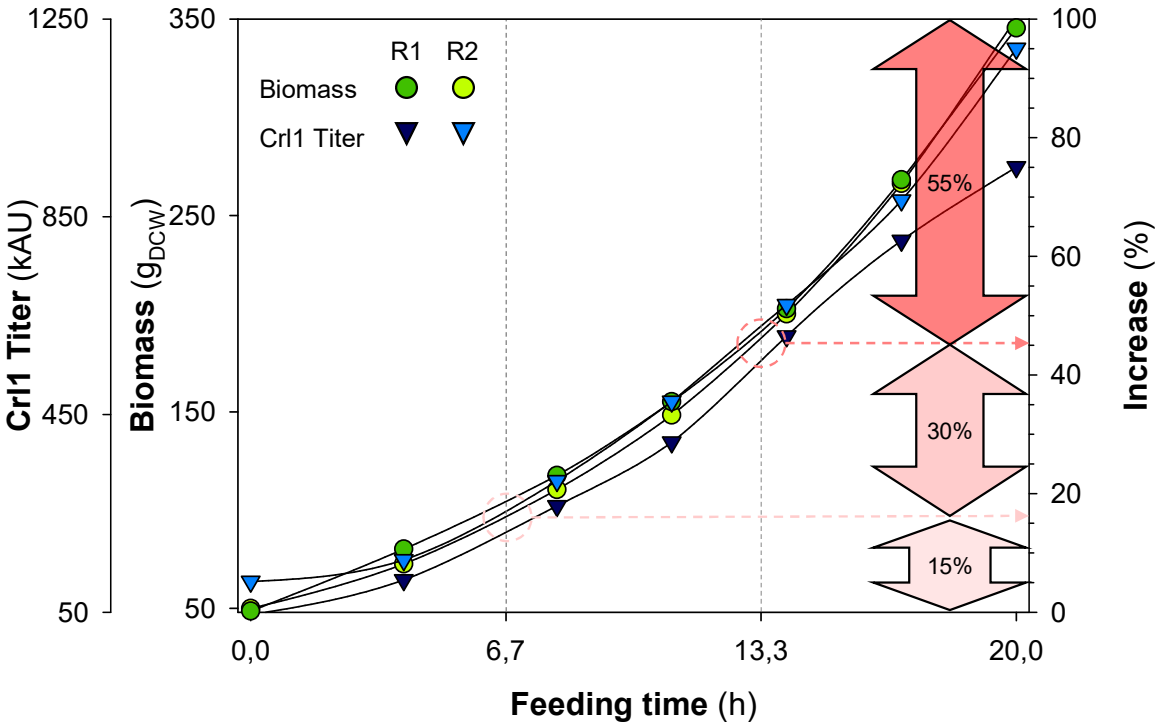


Figure 25. Total biomass generation and Crl1 production in the two duplicates of lab-scale hypoxic fermentations SCC-HPX1 and SCC-HPX2, and their corresponding relative increase at each third of the total feeding time. Total biomass amount (●, g); Total Crl1 Titer (▲, kAU). To illustrate the total increase of each variable, the limits of the Biomass and Crl1 axes have been matched with their initial and final values.

Three different approaches have been evaluated with progressively greater complexity but also with progressively better results to develop a highly precise, accurate, and fully automated *RQ* control. As mentioned in **Chapter 1**, the development and implementation of these control strategies outlined in the following sections have been done within the framework of the "CPV of the future" project, together in collaboration with AIZON (Barcelona, Spain). Especially, the last control strategy based on artificial intelligence (AI) was the final goal of the project.

6.2 Selecting the manipulated variable for a hypoxic control strategy

In contrast to chemostat fermentations, which are characterized by a steady state and constant oxygen consumption, non-stationary carbon-limited fed-batch fermentations exhibit an exponential increase in biomass, product titer, and all other growth-related components, but they are also defined by an exponential increase in oxygen consumption and carbon dioxide production. Then, any control action performed over this process must have exponential behavior. This is generally accepted for *pH* or temperature, for example: the amount of acid or base added to the culture to regulate *pH* increases exponentially during a fed-batch phase, and the heat amount added or removed from the fermenter through the jacket or other cooling systems also increases exponentially if the specific growth rate is kept constant (Dabros et al., 2010; Funke et al., 2010; Katla et al., 2019). According to a classic DO cascade control, the agitation rate is typically the first manipulated variable in the cascade, and it increases exponentially within the defined operational range. As the agitation rate reaches its maximum level, the second variable comes into action, which is often the oxygen proportion in the inlet gas, by mixing air with pure oxygen (or nitrogen) (Baeza, 2017; de Macedo Robert et al., 2019; Rathore et al., 2021). Nonetheless, the combined action of agitation rate and gas mixing allows an exponentially increasing *OTR*, thus maintaining the set-point of DO.

Then, a variable with a high impact on *RQ* should be used for the control strategy implementation. Some examples can be found in the literature where respirometric parameters are used as measured variables to implement a control strategy. Nonetheless, in these studies, the modified variable was the feeding addition, therefore not maintaining the desired μ (Jenzsch et al., 2006; Wang et al., 2007; Ranjan & Gomes, 2009; Mesquita et al., 2019). If μ is to be maintained constant, then another variable or variables should be selected for the *RQ* control strategy.

As shown in **Equations 39**, **40**, and **41**, biological and operational variables affecting *RQ* can be stated as q_{O2}, q_{CO2}, and X.

$$CER = q_{CO2} \cdot X \quad \textbf{Eq. 39}$$

$$OUR = q_{O2} \cdot X \quad \textbf{Eq. 40}$$

$$RQ = {}^{CER}/_{OUR} \quad \textbf{Eq. 41}$$

Where *CER* and *OUR* are the Carbon Evolution Rate ($mols_{CO2} \cdot L^{-1} \cdot h^{-1}$) and Oxygen Uptake Rate ($mols_{O2} \cdot L^{-1} \cdot h^{-1}$); q_{CO2} and q_{O2} are the specific carbon dioxide production rate ($mols_{CO2} \cdot g_{DCW}^{-1} \cdot h^{-1}$) and the specific oxygen consumption rate ($mols_{O2} \cdot g_{DCW}^{-1} \cdot h^{-1}$), respectively; *X* is biomass concentration ($g \cdot L^{-1}$); and finally, *RQ* stands for Respiratory Quotient ($mols_{CO2} \cdot mols_{O2}^{-1}$, but normally expressed as a non-dimensional variable). They are intrinsic to biomass growth and physiologic state and therefore they cannot be manually modified without altering *μ* or altering *RQ*, since they are included in both *CER* and *OUR* terms. *X* is biomass concentration, and since it is meant to grow exponentially, it cannot be used as a control variable either.

However, in an oxygen-limited system, all the oxygen transferred to the culture is supposed to be consumed immediately by the cells, so the oxygen transfer rate (*OTR*) is equal to the oxygen uptake rate (*OUR*), with the same units ($mols_{O2} \cdot L^{-1} \cdot h^{-1}$).

$$OTR = OUR \quad \textbf{Eq. 42}$$

Then, it is possible to modify the *RQ* by modifying the *OTR*, although this can also have an impact on *CER*.

Equation 43 shows the variables involved in the *OTR* term:

$$OTR = k_L a \cdot ({O_2}^{Sat} - O_2) \quad \textbf{Eq. 43}$$

Were k_La is the volumetric mass transfer coefficient (h^{-1}, or more commonly s^{-1}), O_2^{Sat} is the saturation oxygen concentration in the culture broth ($mols_{O2} \cdot L^{-1}$) and O_2 is the actual oxygen concentration ($mols_{O2} \cdot L^{-1}$). In hypoxic conditions, $O_2 \approx 0$, so two parameters can be considered for *OTR* adjustment: k_La and O_2^{Sat}.

Therefore, two different variables can be proposed as manipulated variables to implement the *RQ* control: agitation rate and oxygen proportion in the inlet gas. The former has a significant impact on k_La, since it affects both the turbulent regime in the bioreactor (k_L) and the interfacial area between the gas and liquid phases (a), whereas the latter influences the oxygen saturation in the culture broth (O_2^{Sat}) (Garcia-Ochoa & Gomez, 2009; Garcia-Ochoa et al., 2010). In fact, oxygen proportion is not directly manipulated. Flowrates of air and pure oxygen or pure nitrogen are actually the manipulated variables to modify the oxygen proportion in the inlet gas. From an industrial perspective, the use of pure gases is not very attractive since it entails additional transport and storage costs, safety risks, etc. (Liu et al., 2016). For this reason, agitation rate was initially selected as the manipulated variable to implement the *RQ* control. Aside from that, having only one manipulated variable instead of three facilitates the implementation of an automated *RQ* controller.

Still, a fed-batch with an *RQ* control based on the modification of the inlet gas composition was performed (HPX-GASC), just to ensure the viability of this option, obtaining similar results compared with those fermentations where the *RQ* control was achieved by modifying the agitation rate. Results are shown in the Annex.

6.3 First approach to *RQ* control: manual-heuristic modification of agitation rate (MHC)

As it has been shown in **Section 5.2**, the initial approach to *RQ* control was to perform manual control actions on the agitation rate. This set of experiments had three main objectives: firstly, and more obviously, to ensure that the improvement in RPP under hypoxic conditions observed in the chemostat cultures was also observed in a fed-batch system. Secondly, to assess the viability of a dynamic *RQ* control based on agitation modification. Additionally, these experiments also had a less apparent, yet equally relevant goal, which was to generate enough data to subsequently build an *RQ* control model using AI algorithms, which require large amounts of data in order to be trained (Das et al., 2015; Zhou et al., 2019; Ondracka et al., 2022). This question will be further discussed in the following sections.

As commented in **Section 5.2**, a total of 4 hypoxic fermentations were carried out, 2 of them with SCC and the other 2 with MCC. However, only SCC hypoxic fermentations are discussed in this section, since only SCC was tested with the other control strategies presented in this chapter. Both SCC hypoxic fermentations consisted of an initial batch with glycerol as the sole carbon source, followed by a fed-batch phase with an exponential feeding profile of glucose at μ = 0.10 h^{-1} at a constant air flow rate of 2 L · min^{-1} and an agitation rate from 600 to 1500 rpm. In addition, as reported in **Section 5.2.1**, hypoxic conditions were maintained by keeping the *RQ* approximately constant by manually adjusting the agitation.

The manual modification of the agitation rate followed heuristic rules: as commented in the previous section, increasing the agitation caused a reduction of *RQ*, so control actions were based on this principle. An operational *RQ* range was defined as $1.2 < RQ < 1.6$.

Thus, at the beginning of the fed-batch phase, where hypoxic conditions were to be implemented, an initial agitation rate was set by trial and error to achieve, approximately, $RQ = 1.4$. Then, approximately every hour after the beginning of the fed-batch, *RQ* was evaluated. If $RQ > 1.6$, then the agitation rate was

increased by a step of either 50 or 100 rpm, mainly depending on the researcher's expertise (∆rpm = 50 or 100). Since the inertia of the system is that substrate uptake rate (*SUR*) grows exponentially as substrate addition increases, *CER* also increases exponentially. To maintain a constant *RQ*, *OTR* should also increase exponentially, so the agitation rate should follow the same trend. Reductions in agitation rate during the fed-batch were therefore not expected.

In this first approach to *RQ* control, the on-line calculation of *RQ* was carried out through *BlueVis* software from *BlueSens* (Herten, Germany), with the pre-defined "*CER/OUR/RQ* Plugin". The software allowed the integration of data coming from the O_2 and CO_2 analyzers to calculate the *OUR*, *CER,* and *RQ* on-line, using **Equations 44** and **45**, from the "*BlueVis Handbook"* (BlueSens, 2017).

$$CER = \frac{F_g \cdot P}{V \cdot R \cdot T} \cdot \left(\frac{100 - O_{2_i} - CO_{2_i}}{100 - O_{2_o} - CO_{2_o}} \cdot CO_{2_o} - CO_{2_i}\right) \quad \textbf{Eq. 44}$$

$$OUR = \frac{F_g \cdot P}{V \cdot R \cdot T} \cdot \left(O_{2_i} - \frac{100 - O_{2_i} - CO_{2_i}}{100 - O_{2_o} - CO_{2_o}} \cdot O_{2_o}\right) \quad \textbf{Eq. 45}$$

Where F_g is the inlet gas flowrate (constant at 2 L · min^{-1}), *P* and T are the standard pressure (1.0133 bar) and temperature (0º C) with which the massflow controllers were calibrated, *V* is the volume of culture broth (starting at 2 L and increasing), *R* is the gas constant (8.314 · 10^{-2} bar · L · K^{-1} · mol^{-1}) and CO_{2_i}, CO_{2_o}, O_{2_i}, and O_{2_o} are the carbon dioxide and oxygen concentrations (%) in the inlet (i) and outlet (o) gas. CO_{2_i} and O_{2_i} were assumed to be 0.04% and 20.97%, respectively.

This method was found to have several limitations. Firstly, humidity both in the inlet and outlet gas plays a non-crucial but still significant role in *RQ* calculation, mainly affecting the "inert factor" $\left(\frac{100 - O_{2_i} - CO_{2_i} - H_2O_i}{100 - O_{2_o} - CO_{2_o} - H_2O_o}\right)$, which accounts for the difference between the inlet and the outgas flowrates. With the "*CER/OUR/RQ* Plugin" from BlueVis, humidity could not be taken into account in the on-line *RQ* calculation, as shown in **Equations 44** and **45**. However, the gas analyzer includes a humidity sensor, so humidity measurements were included *a posteriori*

in the *RQ* calculation, when processing all fermentation data after finishing the process.

On the other hand, the other main drawback was the inability to include the analyzer's fine-tune calibration values in the O_2 and CO_2 measurements. As explained in **Section 4.3.6**, gas analyzers were calibrated before each fermentation with a set of different gas mixtures with different O_2 and CO_2 proportions. For each analyzer, a calibration line was set, relating the analog signal of the analyzer to the real concentration of each gas (O_2 or CO_2). However, using the *BlueVis* software with the standard "*CER/OUR/RQ* Plugin*"*, there was no option to include these calibration values to correct the O_2 or CO_2 measurements. Then, the on-line *RQ* calculation was carried out using the digital signal of the gas analyzer, which only considered the manufacturer's initial calibration averaged by a one-point calibration performed once at the beginning of this set of experiments and only with air (20.97% O_2 and 0.04% CO_2). Even so, the calibration information was included in the *RQ* calculation *a posteriori*, during data treatment and processing, as done with humidity measurement. A flowchart of the manual-heuristic control algorithm is shown in **Figure 26**.

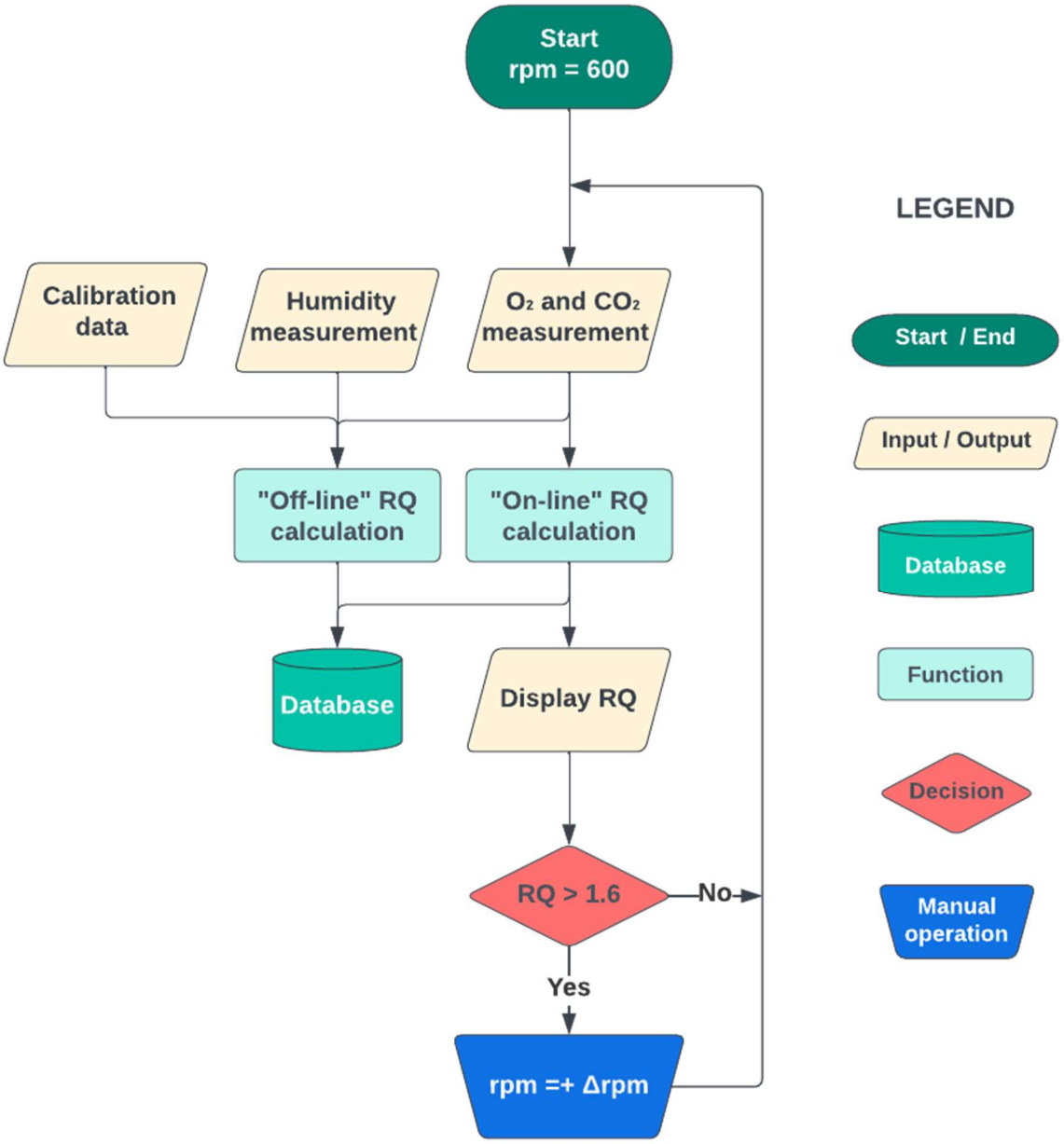


Figure 26. Flowchart of the manual-heuristic control strategy. The initial agitation rate was set at 600 rpm. *RQ* was calculated (on-line *RQ* calculation) using O_2 and CO_2 measurements, and every 15/30/60 min (variable time) the agitation rate was increased by 25/50/100 rpm (variable Δrpm) only if *RQ* > 1.6. This control strategy was manually implemented step by step, using the *BlueVis* software to calculate the *RQ* and the *BiostatB* interface to increase the agitation rate.

After finishing the fermentation, an off-line *RQ* calculation was carried out taking into account humidity measurements and gas analyzer's calibration data.

Because of these two main limitations, there were some significant differences between the *RQ* calculated on-line, which was used to monitor and control the hypoxic conditions during the experiment, and the final *RQ* calculated *a posteriori* and included in both all graphs, tables, and, in general, all results presented in this work. For this reason, although the "On-line *RQ*" was always kept within the range from 1.2 to 1.6, some deviations from this range can be observed in the plotted *RQ* or "Off-line *RQ*", as well as in the final results given in this work.

The results of these experiments are already discussed in **Section 5.2.2**, where a comparison between hypoxic and normoxic fermentations is presented. However, in this current section, only hypoxic fermentations are compared from a bioprocess control perspective. Besides, in order to evaluate the reproducibility of the process, each replicate is plotted separately.

Biomass and ethanol concentrations, as well as the Crl1 titer of SCC hypoxic fermentations are plotted in **Figure 27A**, whereas *RQ* and agitation rate profiles are plotted in **Figure 27B**. As already mentioned, each replicate is plotted separately to evaluate the reproducibility between replicates. These results are also discussed in **Section 5.2.2**, so the discussion does not focus on them here. Therefore, what is to be noted here is the low reproducibility of the process. As pointed out in **Section 5.1.2** and **Section 5.2.3**, the hypoxic level has a significant impact on the physiology of the yeast, so significant differences in *RQ* due to a non-automated and non-accurate controller can lead to significant differences in biomass, Crl1, and ethanol production, and, ultimately, to different process efficiencies.

In general, *RQ* was kept within the desired range, with large oscillations and numerous peaks. As advanced in the previous section, during the first third of the fed-batch phase, almost no control actions were required since the dynamics of the system were slow. From a practical standpoint, they were scheduled to occur during the night hours, which allowed the researcher to avoid spending the entire fed-batch phase in the laboratory. During the intermediate third, control actions were performed more often, and in the last third the agitation was increased every hour, so at the maximum pre-defined frequency.

It is noteworthy that manual control actions were not carried out strictly according to the defined control algorithm. In some cases, control actions were performed not every hour but when they were deemed necessary, i.e., when *RQ* was higher than 1.6 even if agitation had been increased less than one hour earlier. Moreover, in some periods when there was more time availability, agitation was increased stepwise by smaller but more frequent steps, looking for finer *RQ* control. An example of this is the period from t = 10 h to t = 15 h in SCC-HPX1 (R1, black line). During this period, there seemed to be fewer oscillations in *RQ*, although the differences were not so clear. However, it could indicate what is evident from the outset, namely, that smoother and more frequent control actions lead to improved controller performance.

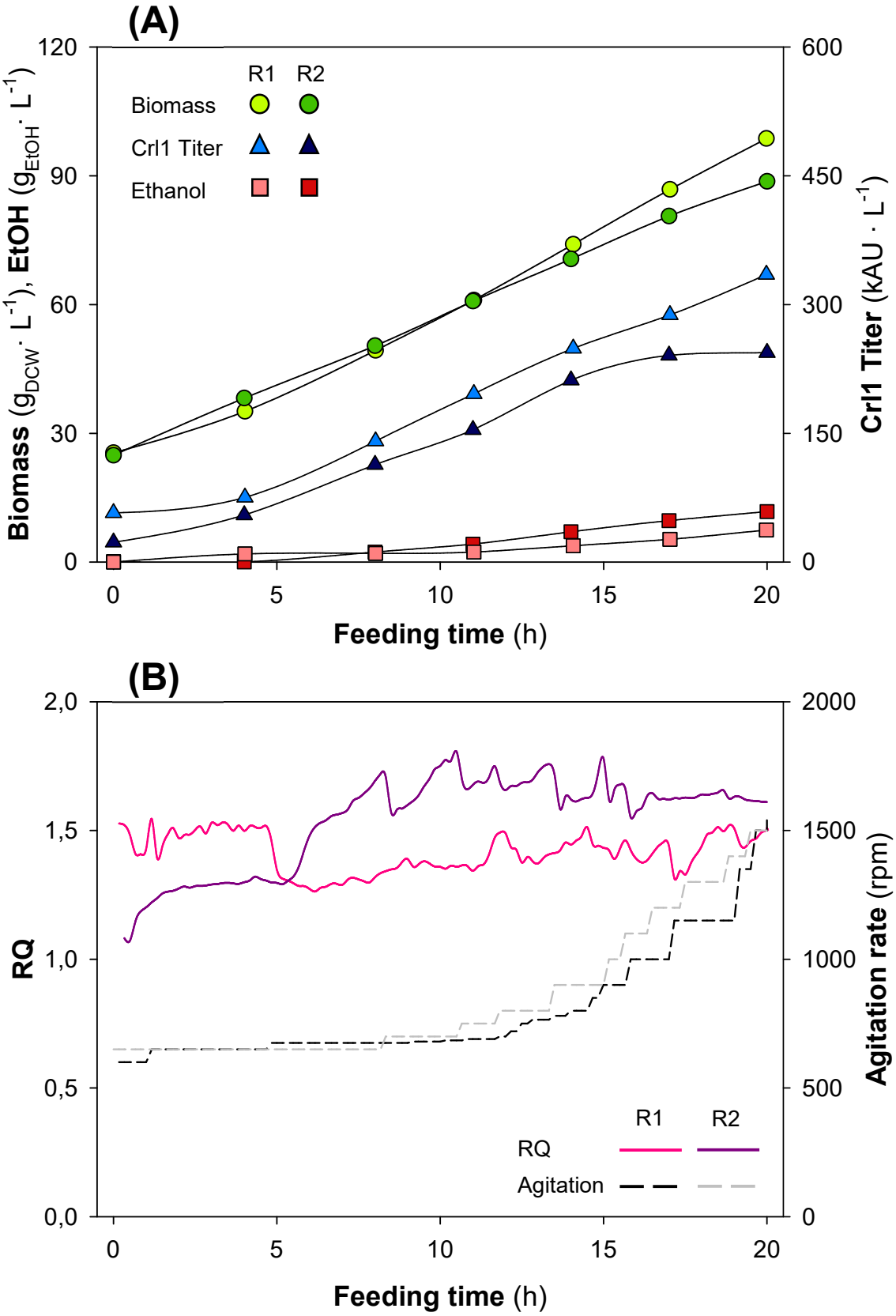


Figure 27. Primary data of the main key process parameters and *RQ* control for the two duplicates (R1 and R2) of the hypoxic fed-batch cultivations with the manual-heuristic control strategy (SCC-HPX1 and SCC-HPX2). **(A)** Biomass concentration (●, g · L^{-1}); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}). **(B)** Off-line *RQ* calculation (continuous lines); agitation rate (discontinuous lines, rpm).

Still, apart from the oscillations, whether large or small, attention must be paid to the low accuracy of the *RQ* controller, since with this strategy it becomes almost impossible to maintain a specific *RQ* set-point. Reasons for that are the two aforementioned drawbacks: the inability to include humidity measurements and calibration values into the on-line *RQ* calculation. This indicates a potential advance in the development of a more effective *RQ* control strategy.

To conclude this section, the three initial goals had been accomplished, namely testing the improvement of RPP in hypoxic conditions in fed-batch, assessing the viability of an *RQ* controller based on agitation modification, and generating enough data for the subsequent development and training of an AI model. Furthermore, the two main shortcomings of this strategy have been identified, which are the inability to include humidity measurements and the external calibration data in the *RQ* measurement, in addition to the obvious inference that the automation of the controller would improve its efficiency.

6.4 Automation of the *RQ* control: integration of external signals and development of a Boolean logic-based controller (BLC)

After having identified the main shortcomings of the manual control strategy, the next step was the implementation of an automated control strategy. In essence, the idea was to automate the same control strategy. Therefore, the whole control system had to be capable of calculating the *RQ* with high precision before deciding whether to increase agitation. In addition, this decision had to be completely automated, as automation allows to perform smoother and more frequent control actions, resulting in a finer *RQ* control without major fluctuations (Craven et al., 2014; Brignoli et al., 2020).

At this point, a software platform capable of integrating data coming from bioreactor and data coming from gas analyzers was strictly necessary. Bearing in mind the idea of a further implementation of an AI-based model, which is fed with as much data as possible, the decision was made to acquire software able to collect all data from all external devices, together with data from the bioreactor. Besides, this data had to be manipulated in order to calculate *RQ* accurately, and finally, the software had to be able to perform control actions based on a simple rule: if *RQ* > *RQ* set-point, then agitation must be increased.

Considering all of the above, the implementation of the so-called soft sensors was crucial. A soft sensor is a software application that can integrate direct measurement data, process it according to a predefined model, and perform indirect measurements. Simple/basic soft sensors to calculate *CER*, *OUR,* and *RQ* are among the most widely used in bioprocessing, although soft sensors to determine *μ* are also frequently described (Luttmann et al., 2012; Barrigón et al., 2013; Stanke & Hitzmann, 2013). Moreover, there is a lot of literature available on the use of soft sensor measurements to implement control strategies (Goodwin, 2000; Sagmeister et al., 2013; Randek & Mandenius, 2018; Beiroti et al., 2019; Rathore et al., 2021; Allampalli et al., 2022).

Among all the options evaluated, the *Eve* software from Infors (Bottmingen, Switzerland) was selected for this purpose, being able to read data not only from bioreactor and gas analyzers but also from microburette (piston pump) and

ethanol sensor. Furthermore, the implementation of soft sensors within the *Eve* interface was very easy, and the coding language was based on C++. The latter was an imperative point, as learning programming skills were beyond the scope of this thesis or the research group's objectives. Thus, a relatively well-established coding language and one that could be modified by outside parties was needed.

After configuring the software, all data coming from the bioreactor and external devices was integrated into *Eve*, therefore the software acted as a “Supervisory Control And Data Acquisition” or *SCADA* system (Ivarsson, 2017; Brunner et al., 2021). In addition, all data was available for the soft sensors and the controller implementation in the *Eve* interface (Macdonald, 2018). **Table 10** includes a list of all the available variables and implemented soft sensors.

Table 10. Set of variables coming from bioreactor and peripheral devices available in the *Eve* interface and soft sensors implemented using these variables as input data.

Variable	Origin	Description	Variable	Origin	Description
TEMP	Biostat B	Broth temperature	*Microburette Flowrate*	Microburette	Substrate flowrate
JTEMP	Biostat B	Jacket temperature	*MeOH Value**	Volatile sensor	Ethanol signal value
STIRR	Biostat B	Agitation rate	*CO_2*	Gas analyzer	%CO_2 (internal calibration)
pH	Biostat B	pH	*O_2*	Gas analyzer	%O_2 (internal calibration)
pO_2	Biostat B	Dissolved oxygen (DO)	*Humidity*	Gas analyzer	% Humidity
BASE	Biostat B	NH_4OH 15% Added	*EXT A*	Gas analyzer	CO_2 sensor signal
GF_AIR	Biostat B	Air flowrate	*EXT B*	Gas analyzer	O_2 sensor signal
GF_O2	Biostat B	Pure oxygen flowrate			

Soft sensor	Inputs	Outputs	Description
ExitCO_2	*CO_2*	*Outgas %CO_2*	Concentration of CO_2 in the outgas. Sensor signal multiplied by calibration values
ExitO_2	*O_2*	*Outgas %O_2*	Concentration of O_2 in the outgas. Sensor signal multiplied by calibration values
Volume	*Microburette flowrate*	*Volume*	Volume estimation considering substrate addition (for OUR and CER calculation)
RQ	*CO_2, O_2, Humidity, GF_AIR, GF_O2, Volume*	*CER, OUR, RQ*	CER, OUR, and RQ calculation
Ethanol	*MeOH Value**	*Ethanol concentration*	Ethanol concentration calculation using MeOH sensor with mean calibration values

Implementing soft sensors enabled the calibration values of the gas analyzers and the humidity measurements to be incorporated into *RQ* calculation to provide a more accurate value on which the control of the bioprocess is based. On the one hand, two simple soft sensors called "*ExitO₂*" and "*ExitCO₂*" were defined. As displayed in **Table 10**, these soft sensors basically convert the analog signal of the analyzer to molar fraction by using the calibration values of each analyzer. On the other hand, a soft sensor called *"RQ"* encompasses the corrected O_2 and CO_2 values and the humidity measurements to calculate *CER*, *OUR,* and *RQ*. **Equations 46** and **47** show the *CER* and *OUR* calculations taking into account the humidity in both the inlet and the outgas steams.

$$CER = \frac{F_g \cdot P}{V \cdot R \cdot T} \cdot \left(\frac{100 - O_{2_i} - CO_{2_i} - H_2O_i}{100 - O_{2_o} - CO_{2_o} - H_2O_o} \cdot CO_{2_o} - CO_{2_e}\right) \qquad \textbf{Eq. 46}$$

$$OUR = \frac{F_g \cdot P}{V \cdot R \cdot T} \cdot \left(O_{2_i} - \frac{100 - O_{2_i} - CO_{2_i} - H_2O_i}{100 - O_{2_o} - CO_{2_o} - H_2O_o} \cdot O_{2_o}\right) \qquad \textbf{Eq. 47}$$

Where *H₂Oi* and *H₂Oₒ* are the humidity in the inlet gas and in the outgas, respectively (%). All other variables have the same units as in **Equations 44** and **45**.

As shown in **Table 10**, two other soft sensors were defined: "*Ethanol concentration"* and "*Volume*". The former calculates ethanol concentration from the probe's analog signal using **Equation 7** from **Section 4.3.5**, and the latter determines the volume of culture broth based on feeding addition. According to a formal definition, except for the "*RQ"* soft sensor, all these parameters are not essentially soft sensors, since they are only simple calculations based on direct measurements, but they were considered soft sensors in the *Eve* environment.

Besides the soft sensors application, and as mentioned previously, the control strategy was built using the same principles as the manual control strategy, with three key upgrades: 1) an increase of the control actions frequency, 2) a reduction of the ∆rpm and 3) the ability to reduce the agitation rate if needed. However, in this case, the control algorithm was designed to be as much automated as possible. Thus, although using the same control law as for the manual control ("If *RQ* > *RQ* set-point, then increase the agitation"), in this case, the control law was

defined as a Boolean logic controller (BLC), having only TRUE or FALSE values. The term “Boolean logic control” has been used with similar control approaches (Premier et al., 2011). When *RQ* > 1.4, the control law was TRUE and therefore the agitation was increased.

The first improvement was the frequency of control actions, which was increased from every 1 hour to every 10 minutes. This frequency was defined based on the response time of the system, from the moment a modification of the agitation was made until a change in the outgas composition was observed. It was estimated empirically with a range of between 5 and 10 minutes based on the results of the manual control fermentations, to avoid the overacting of the controller (Simutis & Lübbert, 2015). Such a time response is mainly due to the fermenter’s headspace, which is not occupied by liquid and is therefore full of gas, and it is, by far, larger than the time response of the gas analyzer (Christensen et al., 1995). Changing the agitation immediately changes the intrinsic O_2 transfer rate, as well as for CO_2 and other gases, but these differences cannot be observed until the headspace gas has been displaced. Furthermore, the volume of the headspace is variable throughout fermentation since the culture broth increases over time. To be conservative, a 10-minute time window was then left between each increase in agitation, so the frequency of control actions was defined at 10 min.

In consistency with the increase in frequency, the second improvement of the automated controller was the gain of rpm. This was drastically reduced with the aim of avoiding big changes in *OTR* and so, avoiding high peaks and fluctuations in *RQ*. However, as advanced in **Section 6.1**, the increase in agitation throughout the feeding phase should not be constant. Ideally, the rpm steps at the beginning should be lower than at the end. Thus, to meet bioprocess requirements, the increase in agitation, defined as Δrpm, was variable over time. In general terms but not in a totally strict manner, this trend was followed: during approximately the first third of the feeding phase, Δrpm = 10. Then, during the middle third, Δrpm = 15, and finally, during the last third of the fermentation, Δrpm = 25. This change in Δrpm was done manually, so the automated control strategy was not strictly fully automated, but it only required two small modifications for the whole feeding phase. As it represented an already significant improvement in terms of

automation over the previous control strategy, it was considered acceptable. Furthermore, these moments were used by the researcher as checkpoints in order to ensure that the controller was functioning correctly.

The last improvement of the automated control strategy was the addition of new functionality allowing for a decrease in agitation if *RQ* was too low. For an *RQ* < 1.3, the controller applied an agitation step of Δrpm = -10. This upgrade gave more precision and accuracy to the controller at the initial third of the fed-batch, because the arbitrarily selected initial agitation rate of 600 rpm could not be appropriate. Apart from these initial hours, it shouldn't be necessary, since as pointed out previously, the inertia of the system is that *RQ* continuously increases as long as substrate addition continues to increase.

Considering that this control strategy was intended to be as much automated as possible, it was decided to avoid enriching the inlet gas with pure oxygen manually at the last stages of the feeding phase, whenever *RQ* could be maintained at the set-point without reaching the physical limit of agitation rate. A flowchart of the Boolean logic control algorithm can be observed in **Figure 28**.

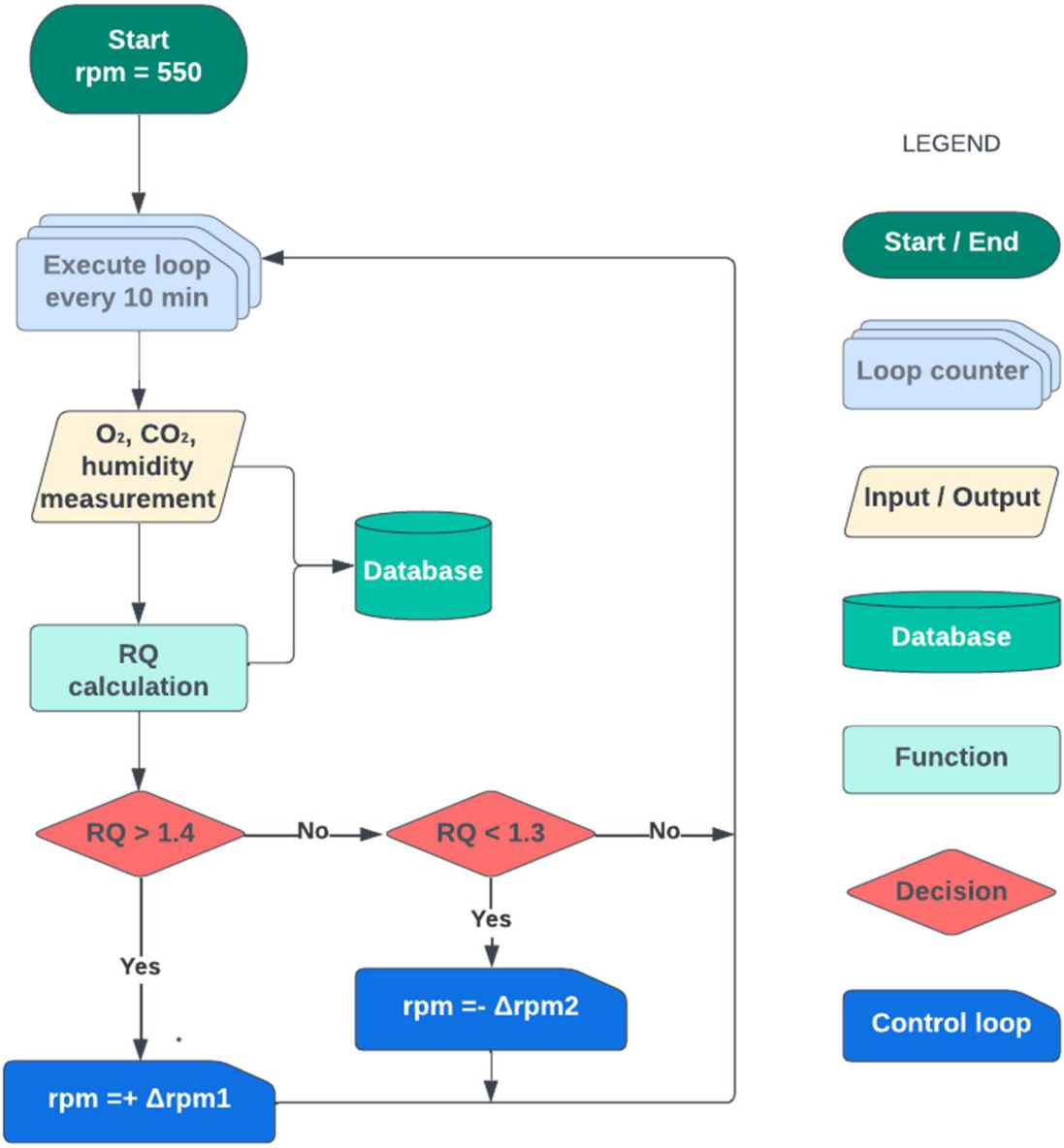


Figure 28. Flowchart of the Boolean-logic control strategy. The initial agitation rate was set at 600 rpm, and every 10 min *RQ* was calculated applying **Equations 46** and **47** and using O_2 and CO_2 measurements (including the calibration data of the analyzers and the humidity in the inlet and offgas steams), and *RQ* data was saved. Then, if *RQ* > 1.4, the agitation rate was increased by 10/15/25 rpm (variable Δrpm1) and if *RQ* < 1.3, it was decreased by 10 rpm (Δrpm2). This control strategy was automated using the *Eve* software, being necessary to manually modify only Δrpm1.

The previously mentioned optimal ∆rpm for each subphase of the feeding phase were found empirically, so several tests were run to fine-tune the controller. These fermentations were performed as part of the thesis of the PhD candidate Albert Sales, and also in the framework of the “CPV of the future” international project. These fermentations consisted of a study of the effect of μ in hypoxic conditions. Specifically, four different μ were tested (μ = 0.03 h^{-1}, μ = 0.065 h^{-1}, μ = 0.10 h^{-1}, μ = 0.12 h^{-1}) with SCC to evaluate the effect of specific growth rate on Crl1 production and other physiologic parameters, maintaining an *RQ* set-point of 1.4 using this control strategy implemented in *Eve*. The results are not included in this work, but they can be found in the literature (Sales et al., 2023).

Once the controller had been adjusted, a hypoxic fermentation was performed in two replicates to test the efficiency of the controller. The fermentation strategy was exactly the same as in the previous section, with the obvious exception of the *RQ* control strategy. Only SCC was tested with this strategy since no differences concerning *RQ* control were expected between clones.

The results of these fermentations are plotted in **Figure 29**. In terms of production, in both fermentations the SCC behaved equally, giving practically identical biomass, Crl1, and ethanol profiles, as shown in **Figure 29A**. However, in the last hours of the second replicate R2 (grey lines) biomass growth and Crl1 production were reduced. At this point, the specific growth rate could not be maintained in its set-point value, so $\mu < 0.10\ h^{-1}$, and glucose accumulation up to 7.5 $g \cdot L^{-1}$ was observed, indicating that at this high biomass concentration and glucose addition, the process becomes slightly unstable. The reason for this reduction of the specific growth rate was unidentified, although the problem was not related to the controller but to the bioprocess. Additionally, higher ethanol concentrations were observed in this second replicate.

Regarding controller efficiency, as can be observed in **Figure 29B**, there is a really huge improvement in terms of *RQ* control performance compared with the manual control results from the previous section. In both replicates *RQ* was kept at 1.4 with only small deviations of approximately ± 0.05, except for the first 1 – 2 hours, when the optimal agitation rate had to be found by the controller, having an initial input of 600 rpm. However, since the controller had the ability to increase

and reduce the agitation rate, this optimal agitation was rapidly found. Interestingly, the *RQ* values deviate from the set point during the last 30 - 45 minutes of the replica R1, because with a ∆rpm = 25 rpm and a control action frequency of 10 minutes, the controller was unable to maintain this *RQ* set point. However, the deviation was only about 0.1 *RQ* units.

As commented in **Section 6.1**, the frequency and magnitude of the control actions vary throughout the feeding phase. However, in this case, it can be seen more clearly. During the initial third of the feeding phase (from approximately t = 0 to t = 7 h), almost no control actions were carried out, and those few that were performed were tiny increases or decreases of agitation to adjust the correct *RQ* value. During the intermediate third of the fed-batch, the controller increased agitation constantly, but still not acting every 10 minutes. From approximately t = 14 h to the end, coinciding with the last third of the feeding phase, ∆rpm was set to 25 rpm, and the controller increased the agitation almost every 10 minutes, i.e., with the maximum frequency, especially in the last 2 hours. The ∆rpm should have been increased even more in the last 30 to 45 minutes to maintain a steady *RQ* set-point, or alternatively, the inlet gas should have been enriched with pure oxygen. However, because of the small deviation from the set-point, it was decided not to make any further changes to this parameter.

Based on these results, the main conclusion is that ∆rpm should be adaptive to the process requirements. Ideally, this adaptation should be automated. Hence, the power of the controller could be varied throughout the process to achieve the desired outcome.

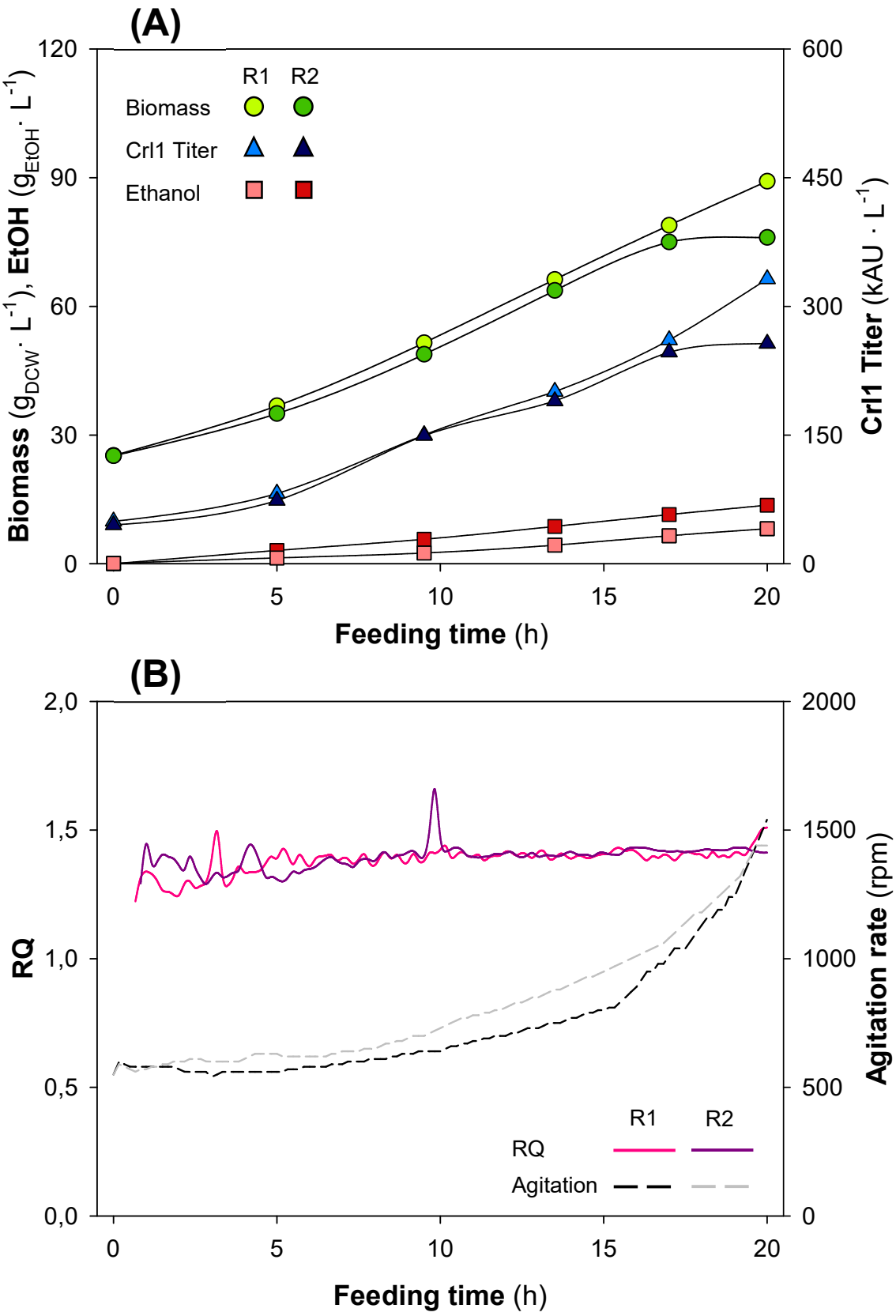


Figure 29. Primary data of the main key process parameters and *RQ* control for the two duplicates (R1 and R2) of the hypoxic fed-batch cultivations with the Boolean-logic control strategy (HPX-BLC1 and HPX-BLC2). **(A)** Biomass concentration (●, g · L^{-1}); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}). **(B)** On-line *RQ* calculation (continuous lines); agitation rate (discontinuous lines, rpm).

The results of these experiments demonstrate that the process is highly reproducible in terms of *RQ* control. Moreover, discarding the last hours of R2, results also indicate that Crl1 production is more reproducible with the automated controller rather than the manual controller, and suggest that the lack of *RQ* oscillations provides a better environment for cell growth and RPP (Xia et al., 2015).

Despite not being shown in the graphs, arabitol and succinic acid were detected in the culture broth, although final concentrations were smaller than those obtained with the manual control fermentations. As a result, this reaffirms that the absence of *RQ* oscillations and, therefore, the lack of metabolic changes ease biomass growth and partially prevent undesirable byproduct formation. Actually, *RQ* oscillations generate instability in terms of glucose metabolization, since the fluxes through the oxidative and the fermentative pathways are highly dependent on the level of oxygen limitation (Baumann et al., 2010). *RQ* oscillations can thus be considered as harmful as fluctuations in substrate concentration, which have been reported to affect production yields with *P. pastoris* as well as with another cell factories (Neubauer et al., 1995; Junne et al., 2011; Lorantfy et al., 2013; Wang et al., 2020).

To sum up, a very efficient *RQ* controller has been satisfactorily implemented, leading to a far more reproducible process and slightly more efficient in terms of production and substrate utilization. Based on the same heuristic rules as the manual control strategy, the automation has led to a finer and more accurate *RQ* control and a much less labor-consuming productive bioprocess. Nonetheless, there is still some room for improvement, namely developing a strategy with a variable Δrpm (scheduled or even adaptive), so that it can be adjusted according to process time. The last implemented control strategy, presented in the next section, is focused on that point.

6.5 Implementing an Adaptive-Proportional Controller (APC) using Artificial Intelligence algorithms: a foundation for Industry 4.0

In the framework of the “CPV of the Future” project, an *RQ* controller based on AI algorithms was developed and implemented as the last step in making the control strategy more complex. As mentioned before, the development of this controller was carried out in collaboration with AIZON. Upon developing the AI Model outlined below, AIZON's team extensively trained it using the data obtained from all hypoxic fermentation experiments described in **Sections 6.3** and **6.4** and also those conducted under hypoxic conditions at different μ (Sales et al., 2023). Also, AIZON's team developed the code necessary to apply an *RQ* control in the *Eve* environment based on the data derived from the AI model.

As stated in **Section 6.4**, having a variable and adaptive ∆rpm was the next and last step in the development of an optimal control strategy to maintain the desired hypoxic conditions through the feeding phase. It was previously determined that this parameter should increase throughout the process, in accordance with the exponential growth of biomass. However, with the automated control strategy presented in **Section 6.4**, this increase was implemented following the researcher's discretion and expertise. Strictly speaking, for a fully automated control strategy, the controller should have the ability to define the correct ∆rpm at any time during the process. Thus, for its relative simplicity and adjustability, an adaptive proportional (Adaptive-P) controller was considered the best option. Due to the relatively slow dynamics of the process compared with the time response of the system, proportional-integrative (PI), proportional-derivative (PD), or proportional-integrative-derivative (PID) controllers were discarded as the first option. Furthermore, the decision criterion was to combine a controller that was as simple as possible with as much data as possible in order to train the AI model used by this controller.

The term proportional (P) control is used in bioprocess control science to describe a controller that, in general terms, measures a variable and takes action based on the difference between the measurement of this variable and the variable's

set-point, multiplying this difference or “error (ε)” by a “Proportional Gain (K_P)”. It can also be more complex and include more terms like weights (Vilanova, 2008). P is the proportional part of a generic PID control (Gnoth et al., 2008; Baeza, 2017). However, depending on the application, not all elements of the PID control are necessary, as they can cause the control to increase in complexity unnecessarily (Rathore et al., 2021). On the other hand, an adaptive controller is a controller that is capable of changing its behavior in response to changes in the process dynamics or disturbances in the system (Bastin & Dochain, 1990; Ferreira et al., 2012).

Obviously, the controlled variable was *RQ*, which had a fixed set-point of 1.4. *RQ* was calculated in the *Eve* environment using the same “*RQ*” soft sensor described in **Section 6.4** and shown in **Table 10**, using “*ExitCO₂*” and “*ExitO₂*” as inputs to the function. By doing so, the calibration values of the gas analyzers could be incorporated into the *RQ* measurement, as explained in **Section 6.4**.

However, unlike the previous control strategy, in this case, the controller was not implemented only through the *Eve* software. Data collected from the bioreactor and external devices, including raw data that resulted from measurements and calculations done by the soft sensors, was continuously uploaded to a cloud platform from AIZON using a “REST API” call function from *Eve*. By using this function, the software sent a data package once per minute to a selected IP address. This made this data available for the AIZON software stored on the cloud platform, which then performed the estimation of the "Adaptive Proportional (P) Gain". The methodology used to calculate this parameter is described in **Section 6.5.1**.

In parallel, AIZON's team developed a program that was essentially an *RQ* controller, implemented in *Phyton*, in order to continuously receive data from the bioreactor, including the *RQ* value from the “*RQ soft sensor*”, and calculate the function output (∆rpm) based on **Equation 48**. This equation basically multiplied the Adaptive-P Gain or “K_P” by the “ε” value that corresponds to the error between the calculated *RQ* and the *RQ* set-point. This function was applied to obtain a ∆rpm, used in **Equation 49** to predict the optimal agitation rate. Then, the updated agitation rate was sent back to the *Eve* software, which applied it in the

bioreactor, also utilizing the REST API call function. Using this procedure, the *Eve* program did not make any decisions regarding the bioprocess control. Rather, all task was given to the controller working in the cloud, and *Eve* only acted as a command executor. The controller function is defined by **Equations 48** and **49**.

$$\Delta rpm = \left(RQ - RQ_{\text{set-point}}\right) \cdot "Adaptive\ P\ Gain" \quad \textbf{Eq. 48}$$

$$rpm = rpm + \Delta rpm \quad \textbf{Eq. 49}$$

Where *Adaptive P Gain* is the proportional parameter of the control function (rpm), and the difference $RQ - RQ_{set\text{-}point}$ is equal to ε. The *Adaptive P Gain* was continuously calculated every minute to predict the ideal agitation required to maintain the desired *RQ*. Furthermore, **Equation 48** enables the controller to lower the agitation, since if $RQ < RQ_{set\text{-}point}$ then the ∆rpm becomes negative, and the controller decreases the agitation following **Equation 49**.

Although a new agitation speed was predicted every minute, the actual modification of the parameter was only carried out with a maximum frequency of 7 minutes in order to account for the system time response, as detailed in **Section 6.4**. Specifically, the controller ran **Equation 48** once a minute to calculate the necessary ∆rpm, but it only applied this modification once every 7 minutes following **Equation 48**. To achieve this, a parameter was defined to account for this performance time window. The inclusion of this parameter also permitted this performance time window to be modified quickly and easily if necessary.

Besides, if the difference between the current and the predicted agitation was lower than 3 rpm (∆rpm < 3), this modification was not implemented, and the controller repeated this procedure the following minute. Therefore, the frequency of control actions was variable, with intervals from at least 7 minutes to intervals of approximately half an hour, depending on the requirements of the process. The result is that both adaptive ∆rpm and frequency in control actions could be applied, thereby overcoming the main weaknesses defined in **Section 6.4** for the automated control strategy. A simplified flowchart of the AI-aided Adaptive-P control algorithm is shown in **Figure 30**.

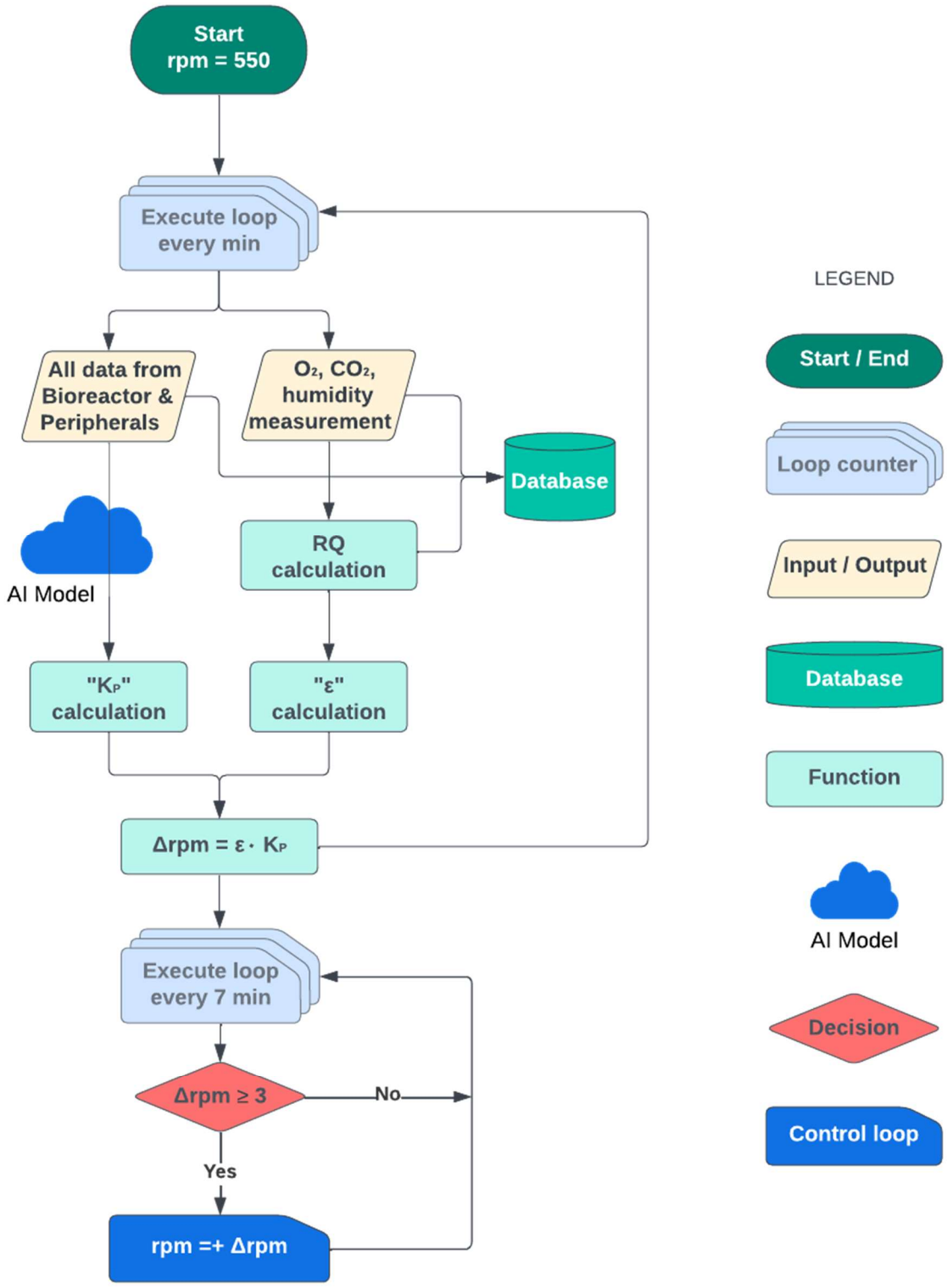


Figure 30. Flowchart of the AI-aided adaptive proportional control strategy. The initial agitation rate was set at 550 rpm, and every min *RQ* was calculated using **Equations 46** and **47**, and *RQ* data was saved together with all data coming from the bioreactor and the peripheral devices. Next, ε (error between *RQ* and *RQ* set-point) was calculated. In parallel, K_P (proportional gain) was calculated through an AI model (random forest regression algorithm) using all data coming from the bioreactor and peripherals. Finally, Δrpm (positive or negative) was calculated using ε and K_P and finally agitation was modified every 7 minutes only if Δrpm ≥ 3, following **Equations 48** and **49**.

To evaluate the performance of the controller toward a disturbance in the system, and based on the facts exposed in **Section 6.1**, the composition of the inlet gas was modified at t = 13.33 h of the feeding phase, coinciding with the beginning of the last third of the fermentation. At that point, the air flowrate was automatically modified from 2.0 L · min^{-1} to 1.8 L · min^{-1}, and a flowrate of 0.2 L · min^{-1} of pure oxygen was added, giving an oxygen concentration of 28.87% in the inlet gas. A reduction in the agitation rate was therefore expected at this point, although the nature of the response was still unknown. This modification was applied by adding a command in the *RQ* controller code implemented in *Phyton.*

When the controller had been implemented and tested with simulated hypoxic conditions, two hypoxic fermentations were conducted to check if the AI-based controller could provide better results than the Boolean logic automated controller. As was done with the previous control strategy, both fed-batches were carried out with the SCC, with a pre-programmed exponential profile of glucose feeding with μ = 0.10 h^{-1} and an *RQ* set-point = 1.4.

Figure 31 illustrates the results of these two AI-controlled fermentations. Regarding biomass, Crl1 and ethanol production of both replicates, shown in **Figure 31A**, they were very similar to those obtained with the BLC. They were especially comparable between them, highlighting the remarkable reproducibility achieved with this controller. However, biomass production slightly decreased in the last hours compared to the previous set of fermentations from **Section 6.4** with the BLC. This could be explained by the increase of *RQ*, which leads to a reduction in $Y_{X/S}$ and thereby a lower biomass generation since higher ethanol profiles were also observed, as shown in figure X in the following section.

In **Figure 31B**, where the agitation and *RQ* profiles are plotted, a very precise *RQ* control can be observed, which is very close to the set-point and exhibits very small oscillations. This strategy performed at least as good as the BLC in terms of controller efficiency. In contrast, the time commitment necessary to keep these fermentations running was significantly decreased compared to the previous strategy and extremely minor compared to the manual control strategy. Since the ∆rpm was automatically predicted every minute and agitation was automatically modified every 7 (or more) minutes, it was not necessary to assess whether the

∆rpm was appropriate at each point in the process, and the researcher only had to be present in the laboratory for sampling and supervising. Due to the fact that the process was controlled without having to be present, there were fewer samples and longer intervals between them.

Meaningfully, the *RQ* control did not perform perfectly during the last third of each fermentation, coinciding with the moment when the gas composition was modified. This is not surprising, since the Artificial Intelligence model for predicting the "Adaptive-P Gain" was not trained with data in which the gas composition was altered. This is further discussed in **Section 6.5.1**. Even so, the *RQ* deviation was very slight, only about 0.1 *RQ* units above the set-point. On the other hand, the controller's response toward a disturbance was really fast, being able to make the *RQ* return to the set-point in less than 1 hour for R2 and less than 30 min for R1.

As already commented in **Section 6.4**, the high *RQ* stability led to a more reproducible Crl1 production as well as a lower byproducts generation, including ethanol as well as arabitol and succinate, compared to the manual control strategy described in **Section 6.3**, emphasizing that oscillations in metabolic conditions adversely affect the process.

In summary, a fully automated and precise *RQ* controller based on AI algorithms has been satisfactorily implemented. High controller efficiency has been achieved, comparable to the BLC in terms of bioprocess productivity and yields, but with improved adaptability and automation. Although it did not perform perfectly when the inlet gas composition was modified, it probably could have been avoided if the model was trained with a wider range of data. Additionally, the performance toward a disturbance was extremely fast and precise, proving AI is a tool that can be used to control bioprocesses efficiently.

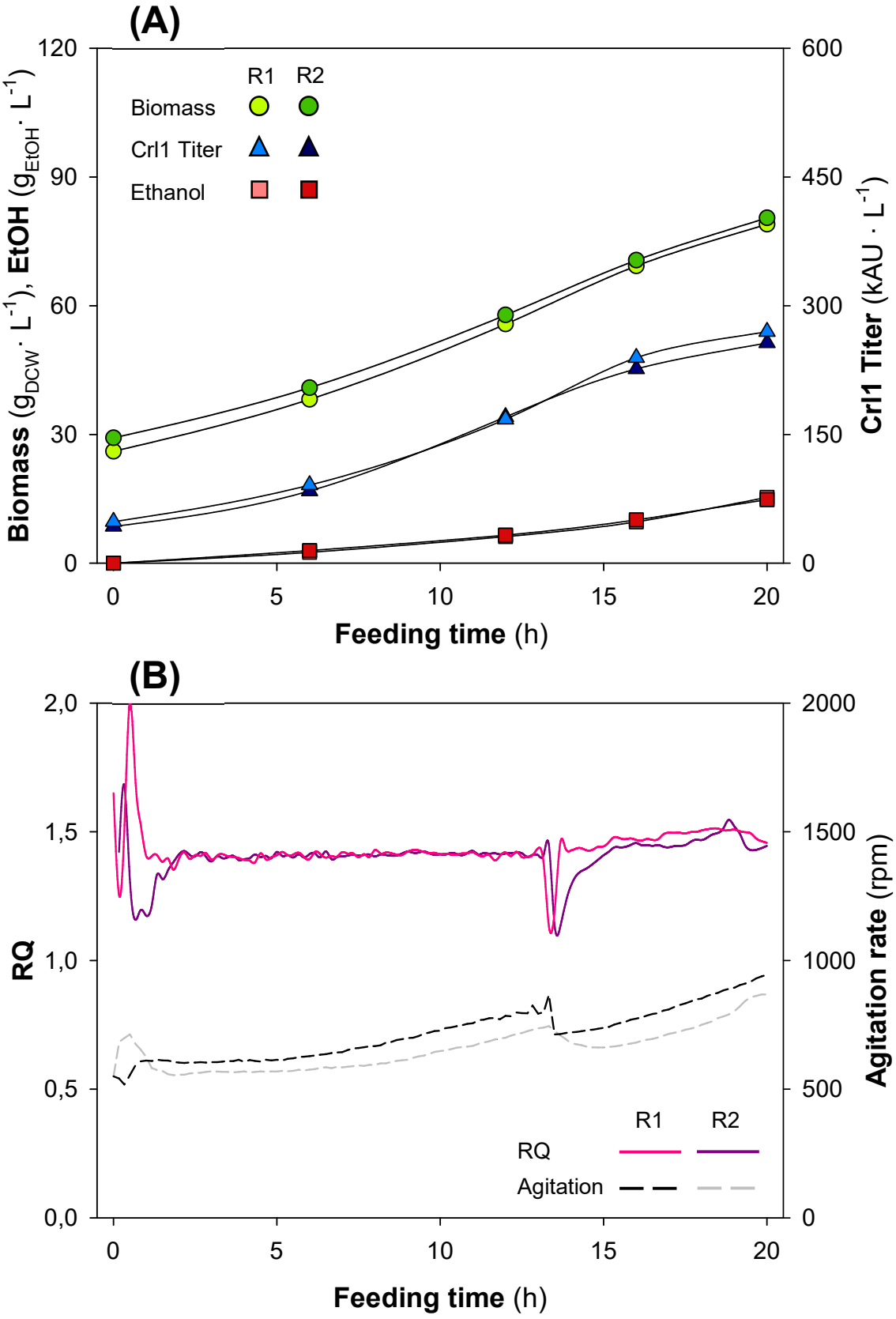


Figure 31. Primary data of the main key process parameters and *RQ* control for the two duplicates (R1 and R2) of the hypoxic fed-batch cultivations with the AI-aided adaptive proportional control strategy (HPX-APC1 and HPX-APC2). **(A)** Biomass concentration (●, g · L^{-1}); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}). **(B)** On-line *RQ* calculation (continuous lines); agitation rate (discontinuous lines, rpm).

6.5.1 Application of Artificial Intelligence algorithms to adaptive bioprocess control

The calculation of the "Adaptive-P Gain" parameter or "K_P" was achieved through the application of AI algorithms. Specifically, a random forest algorithm was applied. The random forest algorithm is a machine learning technique that seeks to make accurate predictions by leveraging the power of multiple decision trees. The algorithm randomly selects features from a large dataset and uses these features to create multiple decision trees, each of which making its own prediction. During the training process, each decision tree iteratively selects features and divides the data into smaller groups until it has made a final prediction. The algorithm continues to create decision trees until it has a set number of trees or until it reaches a certain level of accuracy. The mechanism of operation of the algorithm is schematically shown in **Figure 32**.

One of the main characteristics of the random forest algorithm is its ability to handle a large number of variables, even if many of these variables are not relevant for an accurate prediction. By creating multiple decision trees and combining their predictions, the algorithm can effectively filter out irrelevant features and focus on the most important ones. It is a powerful technique for making data-driven decisions since it is able to decide whether a variable has a relevant impact or not on the estimation of a certain parameter, and it has already been used in bioprocess engineering applications (Mowbray et al., 2021; Singh & Singhal, 2022).

In this case, 100 decision trees were implemented, each one trained with a different subset of the data. Thus, when making a prediction, the algorithm passes the input through each decision tree, and the results from all decision trees are combined to form a final prediction. In this case, when the algorithm was trying to predict a numerical value, such as the optimal agitation rate, the average of the predicted values from all decision trees was taken (Breiman, 2001). It is noteworthy that all sets of data with which the model was trained included the results from the fermentations performed with the MHC from **Section 6.3** and with the BLC, both the included in **Section 6.4** and some of the previously mentioned fermentations carried out with different μ detailed in the bibliography,

where no pure oxygen was used (Sales et al., 2023). Thus, it makes sense that during the last third of the cultivations controlled with the AI-APC, when inlet air was enriched with pure oxygen, the controller did not control the *RQ* as fine as during the rest of the feeding phase.

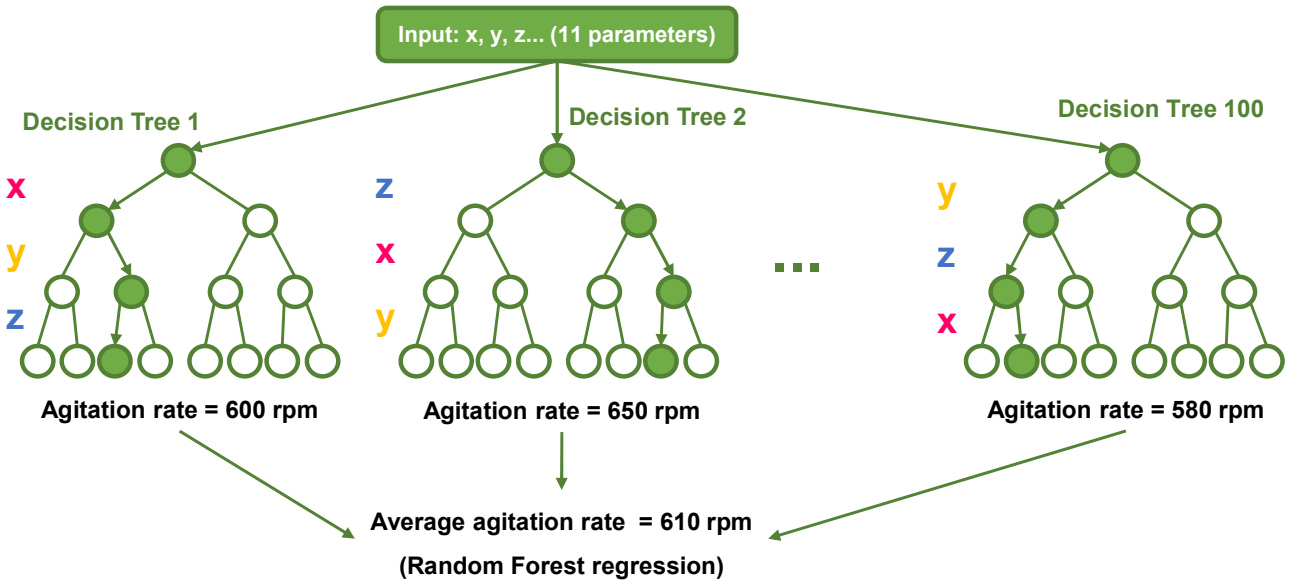


Figure 32. Scheme of a random forest regressor algorithm. 100 decision trees, each one trained with a different subset of data, predict the necessary agitation rate to maintain *RQ* = 1.4, and the final value is obtained by averaging all predictions. In addition, each decision tree may assign different priorities to each parameter or process variable, depending on the subset of data used in its training process.

x, y and z represent different variables or process parameters.

6.6 Performance comparison between the implemented control strategies

To evaluate the proper functioning of each *RQ* controller, a comparison of the three implemented control strategies in terms of bioprocess performance and reproducibility is conducted in this section.

Table 11 shows the numerical values of key process parameters obtained in each hypoxic fermentation. Additionally, the mean value and the standard deviation (SD) are displayed. In parallel, **Figure 33** shows the profiles of the total production of biomass and ethanol as well as Crl1 titers for those fermentations. Similar values for key process parameters were observed for all strategies. Furthermore, biomass and Crl1 production profiles were very similar, reaching a final biomass production of 300 g of DCW and a final Crl1 titer of approximately 1000 kAU. However, ethanol production differed between strategies, with the last one exhibiting higher levels of production. Additionally, an undesired reduction of around 7% of μ was detected with the third strategy. As mentioned before, this could be explained by the increase in *RQ* observed in the last third of the feeding phase, which led to an inversely proportional decrease of $Y_{X/S}$. As mentioned, ethanol production was also higher, as observed in **Figure 33** and reflected in the q_{EtOH} values in **Table 11**.

Table 11. Value of key process parameters obtained in fed-batch hypoxic fermentations with the three control strategies tested. Specific growth rate, μ (h^{-1}); specific substrate consumption rate, q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); biomass-to-substrate yield, $Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$); specific ethanol production rate, q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); respiratory quotient, *RQ*; specific Crl1 production rate, q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); and product-to-biomass yield, $Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$). Values in bold represent the mean and SD (±) between biological replicates (R1 and R2).

	MHC		**BLC**		**AI-APC**	
	R1	R2	R1	R2	R1	R2
μ ($g_{DCW} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.105	0.096	0.103	0.097	0.095	0.091
	0.101 ± 0.006		**0.100 ± 0.004**		**0.093 ± 0.003**	
q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.21	0.22	0.22	0.21	0.22	0.21
	0.21 ± 0.01		**0.22 ± 0.01**		**0.22 ± 0.01**	
$Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$)	0.50	0.45	0.47	0.45	0.43	0.43
	0.47 ± 0.04		**0.46 ± 0.01**		**0.43 ± 0.01**	
q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.020	0.034	0.027	0.030	0.040	0.037
	0.027 ± 0.010		**0.029 ± 0.002**		**0.038 ± 0.002**	
RQ	1.41	1.53	1.38	1.38	1.40	1.44
	1.47 ± 0.09		**1.38 ± 0.01**		**1.42 ± 0.02**	
q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	348	310	335	358	317	311
	329 ± 27		**347 ± 16**		**314 ± 4**	
$Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$)	3.60	3.42	3.83	3.92	3.67	3.83
	3.51 ± 0.12		**3.88 ± 0.06**		**3.75 ± 0.12**	

It is worth noting that *RQ* values from **Table 11** are the mean *RQ* values throughout each fermentation. As a result, the mean and standard deviation values shown in bold are indicative of the variability between replicates, rather than the deviation from the set-point, which is in this case a more suitable indicator of the controller's accuracy, or the deviation from the mean *RQ* value, which is an indicator of the controller's precision.

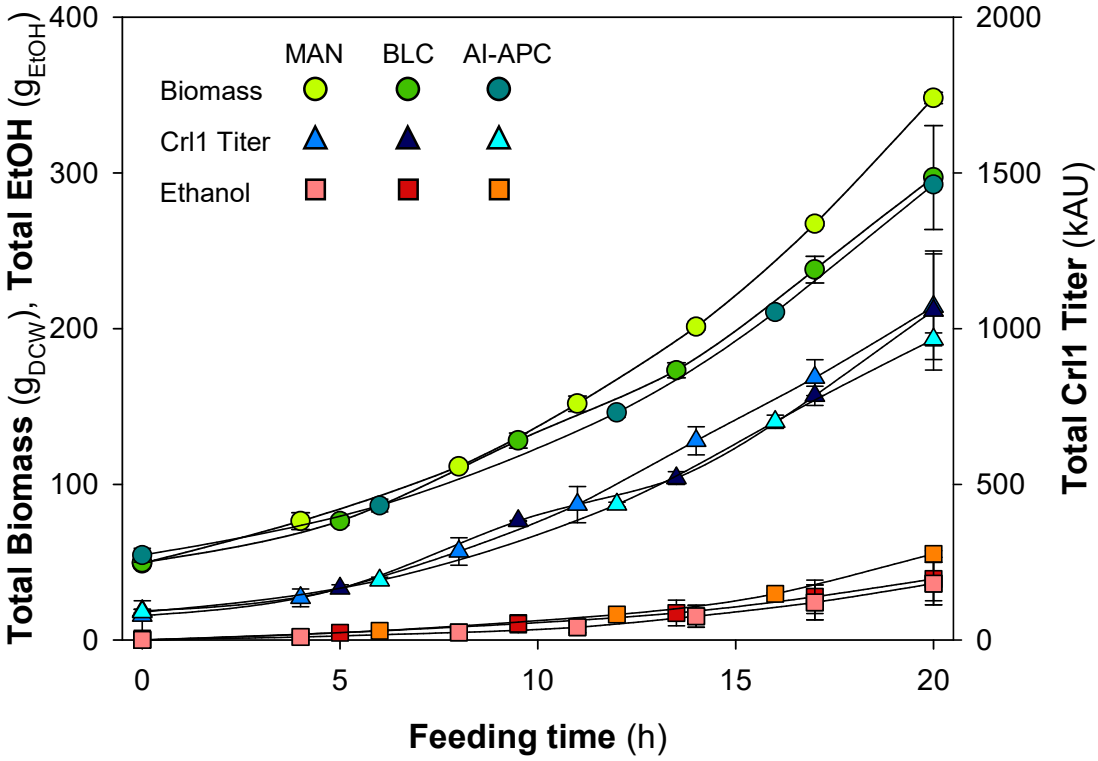


Figure 33. Comparison of primary data of the main key process parameters of the hypoxic fed-batch cultivations with the three presented control strategies (SCC-HPX1, SCC-HPX2, HPX-BLC1, HPX-BLC2, HPX-APC1, and HPX-APC2). Total biomass amount (●, g); total Crl1 Titer (▲, kAU); total ethanol amount (■, g). Error bars represent the SD of biological replicates.

Thereby, in terms of bioprocess reproducibility, the application of two more sophisticated *RQ* controllers led to a reduction in the variability between replicas, especially concerning Crl1 production, since standard deviation was smaller for the BLC and AI-APC. This is also reflected in **Figure 33** by the error bars of the BLC and AI-APC, which are almost unappreciable.

However, to assess the controller's efficiency, its autonomy, accuracy, and precision should also be considered. Regarding autonomy, the manual control strategy was obviously the least efficient. When comparing the BLC and the AI-APC, in the case of the former several ∆rpm updates were required during cultivation, whereas in the case of the latter, this change was applied automatically in accordance with the prediction of an AI model, therefore turning it into the most autonomous.

As mentioned above, the *RQ* deviation from its set-point highlights the controller's accuracy. On the other hand, the deviation from the mean *RQ* can be considered as an estimator of the controller's precision. With this aim, two independent statistical performance indicators were calculated to assess this accuracy and precision: Mean Relative Error (MRE) and Root Mean Square Deviation (RMSD), which are defined by **Equations 50** and **51**. MRE can be considered as a relative value of a mean *RQ* error through the fed-batch phases, having values between 0 (good performance) and 1 (bad performance), and it was used to assess the accuracy of the controllers. On the other hand, RMSD could be seen as a mean standard deviation between the *RQ* values throughout the fed-batch phases and the mean *RQ* value of each fermentation, shown in **Table 11**, and it was used to determine the precision of the controller.

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - y_{sp}|}{y_{sp}} \qquad \textbf{Eq. 50}$$

$$RMSD = \sqrt{\sum_{i=1}^{n}\frac{(y_i - \bar{y})^2}{n}} \qquad \textbf{Eq. 51}$$

Where y_i is the value of the variable (*RQ*) at each time-point ($_i$), y_{sp} is the *RQ* set-point, since the scope is to evaluate the accuracy of the controller (MRE), and $\bar{y}$ is the mean *RQ*, for the analysis of the controller's precision (RMSD). Although they can give redundant information, the analysis of two different statistics can help avoiding data artifacts.

Using all datapoints from each pair of replicates, the values shown in **Table 12** were obtained.

Table 12. Value of Mean Relative Error (MRE) and Root Mean Square Deviation (RMSD) for each hypoxic fed-batch fermentation with the three control strategies tested.

		MHC		BLC		AI-APC	
		R1	R2	R1	R2	R1	R2
MRE	**Accuracy** (how close is *RQ* to the set-point)	0.049	0.141	0.045	0.046	0.035	0.040
RMSD	**Precision** (how close are *RQ* measurements to each other)	0.083	0.194	0.108	0.112	0.083	0.105

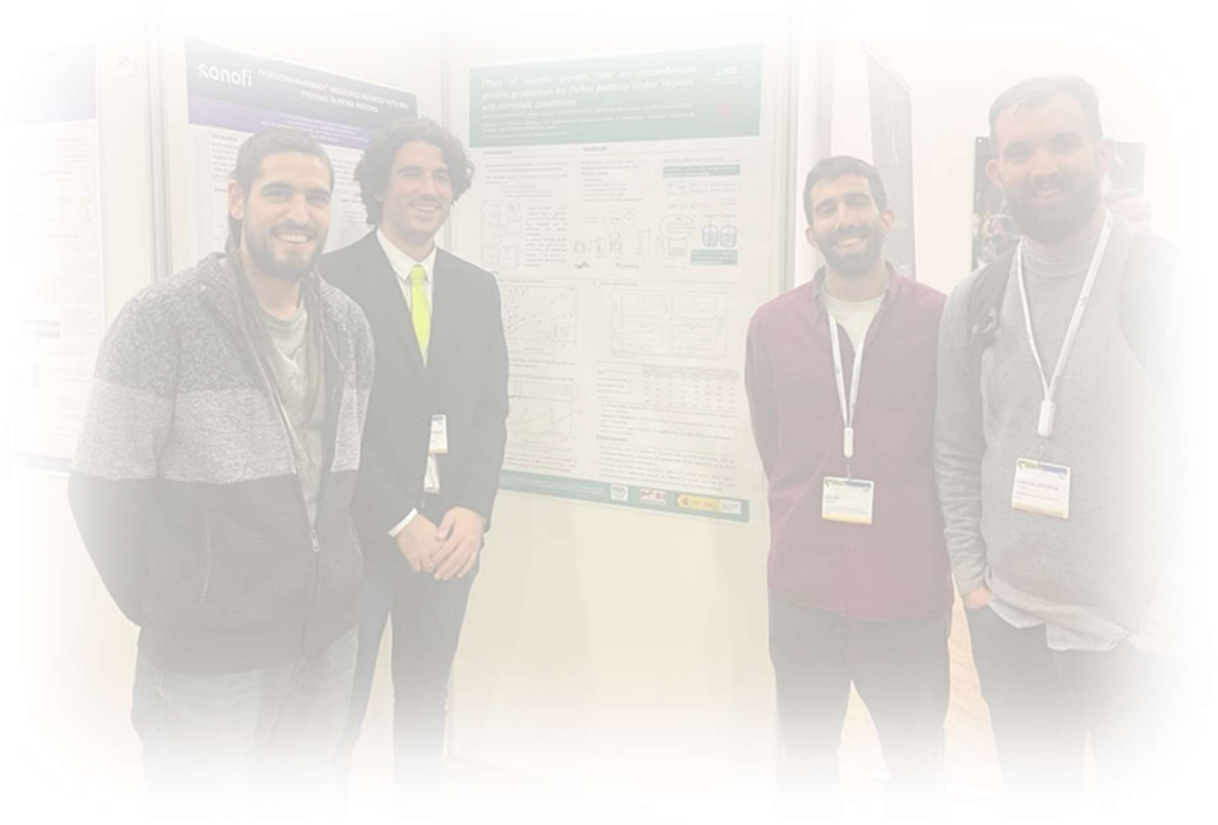
Based on the analysis of these statistics, in terms of accuracy, the two last strategies outperformed the manual control strategy, having lower MRE values. These numbers indicate that the mean error was about 5% (R1) and 14% (R2) with the MHC, about 4.5% with the BLC, and lower than 4% with the AI-APC strategy. On the other hand, in terms of precision, the same trend was observed when comparing RMSD values. As mentioned previously, RMSD could be considered as a mean standard deviation from the mean *RQ* value, becoming lower as the complexity of the controller increases. Additionally, these results also demonstrated the higher reproducibility achieved with the automated control strategies: the Boolean-logic and the AI Adaptive-P controllers.

In addition, as mentioned before during the last third of the AI-APC fermentations the *RQ* was constantly above the set-point, suggesting that the *RQ* control during the rest of the fed-batch was more effective. This point added to the fact that in those fermentations there was a disturbance in the inlet gas composition, which had a direct impact on *RQ* and therefore also on MRE and RMSD, turns the third strategy into the most efficient and accurate.

Ultimately, automation of control actions, their increase in frequency, and their variation in intensity result in a more efficient control of bioprocesses. Additionally, the use of AI algorithms to implement adaptive-P controllers has proved successful in the field of bioprocess engineering.

7. RESULTS III

Evaluation of hypoxic strategy on the production of different recombinant lipases



An adapted version of Chapter 7 has been submitted to be published as a research article in *Microbial Biotechnology:*

Sales, A., Gasset, A., Requena-Moreno, G., Valero, F., Montesinos-Seguí, J. L., Garcia-Ortega, X. (2023). Synergic kinetic and physiological stress control to improve bioprocess efficiency of *Komagataella phaffii* recombinant protein production bioprocesses. *Microbial Biotechnology*

7.1 Implementation of the automated *RQ* control system

Once the automated Boolean logic *RQ* control algorithm had been established, the hypoxic strategy was tested with two different *P. pastoris* strains expressing different proteins to check if the positive effect over recombinant production of Crl1 was also observed with other lipases. As stated in the methodology section, the selected strains were, firstly, an X-33 *P. pastoris* strain producing the recombinant *R. oryzae* lipase (proRol) under the regulation of the P_{GAP}, and secondly, a commercial BG11 *P. pastoris* strain producing *C. antarctica* lipase B (CalB), also under the regulation of the GAP promoter.

It is noteworthy that the mature sequence of the *R. oryzae* lipase (recombinant-Rol or rRol) had never been expressed under the regulation of P_{GAP}, since it presented UPR-related problems which prevented its constitutive expression and secretion. In another work conducted in the research group shortly before this thesis, it was demonstrated that the inclusion of a pro-sequence facilitated the intracellular expression and folding of this lipase, thus allowing for a constitutive expression (López-Fernández et al., 2019, 2021). On the other hand, CalB is a very industrially attractive enzyme that has been extensively expressed in the *Pichia* cell factory (Eom et al., 2013; Looser et al., 2017). However, the use of the commercial *P. pastoris* strain BG11 from BioGrammatics (Carlsbad, CA, USA) presented a challenge, since it was not so well characterized in the research group. These strains were already used in previous studies and more details about their genetic construction can be found in the literature (Nieto-Taype, 2020; Garrigós-Martínez et al., 2021; López-Fernández et al., 2021a).

In contrast to the hypoxic fermentations conducted with SCC and MCC strains producing Crl1 and reported in **Section 5.2**, in this case, the *RQ* control was achieved through the Boolean-logic *RQ* control strategy (BLC) instead of a manual actuation over this manipulated variable, as explained in **Chapter 6**. As also detailed before, this improvement made the process more reliable and reproducible, but it also presented a major drawback: the automatic *RQ* control algorithm had a single manipulated variable (agitation rate) and did not act over the air and pure oxygen flowrates. Therefore, the increment in oxygen demand along the process had to be satisfied only by acting on the agitation rate, rather

than enriching the inlet gas with pure oxygen, since the controller could not act on the gas flowrates. As explained in **Section 4.2.3.2**, with a μ = 0.10 h^{-1}, the oxygen demand was not met only by incrementing the agitation rate but also by incrementing the oxygen proportion in the inlet gas. This was done in all the normoxic fermentations but also in one of the MCC hypoxic fermentations. So, a lower μ should be selected to avoid the use of pure oxygen, especially in the hypoxic fermentations, and also to avoid reaching very high agitation rates. However, in the early stages of the feeding phase, the oxygen demand is much lower than in the final stages, and there is a physical minimum agitation rate from which it should not be lowered since complete homogeneity must be guaranteed to avoid any oxygen or substrate gradients in the culture broth (Alford, 2006; Ask & Stocks, 2022). Therefore, the only way to reduce *OTR* would be to aerate the fermenter with a mix of air with nitrogen. Thus, considering the results from the aforementioned study of the production kinetics (μ - q_P relationship), obtained in the research group but not included in this work, a μ = 0.065 h^{-1} was selected and implemented for these four fermentations (two with each strain, in normoxic and hypoxic conditions). With this μ and the equipment's oxygen transfer capacity, hypoxic conditions could be applied without the use of either pure oxygen or pure nitrogen, and *RQ* could be controlled by solely modifying the agitation rate.

Furthermore, as explained in more detail in **Sections 4.2.3.2** and **5.4.1**, for these fermentations the feeding flowrate profiles were equal for the normoxic and the hypoxic fermentations, whereas the glucose concentrations in the feeding tank were different for each condition. This was done to counteract the reduction in the $Y_{X/S}$ caused by hypoxic conditions and with the goal of achieving the same final volume, as done in the pilot-scale fermentations from **Section 5.4.1**. Nevertheless, in this case, it proved to be more effective. Therefore, biomass concentration profiles for all these normoxic and hypoxic fermentations were practically identical, reaching a final biomass dry cell weight of around 90 - 95 g · L^{-1}. Thus, as pointed out in **Section 5.2.2**, this strategy results in a more comparable and reproducible biomass profiles, rather than modifying the feeding flowrates, which led to comparable biomass productions but slightly different final volumes and so, different biomass concentrations, as shown in **Figures 13** and **14** from **Section 5.2.2**.

7.2 proRol production under normoxic and hypoxic conditions

The time evolution of proRol titer and biomass and ethanol concentrations of the normoxic (dark symbols) and hypoxic (light symbols) are plotted in **Figure 34**. These fermentations were carried out with the X-33 strain.

Biomass profiles of both fermentations are identical, reaching a final biomass concentration of 92 g · L^{-1} in both conditions, thereby confirming that the modification of the feeding substrate concentration (S_0 from **Equation 4**) is a very beneficial strategy to achieve equal biomass productions with same volume profiles in different conditions regarding substrate utilization or even different substrates. Specifically, glucose concentrations were 325 g · L^{-1} and 427 g · L^{-1} for normoxic and hypoxic fermentations, respectively.

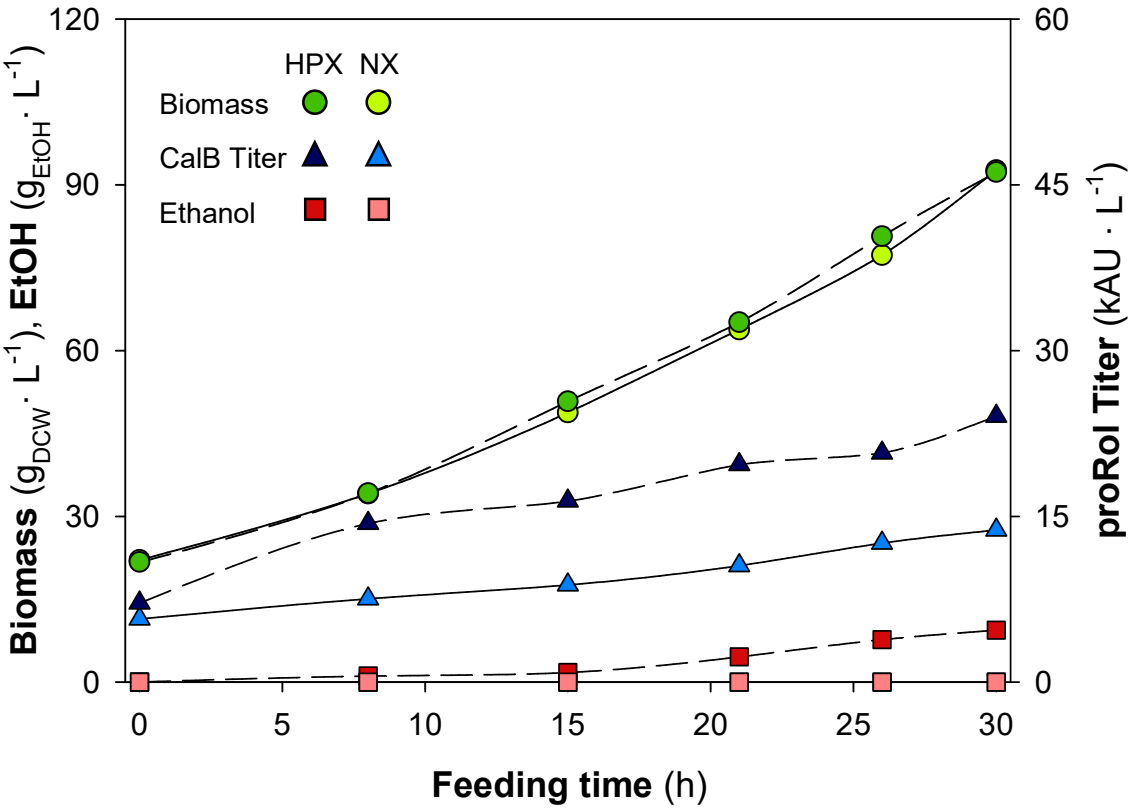


Figure 34. Primary data of the main key process parameters for the fed-batch cultivations grown at hypoxic (dark symbols) and normoxic (light symbols) conditions with the proRol-producing strain (proRol-HPX and proRol-NX). Biomass concentration (●, g · L^{-1}); proRol Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}).

Concerning lipase production, proRol titers at the end of both batch phases were practically identical, thus confirming the reliability of the glycerol-based batch phase. On the other hand, at the end of the feeding phase, it reached a final proRol titer of 24.1 kAU · L^{-1} in hypoxic conditions, clearly higher than the 13.8 kAU · L^{-1} obtained in normoxic conditions. This difference is even more pronounced if the lipase activity produced during the batch phase (around 6 – 7 kAU · L^{-1}) is not considered, obtaining a proRol production through the feeding phase 2.0-fold higher in hypoxic conditions, slightly lower than the increase in Crl1 titer previously observed in **Section 5.2.2**.

Regarding by-products generation, up to 9.4 g · L^{-1} of ethanol and 3.3 g · L^{-1} of arabitol were detected in the culture broth at the end of the feeding phase in hypoxic conditions, confirming the shift from respirative to respiro-fermentative metabolism in these conditions (Baumann et al., 2008, 2010; Carnicer et al., 2009; Adelantado et al., 2017). As reported in **Section 5.2.2**, similar concentrations of ethanol and arabitol were observed with the same *P. pastoris* strain X-33 expressing Crl1 in hypoxic conditions. Neither ethanol nor arabitol were detected in normoxic fermentations.

The proRol production in normoxic conditions was higher but in the same unit range as the one observed with the same strain in previous works performed in the research group, where a glucose-based fed-batch with a μ = 0.045 h^{-1} was implemented (≈ 6 kAU · L^{-1}) (López-Fernández et al., 2021a), indicating a slightly higher production with higher μ and a potential better fermentation performance in the current work. Furthermore, proRol titer was higher than those observed historically in the research group, where Rol was expressed with different strains, using the AOX1 and FLD1 promoters (P_{AOX1} and P_{FLD1}) and applying several optimized strategies such as methanol limited and non-limited fed-batch cultures (MLFB and MNLFB, respectively), DO-stat with different dissolved oxygen set-points, etc. (Cos et al., 2005a, 2005b; Barrigón et al., 2013; Ponte et al., 2016). However, the effect of the pro-sequence of proRol possibly plays an important role in these production differences. A comparison of the proRol and rRol titers obtained in different works is displayed in **Table 13**.

Due to the highly dependent nature of enzymatic activity on assay conditions, it is difficult to compare this process with other RoI-producing bioprocesses developed in other laboratories. So, to avoid misinterpretations, it has only been compared with works conducted in the same research group.

Additionally, the proRol enzymatic activity measured using the methodology described in this work (**Section 4.3.7**) must be correlated with proRol and Rol activity measured considering the methodology used in those other works. Basically, in those other works proRol and Rol activity was measured using a commercial lipase analytic kit from Roche (Basel, Switzerland), which was not based on p-NPB hydrolysis. As explained in more detail in the bibliography, **Equation 52** provides this correlation (López-Fernández, 2022).

$$Act_{(pNPB)} = 0.0178 \cdot Act_{(Roche)} - 0.1101 \qquad \textbf{Eq. 52}$$

Where both $Act_{(pNPB)}$ and $Act_{(Roche)}$ are expressed in kAU · L^{-1}.

Table 13. rRol and proRol production with different bioprocess approaches.

Strain / Protein	Promoter	Feeding strategy	μ (h^{1})	Titer ($kAU \cdot L^{-1}$)	q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	Reference
X-33 / proRol	GAP	Glucose-Limited Fed-Batch (Hypoxia)	0.065	24.1	14.1	Current work (proRol-HPX)
X-33 / proRol	GAP	Glucose-Limited Fed-Batch (Normoxia)	0.065	13.8	8.81	Current work (proRol-NX)
X-33 / proRol	GAP	Glucose-Limited Fed-Batch	0.045	5.95	8.42	López-Fernández et al., 2021 (a)
X-33 / proRol	AOX1	Methanol-Limited Fed-Batch	0.038	2.51	1.11	López-Fernández et al., 2019
X-33 / rRol	AOX1	Methanol-Limited Fed-Batch	0.043	0.37	0.21	López-Fernández et al., 2019
X-33 / proRol	AOX1	Methanol Non-Limited Fed-Batch (set-point: 3 $g_{MetOH} \cdot L^{-1}$)	0.065	6.26	5.37	López-Fernández et al., 2019
X-33 / rRol	AOX1	Methanol Non-Limited Fed-Batch (set-point: 3 $g_{MetOH} \cdot L^{-1)}$	0.046	4.87	5.62	López-Fernández et al., 2019
X-33 / rRol	AOX1	Methanol Non-Limited Fed-Batch (set-point: 3 $g_{MetOH} \cdot L^{-1}$ and 25% DO)	0.034	6.44	4.45	Ponte et al., 2016
X-33 / rRol	AOX1	Methanol Non-Limited Fed-Batch (set-point: 10 $g_{MetOH} \cdot L^{-1}$)	0.025	5.12	3.01	Barrigón et al., 2013
X-33 / rRol	AOX1	Methanol-Limited Fed-Batch	0.015	2.29	0.71	Barrigón et al., 2013
X-33 (MutS) / rRol	AOX1	Methanol Non-Limited Fed-Batch (set-point: 1 - 2 $g_{MetOH} \cdot L^{-1}$)	0.005	3.54	1.37	Cos et al., 2005 (a)
X-33 (Mut+) / rRol	AOX1	Methanol Non-Limited Fed-Batch (set-point: 1 - 2 $g_{MetOH} \cdot L^{-1}$)	0.036	2.56	2.20	Cos et al., 2005 (a)
X-33 / rRol	FLD1	Sorbitol-Limited Fed-Batch	0.010	2.92	0.92	Cos et al., 2005 (b)

Concerning *RQ* control, as can be observed in **Figure 35**, in the hypoxic fermentation the value of *RQ* was controlled between 1.25 – 1.4, close to the setpoint defined at RQ = 1.4 but slightly lower. A peak can be observed at t = 5 h, corresponding to the moment when the broth level surpassed the upper impeller and consequently the *OTR* decreased. At this point, the controller began to raise the agitation by steps of 10 RPM every 10 minutes, and the *RQ* recovered its setpoint value approximately 1.5 hours later, approximately. After that moment, agitation increased slowly and exponentially, according to the addition of glucose since an exponential substrate addition must be accompanied by an exponential increase in oxygen transfer capacity to maintain a constant *RQ* (Knoll et al., 2007; Garcia-Ortega et al., 2017). On the other hand, concerning normoxic fermentation (continuous lines), *RQ* presented values around 1.1, already observed in **Sections 5.1** and **5.2** with the Crl1 SCC and MCC and also in other studies where a glucose feeding strategy was implemented (Nieto-Taype et al., 2020a; Tomàs-Gamisans et al., 2020; Zavec et al., 2020). The agitation increased exponentially to maintain a DO setpoint of 30% of air saturation, although in this case, the agitation rate was higher than in the hypoxic fermentation because completely aerobic conditions were requested. A sudden decrease of the stirring rate, indicated by a grey arrow, can be observed at t = 26 h, corresponding to a manual enrichment with pure oxygen in the inlet gas to avoid reaching too high agitation rates.

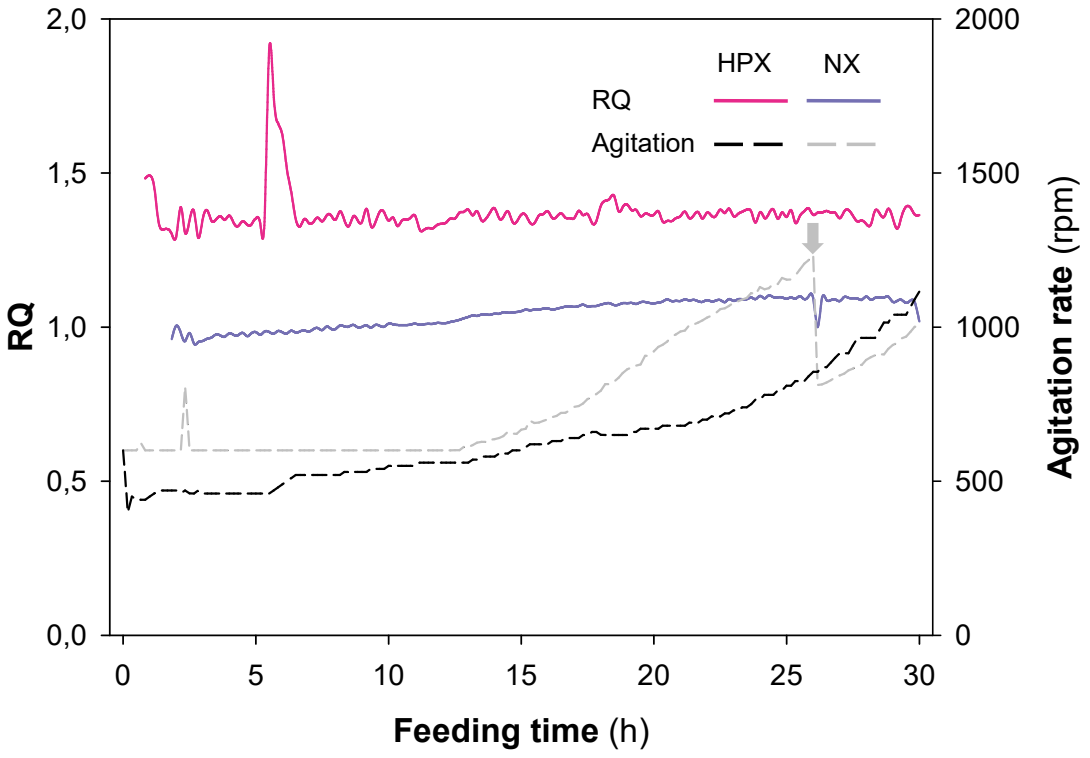


Figure 35. On-line calculation of *RQ* and agitation rate for the fed-batch cultivations grown at hypoxic and normoxic conditions with the proRol-producing strain (proRol-HPX and proRol-NX). Off-line *RQ* (continuous lines); agitation rate (discontinuous lines, rpm).

A gray arrow indicates the moment from which the inlet air was enriched with pure O_2.

In **Table 15** from **Section 7.4**, where a proper comparison between different lipases produced under normoxic and hypoxic conditions is conducted, the mean values of related key process parameters (μ, q_S, $Y_{X/S}$, q_{EtOH}, *RQ,* and, q_P) of the normoxic and hypoxic fermentations of proRol- and CalB-producing strains are presented, as well as oxygen limitation fold-change values for q_P and $Y_{X/S}$.

It can be stated that for this *P. pastoris* X-33 strain producing proRol, *RQ* has been satisfactorily controlled around the setpoint. Furthermore, through the oxygen limitation strategy, an improvement of 2.0-fold in q_P has been achieved. In addition, when implementing hypoxic conditions, a reduction of 0.81-fold in $Y_{X/S}$ is observed, comparable to that obtained in **Section 5.2.2** and in other studies where equivalent hypoxic conditions were implemented (Baumann et al., 2008; Adelantado et al., 2017; Garcia-Ortega et al., 2017).

7.3 CalB production under normoxic and hypoxic conditions

Analogously to **Figure 34**, in **Figure 36** the concentrations of biomass, ethanol, and CalB titer through the feeding phase of the normoxic and hypoxic fermentations with the *P. pastoris* BG11 strain producing recombinant CalB under the P_{GAP} regulation are plotted.

Cell density was comparable in both fermentations, with a slight halt in biomass growth being detectable in the last hours of the normoxic fermentation. Because of this, the final biomass dry cell weight at the end of the feeding phase was higher in hypoxia (97 $g_{DCW} \cdot L^{-1}$) than in normoxia (90 $g_{DCW} \cdot L^{-1}$). Except for the last hours, the biomass concentration was equal for both conditions.

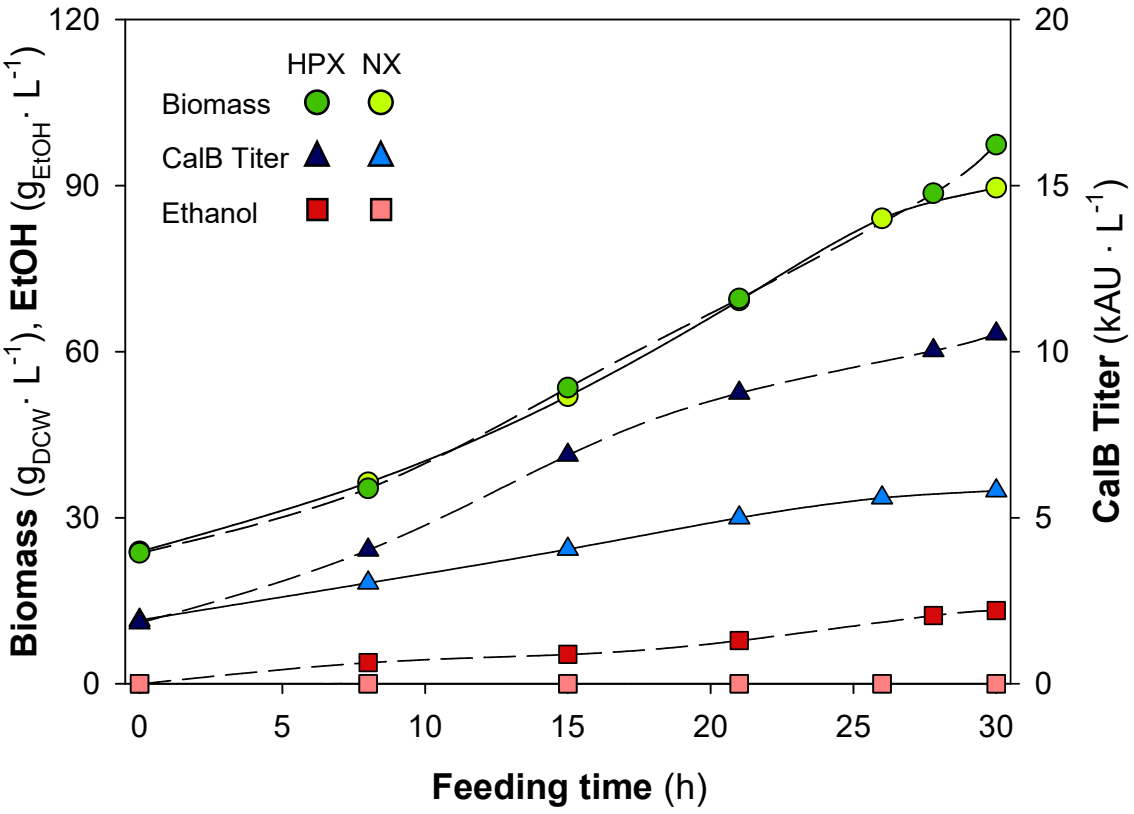


Figure 36. Primary data of the main key process parameters for the fed-batch cultivations grown at hypoxic (dark symbols) and normoxic (light symbols) conditions with the CalB-producing strain (CalB-HPX and CalB-NX). Biomass concentration (●, $g \cdot L^{-1}$); CalB Titer (▲, $kAU \cdot L^{-1}$); EtOH, Ethanol concentration (■, $g \cdot L^{-1}$).

In terms of production, as observed with the proRol fermentations, the enzymatic activity at the end of the batch phase was equal for both fermentations, giving robustness to these results. Considering the feeding phase, the CalB titer was almost duplicated in hypoxic conditions, reaching a value of 10.5 kAU · L^{-1} compared to the 5.8 kAU · L^{-1} obtained in normoxia.

The q_P value for CalB in normoxic conditions can be compared with those obtained in parallel works using the same strain but with glycerol as a carbon source. In those works, an operational range of μ from 0.025 h^{-1} to 0.15 h^{-1} was tested in fed-batch mode, although no fermentation at μ = 0.065 h^{-1} was conducted, thus preventing a direct comparison between those works and the results presented in this thesis. However, the μ of both normoxic and hypoxic fermentations was within the range of the μ tested in those parallel works, and q_P in normoxic conditions was also within the q_P range obtained there (Bernat-Camps et al., 2023). Besides, an interpolation of q_P at μ = 0.065 h^{-1} can be carried out, giving a theoretical value of $q_P \approx 4.70$ AU · g_{DCW}^{-1} · h^{-1}, which is very close to the q_P obtained in this work (q_P = 4.15 AU · g_{DCW}^{-1} · h^{-1}), providing robustness to these results.

On the other hand, q_P in hypoxic conditions was notably higher, even higher than the q_P obtained in normoxia with the highest μ = 0.15 h^{-1} (Bernat-Camps et al., 2023). A numeric comparison is given in **Table 14**. In addition, comparing normoxic results with chemostat glycerol-based cultivations with the same strain, higher titers and q_P were obtained in normoxic fed-batch (Garrigós-Martínez et al., 2021). The same effect has already been observed for Crl1 in **Chapter 5** and was also expected for CalB, since the q_P values are generally higher in fed-batch compared to chemostat, although in some cases this relationship is not fulfilled (Garrigós-Martínez et al., 2019; Nieto-Taype et al., 2020a, 2020b; Gasset et al., 2022).

Moreover, as shown in **Table 14**, q_P values in hypoxic conditions are in the same order of magnitude as those obtained following different optimized fed-batch strategies with different strains expressing CalB under the regulation of different novel methanol-free promoters such as P_{DF}, P_{UPP}, and P_{DH}, even being higher in some cases (Garrigós-Martínez et al., 2021; Bernat-Camps et al., 2023).

Table 14. CalB production with different bioprocess approaches.

*Note that Titer and q_P shown in (Garrigós et al., 2021) have been multiplied by a factor of 2.48 for a molar extinction coefficient (ε) correction.

Strain / Protein	Promoter	Feeding strategy	μ (h^{-1})	Titer ($kAU \cdot L^{-1}$)	q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	Reference
BG11 / CalB	GAP	Glucose-Limited Fed-Batch (Hypoxia)	0.071	10.5	10.8	Current work (CalB-HPX)
BG11 / CalB	GAP	Glucose-Limited Fed-Batch (Normoxia)	0.065	5.81	4.15	Current work (CalB-NX)
BG11 / CalB	GAP	Glycerol-Limited Fed-Batch	0.025	6.69	2.45	Bernat et al., 2023
BG11 / CalB	GAP	Glycerol-Limited Fed-Batch	0.15	4.44	9.44	Bernat et al., 2023
BG11 / CalB	GAP	Glycerol Chemostat	0.046	2.93	3.05	Garrigós et al., 2021
BG11 / CalB	GAP	Glycerol Chemostat	0.10	2.51	5.93	Garrigós et al., 2021
BG11 / CalB	DH	Glycerol-Limited Fed-Batch	0.025	9.82	4.59	Bernat et al., 2023
BG11 / CalB	DF	Glycerol-Limited Fed-Batch	0.087	27.7	32.5	Garrigóset al., 2021
BG11 / CalB	DF	Glycerol-Limited Fed-Batch	0.052	33.4	29.1	Bernat et al., 2023
BG11 / CalB	UPP	Glycerol-Limited Fed-Batch	0.084	26.3	24.7	Garrigós et al., 2021

Again, it is worth mentioning that no comparisons have been made with CalB-producing bioprocesses developed in other laboratories, as CalB enzymatic activity is highly dependent on temperature and pH, and different assay conditions are often used in different laboratory settings (Eom et al., 2013; Looser et al., 2017).

Strikingly, in this case with the BG11 strain, ethanol concentrations detected in hypoxic conditions were significantly higher than those observed in other hypoxic fermentations with the X-33 strain, regardless of neither the protein expressed nor the growth rate. In all the lab-scale X-33 hypoxic fermentations the final ethanol concentration was within the range between 9 – 10 g · L^{-1}, whereas in this case, the final ethanol concentration was about 13 g · L^{-1}. In addition, no arabitol accumulation was detected in this fermentation, indicating a possible divergence in metabolic readjustment under hypoxic conditions due to the differential genetic background. However, the BG11 strain is only a Mut^S variant of BG10 and, according to the reviewed bibliography, BG10 and X-33 strains differ in only 24 point mutations and for the fact that the BG10 strain has its killer plasmids removed, which theoretically do not encode for any known catabolic enzyme (Vogl & Glieder, 2013; Sturmberger et al., 2016). Killer plasmids are DNA fragments typically encoding for toxins to enable the cells to kill other yeasts, and they have been found in several yeasts' genera (Gunge et al., 1981; Hayman & Bolen, 1991).

Regarding *RQ* control, in **Figure 37** the on-line value of the respiratory quotient can be observed (continuous lines). In normoxic conditions, *RQ* was around 1 – 1.1, equivalent to the other normoxic fermentations, while in hypoxia *RQ* was adequately controlled within the range of 1.25 – 1.4 except for the first 3 h when manually adjusting the initial rpm value. The agitation rate of the hypoxic fermentation presented an exponential time profile in contrast to the relatively constant profile observed in the normoxic fermentation. Nonetheless, it must be noted that the normoxic fermentation was performed in a different BiostatB fermenter, with equal configuration but in which the DO control cascade was still not properly configured at that moment, so it was decided to keep a constant agitation and manually increase it by a step of 100 rpm at t = 23.5 h and t = 29.5 h to avoid potential problems related to oxygen availability. Moreover, grey arrows

indicate when inlet gas was enriched with pure oxygen, as done in the previous case with the normoxic fermentation producing proRol. Interestingly, comparing **Figures 35** and **37**, in the case of the CalB hypoxic fermentation no major *RQ* peaks were observed since the position of the impellers was modified to avoid the detrimental effect over *OTR* previously described in **Section 7.2** when the broth level surpassed the upper impeller.

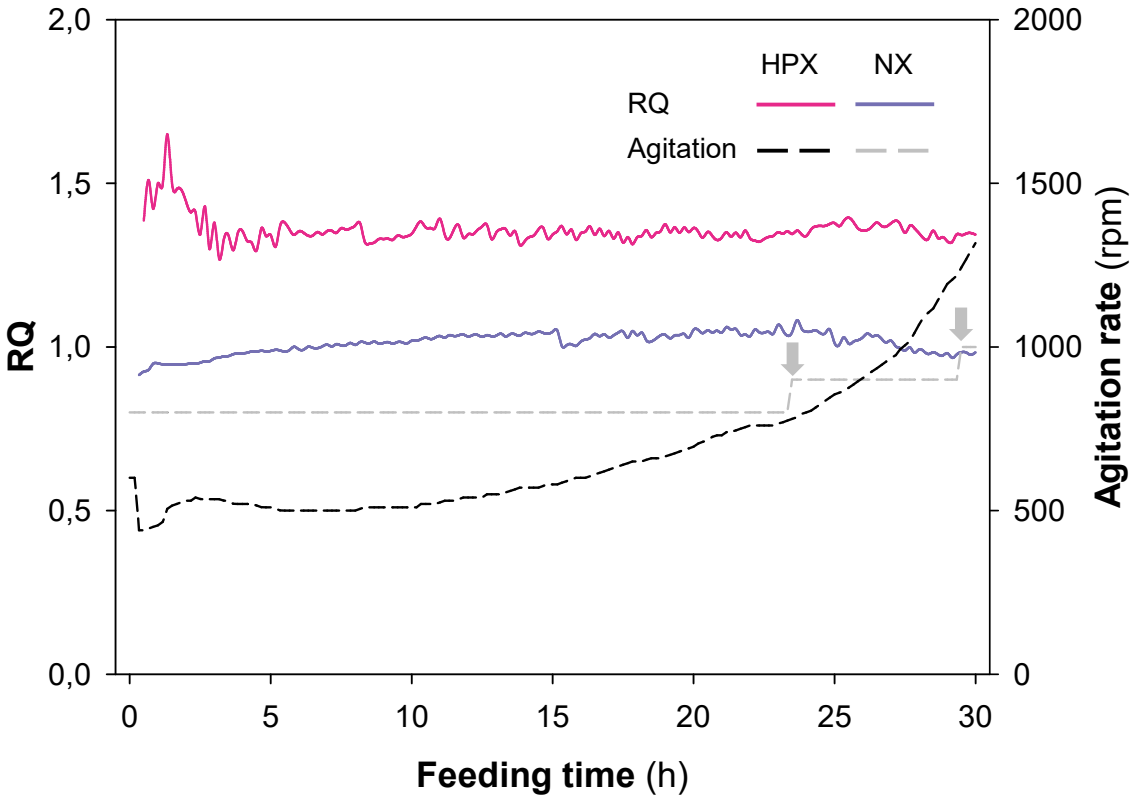


Figure 37. On-line calculation of *RQ* and agitation rate for the fed-batch cultivations grown at hypoxic and normoxic conditions with the CalB-producing strain (CalB-HPX and CalB-NX). Off-line *RQ* (continuous lines); agitation rate (discontinuous lines, rpm).

A gray arrow indicates the moment from which the inlet air was enriched with pure O_2.

It can be concluded that with this CalB-producing BG11 strain, an important increase of 2.6-fold in q_P was achieved, together with a reduction of 0.85-fold in $Y_{X/S}$ already expected. In addition, although the generation by-products were different for this strain rather than for the previously X-33 tested strains (Crl1-SCC, Crl1-MCC, proRol), the *RQ* controller performed very similarly and with comparable good results.

7.4 Comparison of lipase production with hypoxic strategy

As previously pointed out, **Table 15** shows the numeric values of key process parameters obtained for proRol- and CalB-producing strains, together with the value of FC_{OL}-$Y_{X/S}$ and FC_{OL}-q_P, or oxygen-limitation fold-change for $Y_{X/S}$ and q_P. These last two parameters have been added to make a proper comparison of the hypoxic effect, since proRol and CalB have different lipase activities with different orders of magnitude, and so the comparison of q_P becomes useless.

Additionally, results from a hypoxic and a normoxic fermentation with an X-33 strain producing Crl1 (SCC) with a μ = 0.065 h^{-1} have been also included in **Table 15**. This last set of data corresponds to two fermentations not included in this work, and they have been added only for comparison purposes, as previously done in **Section 5.4.2** when a comparison between lab scale and pilot scale fermentations with a μ = 0.065 h^{-1} was conducted. These results have been kindly provided by the PhD candidates Albert Sales and Guillermo Requena (Sales et al., 2023).

Considering results from **Table 15**, similar values for key process parameters were obtained with these two clones producing different lipases. Besides, similar $Y_{X/S}$ ratios between hypoxia and normoxia or FC_{OL}-$Y_{X/S}$ were obtained for three different strains producing three different lipases, confirming that the physiological stress the cells are subjected to is equivalent in all three cases (Crl1, proRol, and CalB).

However, the q_P ratios or FC_{OL}-q_P were notably different for the three strains, indicating that the positive effect of the use of P_{GAP} combined with oxygen limitation on recombinant production is somehow protein- or strain-dependent. Interestingly, the lower FC_{OL}-q_P was obtained for the proRol-producing strain. In agreement with these results, the overexpression of recombinant Rol has been described as having a detrimental impact on the growth and fitness of the host cells (Beer et al., 1998; Minning et al., 2001; Cámara et al., 2016, 2017).

Table 15. Values of key process parameters obtained in fed-batch fermentations in normoxic and hypoxic conditions with all the strains tested. Specific growth rate, μ (h^{-1}); specific substrate consumption rate, q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); biomass-to-substrate yield, $Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$); specific ethanol production rate, q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); respiratory quotient, *RQ*; specific Crl1 production rate, q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); product-to-biomass yield, $Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$); $Y_{X/S}$ oxygen-limitation fold-change, FC_{OL}-$Y_{X/S}$; and q_P oxygen-limitation fold-change, FC_{OL}-q_P.

	X-33 - proRol		BG11 - CalB		X-33 – Crl1	
	NX	**HPX**	**NX**	**HPX**	**NX**	**HPX**
μ (h^{-1})	0.065	0.068	0.065	0.071	0.059	0.064
q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.12	0.16	0.12	0.16	0.13	0.15
$Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$)	0.520	0.421	0.536	0.455	0.48	0.42
q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.00	0.021	0.00	0.020	0.00	0.027
RQ	1.03	1.37	1.01	1.35	1.05	1.43
q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	8.8	17.4	4.2	10.8	128	267
$Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$)	0.134	0.209	0.058	0.120	2.17	4.18
FC_{OL}-$Y_{X/S}$	0.81		0.85		0.86	
FC_{OL}-q_P	2.0		2.6		2.1	

Comparing these results with the reference strain used in this thesis (SCC), it can be stated that the hypoxic strategy had been satisfactorily applied to produce three different lipases that have an important industrial interest (Sharma et al., 2001; Jaeger & Eggert, 2002; Ken Ugo et al., 2017). However, the application of the hypoxic strategy with the BLC or with the AI-APC for the production of other proteins with more important structural and functional differences, also under the regulation of P_{GAP}, would have been very interesting. Nonetheless, as explained in **Chapter 1**, there were limited available strains expressing a recombinant

protein under P_{GAP} regulation in the research group. The construction of new strains was not the focus of this thesis, and due to time limitations, it was discarded.

Besides, although in the past a Fab-expressing X-33 strain had been tested in hypoxic conditions in the research group, it was never tested in fed-batch mode, only in chemostat operation. In those chemostat cultivations, a 3-fold increase in q_P was achieved when applying hypoxic conditions, comparable to those obtained in this work (Garcia-Ortega et al., 2017).

As future research, testing this strategy with different strains and proteins would validate it as an efficient tool to improve recombinant protein production. Furthermore, other glycolytic promoters used in recombinant protein production, such as P_{PGK1}, P_{ENO1}, and P_{TPI1}, should provide similar improvements when applying hypoxic conditions (Stadlmayr et al., 2010; Vogl & Glieder, 2013; De Macedo Robert et al., 2017). Conspicuously, the *PGK1* gene showed similar expression patterns to *TDH3*, as seen in **Section 5.3**, so the use of its promoter (P_{PGK1}) for protein production in hypoxic conditions becomes very promising.

To conclude, it can be stated that these results give evidence that bioprocess optimization is as powerful a tool as strain or promoter engineering, considering the q_P enhancements achieved in this work. With the lipases tested, q_P enhancements were from 2- to 2.6-fold.

8. CONCLUSIONS

Non-conventional production strategies are difficult to be accepted in an industrial context, especially for regulatory issues. Even so, with a few optimizations of the process control parameters, important increases in bioprocess productivities can be achieved. In combination with the wide variety of strain engineering strategies available, this can facilitate improvements of up to some orders of magnitude in the performance of bioprocesses.

In this study, an initial characterization of the hypoxic and gene dosage effects on the production of recombinant Crl1 under P_{GAP} regulation with the yeast *P. pastoris* has been satisfactorily conducted using chemostat cultures applying different oxygen concentrations in the inlet gas. As a conclusion, the optimal hypoxic conditions with regards to Crl1 production have been defined at RQ = 1.4, where a 4.3-fold and a 2.2-fold increase in q_P has been observed with SCC and MCC, respectively. Therefore, from the results obtained, it can be stated that the combination of increased gene dosage and hypoxia has resulted in a synergistic but not summatory effect since the improvement in productivity has resulted in greater with SCC than with MCC.

The definition of the optimal RQ has allowed to transfer this strategy to a dynamic fed-batch culture using the agitation rate as a manipulated variable to control the process in the optimal hypoxic degree. In this first approach, repetitive manual modifications of the agitation rate have been carried out to maintain the RQ at its optimal value. This has led to an improvement in q_P of 2-fold for SCC and 1.5-fold for MCC when applying hypoxic conditions. However, in summary, the control of RQ with manual actions has become imprecise and inaccurate.

Then, an exhaustive analysis of the transcription patterns of several genes of interest related to physiology and recombinant production has been carried out. The combination of chemostat and fed-batch cultivations with transcriptional analysis has contributed to gaining a deep understanding of the process, identifying possible bottlenecks in the transcription, translation, and secretion pathways.

The implementation of a scaling-up procedure with a scale-up factor of 15 has been satisfactorily carried out. Although on the pilot scale, the same productivity

as on the lab scale has not been attained when implementing hypoxic conditions, the results are very promising and have validated the scaling-up criterion consisting of using the *RQ* as the controlled variable. Additionally, a more precise and automated control strategy could likely have improved these results.

Following the characterization and validation of the hypoxic effect, an automated control strategy has been defined and implemented to achieve a more robust and reproducible process, since important deviations from the *RQ* set-point were observed when implementing the manual control, which led to high variability in the process, really unattractive from an industrial point of view. Thus, the signal of different measurement devices, especially the values of CO_2 and O_2 composition in the exhaust gas, have had to be integrated into a *SCADA* system for a proper on-line calculation of the *RQ*. This control strategy, named Boolean logic controller (BLC), has provided a very reliable control strategy with very reproducible results.

After doing so, an innovative control strategy based on the combination of artificial intelligence (AI) algorithms with an Adaptive-Proportional controller (AI-APC) has been developed together with AIZON company and implemented with highly satisfactory results. This represents a clear and successful example of the application of AI in the control of bioprocesses, a field in which these new technologies are still little implemented.

For the implementation of these two improved control strategies, the integration of the signals of all measurement devices in a *SCADA* system and the connection with a digital twin of the fermenter stored on the cloud have been successfully implemented. As a main outcome, the accuracy of the *RQ* control has improved around 3-fold for both control strategies with respect to the manual control strategy. On the other hand, the precision enhancement has been around 1.3-fold for the BLC and 1.5-fold for the AI-APC.

Finally, the hypoxic strategy has been also tested for the production of two different recombinant lipases of interest, *Rhizopus oryzae* lipase and *Candida antarctica* lipase, obtaining comparable improvements in the final titers and

productivities to those observed with Crl1, corroborating the enhancements in recombinant protein production achieved with hypoxic conditions.

Overall, it can be concluded that the combination of strain engineering, bioprocess optimization, omics, scale-up process, integration of several peripheral devices in a *SCADA* system with a digital twin, and the programming of two different control codes based on a very reliable *RQ* calculation from exhaust gas analysis makes this thesis a thorough and comprehensive work as well as a foundation for the application of AI-based strategies in the biotech and pharma industries.

9. ANNEX

This section presents two cultivations that are not included in the set of fermentations presented in **Chapter 5**, but which have been considered to demonstrate relevant information. These two fermentations consisted of two initial tests with the manual-heuristic *RQ* control. One of them (HPX-HIRQ) was run under more severe hypoxic conditions, with an *RQ* set-point ≈ 2.0, to test the potential inhibitory effect of high ethanol concentrations at high *RQ*. The other one (HPX-GASC) was controlled with an *RQ* set-point ≈ 1.4 but instead of acting over agitation rate, the manipulated variables were the flowrates of air, nitrogen, and oxygen, maintaining a total aeration flowrate of 2 $L \cdot min^{-1}$.

Figures 38 and **39** show the evolution of key variables (biomass and ethanol concentrations, as well as Crl1 titer) and the controlled *RQ* and the variables involved in this control (agitation rate or gas flowrate) of these two fermentations. Accordingly, **Table 16** shows the values of related key process parameters of these two fermentations, together with results from SCC-HPX1 for comparison purposes.

In general terms, in HPX-HIRQ the production of biomass was lower than in other hypoxic fermentations, and an important reduction of more than 50% in $Y_{X/S}$ was observed, compared to normoxic conditions. On the other hand, ethanol production was higher, reaching around 23 $g \cdot L^{-1}$ at the end of the fermentation. In consonance, q_{EtOH} was much higher than that observed in other hypoxic fermentations with lower *RQ*, as can be observed in **Table 16**. However, despite achieving lower biomass production, Crl1 titers were very similar to those observed in SCC-HPX1 and SCC-HPX2, suggesting that there was more Crl1 production per unit of biomass in HPX-HIRQ, also confirmed by q_P and $Y_{P/X}$ values observed in **Table 16**. Nevertheless, due to the high ethanol accumulation, the specific growth rate could not be maintained at 0.1 h^{-1}, and glucose accumulation up to 15 $g \cdot L^{-1}$ was detected in the culture broth, thus causing the process to be unstable.

In the case of HPX-GASC, the performance of the fermentation was comparable to those fermentations where *RQ* control was achieved by modifying the agitation rate (SCC-HPX1 and SCC-HPX2), in terms of biomass and ethanol generation and Crl1 production. However, *RQ* control by modifying the gas composition with

different flowrates of air, N_2, and O_2 proved to be more inefficient, since high fluctuations in *RQ* can be observed in **Figure 39**, and the mean *RQ* was slightly higher than the set-point, as shown in **Table 16**. Apart from that, the values of the key process parameters were in the same range as those achieved in SCC-HPX1 and SCC-HPX2, confirming that gas mixing could be used as a manipulated variable to control *RQ*. However, three different variables (air, N_2, and O_2 flowrates) must be manipulated, instead of one.

It must be noted that the reconciliation procedure explained in **Section 4.4.6** was not applied to the values of the key process parameters from HPX-HIRQ and HPX-GASC displayed in **Table 16**.

Table 16. Values of key process parameters obtained in hypoxic fed-batch fermentations HPX-HIRQ, HPX-GASC and SCC-HPX1. Specific growth rate, μ (h^{-1}); specific substrate consumption rate, q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$); biomass-to-substrate yield, $Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$); specific ethanol production rate, q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$); respiratory quotient, *RQ*; specific Crl1 production rate, q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$); and product-to-biomass yield, $Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$).

	HPX-HIRQ	**HPX-GASC**	**SCC-HPX1**
μ (h^{-1})	0.075	0.084	0.105
q_S ($g_S \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.32	0.21	0.21
$Y_{X/S}$ ($g_{DCW} \cdot g_S^{-1}$)	0.24	0.41	0.50
q_{EtOH} ($g_{EtOH} \cdot g_{DCW}^{-1} \cdot h^{-1}$)	0.033	0.011	0.020
RQ	2.06	1.49	1.31
q_P ($AU \cdot g_{DCW}^{-1} \cdot h^{-1}$)	405	319	348
$Y_{P/X}$ ($kAU \cdot g_{DCW}^{-1}$)	5.35	3.78	3.60

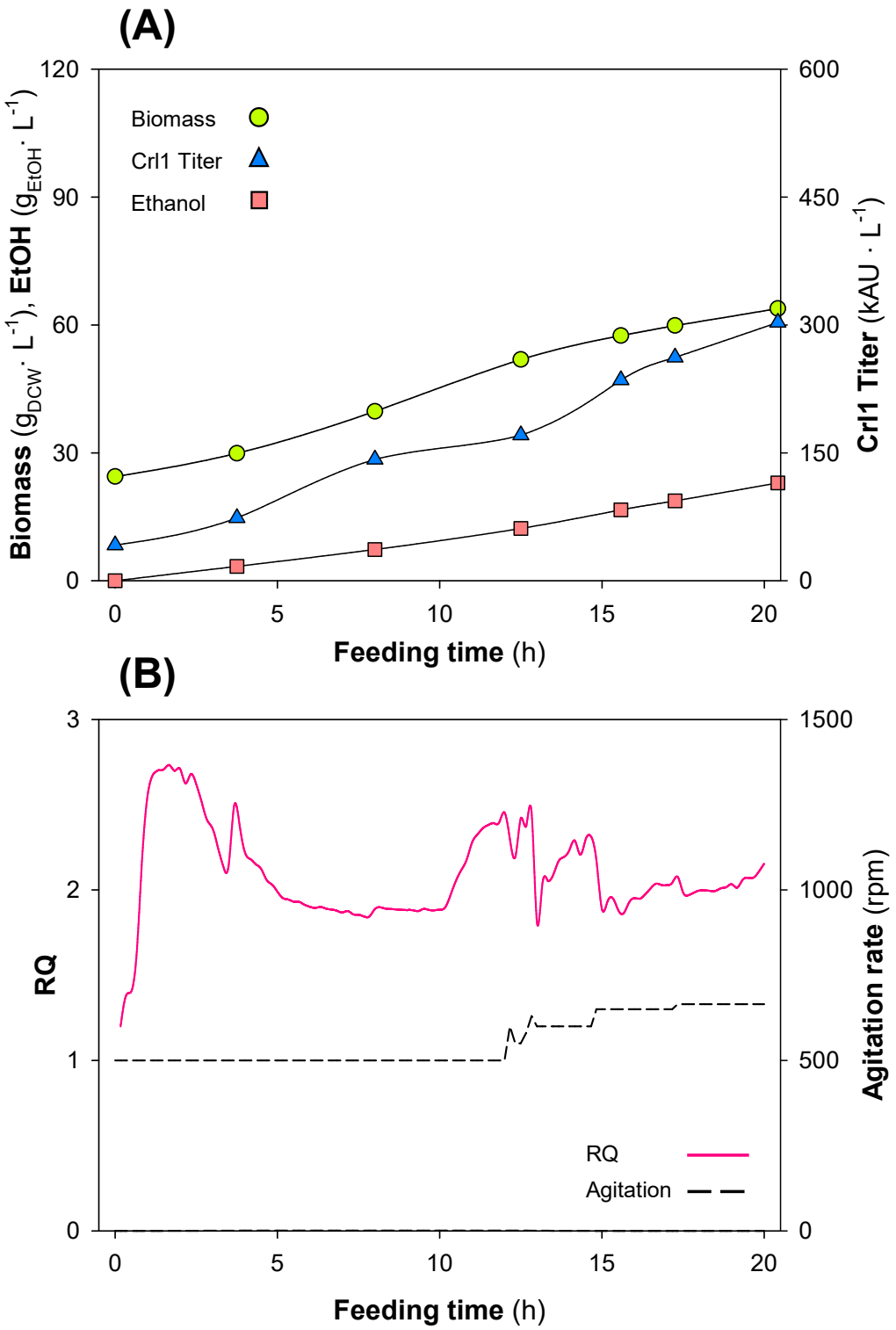


Figure 38. Primary data of the main key process parameters and *RQ* control of HPX-HIRQ. **(A)** Biomass concentration (●, g · L^{-1}); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}). **(B)** Off-line *RQ* calculation (pink lines); agitation rate (dashed lines, rpm).

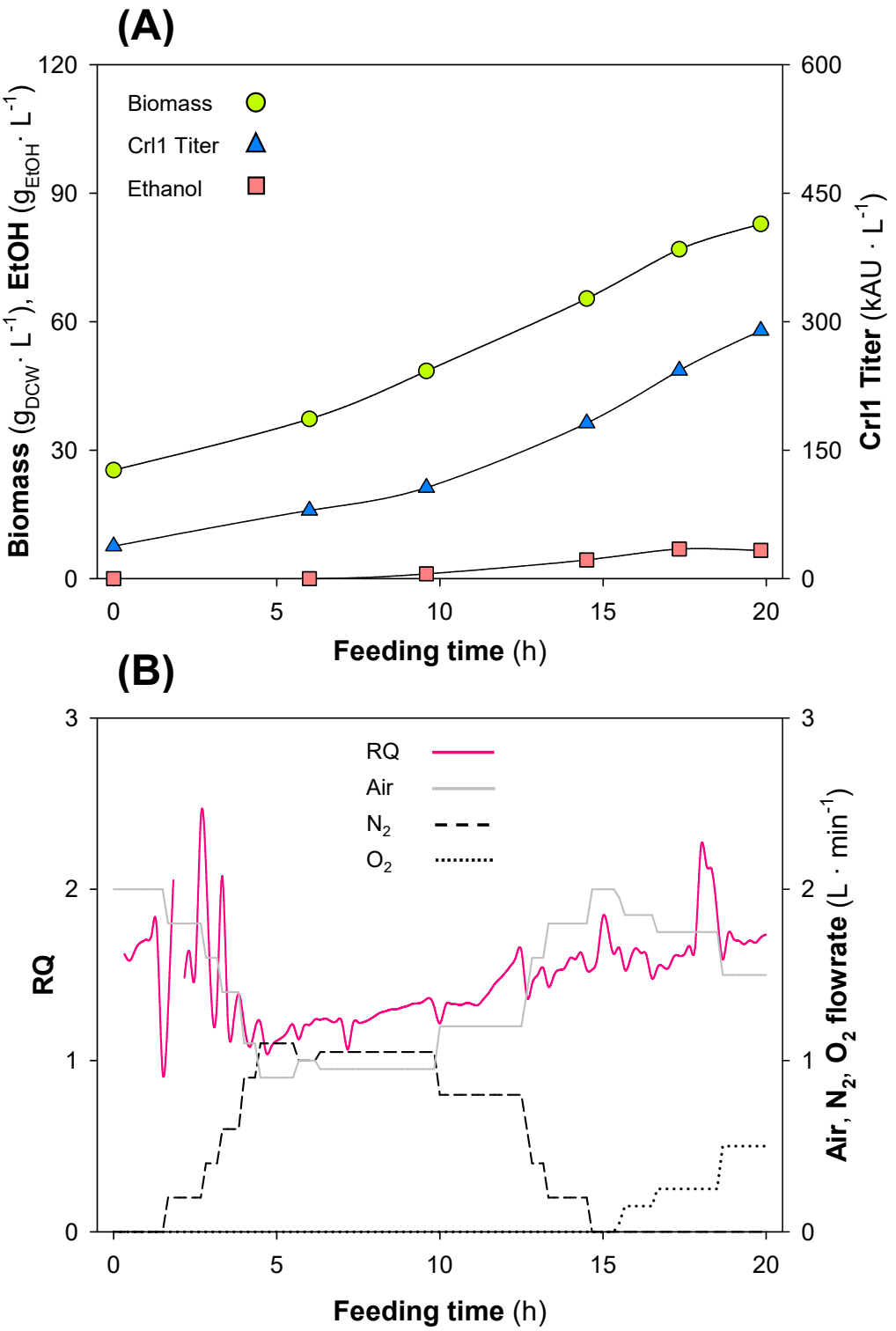
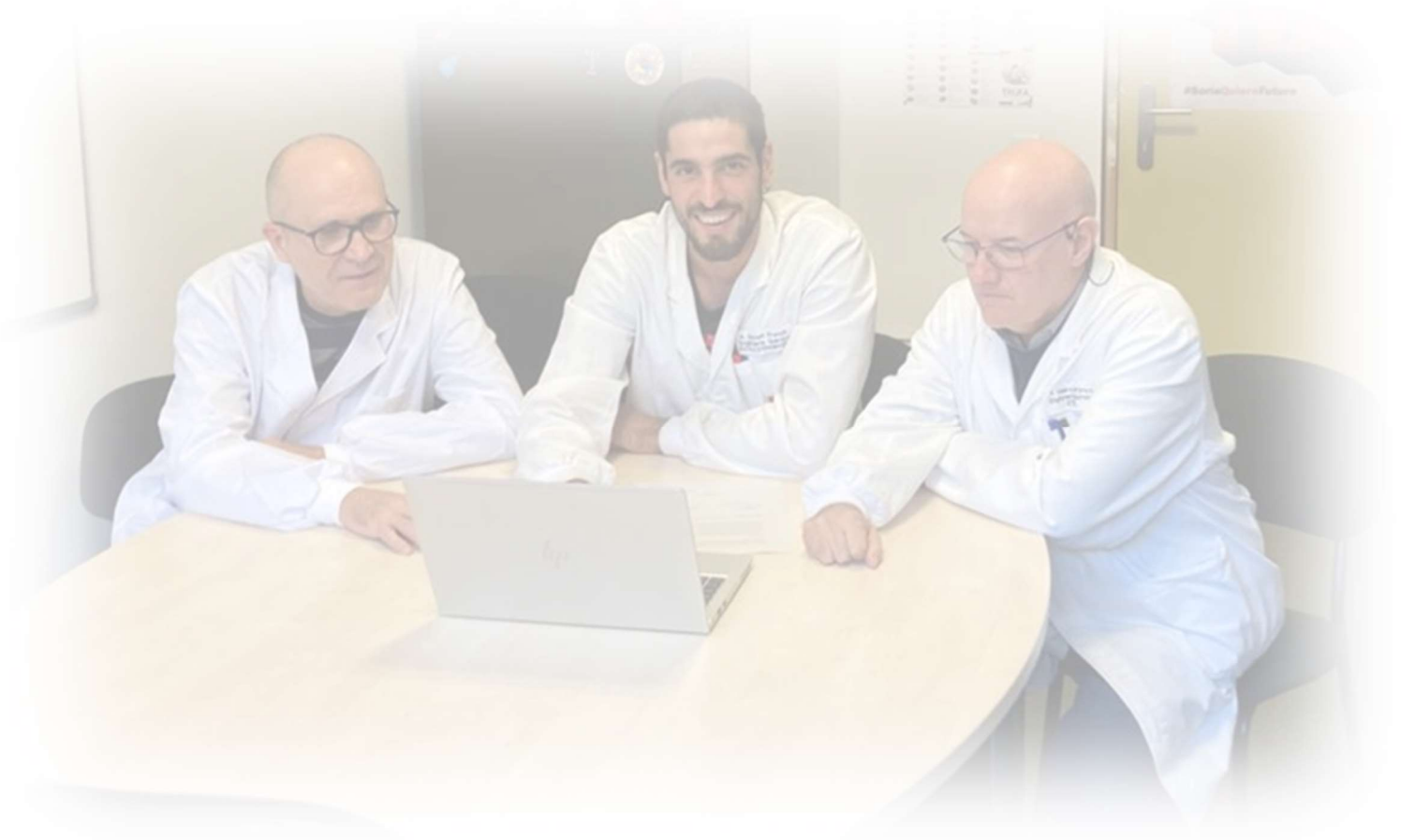


Figure 39. Primary data of the main key process parameters and *RQ* control of HPX-GASC. **(A)** Biomass concentration (●, g · L^{-1}); Crl1 Titer (▲, kAU · L^{-1}); EtOH, Ethanol concentration (■, g · L^{-1}). **(B)** Off-line *RQ* calculation (pink lines); air flowrate (gray lines, L · min^{-1}); nitrogen flowrate (dashed lines, L · min^{-1}); oxygen flowrate (dotted lines, L · min^{-1}).

10. BIBLIOGRAPHY



Abad, S., Nahalka, J., Winkler, M., Bergler, G., Speight, R., Glieder, A., & Nidetzky, B. (2011). High-level expression of *Rhodotorula gracilis* D-amino acid oxidase in *Pichia pastoris*. *Biotechnology Letters*, *33*(3), 557–563. https://doi.org/10.1007/s10529-010-0456-9

Adelantado, N., Tarazona, P., Grillitsch, K., García-Ortega, X., Monforte, S., Valero, F., Feussner, I., Daum, G., & Ferrer, P. (2017). The effect of hypoxia on the lipidome of recombinant *Pichia pastoris*. *Microbial Cell Factories*, *16*(1), 86. https://doi.org/10.1186/s12934-017-0699-4

Ahmad, M., Hirz, M., Pichler, H., & Schwab, H. (2014). Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production. In *Applied Microbiology and Biotechnology* (Vol. 98, Issue 12, pp. 5301–5317). Springer Verlag. https://doi.org/10.1007/s00253-014-5732-5

Alford, J. S. (2006). Bioprocess control: Advances and challenges. *Computers and Chemical Engineering*, *30*(10–12), 1464–1475. https://doi.org/10.1016/j.compchemeng.2006.05.039

Allampalli, P., Rathinavelu, S., Mohan, N., & Sivaprakasam, S. (2022). Deployment of metabolic heat rate based soft sensor for estimation and control of specific growth rate in glycoengineered *Pichia pastoris* for human interferon alpha 2b production. *Journal of Biotechnology*, *359*, 194-206.

Ask, M., & Stocks, S. M. (2022). Aerobic bioreactors: condensers, evaporation rates, scale-up and scale-down. *Biotechnology Letters*, *44*(7), 813-822.

Ata, Ö., Prielhofer, R., Gasser, B., Mattanovich, D., & Çalık, P. (2017). Transcriptional engineering of the glyceraldehyde-3-phosphate dehydrogenase promoter for improved heterologous protein production in *Pichia pastoris*. *Biotechnology and Bioengineering*, *114*(10), 2319–2327. https://doi.org/10.1002/bit.26363

Ata, Ö., Rebnegger, C., Tatto, N. E., Valli, M., Mairinger, T., Hann, S., Steiger, M. G., Çalık, P., & Mattanovich, D. (2018). A single Gal4-like transcription factor activates the Crabtree effect in *Komagataella phaffii*. *Nature Communications*, *9*(1), 4911. https://doi.org/10.1038/s41467-018-07430-4

Aw, R., & Polizzi, K. M. (2013). Can too many copies spoil the broth? *Microbial Cell Factories*, *12*(1), 128. https://doi.org/10.1186/1475-2859-12-128

Baeshen, M. N., Al-Hejin, A. M., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., Saini, K. S., Baeshen, N. A., & Redwan, E. M. (2015). Production of biopharmaceuticals in *E. Coli*: Current scenario and future perspectives. In *Journal of Microbiology and Biotechnology* (Vol. 25, Issue 7, pp. 953–962). Korean Society for Microbiology and Biotechnology. https://doi.org/10.4014/jmb.1412.12079

Baeza, J. A. (2017). Principles of Bioprocess Control. In *Current Developments in Biotechnology and Bioengineering: Bioprocesses, Bioreactors and Controls*. Elsevier B.V. https://doi.org/10.1016/B978-0-444-63663-8.00018-5

Baghban, R., Farajnia, S., Rajabibazl, M., Ghasemi, Y., Mafi, A. A., Hoseinpoor, R., Rahbarnia, L., & Aria, M. (2019). Yeast Expression Systems: Overview and Recent Advances. In *Molecular Biotechnology* (Vol. 61, Issue 5, pp. 365–384). Humana Press Inc. https://doi.org/10.1007/s12033-019-00164-8

Bao, J., Guo, D., Li, J., & Zhang, J. (2019). The modelling and operations for the digital twin in the context of manufacturing. *Enterprise Information Systems*, *13*(4), 534-556.

Barreiro, A., Prenafeta, A., Bech-Sabat, G., Roca, M., Perozo Mur, E., March, R., González-González, L., Madrenas, L., Corominas, J., Fernández, A., Moros, A., Cañete, M., Molas, M., Pentinat-Pelegrin, T., Panosa, C., Moreno, A., Puigvert Molas, E., Pol Vilarrassa, E., Palmada, J., … Ferrer, L. (2023). Preclinical evaluation of a COVID-19 vaccine candidate based on a recombinant RBD fusion heterodimer of SARS-CoV-2. *IScience*, *26*(3). https://doi.org/10.1016/j.isci.2023.106126

Barrero, J. J. (2020). Overcoming the secretory limitations in *Pichia pastoris* for recombinant protein production.

Barrero, J. J., Casler, J. C., Valero, F., Ferrer, P., & Glick, B. S. (2018). An improved secretion signal enhances the secretion of model proteins from *Pichia pastoris*. *Microbial Cell Factories*, *17*(1). https://doi.org/10.1186/s12934-018-1009-5

Barrigón, J. M., Montesinos, J. L., & 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? *Biochemical Engineering Journal*, *75*, 47–54. https://doi.org/10.1016/j.bej.2013.03.018

Barrigon, J. M., Valero, F., & Montesinos, J. L. (2015). A macrokinetic model-based comparative meta-analysis of recombinant protein production by *Pichia pastoris* under AOX1 promoter. *Biotechnology and Bioengineering*, *112*(6), 1132–1145. https://doi.org/10.1002/bit.25518

Bastin, G., & Dochain, D. (1990). Adaptive Control of Bioreactors. *On-Line Estimation and Adaptive Control of Bioreactors*, 251–337. https://doi.org/10.1016/b978-0-444-88430-5.50009-8

Baumann, K., Carnicer, M., Dragosits, M., Graf, A. B., Stadlmann, J., Jouhten, P., Maaheimo, H., Gasser, B., Albiol, J., Mattanovich, D., & Ferrer, P. (2010). A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. *BMC Systems Biology*, *4*(1), 141. https://doi.org/10.1186/1752-0509-4-141

Baumann, K., Maurer, M., Dragosits, M., Cos, O., Ferrer, P., & Mattanovich, D. (2008). Hypoxic fed-batch cultivation ofPichia pastoris increases specific and volumetric productivity of recombinant proteins. *Biotechnology and Bioengineering*, *100*(1), 177–183. https://doi.org/10.1002/bit.21763

Beck, A., Cochet, O., & Wurch, T. (2010). Glycofi's technology to control the glycosylation of recombinant therapeutic proteins. In *Expert Opinion on Drug Discovery* (Vol. 5, Issue 1, pp. 95–111). https://doi.org/10.1517/17460440903413504

Beiroti, A., Aghasadeghi, M. R., Hosseini, S. N., & Norouzian, D. (2019). Application of recurrent neural network for online prediction of cell density of recombinant *Pichia pastoris* producing HBsAg. *Preparative Biochemistry and Biotechnology*, 49(4), 352-359.

Benjamin, S., & Pandey, A. (1998). *Candida rugosa* lipases: Molecular biology and versatility in biotechnology. *Yeast*, *14*(12), 1069–1087. https://doi.org/10.1002/(SICI)1097-0061(19980915)14:12<1069::AID-YEA303>3.0.CO;2-K

Bernat-Camps, N., Ebner, K., Schusterbauer, V., Fischer, J. E., Nieto-Taype, M. A., Valero, F., Glieder, A., & Garcia-Ortega, X. (2023). Enabling growth-decoupled *Komagataella phaffii* recombinant protein production based on the methanol-free PDH promoter. *Frontiers in Bioengineering and Biotechnology*, *11*. https://doi.org/10.3389/fbioe.2023.1130583

Biggs, A. I. (1954). A spectrophotometric determination of the dissociation constants of p-nitrophenol and papaverine. *Transactions of the Faraday Society*, *50*, 800-802.

Bilal, M., Iqbal, H. M. N., Guo, S., Hu, H., Wang, W., & Zhang, X. (2018). State-of-the-art protein engineering approaches using biological macromolecules: A review from immobilization to implementation viewpoint. In *International Journal of Biological Macromolecules* (Vol. 108, pp. 893–901). Elsevier B.V. https://doi.org/10.1016/j.ijbiomac.2017.10.182

BlueSens (2017). *BlueVis 4.0 Handbook* (Rev. 170405002).

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

Brignoli, Y., Freeland, B., Cunningham, D., & Dabros, M. (2020). Control of specific growth rate in fed-batch bioprocesses: Novel controller design for improved noise management. *Processes*, *8*(6), 5–8. https://doi.org/10.3390/PR8060679

Brunner, V., Siegl, M., Geier, D., & Becker, T. (2021). Challenges in the Development of Soft Sensors for Bioprocesses: A Critical Review. *Frontiers in Bioengineering and Biotechnology*, *9*(August), 1–21. https://doi.org/10.3389/fbioe.2021.722202

Bustos, C., Quezada, J., Veas, R., Altamirano, C., Braun-Galleani, S., Fickers, P., & Berrios, J. (2022). Advances in Cell Engineering of the *Komagataella phaffii* Platform for Recombinant Protein Production. In *Metabolites* (Vol. 12, Issue 4). MDPI. https://doi.org/10.3390/metabo12040346

Çalık, P., Hoxha, B., Çalık, G., & Özdamar, T. H. (2018). Hybrid fed-batch bioreactor operation design: control of substrate uptake enhances recombinant protein production in high-cell-density fermentations. *Journal of Chemical Technology & Biotechnology*, *93*(11), 3326–3335. https://doi.org/10.1002/jctb.5696

Cámara, E., Albiol, J., & Ferrer, P. (2016). Droplet digital PCR-aided screening and characterization of *Pichia pastoris* multiple gene copy strains. *Biotechnology and Bioengineering*, *113*(7), 1542–1551. https://doi.org/10.1002/bit.25916

Cámara, E., Landes, N., Albiol, J., Gasser, B., Mattanovich, D., & Ferrer, P. (2017). Increased dosage of AOX1 promoter-regulated expression cassettes leads to transcription attenuation of the methanol metabolism in *Pichia pastoris*. *Scientific Reports*, *7*(1), 44302. https://doi.org/10.1038/srep44302

Carnicer, M., Baumann, K., Töplitz, I., Sánchez-Ferrando, F., Mattanovich, D., Ferrer, P., & Albiol, J. (2009). Macromolecular and elemental composition analysis and extracellular metabolite balances of *Pichia pastoris* growing at different oxygen levels. *Microbial Cell Factories*, *8*(1), 65. https://doi.org/10.1186/1475-2859-8-65

Carnicer, M., Ten Pierick, A., Van Dam, J., Heijnen, J. J., Albiol, J., Van Gulik, W., & Ferrer, P. (2012). Quantitative metabolomics analysis of amino acid metabolism in recombinant *Pichia pastoris* under different oxygen availability conditions. *Microbial Cell Factories*, *11*, 1–10. https://doi.org/10.1186/1475-2859-11-83

Chang, S.-W., Lee, G.-C., & Shaw, J.-F. (2006). Codon Optimization of Candida rugosa lip 1 Gene for Improving Expression in *Pichia pastoris* and Biochemical Characterization of the Purified Recombinant LIP1 Lipase. *Journal of Agricultural and Food Chemistry*, *54*(3), 815–822. https://doi.org/10.1021/jf052183k

Christensen, L. H., Schulze, U., Nielsen, J., & Villadsen, J. (1995). Acoustic off-gas analyser for bioreactors: Precision, accuracy, and dynamics of detection. *Chemical Engineering Science*, *50*(16), 2601–2610. https://doi.org/10.1016/0009-2509(95)00108-H

Clementschitsch, F., & Bayer, K. (2006). Improvement of bioprocess monitoring: Development of novel concepts. *Microbial Cell Factories*, *5*, 1–11. https://doi.org/10.1186/1475-2859-5-19

Cos, O., Ramon, R., Montesinos, J. L., & Valero, F. (2006). A simple model-based control for *Pichia pastoris* allows a more efficient heterologous protein production bioprocess. *Biotechnology and Bioengineering*, *95*(1), 145–154. https://doi.org/10.1002/bit.21005

Cos, O., Resina, D., Ferrer, P., Montesinos, J. L., & Valero, F. (2005b). Heterologous production of *Rhizopus oryzae* lipase in *Pichia pastoris* using the alcohol oxidase and formaldehyde dehydrogenase promoters in batch and fed-batch cultures. *Biochemical Engineering Journal*, *26*(2–3), 86–94. https://doi.org/10.1016/j.bej.2005.04.005

Cos, O., Serrano, A., Montesinos, J. L., Ferrer, P., Cregg, J. M., & Valero, F. (2005a). Combined effect of the methanol utilization (Mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. *Journal of Biotechnology*, *117*(1), 321–335. https://doi.org/10.1016/j.jbiotec.2004.12.010

Craven, S., Whelan, J., & Glennon, B. (2014). Glucose concentration control of a fed-batch mammalian cell bioprocess using a nonlinear model predictive controller. *Journal of Process Control*, *24*(4), 344–357. https://doi.org/10.1016/j.jprocont.2014.02.007

Cregg, J. M., Madden, K. R., Barringer, K. J., Thill, G. P., & Stillman, C. A. (1989). Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. *Molecular and Cellular Biology*, *9*(3), 1316–1323. https://doi.org/10.1128/MCB.9.3.1316

Curvers, S., Brixius, P., Klauser, T., Thommes, J., Weuster-Botz, D., Takors, R., & Wandrey, C. (2001). Human Chymotrypsinogen B Production with *Pichia pastoris* by Integrated Development of Fermentation and Downstream Processing. Part 1. Fermentation. *Biotechnology Progress*, *17*(3), 495–502. https://doi.org/10.1021/bp000164j

Dabros, M., Schuler, M. M., & Marison, I. W. (2010). Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. *Bioprocess and Biosystems Engineering*, *33*(9), 1109–1118. https://doi.org/10.1007/s00449-010-0438-2

Daniyan, I.A., Daniyan, O. L., Adeodu, A. O., & Aribidara, A. A. (2014). (2014). Towards Sustainable Development in the Third World: Design of a Large-Scale Biodiesel Plant. *Journal of Bioprocessing and Chemical Engineering*, 1–9.

Darby, R. A., Cartwright, S. P., Dilworth, M. V., & Bill, R. M. (2012). Which yeast species shall I choose? *Saccharomyces cerevisiae* versus *Pichia pastoris*. *Recombinant protein production in yeast: methods and protocols*, 11-23.

Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of Artificial Intelligence in Machine Learning: Review and Prospect. *International Journal of Computer Applications*, *115*(9), 31–41. https://doi.org/10.5120/20182-2402

DaSilva, E. J. (2004). The colours of biotechnology: science, development and humankind. *Electronic journal of biotechnology*, *7*(3), 01-02.

De Brabander, P., Uitterhaegen, E., Delmulle, T., De Winter, K., & Soetaert, W. (2023). Challenges and progress towards industrial recombinant protein production in yeasts: A review. *Biotechnology Advances*, 108121.

De Macedo Robert, J., Garcia-Ortega, X., Montesinos-Seguí, J. L., Guimaraes Freire, D. M., & Valero, F. (2019). Continuous operation, a realistic alternative to fed-batch fermentation for the production of recombinant lipase B from *Candida antarctica* under the constitutive promoter PGK in *Pichia pastoris*. *Biochemical Engineering Journal*, *147*, 39–47. https://doi.org/10.1016/j.bej.2019.03.027

De Macedo Robert, J., Lattari, F. S., Machado, A. C., de Castro, A. M., Almeida, R. V., Torres, F. A. G., ... & Freire, D. M. G. (2017). Production of recombinant lipase B from *Candida antarctica* in *Pichia pastoris* under control of the promoter PGK using crude glycerol from biodiesel production as carbon source. *Biochemical Engineering Journal*, *118*, 123-131

De Schutter, K., Lin, Y. C., Tiels, P., Van Hecke, A., Glinka, S., Weber-Lehmann, J., Rouzé, P., Van De Peer, Y., & Callewaert, N. (2009). Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nature Biotechnology*, *27*(6), 561–566. https://doi.org/10.1038/nbt.1544

Doran, P.M. (2013) Bioprocess Engineering Principles, 2nd ed.; Elsevier: New York.

Dubey, K. K., Ray, A. R., & Behera, B. K. (2008). Production of demethylated colchicine through microbial transformation and scale-up process development. *Process Biochemistry*, *43*(3), 251-257.

Duong-Trung, N., Born, S., Kim, J. W., Schermeyer, M. T., Paulick, K., Borisyak, M., ... & Martinez, E. (2023). When bioprocess engineering meets machine learning: A survey from the perspective of automated bioprocess development. *Biochemical Engineering Journal*, *190*, 108764.

Embury, M., & Clayton, J. (2017). Standardization and certification challenges for biopharmaceutical plants. *Computer*, *50*(9), 83-86.

Eom, G. T., Lee, S. H., Song, B. K., Chung, K. W., Kim, Y. W., & Song, J. K. (2013). High-level extracellular production and characterization of *Candida antarctica* lipase B in *Pichia pastoris*. *Journal of Bioscience and Bioengineering*, *116*(2), 165–170. https://doi.org/10.1016/j.jbiosc.2013.02.016

Ergün, B. G., Gasser, B., Mattanovich, D., & Çalık, P. (2019). Engineering of alcohol dehydrogenase 2 hybrid-promoter architectures in *Pichia pastoris* to enhance recombinant protein expression on ethanol. *Biotechnology and Bioengineering*, *116*(10), 2674–2686. https://doi.org/10.1002/bit.27095

European Commission. (2017). New techniques in agricultural biotechnology.

European Medicines Agency. (2014). Guideline on process validation for the manufacture of biotechnology-derived active substances and data to be provided in the regulatory submission. *Committee for Medicinal Products for Human Use (CHMP) Guideline*, *44*(April), Stability studies on rFVIIIFc drug substance. 1–14.

Ferreira, A. R., Ataíde, F., Von Stosch, M., Dias, J. M. L., Clemente, J. J., Cunha, A. E., & Oliveira, R. (2012). Application of adaptive DO-stat feeding control to *Pichia pastoris* X33 cultures expressing a single chain antibody fragment (scFv). *Bioprocess and Biosystems Engineering*, *35*(9), 1603–1614. https://doi.org/10.1007/s00449-012-0751-z

Ferrer, P., Montesinos, J. L., Valero, F., & Solà, C. (2001). Production of Native and Recombinant Lipases by *Candida rugosa*: A Review. *Applied Biochemistry and Biotechnology*, *95*(3), 221–256. https://doi.org/10.1385/ABAB:95:3:221

Fina, A., Brêda, G. C., Pérez-Trujillo, M., Freire, D. M. G., Almeida, R. V., Albiol, J., & Ferrer, P. (2021). Benchmarking recombinant *Pichia pastoris* for 3-hydroxypropionic acid production from glycerol. *Microbial Biotechnology*, *14*(4), 1671–1682. https://doi.org/10.1111/1751-7915.13833

Fina, A., Heux, S., Albiol, J., & Ferrer, P. (2022). Combining Metabolic Engineering and Multiplexed Screening Methods for 3-Hydroxyprionic Acid Production in *Pichia pastoris*. *Frontiers in Bioengineering and Biotechnology*, *10*. https://doi.org/10.3389/fbioe.2022.942304

Food and Drug Administration. (2011). Guidance for Industry Process Validation: General Principles and Practices Guidance for Industry.

Food and Drug Administration. (2019). Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD).

Frick, O., & Wittmann, C. (2005). Characterization of the metabolic shift between oxidative and fermentative growth in *Saccharomyces cerevisiae* by comparative 13C flux analysis. *Microbial Cell Factories*, *4*. https://doi.org/10.1186/1475-2859-4-30

Fröhling, M., & Hiete, M. (2020). The sustainability and life cycle assessments of industrial biotechnology: an introduction. In *Advances in Biochemical Engineering/Biotechnology* (Vol. 173, pp. 3–9). Springer. https://doi.org/10.1007/10_2020_123

Funke, M., Buchenauer, A., Mokwa, W., Kluge, S., Hein, L., Müller, C., Kensy, F., & Büchs, J. (2010). Bioprocess Control in Microscale: Scalable Fermentations in Disposable and User-Friendly Microfluidic Systems. *Microbial Cell Factories*, *9*, 1–13. https://doi.org/10.1186/1475-2859-9-86

Galvanauskas, V., Simutis, R., Levišauskas, D., & Urniežius, R. (2019). Practical solutions for specific growth rate control systems in industrial bioreactors. *Processes*, *7*(10), 1–14. https://doi.org/10.3390/pr7100693

Garcia-Ochoa, F., & Gomez, E. (2009). Bioreactor scale-up and oxygen transfer rate in microbial processes: An overview. *Biotechnology Advances*, *27*(2), 153–176. https://doi.org/10.1016/j.biotechadv.2008.10.006

Garcia-Ochoa, F., Gomez, E., Santos, V. E., & Merchuk, J. C. (2010). Oxygen uptake rate in microbial processes: An overview. *Biochemical Engineering Journal*, *49*(3), 289–307. https://doi.org/10.1016/j.bej.2010.01.011

Garcia-Ortega, X., Adelantado, N., Ferrer, P., Montesinos, J. L., & Valero, F. (2016). A step forward to improve recombinant protein production in *Pichia pastoris*: From specific growth rate effect on protein secretion to carbon-starving conditions as advanced strategy. *Process Biochemistry*, *51*(6), 681–691. https://doi.org/10.1016/j.procbio.2016.02.018

García-Ortega, X., Cámara, E., Ferrer, P., Albiol, J., Montesinos-Seguí, J. L., & Valero, F. (2019). Rational development of bioprocess engineering strategies for recombinant protein production in *Pichia pastoris* (*Komagataella phaffii*) using the methanol-free GAP promoter. Where do we stand? *New Biotechnology*, *53*(August 2018), 24–34. https://doi.org/10.1016/j.nbt.2019.06.002

Garcia-Ortega, X., Ferrer, P., Montesinos, J. L., & Valero, F. (2013). Fed-batch operational strategies for recombinant Fab production with *Pichia pastoris* using the constitutive GAP promoter. *Biochemical Engineering Journal*, *79*, 172–181. https://doi.org/10.1016/j.bej.2013.07.013

Garcia-Ortega, X., Valero, F., & Montesinos-Seguí, J. L. (2017). Physiological state as transferable operating criterion to improve recombinant protein production in *Pichia pastoris* through oxygen limitation. *Journal of Chemical Technology & Biotechnology*, *92*(10), 2573–2582. https://doi.org/10.1002/jctb.5272

Garrigós-Martínez, J., Nieto-Taype, M. A., Gasset-Franch, A., Montesinos-Seguí, J. L., Garcia-Ortega, X., & Valero, F. (2019). Specific growth rate governs AOX1 gene expression, affecting the production kinetics of Pichia pastoris (Komagataella phaffii) PAOX1-driven recombinant producer strains with different target gene dosage. *Microbial Cell Factories*, *18*(1), 187. https://doi.org/10.1186/s12934-019-1240-8

Garrigós-Martínez, J., Vuoristo, K., Nieto-Taype, M. A., Tähtiharju, J., Uusitalo, J., Tukiainen, P., Schmid, C., Tolstorukov, I., Madden, K., Penttilä, M., Montesinos-Seguí, J. L., Valero, F., Glieder, A., & Garcia-Ortega, X. (2021). Bioprocess performance analysis of novel methanol-independent promoters for recombinant protein production with *Pichia pastoris*. *Microbial Cell Factories*, *20*(1), 74. https://doi.org/10.1186/s12934-021-01564-9

Gasser, B., & Mattanovich, D. (2018). A yeast for all seasons – Is *Pichia pastoris* a suitable chassis organism for future bioproduction? *FEMS microbiology letters*, 365(17), fny181.

Gasset, A., Garcia-Ortega, X., Garrigós-Martínez, J., Valero, F., & Montesinos-Seguí, J. L. (2022). Innovative Bioprocess Strategies Combining Physiological Control and Strain Engineering of *Pichia pastoris* to Improve Recombinant Protein Production. *Frontiers in Bioengineering and Biotechnology*, *10*(January), 1–12. https://doi.org/10.3389/fbioe.2022.818434

Gnoth, S., Jenzsch, M., Simutis, R., & Lübbert, A. (2008). Control of cultivation processes for recombinant protein production: A review. *Bioprocess and Biosystems Engineering*, *31*(1), 21–39. https://doi.org/10.1007/s00449-007-0163-7

Gomes, A. M. V., Carmo, T. S., Carvalho, L. S., Bahia, F. M., & Parachin, N. S. (2018). Comparison of yeasts as hosts for recombinant protein production. In *Microorganisms* (Vol. 6, Issue 2). MDPI AG. https://doi.org/10.3390/microorganisms6020038

Goodwin, G. C. (2000). Predicting the performance of soft sensors as a route to low cost automation. *Annual Reviews in Control*, *24*, 55–66. https://doi.org/10.1016/S1367-5788(00)00012-2

Guerfal, M., Ryckaert, S., Jacobs, P. P., Ameloot, P., Craenenbroeck, K. Van, Derycke, R., & Callewaert, N. (2010). The HAC1 gene from Pichia pastoris: characterization and effect of its

overexpression on the production of secreted, surface displayed and membrane proteins. In *Microbial Cell Factories* (Vol. 9). http://www.microbialcellfactories.com/content/9/1/49

Gunge, N., Tamaru, A., Ozawa, F., & Sakaguchi, K. (1981). Isolation and characterization of linear deoxyribonucleic acid plasmids from *Kluyveromyces lactis* and the plasmid-associated killer character. *Journal of Bacteriology*, *145*(1), 382-390.

Gupta, S. K., & Shukla, P. (2017). Microbial platform technology for recombinant antibody fragment production: A review. In *Critical Reviews in Microbiology* (Vol. 43, Issue 1, pp. 31–42). Taylor and Francis Ltd. https://doi.org/10.3109/1040841X.2016.1150959

Hayman, G. T., & Bolen, P. L. (1991). Linear DNA plasmids of *Pichia inositovora* are associated with a novel killer toxin activity. *Current genetics*, *19*, 389-393.

Heux, S., Meynial-Salles, I., O'Donohue, M. J., & Dumon, C. (2015). White biotechnology: State of the art strategies for the development of biocatalysts for biorefining. In *Biotechnology Advances* (Vol. 33, Issue 8, pp. 1653–1670). Elsevier Inc. https://doi.org/10.1016/j.biotechadv.2015.08.004

Hewitt, C. J., & Nienow, A. W. (2007). The Scale-Up of Microbial Batch and Fed-Batch Fermentation Processes. *Advances in Applied Microbiology*, *62*(07), 105–135. https://doi.org/10.1016/S0065-2164(07)62005-X

Heyland, J., Fu, J., Blank, L. M., & Schmid, A. (2011). Carbon metabolism limits recombinant protein production in *Pichia pastoris*. *Biotechnology and Bioengineering*, *108*(8), 1942–1953. https://doi.org/10.1002/bit.23114

Higgins, D. R., & Cregg, J. M. (1998). Methods in molecular biology: Pichia protocols. *Production of recombinant proteins in fermentor cultures of the yeast Pichia pastoris. Curr Opin Biotechnol*, *13*, 329-332.

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*. *Biotechnology and Bioengineering*, *85*(4), 367–375. https://doi.org/10.1002/bit.10904

Hoskisson, P. A., & Hobbs, G. (2005). Continuous culture - Making a comeback? In *Microbiology* (Vol. 151, Issue 10, pp. 3153–3159). https://doi.org/10.1099/mic.0.27924-0

Hudson, J. (2017). Genetically modified products and GMO foods: A game of chance. In *Developing New Functional Food and Nutraceutical Products* (pp. 481–494). Elsevier Inc. https://doi.org/10.1016/B978-0-12-802780-6.00027-4

Ivarsson, C. J. (2017). Visualization for Advanced Big Data Analysis. *Genetic Engineering and Biotechnology News*, *37*(2), 32. https://doi.org/10.1089/gen.37.02.17

Jaeger, K.-E., & Eggert, T. (2002). Lipases for biotechnology. *Current Opinion in Biotechnology*, *13*(4), 390–397. https://doi.org/10.1016/S0958-1669(02)00341-5

Jenzsch, M., Simutis, R., & Luebbert, A. (2006). Generic model control of the specific growth rate in recombinant *Escherichia coli* cultivations. *Journal of Biotechnology*, *122*(4), 483–493. https://doi.org/10.1016/j.jbiotec.2005.09.013

Jordà, J., de Jesus, S. S., Peltier, S., Ferrer, P., & Albiol, J. (2014). Metabolic flux analysis of recombinant *Pichia pastoris* growing on different glycerol/methanol mixtures by iterative fitting of NMR-derived 13C-labelling data from proteinogenic amino acids. *New Biotechnology*, *31*(1), 120–132. https://doi.org/10.1016/j.nbt.2013.06.007

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. *Microbial cell factories*, *11*(1), 1-14.

Junne, S., Klingner, A., Kabisch, J., Schweder, T., & Neubauer, P. (2011). A two-compartment bioreactor system made of commercial parts for bioprocess scale-down studies: Impact of oscillations on *Bacillus subtilis* fed-batch cultivations. *Biotechnology Journal*, *6*(8), 1009–1017. https://doi.org/10.1002/biot.201100293

Kafarski, P. (2012). Rainbow code of biotechnology. *Chemik*, *66*(8), 811-816.

Kalender, Ö., & Çalık, P. (2020). Transcriptional regulatory proteins in central carbon metabolism of *Pichia pastoris* and *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology*, *104*(17), 7273–7311. https://doi.org/10.1007/s00253-020-10680-2

Karky, R. B., & Perry, M. (2019). Disharmonization in the regulation of transgenic plants in Europe. *Biotechnology Law Report*, *38*(6), 350–375. https://doi.org/10.1089/blr.2019.29135.rbk

Katla, S., Mohan, N., Pavan, S. S., Pal, U., & Sivaprakasam, S. (2019). Control of specific growth rate for the enhanced production of human interferon α2b in glycoengineered *Pichia pastoris*: process analytical technology guided approach. In *Journal of Chemical Technology and Biotechnology* (Vol. 94, Issue 10). https://doi.org/10.1002/jctb.6118

Keller, P. (2011). *Statistical process control demystified*. McGraw-Hill Education.

Ken Ugo, A., Vivian Amara, A., CN, I., & Kenechuwku, U. (2017). Microbial Lipases: A Prospect for Biotechnological Industrial Catalysis for Green Products: A Review. *Fermentation Technology*, *06*(02). https://doi.org/10.4172/2167-7972.1000144

Kim, H., Yoo, S. J., & Kang, H. A. (2015). Yeast synthetic biology for the production of recombinant therapeutic proteins. In *FEMS Yeast Research* (Vol. 15, Issue 1). Oxford University Press. https://doi.org/10.1111/1567-1364.12195

Kleter, G. A., & Kok, E. J. (2010). Safety assessment of biotechnology used in animal production, including genetically modified (GM) feed and GM animals-a review*. In *Animal Science Papers and Reports* (Vol. 28, Issue 2).

Knoll, A., Bartsch, S., Husemann, B., Engel, P., Schroer, K., Ribeiro, B., Stöckmann, C., Seletzky, J., & Büchs, J. (2007). High cell density cultivation of recombinant yeasts and bacteria under non-pressurized and pressurized conditions in stirred tank bioreactors. *Journal of Biotechnology*, *132*(2), 167–179. https://doi.org/10.1016/j.jbiotec.2007.06.010

Koutinas, A. A., Vlysidis, A., Pleissner, D., Kopsahelis, N., Lopez Garcia, I., Kookos, I. K., Papanikolaou, S., Kwan, T. H., & Lin, C. S. K. (2014). Valorization of industrial waste and by-product streams via fermentation for the production of chemicals and biopolymers. In *Chemical Society Reviews* (Vol. 43, Issue 8, pp. 2587–2627). Royal Society of Chemistry. https://doi.org/10.1039/c3cs60293a

Kroukamp, H., den Haan, R., van Zyl, J. H., & van Zyl, W. H. (2018). Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*. In *Biofuels, Bioproducts and Biorefining* (Vol. 12, Issue 1, pp. 108–124). John Wiley and Sons Ltd. https://doi.org/10.1002/bbb.1824

Kumar, A., Udugama, I. A., Gargalo, C. L., & Gernaey, K. V. (2020). Why Is Batch Processing Still Dominating the Biologics Landscape? Towards an Integrated Continuous Bioprocessing Alternative. *Processes*, *8*(12), 1641. https://doi.org/10.3390/pr8121641

Kurtzman, C. P. (2005). Description of *Komagataella phaffii* sp. nov. and the transfer of *Pichia pseudopastoris* to the methylotrophic yeast genus *Komagataella. International Journal of Systematic and Evolutionary Microbiology*, *55*(2), 973–976. https://doi.org/10.1099/ijs.0.63491-0

Kurtzman, C. P. (2009). Biotechnological strains of *Komagataella* (*Pichia*) pastoris are *Komagataella phaffii* as determined from multigene sequence analysis. *Journal of Industrial Microbiology and Biotechnology*, *36*(11), 1435–1438. https://doi.org/10.1007/s10295-009-0638-4

Ladisch', M. R., & Kohlmann, K. L. (1992). Recombinant Human Insulin. In *Biotechnol. Prog* (Vol. 1092).

Lara, A. R., Galindo, E., Ramírez, O. T., & Palomares, L. A. (2006). Living with heterogeneities in bioreactors: understanding the effects of environmental gradients on cells. *Molecular biotechnology*, *34*(3), 355-381.

Li, P., Anumanthan, A., Gao, X. G., Ilangovan, K., Suzara, V. V., Düzgüneş, N., & Renugopalakrishnan, V. (2007). Expression of recombinant proteins in *Pichia pastoris*. In *Applied Biochemistry and Biotechnology* (Vol. 142, Issue 2, pp. 105–124). https://doi.org/10.1007/s12010-007-0003-x

Liu, W. C., Gong, T., Wang, Q. H., Liang, X., Chen, J. J., & Zhu, P. (2016). Scaling-up Fermentation of *Pichia pastoris* to demonstration-scale using new methanol-feeding strategy and increased air pressure instead of pure oxygen supplement. *Scientific Reports*, *6*(September 2015), 1–12. https://doi.org/10.1038/srep18439

Looser, V., Bruhlmann, B., Bumbak, F., Stenger, C., Costa, M., Camattari, A., Fotiadis, D., & Kovar, K. (2015). Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. *Biotechnology Advances*, *33*(6), 1177–1193. https://doi.org/10.1016/j.biotechadv.2015.05.008

Looser, V., Lüthy, D., Straumann, M., Hecht, K., Melzoch, K., & Kovar, K. (2017). Effects of glycerol supply and specific growth rate on methanol-free production of CalB by *P. pastoris*: functional characterisation of a novel promoter. *Applied Microbiology and Biotechnology*, *101*(8), 3163–3176. https://doi.org/10.1007/s00253-017-8123-x

López-Fernández, J. (2022). Heterologous Rhizopus oryzae lipase expressed in *Komagataella phaffii*: improved production, biochemical characterization and industrial applications.

López-Fernández, J., Barrero, J. J., Benaiges, M. D., & Valero, F. (2019). Truncated prosequence of *Rhizopus oryzae* lipase: Key factor for production improvement and biocatalyst stability. *Catalysts*, *9*(11). https://doi.org/10.3390/catal9110961

López-Fernández, J., Benaiges, M. D., & Valero, F. (2020). *Rhizopus oryzae* Lipase, a Promising Industrial Enzyme: Biochemical Characteristics, Production and Biocatalytic Applications. *Catalysts*, *10*(11), 1277. https://doi.org/10.3390/catal10111277

López-Fernández, J., Benaiges, M. D., & Valero, F. (2021a). Constitutive expression in *Komagataella phaffii* of mature *Rhizopus oryzae* lipase jointly with its truncated prosequence improves production and the biocatalyst operational stability. *Catalysts*, *11*(10). https://doi.org/10.3390/catal11101192

López-Fernández, J., Dolors Benaiges, M., & Valero, F. (2021b). Second- and third-generation biodiesel production with immobilised recombinant *Rhizopus oryzae* lipase: Influence of the support, substrate acidity and bioprocess scale-up. *Bioresource Technology*, *334*. https://doi.org/10.1016/j.biortech.2021.125233

López-Fernández, J., Moya, D., Dolors Benaiges, M., Valero, F., & Alcalà, M. (2022). Near Infrared Spectroscopy: A useful technique for inline monitoring of the enzyme catalyzed biosynthesis of third-generation biodiesel from waste cooking oil. *Fuel*, *319*. https://doi.org/10.1016/j.fuel.2022.123794

Lorantfy, B., Jazini, M., & Herwig, C. (2013). Investigation of the physiological response to oxygen limited process conditions of *Pichia pastoris* Mut+ strain using a two-compartment scale-down system. *Journal of Bioscience and Bioengineering*, *116*(3), 371–379. https://doi.org/10.1016/j.jbiosc.2013.03.021

Love, K. R., Dalvie, N. C., & Love, J. C. (2018). The yeast stands alone: the future of protein biologic production. In *Current Opinion in Biotechnology* (Vol. 53, pp. 50–58). Elsevier Ltd. https://doi.org/10.1016/j.copbio.2017.12.010

Luo, Y., Kurian, V., & Ogunnaike, B. A. (2021). Bioprocess systems analysis, modeling, estimation, and control. *Current Opinion in Chemical Engineering*, *33*, 100705. https://doi.org/10.1016/j.coche.2021.100705

Luttmann, R., Bracewell, D. G., Cornelissen, G., Gernaey, K. V., Glassey, J., Hass, V. C., Kaiser, C., Preusse, C., Striedner, G., & Mandenius, C. F. (2012). Soft sensors in bioprocessing: A status report and recommendations. *Biotechnology Journal*, *7*(8), 1040–1048. https://doi.org/10.1002/biot.201100506

Lynd, L. R., Wyman, C. E., & Gerngross, T. U. (1999). Biocommodity engineering. *Biotechnology Progress*, *15*(5), 777–793. https://doi.org/10.1021/bp990109e

Ma, J. K. C., Drossard, J., Lewis, D., Altmann, F., Boyle, J., Christou, P., Cole, T., Dale, P., van Dolleweerd, C. J., Isitt, V., Katinger, D., Lobedan, M., Mertens, H., Paul, M. J., Rademacher, T., Sack, M., Hundleby, P. A. C., Stiegler, G., Stoger, E., … Fischer, R. (2015). Regulatory approval and a first-in-human phase I clinical trial of a monoclonal antibody produced in transgenic tobacco plants. *Plant Biotechnology Journal*, *13*(8), 1106–1120. https://doi.org/10.1111/pbi.12416

Macauley-Patrick, S., Fazenda, M. L., McNeil, B., & Harvey, L. M. (2005). Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, *22*(4), 249–270. https://doi.org/10.1002/yea.1208

Macdonald, G. J. (2018). Biosensors Track More Metrics, Channel More Data. *Genetic Engineering & Biotechnology News*, *38*(15), 20-22.

Manzano, T. (Apr 7, 2020). PDA Study Explores Role of A.I. in CPV. https://www.pda.org/pda-letter-portal/home/full-article/pda-study-explores-role-of-a.i.-in-cpv

Manzano, T., Fernàndez, C., Ruiz, T., & Richard, H. (2021). Artificial Intelligence Algorithm Qualification: A Quality by Design Approach to Apply Artificial Intelligence in Pharma. *PDA Journal of Pharmaceutical Science and Technology*, *75*(1), 100-118.

Marques, M. P. C., Cabral, J. M. S., & Fernandes, P. (2010). Bioprocess scale-up: Quest for the parameters to be used as criterion to move from microreactors to lab-scale. *Journal of Chemical Technology and Biotechnology*, *85*(9), 1184–1198. https://doi.org/10.1002/jctb.2387

Marx, H., Mecklenbräuker, A., Gasser, B., Sauer, M., & Mattanovich, D. (2009). Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. *FEMS Yeast Research*, *9*(8), 1260–1270. https://doi.org/10.1111/j.1567-1364.2009.00561.x

Mattanovich, D., Graf, A., Stadlmann, J., Dragosits, M., Redl, A., Maurer, M., Kleinheinz, M., Sauer, M., Altmann, F., & Gasser, B. (2009). Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microbial Cell Factories*, *8*. https://doi.org/10.1186/1475-2859-8-29

Mattanovich, D., Sauer, M., & Gasser, B. (2017). Industrial microorganisms: *Pichia pastoris*. *Industrial biotechnology: microorganisms*, *2*, 687-714.

Mehta, A., Guleria, S., Sharma, R., & Gupta, R. (2020). The lipases and their applications with emphasis on food industry. In *Microbial Biotechnology in Food and Health* (pp. 143–164). Elsevier. https://doi.org/10.1016/B978-0-12-819813-1.00006-2

Mesquita, T. J. B., Sargo, C. R., Fuzer, J. R., Paredes, S. A. H., Giordano, R. D. C., Horta, A. C. L., & Zangirolami, T. C. (2019). Metabolic fluxes-oriented control of bioreactors: A novel approach to tune micro-Aeration and substrate feeding in fermentations. *Microbial Cell Factories*, *18*(1), 1–17. https://doi.org/10.1186/s12934-019-1198-6

Minning, S., Schmidt-Dannert, C., & Schmid, R. D. (1998). Functional expression of *Rhizopus oryzae* lipase in *Pichia pastoris*: high-level production and some properties. *Journal of Biotechnology*, *66*(2-3), 147-156.

Minning, S., Serrano, A., Ferrer, P., Solá, C., Schmid, R. D., & Valero, F. (2001). Optimization of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *Journal of Biotechnology*, *86*(1), 59–70. https://doi.org/10.1016/S0168-1656(00)00402-8

Mowbray, M., Savage, T., Wu, C., Song, Z., Cho, B. A., Del Rio-Chanona, E. A., & Zhang, D. (2021). Machine learning for biochemical engineering: A review. *Biochemical Engineering Journal*, *172*, 108054.

Neubauer, P., Häggström, L., & Enfors, S. O. (1995). Influence of substrate oscillations on acetate formation and growth yield in *Escherichia coli* glucose limited fed-batch cultivations. *Biotechnology and Bioengineering*, *47*(2), 139-146.

Neubauer, P., & Junne, S. (2016). Scale-Up and Scale-Down Methodologies for Bioreactors. *Bioreactors*, 323–354. https://doi.org/10.1002/9783527683369.ch11

Nieto-Taype, M. A. (2020). Combining bioprocess and strain engineering strategies as efficient tools for the optimization of recombinant protein production in *Pichia pastoris*.

Nieto-Taype, M. A., Garcia-Ortega, X., Albiol, J., Montesinos-Seguí, J. L., & Valero, F. (2020b). Continuous Cultivation as a Tool Toward the Rational Bioprocess Development With *Pichia Pastoris* Cell Factory. *Frontiers in Bioengineering and Biotechnology*, *8*(June), 1–21. https://doi.org/10.3389/fbioe.2020.00632

Nieto-Taype, M. A., Garrigós-Martínez, J., Sánchez-Farrando, M., Valero, F., Garcia-Ortega, X., & Montesinos-Seguí, J. L. (2020a). Rationale-based selection of optimal operating strategies and gene dosage impact on recombinant protein production in *Komagataella phaffii* (*Pichia pastoris*). *Microbial Biotechnology*, *13*(2), 315–327. https://doi.org/10.1111/1751-7915.13498

Nishu, N., Masih, S., Kamal, S., Jain, P., & Khan, Z. K. (2020). Transgenic animals in research and industry. In *Animal Biotechnology: Models in Discovery and Translation* (pp. 463–480). Elsevier. https://doi.org/10.1016/B978-0-12-811710-1.00021-5

Ogata, K., Nishikawa, H., & Ohsugi, M. (1969). A yeast capable of using methanol. *Agricultural and Biological Chemistry*.

Ondracka, A., Gasset, A., García-Ortega, X., Hubmayr, D., van Wijngaarden, J. B. G., Montesinos-Seguí, J. L., Valero, F., & Manzano, T. (2022). CPV of the Future: AI-powered continued process verification for bioreactor processes. *PDA Journal of Pharmaceutical Science and Technology*, pdajpst.2021.012665. https://doi.org/10.5731/pdajpst.2021.012665

Onyeaka, H., Nienow, A. W., & Hewitt, C. J. (2003). Further studies related to the scale-up of high cell density *Escherichia coli* fed-batch fermentations: the additional effect of a changing microenvironment when using aqueous ammonia to control pH. *Biotechnology and Bioengineering*, *84*(4), 474-484.

Pirt, S. J. (1982). Maintenance energy: a general model for energy-limited and energy-sufficient growth. *Archives of microbiology*, *133*, 300-302.

Ponte, X., Montesinos-Seguí, J. L., & Valero, F. (2016). Bioprocess efficiency in *Rhizopus oryzae* lipase production by *Pichia pastoris* under the control of PAOX1 is oxygen tension dependent. *Process Biochemistry*, *51*(12), 1954–1963. https://doi.org/10.1016/j.procbio.2016.08.030

Porter, J. L., Ukhairul, R., Rusli, A., & Ollis, D. L. (2016). *Directed Evolution of Enzymes for Industrial Biocatalysis*. www.chembiochem.org

Potgieter, T. I., Kersey, S. D., Mallem, M. R., Nylen, A. C., & D'Anjou, M. (2010). Antibody expression kinetics in glycoengineered *Pichia pastoris*. *Biotechnology and Bioengineering*, *106*(6), 918–927. https://doi.org/10.1002/bit.22756

Potvin, G., Zhang, Z., Defela, A., & Lam, H. (2016). Screening of Alternative Carbon Sources for Recombinant Protein Production in *Pichia pastoris*. *International Journal of Chemical Reactor Engineering*, *14*(1), 251–257. https://doi.org/10.1515/ijcre-2015-0092

Premier, G. C., Kim, J. R., Michie, I., Dinsdale, R. M., & Guwy, A. J. (2011). Automatic control of load increases power and efficiency in a microbial fuel cell. *Journal of Power Sources*, *196*(4), 2013–2019. https://doi.org/10.1016/j.jpowsour.2010.09.071

Prielhofer, R., Cartwright, S. P., Graf, A. B., Valli, M., Bill, R. M., Mattanovich, D., & Gasser, B. (2015). *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. *BMC Genomics*, *16*(1), 1–17. https://doi.org/10.1186/s12864-015-1393-8

Prielhofer, R., Reichinger, M., Wagner, N., Claes, K., Kiziak, C., Gasser, B., & Mattanovich, D. (2018). Superior protein titers in half the fermentation time: Promoter and process engineering for the glucose-regulated GTH1 promoter of *Pichia pastoris*. *Biotechnology and Bioengineering*, *115*(10), 2479–2488. https://doi.org/10.1002/bit.26800

Puxbaum, V., Mattanovich, D., & Gasser, B. (2015). Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. *Applied Microbiology and Biotechnology*, *99*(7), 2925–2938. https://doi.org/10.1007/s00253-015-6470-z

Rabbani, G., Ahmad, E., Khan, M. V., Ashraf, M. T., Bhat, R., & Khan, R. H. (2015). Impact of structural stability of cold adapted *Candida antarctica* lipase B (CaLB): In relation to pH, chemical and thermal denaturation. *RSC Advances*, *5*(26), 20115–20131. https://doi.org/10.1039/c4ra17093h

Randek, J., & Mandenius, C. F. (2018). On-line soft sensing in upstream bioprocessing. *Critical Reviews in Biotechnology*, *38*(1), 106–121. https://doi.org/10.1080/07388551.2017.1312271

Ranjan, A. P., & Gomes, J. (2009). Simultaneous dissolved oxygen and glucose regulation in fed-batch methionine production using decoupled input-output linearizing control. *Journal of Process Control*, *19*(4), 664–677. https://doi.org/10.1016/j.jprocont.2008.07.008

Raschmanová, H., Zamora, I., Borčinová, M., Meier, P., Weninger, A., Mächler, D., Glieder, A., Melzoch, K., Knejzlík, Z., & Kovar, K. (2019). Single-Cell Approach to Monitor the Unfolded Protein Response During Biotechnological Processes With *Pichia pastoris*. *Frontiers in Microbiology*, *10*(FEB), 1–18. https://doi.org/10.3389/fmicb.2019.00335

Rathore, A. S., Mishra, S., Nikita, S., & Priyanka, P. (2021). Bioprocess control: Current progress and future perspectives. *Life*, *11*(6). https://doi.org/10.3390/life11060557

Read, E. K., Shah, R. B., Riley, B. S., Park, J. T., Brorson, K. A., & Rathore, A. S. (2010). Process analytical technology (PAT) for biopharmaceutical products: Part II. Concepts and applications. *Biotechnology and Bioengineering*, *105*(2), 285-295.

Rigoldi, F., Donini, S., Redaelli, A., Parisini, E., & Gautieri, A. (2018). Review: Engineering of thermostable enzymes for industrial applications. In *APL Bioengineering* (Vol. 2, Issue 1). American Institute of Physics Inc. https://doi.org/10.1063/1.4997367

Ritacco, F. V., Wu, Y., & Khetan, A. (2018). Cell culture media for recombinant protein expression in Chinese hamster ovary (CHO) cells: History, key components, and optimization strategies. In *Biotechnology Progress* (Vol. 34, Issue 6, pp. 1407–1426). John Wiley and Sons Inc. https://doi.org/10.1002/btpr.2706

Sagmeister, P., Wechselberger, P., Jazini, M., Meitz, A., Langemann, T., & Herwig, C. (2013). Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development. *Chemical Engineering Science*, *96*, 190–198. https://doi.org/10.1016/j.ces.2013.02.069

Sales, A., Gasset, A., Requena-Moreno, G., Valero, F., Montesinos-Seguí, J. L., Garcia-Ortega, X. (2023). Synergic kinetic and physiological stress control to improve bioprocess efficiency of *Komagataella phaffii* recombinant protein production bioprocesses. *Microbial Biotechnology*. (**under revision**)

Schillberg, S., & Spiegel, H. (2022). *Recombinant Proteins in Plants Methods and Protocols Methods in Molecular Biology*. http://www.springer.com/series/7651

Sharma, R., Chisti, Y., & Banerjee, U. C. (2001). Production, purification, characterization, and applications of lipases. *Biotechnology Advances*, *19*(8), 627–662. https://doi.org/10.1016/S0734-9750(01)00086-6

Siehl, D. L., Castle, L. A., Gorton, R., Chen, Y. H., Bertain, S., Cho, H. J., Keenan, R., Liu, D., & Lassner, M. W. (2005). Evolution of a microbial acetyltransferase for modification of glyphosate: A novel tolerance strategy. *Pest Management Science*, *61*(3), 235–240. https://doi.org/10.1002/ps.1014

Simutis, R., & Lübbert, A. (2015). Bioreactor control improves bioprocess performance. *Biotechnology Journal*, *10*(8), 1115–1130. https://doi.org/10.1002/biot.201500016

Singh, A., & Singhal, B. (2022). *Role of Machine Learning in Bioprocess Engineering: Current Perspectives and Future Directions*.

Soetaert, W., & Vandamme, E. (2006). The impact of industrial biotechnology. In *Biotechnology Journal* (Vol. 1, Issues 7–8, pp. 756–769). https://doi.org/10.1002/biot.200600066

Sørensen, H. (2010). Towards universal systems for recombinant gene expression. *Microbial Cell Factories*, *9*(1), 27. https://doi.org/10.1186/1475-2859-9-27

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. *Journal of Biotechnology*, *150*(4), 519–529. https://doi.org/10.1016/j.jbiotec.2010.09.957

Stanke, M., & Hitzmann, B. (2013). Automatic control of bioprocesses. *Measurement, Monitoring, Modelling and Control of Bioprocesses*, 35-63.

Stephanopoulos, G., Aristidou, A. A., & Nielsen, J. (1998). *Metabolic engineering: principles and methodologies*.

Sturmberger, L., Chappell, T., Geier, M., Krainer, F., Day, K. J., Vide, U., Trstenjak, S., Schiefer, A., Richardson, T., Soriaga, L., Darnhofer, B., Birner-Gruenberger, R., Glick, B. S., Tolstorukov, I., Cregg, J., Madden, K., & Glieder, A. (2016). Refined *Pichia pastoris* reference genome sequence. *Journal of Biotechnology*, *235*, 121–131. https://doi.org/10.1016/j.jbiotec.2016.04.023

Takors, R. (2012). Scale-up of microbial processes: Impacts, tools and open questions. *Journal of Biotechnology*, *160*(1–2), 3–9. https://doi.org/10.1016/j.jbiotec.2011.12.010

Tomàs-Gamisans, M., Andrade, C. C. P., Maresca, F., Monforte, S., Ferrer, P., & Albiol, J. (2020). Redox Engineering by Ectopic Overexpression of NADH Kinase. *86*(6), 1–15. https://doi.org/https://doi.org/10.1128/AEM .02038-19

Valero, F. (2013). Bioprocess Engineering of *Pichia pastoris*, an Exciting Host Eukaryotic Cell Expression System. *Protein Engineering - Technology and Application*, 3–32. https://doi.org/10.5772/56407

Van der Heijden, R. T. J. M., Romein, B., Heijnen, J. J., Hellinga, C., & Luyben, K. Ch. A. M. (1994a). Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. *Biotechnology and Bioengineering*, *43*(1), 11–20. https://doi.org/10.1002/bit.260430104

Van der Heijden, R. T. J. M., Romein, B., Heijnen, J. J., Hellinga, C., & Luyben, K. C. A. (1994b). Linear constrain relations in biochemical reaction systems III. Sequential application of data reconciliation for sensitive detection of systematic errors. *Biotechnology and Bioengineering*, *44*(7), 781-791.

Vanleeuw, E., Winderickx, S., Thevissen, K., Lagrain, B., Dusselier, M., Cammue, B. P. A., & Sels, B. F. (2019). Substrate-Specificity of *Candida rugosa* Lipase and Its Industrial Application. *ACS Sustainable Chemistry and Engineering*, *7*(19), 15828–15844. https://doi.org/10.1021/acssuschemeng.9b03257

Vardar-Sukan, F. (1998). Foaming: Consequences, prevention and destruction. *Biotechnology Advances*, *16*(5–6), 913–948. https://doi.org/10.1016/S0734-9750(98)00010-X

Veloso, A. C., & Ferreira, E. C. (2017). Online Analysis for Industrial Bioprocesses: Broth Analysis. In *Current Developments in Biotechnology and Bioengineering: Bioprocesses, Bioreactors and Controls*. Elsevier B.V. https://doi.org/10.1016/B978-0-444-63663-8.00023-9

Vemuri, G. N., Eiteman, M. A., Mcewen, J. E., Olsson, L., & Nielsen, J. (2007). Increasing NADH oxidation reduces overflow metabolism in *Saccharomyces cerevisiae*. In *PNAS February* (Vol. 13). www.pnas.orgcgidoi10.1073pnas.0607469104

Verheijen, P. (2009). Data Reconciliation and Error Detection. In *The Metabolic Pathway Engineering Handbook* (pp. 8-1-8–14). CRC Press. https://doi.org/10.1201/9781439802977.ch8

Vilanova, R. (2008). IMC based Robust PID design: Tuning guidelines and automatic tuning. *Journal of Process Control*, *18*(1), 61-70.

Vogl, T., & Glieder, A. (2013). Regulation of *Pichia pastoris* promoters and its consequences for protein production. *New Biotechnology*, *30*(4), 385–404. https://doi.org/10.1016/j.nbt.2012.11.010

Wang, G., Haringa, C., Tang, W., Noorman, H., Chu, J., Zhuang, Y., & Zhang, S. (2020). Coupled metabolic-hydrodynamic modeling enabling rational scale-up of industrial bioprocesses. *Biotechnology and Bioengineering*, *117*(3), 844–867. https://doi.org/10.1002/bit.27243

Wang, G., Huang, M., & Nielsen, J. (2017). Exploring the potential of *Saccharomyces cerevisiae* for biopharmaceutical protein production. In *Current Opinion in Biotechnology* (Vol. 48, pp. 77–84). Elsevier Ltd. https://doi.org/10.1016/j.copbio.2017.03.017

Wang, N. S., & Stephanopoulos, G. (1983). Application of macroscopic balances to the identification of gross measurement errors. *Biotechnology and Bioengineering*, *25*(9), 2177–2208. https://doi.org/10.1002/bit.260250906

Wang, Z., Tan, T., & Song, J. (2007). Effect of amino acids addition and feedback control strategies on the high-cell-density cultivation of Saccharomyces cerevisiae for glutathione production. *Process Biochemistry*, *42*(1), 108–111. https://doi.org/10.1016/j.procbio.2006.07.008

Waterham, H. R., Digan, M. E., Koutz, P. J., Lair, S. V., & Cregg, J. M. (1997). Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene*, *186*(1), 37–44. https://doi.org/10.1016/S0378-1119(96)00675-0

Wehbe, O., Yaman, O. U., & Çalık, P. (2020). Ethanol fed-batch bioreactor operation to enhance therapeutic protein production in *Pichia pastoris* under hybrid-architectured ADH2 promoter. *Biochemical Engineering Journal*, *164*(September), 107782. https://doi.org/10.1016/j.bej.2020.107782

Westbrook, A. W., Miscevic, D., Kilpatrick, S., Bruder, M. R., Moo-Young, M., & Chou, C. P. (2019). Strain engineering for microbial production of value-added chemicals and fuels from glycerol. In *Biotechnology Advances* (Vol. 37, Issue 4, pp. 538–568). Elsevier Inc. https://doi.org/10.1016/j.biotechadv.2018.10.006

Xia, J., Wang, G., Lin, J., Wang, Y., Chu, J., Zhuang, Y., & Zhang, S. (2015). Advances and Practices of Bioprocess Scale-up. In: Bao J., Ye Q., Zhong JJ. (eds) Bioreactor Engineering Research and Industrial Applications II. *Advances in Biochemical Engineering/Biotechnology*, *152*(July 2015), 127–141. https://doi.org/10.1007/10

Yang, K. K., Wu, Z., & Arnold, F. H. (2019). Machine-learning-guided directed evolution for protein engineering. In *Nature Methods* (Vol. 16, Issue 8, pp. 687–694). Nature Publishing Group. https://doi.org/10.1038/s41592-019-0496-6

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

Yao, T., & Asayama, Y. (2017). Animal-cell culture media: History, characteristics, and current issues. In *Reproductive Medicine and Biology* (Vol. 16, Issue 2, pp. 99–117). John Wiley and Sons Ltd. https://doi.org/10.1002/rmb2.12024

Yin, J., Li, G., Ren, X., & Herrler, G. (2007). Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. In *Journal of Biotechnology* (Vol. 127, Issue 3, pp. 335–347). https://doi.org/10.1016/j.jbiotec.2006.07.012

Zahrl, R. J., Peña, D. A., Mattanovich, D., & Gasser, B. (2017). Systems biotechnology for protein production in *Pichia pastoris*. *FEMS Yeast Research*, *17*(7), 1–15. https://doi.org/10.1093/femsyr/fox068

Zavec, D., Gasser, B., & Mattanovich, D. (2020). Characterization of methanol utilization negative *Pichia pastoris* for secreted protein production: New cultivation strategies for current and future applications. *Biotechnology and Bioengineering*, *117*(5), 1394–1405. https://doi.org/10.1002/bit.27303

Zhang, A. L., Luo, J. X., Zhang, T. Y., Pan, Y. W., Tan, Y. H., Fu, C. Y., & Tu, F. Z. (2009). Recent advances on the GAP promoter derived expression system of *Pichia pastoris*. *Molecular Biology Reports*, *36*(6), 1611–1619. https://doi.org/10.1007/s11033-008-9359-4

Zheng, J., Guo, N., Lin, F. Lai, Wu, L. Shuang, & Zhou, H. Bo. (2014). Screening of multi-copy mannanase recombinants of *Pichia pastoris* based on colony size. *World Journal of Microbiology and Biotechnology*, *30*(2), 579–584. https://doi.org/10.1007/s11274-013-1479-x

Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge Intelligence: Paving the Last Mile of Artificial Intelligence With Edge Computing. *Proceedings of the IEEE*, 1–25. https://doi.org/10.1109/JPROC.2019.2918951

Zhu, T., Guo, M., Sun, C., Qian, J., Zhuang, Y., Chu, J., & Zhang, S. (2009). A systematical investigation on the genetic stability of multi-copy *Pichia pastoris* strains. *Biotechnology Letters*, *31*(5), 679–684. https://doi.org/10.1007/s10529-009-9917-4

Zhu, T., Guo, M., Tang, Z., Zhang, M., Zhuang, Y., Chu, J., & Zhang, S. (2009). Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast *Pichia pastoris*. *Journal of Applied Microbiology*, *107*(3), 954–963. https://doi.org/10.1111/j.1365-2672.2009.04279.x