Universidade de São Paulo

Escola Politécnica

RAFAEL DA GAMA FERREIRA

Development and *in silico* evaluation of an expression platform based on *E. coli* for the production of a recombinant beta-glucosidase

São Paulo

2019

RAFAEL DA GAMA FERREIRA

Development and *in silico* evaluation of an expression platform based on *E. coli* for the production of a recombinant beta-glucosidase

Corrected Version

(The original version can be found in the school that harbors the Graduate Program)

Ph. D. Thesis presented to the Graduate Program in Chemical Engineering at Escola Politécnica da Universidade de São Paulo to obtain the degree of Doctor of Science.

Concentration area: Bioprocess Engineering

Advisor: Prof. Dr. Adriano Rodrigues Azzoni

São Paulo

2019

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catalogação-na-publicação

Ferreira, Rafael da Gama

Development and in silico evaluation of an expression platform based on E. coli for the production of a recombinant beta-glucosidase / R. G. Ferreira -- versão corr. -- São Paulo, 2019.

188 f.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Química.

1.Biotecnologia 2.Bioprocessos 3.Biocombustíveis 4.Enzimas 5.Microbiologia Industrial I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Química II.t.

FERREIRA, R. G. **Development and *in silico* evaluation of an expression platform based on *E. coli* for the production of a recombinant beta-glucosidase.** Corrected version. 2019. 188 ll. Ph. D. Thesis (Chemical Engineering) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2019.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

AGRADECIMENTOS

Agradeço primeiramente ao prof. Adriano, que abriu as portas do seu laboratório e da Biologia Molecular para mim, que muito me ensinou ao longo de todo o trabalho e que, mais que um orientador, sempre foi um verdadeiro amigo.

Em segundo lugar, agradeço à profª Sindélia, que também me acolheu em seu laboratório de forma generosa e que, se não foi coorientadora no papel, certamente o foi na prática, contribuindo imensamente do início ao fim deste trabalho.

Agradeço também a todos os colegas do GEnBio e agregados que tive a oportunidade de conhecer ao longo destes 66 meses de doutoramento, seja pelo auxílio direto no trabalho, pelas discussões proveitosas ou pela amizade sincera: Bianca, Bruno, Camila, Carla, Daniel, Daniela, Davi, Dielle, Edmar, Fernando, Fernanda, Gabriel, Gabriela, Georges, Guilherme Fabri, Guilherme Del Padre, Isabella, Júlia, Karina, Letícia Parizotto, Letícia Veloso, Luan, Luís, Maira, Margareth, Matheus, Pedro Henrique, Priscila, Rafael Risnik, Sara, Sebastian, Suzanna, Thais, Thiago e Wesley, muito obrigado! Agradeço igualmente aos colegas do Laboratório de Hidrolases Bacterianas do CTBE, Ana, Karla e Rafael Alves, por terem feito com que me sentisse em casa tão rapidamente e por terem me ajudado sempre que precisei.

Meu muito obrigado também às técnicas de laboratório do GEnBio, Andrea e Orlinda, pela amizade duradoura e assistência cotidiana no trabalho; da mesma forma, agradeço aos técnicos do Lab. de Biossíntese de Hidrolases Bacterianas, Aline, Mateus e Márcia, pela hospitalidade e pelo auxílio indispensável na parte experimental. Em particular, sou grato à Aline por ter me ensinado as técnicas básicas de Biologia Molecular de maneira gentil e paciente.

Sou grato ainda ao prof. Thiago, pela amizade e cortesia de ceder seu laboratório para a realização do ensaio em biorreator; à profª Rosane, do IPT, pela simpatia e gentileza de nos conferir acesso ao software SuperPro Designer; ao prof. Aldo, pelo invariável bom humor e assistência em diversas ocasiões; e ao Dr. Demetri Petrides, pelo valioso auxílio na utilização do programa SuperPro Designer e revisão generosa do nosso flowsheet.

Agradeço em especial à minha família e, em particular, ao meu pai, à minha mãe e à minha irmã, pelo amor e apoio incondicionais, sempre. Sem a educação que os meus pais me proporcionaram e os valores que me transmitiram, jamais teria chegado até aqui.

Além disso, manifesto meu profundo agradecimento à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), que me confiou uma bolsa de estudos e verba de reserva técnica indispensáveis ao desenvolvimento deste trabalho de doutorado (Processo n° 2014/13974-6). Agradeço também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Código Financeiro 001, pela bolsa que me concedeu (Processo n° 3300201) até a sua substituição pela bolsa da FAPESP e também pelo apoio prestado ao trabalho através do programa CAPES-PROEX. Agradeço igualmente ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo amparo fornecido por meio de projetos concedidos ao prof. Adriano Rodrigues Azzoni, meu orientador (Processos n° 444412/2014-0 e n° 307739/2015-5). Por fim, agradeço ao contribuinte paulista e ao contribuinte brasileiro que, em última análise, financiaram este trabalho.

ABSTRACT

FERREIRA, R. G. **Development and *in silico* evaluation of an expression platform based on *E. coli* for the production of a recombinant beta-glucosidase.** Corrected version. 2019. 188 ll. Ph. D. Thesis (Chemical Engineering) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2019.

The enzymatic conversion of lignocellulosic biomass into fermentable sugars is a promising approach for producing renewable fuels and chemicals. However, the cost of the fungal enzymes usually employed in this process remains a significant bottleneck for manufacturing low value-added products from biomass. A potential route to increase hydrolysis yield, and thereby to reduce hydrolysis cost, would be to supplement the fungal enzymes with their lacking enzymatic activities, such as β-glucosidase. To produce such enzymes at a low cost, the bacterium *Escherichia coli* is a strong contender, owing to its ability to grow rapidly on simple and inexpensive media, and to achieve high levels of productivity. Nevertheless, there is hardly any techno-economic analysis of low-value protein production using *E. coli* in the literature, and, more generally, there are very few techno-economic analyses of low-value protein production ever reported, with the exception of cellulase production by *Trichoderma reesei*. In particular, the biotechnological application of recombinant *E. coli* platforms equipped with toxin-antitoxin systems to ensure plasmid stability remains largely unexplored, and its economic impact, unknown. As such, this work presents a comprehensive techno-economic analysis of the industrial production of a low-cost enzyme (β-glucosidase) using both *E. coli* BL21(DE3) and *E. coli* SE1, a modified BL21(DE3) strain equipped with a toxin-antitoxin system for plasmid maintenance. Moreover, this study describes the actual cloning and expression of a β-glucosidase enzyme into *E. coli* BL21(DE3) and *E. coli* SE1, and the development of a novel inoculum production scheme that exploits the features of the SE1 strain, based on repeatedly recycling a fraction of the inoculum cells. The results of the techno-economic analysis project an enzyme production cost of 316 US$/kg in the baseline scenario, which is considerably higher than the values reported in the literature for the fungal cocktails. The facility-dependent cost, which is strongly associated with the cost of equipment, accounts for roughly half of the estimated cost, while the cost of raw

materials, especially IPTG and glucose, and the cost of consumables are all quite significant. However, the simulation of multiple scenarios and optimization measures suggest that the enzyme cost can be substantially reduced on many fronts, such as: substituting the carbon source for cheaper alternatives; reducing the amount of IPTG used for induction; using an *E. coli* strain capable of extracellular production; or eliminating the steps of concentration and stabilization of the enzyme, in the case of on-site enzyme utilization. Developing *E. coli* strains capable of high rEnzyme volumetric productivities can also significantly reduce the cost of the enzyme, up to approximately 135 US$/kg in the scenario of highest productivity. In addition, based on the experimental results with the *E. coli* SE1 system, an inoculum recycle strategy that avoids the need of an extensive seed train was simulated, resulting in a significant reduction of the enzyme cost. Finally, the combination of multiple process improvements could lead to an enzyme cost near 20 US$/kg of protein, which comes close to the cost of fungal cellulases and demonstrates the great biotechnological potential of recombinant *E. coli* platforms.

Keywords: Cellulases. Beta-Glucosidase. Plasmid Stability. Toxin-Antitoxin. Bioprocess Simulation. Techno-Economic Analysis. Inoculum Production.

RESUMO

FERREIRA, R. G. **Desenvolvimento e avaliação *in silico* de uma plataforma de expressão baseada em *E. coli* para a produção de beta-glicosidase recombinante.** Versão corr. 2019. 188 f. Tese (Doutorado em Engenharia Química) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2019.

A conversão enzimática de biomassa lignocelulósica em açúcares fermentescíveis é uma via promissora para a produção de combustíveis e produtos químicos renováveis. No entanto, o custo das enzimas fúngicas usualmente empregadas nesse processo permanece um gargalo significativo para a fabricação de produtos de baixo valor agregado a partir de biomassa. Uma possível estratégia para aumentar o rendimento da hidrólise e, assim, reduzir seu custo, seria suplementar as enzimas fúngicas com suas atividades enzimáticas deficientes, tais como a enzima β-glicosidase. Para produzir tais enzimas a um baixo custo, a bactéria *Escherichia coli* é uma forte candidata, dada a sua capacidade de crescer rapidamente em meios simples e baratos e de alcançar altos níveis de produtividade. No entanto, na literatura quase não há análises técnico-econômicas de produção de proteínas de baixo valor agregado utilizando *E. coli* e, de forma mais geral, há muito poucas análises técnico-econômicas de produção de proteínas de baixo valor agregado publicadas, com exceção da produção de celulases por *Trichoderma reesei*. Em particular, a aplicação biotecnológica de plataformas recombinantes baseadas em *E. coli* dotadas de sistemas toxina-antitoxina para garantir a estabilidade plasmidial segue em larga medida inexplorada, e seu impacto econômico, desconhecido. Assim, este trabalho apresenta uma análise técnico-econômica abrangente da produção industrial de uma enzima de baixo custo (β-glicosidase) usando *E. coli* BL21 (DE3) e *E. coli* SE1, uma cepa de BL21 (DE3) modificada que possui um sistema toxina-antitoxina para manutenção plasmidial. Além disso, este estudo descreve a clonagem e expressão de uma β-glicosidase em *E. coli* BL21 (DE3) e *E. coli* SE1, assim como o desenvolvimento de um novo método de produção de inóculo que tira proveito das peculiaridades da linhagem SE1, baseado em reciclar repetidamente uma fração das células do inóculo. Os resultados da análise técnico-econômica apontam para um custo de produção da enzima de 316 US$/kg no cenário-base, valor consideravelmente superior àqueles relatados na literatura para os coquetéis

fúngicos. Os custos de overhead da planta, que estão fortemente associados ao custo de aquisição dos equipamentos, são responsáveis por aproximadamente metade do custo total, enquanto o custo de matérias-primas, especialmente IPTG e glicose, e o custo de consumíveis são bastante significativos. Porém, a simulação de múltiplos cenários e medidas de otimização sugerem que o custo da enzima pode ser substancialmente reduzido em muitas frentes, tais como: a substituição da fonte de carbono por alternativas mais baratas; a redução da quantidade de IPTG usado para indução; a utilização de cepas capazes de produzir a enzima extracelularmente; ou a eliminação das etapas de concentração e estabilização da enzima, em caso de utilização da enzima *in situ*. O desenvolvimento de cepas de *E. coli* capazes de atingir altas produtividades volumétricas de rEnzima também pode reduzir significativamente o seu custo, chegando a US$ 135/kg no cenário de maior produtividade. Com base nos resultados experimentais com a linhagem *E. coli* SE1, uma estratégia de reciclagem de inóculo que evita a necessidade de um extenso trem de inoculação também foi simulada, gerando significativa diminuição do custo da enzima. Por fim, a combinação de múltiplas melhorias no processo poderia levar a um custo de enzima em torno de 20 US$/kg de proteína, valor que se aproxima do custo das celulases fúngicas e que demonstra o grande potencial biotecnológico de plataformas de expressão baseadas em *E. coli* recombinante.

Palavras-chave: Celulases. Beta-Glicosidase. Estabilidade Plasmidial. Toxina-Antitoxina. Simulação de Bioprocessos. Análise Técnico-Econômica. Produção de Inóculo.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

AFEX	ammonia fiber expansion
BGL	β-glucosidase
BL21(DE3)/BGL	*E. coli* BL21(DE3) carrying the plasmid pStaby1.2-TpBgl1
CAPEX	Capital Expenditures
CBH#	cellobiohydrolase type #
CBM	Carbohydrate-binding module
CBP	Consolidate Bioprocessing
CS	Carbon-steel alloy
CTBE	Brazilian Bioethanol Science and Technology Laboratory
DCW	Dry Cell Weight
DFC	Direct Fixed Capital
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EG	Endoglucanase
FS#	Feeding Solution #
GH#	Glycosyl Hydrolase family #
HEPES	4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid
IPT	Institute of Technological Research
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny Broth medium
LB-amp	Lysogeny Broth medium with 50 mg/L of ampicillin sodium
LB-kan	Lysogeny Broth medium with 30 mg/L of kanamycin sulfate
LHW	liquid hot water
LPMO	Lytic Polysaccharide Monooxygenase
MESP	Minimum Ethanol Selling Price
MW	Molecular Weight
NREL	National Renewable Energy Laboratory

OD	optical density
OD_{600}	optical density read at a wavelength of 600 nm
OPEX	Operating Expenditures
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
pDNA	plasmid DNA
PHB	poly-3-hydroxybutyrate
PMS	Process Modeling and Simulation
pNP	*p*-nitrophenol
pNPG	*p*-nitrophenyl glucopyranoside
PSK	post-segregational killing
rEnzyme	recombinant enzyme
RNAi	RNA interference
ROI	return on investment
rProtein	recombinant protein
SDS	Sodium Dodecyl Sulfate
SE1/BGL	*E. coli* SE1 carrying the plasmid pStaby1.2-TpBgl1
SHF	Separate Hydrolysis and Fermentation
SmC	Submerged Culture
SS304	Stainless Steel of grade 304
SS316	Stainless Steel of grade 316
SSC	Solid-State Culture
SSCF	Simultaneous Saccharification and Cofermentation
SSF	Simultaneous Saccharification and Fermentation
TA	toxin-antitoxin
TAE	tris-acetate-EDTA
Tris	tris(hydroxymethyl)aminomethane
TSAP	Thermosensitive Alkaline Phosphatase
UPC	Unit Production Cost
USP	Universidade de São Paulo

LIST OF SYMBOLS

k_{cat}	turnover number
K_i	inhibitor constant
K_M	Michaelis constant
N	number of colony-forming units per plate
T	temperature
t	time
X	cell concentration in terms of dry cell weight
$Y_{X/S}$	yield of dry cell weight on the limiting substrate
θ	volume fraction of the seed fermenter transferred to the main fermenter
λ	wavelength
μ	specific growth rate
φ	ratio of recombinant protein over total soluble proteins
ψ	plasmid stability

LIST OF UNITS

CBU	activity unit of β-glucosidase on cellobiose
CFU	colony-forming unit
EUR	euro
FPU	activity unit of cellulases on filter-paper
gDW	gram dry weight
IDR	Indonesian rupiah
INR	Indian rupee
R$	Brazilian real
SEK	Swedish krona

ton	metric ton (1000 kg)
U	enzyme activity unit (1 µmol of converted substrate per minute, under specified conditions)
US$	United States dollar
VVM	gas volume per unit of liquid volume per minute

TABLE OF CONTENTS

1 Introduction

The continuous depletion of oil reserves, as well as the pressing concerns about global warming, have sparked interest for renewable fuels and chemicals in the last decades (1). Although several pathways to obtain such fuels and chemicals have been studied in both academia and industry, many of them are based on the enzymatic deconstruction of lignocellulosic biomass, such as agricultural wastes, and subsequent biological transformation into the desired compound (2,3).

Given that lignocellulosic biomass is a complex material, composed of three main fractions (cellulose, hemicellulose and lignin), various enzymes are necessary for its effective degradation (4,5). The breakdown of the cellulose fraction is particularly attractive for subsequent microbiological processes such as ethanol fermentation, given that it generates glucose, a carbon and energy source widely used in Nature. It can be accomplished by the concerted efforts of three major types of hydrolytic enzymes: endoglucanases, cellobiohydrolases and β-glucosidases, collectively called "cellulases" (6). This process of biochemical decomposition of cellulosic biomass, called enzymatic hydrolysis, is not economically viable yet, largely owing to the cost of producing these enzymes (7,8).

Nevertheless, only a few techno-economic studies that evaluate the production cost of cellulases have been reported in the literature. Those that did delve into the enzyme cost have concentrated on the economics of the cellulase mixture produced by filamentous fungus *Trichoderma reesei* (7,9–11), which is generally regarded as the most efficient producer of cellulases (12) and, accordingly, stands as the major producer of cellulases in industry today (13–15). However, it is widely known that the composition of the enzymatic cocktail secreted by *T. reesei* is not optimal for the industrial degradation of cellulosic biomass in terms of cellulase activities, notably owing to its low amount of β-glucosidase (EC 3.2.1.21) activity (16–18). For this reason, supplementing the fungal cocktail with lacking enzyme activities is crucial for enzyme cost reduction (19), provided that the increase in the hydrolysis yield outweighs the cost of the supplementing enzymes. Moreover, producing these supplementing enzymes on-site (that is, in the same location as the fermentation plant) may be more economical than off-site, since the former case

avoids transportation and formulation costs (9,20). Yet, to our knowledge, no techno-economic analysis of such an enzyme, particularly one produced on-site, has been carried out to date. In fact, there are surprisingly few techno-economic analyses of microbial processes to produce any low or intermediate value-added protein other than fungal cellulases. Furthermore, given that approximately 90% of all industrial enzymes are produced by recombinant microorganisms (21,22), it is also remarkable that so few analyses of protein production by recombinant microorganisms have been published.

Escherichia coli is the most common host for recombinant expression of cellulases in academia (6) and one of the most frequently used bacteria to produce recombinant proteins (rProteins) in general (23–29). A major advantage of *E. coli* is its ability to grow rapidly and achieve high cell densities on simple and inexpensive media, as well as attain high titers of recombinant protein. Another crucial benefit of *E. coli* is the large body of accumulated knowledge concerning the physiology and genetics of that species, together with the great number of tools and techniques for its genetic manipulation. In fact, overexpression of rProteins in *E. coli* typically relies on the use of artificial, multicopy plasmids.

However, the use of plasmid DNA for rProtein production faces a significant challenge: if left unchecked, recombinant cells tend to lose their plasmids and, as a result, rProtein productivity declines precipitously. This phenomenon occurs because the very expression of rProteins, and even plasmid maintenance and replication by themselves, constitute a metabolic burden for the microorganism, so that cells that lose plasmid molecules during division tend to outcompete plasmid-harboring cells (25,30,31). To prevent this problem, plasmid stability is usually ensured by adding an antibiotic to the culture medium while having the corresponding antibiotic resistance gene inserted into the plasmid (23,27,32). Despite being a simple and practical method, particularly for academic research, this system is not without its downsides: it risks spreading antibiotic-resistance genes to pathogens (23,27,33) and faces decreasing performance over the course of cell culture (33,34). Potential antibiotic-free alternatives for plasmid maintenance are toxin-antitoxin systems, which are based on two genes: one that codes for a lethal toxic protein (the “toxin”), which may be present in the plasmid or the chromosome; and another gene that gives rise to a protein (or a non-coding RNA) that counteracts the toxin (the “antitoxin”), inserted into the plasmid (35–39). As such, cells

that receive no plasmid copy upon division immediately die, thus ensuring high levels of plasmid stability over time. In addition, such plasmid addiction systems provide interesting opportunities for inoculum production of recombinant strains, given that, in theory, cells could be passaged repeatedly without losing their plasmids.

Notwithstanding the numerous advantages of producing proteins with recombinant *E. coli*, there are very few techno-economic analyses of rProtein production processes using that expression system in the literature. Moreover, those that do exist usually concern the production of high value-added proteins, such as hormones and antibody fragments, whose purity requirements and production costs can be much larger than those of industrial enzymes such as cellulases. Furthermore, to the best of our knowledge, there is no published techno-economic analysis of a rProtein production process that considers multiple process configurations and process variables such as biomass productivity, rProtein content, amount of inducer, cost of antibiotics, fermenter material and seed train size. In addition, the biotechnological use of recombinant platforms equipped with toxin-antitoxin systems remains largely unexplored, and their economic impact, unknown. In particular, the new possibilities for industrial inoculum production provided by such systems have not been evaluated, neither experimentally nor *in silico*.

1.1 Objective

The major aims of this work were to clone and express a cellulolytic enzyme (β-glucosidase) in *E. coli* SE1, a modified BL21(DE3) strain equipped with a toxin-antitoxin system; to develop a novel inoculum production strategy, exploiting the features of *E. coli* SE1; and to design, model, simulate and economically evaluate multiple process configurations for enzyme production based on the *E. coli* SE1 platform, including the implementation of the alternative inoculum production scheme proposed earlier. Next, these objectives are described in greater detail.

1.1.1 Specific Objectives

a) To clone a cellulolytic enzyme (β-glucosidase) in *E. coli* SE1, a modified *E. coli* BL21(DE3) strain equipped with a toxin-antitoxin system for plasmid segregational stability, through the method of restriction digestion; to express that enzyme and verify its activity; and perform the same procedures with the conventional *E. coli* BL21(DE3), for the purpose of comparison;
b) To carry out cultures of *E. coli* SE1 and BL21(DE3) on mineral medium, in shake-flasks, to determine the kinetic parameters and enzymatic activity of both strains and compare their performance;
c) To develop and test an unconventional inoculum production method exploiting *E. coli* SE1 features, based on recycling a fraction of the cells after each passage, while regularly assessing plasmid stability through agar plate counts;
d) To design, model, simulate and economically evaluate an enzyme production process based on the recombinant *E. coli* SE1 platform, as well as multiple alternative process configurations, including the implementation of the cell recycle strategy for inoculum generation, so as to estimate the enzyme production cost, as well as to identify the most critical process parameters and optimal process configurations.

In Section 2, we present a literature review covering the enzymatic hydrolysis of lignocellulose, the special role of β-glucosidase in this process, the recombinant production of proteins using *E. coli,* as well as the plasmid stability systems required for such production; moreover, we review the literature of modeling, simulation, and economic analysis of bioprocesses, with an emphasis on the production of industrial enzymes. In Section 3, we describe the cloning and expression of β-glucosidase in *E. coli* strains SE1 and BL21(DE3) (specific objectives a and b above); in addition, we evaluate plasmid stability over repeated passages for both strains (specific objective c). In Section 4, we design and simulate an industrial process for β-glucosidase production, as well as perform a thorough techno-economic assessment of that process (specific objective d). Lastly, in Section 5, we present the major conclusions from both the experimental work and the *in silico* analysis.

2 LITERATURE REVIEW

2.1 Enzymatic Hydrolysis of Lignocellulose

2.1.1 Lignocellulose Composition and Structure

Lignocellulose is the most abundant renewable carbon source on Earth (40), accounting for more than 60% of all the plant biomass (41). It consists of a composite material primarily formed by two types of polysaccharides, cellulose and hemicellulose, and a complex aromatic polymer, lignin (42,43).

Cellulose consists of a homopolymer of D-glucose in which the units bind through β-(1 → 4) glycosidic bonds, giving rise to linear chains. These chains interact with each other through hydrogen bonds, forming highly crystalline microfibrils that provide cellulose its high tensile strength (44). While amylaceous substances (which are also homopolymers of D -glucose but with α (1 → 4) and α (1 → 6) bonds) require only 60 °C - 70 °C to pass from crystalline structure to amorphous structure, cellulose requires 320 °C and pressures of 25 MPa for such transition (45). The cellulose microfibrils are surrounded by a matrix of hemicellulose and lignin (44), so that cellulose binds to hemicellulose through hydrogen bonds (46) and hemicellulose binds to lignin through covalent bonds of the ester or ether kind (45–47).

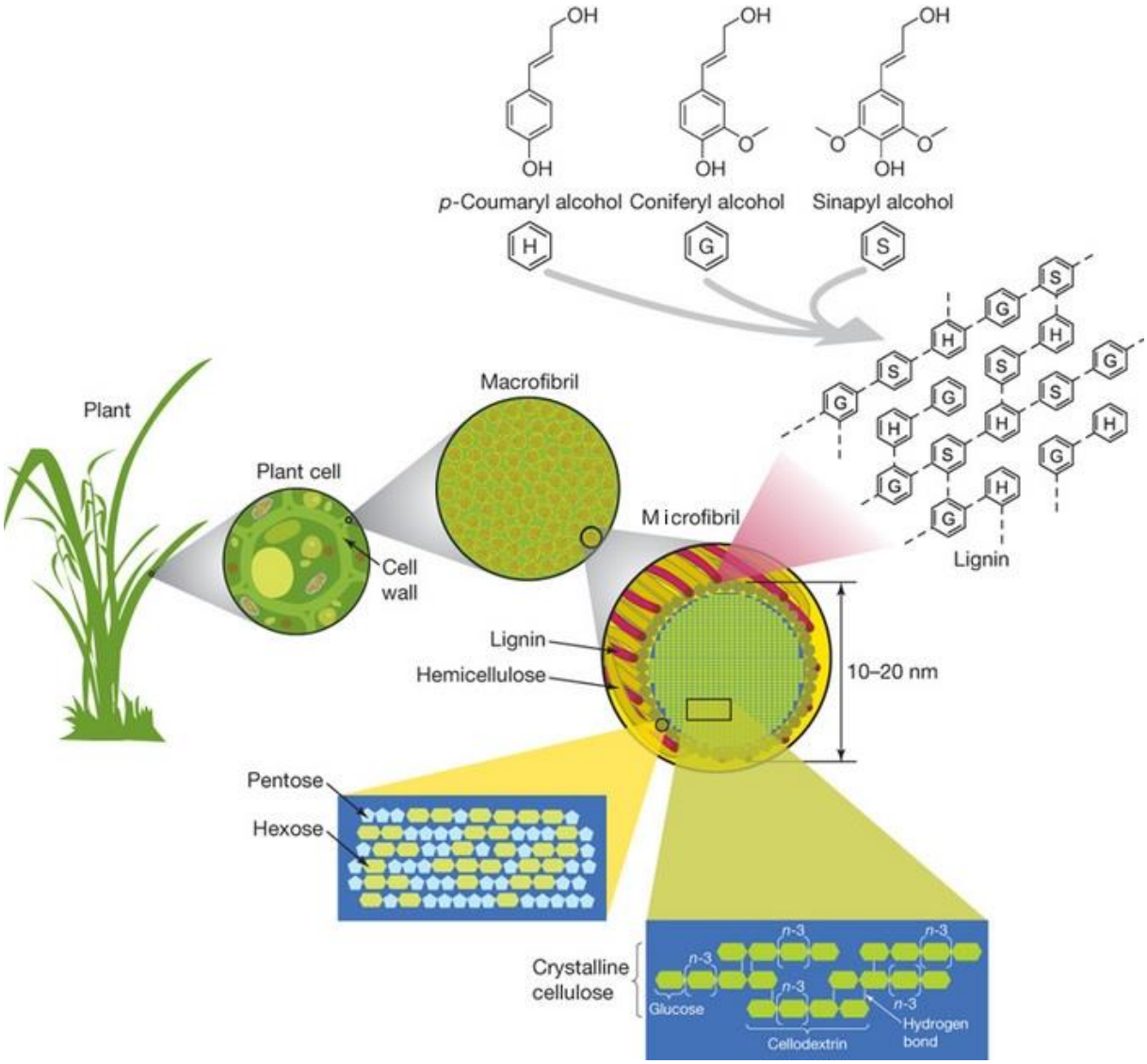
Hemicellulose is a group of branched heteropolymers composed of pentoses (β-D-xylose and α-L-arabinose), hexoses (β-D-glucose, α-D-galactose and β-D-mannose) and uronic acids (α-D-glucuronic, α-D-galacturonic and α-D-4-O-methylgalacturonic acids); in addition, the hydroxyl groups of the sugar monomers may be partially substituted by acetyl groups. Although the composition of the hemicellulosic fraction varies according to plant species, stage of development and tissue, xylose is generally the most abundant sugar; in fact, the major hemicellulose chains are typically composed of xylose units linked to each other through β (1 → 4) bonds (xylans) with side substitutions at the 2-O and / or

3-O positions by arabinose (48). It is worth noting that, although hydrogen bonds and other intermolecular interactions among hemicellulose chains and between hemicellulose and cellulose do exist, hemicellulose is essentially amorphous because of its heterogeneity (49).

Lignin is a complex aromatic heteropolymer composed of three phenolic monomers derived from cinnamic acid (*trans*-phenylpropenoic acid), namely *p*-coumaryl, coniferyl and synapylic alcohol (45); these monomers, called monolignols, are present in characteristic ratios, according to the species, tissue and stage of development of the plant, and bind to each other through ether-like or C-C bonds. Moreover, carboxylic acids derived from two of the monolignols - the *p*-coumaric and coniferic acids - bind to certain monosaccharides from hemicellulose, especially arabinoses, and cellulose, through ether-like and ester-like bonds; thus, lignin, hemicellulose and cellulose become intertwined, which provides mechanical resistance to the cell wall and, by extent, to the entire plant. In addition, the aromatic character of the monolignols make lignin highly hydrophobic and, thereby, impermeable. The rigidity and impermeability of lignin allow the transport of water and nutrients through the plant vessels (in fact, the presence of lignin is characteristic of vascular plants), and also provides resistance to biological degradation.

Lignin is present in greater amounts in woody biomass, such as tree trunks and branches; softwoods have the highest levels of lignin (30-60%), followed by hardwoods (30-55%). Herbaceous plants and agricultural residues, on the other hand, usually have lower levels of lignin; sugarcane bagasse, in particular, typically contains 19 – 24% of lignin (45,50,51).

Figure 1 – Multiple levels of lignocellulose structure.



Caption: Diagram showing the multiple levels of structure of lignocellulosic biomass, from plant leaves to the molecules that constitute plant cell walls. Plant cell walls are comprised of fibers (macrofibrils), which themselves consist of bundles of smaller fibers (microfibrils). In these microfibrils, a bundle of parallel cellulose chains is enveloped by lignin and hemicellulose chains. Cellulose chains are composed of glucose molecules; hemicellulose, of various pentose and hexose sugars; and lignin, of three phenolic monomers. Source: reproduced from ref. (52) with a minor correction ("Microfibril" replaced "Macrofibril" in the original).

Owing to the intricate and compact structure of cellulose, hemicellulose and lignin, as well as the crystalline character of cellulose and low solubility of lignocellulose as a whole, lignocellulosic biomass is highly resistant to enzymatic degradation; in the context of lignocellulose hydrolysis, it is said to be *recalcitrant*.

2.1.2 Sources of Lignocellulose

Four main sources of lignocellulosic biomass can be distinguished: “energy crops”, agricultural residues, residues from the forestry and paper industry, and other industrial/municipal wastes. Whereas the latter three categories concern particular kinds of waste currently generated by different economic activities, the category of “energy crops” consists of plants that could potentially be grown for the express purpose of lignocellulose hydrolysis, with the ultimate aim of producing fuels and/or chemicals. Good candidates would be perennial or short rotation crops that produce high-yields with low nutrient and maintenance requirements, mainly represented by herbaceous plants, such as switchgrass, miscanthus and giant reed, and fast-growing woody crops, such as willow, poplar and eucalyptus (53,54).

In Brazil, most academic and industrial research has focused on agricultural residues, and, in particular, on sugarcane bagasse, given that it is an abundant agricultural waste and that its processing could be integrated to existing sugar-ethanol plants to produce ethanol. Sugarcane bagasse contains, on average, 32-44% cellulose, 27-32% hemicellulose, 19-24% lignin and 4.5-9% ashes (55). In terms of monosaccharides other than glucose, sugarcane bagasse is composed of: 20.5-25.6% xylose; 2.3-6.3% arabinose; 1.6% galactose and 0.5-0.6% mannose (all values consider bagasse dry weight) (48); i.e., most non-glucose sugars present are pentose sugars, and xylose, in particular, is largely predominant.

2.1.3 Hydrolysis of Lignocellulose

Cellulose and hemicellulose chains have to be converted into oligo or monosaccharides so that they can be taken up by microorganisms and used as carbon and energy sources. Formally, such breakdown is carried out by water molecules, which gives the process the name of hydrolysis (sometimes also called “saccharification” in the context of biofuel production). However, the breakdown of cellulose in pure water is not kinetically feasible

at room temperature and neutral pH (in fact, the half-life of cellulose is estimated to be several million years old (56)). Therefore, in practice, hydrolysis needs to be catalyzed, either by a chemical (acid) or biochemical (enzymatic) agent.

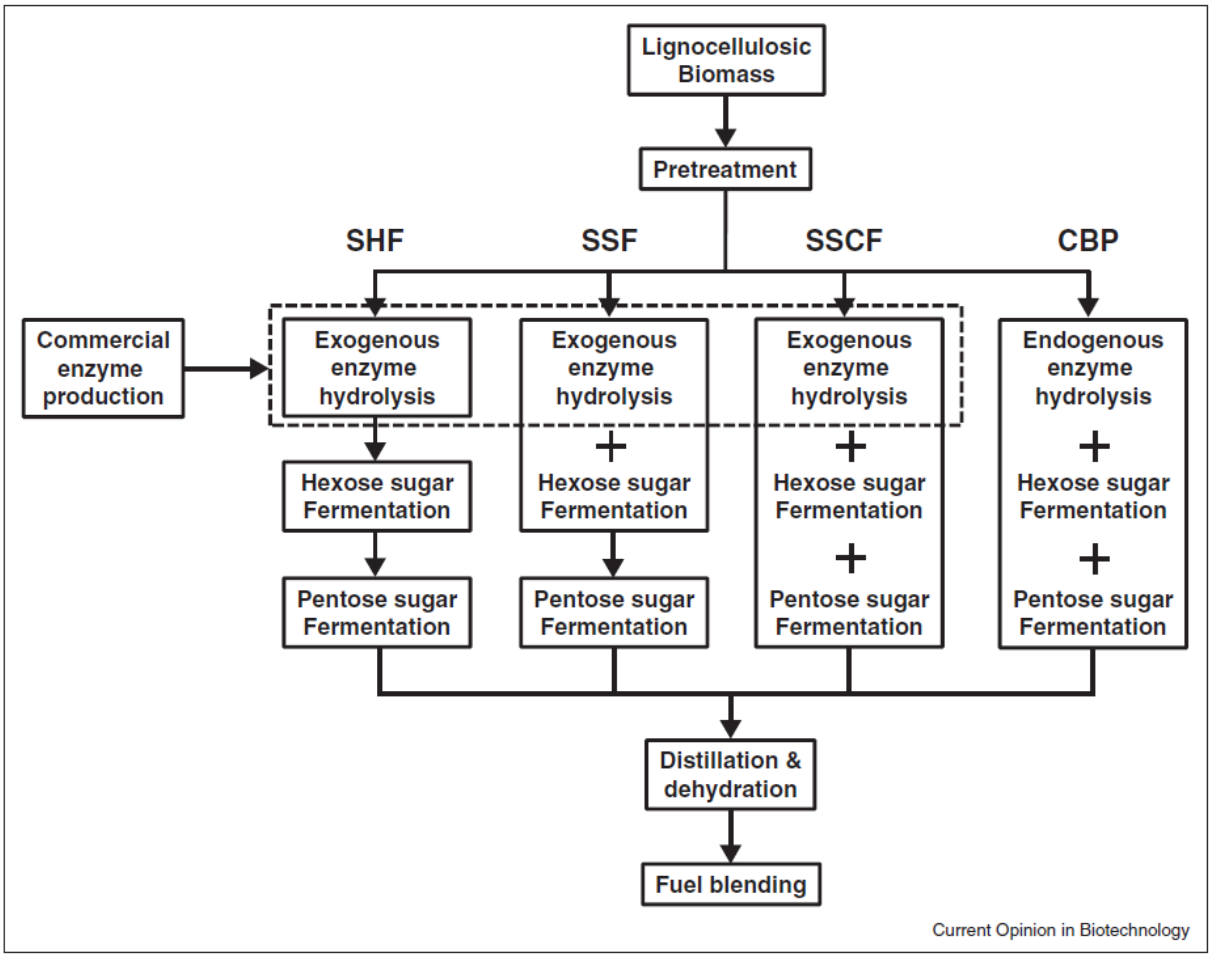
Chemical or acid hydrolysis employs strong acids, such as sulfuric or hydrochloric acid, either at low concentrations and high temperatures (2-5%, 160-200 °C, 10 atm), or at high concentrations and moderate temperatures (10-30%, 50 °C, 1 atm). The major drawback of dilute acid hydrolysis is the unintended degradation of biomass components, forming toxic or inhibitory substances that harm subsequent steps (such as ethanol fermentation by yeast). The main disadvantage of concentrated acid hydrolysis, on the other hand, is the acid cost, which makes acid recovery after hydrolysis indispensable (57).

Enzymatic hydrolysis, as its name implies, is based on the action of enzymes, usually produced by microorganisms; it can achieve high yields at mild conditions (50 °C, pH 5) without generating toxic or inhibitory byproducts (5,57). Since this technology is less mature than the chemical pathway, it also has more room for improvement and thus for cost reduction (58). For all these reasons, the enzymatic route has been generally favored in both academic and industrial settings over chemical hydrolysis.

The complete breakdown of cellulose generates glucose, which is a universal carbon source that can be used by a variety of microorganisms to produce ethanol or other chemicals. The breakdown of hemicellulose, on the other hand, generates a number of different sugars, mainly pentoses, and, in particular, xylose, which can also be used by certain microorganisms, though not as universally or effectively as glucose. Consequently, some process schemes propose the separation of cellulose from hemicellulose prior to hydrolysis, so that two distinct liquors are generated - one composed of hexoses (essentially glucose) and the other of pentoses (primarily xylose) - which can then be fermented by distinct microorganisms in separate vessels. When the hydrolysis reaction occurs separately from the fermentation step, and the glucose liquor is fermented separately from the pentose liquor, the process is referred to as Separate Hydrolysis and Fermentation (SHF); on the other hand, when hydrolysis occurs separately from fermentation but both glucose and pentose sugars are fermented together (either by one microorganism or a microbial consortium), the process is referred to as Separate Hydrolysis and Cofermentation (SHCF). Yet another possibility is to carry out the

hydrolysis reaction and the fermentation process in the same vessel, either by using a single microorganism/consortium capable of synthesizing the hydrolytic enzymes as well as fermenting the resulting sugars, or by adding the hydrolytic enzymes and the fermenting organism to the same reactor; the former case is referred to as Consolidated Bioprocessing (CBP), while the latter is referred to as Simultaneous Saccharification and Cofermentation (SSCF) or Simultaneous Saccharification and Fermentation (SSF), depending on whether or not pentose sugars are fermented (48,59–61). All these process configurations are represented in Figure 2.

Figure 2 – Basic process configurations for lignocellulose hydrolysis and fermentation.



Caption: SHF: Separate Hydrolysis and Fermentation; SSF: Simultaneous Saccharification and Fermentation; SSCF: Simultaneous Saccharification and Cofermentation; CBP: Consolidated Bioprocessing. Source: ref. (60).

Separating hydrolysis from fermentation has the advantages of allowing optimum conditions (temperature, pH, medium composition) for each process, and of reducing medium viscosity prior to the fermentation step, which favors mixing and mass and heat

transfer. Besides, SHF avoids certain products of the fermentation step, such as ethanol, from inhibiting the hydrolytic enzymes. On the other hand, carrying out hydrolysis and fermentation in the same vessel reduces end-product inhibition of the hydrolysis reaction, given that the sugars produced by hydrolysis are continually consumed by the fermenting microorganism. Furthermore, using a single vessel for both processes presumably reduces equipment cost (60,62).

Owing to the cross-linked and heterogeneous structure of lignin, its biological deconstruction is relatively slow, which makes this process commercially unattractive (4,63); on the other hand, it is useful for the purpose of isolating the polysaccharide fractions from the lignocellulosic biomass (64). Today, around 95% of all the lignin produced is burned for energy generation, while the remaining 5% are used to produce binders, surfactants, dispersants and additives (65). Since the amount of lignin generated by lignocellulosic fuels and chemicals will significantly exceed their energetic needs, many new avenues of lignin valorization have been recently studied, such as the production of fine chemicals, solvents and plastics (66). In fact, lignin valorization has proved crucial for the economic feasibility of a biofuel production process analyzed recently (67); in that work, lignocellulosic polysaccharides were used to produce hydrocarbon fuels, while lignin was employed to coproduce adipic acid, a building block chemical primarily used to produce nylon and other polymeric materials (67,68).

2.1.4 Pretreatment of Lignocellulose

For enzymatic hydrolysis to be efficient, lignocellulosic biomass must be first subjected to one or more physical, chemical, physicochemical, and/or biochemical processes, so-called pretreatment methods. Ideally, the pretreatment step should separate cellulose, hemicellulose and lignin from one another, reduce the crystallinity and degree of polymerization of cellulose, increase the biomass specific surface available to the enzymes (by reducing particle size or increasing particle porosity), increase pore size so that enzymes penetrate more easily, and remove acetyl groups that interfere with enzyme-substrate recognition (69–72).

2.1.4.1 Physical Pretreatment Methods

Mechanical size reduction processes, such as chipping, milling and grinding, are usually the first pretreatment step. They increase the surface area accessible to enzymes by reducing particle size and increasing porosity; besides, they decrease cellulose's crystallinity (4,70,73–76), cellulose's degree of polymerization (73–75), and improve general mass transfer properties by reducing particle size (4). Other physical pretreatment methods include extrusion and irradiation (gamma, electron beam or microwave radiation) (4,73,75–77).

2.1.4.2 Physicochemical Pretreatment Methods

Physicochemical methods are also often employed to pretreat lignocellulosic biomass; wet oxidation, liquid hot water (LHW), steam explosion, ammonia fiber explosion (AFEX) and supercritical CO_2 explosion are some of the most common (74,78). All these methods rely on high temperatures and, except for LHW and wet oxidation, on sudden pressure changes, to increase biomass porosity, reduce particle size, and disrupt or remove the hemicellulose and lignin fractions (72,79). Both hydrothermal treatments (LHW and steam explosion) and CO_2 promote hemicellulose hydrolysis through an acid-catalyzed mechanism, thanks to water autohydrolysis (48,61,80) in the case of hydrothermal treatments, and to carbonic acid formation (61,74,81), in the case of supercritical CO_2 explosion. However, these three methods do not significantly remove the lignin fraction (72,76,79). On the contrary, AFEX is an alkaline treatment and, as such, breaks up the ester-like bonds between lignin and the carbohydrate fractions, efficiently removing lignin; moreover, this method removes a minor part of the hemicellulosic fraction, and significantly reduces the degree of crystallinity of cellulose (57,61,70,72). Wet oxidation, which consists of a high-temperature treatment using water and air/oxygen as an oxidant agent, brings about hydrolytic and oxidation reactions that remove both lignin and hemicellulose from the biomass (73,74,80).

2.1.4.3 Chemical Pretreatment Methods

Several chemical pretreatment methods have also been evaluated, such as dilute acid, alkali, organosolv, ozonolysis, and ionic liquids (74,75,78). Dilute acid treatments, which are usually carried out with mineral acids (4,82) and under high temperatures (57,61,70), effectively hydrolyze and remove hemicellulose, increase biomass porosity, significantly disrupt lignin (75,76) and reduce cellulose's degree of polymerization as well (70). Alkaline treatments, on the other hand, remove the lignin fraction and part of the hemicellulosic fraction (70,75,76), reduce cellulose's crystallinity and degree of polymerization, remove acetyl groups and uronic acid substitutions, and make cellulose swell, thereby increasing the biomass specific surface (70). Organosolv treatments make use of organic solvents such as ethanol, methanol, ethylene glycol or acetone, or their aqueous mixtures (57,61), over a temperature range of 100-250 °C (57,61,70). They cleave lignin internal bonds and lignin-carbohydrate bonds, therefore extensively removing lignin (57,61,70), and they can also remove hemicellulose to a large extent, particularly when water and/or catalytic amounts of acids are used (57,70,74,83). Besides, organosolv treatment increases biomass porosity (61,70,83) and reduces cellulose's degree of polymerization (70). Ozonolysis, a treatment based on the strong oxidative power of ozone gas, significantly removes lignin (74,75,83). Ionic liquids are salts with low melting points (< 100 °C, often liquid at ambient temperature) typically made up of a large organic cation and a small (inorganic or organic) anion (57,74,75,83,84). They are able to break up the intermolecular interactions that hold the lignocellulose structure together, effectively separating cellulose, lignin and hemicellulose (57,61,70,83,84). Moreover, in this process they decrystallize the cellulose fraction (70,72,83–85), increase biomass porosity (70,72,83) and reduce cellulose's degree of polymerization (84). Thus, these so-called green solvents are regarded as promising tools for biomass pretreatment in the long term (61,70,74,83,84).

It is worth noting that most physicochemical and chemical treatments degrade part of the lignocellulose, particularly hemicellulose sugars and lignin monomers, transforming them into toxic or inhibitory substances such as furan derivatives, phenolic compounds, and organic acids (45,73–75,81,86–89). The character and amount of these compounds

are specific to each pretreatment method, as well as to its operating conditions and the biomass source. In order to tackle this issue, several strategies have been put forward, such as selecting or engineering feedstocks that generate fewer toxic compounds (82); selecting or engineering microorganisms that tolerate well the toxic compounds; reducing the amount of toxic substances in the pretreated biomass through chemical or biological processes (so-called detoxification processes) prior to the fermentation step (72,74,75,82); or replacing/complementing physicochemical/chemical pretreatment with methods that do not generate inhibitors, such as biological pretreatment processes (75,82,90).

2.1.4.4 Biological Pretreatment Methods

Biological pretreatment methods, much like enzymatic hydrolysis, are predicated on the action of proteins that cleave the covalent and non-covalent bonds that intertwine cellulose, hemicellulose and lignin. These proteins may be produced by microorganisms *in situ* (that is, directly secreted onto the biomass) or in a separate operation (90). Particularly useful for pretreatment purposes are feruloyl esterases, which cleave ester bonds that bind hemicellulose to lignin; enzymes that hydrolyze hemicellulose chains, such as xylanases (40,61,91,92); and lignin-degrading enzymes, especially lignin peroxidases, manganese dependent peroxidases, versatile peroxidases and laccases (40,61,64,90,92–95). Some microbial proteins that have no significant catalytic activity, but nonetheless disrupt cellulose and/or hemicellulose intermolecular interactions, such as swollenin, also help to deconstruct the lignocellulose matrix and thus can be of interest for pretreatment purposes (49,63). In the literature, most biological pretreatment processes are performed using soft rot fungi, particularly white rot species (63,64,72,83,90,93), and under solid state culture (64,78,83,90,93).

Biological pretreatment methods have some important advantages over physicochemical and chemical methods: since they are carried out under milder conditions of temperature, pressure and pH, they save energy and are less dangerous and detrimental to the environment; and they do not generate toxic or inhibitory substances that hamper

fermentation processes, as do physicochemical and chemical methods (49,73,75,76,83,90,93). Nevertheless, they are not cost-effective; on one hand, if enzymes are produced separately, the process becomes too expensive; and on the other hand, if a microbial treatment is employed, the process takes too long (from weeks to months), and the microorganisms may inadvertently consume a significant amount of sugars (4,63,73,75,90,93).

2.1.5 Enzymes for Lignocellulose Hydrolysis

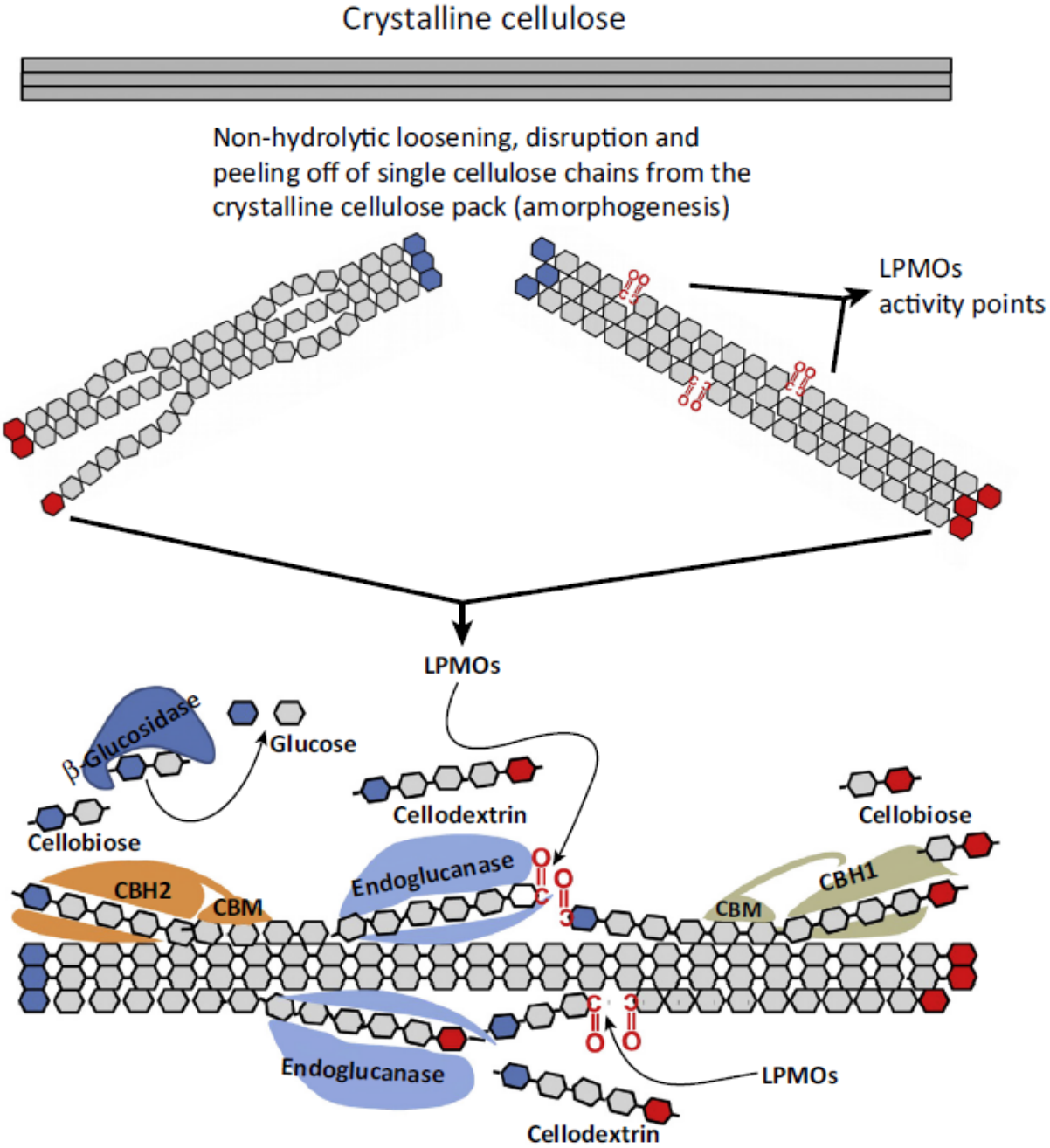
2.1.5.1 Cellulases

The cellulose fraction of biomass can be broken down by two basic enzymatic systems: either a set of free extracellular enzymes, produced by bacteria or fungi, or a multienzymatic complex attached to the surface of the cell, called a cellulosome, produced by certain anaerobic bacteria (42,44,69). In any case, multiple enzymatic activities are required to hydrolyze cellulose into soluble monosaccharides that can be assimilated by cells. The classical model of enzymatic hydrolysis consists of three major types of glycosyl-hydrolases (GHs): (i) endoglucanases (EGs; EC 3.2.1.4), which randomly cleave inner bonds of cellulose chains, preferentially in amorphous regions (ii) cellobiohydrolases (CBHs; also called exoglucanases), which cleave bonds at a distance of 2 glucose units from the end of the chain, in a processive manner (there are two subtypes of CBHs: those acting on reducing ends, CBH1 (EC 3.2.1.176), and those acting on non-reducing ends, CBH2), thus generating a disaccharide of glucose (cellobiose); and (iii) β-glucosidases (BGLs; EC 3.2.1.21; also called cellobiases), which convert short oligosaccharides, especially the cellobiose generated by CBHs, into glucose; collectively, these three GH classes are called cellulases (40,42–44,51,96). Cellulases act on cellulose synergistically: EG generates new chain ends on which CBHs can act, and the activity of CBHs exposes new amorphous regions to EG (97,98). Furthermore, BGL eliminates cellobiose, which is a strong inhibitor of EGs and CBHs, from the reaction mixture (16,17,99).

Recently, it was discovered that a different type of enzyme, called Lytic Polysaccharide Monooxygenase (LPMO), also takes part in cellulose hydrolysis, acting synergistically with conventional cellulases. LPMOs (AA9 or AA10 in the Carbohydrate-Active Enzymes database (100)) are copper-dependent enzymes that oxidatively cleave internal glycosidic bonds of cellulose at C1 and/or C4 positions, even on crystalline regions (13,40,43,96,101,102); moreover, they were shown to act on other polysaccharides, such as chitin, xylan, xyloglucans, β-glucans, glucomannans and starch (43,101). They are thought to boost cellulose hydrolysis primarily by making new chain ends available to the other enzymes, particularly in crystalline regions that would be otherwise highly resistant to cellulases (13,43,96,101,102). It is worth mentioning that LPMOs are already present in the latest commercial cellulase cocktails (43,96).

It should be noted that many cellulases (and hemicellulases) that act on insoluble substrates have a modular structure: they are constituted by a catalytic domain connected by a flexible peptide linker to a so-called carbohydrate-binding module (CBM), which anchors the enzyme to the solid substrate. CBMs assist biomass hydrolysis by effectively increasing the concentration of their enzymes near the substrate surface and, depending on their amino acid sequence and resulting shape, they provide specificity to a certain substrate or substrate region (such as reducing or non-reducing ends). The concerted action of all the four types of cellulase previously described, including CBHs that possess CBMs, is schematically presented in Figure 3.

Figure 3 – Enzymatic breakdown of cellulose by a set of extracellular enzymes.



Caption: Concerted action of several types of cellulases to break down cellulose. Lytic Polysaccharide Monooxygenases (LMPOs) cleave crystalline chains, generating amorphous regions; endoglucanases randomly break bonds in the middle of cellulose chains, thereby creating reducing and non-reducing ends; cellobiohydrolases of type 1 (CBH1) and 2 (CBH2) cleave on the reducing and non-reducing ends, respectively, generating cellobiose; and β-glucosidases split cellobiose into two glucose molecules. Note that CBHs are represented here as having carbohydrate-binding modules (CBMs). Cellodextrin refers to any short glucose chain produced during the process of cellulose hydrolysis. Source: ref. (40).

2.1.5.1.1 β-glucosidases

As mentioned earlier, BGLs are hydrolytic enzymes that catalyze the cleavage of glycosidic bonds present in short oligomers of β-D-glucose (up to 6 sugar units), particularly cellobiose, to produce glucose. These enzymes are found in all kingdoms of life, including bacteria, archaea, fungi, plants, insects and mammals, in which they have multifarious functions (103); in microorganisms, they take part in the breakdown of polysaccharides, as described earlier; in animals, they participate in the metabolism of glycolipids and exogenous glycosides; and in plants, they have numerous functions, such as cell wall catabolism and lignification, defense, secondary metabolism (16,103,104), symbiosis, seed development, activation of phytohormones (103,104), fruit ripening (16), signaling and release of aromatic compounds (104).

Many BGLs also act on substrates other than β-D-glucosides, such as β- D -galactosides, α-L-arabinosides, β-D-xylosides and β-D-fucosides, as well as alkyl-, amino-, and aryl-β-D-glucosides. In fact, these enzymes may be grouped into three categories depending on their substrate specificity: (i) aryl-β-glucosidases, which display preference for glucosides bonded to aromatic groups; (ii) cellobiases, which preferentially hydrolyze cellobiose; and (iii) broad spectrum BGLs, which hydrolyze a wide variety of substrates. To this date, most documented BGLs belong to the third group (16,103).

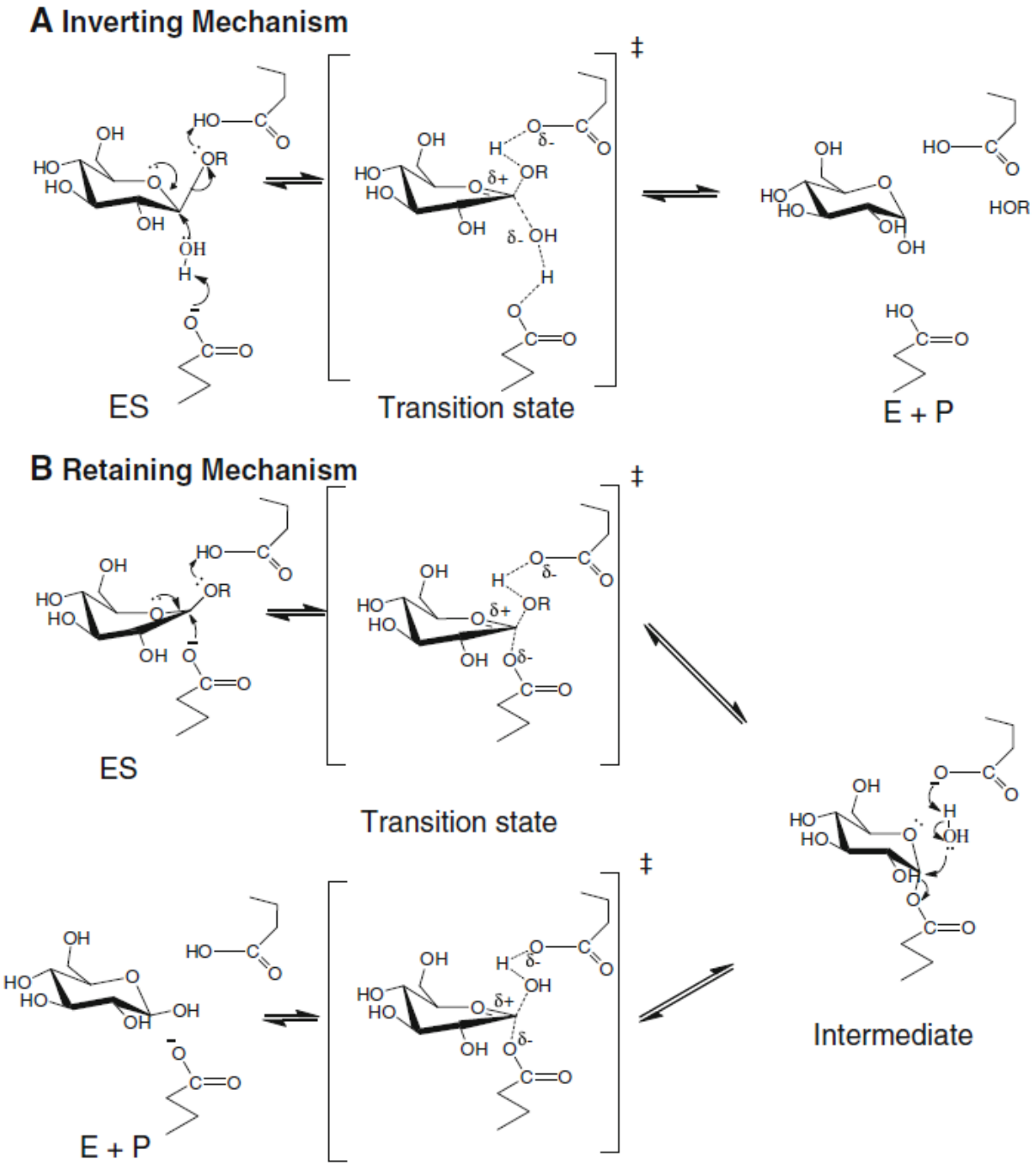
Nevertheless, the substrate specificity of BGLs and, more generally, of glycosyl hydrolases (GHs), provides little to no information about protein structure, catalytic mechanism, and evolutionary relationships. Moreover, a classification system based on substrate specificity does not suit properly enzymes that act on multiple substrates, a common characteristic of BGLs and cellulases in general. For these reasons, in the 1990s Henrissat and Davies (105) proposed a new classification system for GHs based on amino acid sequence identity and 3D structure similarity: GHs with similar amino acid sequences were grouped into families named GH1, GH2, GH3, etc.; and these families were further grouped into clans named GH-A, GH-B, GH-C, etc., according to 3D structure similarity as measured by hydrophobic cluster analysis. In this system, BGLs can be found in families GH1, GH2, GH3, GH5, GH9, GH30, GH39 and GH116, although most of them fall in family

GH1; moreover, most BGLs are assigned to clan GH-A, which means that they form a $(\beta/\alpha)_8$ (TIM) barrel structure (105).

With a few exceptions, the basic catalytic mechanism of all BGLs is the same: a nucleophilic substitution by an H_2O molecule at the sugar anomeric carbon, catalyzed by two amino acids, one acidic and one nucleophilic, carried out in two steps. In the first step, called deglycosylation, the acidic amino acid donates a proton to the oxygen of the glycosidic bond, while the nucleophilic amino acid attacks the sugar anomeric carbon; this step removes a reducing sugar, named aglycone in this context, and generates a sugar (glycone) covalently bonded to the enzyme. In the second step, named glycosylation, water acts as a nucleophile and attacks the anomeric carbon; the bond between the nucleophilic amino acid and the anomeric carbon is cleaved and, at the same time, a proton of the water molecule is removed by the other participating amino acid. As a result, a glucose is produced, and the chirality of its anomeric carbon is preserved. For this reason, this mechanism is said to be a retaining mechanism; it is illustrated in Figure 4, along with the inverting mechanism, that is employed by a few BGL enzymes (103,104).

However, if a nucleophile other than water, such as a monosaccharide, a disaccharide or an alcohol, is able to attack the sugar in the second step, an elongated product is generated; this side reaction is denominated transglycosylation. Under certain reaction conditions, such as low water activity or high sugar concentration, BGLs may also promote the reverse of hydrolysis, generating elongated products as well (16,103,104); however, reverse hydrolysis is under thermodynamic control, whereas transglycosylation is under kinetic control (16). In the context of biomass hydrolysis, these reactions are generally considered undesirable and thus should be minimized.

Figure 4 – Distinct reaction mechanisms of β-glucosidases.



Caption: Two distinct hydrolysis mechanisms of β-glucosidases. The inverting mechanism (A) changes the chirality of the sugar anomeric carbon, whereas the retaining mechanism (B) preserves its chirality. Most β-glucosidases known to date make use of the retaining mechanism. Source: ref. (104).

To be eligible for the highly demanding industrial process of enzymatic hydrolysis, enzymes must exhibit high levels of activity and catalytic efficiency, as well as low sensitivity to common inhibitors. The latter property is especially crucial when dealing

with a high-solids reaction mixture, which is increasingly considered essential for the economic feasibility of biomass hydrolysis (106). In this regard, BGLs (as other enzymes) usually suffer from end-product inhibition, that is, they are inhibited by significant concentrations of glucose in the medium. However, numerous BGLs are curiously unaffected by glucose, and several BGLs even display higher activity in the presence of glucose, that is, they are stimulated by their end-product. These BGLs are said to be glucose-tolerant . The precise mechanism of glucose tolerance seems to vary from BGL to BGL; some are apparently subject to allosteric activation, while others derive their glucose tolerance from transglycosylation reactions with the inhibitor molecule (107,108), or yet from binding of the inhibitor molecule to non-productive sites near the catalytic site. In any case, this characteristic is quite attractive for biomass hydrolysis, particularly under high solids, given that the glucose generated by a glucose-tolerant enzyme would further enhance the process, instead of hamper it (108).

The BGLs with the highest values of catalytic efficiency, measured on *p*-nitrophenyl glucopyranoside (pNPG) or cellobiose, are listed in Table 1, which also computes the inhibitory (or stimulant) effect of glucose through K_i values. It is worth noting that several of the enzymes with highest activity on pNPG originate from hyperthermophilic microorganisms, particularly bacteria from genus *Thermotoga*. Hyperthermophilic enzymes exhibit optimal activity at temperatures over 70 °C, sometimes surpassing 90 °C, and are often highly stable at high temperatures as well. From a process engineering perspective, they present numerous advantages: high temperatures tend to kinetically favor high levels of activity (109), reduce medium viscosity (especially under high solids), minimize the risk of contamination and, in the particular case of biomass hydrolysis, lower energy consumption (given that the lignocellulosic biomass, usually pretreated at high temperatures, does not need to be cooled before the enzymatic step) (109,110). It is also interesting that a few of these hyperthermophilic enzymes are glucose-tolerant as well. On the other hand, the highest levels of catalytic efficiency measured on cellobiose are found, by and large, in filamentous fungi such as *Penicillium* and *Talaromyces*. Notably, several of them are thermophilic too. Although not evident from Table 1, it appears that glucose inhibition is generally stronger in these enzymes. In other words, there is a trade-off between catalytic efficiency on cellobiose and glucose tolerance (106).

Table 1 – Properties of β-glucosidases with high catalytic efficiency on *p*-nitrophenyl glucopyranoside or cellobiose.

Microbial source	Substrate	Activity (U/mg protein)	*T*(°C)	k_{cat}/K_M (s^{-1}·mM^{-1})	K_i (mM)	Reference
Thermotoga naphthophila RKU-10 (GH3)	pNPG	1.30E+05	85	2.70E+06	150	(111)
Thermotoga naphthophila RKU-10 (GH1)	pNPG	2.55E+05	95	1.02E+06	stimulated	(112)
Thermotoga petrophila (GH1)	pNPG	3.04E+04	90	3.08E+04	stimulated	(113,114)
hydrothermal spring (Archaea)	pNPG	3.20E+03	90	2.62E+04	150	(115)
Penicillium italicum	pNPG	6.18E+01	60	1.58E+04	8.9	(116)
Dictyoglomus turgidum	pNPG	1.60E+02	80	1.00E+04	750	(117)
Talaromyces amestolkiae (BGL3)	pNPG	4.24E+01	70	9.71E+03	—	(118)
Talaromyces leycettanus	pNPG	9.05E+02	75	9.10E+03	14	(119)
Periconia sp.	pNPG	6.72E+02	40	6.27E+03	20	(120)
Fomitopsis palustris	pNPG	1.91E+02	50	6.16E+03	0.35	(116)
Neosartorya fischeri	pNPG	2.19E+03	80	5.59E+03	13.4	(121)
Rhizomucor miehei	pNPG	6.22E+01	70	5.35E+03	8	(122)
Penicillium funiculosum	pNPG	1.35E+03	60	4.51E+03	stimulated	(123)
Pyrococcus furiosus	pNPG	4.46E+02	95	4.51E+03	20	(124)
Humicola insolens	pNPG	4.62E+01	65	4.24E+03	3.27	(125)
Anoxybacillus sp. DT3-1	pNPG	1.16E+03	70	3.69E+03	tolerant	(126)
Bacillus polymyxa	pNPG	2.42E+03	37	3.36E+03	19	(127)

Microbial source	Substrate	Activity (U/mg protein)	T(°C)	k_{cat}/K_M (s^{-1}·mM^{-1})	K_i (mM)	Reference
Talaromyces amestolkiae (BGL2)	pNPG	3.54E+01	70	2.56E+03	—	(118)
Caldicellulosirup-tor owensensis	pNPG	1.62E+03	75	2.47E+03	—	(128)
Thermoascus aurantiacus	pNPG	1.90E+05	60	2.13E+03	0.29	(129)
Thermotoga neapolitana	pNPG	—	90	2.00E+03	—	(130)
Stereum hirsutum	pNPG	1.73E+03	65	1.98E+03	29	(131)
Fomitopsis pinicola	pNPG	1.42E+03	50	1.66E+03	14.8	(132)
Thermotoga maritima + *Cellvibrio gilvus* (chimera)	pNPG	5.50E+00	30	1.64E+03	—	(133)
Aspergillus oryzae	pNPG	1.76E+02	60	1.27E+03	2.9	(134)
Sulfolobus sofataricus	pNPG	—	75	1.08E+03	—	(135)
Aspegillus niger	pNPG	366	40	4.17E+02	—	(136)
Penicillium italicum	cellobiose	—	60	6.36E+03	11.3	(116)
Penicillium funiculosum	cellobiose	1354	60	3.61E+03	stimulated	(123)
Talaromyces amestolkiae (BGL3)	cellobiose	—	70	3.31E+03	—	(118)
Periconia sp.	cellobiose	—	70	1.94E+03	—	(120)
Thermoascus aurantiacus	cellobiose	—	55	1.55E+03	—	(106)
Humicola insolens	cellobiose	45.6	50	1.31E+03	3.27	(125)
Termite gut	cellobiose	—	50	1.22E+03	tolerant	(137)
Prunus armeniaca (wild apricot)	cellobiose	—	35	9.17E+02	—	(138)
Acremonium thermophilum	cellobiose	—	55	7.65E+02	—	(106)

Microbial source	Substrate	Activity (U/mg protein)	T(°C)	k_{cat}/K_M (s^{-1}·mM^{-1})	K_i (mM)	Reference
Talaromyces amestolkiae	cellobiose	—	70	6.19E+02	—	(118)
Aspergillus oryzae	cellobiose	—	50	5.10E+02	5	(134)
Aspergillus niger (Novozymes 188)	cellobiose	—	50	4.85E+02	1.94	(139)

Caption: Microbial sources and kinetic properties of β-glucosidases with high catalytic efficiency (k_{cat}/K_M) on *p*-nitrophenyl glucopyranoside (pNPG) or cellobiose. The enzymes are listed in order of decreasing k_{cat}/K_M for each substrate. Symbols: T, optimal temperature; k_{cat}, turnover number; K_M, Michaelis constant; K_i, inhibitor constant for glucose. Source: adapted and expanded from ref. (106).

2.1.5.2 Hemicellulases

Complete hydrolysis of hemicellulose requires another set of enzymes, collectively called hemicellulases, which depend on the nature of the hemicellulose in question. In general, endo-1,4-β-xylanases (which randomly hydrolyze xylan chains at internal positions - EC 3.2.1.8), exo-β-1,4-xylanases (which hydrolyze reducing ends of xylan, releasing xylose and xylobiose units - still without EC number), 1,4-β-xylosidases (which hydrolyze xylose units from the non-reducing ends of xylan - EC 3.2.1.37), α-L-arabinofuranosidases (which hydrolyze arabinose units hanging on the xylan chain - EC 3.2.1.55), α-glucuronidases (which hydrolyze glucuronic acid units - EC 3.2.1.139), acetyl xylanoesterases (which deacetylate xylan chains - EC 3.1.1.72) and feruloyl / cumaroyl esterases (which hydrolyze ester bonds between coniferic or *p*-coumaric acid from lignin and sugars from hemicellulose - EC 3.1.1.73) are useful enzyme activities for hemicellulose deconstruction (4,51,75,140–143). In the case of sugarcane bagasse, whose hemicellulosic fraction consists mainly of xylans (48), xylanases and feruloyl esterases are particularly useful, since they digest hemicellulose and disentangle hemicellulose from lignin, respectively (91).

2.1.5.3 Ligninases

As alluded in the pretreatment section, the enzymatic degradation of lignin is a distinct challenge altogether. Lignin biosynthesis involves the oxidative coupling of monolignols, which produces a heterogeneous and irregular polymer network, cross-linked through aryl ether and C-C bonds, which are less reactive than most bonds present in other biological polymers (144–146). Consequently, depolymerization of lignin is complex, requiring a variety of non-hydrolytic, oxidoreductase enzymes, notably Lignin Peroxidases (EC 1.11.1.14), Manganese Peroxidases (EC 1.11.1.13), Versatile Peroxidases (EC 1.11.1.16) and laccases (EC 1.10.3.2), to be accomplished (40,64,90,92,145,147). Other relevant enzyme activities for lignin degradation are glyoxal oxidase (EC 1.2.3.5), aryl-alcohol oxidase (EC 1.1.3.7) (92,93,148), cellobiose dehydrogenase (EC 1.1.99.18) (93,148), manganese-independent peroxidase (EC 1.1.1.7), pyranose 2-oxidase (EC 1.1.3.4), cellobiose/quinone oxidoreductase (EC 1.1.5.1) (93) and feruloyl esterase (EC 3.1.1.73) (92).

2.1.5.4 Expansin-like Proteins

Over the last decade, proteins that exhibit no catalytic activity, but nonetheless help to break down lignocellulose, have been discovered. They act by weakening and disrupting intermolecular interactions among the cellulose chains, thus making the cellulose microfibrils swell and disperse, which renders the lignocellulosic biomass more accessible to the catalytic proteins. The first proteins identified with such properties were discovered in plants, and called expansins; by loosening the lignocellulosic structure of the plant cell wall, expansins allow the plant to grow (149–151). Later, proteins with similar properties were found in fungi and bacteria, notably swollenin, in *Trichoderma reesei* (13,40,43,96,150), loosenin, in *Bjerkandera adusta* (150,152,153) and protein BsEXLX1, in *Bacillus subtilis* (154). Moreover, it has been found that supplementing cellulases with microbial expansin-like proteins enhances the efficiency of lignocellulose hydrolysis (96,150,151). Thus, the use of expansins or functionally similar microbial

proteins may be regarded as a mild biological pretreatment process, presumably reducing the intensity of the chemical/physicochemical pretreatment needed for hydrolysis.

2.2 Microorganisms for Producing Lignocellulose-Degrading Enzymes

The microorganisms most commonly used for producing cellulases, both in industrial and academic settings, are filamentous fungi of the genus *Trichoderma*, particularly *T. reesei* (8,42–44,155). *T. reesei* was first isolated from rotting US army equipment in the Solomon Islands, during World War II. Its ability to produce large amounts of cellulose-degrading enzymes, at first regarded as a problem, was soon recognized as a potentially useful trait and, since then, the species has been modified to secrete ever increasing levels of cellulases (13,156), with recent reports of industrial strains capable of secreting upwards of 100 g/L of protein (8,44,157,158). *T. reesei* produces at least 2 types of CBH (CBH-I and CBH-II), 5 types of EG and 2 types of BGL; however, CBHs and EGs account for roughly 65 – 78 % and 21 – 32 % of the total secreted protein, respectively (159), while the amount of BGL and other hydrolytic enzymes is small (6,159). Consequently, the enzymatic cocktail from *T. reesei* is often supplemented with BGL from other sources (16,17). Although BGL can be produced by various yeasts and bacteria, including recombinant strains of *E. coli* and *S. cerevisiae*, it is usually produced by filamentous fungi, particularly *Aspergillus sp.* and *Penicillium sp.* (16,155). Hemicellulases and ligninases are also synthesized by a wide variety of bacteria and fungi; industrially, however, hemicellulases are mainly produced by filamentous fungi from genera *Trichoderma* and *Aspergillus* (4,5,48,141), while ligninases are most often produced by white rot fungi such as *Phanerochaete chrysosporium* (40,61,145).

Clearly, there is not a single microorganism in Nature capable of producing a complete and balanced set of enzymes that efficiently degrades all kinds of lignocellulosic biomass (30,59,86,148), which was to be expected, given that, in the natural environment, plant biomass is degraded by an entire community of organisms (in fact, developing a single organism capable of decomposing lignocellulose into sugars alone is a major goal of consolidated bioprocessing). Thus, various strategies for improving the industrial process

of lignocellulose degradation have been proposed: supplementing the fungal cellulase mixture with lacking enzyme activities and/or accessory proteins (16,17,19,30,51,63,91,160–163); endowing cellulolytic organisms such as *T. reesei or Clostridium* with lacking enzyme activities or better enzymes, through genetic and protein engineering (4,6,155,164–169,12,13,16,42,48,86,142,143); co-culturing *T. reesei* with good producers of lacking enzymes, such as *Aspergillus sp.* (16,30,48,142,163,164,170); or yet endowing cellulolytic enzymes to fermenting, non-cellulolytic organisms such as *E. coli* or *S. cerevisiae* (e. g. providing BGL to *S. cerevisiae*, so that it can directly consume cellobiose) (4,6,168,171,13,42,44,48,60,142,143,167).

2.2.1 Recombinant Protein Production Using *E. coli*

Escherichia coli is one of the most widely used hosts for the production of recombinant proteins, with applications in the pharmaceutical and industrial sectors (24,25,27,31,172). *E. coli*'s major advantages for producing proteins are its short generation time, low nutritional requirements, the availability of strains with low proteolytic activities, and *E. coli*'s excellent genetic and physiological characterization (25,27,32,33), which enables relatively straightforward genetic manipulations (particularly through recombinant DNA technology), as well as the possibility of expressing proteins in high cell density cultures with high levels of productivity (26,27), using simple (173) and inexpensive media (27,39,141). In fact, a biotechnology company that uses *E. coli* for rProtein production has reported a median titer of 11 g/L and a range of 3 to 20 g/L for proteins expressed in soluble form, using a proprietary strain (174). Nevertheless, this host presents some drawbacks: the expression of the protein of interest normally occurs in the cytoplasm, which can make downstream processing more complex and expensive and can also impair the formation of disulfide bridges, leading to inadequate folding and often precipitation of the protein under the form of insoluble, inactive particles called inclusion bodies; its cell machinery may stall or truncate proteins owing to codon bias, bringing about low protein expression and/or function loss; *E. coli* is not able to perform post-translational modifications on proteins, particularly glycosylation (25,27,28,31,33), which are sometimes functionally important; and the

lipopolysaccharides that constitute its outer membrane elicit strong immune responses in humans and other mammals (24,25,27,28,33). The latter four issues, however, are usually more relevant when producing eukaryotic proteins for human or animal use, and do not preclude *E. coli* from accounting for one-third of the recombinant proteins approved by the FDA for therapeutic purposes (25,33).

Another drawback of *E. coli*, not limited to recombinant protein production, is the accumulation of acetate under typical batch culture conditions, which inhibits cell growth and harms protein productivity (24,25,27,175,176). This challenge can be overcome, however, by conducting cell culture in a fed-batch regime and constraining the specific growth rate, by means of aeration, temperature and feed rate control (24,176). The fed-batch mode is indeed a widely-employed process regime in the biotechnology industry, for it allows achieving higher volumetric productivities than the batch regime, while avoiding the stability and control issues of a truly continuous regime (177). In the case of recombinant protein production employing plasmid vectors, however, plasmid segregational stability is a major concern, since cells tend to lose their plasmid DNA over time if there is no system in place to prevent this process. Plasmid loss occurs because the maintenance, replication and expression of plasmids, particularly in the case of high level expression, entail a severe metabolic burden for the cell, so that cells that happen to lose their plasmid(s) during cell division have a competitive advantage over the plasmid-bearing cells and, consequently, the former tend to outgrow the latter over time (25,178).

2.2.2 Plasmid Segregational Stability in Bacteria

Plasmid segregational stability in bacterial cultures is usually ensured by inserting an antibiotic resistance gene (such as kanamycin or ampicillin resistance) into the plasmid, and, at the same time, adding the corresponding antibiotic to the culture medium (27,31,37,39); consequently, upon cell division, if a daughter cell receives no plasmid copy from the mother cell, it dies. This method, however, presents some drawbacks for producing recombinant proteins on a large scale, especially the high cost of antibiotics (36,39,179), the risk of spreading antibiotic resistance genes to pathogens through

horizontal gene transfer (23,27,33), the decline in performance over the course of cell culture (33,34,39), and, particularly in the case of products for human use, the requirement to remove the antibiotics from the final product (36,39). Thus, some alternative approaches have been developed, which we group here in four broad categories: complementation systems, toxin-antitoxin systems, partitioning systems and chromosomal integration methods. Complementation systems and toxin-antitoxin systems, in particular, are collectively referred to as plasmid addiction systems (32).

Complementation systems are the most common method of preserving a plasmid without using antibiotics. They are based on relocating an essential gene from the host genome to the plasmid. In practice, a copy of a gene that is essential for the microorganism is inserted into the plasmid, while the original gene, normally present in the genome, is deleted or knocked out. As a consequence, bacteria that lose the plasmid during cell division die (37). For instance, Voss and Steinbüchel (180) deleted the KDPG-aldolase gene in *Ralstonia eutropha*, rendering the bacteria unable to catabolize certain carbon sources through the Entner-Doudoroff pathway, and, at the same time, restored the KDPG-aldolase gene in a plasmid vector which also contained the cyanophycin synthetase gene. As a result, they were able to produce cyanophycin to the level of 40% (w/w), much higher than the level of 6.2% obtained with an antibiotic resistance system.

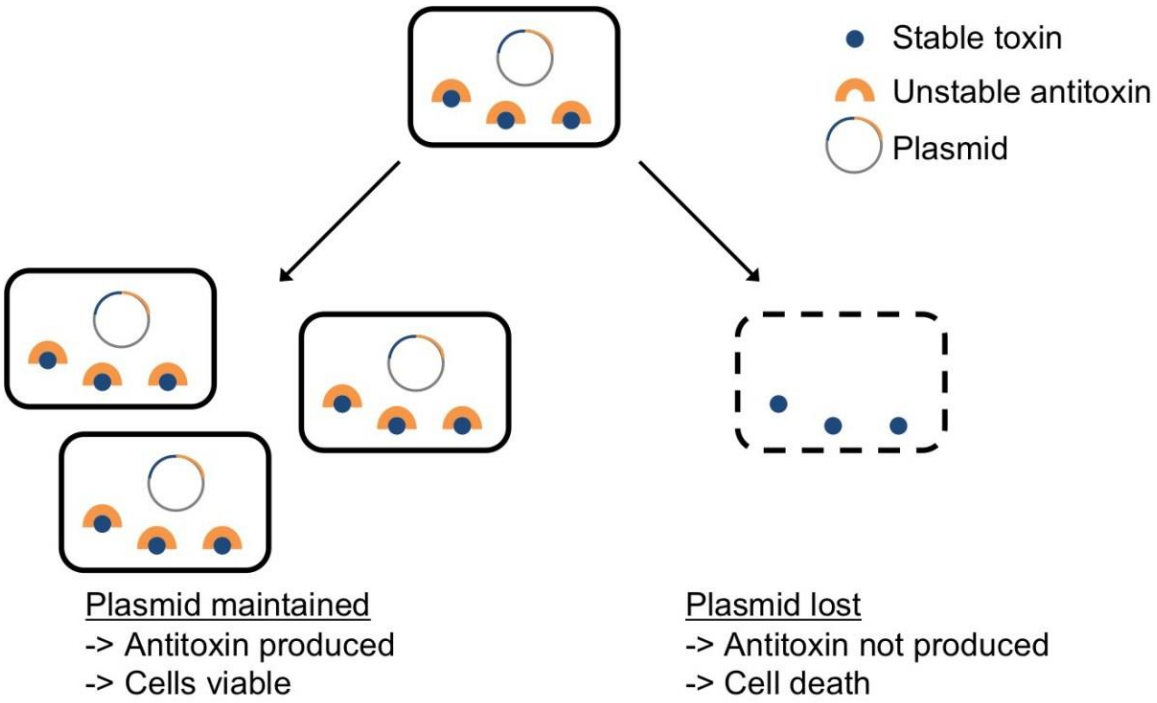
A particular type of complementation system is that of auxotrophic mutants, which are unable to synthesize a certain essential molecule (typically an amino acid or nucleotide) and grow but, if such molecule is supplied in the culture medium, they are able to grow again (23,31,36,37). In the context of complementation systems, auxotrophy is created by disrupting or deleting a key anabolic gene on the chromosome and simultaneously introducing a plasmid carrying a functional copy of the same gene. To cite a few examples, Ikeda et al. (181) introduced serine auxotrophy into a *Corynebacterium glutamicum* platform for tryptophan production; the product levels increased from 43 to 50 g/L in comparison with the antibiotic system; Squires et al. (182) introduced uracil and proline auxotrophy into *Pseudomonas fluorescens* to express a single chain antibody, obtaining the same levels as the antibiotic system; Vidal et al. (183) developed an *E. coli* system auxotrophic for glycine and capable of producing rhamnulose 1-phosphate, achieving a productivity level also comparable to the conventional system; and Kroll et al. (184)

employed a lysine-auxotrophic *E. coli* to produce cyanophycin, which reached 18% of dry cell weight at pilot scale.

A more elaborate version of the complementation system, called operator-repressor titration, has been recently described in the literature; in this system, the essential gene is present and functional in the genome, but repressed by an operator/repressor system artificially inserted into the genome. The system works in such a way that the plasmid relieves the repression of the essential gene, thus allowing the host cell to survive (36–38). Another variation of the complementation system is based on the regulation of an essential gene by a repressor whose expression is inhibited by a non-coding RNA (through antisense RNA interaction) encoded in the plasmid. An example described in the literature involves an *E. coli* strain whose essential *MurA* gene, involved in the formation of peptidoglycan, is under control of the Tet repressor protein. The expression of the repressor protein is inhibited by an RNAi naturally produced by plasmids with a ColE1 origin of replication. Thus, plasmid stability is ensured in the cell with no need to insert any DNA sequence into the plasmid for that purpose (37,38).

Another kind of plasmid stability strategy is the so-called toxin-antitoxin (TA) system, also known as post-segregational killing (PSK) system. It is based on two genes: one present in the plasmid or chromosome, that codes for a lethal toxic protein (the "toxin"), and the other inserted into the plasmid, that gives rise to another protein (or to a non-coding RNA) that inhibits the toxin (the "antitoxin") (23,35–38,185,186). A prime example of such a toxin-antidote mechanism is the *ccdA*/*ccdB* system, originally found in the F plasmid of *E. coli* (186). The *ccdB* gene codes for a protein that binds to DNA gyrase, and the *ccdA* gene, for an antidote protein. Thus, only the plasmid-bearing bacteria are able to divide (23,36,37,186). Another illustrative example is the *hok*/*sok* system, also discovered in an *E. coli* plasmid, in which the translation of the *hok* (host killing) gene produces a lethal toxin to the host, while the transcription of the *sok* (suppression killing) gene produces an RNA that binds to the *hok* mRNA, completely repressing the *hok* gene (37,186). It is worth noting that, in TA systems, the toxin must be relatively more stable than the antitoxin, so that, if the cell loses the plasmid during cell division, the residual antitoxin molecules in the cell degrade faster than the toxin molecules, and the cell dies (35,185,186). Figure 5 schematically illustrates this principle.

Figure 5 – Schematic diagram of toxin-antitoxin systems of plasmid maintenance.



Caption: Functioning mechanism of toxin-antitoxin systems of plasmid maintenance. In this illustration, the mother cell (on top) has a plasmid harboring a gene that codes for a stable toxin (blue) and a gene that codes for an unstable antitoxin (orange). Upon cell division, cells that inherit the plasmid survive, whereas cells that, by chance, lose the plasmid die because of the lasting toxin molecules. Source: ref. (187).

Successful applications of TA systems include the production of considerable levels of human interferon-β protein (188), silk-elastin-like protein (189) and BGL (190) in *ccdA*/*ccdB* systems based on *E. coli*, reaching 1.8 g/L in the case of silk-elastin-like protein and 1.0 g/L of protein in the case of BGL. Likewise, miniantibodies (191), astaxanthin (192) and poly-3-hydroxybutyrate (PHB) (193) were produced in *E. coli* with the *hok*/*sok* system; in the first case, a high titer of functional miniantibodies was obtained (3.3 g/L); in the second case, a level similar to that of the antibiotic system was achieved (385 mg/L); and in the last case, a very high level of PHB was attained (81 g/L).

A completely distinct plasmid stability mechanism exists for low-copy number plasmids: active partitioning. It is analogous to chromosome segregation, and requires three elements: a centromere-like DNA sequence present in the plasmid, an adaptor protein that binds to this sequence, and a motor protein (an ATPase or GTPase) that binds to the adaptor protein and, thanks to the energy provided by the hydrolysis of a nucleoside triphosphate, moves DNA inside the cell (194,195), in such a way that both daughter cells are guaranteed to receive plasmid copies. No similar mechanism has been discovered for

high copy number plasmids yet; it is generally thought that random distribution upon cell division is sufficient to ensure the presence of plasmid copies in both daughter cells, although this theory has been questioned by works that show the formation of clusters of plasmid molecules inside the cell, as well as the tendency of plasmids to localize at the cell poles, away from the nucleoid (194). Another issue related to multicopy plasmid stability is the process of multimerization, which, in practice, reduces the number of pDNA units to be transmitted during cell division, therefore hastening plasmid loss. To counter this process, multicopy natural plasmids such as ColE1 have a site-specific recombination system (Xer-*cer*) that converts multimers into monomers (196). Wilms et al. took advantage of this system by introducing the multimer resolution *cer* gene into a plasmid-based *E. coli* platform for production of carbamoylase, and obtained high levels of plasmid stability and enzyme yield (up to 3.8 g/L) (197).

A more definitive solution to plasmid segregational instability is chromosomal integration of the gene(s) of interest. This can be accomplished by a wide variety of techniques, such as homologous recombination, site-specific recombination, and transposon-mediated methods (198,199). However, chromosomal integration generally suffers from weaker expression levels than multicopy plasmids, given that most methods only allow the integration of one or a few gene copies at a time; in other words, chromosomal integration is undermined by low gene dosage. To circumvent this issue, some strategies have been developed, such as using strong promoters or carrying out integration multiple times (198,200). For example, Srinivasan et al. integrated one to three copies of an organophosphohydrolase gene into *Ralstonia eutropha*, and observed a linear increase in enzyme expression with gene dosage; the strain with a triple copy attained 4.3 g/L of enzyme (201). In another work (202), the same research group integrated the T7 RNA polymerase system as well as a single copy of the organophosphohydrolase gene under control of the T7 promoter into the bacterial chromosome; this time, they obtained an enzyme titer of 11.6 g/L. Also illustrative is the work of Gao et al. (203), which presents the construction of a pathway for poly(3-hydroxypropionate) synthesis from glucose in *E. coli* employing a hybrid strategy, which involved chromosomal integration and a plasmid stabilized by tyrosine auxotrophy; a product level of 25.7 g/L was achieved, about 2.5× higher than a two-plasmid strategy. In yet another example (204), Tyo et al. devised a high-throughput chromosomal integration technique, named CIChE, to introduce ~40

consecutive gene copies into *E. coli*'s genome; the technique was applied for lycopene and PHB production, attaining similar or higher titers than multicopy plasmids.

Lastly, a unique approach for increasing plasmid stability deserves mention: a cell immobilization strategy to trap two divergent plasmids in *E. coli* for production of L-carnitine. The cells were immobilized in carrageenan beads and submitted to continuous flow, with a significant increase in productivity when compared to free cells (from 0.07 to 0.71 $g \cdot L^{-1} \cdot h^{-1}$) (205).

2.3 Modeling and Simulation of Bioprocesses

Bioprocess modeling and simulation constitute a relatively new application of the Process Systems Engineering field, which "deals with the analysis, design, optimization, operation and control of complex process systems, as well as the development of model-based methods and tools that allow the systematic development of processes and products across a wide range of systems involving physical and chemical change" (206), in which biological change also takes place. In fact, process modeling and simulation have emerged in the late 1950s in the petroleum and chemical industry, when software capable of simulating chemical processes was developed in-house by petrochemical companies. Since then, process modeling and simulation have evolved significantly, and today are routinely used in these industries with the aid of commercial software such as Aspen Plus (207,208).

Process modeling consists of creating a representation of a process system with a predefined goal. This representation is generally mathematical, comprising a set of equations that describe certain aspects of the process system that are deemed relevant for the process goal. In general, models can be classified as mechanistic or empirical. Mechanistic models stem from physical or chemical mechanisms underlying the system, such as equation of mass and energy conservation; empirical models, on the other hand, are developed exclusively from experimental data, relating input and output values without knowledge of the underlying physical or chemical phenomena. In practice, most process models are hybrid, in the sense that they employ both mechanistic and empirical

equations (209). Moreover, when modeling chemical or biotechnological processes, it is convenient to define each major physical, chemical or biological transformation associated with a piece of equipment as a unit operation, so that the process system consists of a series of interconnected unit operations. As such, a process model can be created by first modeling each unit operation individually, and then linking the unit operation models together (through equations of mass and energy conservation).

Process simulation is subtly distinct from process modeling; it consists of the application of a process model to a particular set of input variables, and is normally carried out using computer-aided tools, so-called process simulators. As a consequence, process modeling and simulation (PMS) allow one to mimic a great number of experiments which would be expensive, time-consuming or downright impossible to be performed in practice. This, in turn, helps to design and compare prospective processes and to analyze, improve, operate or control existing ones. Moreover, a wealth of process simulation data such as mass and energy input and output values, type and size of equipment, unit operation times etc. can be associated with economic data and models to generate economic evaluations of production processes. Similarly, process simulation information can be combined with environmental data and tools to produce process environmental assessments, such as life cycle analyses. Consequently, a process can be designed, analyzed, improved or compared with various process configurations or alternatives with the aim of obtaining an economical, environmentally sound and technically feasible process (210). In addition, these design and analytical abilities of PMS are especially useful when dealing with biotechnological processes, given the sheer number of processing routes available (various hosts, strains, substrates, metabolic pathways, products and coproducts are available, not to mention the possibilities of modification at every level, through protein engineering, genetic engineering, metabolic engineering, etc.) (206,207).

Process improvement, in particular, is greatly facilitated by process modeling and simulation. It can be achieved by process optimization, which is the variation and selection of process parameters to improve one or several output values, usually with the aid of computational tools (211); process integration, which constitutes an approach to reduce the use of material and energy resources in a process, especially by recycling material streams and implementing heat recovery among unit operations (212); and process intensification, which may be defined as any drastic improvement in process

efficiency that substantially reduces equipment size, resource consumption or waste formation for a given throughput (213) (as such, process intensification can be regarded as an extension of process integration).

2.3.1 Challenges of Applying Modeling and Simulation to Biotechnological Processes

Bioprocesses have been growing in use since the 1990s, especially in sectors that manufacture high value-added products (such as pharmaceuticals and their intermediates) (207). More recently, however, the interest in the use of bioprocesses for producing fuels, plastics and bulk chemicals has been reinvigorated, for both economic and environmental reasons (179,207,214). However, contrary to petrochemical companies, biotechnological companies have been somewhat slow to adopt PMS tools for process development and improvement (215,216). One reason for this reticence is that a great number of biotechnology companies manufacture patent-protected, high-value-added products, especially pharmaceuticals and, as such, heavily rely on fast development and implementation of production processes, so as to make the most of the limited period of patent exclusive rights, rather than on process improvement (207). Moreover, minimizing the operational cost of producing a patent-protected, high value-added biopharmaceutical product is less critical than that of a low value-added chemical product, given that R&D and clinical approval have a large contribution to the final production cost in the former case. In addition, continuous process improvement in pharmaceutical biotechnology is discouraged by strict government regulation: once a pharmaceutical product and its production process are authorized, the production process can no longer be changed without a new and expensive approval process (217).

Many intrinsic characteristics of bioprocesses have also hindered the application of PMS to the biotechnology sector: biological phenomena are often nonlinear and time-dependent, usually conducted in batch or fed-batch mode, as opposed to chemical processes, conducted in continuous, steady-state mode (211); raw materials and process streams have complex and sometimes variable or ill-defined composition (particularly in the case of biorefinery processes, that use plant-based feedstocks) (208,211); live cells

are susceptible to change over time (due to genetic events and physiological alterations) (217) and are heterogeneous over space as well (216); physical and chemical properties of biological components and mixtures are complex and often unknown (207,208,211); models of bioprocess unit operations are often highly empirical, thus having limited range of validity and predictive power (208,211,216,218); these models often exhibit low reproducibility and are not transferable even to similar processes without a new, complete round of parameter estimation (216); and the primary experimental data available to bioprocess models are limited to nutrients, biomass and a few metabolites, which can be measured online, whereas the impressive advances in genomics, transcriptomics, proteomics and metabolomics achieved in recent years have not, by and large, been incorporated in these models (211,216).

Yet another reason for the lack of enthusiasm for PMS in the biotechnology field has been cultural, given the widespread belief that biology alone can solve the problems of bioprocess development. Undeniably, biology has progressed tremendously during the last decades, which allowed astonishing improvements of enzymes (protein engineering) and cells (genetic and metabolic engineering, synthetic biology, etc.) (215). Nevertheless, greater integration between these lower levels of development (enzymes and cells) and the bioprocess engineering level is indispensable, given that the results of development efforts at lower levels parametrize and constrain development at higher levels and, conversely, knowledge from higher levels of development can inform and guide development at lower levels. As a consequence, the application of PMS and economic and environmental evaluations on early stages of bioprocess development is vital to spur closer integration among biotechnology areas and thus create biotechnological processes which are technically, economically and environmentally viable (219).

In the literature, the PMS tools most commonly used to model and simulate bioprocesses are SuperPro Designer® (Intelligen, USA) and Aspen Plus® (AspenTech, USA). The former is more focused on biotechnological processes, while the latter is a more generic chemical process simulator. Both programs have the ability to simulate batch processes and unit operations typical of bioprocesses, such as crossflow filtration and chromatography. Moreover, both products have built-in tools for economic analysis, and have been used to model, simulate and economically evaluate the production of industrial enzymes (cf. Section 2.4.3).

2.3.2 Strategies for Bioprocess Intensification

As indicated earlier, improvements through optimization, integration and intensification are required to make bioprocesses economically and environmentally attractive, especially for production of biobased fuels and bulk chemicals, whose profit margins are usually tight. Bioprocess intensification approaches, in particular, open the door to cost reductions of several orders of magnitude. Although intensification possibilities are countless, they can be categorized according to the target of improvement, as shown in Table 2. Within this framework proposed by Woodley (215), improvements at the level of enzyme or cell are grouped in the "Biocatalyst Engineering" category, whereas improvements at the level of unit operations are grouped in the "Process Engineering" category. Typical examples of intensification through Biocatalyst Engineering would be increasing enzyme activity by means of Protein Engineering or increasing cell tolerance to a certain metabolite (e.g. ethanol) by means of Evolutionary Engineering. On the other hand, intensification through Process Engineering can be divided in two major approaches: compartmentalization, e. g. using two liquid phase systems to allow high concentrations of a hydrophobic substrate and/or product (for instance, toluene oxidation by *Pseudomonas putida*); and hybrid unit operations, e. g. simultaneous fermentation and *in situ* product removal (for instance, ABE fermentation and gas stripping of organic products from *Clostridium*) (215).

Naturally, the combination of diverse intensification approaches offers even further intensification opportunities; for instance, Shen et al. (220) constructed a metabolic pathway for synthesis of 1-butanol in *E. coli* using genes from multiple organisms with the aid of molecular biology techniques, and then cultivated the bacteria in a bioreactor fed with a nitrogen gas flow so as to shift bacterial metabolism to anaerobiosis and extract the alcohol product *in situ* at the same time (1-butanol is toxic to the bacteria above 10 g/L). As a result, high values of titer (30 g/L) and productivity (0.18 g/L/h) were obtained.

Table 2 – Classification of bioprocess intensification strategies

Field	Target	Approach	Examples
Biocatalyst engineering (Biology)	Enzyme	Protein Engineering	- Increase enzyme activity - Reduce end-product inhibition
	Cell	Genetic Engineering Metabolic Engineering Synthetic Biology Evolutionary Engineering Gene Editing	- Increase cell tolerance to metabolite concentration in the medium - Move gene or metabolic pathway to a better host
Process Engineering	Reactor	Compartmentalization	- Use of two-liquid phase system to allow operation at high concentrations of poorly water-soluble substrate and product
	Multiple unit operations	Hybrid unit operations	- Simultaneous fermentation and gas stripping to reduce product toxicity and increase product concentration

Source: adapted from (215).

2.4 Economic Analysis of Bioprocesses

The systematic economic assessment of a bioprocess is crucial for its market success, especially if the product in question is a biobased fuel or bulk chemical, which face narrow profit margins and have to compete with inexpensive petrochemical analogs (221). For analytical purposes, the economic assessment of a production process is decomposed in two major areas: the cost analysis and the profitability analysis.

2.4.1 Cost Analysis

The costs of a production process are divided in two major categories: *capital costs* (also called capital expenditures or CAPEX) and *operating costs* (also called operating expenditures or OPEX). Capital costs are related to the investment necessary to build and start up the plant, whereas operating costs account for ongoing production costs such as

those of raw materials and utilities. The components of each cost category are detailed next, following the definitions of Heinzle, Biwer and Cooney (210).

Capital Costs:

a) Direct Fixed Capital (DFC): includes the purchase and installation costs of all the equipment employed in the process, such as reactors, storage tanks, filters, centrifuges and chromatographic columns; all the process piping and electrical installation; and the cost of buildings, engineering and construction.
b) Working Capital: corresponds to the money the company needs to maintain its day-to-day operations.
c) Startup and Validation Costs: the process of starting up a new plant or production process can be extremely complex and critical, especially when dealing with a novel and intricate process. Consequently, these costs may need to be taken into consideration.
d) Up-Front R&D and Royalties Cost: comprises the costs of R&D and, in the case of using a patent from another corporation or entity, the royalties that are due, before the beginning of production.

Operating Costs:

a) Costs of raw materials: encompasses the costs of all the chemical reagents used in the process, including the components of culture media, buffer solutions and cleaning solutions.
b) Labor Cost: comprises the cost of all the personnel related to manufacturing, including managers and other administrative positions.
c) Utilities cost: involves the cost of process water, electricity, cooling and heating agents such as steam and chilled water.
d) Cost of waste treatment and disposal: involves the cost of treatment and disposal of aqueous and solid wastes
e) Cost of Consumables: covers the cost of materials that are subject to fouling or damage over time, and thus have to be periodically discarded and replaced. Typical examples are filtration membranes, chromatographic resins and disposable reactors.

f) Cost of Laboratory and Quality Control: includes all the costs associated with the off-line analysis of physical, chemical and biological properties of the final product, as well as the analysis of raw materials and intermediate samples.
g) Facility-dependent Cost (also known as Facility Overhead Cost): comprehends the cost of depreciation and maintenance of equipment, insurance, local taxes and other factory expenses not directly related to manufacturing, such as accounting, security, cafeteria, etc.
h) Miscellaneous: embraces R&D, validation, marketing and sales activities associated with the bioproduct.

A crucial parameter to economically evaluate a production process is the *unit production cost* (UPC) of the final product. It is defined as the *annual operating cost* (the sum of all operating costs over one year) divided by the *annual production rate* (the total amount of product generated over one year) (210). As such, the UPC does not directly account for the capital costs, although some of the capital costs end up seeping into the facility-dependent cost.

2.4.2 Profitability Analysis

The revenue side of the economic analysis depends primarily on the selling price of the final product (and coproducts, if any), and its annual production rate; the multiplication of these values leads to the *annual revenues*. To determine if a process is in fact economically attractive, the following parameters can be calculated (as per (210)):

a) Gross Profit: defined as the difference between the annual revenues and the annual operating cost.
b) Gross Margin: defined as the gross profit divided by the annual revenues
c) Net Profit: equal to the gross profit minus income taxes; the annual depreciation may be added to the net profit as well.
d) Return on Investment (ROI): consists of the net profit divided by the total capital costs.

e) Payback Time: is the time required for the total capital costs to be counterbalanced by the accumulated annual profits. It is calculated by dividing the total capital costs by the net profit, and given in number of years.

In order to make the biobased production of fuels and chemicals profitable, there is a growing consensus that it is indispensable to maximize the annual revenues by drawing upon the value of coproducts and cogeneration of energy; in other words, a true *biorefinery approach* is called for (222–224).

2.4.3 Economic Analyses of the Production of Industrial Enzymes

More than half of the economic evaluations of enzyme production published during the last 20 years concern *Trichoderma* cellulase production, as can be seen in Table 3. Works concerning enzymes other than cellulases are mostly based on solid-state cultivation of filamentous fungi. A good number of these consider very small production scales, such as shake-flasks and tray reactors, which makes them of limited use for cost projections on industrial scales. Several articles also present cost results in terms of enzymatic activity only, without indicating enzyme specific activity, which makes comparisons between different enzymes impractical. Moreover, some analyses present their cost results in local currencies, without indicating the exchange rate to the U. S. dollar at the time of the study, which further complicates direct comparisons. In any case, the enzyme cost estimated in those works that consider industrially significant scales varies widely: from 3.3 US$/kg of protein (225) to the range of 70-217 US$/kg of protein (226).

With respect to the economic analyses of cellulase production, most assume production by filamentous fungus *Trichoderma reesei* through submerged fermentation in very large reactors (usually 300 m^3 and even 940 m^3 in one case), using either Aspen Plus or SuperPro Designer to model and simulate the bioprocess. In several of these works, the assessment of cellulase production is actually a minor part of an extensive economic evaluation of lignocellulosic ethanol (or another end-product) production, assuming that cellulases are produced on-site. In fact, a good number of these analyses report cellulase cost in terms of dollars per gallon (or liter) of ethanol, making it difficult to compare

values from different sources, given that this unit depends not only on the enzyme cost per se, but also on many other factors such as choice of feedstock, enzyme loading and process yield (10). In any case, within those works that report enzyme cost in dollars per kg of protein, estimations range from 3.8 to 10.1 US$/kg, with the exception of the analysis of Zhuang et al. (227), which estimated 15.0 US$/kg and 40.0 US$/kg for a solid state fermentation process and a submerged fermentation process, respectively, using *Clostridium thermocellum* instead of *T. reesei*. It is also worth noting that many evaluations of cellulase production found in the literature were either performed by the National Renewable Energy Laboratory (NREL) or based on the Aspen Plus model made freely available by NREL, which may somewhat skew the economic landscape of cellulase production.

Table 3 – Cost of industrial enzymes found in the literature from 1999 to 2018.

Ref.	Enzyme	Enzyme Cost	Production Scale	Microorganism	Production Mode	Process Simulator
(20)	Cellulase	Off-site (glucose-fed) 0.78 US$/gal EtOH On-site (glucose-fed) 0.58 Integrated (cellulose-fed) 0.23	3105 ton/yr	*T. reesei* (inferred)	SmC/ on- and off-site	None
(7)	Cellulase	4.24 US$/kg	300 m^3 reactor (inferred)	*T. reesei*	SmC/ on-site (NREL-based)	Aspen Plus
(228)	Cellulase	0.078 EUR/L EtOH (off-site)	Unclear	*T. reesei*	SmC/ on- and off-site	Aspen Plus
(229)	Cellulase	3.8 - 6.7 US$/kg on-site 4.0 - 8.8 US$/kg (off-site)	300 m^3 reactor (inferred)	*T. reesei*	SmC/ on- and off-site (NREL-based)	Aspen Plus
(10)	Cellulase	10.14 US$/kg	300 m^3 reactor	*T. reesei* (inferred)	SmC/ off-site	SuperPro Designer
(230)	Cellulase	0.42-0.53 SEK/ L EtOH MESP 4.71 - 4.82 SEK/L	37-121 m^3 reactor	*T. reesei*	SmC/ on-site	Aspen Plus
(227)	Cellulase	16 US$/kg (SSC) 40 US$/kg (SmC) >90 US$/kg (market price)	940 m^3 reactor	*C. thermocellum*	SmC and SSC/ on- and off-site	SuperPro Designer
(231)	Cellulase	0.12 £/L EtOH	300 m^3 reactor (inferred)	*T. reesei*	SmC/ On-site (NREL-based)	SuperPro Designer, MATLAB
(232)	Cellulase	5.38 US$/kg	300 m^3	*T. reesei*	SmC/ On-site	Aspen Plus

Ref.	Enzyme	Enzyme Cost	Production Scale	Microorganism	Production Mode	Process Simulator
					(NREL)	
(233)	Cellulase	4.23 US$/kg	300 m^3	*T. reesei*	SmC/ On-site (NREL)	Aspen Plus
(234)	Cellulase	5.38 US$/kg	300 m^3	*T. reesei*	SmC/ On-site (NREL)	Aspen Plus
(9)	Cellulase	4.24 US$/kg	300 m^3	*T. reesei*	SmC/ On-site (NREL)	Aspen Plus
(67)	Cellulase	6.16 US$/kg	300 m^3	*T. reesei*	SmC/ On-site (NREL)	Aspen Plus
(235)	Manganese peroxidase and laccase	2.27 EUR/kU 1.08 EUR/kU (only operating cost)	30 L	White rot fungi (*Irpex lacteus* and *Ganoderma lucidum*)	SmC	None
(236)	Lipase	1.06×10^6 IDR/kg Approx. 15 US$/kg	4.3 ton/yr	*Aspergillus niger*	SSC	None
(225)	Protease	1.66 US$/kg (50% moisture)	30.6 ton/yr (50% moisture)	*Brevibacterium luteolum*	SSC	None
(237)	Pectinase (poly-galacturonase)	858-1605 INR/L 3000 U/L Approx. 23 US$/mg	30 m^3 of concentrate/yr 75 L reactor in SmC case	*Aspergillus carbonarius*	SmC & SSC	None

Ref.	Enzyme	Enzyme Cost	Production Scale	Microorganism	Production Mode	Process Simulator
(238)	Laccase	0.40 – 70 EUR/kU	Shake-flasks and trays	White-rot fungi (*Trametes pubescens*)	SmC & SSC	None
(239)	Lipase	42-131 US$/L	100 m^3 of concentrate/yr	*Penicillium restrictum*	SmC & SSC	Own program
(240)	Amylases, cellulases, xylanases, proteases	10.4 US$/kg	644 kg (50%w enzymes)/yr	*Aspergillus awamori*	SSC	SuperPro Designer
(226)	Amylases, cellulases, xylanases, proteases	70-217 US$/kg (no coproduct revenues) 57-139 US$/kg (with coproduct revenues)	100 ton/yr	*Aspergillus awamori*	SSC	SuperPro Designer
(241)	Laccase	0.16-0.25 EUR/kU	200 L	*Pichia pastoris*	SmC	SuperPro Designer

Caption: Production cost of industrial enzymes obtained from techno-economic analyses published in the literature from 1999 to 2018. For each reference, the type of enzyme and its production cost are indicated, as well as the production scale, the microorganism employed for enzyme expression, the process mode (submerged culture or solid-state culture, abbreviated as SmC and SSC, respectively) and the process simulator used (if any). Source: this work.

3 CLONING AND EXPRESSION OF β-GLUCOSIDASE IN *E. COLI* BL21(DE3) AND SE1

3.1 Methodology

3.1.1 Microorganisms

Strains DH10B, DH5α, BL21(DE3), CYS21 and SE1 of *Escherichia coli* were employed in the experimental work. CYS21 and SE1 derive from DH10B and BL21(DE3), respectively, and are part of the StabyExpress™ T7 kit acquired from Delphi Genetics by Prof. Sindélia Freitas, a collaborator who was then working at the Brazilian Bioethanol Science and Technology Laboratory (CTBE). These two strains distinguish themselves by possessing the *ccdB* gene integrated into their chromosome. Moreover, they were supplied to us carrying plasmids: CYS21 + pStaby and SE1 + pET28a-*cel5A*-*ccdA*. The other three strains, DH10B, DH5α and BL21(DE3), were obtained from the cell bank of our own research group in the Department of Chemical Engineering of the University of Sao Paulo.

3.1.2 Plasmids

Subcloning started from two plasmids: a pET28a plasmid (Novagen) carrying the *TpBgl1* gene (inserted in the multiple cloning site) provided by Prof. Fábio Squina and generated at the Brazilian Bioethanol Science and Technology Laboratory (CTBE); and the pStaby1.2 plasmid, which is part of the StabyExpress™ T7 kit (DelphiGenetics), and consists essentially of a pET21a plasmid combined with the *ccdA* gene between the lac operon and the T7 promoter. The pET28a-*TpBgl1* plasmid was provided to us under the form of purified pDNA, while the pStaby1.2 plasmid was provided to us under the form of CYS21 cells harboring the plasmid. As mentioned above, the plasmid pET28a-*cel5A*-*ccdA* was

also present in the SE1 cells provided to us, although it was not used in this work (in fact, it was later eliminated from the SE1 cells).

3.1.3 Cloning of β-glucosidase into *E. coli* BL21(DE3) and SE1

Cloning of the *TpBgl1* gene from hyperthermophilic bacterium *Thermotoga petrophila* into strains BL21(DE3) and SE1 of *E. coli* was carried out through the method of restriction enzyme digestion, as detailed in the next sections.

3.1.3.1 Generation of Chemically Competent *E. coli* DH10B Cells

The pET28a-*TpBgl1* plasmid was transformed into *E. coli* DH10B cells for the purpose of preservation and amplification. First, chemocompetent DH10B cells were created following the calcium chloride protocol of ref. (242) (except that the buffer component of the calcium chloride solution was changed from PIPES to HEPES, and the volume of bacterial culture, scaled-down from 400 mL to 100 mL). At the end of the procedure, the chemocompetent cells in microfuge tubes were stored in the -80 °C freezer.

3.1.3.2 Transformation of the pET28a-*TpBgl1* Plasmid into Competent *E. coli* DH10B Cells

For bacterial transformation, competent cells were thawed on ice (20-30 min) and a mass of 20 ng of pure plasmid was then added to 50 µL of competent cells in a microfuge tube, also on ice. The mixture was incubated on ice for 30 min, and then heat-shocked at 42 °C, in a water bath, for 45 s; next, the tube was placed back on ice for 2 minutes. A volume of 950 µL of SOC medium (Table 4) was then added, and the tube, incubated in an orbital

shaker at 37 °C, 150 rpm, for 1h. At last, a sample of 100 µL was plated onto Lysogeny Broth (LB) agar plates (Table 4) with 30 mg/L of kanamycin sulfate, and the plates, incubated at 37 °C overnight. In the next morning, the occurrence of colonies indicated a successful transformation.

Table 4 – Composition of complex media: SOC, LB and 2×YT.

	Concentration		
Component	**SOC** (243)	**LB** (244)	**2×YT** (245)
Tryptone	20 g/L	10 g/L	16 g/L
Yeast Extract	5 g/L	5 g/L	10 g/L
NaCl	10 mM	10 g/L	10 g/L
KCl	2.5 mM	q. s.	q. s.
$MgCl_2$	10 mM	-	-
Glucose	20 mM	-	-
Bacteriological Agar[α]	-	15 g/L	-
Ultrapure Water[β]	q. s.	q. s.	q. s.

[α] Only when preparing plates
[β] Type I water, obtained using a Milli-Q® Reference Water Purification System
Source: this work.

3.1.3.3 Preservation of the pET28a-*TpBgl1* Plasmid in a Glycerol Stock of *E. coli* DH10B Cells

Single colonies of *E. coli* DH10B + pET28a-*TpBgl1* were picked from the transformation plate and inoculated into 10 mL of LB medium with 30 mg/L of kanamycin, in 50 mL centrifuge tubes (see Table 5 for the relationships among strains, plasmids and selection markers). The tubes were then incubated in an orbital shaker at 37 °C and 250 rpm, overnight. In the next morning, when the culture attained late exponential phase/early stationary phase (DO ~ 1.5-2.0), the culture broth was mixed with an equal volume of

30% (v/v) glycerol solution, dispensed in cryotubes, and finally stored in the -80 °C freezer.

Table 5 – *E. coli* strains and plasmids in this work

***E. coli* strain**	**Plasmid**	**Selection Marker**	**Marker Concentration**
DH10B	pET28a-*TpBgl1*	kanamycin	30 mg/L
CYS21 (*ccdB*$^+$)	pStaby1.2	ampicillin (redundant)	—
DH5α	pStaby1.2-*TpBgl1*	ampicillin	50 mg/L
BL21(DE3)	pStaby1.2-*TpBgl1*	ampicillin	50 mg/L
SE1(*ccdB*$^+$)	pStaby1.2-*TpBgl1*	ampicillin (redundant)	—

Caption: *E. coli* strains and plasmids employed or produced in this work. The selection markers corresponding to each expression platform, as well as their concentration in culture media, are also indicated. Source: this work.

3.1.3.4 Amplification, Extraction and Purification of Plasmids pET28a-*TpBgl1* and pStaby1.2

Plasmids pET28a-*TpBgl1* and pStaby1.2, carried by strains DH10B and CYS21, respectively, were amplified, extracted and purified as per the QIAprep Spin miniprep kit high-yield protocol (246). This protocol employs 5 mL of 2×YT medium (Table 4) instead of LB medium to grow cells, and then proceeds to the standard protocol of the QIAprep Spin Miniprep kit (QIAGEN) to extract and purify the plasmid DNA (pDNA). It should be noted that kanamycin sulfate was added to the level of 30 mg/L for the growth step of *E. coli* DH10B + pET28a-*TpBgl1*, so as to guarantee plasmid maintenance. At the end of the purification protocol, a sample of 1-2 µL was taken and analyzed with a nanospectrophotometer (DeNovix DS-11) to verify the purity and concentration of the pDNA.

3.1.3.5 Restriction Digestion of Plasmids pET28a-*TpBgl1* and pStaby1.2

For actually subcloning the *TpBgl1* gene into the pStaby1.2 plasmid, a double, sequential restriction enzyme digestion of plasmids pET28a-*TpBgl1* and pStaby1.2 was performed, using restriction enzymes NheI and EcoRI (both from Thermo Fisher Scientific). The protocol suggested by the "Double Digest Calculator" on the Thermo Fisher Scientific website (247) was observed, except that the reaction volume was scaled-up and that longer incubation times were employed, as explained below.

The first digestion was performed in a 1.5 mL microfuge tube, with 1× Tango buffer (a low salt concentration buffer from Thermo Fisher Scientific), 1.5 µL of restriction enzyme NheI (15 U), and pDNA (1.5 µg), making up a total volume of 30 µL. Initially, the tube was incubated in a water bath at 37°C for 2h. When the first digestion was completed, 4.1 µL of 10× concentrated Tango buffer and 1.5 µL of EcoRI (15 U) were added to the reaction mixture, so that the Tango buffer concentration was doubled (2×) and the final volume was increased to 35.6 µL. Next, the microfuge tube was again incubated at 37 °C for 2h. The reaction was then interrupted by transferring the microfuge tube to a dry bath at 65 °C, and incubating it there for 20 min. Finally, thermosensitive alkaline phosphatase (TSAP, 1 MBU*/µL, from Promega) was added to the digested vector (1.5 MBU = 1.5 µL), and the mixture, incubated in a water bath at 37 °C for 15 min; the TSAP enzyme was also inactivated in a dry bath, at 74 °C for 20 min†.

A single enzyme digestion, using only the enzyme NheI and scaling-down the reaction volume to 20 µL, was also performed, in order to generate linear versions of the plasmids pET28a-*TpBgl1* and pStaby1.2. In this case, after the 2-h incubation with NheI, the tubes were directly transferred to the 65 °C dry bath and incubated for 20 min.

---

* MBU stands for Molecular Biology Unit. One Molecular Biology Unit is defined as the amount of enzyme required to dephosphorylate 1 µg of linearized pGEM®-3Zf(+) DNA in 15 minutes at 37 °C in any 1× Promega Restriction Enzyme Buffer, except Buffer F.

† The use of TSAP was actually superfluous, since the digested vector did not have blunt or otherwise complementary ends.

3.1.3.6 Agarose Gel Electrophoresis of pET28a-*TpBgl1*

The products of double and single restriction enzyme digestion described above were subjected to agarose gel electrophoresis, so as to confirm the correct cutting patterns and, in the case of the pET28a-*TpBgl1* plasmid, to separate the DNA insert (that is, the *TpBgl1* gene) as a distinct gel band for later extraction and purification. For the most part, the protocol of ref. (248) was observed; mini-gels containing 10 g/L of low electroendosmosis agarose were prepared using 50 mL of TAE buffer and 5 µL of the fluorescent stain SYBR Safe (10,000×, Thermo Fisher Scientific). Before being loaded onto the gel, DNA samples were mixed with loading buffer 6× (Sinapse Biotecnologia). The DNA molecular weight (MW) standard used was the 1 kb Plus DNA Ladder from Thermo Fisher Scientific. Finally, electrophoresis was conducted under 80-100V, for approximately 1 h, employing a horizontal electrophoresis chamber (Loccus Biotecnologia LCH 12×14).

3.1.3.7 DNA Insert Extraction and Purification

The DNA band identified as the insert in the agarose gel was cut out with the aid of a sterile scalpel, and then extracted and purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences), following the manufacturer's protocol for gel band extraction and purification.

3.1.3.8 Ligation Reaction

The double-digested pStaby1.2 plasmid and the purified DNA insert were ligated with the enzyme T4 DNA ligase (Promega); three 0.2-mL microtubes were prepared, with different vector : insert molar ratios: 1:1, 1:3 and 1:0 (negative control). The total amount of DNA (insert plus vector) used in all reactions was equal to 100 ng; and the total reaction

volume, 10 μL. The reaction tubes were then incubated overnight in the refrigerator (4-8 °C).

3.1.3.9 Transformation of the Ligation Product (pStaby1.2-*TpBgl1*) into *E. coli* DH5α Cells

The product of the ligation reaction was transformed into chemocompetent DH5α cells, which were derived through the same calcium chloride protocol described earlier (Section 3.1.3.1). The transformation of these cells with the ligation product also followed the corresponding protocol described earlier (Section 3.1.3.2), save that 1-5 μL of ligation product was directly added to the competent cells, instead of 20 ng of purified pDNA, and that LB plates with 50 mg/L of ampicillin (rather than kanamycin) were used at the final step. Four of the resulting colonies were then picked and inoculated into LB medium with 50 mg/L of ampicillin (henceforth abbreviated as LB-amp) to produce a glycerol stock, following the protocol of Section 3.1.3.3. Next, each LB glycerol stock was scraped with an inoculation loop, with which new LB-amp plates were streaked. At last, the ligation product was amplified, extracted and purified according to the protocols detailed in Section 3.1.3.4, except that ampicillin 50 mg/L was used instead of kanamycin 30 mg/L.

3.1.3.10 Restriction Digestion of the Ligation Product (pStaby1.2-*TpBgl1*)

To verify that the subcloning procedure was successful, the purified ligation product was subjected to single and double restriction enzyme digestion (in separate reaction tubes), following the protocol described earlier (Section 3.1.3.5).

3.1.3.11 Agarose Gel Electrophoresis of the Ligation Product (pStaby1.2-*TpBgl1*)

The products of single and double restriction digestion were submitted to agarose gel electrophoresis as described earlier (Section 3.1.3.6), to confirm the presence of the DNA insert (*TpBgl1*) in the pStaby1.2 plasmid.

3.1.3.12 Transformation of pStaby1.2-*TpBgl1* into *E. coli* BL21(DE3) Cells

Having confirmed that the ligation product was in fact the pStaby1.2-*TpBgl1* plasmid, it was transformed into chemocompetent BL21(DE3) cells, which were generated through the same calcium chloride protocol described earlier (Section 3.1.3.1). The transformation of these cells with the plasmid also followed the protocol already presented (Section 3.1.3.2), with the exception that LB-amp plates instead of LB-kan plates were used at the final step.

3.1.3.13 Preservation of *E. coli* BL21(DE3) + pStaby1.2-*TpBgl1* Cells in an LB Glycerol Stock

A single colony from the transformation plate was picked and inoculated into LB-amp to produce a glycerol stock, following the protocol already described (Section 3.1.3.3).

3.1.3.14 Transformation and Selection of pStaby1.2-*TpBgl1* in *E. coli* SE1

Given that we did not have competent *E. coli* SE1 cells devoid of plasmid, but rather SE1 + pET28a-*cel5A*-*ccdA*, the following procedure was carried out to obtain SE1 cells harboring only the pStaby1.2-*TpBgl1*:

a) SE1 cells carrying the pET28a-*cel5A*-*ccdA* plasmid were made competent through the calcium chloride protocol described earlier (Section 3.1.3.1);
b) these cells were then transformed with the pStaby1.2-*TpBgl1* plasmid and plated onto an LB-amp plate, following protocols already presented (Section 3.1.3.2);
c) a single colony from the transformation plate was picked to seed 50 mL of LB containing 100 mg/L of ampicillin in a 250 mL flask, and the cells were incubated in an orbital shaker at 37 °C, 250 rpm, overnight;
d) in the next morning, the culture broth was serially diluted by a factor of 10^7 (7 dilutions of 10x) using Phosphate Buffered Saline (PBS; see Table 6), and 100 µL were plated onto an LB-amp plate, which was incubated at 37 °C overnight;
e) finally, the resulting plate was replica-plated (249) onto two kanamycin plates and one ampicillin plate. The ampicillin plate served as a positive control, whereas the kanamycin plates served to indicate which colonies, if any, harbored the pET28a-*cel5A*-*ccdA* plasmid (which confers kanamycin resistance), on top of the pStaby1.2-*TpBgl1* plasmid. The replica plates were incubated at 30 °C overnight;
f) on the next day, a single colony present in the LB-amp plate and with no correspondent in the LB-kan plate was then picked and inoculated into LB medium to produce a glycerol stock, following the corresponding protocol already described (Section 3.1.3.3).

3.1.3.15 Expression of β-glucosidase in *E. coli* BL21(DE3) and SE1 Grown in LB

A qualitative assay of β-glucosidase (BGL) expression by BL21(DE3) + pStaby1.2-*TpBgl1* and SE1 + pStaby1.2-*TpBgl1* cells, henceforth referred to as BL21(DE3)/BGL and

SE1/BGL, respectively, was carried out in LB medium. First, a plate of each strain was obtained by scraping the corresponding glycerol stock with an inoculation loop and then streaking an LB plate, which was subsequently incubated at 37 °C for 16 h. A single colony was then picked and inoculated into 50 mL of LB in a 250-mL Erlenmeyer flask. Next, the flasks were incubated in the shaker at 37 °C and 250 rpm, for 16h. Then, 1 mL of the broth was used to seed a new flask containing 50 mL of LB, also incubated at 37 °C and 250 rpm. Its OD_{600} was followed until mid-to-late exponential phase ($OD_{600} \sim 0.6$) was reached, at which point 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce the expression of the recombinant enzyme. The culture was further incubated for 3h under the same conditions; afterwards, a sample of 10 mL was taken from the broth and centrifuged at 12,000 g for 5 min, at 4 °C. Finally, the supernatant was decanted and discarded, and the pellet, stored at -20 °C. It should be noted that, when dealing with the BL21(DE3)/BGL strain, ampicillin 50 mg/L was added to the medium at every growth step (including plating).

Table 6 – Composition of Phosphate Buffered Saline (PBS).

Component	**Concentration (g/L)**
NaCl	8.00
KCl	0.20
Na_2HPO_4	1.44
KH_2PO_4	0.24
HCl	to pH = 7.2
Ultrapure Water	q. s.

Source: ref. (250).

3.1.3.16 Cell Lysis and β-glucosidase Extraction

The cell pellet frozen at -20°C was thawed on ice and resuspended in chilled lysis buffer (Table 7), using the same volume as the original sample volume. The suspension was then

incubated on ice for 10 min and sonicated for 5 minutes (Branson Analog Sonifier 250A; 5 cycles of 1 minute ON + 1 minute OFF, duty cycle = 50%, output = 5, using the built-in ½ inch horn), on ice. The resulting cell lysate was then centrifuged at 12,000 *g* and 4 °C, for 5 min, and the supernatant (from this point on referred to as cell extract) was recovered by decantation.

Table 7 – Composition of the cell lysis buffer

Component	**Concentration**
Tris-HCl pH 7.5	50 mM
NaCl	100 mM
EDTA disodium	1 mM
DTT	1 mM
Hen Egg White Lysozyme	0.3 g/L
SigmaFAST™ S8830 EDTA-free Protease Inhibitor Cocktail[β]	1 tablet/200 mL
Ultrapure water	q. s.

[β]Only used in quantitative expression assays. Source: this work.

3.1.3.17 β-glucosidase Activity Assay

BGL activity was evaluated through the *p*-nitrophenyl β-D-glucopyranoside (pNPG) assay. Cell extract (4 µL), pNPG (1 mM) and 50 mM citrate buffer pH 6.0 were mixed, making up a volume of 400 µL, in a screw-top microfuge tube. The tube was then incubated in a dry bath at 90 °C, for 4 min; the reaction was terminated by transferring the tubes to an ice bath and immediately adding 400 µL of Na_2CO_3 1 M. Finally, the absorbance was measured in a spectrophotometer (Quimis Q898U2M5), at a wavelength of λ = 405 nm.

This protocol was adapted from (114), scaling up the reaction volume and selecting pH and temperature conditions within the optimal ranges determined by the authors.

Table 8 – Composition of M9-agar medium.

Component	Concentration	Sterilization Method	Stock Solution
KH_2PO_4	3.0 g/L	Autoclaved	10×
NH_4Cl	1.0 g/L		
$Na_2HPO_4 \cdot 7H_2O$	12.8 g/L		
NaCl	0.5 g/L		
Glucose	20 g/L	Filter-sterilized	500 g/L
$CaCl_2$	0.1 mM	Autoclaved	1 M
$MgSO_4 \cdot 7H_2O$	2.0 mM	Autoclaved	1 M
Bacteriological Agar	15 g/L	Autoclaved	-
Distilled Water	q. s.	Autoclaved	-

Note: distilled water is used instead of ultrapure water in this medium. Source: ref. (252).

3.1.3.18 Adaptation of *E. coli* BL21(DE3)/BGL and SE1/BGL Cells to Defined Media

For *E. coli* first cultivated in LB to grow vigorously in defined media, a medium adaptation step is required (251). As such, both SE1/BGL and BL21(DE3)/BGL strains were plated onto M9-agar (Table 8), by scraping the corresponding frozen LB glycerol stock with an inoculating loop and streaking an M9-agar plate. Next, the M9 plates were incubated at 37 °C for 48 h and then a single colony was picked to inoculate 10 mL of HDF medium (Table 9) in a 50 mL centrifuge tube, which was incubated in a shaker overnight at 37 °C and 250 rpm. In the next morning, the OD_{600} was measured and a certain volume of culture broth was transferred to a new 250 mL flask with 50 mL of HDF medium so that the initial OD_{600} would be 0.1. The culture was then grown at 37 °C and 250 rpm until mid-exponential phase ($OD_{600} \sim 1.5 – 2.0$). After that, the culture broth was centrifuged at 4

°C, 1,600 *g*, for 10 minutes, and resuspended in a volume of HDF with 15% glycerol (v/v), so that the theoretical OD_{600} would be 4.0. At last, the cells were dispensed in microtubes, in 330 µL aliquots, and stored in the -80 °C freezer. Again, when dealing with the BL21(DE3)/BGL strain, ampicillin 50 mg/L was added to the medium at every growth step (including plating).

Table 9 – Composition of HDF medium with 10 g/L of glucose.

Component	**Concentration (g/L)**	**Sterilization Method**	**Stock Solution**
KH_2PO_4	13.3	Autoclaved	20×
$(NH4)_2HPO4$	4.0		
Citric Acid	1.7		
Fe (III) Citrate	0.1		
NH_3 (aqueous solution, 28-30% w/w)	to pH = 7.0		
Glucose	10.0	Filter-sterilized	500 g/L
$Zn(CH3COO)_2{\cdot}2H_2O$	0.0338	Filter-sterilized	200×
$MnCl_2{\cdot}4H_2O$	0.0150		
EDTA Disodium	0.0141		
H_3BO_3	0.0030		
$CoCl_2{\cdot}6H_2O$	0.0025		
$Na_2MoO_4{\cdot}2H_2O$	0.0021		
$CuCl_2{\cdot}2H_2O$	0.0015		
$MgSO_4{\cdot}7H_2O$	1.2	Autoclaved	100×
Thiamine·HCl	0.0045	Filter-sterilized	2000×
Ultrapure Water	q. s.	Autoclaved	-

Note: ampicillin sodium salt (50 mg/L) was added as needed. Source: ref. (253).

3.1.4 Assays in HDF Medium

3.1.4.1 Preculture in HDF Medium

A tube from the HDF glycerol stock was thawed on ice and a volume of 50 μL was added to a 250 mL Erlenmeyer flask containing 50 mL of HDF medium; the preculture was then incubated in an orbital shaker for 16h at 37°C, 250 rpm (with 50 mg/L of ampicillin in the case of BL21(DE3)/BGL).

3.1.4.2 Growth Kinetics in HDF Medium

After growing a preculture in HDF medium for 16h (Section 3.1.4.1), the OD_{600} was measured and a certain volume added to fresh HDF medium, so as to produce a solution of 100 mL with an initial OD_{600} = 0.1, using a 500 mL Erlenmeyer flask. The flask was then incubated in an orbital shaker at 37 °C, 250 rpm, and cell growth was followed hourly through the values of OD_{600}, until the stationary phase was reached. Once again, when dealing with the BL21(DE3)/BGL strain, ampicillin 50 mg/L was added to the medium at every growth step.

3.1.4.3 Expression Assay in HDF Medium

Enzyme expression was quantitatively assessed in HDF medium, at 37 °C and 250 rpm. After growing a preculture in HDF for 16h (Section 3.1.4.1), the OD_{600} was measured and a certain volume added to fresh HDF medium, so as to produce a solution of 100 mL with an OD_{600} = 0.1, using a 500 mL Erlenmeyer flask. Cell growth was then followed hourly through the values of OD_{600}; after 5h, when the bacterial culture attained mid-exponential

phase (OD_{600} ~ 1.5 – 2.0), IPTG was added to a final concentration of 1 mM. After that, the OD_{600} continued to be measured hourly until t = 9 h, when the culture was interrupted and a 10 mL sample was taken. This sample was centrifuged at 12,000 *g* for 5 min, at 4 °C; the supernatant was decanted and discarded, and the pellets, stored at -20 °C. As always, ampicillin was added to the level of 50 mg/L at every growth step of BL21(DE3)/BGL.

3.1.4.4 Cell Lysis and β-glucosidase Extraction

Cell disruption was performed following the protocol presented in Section 3.1.3.16.

3.1.4.5 β-glucosidase Activity Assay

The evaluation of BGL activity was carried out observing the protocol described in Section 3.1.3.17, except that the cell extract was first diluted in lysis buffer 5- to 10-fold, so that the final absorbance fell within the range of 0.2 – 0.8.

3.1.4.6 Determination of the Extinction Coefficient of *p*-nitrophenol

To determine the molar extinction coefficient of *p*-nitrophenol (pNP), which is the product of BGL activity on pNPG, a standard curve of absorbance at λ = 405 nm against pNP molar concentration was created. For this purpose, pNP was first dissolved in 400 µL of 50 mM citrate buffer at pH 6.0, and then 400 µL of Na_2CO_3 1 M were added to the solution. The range of pNP concentration values extended from 0.03 mM to 1.50 mM. The resulting standard curve is given in Figure 26, Appendix A.

3.1.4.7 Bradford Assay

The Braford assay was used to determine the total amount of protein in samples of crude enzyme extracts. In particular, the microassay version of the method (254) was used, with hen egg white lysozyme as the protein standard, ranging from 3 to 18 μg; and a commercial Bradford reagent (Sigma-Aldrich). The resulting standard curve is given in Figure 25, Appendix A. Moreover, when analyzing data from expression assays, the mass of protein calculated for each sample of cell extract was corrected to account for the amount of lysozyme extrinsically added to the lysis buffer; that is, the amount of lysozyme added to the lysis buffer was subtracted from the amount of protein predicted by the standard curve.

3.1.4.8 Denaturing Polyacrylamide Gel Electrophoresis (SDS-PAGE)

To further verify the expression of the BGL enzyme in SE1 and BL21(DE3), cell extract samples were submitted to denaturing polyacrylamide gel electrophoresis (SDS-PAGE), using a Mini-PROTEAN® II electrophoresis module (Bio-Rad). The manufacturer's protocol for casting a 12% separating gel (255) was observed, except that a 40% acrylamide/bis-acrylamide 37.5:1 solution was used instead of a 30% acrylamide/bis-acrylamide 37.5:1 solution (the volume of acrylamide/bis-acrylamide solution was adjusted accordingly). Protein samples were diluted in 4× sample buffer (Table 10), in a screw-top microfuge tube, on ice; thereafter, the tube was incubated at 100°C in a dry bath for 5 min, and placed back on ice. A sample volume of 20 μL, corresponding to 26-28 μg of protein in the case of cell extracts, was loaded into each well. In addition, a volume of 10 μL of Pierce™ Prestained Protein MW Marker (Thermo Fisher Scientific) was introduced into the first well, serving as the MW standard. Electrophoresis was conducted under 180 V, for approximately 75 min. Afterwards, the gel was stained by submersion in a Coomassie blue solution (Table 11), for 1 h, and kept submerged in preservation solution (Table 11) overnight. Next, the gel was destained by submersion in a destaining solution (Table 11) for 2-3 h, exchanging the solution every hour (with fresh destaining

solution). Finally, the gel was placed in preservation solution (Table 11) until it could be photographed.

Table 10 Composition of 4× sample buffer for SDS-PAGE.

Component	Concentration
Tris-HCl pH 6.8	200 mM
glycerol	40% (v/v)
SDS	80 g/L
2-mercaptoethanol	10% (v/v)
Bromophenol blue	2 g/L
Ultrapure water	q. s.

Source: slightly modified from (256).

Table 11 – Composition of staining, destaining and preservation solutions for polyacrylamide gels.

Component	Staining	Destaining	Preservation
Acetic Acid	10% (v/v)	10% (v/v)	7% (v/v)
Methanol	40% (v/v)	40% (v/v)	-
Coomassie Brilliant Blue R-250	1 g/L	-	-
Ultrapure water	q. s.	q. s.	q. s.

Sources: (257) for staining and destaining solutions, (258) for preservation solution.

3.1.4.9 Cell Recycle Assay

For both BL21(DE3)/BGL and SE1/BGL strains, a preculture was grown in HDF medium for 16h following the protocol described in Section 3.1.4.1, and then a certain volume of the broth was used to seed 100 mL of fresh HDF medium in a 500-mL flask, so that the initial OD_{600} would be 0.1. The flasks were then incubated in the shaker at 30 °C and 250

rpm. After 12h, the OD_{600} was measured, a sample (40 μL) was taken for analysis of plasmid stability, and another sample (1.5 mL) was taken to seed 100 mL of fresh HDF medium (in a clean 500-mL flask), throwing out the remaining broth. This procedure was repeated every 12h, for 84h. Ampicillin was added to the level of 50 mg/L at every growth step of BL21(DE3)/BGL.

3.1.4.10 Analysis of Plasmid Stability

A volume of 40 μL of culture broth was serially diluted by a factor of 10^7 (5 dilutions of 25×) in Phosphate Buffered Saline (PBS; see Table 6) and a volume of 100 μL of the final (10^{-7}) dilution was spread on an LB agar plate without antibiotics. The plates were then incubated at 37 °C for 16-24h; afterwards, they were replica-plated (249) onto LB agar plates containing ampicillin (100 mg/L), and those were incubated at 30 °C, for 8-12h. Plasmid stability was defined as the number of colonies grown on the ampicillin plate divided by the number of colonies grown on the plate without antibiotics, always excluding those colonies up to 0.5 cm from the border of the plate (which are not captured by the replica-plating technique).

3.1.4.11 Growth Kinetics of *E. coli* SE1/BGL in a Bioreactor

A growth kinetics assay was performed in a 2-L Labfors-5 bioreactor (Infors HT), using the SE1/BGL strain and 1.2 L of HDF medium. A preculture of SE1/BGL was carried out by thawing an HDF/glycerol stock on ice and adding a volume of 100 μL to a 500 mL Erlenmeyer flask containing 100 mL of HDF medium, which was then incubated in an orbital shaker for 16h at 37°C, 250 rpm. The OD_{600} was then measured and a preculture volume was introduced into the bioreactor, so that a starting OD_{600} of 0.1 was obtained. Growth in the bioreactor was conducted in batch mode; the concentration of dissolved oxygen in the medium was set to 30% of saturation, achieved by feeding 0.6 L/min of compressed air to the vessel and by mixing with a cascade-controlled stirrer (rotation

range of 200-1000 rpm). The pH was set to 7.0 and controlled with a 200-mL ammonia solution, which was prepared by diluting one part of aqueous ammonia (NH_3 28-30%) in three parts of ultrapure water (i.e. 25% v/v). The culture broth was sampled hourly with a syringe, first drawing 5-10 mL to be discarded (due to tubing dead volume), and then taking extra 3-5 mL for OD_{600} measurement and quantitation of glucose. For quantitation of glucose, the sample was centrifuged at 12,000 *g* and 4°C for 5 min, and the supernatant, recovered by decantation; after that, the sample was stored at -20 °C until it could be analyzed. Finally, the concentration of glucose was determined by means of a glucose oxidase – peroxidase colorimetric kit (kit Glicose – n° 434) from Gold Analisa Diagnóstica, following the manufacturer's endpoint method (259).

3.1.4.12 Determination of Dry Cell Weight

Standard curves of dry cell weight against OD_{600} were created for both BL21(DE3)/BGL and SE1/BGL strains. For that purpose, 100 mL of preculture in HDF medium were prepared as per Section 3.1.4.1. Next, a 2-L Erlenmeyer flask containing 400 mL of HDF medium (with 50 mg/L of ampicillin in the case of BL21(DE3)/BGL) were inoculated with a volume of preculture such that the initial OD_{600} would be equal to 0.2, and incubated in an orbital shaker at 37 °C and 250 rpm. From t = 3 h to t = 9h, a sample volume was taken hourly for determination of OD_{600} and dry cell weight. The sample volumes were individually calculated for each point so as to contain approximately 30 mg of dry cell weight, based on past OD_{600} values and a ratio of *E. coli* dry cell weight to OD_{600} equal to 0.39 g/L found in the literature (260). With the aid of a vacuum filtration apparatus (Lab System 2 from Millipore), each sample volume was then filtered through a 0.22-µm membrane filter (type GSWP, also from Millipore), previously dried for 20 – 30 h in a laboratory oven at 95 - 105 °C, cooled for 10 – 15 min in a desiccator and weighed on an analytical balance. Then, the filter cake was washed with ultrapure water (using a volume approximately equal to the sample volume), dried in a regular microwave oven at 200 W for 15 min, cooled for 10 - 15 min in a desiccator and finally weighed on an analytical balance. The concentration of cells in terms of dry cell weight was then obtained by dividing the dry cell weight (equal to the difference between the membrane filter weight

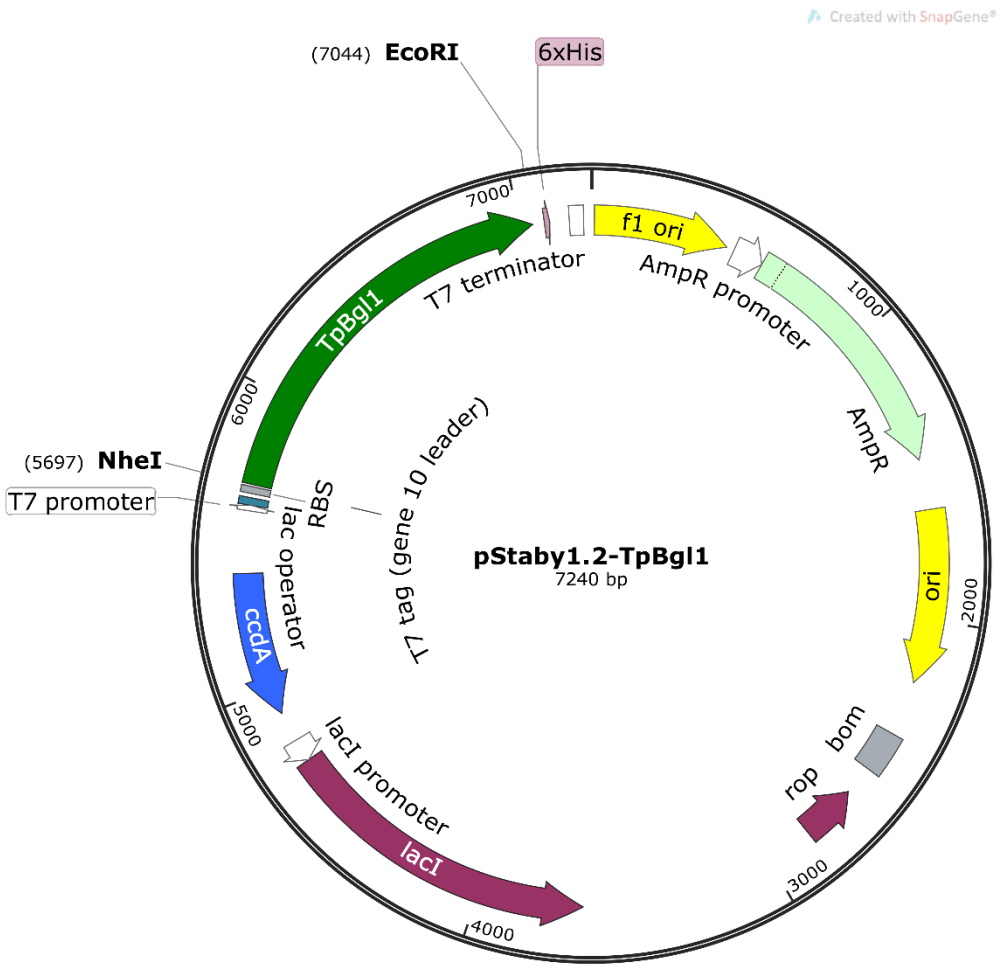
before and after filtration) by the corresponding sample volume. Lastly, each value of cell concentration in terms of dry cell weight was associated with its respective value of OD_{600} to develop a standard curve. This protocol is a slightly modified version of that from Olsson and Nielsen (261). The standard curves for both strains are given in Figure 28, Appendix A.

3.2 Results and Discussion

3.2.1 Cloning of the *TpBgl1* gene

The subcloning of the *TpBgl1* gene into the pStaby1.2 plasmid through the method of restriction digestion aimed at producing a plasmid carrying both the *TpBgl1* (BGL) gene and the *ccdA* (antitoxin) gene; the plasmid map of the expected ligation product is shown in Figure 6.

Figure 6 – Map of the plasmid pStaby1.2-TpBgl1.



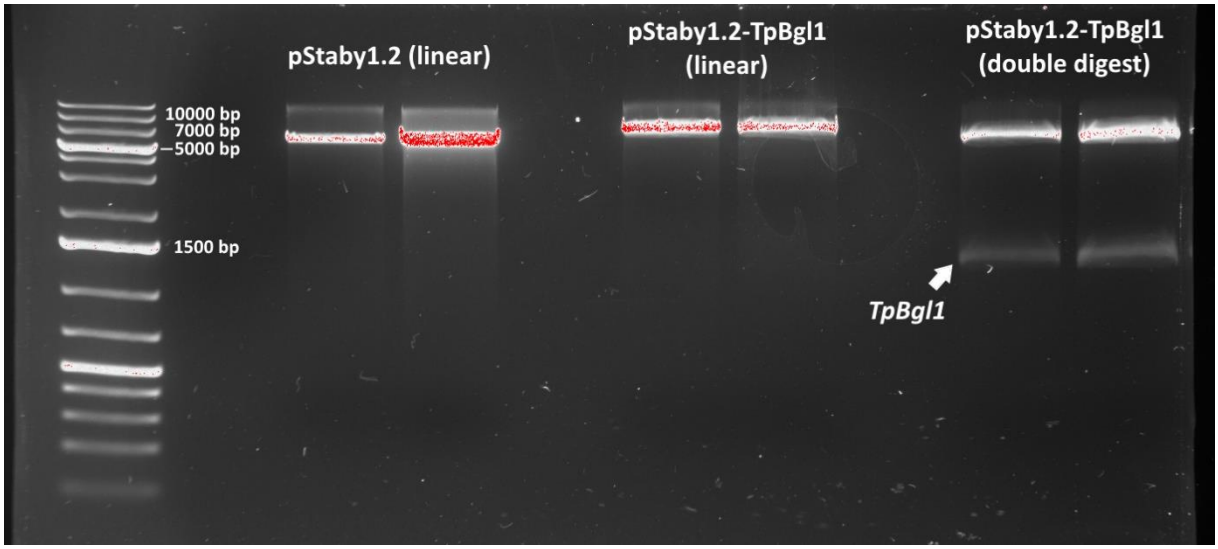
Caption: Map of the plasmid obtained in this work by cloning the *TpBgl1* gene into the pStaby1.2 plasmid. Note the presence of the *ccdA* gene, which codes for an antitoxin that counteracts the ccdB toxin. Source: this work.

The experimental procedure was successful, as attested by the agarose gel generated by the restriction digestion of the pStaby1.2 plasmid and of the ligation product, presented in Figure 7. The first two lanes to the right of the DNA ladder show duplicates of the pStaby1.2 plasmid digested by NheI only; the next two lanes show the ligation product digested by NheI only; and the last two lanes show the ligation product digested by both NheI and EcoRI. The theoretical sizes of the pStaby plasmid, the pStaby1.2-*TpBgl1* plasmid and the *TpBgl1* gene are respectively 5932, 7240 and 1350 bp, which are in good agreement with the observed bands.

The pStaby1.2-*TpBgl1* plasmid was subsequently transformed into *E. coli* BL21(DE3) by means of a standard protocol (see Section 3.1.3.12), and also into *E. coli* SE1, through a

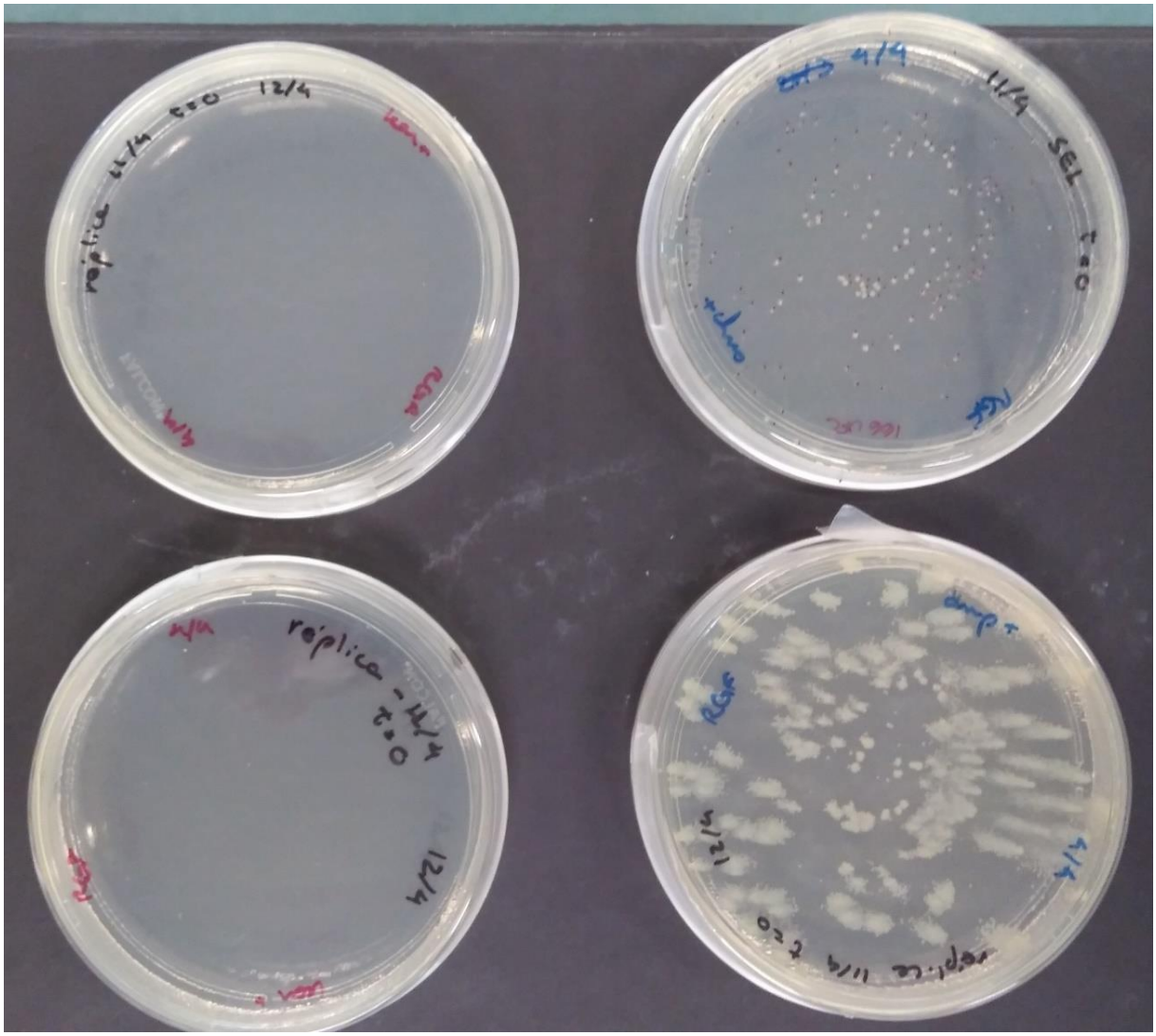
more complex procedure (see Section 3.1.3.14). This elaborate protocol was developed because we did not possess the competent SE1 cells devoid of plasmid provided by Delphi Genetics. These are special competent cells because the SE1 strain, by its very nature, should not survive without a plasmid containing the *ccdA* gene. Nevertheless, we did have SE1 + pET28a-*cel5A*-*ccdA* cells, that is, SE1 cells carrying a pET28a plasmid which contained the *cel5A* and *ccdA* genes, generated by the group of Prof. Sindelia Freitas at CTBE. This plasmid, as any pET28a-based plasmid, also carried a resistance gene against the antibiotic kanamycin (*kanR*). Besides, we had available the pStaby1.2-*TpBgl1* plasmid, which also carried the *ccdA* gene, as well as the ampicillin resistance gene *ampR*. The difference between the antibiotic resistance genes present in those two plasmids ultimately allowed us to devise a protocol to transform the SE1 + pET28a-*cel5A*-*ccdA* strain with the pStaby1.2-*TpBgl1* and, subsequently to make the cells lose the pET28a-*cel5A*-*ccdA* plasmid.

Figure 7 – Restriction digestions of plasmids pStaby1.2 and pStaby1.2-TpBgl1.



Caption: from the left to the right: 1st lane: MW marker; 2nd and 3rd lanes: digestion of the pStaby1.2 vector with a single restriction enzyme (duplicates); 4th and 5th lanes: digestion of the pStaby1.2-TpBgl1 product with a single restriction enzyme (duplicates); 6th and 7th lanes: sequential digestion of the pStaby1.2-TpBgl1 product with two restriction enzymes (duplicates). The MW of the pStaby1.2 plasmid is 5932 bp and that of the TpBgl1 gene is 1350 bp. Source: this work.

Figure 8 – LB-agar plates with colonies of SE1/BGL replicated on kanamycin and ampicillin containing plates.



Caption: the plate in the top right-hand corner is the "original" one, containing ampicillin; the plate right below it is a positive control replica, also with ampicillin (blurred due to overgrowth and, possibly, excessive moisture); the two plates on the left side are replica-plates containing kanamycin. Note that the single dot in the plate in the bottom left-hand corner is not a colony, but a small bubble in the solid medium. Source: this work.

Figure 8 shows the kanamycin and ampicillin plates produced by that protocol; the presence of SE1 colonies on the ampicillin plates, but not on the kanamycin plates, indicates that the SE1 cells acquired the pStaby1.2-*TpBgl1* plasmid and lost the pET28a plasmid, as predicted. Plasmid loss over time is in fact expected in the absence of a selective pressure and, given that the second plasmid introduced into *E. coli* SE1 (pStaby1.2-*TpBgl1*) responds to both the pressure of the chromosomal *ccdB* toxin and that of ampicillin, whereas the first plasmid (pET28a) only responds to the pressure of

the *ccdB* toxin, it is natural that the SE1 cells eventually lose the first plasmid. This is particularly inevitable in the present case because the two plasmids share replication and maintenance sequences, which makes them highly incompatible (262).

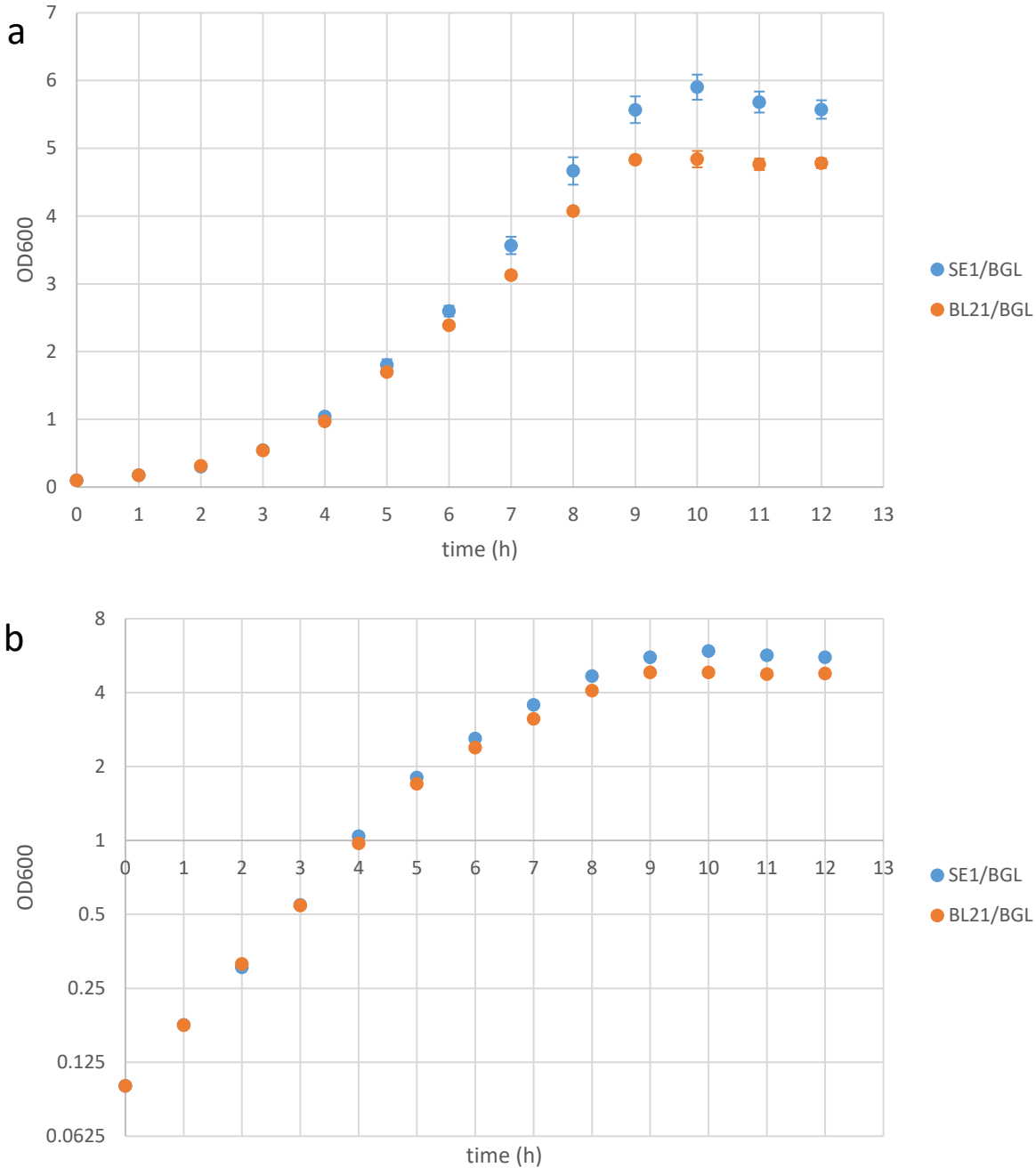
In order to further prove that the pStaby1.2-*TpBgl1* plasmid was present in the transformed SE1 and BL21 strains, a qualitative BGL activity assay was carried out in LB medium, using pNPG as the substrate. As a result, the color of the pNPG solution changed markedly in the presence of cell extracts from SE1 and BL21(DE3), going from transparent to bright yellow, and the OD measured at 405 nm was > 1 for both strains, thus demonstrating that both exhibited BGL activity and carried the pStaby1.2-*TpBgl1* plasmid (even more so considering that this assay was performed at 90 °C, way higher than the optimal temperature range of *E. coli*, which is a mesophilic microbe).

3.2.2 Growth Kinetics in HDF Medium Using Shake-Flasks

The growth kinetics of the strains BL21/BGL and SE1/BGL in HDF medium at 37 °C is shown in Figure 9a. For both strains, no lag phase is noticeable; the curves are essentially identical until t = 5 h, after which the BL21/BGL strain grows at a lower rate and reaches a lower maximum OD_{600} as well: 4.84 ± 0.12 against 5.90 ± 0.19 for SE1/BGL, both attained at t = 10 h. By examining the curves in a semi logarithmic chart (Figure 9b), we observe two distinct phases of exponential growth: the first from 0 to 5 h, and the second from 5 h to 8 h, again for both strains. This diauxic behavior is typical of batch cultures of B-derived strains growing on media in which glucose is the limiting substrate; during the first phase (I), cells readily consume glucose and excrete acetate, a metabolic byproduct; when glucose is depleted, cells start to consume the acetate previously excreted, which characterizes the second growth phase (II) (263). The specific growth rates for each phase were calculated as follows: μ_I = 0.58 h^{-1}, μ_{II} = 0.32 h^{-1} for SE1/BGL; and μ_I = 0.57 h^{-1}, μ_{II} = 0.29 h^{-1} for BL21/BGL. These results largely agree with those of Katayama (264), who employed the same strains, medium and growth conditions, but different plasmids and antibiotics (a pET28a containing the *ccdA* gene and a 1432-bp *cel5A* gene from *B.*

subtilis; and kanamycin 30 mg/L); the maximum values of OD_{600} obtained by that author were 6.2 ± 0.1 for SE1+pET28a-*cel5A*-*ccdA* and 4.2 ± 0.1 for BL21+pET28a-*cel5A*-*ccdA*.

Figure 9 – Growth kinetics of *E. coli* BL21/BGL and SE1/BGL in shake-flasks and HDF medium.



Caption: Growth kinetics, in terms of optical density at 600 nm (OD_{600}), of *E. coli* strains BL21/BGL and SE1/BGL in shake flasks, using HDF medium with 10 g/L of glucose, at 37 °C and 250 rpm. (a) Linear plot (b) Semi-logarithmic plot. Source: this work.

The superior performance of the SE1/BGL strain over the BL21/BGL strain is somewhat surprising, given that these strains are nearly identical. In fact, according to the literature (265), the genotype of the BL21(DE3) strain is:

> Strain B F⁻ *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3[*lacI lacUV5*-T*7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)

whereas the genotype of the SE1 strain, according to the StabyExpress™ kit manual, is:

> Strain B F⁻ ***Cm^R*** *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3[*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S) ***ccdB$^+$***

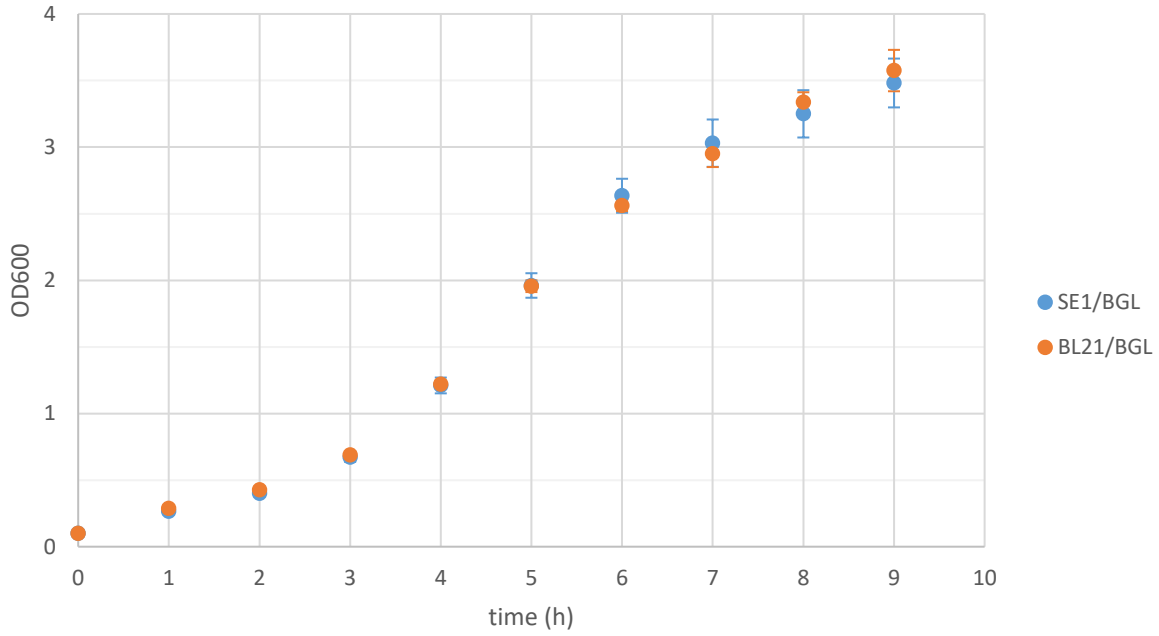
As such, the sole differences between the strains are the additional genes *ccdB* and *Cm^R* present in the SE1 strain; the *ccdB* gene gives rise to the ccdB toxin, while the *Cm^R* gene provides resistance to chloramphenicol (an antibiotic not used in this work). Although no direct comparison of the two strains grown on defined medium was found in the literature other than the aforementioned ref. (264), the few works that do compare SE1 and BL21 (189,266) have not found significant differences between the strains in terms of viable cell count or dry cell weight. Provided that there are no modifications in the SE1 strain undisclosed by its manufacturer, we can only assume that the addition of antibiotics (ampicillin) to the BL21/BGL cells is the reason why they attain a lower value of maximum OD_{600}, possibly by bringing on a heavier metabolic burden to the cell than the *ccdA*/*ccdB* TA system.

3.2.3 Enzyme Expression in HDF Medium

In order to express the recombinant BGL enzyme, cell growth was carried out under the same conditions as in the previous section; that is, HDF medium, 37°C and 250 rpm. However, BGL expression was induced by adding IPTG at t = 5 h, by the end of the "first" exponential phase. As a result, the growth curves of both strains become virtually identical throughout the bacterial culture (see Figure 10), in such a manner that the specific growth rate decreases after IPTG addition, in comparison with non-induced growth. In fact, after approximately one hour of induction, growth becomes roughly

linear, reaching an OD_{600} ~ 3.5 at t = 9h. This deleterious effect of induction on cell growth is well-known, and results from the large metabolic burden imposed by high levels of recombinant protein expression (31).

Figure 10 – Cell growth of *E. coli* BL21/BGL and SE1/BGL *versus* time, with induction of β-glucosidase expression.

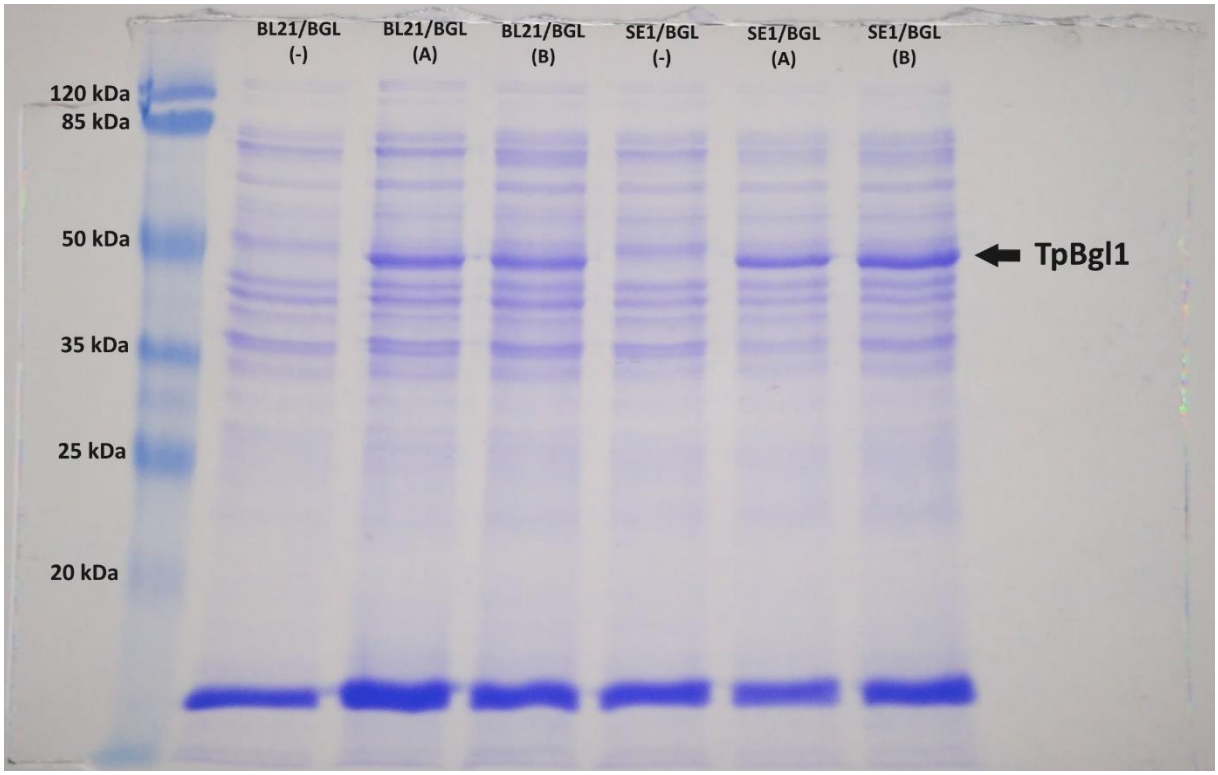


Caption: Plot of cell growth of *E. coli* strains BL21/BGL and SE1/BGL *versus* time, in which the expression of BGL was induced by addition of IPTG at t = 5 h. The assays were performed in shake-flasks, HDF medium with 10 g/L of glucose, at 37 °C and 250 rpm. In the case of BL21/BGL, the medium also contained ampicillin (50 mg/L). Source: this work.

The SDS-PAGE gel of the crude cell lysate, shown in Figure 11, again confirms the expression of the BGL enzyme, denoted by the strong protein band near 50 kDa (the theoretical MW of the BGL enzyme is 51.5 kDa) generated by both SE1/BGL and BL21/BGL samples after induction (t = 9 h), as opposed to the SE1/BGL and BL21/BGL samples before induction (t = 5 h). Furthermore, Table 1212 shows that high levels of BGL activity were detected in the cell lysates obtained from the SE1/BGL and BL21/BGL samples after induction, unmistakably demonstrating that the cloning and expression of the BGL enzyme was successful. Values of total soluble proteins (obtained by the Bradford method), OD_{600} and dry cell weight for the same samples are gathered in Table 1212 as well, revealing that the SE1/BGL system produced a higher activity than the BL21/BGL system, in spite of similar levels of dry cell weight and total soluble proteins. These results

concur with those of (189,264,266), who also observed higher expression levels for the SE1 toxin-antitoxin system than for the BL21 + antibiotic system.

Figure 11 – SDS-PAGE of *E. coli* SE1/BGL and BL21/BGL cell lysates.



Caption: from the left to the right: (1) MW standard (2) BL21/BGL before induction (3) BL21/BGL after induction (replicate A) (4) BL21/BGL after induction (replicate B) (5) SE1/BGL before induction (6) SE1/BGL after induction (replicate A) (7) SE1/BGL after induction (replicate B). The intense bands close to 50 kDa in the post-induction lanes correspond to the BGL enzyme (theoretical MW = 51.5 kDa). The strong bands close to the bottom of the gel correspond to the lysozyme added to the lysis buffer. Source: this work.

Nevertheless, the crude enzyme activity values per mg of total soluble protein obtained here are somewhat lower than that of Haq et al. (113), who attained a figure approximately 5× higher (17.7 kU/mg of total soluble protein) using a BL21(DE3) CodonPlus strain with a pET21a plasmid carrying the same BGL gene. However, those authors cultivated the bacteria in LB and performed induction for 72 h at 22 °C, so that the volumetric productivity of the present work is actually higher than that of Haq et al. (113) (approximately 0.39 U mg^{-1} h^{-1} versus 0.25 U mg^{-1} h^{-1}). It is also worth noting that those authors measured enzyme activity in the extracellular medium, instead of activity in the cell lysate, since they found that extracellular activity was higher. We, on the contrary, have found extracellular activity to be negligible (data not shown).

Table 12 – Concentration of cells, total proteins and enzyme activity from *E. coli* BL21/BGL and SE1/BGL cultures in HDF medium.

	OD_{600}	cell concentration	total soluble proteins		BGL activity	
		gDW/L	g/L of lysate	g/gDW	kU/mL of lysate	kU/mg of total soluble proteins
BL21/BGL	3.58	1.49	1.00	0.67	3.49	3.48
SE1/BGL	3.48	1.30	1.09	0.84	4.35	3.99

Caption: Concentration of cells in terms of optical density at 600 nm (OD_{600}) and dry cell weight, concentration of total soluble proteins in the cell lysate, and β-glucosidase activity in the cell lysate, for *E. coli* BL21/BGL and SE1/BGL. Both strains were cultivated in shake-flasks, using HDF medium with 10 g/L of glucose (and ampicillin 50 mg/L in the case of BL21/BGL), at 37 °C and 250 rpm; enzyme expression was induced by adding IPTG to 1 mM. The data in the table refer to the final culture samples (t = 9 h). Source: this work.

Cota et al. (114) have also produced the same BGL enzyme, but used a BL21(DE3) pRARE2 strain with a pET28a plasmid, cultivated the cells in LB and carried out induction for 16 h at 30°C, achieving up to 9.7 U/nmol of purified rProtein; given that the molecular weight of the BGL is equal to 51.5 kDa, that activity value translates to 188 U/mg of purified rProtein, which seems inconsistent with the values of crude enzyme activity obtained in the present study and in the study of Haq et al. (113). In effect, Haq et al. (113) also purified the enzyme and attained a specific activity of 30.4 kU/mg of pure rProtein, which is approximately 160× higher than the maximum specific activity found by Cota et al. (114). In any case, the specific activity values obtained in the present work are among the highest found in the literature for any BGL (cf. Table 1), even disregarding the fact that we have not purified the enzyme and therefore present enzyme activity in terms of total soluble proteins.

Another BGL found in the *Thermotoga petrophila* genome, belonging to glycosyl hydrolase family 3 (GH3) and referred to as TpBgl3, has also been produced in recombinant *E. coli* (114,190). Cota et al. (114) employed the same protocols used to produce TpBgl1, obtaining a maximum specific activity of 9.2 U/nmol of pure rProtein, similar to the value they found for TpBgl1 and equivalent to 113 U/mg. They also tested the activity of both enzymes on various phenol-derived carbohydrates, discovering that the GH1 BGL had a much broader substrate specificity than the GH3 enzyme (TpBgl1 activities against *p*-nitrophenyl β-D-fucopyranoside, *p*-nitrophenyl β-D-galactopyranoside and

p-nitrophenyl β-D-cellobioside were 94%, 36% and 35%, respectively, in relation to BGL activity on pNPG). In addition, they kinetically characterized both enzymes, concluding that the GH3 enzyme had a 49% higher catalytic efficiency than the GH1 enzyme, but much higher sensitivity to glucose inhibition; in fact, while TpBgl3 showed a K_i of 30 mM for glucose, TpBgl1 was stimulated by glucose concentrations of up to 1 M, displaying approximately 50% higher activity in the presence of 0.6 M of glucose. All in all, those authors conclude that the TpBgl1 enzyme has potential applications in the biofuel and food industries. Shi et al. (190) also produced the TpBgl3 enzyme, but in fed-batch culture, semi-defined medium and using *E. coli* BL21(DE3) with a pET20b plasmid, achieving up to 560 U/ml of cell lysate when induction conditions were optimized. Altogether, the results of the present work and the information on the literature lead us to believe that the recombinant production of the GH1 BGL from *T. petrophila* in *E. coli*, particularly if a plasmid addiction system such as the StabyExpress™ system is used, has a promising future in biomass hydrolysis.

3.2.4 Cell Recycle Assay

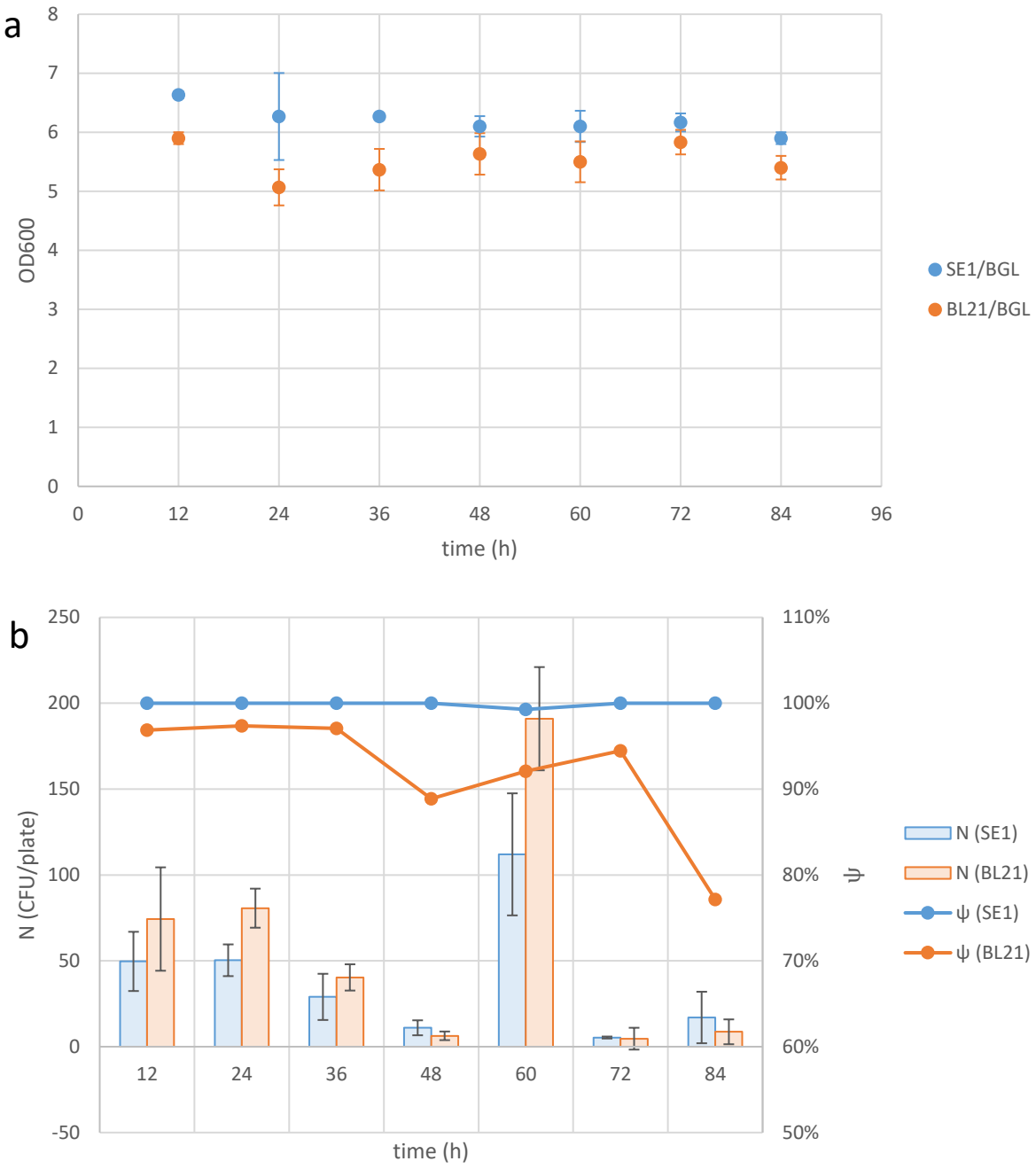
We have performed a "cell recycle" assay in which cells were grown at 30°C for 12h, and then a small fraction of the broth (1.5%) was used to seed fresh HDF medium; this procedure was repeated every 12h for 84h. The objective of this assay was to evaluate the possibility of replacing a seed train by a sole seed bioreactor, bearing in mind a large-scale industrial setting for plasmid-based rProtein production. In principle, this strategy would be regarded as industrially unfeasible because BL21(DE3) and similar *E. coli* expression strains tend to lose their plasmids over time, given that plasmid segregational stability, usually ensured by antibiotics (23,37,267), is not perfectly stringent under intensive culture conditions (36,266).

Figure 12a shows the value of OD_{600} at the end of each 12-h growth cycle. The OD_{600} becomes reasonably stable after 36h for both strains, maintaining a value of approximately 6 for SE1/BGL and slightly lower for BL21/BGL. Curiously, the OD_{600} value of the first point for SE1/BGL, and of all points for BL21/BGL, are higher than those

obtained in the growth kinetics assay at 37 °C. This may be related to the lower temperature employed (30 °C) or to further adaptation of the strains to the HDF medium.

Figure 12b shows the variation of the number of colonies per LB plate (excluding those near the border; see Section 3.1.4.10), represented by the letter *N*, as well as the variation of plasmid stability, represented by the Greek letter ψ, over time, for both strains. The colony count indicates the viability of cells, and shows wide variation between different time points, in contrast with the OD_{600} values seen above. In particular, only three data points – 12h, 24h and 60h – attained at least 50 CFUs per plate, which is considered the minimum necessary for analysis of plasmid stability (30,268), for statistical reasons. Two factors may have contributed to these results: the intrinsic uncertainty of the serial dilution technique, even more so considering the large dilution factor we used (25×); and the fact that the cells might have overgrown at each growth stage, particularly the SE1 strain, which showed a lower count of CFUs than the BL21 strain on all statistically significant points, despite exhibiting higher levels of OD_{600}. It is also worth noting that the OD_{600} values of the SE1 strain slightly decrease as the assay progresses, further supporting the hypothesis that the cells were being overgrown.

Figure 12 – Values of OD_{600} for *E. coli* BL21/BGL and SE1/BGL at the end of each growth step in the cell recycle assay.



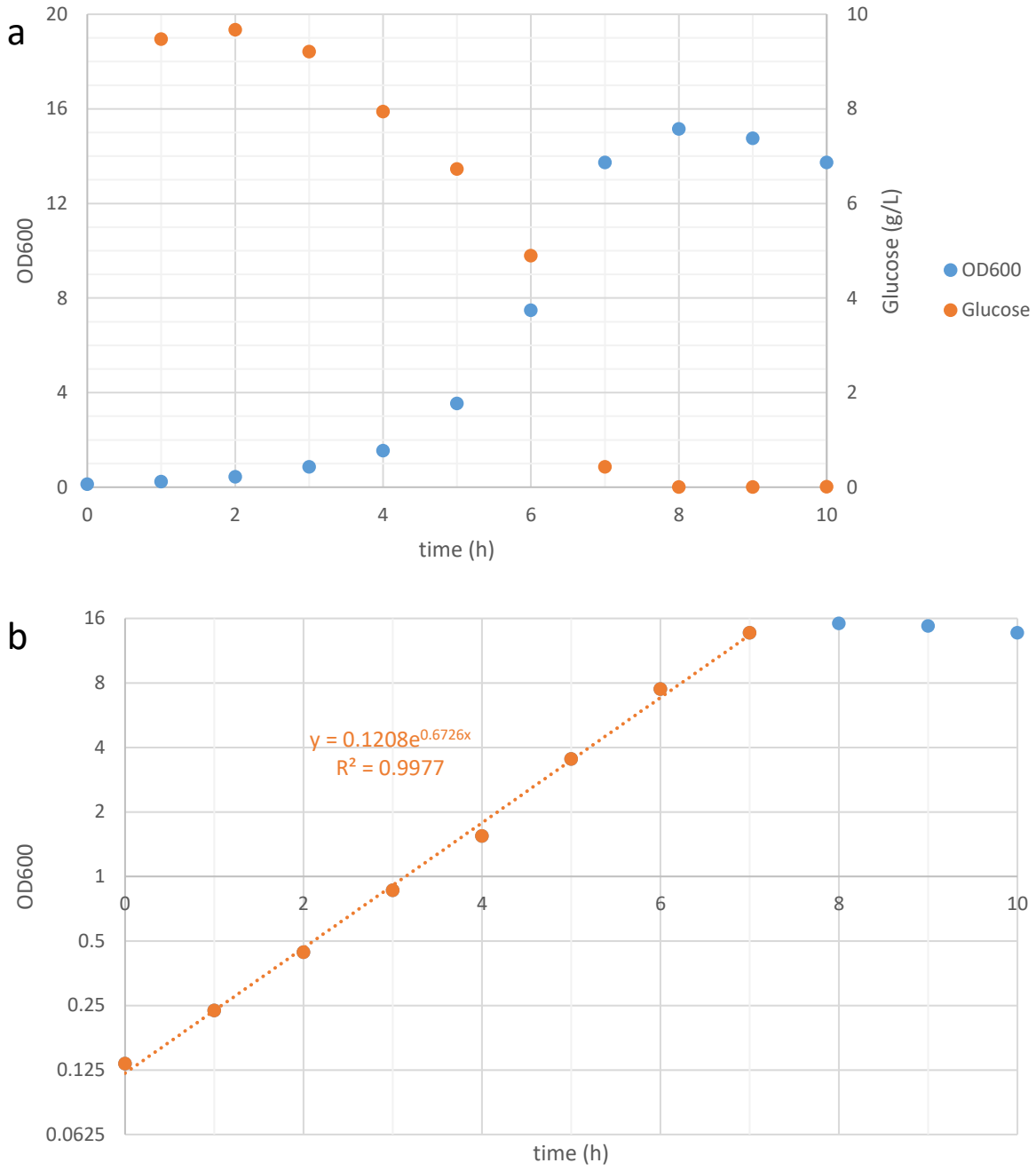
Caption: Values of OD_{600}, colony count (N) and plasmid stability (ψ) for *E. coli* BL21/BGL and SE1/BGL in the cell recycle assay. Bacterial cells were cultivated in shake-flasks, HDF medium, 30 °C and 250 rpm; every 12 h, 1.5% of the broth was used to seed fresh HDF medium in a new flask, while the remaining broth was discarded. (a) OD_{600} at the end of each growth step (b) plasmid stability and colony count at the end of each growth step. Plasmid stability was obtained by spreading a sample on an LB plate without antibiotics, and then replica-plating it on an LB-amp plate. The colony count per LB plate is given to show the statistical quality of each point. Source: this work.

As a consequence, the analysis of plasmid segregational stability, which depends on the number of CFUs per plate, is impaired by the low statistical quality of several data points. In any case, it seems clear that plasmid maintenance is virtually 100% in the SE1 strain, but not so in the BL21 strain. Nonetheless, plasmid stability is still very high for *E. coli* BL21 in the beginning of the assay – of 97% for 12 h, 24 h and 36 h – and decreases moderately to 92% at t = 60 h. These results corroborate other works in the literature. (189,264,266), which found that the SE1+pStaby system preserves plasmid stability at approximately 100%, as opposed to the BL21(DE3) + antibiotic system, which suffers from plasmid loss over time. Moreover, they validate the assumption that the SE1 strain would be ideal for the proposed inoculum production scheme. However, in that system plasmid structural stability and allele segregation would have to be regularly monitored, so as to ensure that high levels of cell growth and recombinant protein expression are maintained over long periods of time (204).

3.2.5 Growth Kinetics in HDF Medium Using a Bioreactor

Having observed the superior performance of the SE1/BGL system over the BL21(DE3)/BGL system in terms of cell growth, plasmid stability and recombinant expression, we carried out a simple growth assay using the SE1/BGL strain in a 2-L bioreactor, under controlled pH, temperature and aeration conditions, and with the same defined medium (HDF with 10 g/L of glucose) chosen for the shake-flask assays. The cell growth curve is presented in Figure 13a, in linear scale, and in Figure 13b, in logarithmic scale; Figure 13a also shows the variation of the glucose concentration over time.

Figure 13 – Kinetics of cell growth and glucose consumption of *E. coli* SE1/BGL in a 2-L bioreactor.



Caption: Cell growth and glucose consumption of *E. coli* SE1 over time, in a bioreactor culture. The culture was carried out in batch mode, in a 2-L bioreactor containing HDF medium with glucose 10 g/L, at 37 °C, pH 7.0 and oxygen set to 30% of saturation. (a) Linear plot of the optical density at 600 nm (OD_{600}) and glucose concentration *versus* time (b) Semi logarithmic plot of OD_{600} *versus* time, showing the exponential phase. Source: this work.

Firstly, we observe in Figure 13a that the SE1 strain attained a much higher value of OD_{600}, equal to 15.2, and did so in a shorter time (t = 8 h), than in the corresponding shake-flask

assays (OD_{600} = 6.1 at t = 9 h). Moreover, if we inspect the curve in logarithmic scale (Figure 13b), we identify no lag phase, but a single and continuous exponential phase extending from 0 to 7 h, from which a specific growth rate of μ = 0.67 h^{-1} can be calculated; this value is also measurably higher than that obtained in the shake-flask assay (μ = 0.59 h^{-1}). All in all, these results show that the SE1 strain displays a superior growth performance under controlled pH and aeration conditions, as expected; furthermore, they suggest that no significant concentration of acetate was produced, given that no diauxic behavior was detected this time. In addition, these values of maximum OD_{600} and specific growth rate are higher than those obtained by Katayama (264), who cultivated the SE1 strain with pET28a-*ccdA*-derived plasmids in a bioreactor under the same conditions and using the same medium as the present work, except for employing glycerol instead of glucose as the carbon source (maximum OD_{600} = 11.2–11.4, μ = 0.54–0.55 h^{-1}). To a certain extent, this was to be expected since *E. coli* grows faster on glucose than on glycerol (175). However, the value of specific growth rate derived in the present work is also higher than that obtained by Silva et al. (30), who cultivated the BL21(DE3) strain with pET28a-derived plasmids in a bioreactor under the same conditions and using the same medium, except for employing a higher concentration of glucose (μ = 0.56–0.59 h^{-1}). More generally, the maximum specific growth rate achieved here is in line with values obtained by various strains of *E. coli* growing at 37 °C on defined medium with glucose as the limiting substrate, which fall in the range of 0.5–1.0 h^{-1} (269).

Using the standard curves relating OD_{600} and dry cell weight for both *E. coli* strains (cf. Figure 28, Appendix A), we can also plot a curve of dry cell weight versus glucose concentration to calculate the yield of biomass on glucose. Such plot, limited to the exponential phase (cf. Figure 27, Appendix A), provides a yield value of $Y_{X/S} = 0.56$ g/g, which is somewhat higher than the average value found on the literature, of 0.50 g/g (176). However, significant errors are likely embedded in our figure: not only is the quantitation of dry cell weight intrinsically prone to error, but we have observed that the HDF medium gradually precipitates over time, which presumably interferes with measurements of both OD_{600} and dry cell weight.

4 TECHNO-ECONOMIC ANALYSIS OF INDUSTRIAL β-GLUCOSIDASE PRODUCTION

The following section aimed to model, simulate and economically evaluate the production of the BGL enzyme cloned into *E. coli*, to be used on-site as a supplementary enzyme in lignocellulose hydrolysis. Process parameters that are known to be significant, such as the process scale, biomass productivity and recombinant protein specific productivity, were evaluated. Furthermore, process characteristics that are rarely emphasized in the literature, such as the seed train expansion factor, bioreactor material, and cost contributions of the inducer and antibiotic compounds, were also examined. It should be noted that this chapter is extensively based on a peer-reviewed article published in 2018 by this author along with his thesis advisor and Dr. Sindelia Freitas, in the Biotechnology for Biofuels journal (270).

4.1 Methods

4.1.1 Design Basis

The production scale was based on the assumption that enzyme manufacturing would be integrated with a sugarcane-based 1G+2G ethanol plant (that is, on-site), using a 100 m^3 bioreactor. With respect to product specifications, it was assumed that the enzyme should be stabilized in a citrate buffer of pH 5.8 and concentrated to a titer of 15 g/L. The main parameters used for the design are summarized in Table 13.

The annual enzyme production rate should be sufficient to hydrolyze a significant fraction of the total amount of sugarcane bagasse generated by the sugarcane processing plant. To calculate this fraction, we assumed an average sugarcane plant in Brazil that processes 2 million tons of sugarcane/year. Besides, we considered a bagasse/sugarcane fraction of 26% (w/w) and a moisture content of 50% (271). We also assumed that a cellulase loading of 10 FPU/gDW of bagasse (19,158) was employed for hydrolysis, and that BGL

was supplemented to the level of 0.2 CBU/FPU[‡]. This level of BGL supplementation is based on a study of Pryor and Nahar (17), which established the minimal amount of BGL required for efficient cellulose hydrolysis. Finally, we took into account the specific activity of the BGL in question, equal to 2.3 CBU/mg of enzyme (113); this parameter is necessary to convert the amount of BGL from activity units (CBU) to mass units (kg).

Table 13 – Main parameters used in the design of the enzyme production process.

Parameter	**Assumption**
Final enzyme titer (after primary recovery and concentration)	15 g/L
Annual operating time of the rEnzyme production plant	7920 h (330 days)
Nominal volume of the main fermenter	100 m^3

Source: this work.

4.1.2 Modeling and Simulation Software

SuperPro Designer v9.5 (Intelligen, Inc.) was employed to model and simulate both the baseline BGL (rEnzyme) production process and the variations of this process with respect to technical and economic parameters. The program was also used to perform the economic assessment of the process.

4.1.3 Upstream Section

Seed trains with expansion factors of 10-, 20- and 100-fold (that is, 10%, 5% and 1% of inoculum volume, respectively) were evaluated. The expansion factor of 20 was used as a

[‡] CBU: BGL activity unit measured on cellobiose; FPU: Cellulase activity unit measured on filter-paper.

reference when evaluating other parameters, such as protein productivity. In all cases, the number of seed fermenters was determined by the initial effective (batch) volume of the main fermenter and the expansion factor chosen, considering that inoculum volumes of up to 20 L could be generated in a laboratory setting (not explicitly modeled). The medium used in the seed fermenters, as detailed in the next section, was the same in all cases.

An alternative strategy for producing the inoculum was also evaluated. In this scheme, only one seed fermenter is employed; after each seed fermentation, a certain fraction of the volume is retained, while the rest is transferred to the main fermenter. In the main fermenter, the fed-batch culture and induction are conducted normally, as described in the next section. At the same time, the seed fermenter containing the residual fraction of cell broth is replenished with fresh culture medium, and a new seed culture begins. In this inoculum design, the size of the seed fermenter was determined by the volume of batch culture produced, which was set to equal that in the conventional inoculum scheme, adjusted by the fraction of recycled broth volume. The precise fraction of recycled volume was determined by fermentation conditions in the seed fermenter and in the main fermenter, as explained later in Section 4.1.4.5. This inoculum strategy was termed "inoculum recycle".

4.1.4 Fermentation Section

4.1.4.1 Microorganism

E. coli BL21(DE3) harboring a pET28-a (+) plasmid (Novagen) carrying the *TpBgl1* gene under control of the *lac* operator and T7 promoter was used as the expression system. The *TpBgl1* gene codes for a thermostable, monomeric BGL, 52 kDa in size, that was originally found in the hyperthermophilic, gram-negative bacterium *Thermotoga petrophila* (113). This plasmid also contains the kanamycin resistance gene *kanR*.

4.1.4.2 Culture Media and Fermentation Conditions

The fermentation process was based on the fed-batch process proposed by Strittmatter et al. (272) and Horn et al. (191), which employs a defined medium containing glucose as the main carbon source and ammonia as the main nitrogen source. However, the replacement of glucose with glycerol was suggested by the original authors and was also evaluated. The entire process is carried out at 26°C in a pressurized (150 kPa) stainless steel vessel, as summarized in Table 14. Initially, the microorganism consumes the substrates present in the batch medium (described in Table 15), in which the carbon source (glucose or glycerol) is the only limiting substrate. When the carbon source concentration approaches a critical value (1.5 g/L), feeding solution 1 (FS1, also described in Table 15, together with feeding solutions FS2 and FS3) is added so as to maintain the carbon source concentration at a constant level. Therefore, the microbial culture process consists of a batch phase followed by a fed-batch phase. The control of glucose concentration, together with the use of a rather low growth temperature (26°C) and a low-acetate-producing *E. coli* strain, such as BL21(DE3) (273), prevents the excessive production of acetate, thus allowing the bacteria to grow steadily at a constant specific growth rate, approximately 0.23 h^{-1}, throughout the process.

Table 14 – Main fermentation parameters used in the baseline case.

Parameter	**Value**
Fermenter Maximum Working Volume	80%
Temperature	26 °C
Overpressure	150 kPa
Fermenter Material	Stainless Steel of grade 316 (SS316)
Specific growth rate (μ)	0.23 h^{-1}
Air Flow Rate	1 VVM$^{\gamma}$

$^{\gamma}$ At standard conditions for temperature and pressure.
Source: this work.

Table 15 – Composition of media for fed-batch culture of *E. coli*.

Component	Batch Medium (mg/L)	FS1 (mg/L)	FS2 (mg/L)	FS3 (mg/L)	Others
Glucose (Glycerol)	25×10^3 (27.8×10^3)	670×10^3 (744.4×10^3)			
$MgSO_4 \times 7H_2O$	1.5×10^3	19.8×10^3			
K_2HPO_4	16.6×10^3				
Citric acid	2.1×10^3				
$(NH_4)_2HPO_4$	4×10^3		227×10^3		
$(NH_4)H_2PO_4$			169.5×10^3		
Fe(III) citrate hydrate	75			5×10^3	
H_3BO_3	3.8			250	
$MnCl_2 \times 4H_2O$	18.8			125	
$EDTA \times 2H_2O$	10.5			700	
$CuCl_2 \times 2H_2O$	1.9			125	
$Na_2MO_4 \times 2H_2O$	3.1			213	
$CoCl_2 \times 6H_2O$	3.1			213	
$Zn(CH_3COO)_2 \times 2H_2O$	10			668	
Aqueous NH_3 (NH_4OH)					25% v/v[δ]
Kanamycin Sulfate	30 mg/L[ε]				
IPTG					1 mM[ε]

[δ] One volume of NH_3 28-30% w/w solution was diluted with three volumes of ultrapure water.
[ε] Concentration given with respect to the final volume of culture broth.
Caption: Composition of media for high cell density culture of *E. coli* with recombinant protein expression. The process starts as a batch culture, using the components listed in the Batch Medium column. When the carbon source approaches depletion, the addition of Feeding Solution 1 (FS1) begins, in order to keep the concentration of the carbon source constant. Feeding solutions 2 and 3 are introduced afterwards, when phosphorus and other elements become scarce. Ammonia is added as needed, to maintain the pH at 7; and IPTG is supplied to induce rProtein expression when the cell concentration attains 40 gDW/L. Source: ref. (191).

Feeding solution 2 (FS2) and feeding solution 3 (FS3) are added at constant rates towards the end of the fed-batch process to avoid any contingent nitrogen or trace metal

limitations. The proposed process also employs pH control, which is achieved by the addition of aqueous ammonia (25% v/v of aqueous NH_3 28%) to maintain a pH of 6.8. The amount of added base was calculated to precisely fulfill the nitrogen demand of the cells (which was not entirely satisfied by the batch medium and FS2).

It is worth noting that, in the simulation, the volume of FS1 used was that which provided the total amount of glucose required by the process, which was determined by the final biomass concentration (see Section 4.1.4.3) and global yield of biomass on glucose (see later on this section) assumed for the bacterial culture. The volumes of FS2 and FS3, on the other hand, were proportional to the volume of the batch medium used, following the ratios suggested by Horn et al. (191), i.e., 1 L of FS2 and 50 mL of FS3 for 8 L of batch medium. Moreover, in the simulator, the batch medium, FS1, FS2 and FS3 were all stored together in a blending tank, sterilized and introduced into the main fermenter as a single solution, for the sake of simplicity.

Cell growth was modeled with stoichiometric equations based on the atomic balance of C, H, O and N. It was assumed that the sole reactants in the growth equation were the main carbon source (glucose or glycerol), ammonia and oxygen, whereas the only products were the *E. coli* biomass, with an empirical formula of $CH_{1.8}O_{0.5}N_{0.2}$, according to the literature (210); carbon dioxide; and water. As such, the formation of organic acids was neglected, and BGL production was assumed to be intracellular.

The stoichiometric coefficient of the carbon source was set to 1 (on a molar basis), and the stoichiometric coefficient of the biomass was considered to be equal to the biomass yield from the carbon source. Values of 0.50 g/g and 0.45 g/g of biomass yield on the carbon source were assumed for the seed fermenters (batch) using glucose and glycerol, respectively. These values are close to those found by Korz et al. (274), who carried out a high cell density culture of *E. coli* on a defined medium (without rProtein expression). On the other hand, the yield values for the main fermenter were estimated by applying a 20% reduction to the batch yield values; as such, they were equal to 0.40 g/g for glucose and 0.36 g/g for glycerol. Based on our experience, the biomass yield can be reduced by as much as 50% in fed-batch mode compared to batch culture mode (30). Furthermore, these lower yield values are in reasonable agreement with those obtained by Wyre and Overton (34), who employed somewhat similar conditions (a fed-batch culture of *E. coli*

with IPTG induction). All these assumptions led to the stoichiometric equations presented in Table 16.

In the baseline scenario, kanamycin sulfate was also added to the main fermenter and the seed fermenters at a final concentration of 30 mg/L. In the main fermenter, IPTG was supplied to achieve a final concentration of 1 mM. Medium sterilization was performed in a continuous heat sterilizer. In contrast, thermo-sensitive compounds such as IPTG and kanamycin were assumed to be filter-sterilized in a laboratory (not explicitly modeled) and introduced directly into the fermenter. A compressor and an air filter were also included in the process model to generate a sterile air feed of 1 VVM. This flow rate was considered sufficient to maintain an oxygen pressure of 20%. Besides, the fermenter was stirred to ensure appropriate oxygen transfer and nutrient mixing, for which a power consumption of 3.0 kW/m^3 was assumed. Finally, an air filter was included downstream from the gas outlet of the fermenter to ensure the biosafety of the process. The fermentations in the seed fermenters were conducted in a similar fashion to the main fermentation batch phase, using the batch medium detailed in Table 15 as well as 30 mg/L of kanamycin and ammonium hydroxide for pH control and nitrogen supplementation; the fermentation process in each seed fermenter is described by the corresponding stoichiometric equations in Table 16.

Table 16 – Stoichiometric equations used to describe cell growth.

Substrate	**Culture**	**Equation**
Glucose	Seed Culture	$180.16\ C_6H_{12}O_6 + 69.10\ O_2 + 12.46\ NH_3 \rightarrow 90.08\ CH_{1.8}O_{0.5}N_{0.2} + 103.08\ CO_2 + 68.55\ H_2O$
	Main Culture	$180.16\ C_6H_{12}O_6 + 93.67\ O_2 + 9.97\ NH_3 \rightarrow 72.06\ CH_{1.8}O_{0.5}N_{0.2} + 135.27\ CO_2 + 76.46\ H_2O$
Glycerol	Seed Culture	$92.09\ C_3H_8O_3 + 55.45\ O_2 + 5.73\ NH_3 \rightarrow 41.44\ CH_{1.8}O_{0.5}N_{0.2} + 57.96\ CO_2 + 53.87\ H_2O$
	Main Culture	$92.09\ C_3H_8O_3 + 66.76\ O_2 + 4.59\ NH_3 \rightarrow 33.15\ CH_{1.8}O_{0.5}N_{0.2} + 72.78\ CO_2 + 57.51\ H_2O$

Caption: Stoichiometric equations used to describe cell growth using either glucose or glycerol as carbon source. The seed cultures are performed in batch mode and, as such, produce high yields of biomass on the carbon source; the main cultures, on the other hand, consist of fed-batch processes and thus generate lower yields. Note that these stoichiometric coefficients are mass-based. Source: this work.

Besides the substitution of glucose for glycerol, other modifications of the fermentation section were investigated (always assuming that everything else remained equal):

a) a tenfold reduction in the amount of IPTG added for induction;
b) a tenfold reduction in the cost of glucose cost;
c) dispensing with the use of antibiotics (kanamycin), assuming the use of a recombinant *E. coli* strain equipped with a TA system, such as SE1;
d) purchase of an extra main bioreactor, identical to the original one, to be employed in stagger mode.

More complex variations implemented on the fermentation section are described in the next sections.

4.1.4.3 Volumetric Productivity

Assuming that the specific growth rate remained constant and equal to 0.23 h^{-1} throughout the main culture and throughout the seed cultures, the duration of the main culture was calculated to be 19 h when the inoculum volume was 10%, 22 h when the volume was 5%, and 29 h when the volume was 1%. It should be noted that these calculations assumed that the seed fermenters employed the same batch medium previously described, that the seed fermenters had the biomass yields indicated earlier, and that the biomass attained at the end of the main culture was 100 gDW/L. However, given the uncertainty associated with microbial cultures and with the expression of recombinant proteins, compounded with the uncertainty related to scaling up such processes, different scenarios of final biomass concentration and soluble rEnzyme content were evaluated, as shown in Table 17. The highest value of the final biomass concentration, 120 g/L, was extracted from the graphical data of Horn et al. (191). Lower values, 80 g/L and 100 g/L, were also considered, since scaling up might diminish biomass yields, especially in the case of aerobic processes (191). For the rEnzyme content, values of 2%, 10% and 20% of the total protein, corresponding to 1%, 5% and 10% of the dry

cell weight, respectively, were evaluated. It should be stressed that except for the rEnzyme content, the overall cell composition was presumed to be the same in all cases, in accordance with the lysis equation presented later. The scenario with intermediate values of the final biomass and total rEnzyme content (Table 17), i.e., 100 g/L and 5%, respectively, was assumed to be the baseline scenario when evaluating the effects of different parameters.

Table 17 – Values of biomass and rEnzyme content considered for economic analysis.

Scenario	Biomass ($gDW\ L^{-1}$)	rEnzyme/ Total Protein	rEnzyme/ Biomass	rEnzyme Volumetric Productivity ($g\ L^{-1}\ h^{-1}$)
1	80	2%	1%	0.04
2	80	10%	5%	0.18
3	80	20%	10%	0.36
4	100	2%	1%	0.05
5	100	10%	5%	0.23
6	100	20%	10%	0.45
7	120	2%	1%	0.05
8	120	10%	5%	0.27
9	120	20%	10%	0.54

Caption: Various values of final biomass concentration and rEnzyme content that were considered for economic analysis. Total soluble proteins are assumed to constitute 50% of the cell dry weight. Furthermore, a seed train with a 20× expansion factor was assumed for all these scenarios, leading to a main culture of 22 h. Note that all percentages here are mass based. Source: this work.

4.1.4.4 Extracellular Enzyme Production

In a set of alternative scenarios, we considered that the enzyme was secreted to the broth, rather than stored inside the cell, as appears to be possible in the case of this particular BGL (113). The enzyme yield was assumed to be the same as that of intracellular production in the baseline scenario. As such, the equation of cell growth for extracellular

production was derived from the equation of the baseline (intracellular) case, simply replacing a fraction (5%) of the biomass with the enzyme:

$$180.16\ C_6H_{12}O_6 + 93.67\ O_2 + 9.97\ NH_3 \rightarrow 68.46\ CH_{1.8}O_{0.5}N_{0.2} + 135.27\ CO_2 + 76.46\ H_2O + 3.60\ \text{rEnzyme}$$

The consequences of extracellular enzyme production on the Downstream Section are discussed later.

4.1.4.5 Fermentation Modifications When Using Inoculum Recycle

The duration of the fermentation in the main fermenter is changed when employing the inoculum recycle strategy. Let index A refer to the seed fermenter and index B refer to the main fermenter; in both cases, the cell concentration X follows an exponential law:

$$X_A = X_{A0} \cdot e^{\mu t_A} \quad (1)$$

$$X_B = X_{B0} \cdot e^{\mu t_B} \quad (2)$$

Where X_{AO} and X_{BO} are the initial cell concentration in culture A and B, respectively; t_A and t_B are the duration of cell growth in culture A and B, respectively; and $\mu = 0.23\ \text{h}^{-1}$ is the specific growth rate in both cultures. If V_{AO} is the total seed culture volume and θ is the volume fraction transferred to the main fermenter, then $(1-\theta)V_{A0}$ is the volume that remains in the seed fermenter after one batch. The initial cell concentration X_{AO} for the next batch then becomes

$$X_{A0} = X_A \cdot (1-\theta)V_{A0}/V_{A0} = X_A \cdot (1-\theta) \quad (3)$$

Substituting Equation (3) in Equation (1), we obtain

$$X_A = X_A \cdot (1-\theta) \cdot e^{\mu t_A} \leftrightarrow 1 = (1-\theta) \cdot e^{\mu t_A} \quad (4)$$

On the other hand, if the initial volume in the main fermenter is V_{BO}, then its initial cell concentration is given by

$$X_{B0} = X_A \cdot \theta \cdot V_{A0}/V_{B0} \quad (5)$$

If we consider that the initial volume in the main fermenter is exactly equal to the volume transferred from the seed fermenter, then

$$V_{B0} = \theta \cdot V_{A0} \rightarrow X_{B0} = X_A \quad (6)$$

Substituting Equation (6) in Equation (2), we get

$$X_B = X_A \cdot e^{\mu t_B} \quad (7)$$

Where X_B if the final cell concentration in the main fermenter, defined as 100 g/L in the baseline scenario; X_A is the final cell concentration in the seed fermenter, equal to 12.5 g/L (determined by the amount of glucose in the medium, 25 g/L, and the yield of cell mass on glucose, 0.50 g/g). The variables t_B in eq. (7), and θ, t_A in eq. (4) remain unknown; however, if we consider that, ideally, the seed culture and the main culture must be synchronized so that the total process cycle time is minimized, then $t_A = t_B$. As a result, we can divide Equation (4) by Equation (7), to obtain:

$$1/100 = (1 - \theta)/12.5 \leftrightarrow \theta = 0.875 = 87.5\% \quad (8)$$

That is, the volume fraction transferred from the seed fermenter to the main fermenter must be equal to 87.5%. Finally, substituting this result into either Equation (4) or (7), the duration of the seed culture/main culture can be calculated: $t_A = t_B = 9.04$ h.

4.1.5 Downstream Section

As the fermentation process ends, we assumed that the fermentation broth is collected in a tank and then passed through a high-pressure homogenizer, where the bacterial cells are lysed to harvest the intracellular protein product. In the simulator, the homogenization process was modeled as a pseudo-reaction in which the biomass is converted into its main components. Based on the *E. coli* composition from Milo and Phillips (275), we developed the following (mass-based) lysis equation:

$$1\ CH_{1.8}O_{0.5}N_{0.2} \rightarrow 0.20\ \text{Cell Debris} + (1-\varphi)(0.50)\ \text{Contaminating Proteins} + \varphi(0.50)\ \text{rEnzyme} + 0.20\ \text{DNA} + 0.03\ \text{RNA} + 0.03\ \text{Glycogen} + 0.03\ \text{Metabolites} + 0.01\ \text{Inorganic Ions}$$

In this equation, φ is the rEnzyme content (rEnzyme/total protein) relative to the total soluble protein. The cell debris are then separated from the mixture by first using a disk-stack centrifuge, which removes 70% of the debris, and then with a dead-end filter, which removes the residual debris. Finally, the enzyme solution is concentrated to the desired titer (15 g of enzyme/L) and stabilized in a citrate buffer solution (citric acid + sodium citrate, pH 5.8) using a diafiltration system. In addition, some changes to this downstream configuration were investigated: (i) the substitution of the centrifuge for a microfilter; (ii) the substitution of diafiltration for simple ultrafiltration. An alternative scenario in which enzyme production was extracellular was also examined; in this case, the high-pressure homogenizer was omitted, so that the fermentation broth went straight to the centrifuge, where the extracellular medium (containing the enzyme) was separated from the whole cells. The remaining of the Downstream Section was left unchanged.

The main parameters of the downstream unit operations are given in Table 21, Appendix B. They were held constant across all scenarios, with the exception of the concentration of solids in the centrifugation sludge, which was varied from 200 g/L to 800 g/L, with 200 g/L increments, in both intracellular and extracellular production scenarios.

4.1.6 Changes to Process Scale, Operating Time, and Equipment Material

As indicated earlier, variations of many parameters and unit operations have been evaluated within each process section, always using the baseline scenario as the starting point. In addition, some changes to the baseline scenario that concern all process sections have been assessed:

a) Process Scale: the nominal volume of the main fermenter was taken as the measure of process scale, and it was varied from 25 m^3 to 150 m^3, with 25 m^3 increments.

Moreover, sensitivity analyses concerning other process parameters were done on three process scales: 50 m^3, 100 m^3 and 150 m^3;

b) Process operating time: the reduction of the process operating time from 330 days to 214 days per year was considered;
c) Material of bioreactors and storage tanks: replacing the material of the fermenters and storage tanks (stainless steel of grade 316) with either stainless steel of grade 304 or carbon-steel was also examined.

4.1.7 High-titer Scenarios

Some process modifications applied to the baseline case were also applied to the scenario of highest enzyme titer indicated in Table 17 of Section 4.1.4.3 (final cell concentration of 120 g/L, and rProtein/dry cell weight ratio of 10%), using the same basic unit operations and process parameters, except for the removal of the enzyme concentration and stabilization steps (dead-end filtration and diafiltration). These modifications were carried out individually and also combined, as shown in the first seven rows of Table 18. Furthermore, analogous high-titer scenarios with extracellular enzyme production (instead of intracellular production) were also simulated, individually and combined, as indicated in the last seven rows of Table 18.

Table 18 – Simulated scenarios of high-titer enzyme production.

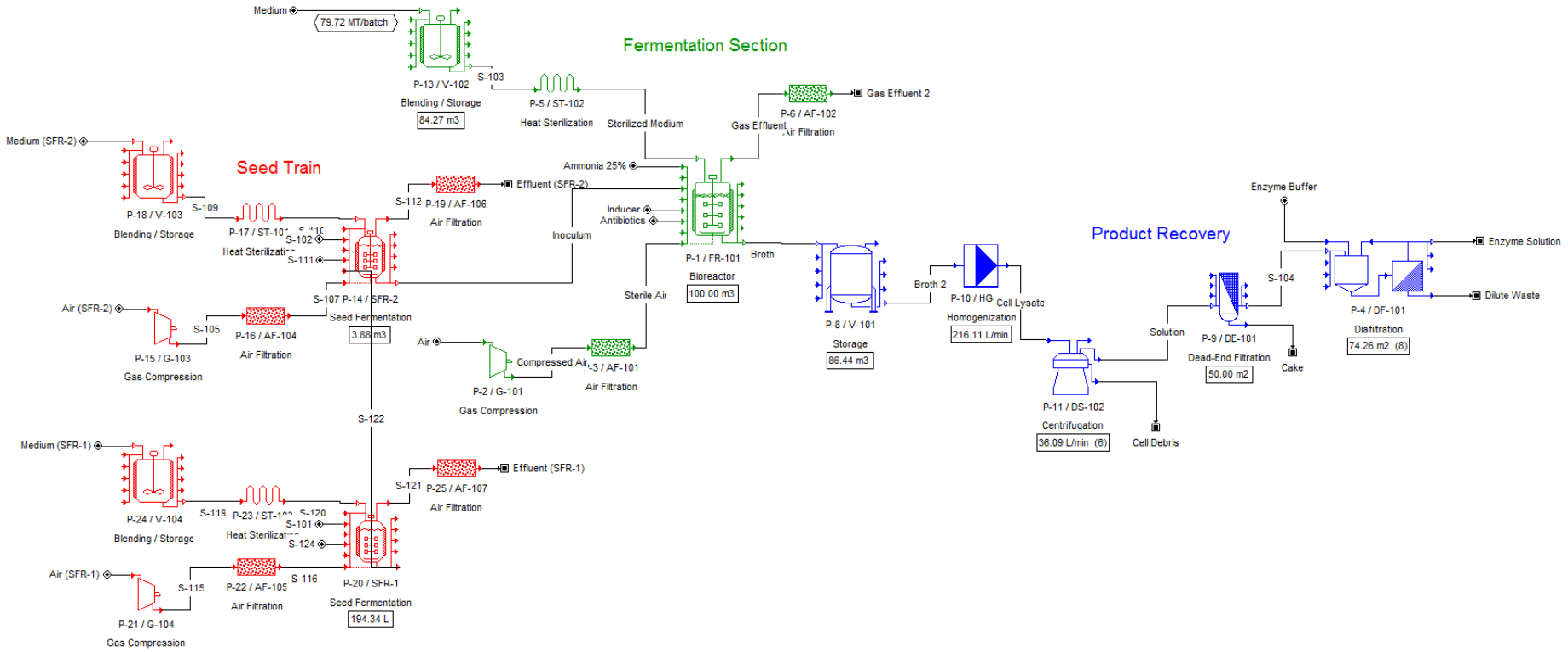
Production Type	Fermenter Material	IPTG Amount	Glucose Cost	Extra Fermenter	Inoculum Recycle	Concentration step
Intracellular	SS316	1X	1X	no	no	no
Intracellular	CS	1X	1X	no	no	no
Intracellular	SS316	0.1X	1X	no	no	no
Intracellular	SS316	1X	0.1X	no	no	no
Intracellular	SS316	1X	1X	yes	no	no
Intracellular	SS316	1X	1X	no	yes	no
Intracellular	CS	0.1X	0.1X	no	yes	no
Extracellular	SS316	1X	1X	no	no	no
Extracellular	CS	1X	1X	no	no	no
Extracellular	SS316	0.1X	1X	no	no	no
Extracellular	SS316	1X	0.1X	no	no	no
Extracellular	SS316	1X	1X	yes	no	no
Extracellular	SS316	1X	1X	no	yes	no
Extracellular	CS	0.1X	0.1X	no	yes	no

Note: all simulated scenarios listed above consider a final cell concentration of 120 g/L and an rEnzyme/dry cell weight ratio of 10% (w/w), corresponding to an enzyme production titer equal to 12 g/L. Abbreviations: SS316 refers to stainless steel of grade 316; CS refers to carbon-steel. Source: this work.

4.1.8 Economic Analysis

The economic assessment was performed using SuperPro Designer and was based on a plant located in the state of São Paulo, Brazil. For the cost of equipment, the international prices provided by SuperPro Designer software were used; other key economic data required for the economic assessment and their corresponding sources are listed in Table 22 and Table 23 in Appendix B.

Figure 14 – Flowsheet of the recombinant β-glucosidase production process.



Caption: Process flow diagram of the recombinant β-glucosidase production process using *E. coli*, developed with SuperPro Designer. This flowsheet represents the baseline scenario. Note that the main process sections are indicated by different colors: the Upstream Section is displayed in red; the Fermentation Section, in green; and the Downstream Section, in blue. Source: this work.

4.2 Results and Discussion

4.2.1 Process Analysis

Following the design basis presented in Section 4.1.1, the process flow chart presented in Figure 14 was created using SuperPro Designer. This baseline process requires a seed train composed of two smaller fermenters, with volumes of approximately 0.2 m^3 and 4 m^3. During each process cycle, 80 tons of culture medium are fed to the main fermenter, and approximately the same amount of cell broth is found at the end of the microbial culture. This cell broth is collected in a storage tank and then lysed in a pressure homogenizer (throughput of 216 L/min). Next, the cell debris thereby generated are removed using six disk-stack centrifuges (throughput of 36 L/min), and the remaining solids are removed by dead-end filtration (filter area of 50 m^2). Finally, the enzyme extract is concentrated to 15 g/L and stabilized in a citrate buffer using a diafiltration system composed of eight ultrafilters (each with 74 m^2 filter area), such that 22 tons of concentrated enzyme solution, containing 334 kg/enzyme, are produced per process cycle. A total of 264 cycles are performed per year, making up 88.3 tons of enzyme/year (cf. Table 19). This annual production rate would be sufficient to hydrolyze approximately 40% of all the sugarcane bagasse produced annually by an average Brazilian sugarcane plant, considering the assumptions detailed in Section 4.1.1.

Table 19 – Executive Summary of the Enzyme Production Process

Number of Cycles	264	batches/yr
Total Capital Investment	70.8	million US$
Operating Cost	27.2	million US$/yr
Annual Production Rate	88.3	ton of enzyme/yr
Unit Production Cost	316	US$/kg of enzyme
Production per Cycle	334	kg of enzyme/batch

Caption: Executive summary of the β-glucosidase production process. These data refer to the baseline scenario. Source: this work.

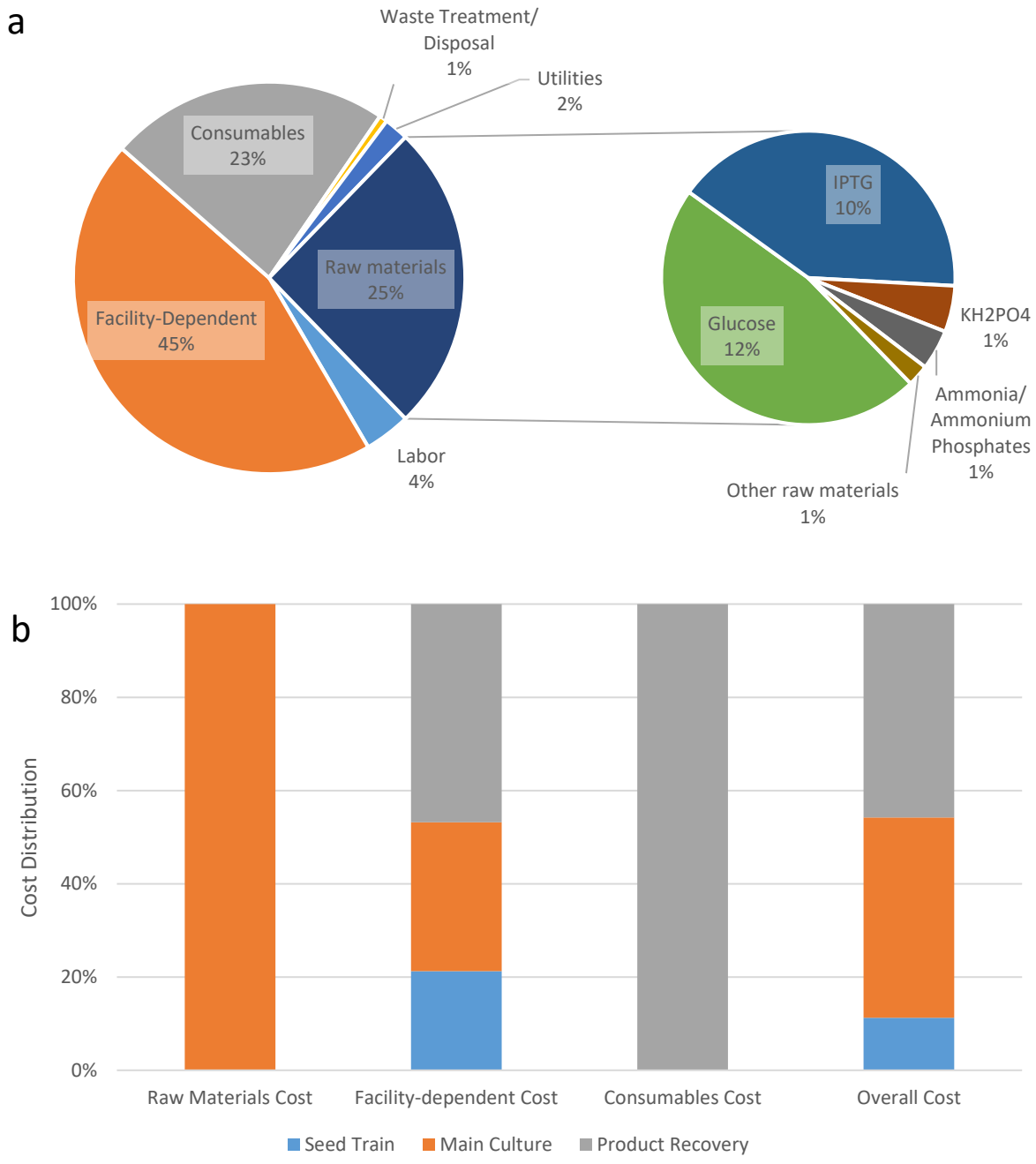
4.2.2 Economic Analysis

As shown in Table 19, the unit production cost was estimated in 316 US$/kg of enzyme, while the total capital costs were estimated in 70.8 million US$ for the baseline scenario (see also Table 26 in Appendix C for a complete list of capital costs). This enzyme production cost is approximately 32 times higher than the estimated cost of the fungal enzyme mixture (10 US$/kg protein), as provided by Klein-Marcuschamer et al. (10), and also higher than other similar enzymes available in the literature (cf. Table 3 in Section 2.4.3). In the context proposed here, of an on-site production of β-glucosidase for fungal cocktail supplementation, such costly BGL would increase the final cost of the fungal cocktail by 137%. We estimated this value considering a fungal enzymatic cocktail with a specific filter paper activity of 0.5 FPU/mg of protein, a BGL specific activity of 2.3 CBU/mg, and that BGL is supplemented at a ratio of 0.2 CBU/FPU. This cost increase would not be justifiable in view of the observed effect of BGL supplementation on the hydrolytic yield (17). Nonetheless, the baseline scenario was intentionally constructed based on conservative assumptions in order to better identify the advantages and limitations of the *E. coli* recombinant system in this context. Moreover, some simulated scenarios that will be described later achieved substantially lower values of BGL cost than the baseline scenario. Thus, in next sections, we investigate all the simulated scenarios and the main drivers of enzyme cost with an eye towards possible cost reduction measures.

4.2.2.1 Cost Composition

Figure 15a shows that in the baseline case, facility-dependent costs, which include plant maintenance, depreciation, insurance, local taxes and overhead costs not directly associated with the process (such as accounting, payroll, fire protection, and security), make up 45% of the unit production costs of the enzyme. Raw materials and consumables (filter cartridges and membranes) account for 25% and 23%, respectively.

Figure 15 – Composition of the enzyme cost in the baseline scenario.



Caption: (a) Composition of the enzyme cost in the baseline scenario for recombinant β-glucosidase production, in terms of cost categories (b) Cost distribution across the three major process sections of the production plant. Source: this work.

It is worth noting that the facility-dependent cost is essentially proportional to the total equipment purchase cost, which was estimated using SuperPro Designer pricing models, based on the U.S. market (cf. Table 25 in Appendix C for a full list of equipment specifications and purchase costs). As a consequence, the facility-dependent cost may be

somewhat overestimated. In fact, Macrelli et al. (276) applied a factor of 0.82 to adjust the fixed capital cost of a bioethanol plant in the U.S. Gulf Coast to Brazilian conditions. Given that the real/dollar currency exchange rate has increased dramatically in 2016 and that the proportion of equipment that would be imported is unknown, we decided not to apply any adjustment factor to the equipment cost provided by the software.

Within the costs of raw materials, glucose and IPTG account for approximately 47% and 41%, respectively, and nitrogen- and phosphorus-rich compounds are together responsible for 10%. The costs of trace elements and, rather surprisingly, kanamycin seem to be negligible. These results confirm the common-sense idea that the use of less expensive carbon sources and induction strategies are important to reduce the enzyme cost.

Regarding the carbon source, it should be stressed that the cost of the glucose used was the market price of the compound. Therefore, it is reasonable to consider whether a glucose-rich liquor generated in the same (2G) plant could, at least in part, replace the purchased glucose, considerably reducing the cost of the carbon source. Similarly, one may envision replacing the purchased glucose with a glucose-fructose syrup obtained by inverting (hydrolyzing) sucrose in a 1G-plant setting or with glycerol, as suggested by Horn et al. (191). The cost of glycerol, in particular, has decreased dramatically during the past decade, mainly because glycerol is a byproduct of biodiesel production, which has greatly increased during the same period. Xylose-rich liquors generated from the hemicellulose hydrolysis process are also a low-cost carbon source that are not well utilized by the conventional ethanol producing organism *S. cerevisiae*. Naturally, the use of these alternative and raw carbon sources could negatively impact the biomass and/or enzyme yields, since these carbon sources usually contain inhibitors of microbial metabolism.

With respect to the cost of induction, IPTG is widely considered to be too costly for the production of inexpensive recombinant proteins, especially at the concentrations at which IPTG is typically used in the laboratory (such as 1 mM). Our results confirm this perception. In fact, the cost contribution of IPTG is comparable to that of the main carbon source, which is 3 orders of magnitude less expensive. However, there are indications that lower IPTG concentrations may give rise to similar or sometimes better volumetric

productivity of recombinant proteins, depending on the specific culture and expression conditions (34). Since IPTG alone accounts for 10% of the unit production cost, reducing the amount of IPTG by one order of magnitude could, theoretically, reduce the enzyme cost by 9%. Alternative induction methods could also be explored, such as replacing the *lac* system with a thermal induction system, which may be particularly convenient for thermotolerant enzymes whose quality is not affected by this additional stress. The use of the *lac* induction system with lactose instead of IPTG may be economically advantageous as well, though probably not for the particular BGL enzyme of the present study, which exhibits high β-galactosidase activity (114).

The effects of reducing the amount of IPTG (by tenfold), reducing the cost of glucose (by tenfold), eliminating kanamycin from the process, and replacing glucose with glycerol were evaluated, assuming in the first three cases that biomass yield and protein productivity are unaffected. According to the simulations, kanamycin elimination and glycerol substitution make little difference in terms of cost, whereas IPTG load and glucose cost reduction have a significant positive impact, reducing the enzyme cost by approximately 10% and 11%, respectively.

Additionally, the cost of consumables, which is quite significant (23%), is mainly due to the cost of the ultrafiltration membranes (80%) used in the diafiltration system and also the cost of the dead-end filtration cartridges (20%). Since the reason to use the dead-end filter is to avoid the fouling of the ultrafiltration membrane, one can conclude that, in our proposed process, the operation of the diafiltration unit has a large direct and indirect economic impact on the cost. In the pursuit of alternative units that are less expensive to operate, it might be interesting to concentrate the enzyme using different methods, such as by precipitation followed by centrifugation. However, the choice of the precipitation agent and the impact of this agent on enzyme activity, recovery yields, process complexity, and the environment should be experimentally evaluated. Our research group has previously evaluated the potential of glycosyl hydrolase precipitation using ethanol under different temperature and pH conditions. In our experience, cellulases can be almost fully recovered using 90% (v/v) ethanol at 25°C and pH 6.5, indicating the potential of this solvent for use in a 2G ethanol plant (277).

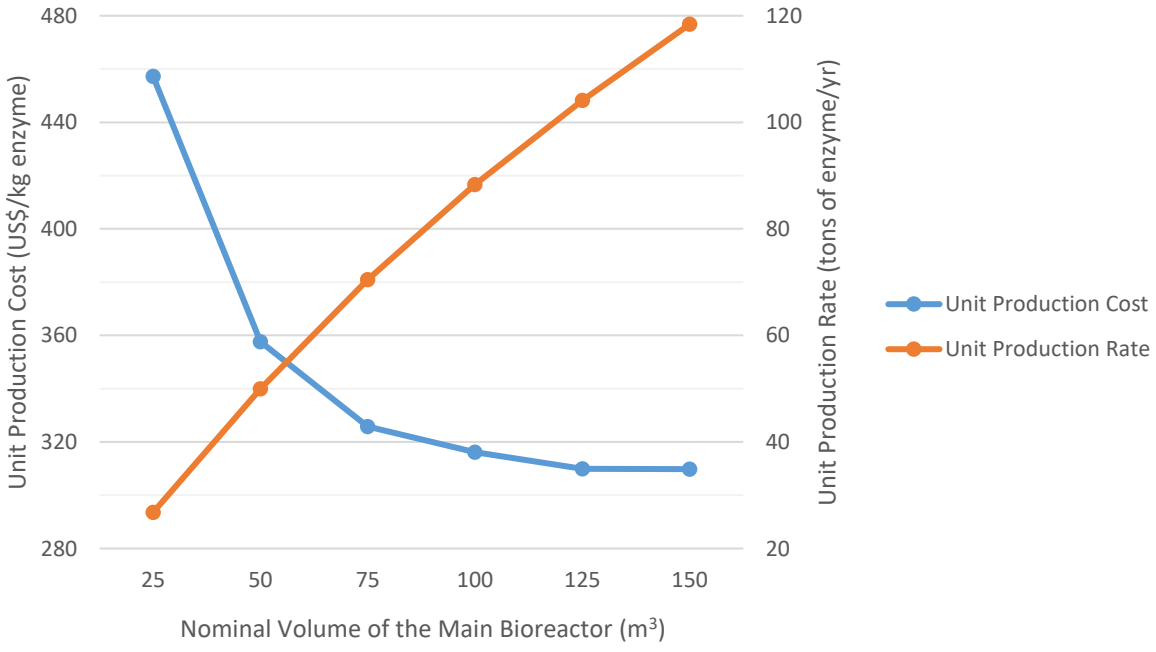
The cost composition presented above refers to the overall process; however, this composition is far from uniform along the process, as seen in Figure 15b. The facility-

dependent cost, for example, is almost evenly distributed among the three main process sections, whereas the cost of raw materials is almost entirely due to the main fermenter feed, and the cost of consumables is entirely due to the downstream section (because the consumables are associated with the operation of filtration equipment). Overall, the fermentation section is the costliest section, followed by the downstream section. Nonetheless, the costs associated with the upstream section are significant (~11%).

4.2.2.2 Effects of Scale, Operating Time and Scheduling

In the baseline case, enzyme costs were calculated considering a process scale that corresponds to a main fermenter volume of 100 m^3. However, it is useful to evaluate the change in enzyme costs associated with a change in the scale of the process, especially considering that the percentage of bagasse set apart for 2G ethanol production in a plant would also depend on the relative prices of ethanol and electricity.

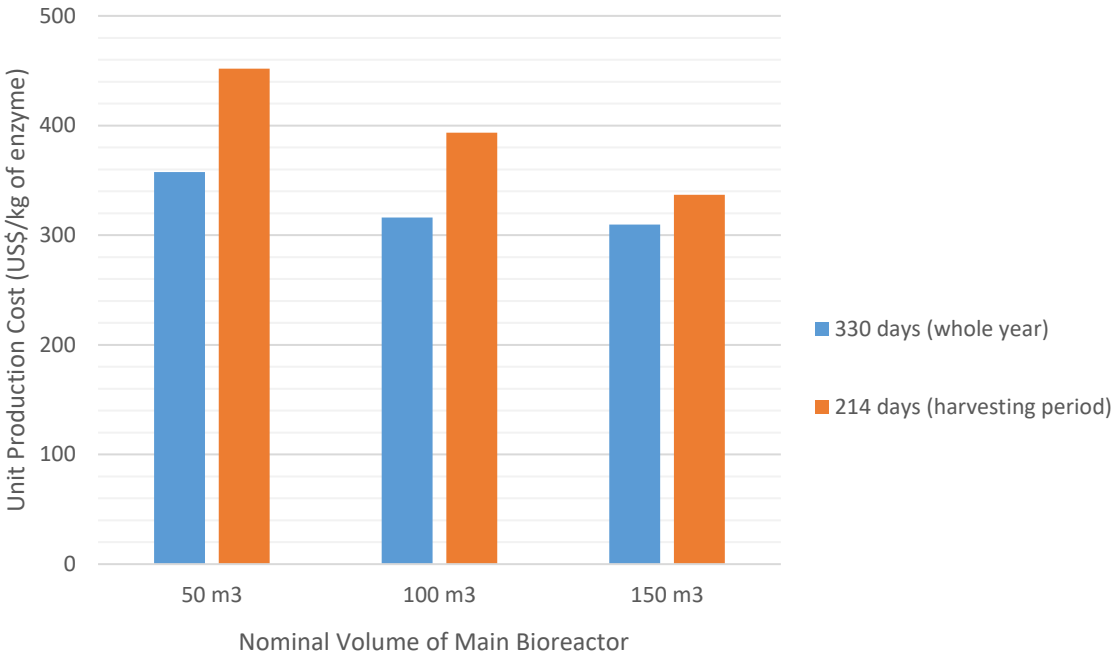
Figure 16 – Variation of enzyme cost and annual production rate with process scale.



Caption: Plot of the unit production cost and of the unit production rate of the β-glucosidase enzyme against the nominal volume of the main bioreactor, here representing the process scale. Source: this work.

As shown in Figure 16, the amount of enzyme produced grows almost linearly with the scale of the process (represented by the main fermenter volume) throughout the range analyzed (from 25 m^3 to 150 m^3). The unit production cost of the enzyme, in contrast, decreases in a non-linear manner as the process scale increases from 25 m^3 to 150 m^3, decreasing drastically at the lower end of the scale and becoming almost flat at the higher end of the scale. The shape of this curve is typical of the phenomenon of economy of scale, largely because the facility-dependent cost becomes relatively smaller as the scale of the process increases. It should be mentioned, however, that the scale-down and scale-up of the process were performed using the software by simply adjusting the process throughput without considering any variations in biomass or enzyme yield that might arise from problems with oxygen transfer or other transport phenomena. Similar to the production scale, the annual operating time of the process strongly affects the unit production cost of the enzyme and is a particularly relevant parameter in the case of an on-site enzyme production plant dedicated to the hydrolysis of lignocellulosic biomass because the harvest of sugarcane does not occur throughout the calendar year, but only between April and November, for approximately 7 months. Since sugarcane bagasse contains a high degree of moisture (50%), this material cannot be stored for long periods. In this context, Santos et al. (278) reported that natural bagasse loses approximately 30% of its calorific content in 150 days. Consequently, if the enzyme production unit were used only for bagasse hydrolysis and if there was no bagasse storage at all, the enzyme plant would remain idle for approximately five months per year (279).

Figure 17 – Variation of enzyme cost with the annual operating time of the plant.



Caption: Unit production cost of the enzyme on three process scales, either considering that the enzyme production plant is active during the whole year (330 days) or only during the sugarcane harvest in the state of Sao Paulo (~214 days). Source: this work.

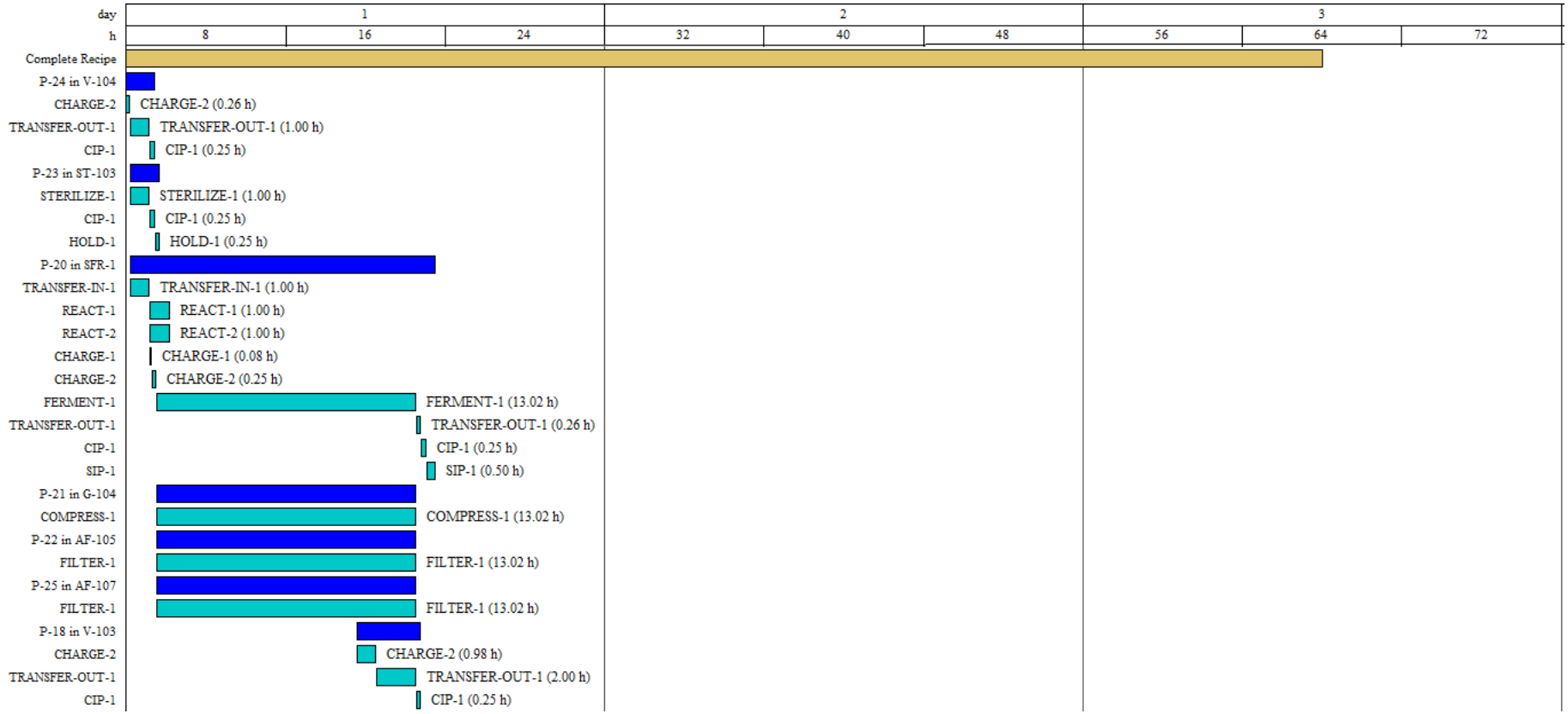
Figure 17 shows how the enzyme cost increases, markedly, from approximately 316 to 393 US$/kg (for a 100-m^3 fermenter), if the annual operating time of the plant is limited to the sugarcane harvesting period. Therefore, it seems evident that a strategy for the safe, long-term storage of bagasse would have to be developed. To this aim, drying the bagasse would probably be vital, given that this residue becomes essentially stable with respect to microbial activity if its moisture content is decreased to 20% or less (280,281).

Together with the annual operating time, the minimum cycle time of the process determines the maximum number of batches per year. In the baseline case, the minimum cycle time was found to be 30 h, corresponding to the time required to pump the culture medium into the main fermenter, to carry out the fed-batch culture per se, and to transfer the culture broth to a storage tank as the culture ends; in other words, the set of operations that occur in the main fermenter constitute the scheduling bottleneck of the process. This becomes clear by examining the Gantt chart of the whole process, shown in Figure 18, in which the label FR-101 designates the main fermenter (a brief description of all the unit operations and their respective labels are provided in Table 24, Appendix C). Given a minimum cycle time of 30 h and an annual operating time of 330 days (7920

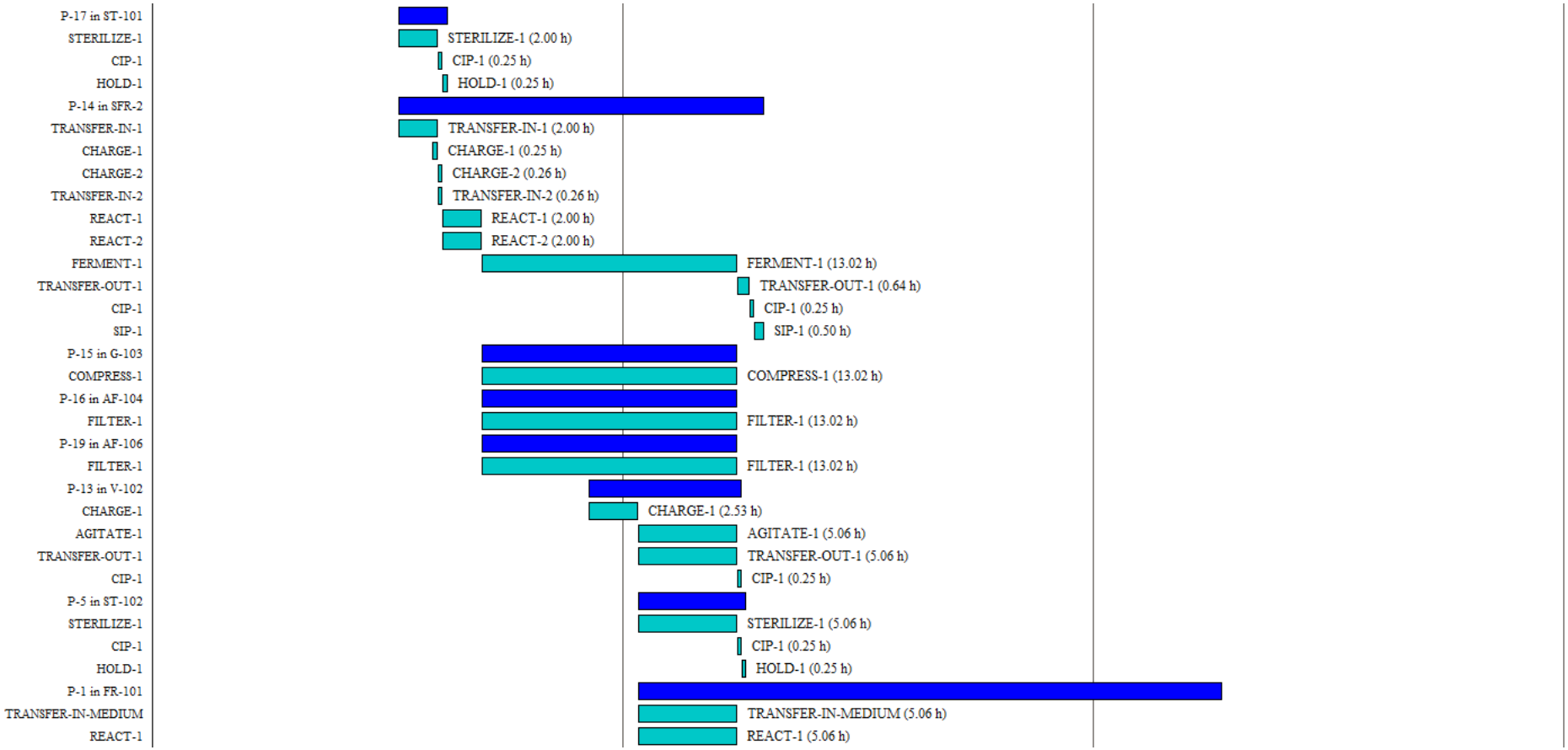
h), the maximum number of batches per year is equal to 264 (cf. in Table 19). It is worth noting that the complete duration of a single batch is not equal to the minimum process cycle time, but longer: 60.15h. Such difference exists because a new batch starts before the current one ends.

A potential strategy to reduce the unit production cost of the enzyme is to increase the number of batches per year by reducing the minimum cycle time. In practice, this can be accomplished by acquiring another main fermenter, that would be used in stagger mode with the first one. As a result, the minimum cycle time is reduced to the fed-batch culture duration, that is, 22 h. The maximum number of batches then increases from 264 to 357; the annual production rate increases from 88.3 ton to 119.4 ton; and the unit production cost of the enzyme decreases from 316 to 296 US$/kg. However, this process scheme also implies a large increase in the total capital investment, from 70.8 million US$ to 80.4 million US$, due to the purchase and installation of another 100-m^3 fermenter.

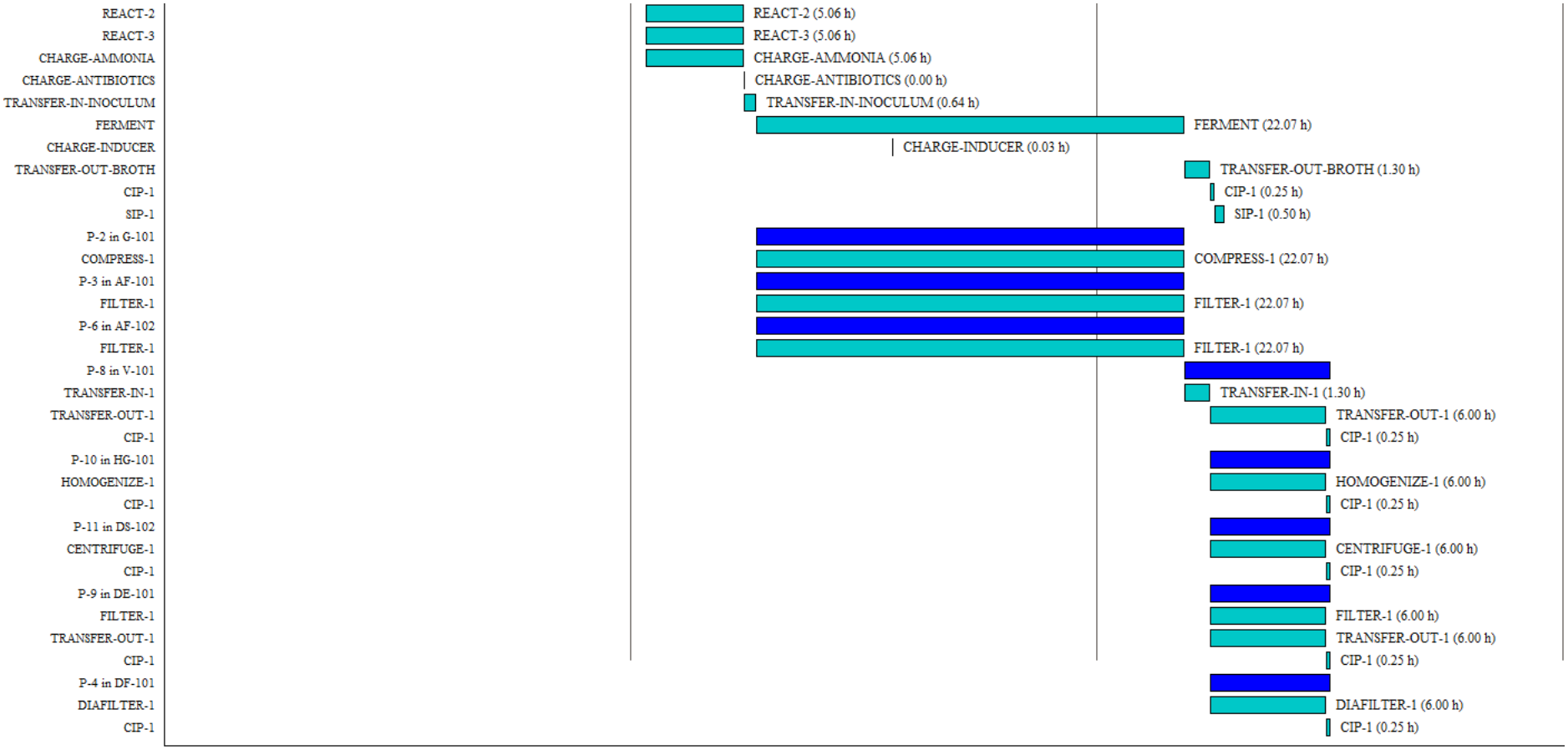
Figure 18 – Gantt chart of the enzyme production process.



Caption: Gantt chart of a single batch of the baseline process for enzyme production. Each acronym refers to a piece of equipment or to an operation carried out in a piece of equipment. The meaning of each acronym is given in Appendix C. Note: Chart split into three pages (pp. 124–126). Source: this work.



Caption: Gantt chart of a single batch of the baseline process for enzyme production. Each acronym refers to a piece of equipment or to an operation carried out in a piece of equipment. The meaning of each acronym is given in Appendix C. Note: Chart split into three pages (pp. 124–126). Source: this work.



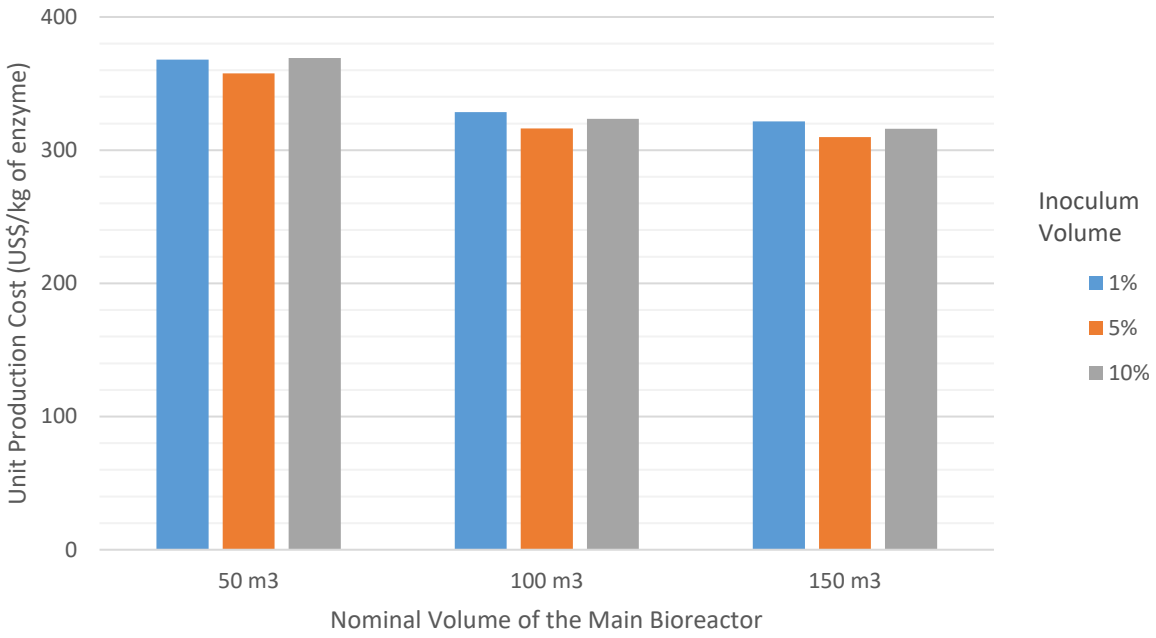
Caption: Gantt chart of a single batch of the baseline process for enzyme production. Each acronym refers to a piece of equipment or to an operation carried out in a piece of equipment. The meaning of each acronym is given in Appendix C. Note: Chart split into three pages (pp. 124–126). Source: this work.

4.2.2.3 Upstream Section

Often overlooked in the literature, the seed train is a key part of any industrial bioprocess, since it is responsible for the propagation of the microorganism from small volumes to large bioreactors. Ideally, the seed train should preserve the desirable characteristics and the viability of the microorganism while avoiding any contamination. Here, the effect of inoculum volume on the cost of the recombinant enzyme for different process scales was simulated, and the result is presented in Figure 19, as represented by the volumes of the main fermenter. For all three scales evaluated, an inoculum volume of 5% led to the lowest cost, although the advantage with respect to inoculum volumes of 1% or 10% appears modest. At the lowest scale (50 m^3), an inoculum volume of 10% led to the highest cost, whereas at the 100 m^3 and 150 m^3 scales, an inoculum volume of 1% led to the highest cost. These results reveal the countervailing effects of inoculum size on enzyme cost: on one hand, larger inoculum volumes reduce the duration of the main culture, thereby increasing the number of batches per year and reducing the enzyme production cost. On the other hand, larger inoculum volumes require larger and more numerous seed bioreactors, thereby increasing the facility-dependent cost and the enzyme production cost.

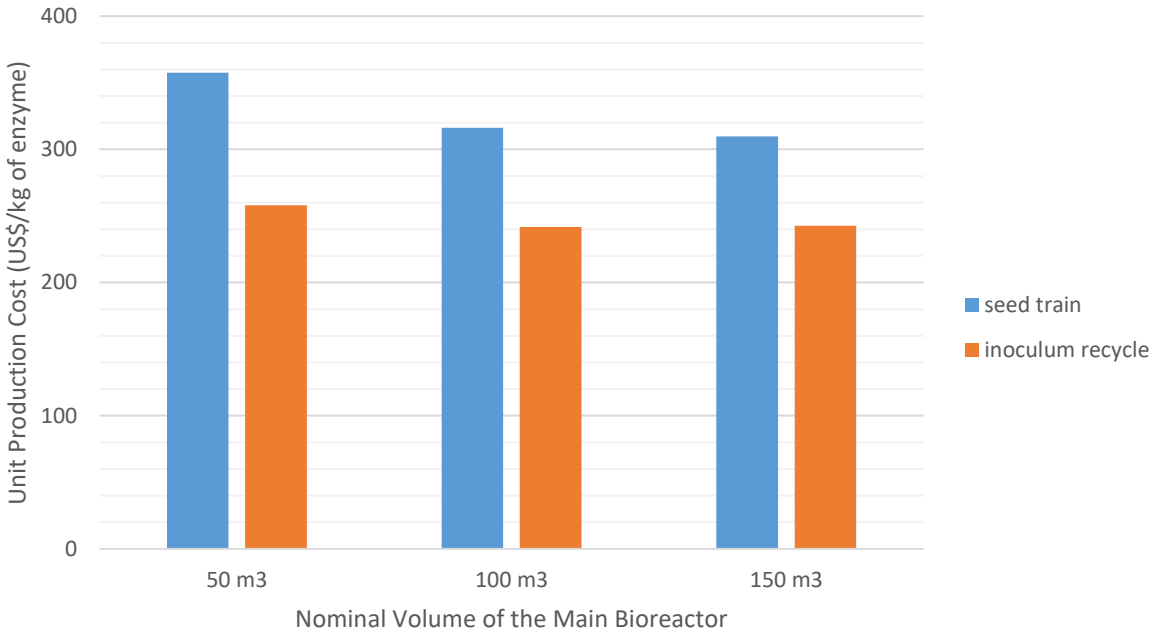
An alternative inoculum production scheme, in which a fraction of the seed culture is retained in the seed fermenter to jumpstart the next seed culture, was also evaluated. The simulation results are given in Figure 20, and show large cost reductions (> 20%) on all assessed scales, demonstrating the great potential of this process intensification strategy. It is worth noting that the cost cuts occur in spite of a slight increase in the total equipment cost of the process (data not shown), since the seed fermenter employed in the inoculum recycle strategy is substantially larger than the seed fermenters used in the conventional seed train scheme. Moreover, the inoculum recycle scheme approximately doubles the annual production rate on all scales (data not shown), because it approximately halves the main fermentation duration, which is the longest step in the process.

Figure 19 – Variation of the enzyme production cost with the inoculum volume.



Caption: Variation of the unit production cost of the enzyme with the inoculum volume, for different process scales. Source: this work.

Figure 20 – Comparison of inoculum production strategies: seed train *versus* inoculum recycle.



Caption: Comparison of two different strategies for inoculum production: a conventional seed train scheme using 5% of inoculum volume against a novel inoculum recycle method. Three process scales, designated by the nominal volume of the main bioreactor, are presented. Source: this work.

However, certain technical challenges must be addressed for the successful implementation of this strategy on industrial scales. Firstly, on a large scale, the mere transfer of the cell broth from the seed fermenter to the main fermenter takes a significant time, and presumably occurs under uncontrolled conditions of temperature, pH, and substrate availability, as well as significant pressure; all these conditions can provoke a variety of physiological responses in the cells, such as decreases in the specific growth rate or plasmid copy number, and ultimately lead to lower rProtein productivities. Secondly, various sources of genetic instability may afflict both chromosomal and plasmidial DNA over time, so that the recombinant platform would have to be regularly monitored; in fact, a brand-new inoculum would likely have to be produced from time to time. All these issues could obviously undermine the cost reductions projected in the simulations, and consequently must be thoroughly examined on smaller scales before considering industrial applications. It also bears stressing that this inoculum recycle scheme is entirely dependent on the use of recombinant platforms displaying extremely high plasmid segregational stability (near 100%), such as the SE1 strain; otherwise, plasmid-free cells would eventually overtake the culture.

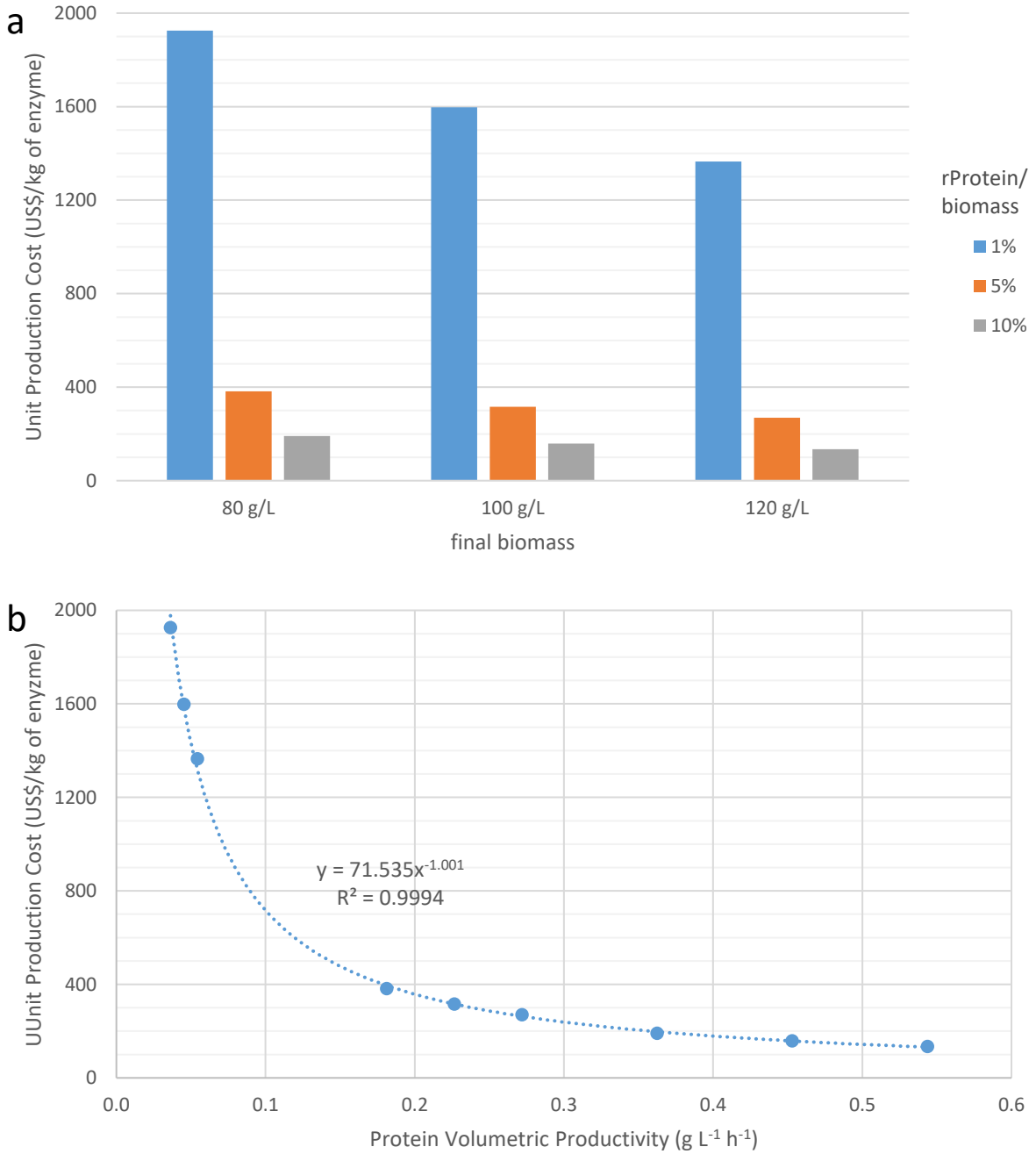
4.2.2.4 Fermentation Section

The effect of the biomass concentration at the end of the fermentation and the effect of the recombinant protein (rEnzyme) content are presented in Figure 21a. Clearly, both variables have a dramatic influence on enzyme cost; the case where the lowest biomass is coupled with the lowest rEnzyme content (1926 US$/kg) and the case where the highest biomass is coupled with the highest rEnzyme content (135 US$/kg) are 1.4 orders of magnitude apart. These results confirm the importance that has generally been ascribed to the volumetric productivity of fermentation processes, which is defined as the mass of the product at the end of the process divided by the final broth volume and the duration of fermentation. To better visualize the effect of this parameter, the biomass concentration and rEnzyme content data were combined and converted into volumetric productivity and plotted against enzyme cost, as shown in Figure 21b. The chart shows that the enzyme cost does indeed decrease rapidly with volumetric productivity, and the

data are very well approximated by a power law in which the cost is inversely proportional to the volumetric productivity. However, it is well known that rEnzyme content may negatively correlate with biomass concentration as a result of the so-called metabolic burden of recombinant protein synthesis and as a result of the toxic properties of inducer molecules such as IPTG (191). Consequently, these findings indicate the need for a better understanding of the tradeoffs involved in recombinant protein production and, in particular, for experimentally identifying optimum induction conditions and the final biomass concentration range.

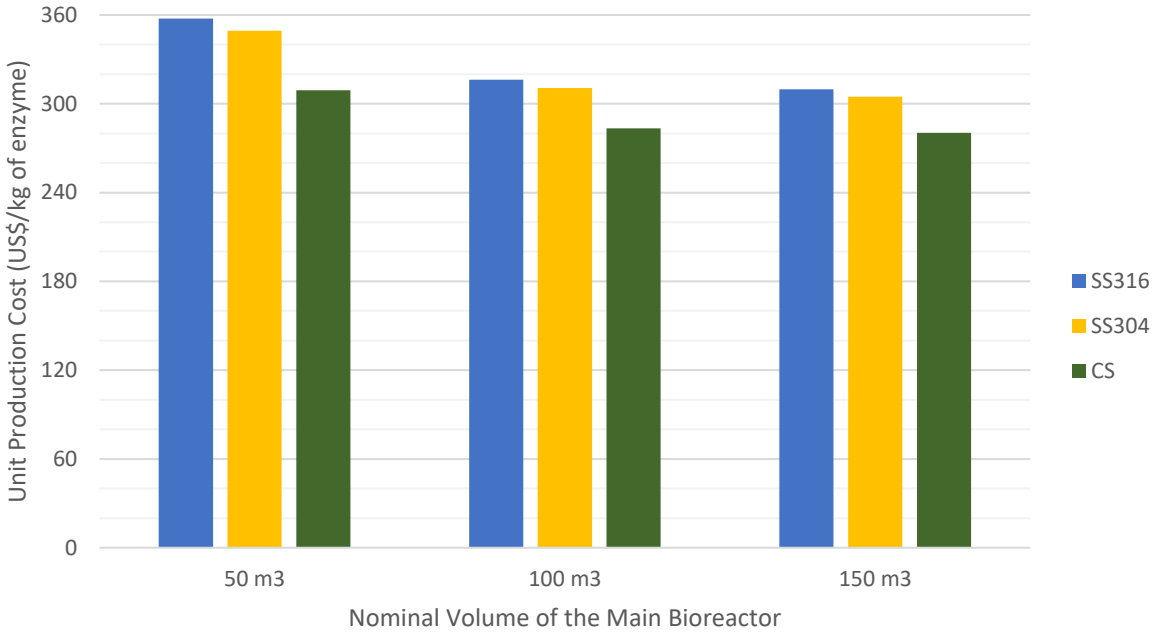
A parameter frequently overlooked in techno-economic analyses of bioprocesses is the material of the fermenters and storage tanks, which must resist the frequently corrosive products and by-products of fermentation. Grade 316 stainless steel is considered to be the standard bioreactor material for most biotechnological processes (282). In fact, grade 316 steel was the material used in the simulations of fungal cocktail production performed by Humbird et al. (9). However, these authors proposed the use of a lower grade of stainless steel, 304, for a *Zymomonas* fermenter. Furthermore, carbon steel has a precedent in the American corn ethanol industry (9) and the Brazilian sugarcane ethanol industry (283). Thus, simulations were conducted assuming that the fermenter, the seed fermenters and the storage tanks were made of stainless steel of higher grade (SS316), stainless steel of lower grade (SS304) or a carbon-steel alloy (CS); the simulations were conducted using cost models from SuperPro Designer. The results are shown in Figure 22. In comparison with standard SS316, the use of SS304 had a negligible effect on the enzyme cost, whereas the substitution of SS316 for CS was significant, decreasing the cost from 316 US$/kg to 283 US$/kg on a process scale of 100 m^3. However, these results do not account for possible requirements for corrosion prevention, such as the application of special coatings (283) or the need to design a thicker fermenter (9). Therefore, the replacement of SS316 with less expensive materials is not warranted by these results.

Figure 21 – Variation of enzyme cost with biomass and rEnzyme content.



Caption: (a) Variation of the unit production cost of the enzyme with the biomass reached at the end of the bacterial culture, for different values of enzyme content (weight fraction of rProtein/biomass) (b) Combined effect of the previous parameters (final biomass and enzyme content) with the fermentation duration, under the form of volumetric productivity, on enzyme cost. Source: this work.

Figure 22 – Variation of enzyme cost with equipment material.

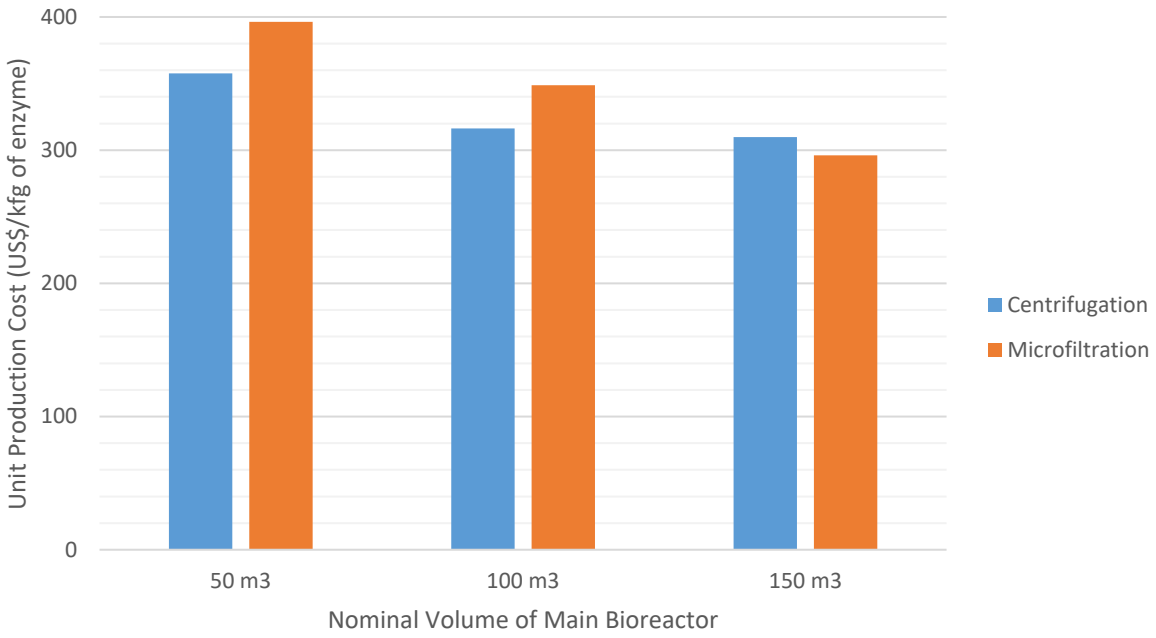


Caption: Variation of the unit production cost of the enzyme with the material of the bioreactors and storage tanks, on three process scales. Labels: SS316, stainless steel of grade 316; SS304, stainless steel of grade 304; and CS, carbon-steel alloy. Source: this work.

4.2.2.5 Downstream Section

Despite being extremely streamlined, the downstream section accounts for nearly half of the enzyme cost, as mentioned earlier. This can be partially attributed to the contribution of the downstream section to the facility-dependent cost (43% of the total). For this reason, some changes in downstream equipment were evaluated. Firstly, centrifugation was replaced with microfiltration. The choice between these unit operations is commonly encountered in the biotechnology industry, and although centrifugation is generally considered to be more cost effective at smaller scales (284), the choice must be evaluated on a case-by-case basis. The results, presented in Figure 23, indicate that centrifugation is the best option to separate cell debris at both the baseline scale (100 m^3) and the smaller (50 m^3) scale, whereas microfiltration was less costly at the larger scale (150 m^3). Nevertheless, the difference in cost between these two separation methods was negligible at the smaller scale and quite modest (approximately 5%) at the largest scale.

Figure 23 – Comparison of centrifugation and microfiltration for cell debris removal.



Caption: Unit production cost of the enzyme when using either centrifugation or microfiltration for removal of cell debris after cell disruption, assessed at three different process scales. Source: this work.

The substitution of diafiltration by a simpler ultrafiltration was also assessed, considering that enzyme stabilization in an appropriate buffer may not be strictly necessary in the case of on-site enzyme application. As a result, the enzyme cost was noticeably reduced, by 8%, to 292 US$/kg.

Alternative downstream operations, such as precipitation using ethanol, might also be used to concentrate enzymes from the clarified lysate (285). However, additional unit operations for precipitate recovery (centrifugation or filtration) and resuspension have to be considered. Aqueous-two phase systems (ATPS) have been effectively used for separation and purification of industrial proteins as well. An elegant review on this topic was published by Asenjo and Andrews (286). The production and purification of chymosin from recombinant *Aspergillus* supernatant is the most successful industrial application of this technology. Silvério et al. (287) studied the separation and purification of laccase from a complex fermented medium using an ATPS system with a thermo-separating polymer. Despite the advantage of the recovery and reutilization of the polymer, a high loss in activity was observed (88%) when compared to the classical PEG-salt systems. In general, precipitation and ATPS are adequate if some increase in purity

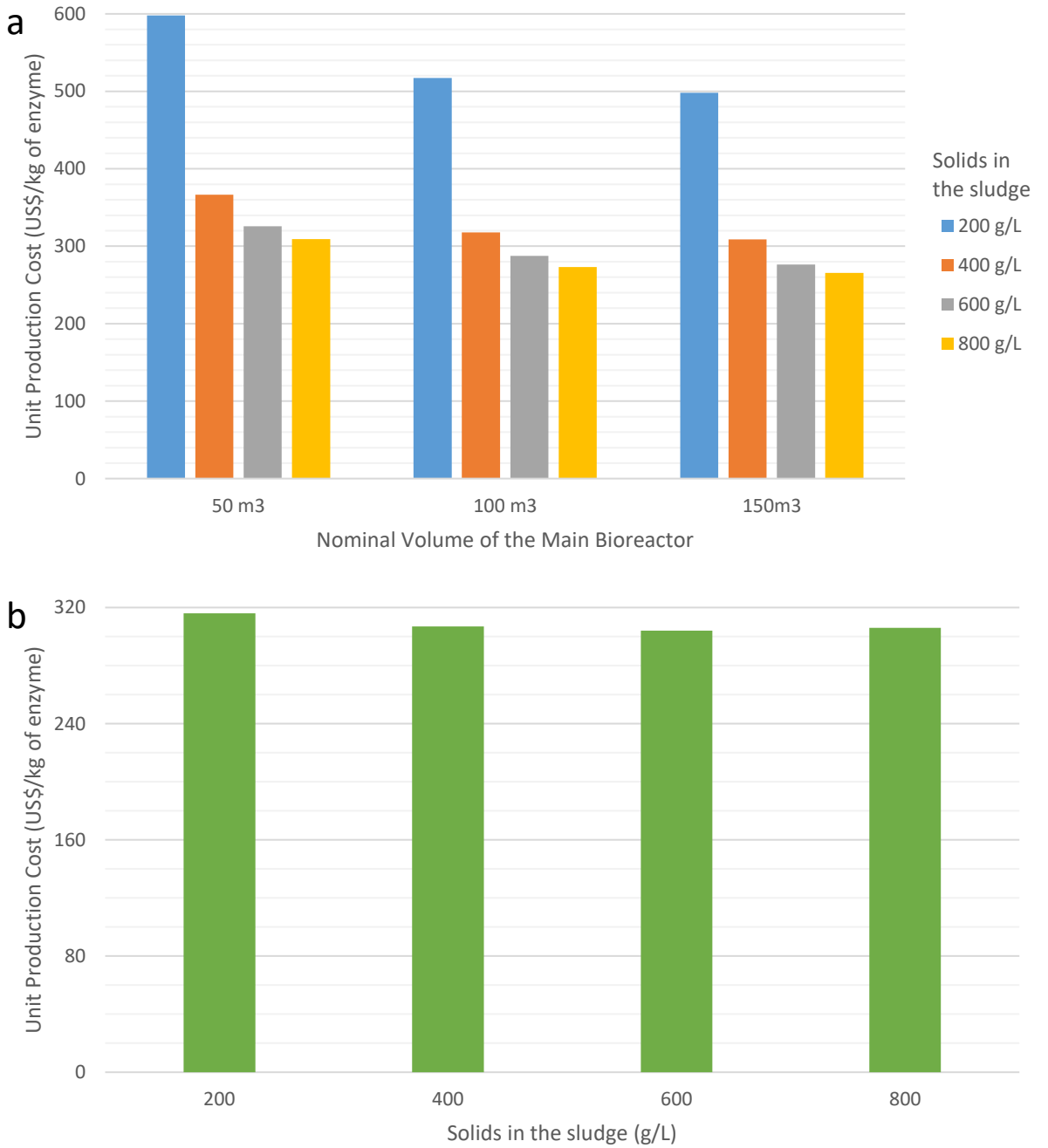
level of the protein is needed (210). Although a more comprehensive study of the possible downstream process designs would be very interesting, in this work we have focused on the most common downstream process configurations described in similar studies, such as those listed in Table 3 in Section 2.4.3.

Finally, an important factor responsible for the low cost of fungal cellulases is the extracellular production (secretion) of the enzymes, which leads to simplified downstream processing. The use of cheaper unit operations for cell removal, such as the vacuum drum filter (well suited to fungal biomass separation) also contributes to the reduced fungal enzyme cost. Klein-Marcuschamer et al. (10), for instance, have used a vacuum drum filter to separate *T. reesei* cells from the enzyme mixture. The NREL models (9,67,232,234) go even further, by not separating the fungal biomass from the enzyme mixture at all; consequently, the dead fungal cells are introduced into the reactor of enzymatic hydrolysis along with the enzyme mixture.

Extracellular protein production by gram-negative bacteria such as *E. coli* is uncommon; however, when producing the TpBgl1 enzyme from *Thermotoga petrophila* in recombinant *E. coli*, Haq et al. (113) observed higher BGL activity in the extracellular medium than in cell lysate, even though no particular strategy was put in place to favor protein secretion. Our experimental results showed, on the contrary, extracellular BGL activity to be negligible, possibly owing to the choice of widely different culture and expression conditions (cf. Section 3.2.3). In any case, we found interesting to simulate scenarios where the enzyme was secreted at the same expression level assumed for intracellular expression in the baseline case. Here, however, centrifugation was employed to separate the whole cells from the liquid phase, which contains the enzyme. The positive impact of enzyme secretion on cost reduction was confirmed, due to simplification in the downstream section; however, the enzyme cost was found to strongly depend on the solids content of the centrifuge sludge, as observed in Figure 24a. Considering a process scale of 100 m^3 and a solids content of 600 g/L (24.4% of DCW) or 800 g/L (27.2% of DCW) in the centrifugate, cost reductions are obtained in relation to the baseline scenario (288 US$/kg and 273 US$/kg, respectively, against 316 US$/kg in the baseline case). A concentration of solids of 400 g/L (20.2% of DCW) in the sludge, on the other hand, leads to an enzyme cost of 318 US$/kg, close to that of the intracellular baseline scenario; and a concentration of 200 g/L (13.3% of DCW) gives rise to a much higher cost, of 517

US$/kg. Such behavior, also observed on the other process scales, occurs because low concentrations of solids in the sludge entail large enzyme losses in the sludge, thus reducing the overall process yield.

Figure 24 – Variation of enzyme cost with solids concentration in the centrifugation sludge.



Caption: Variation of the unit production cost of the enzyme with the concentration of solids in the centrifugation sludge, in (a) extracellular production (b) intracellular production. In the extracellular case, three process scales are presented. Source: this work.

The effect of the concentration of solids in the centrifugate was also evaluated for intracellular enzyme production on the intermediate (100 m^3) production scale; the results, summarized in Figure 24b, show that enzyme cost is much less sensitive to the concentration of solids in this case (which validates our previous intracellular results, which applied a fixed concentration of solids equal to 200 g/L). This distinct response between intracellular and extracellular production is explained by the fact that whole bacterial cells have a very high water content (assumed to be 70%), as opposed to cell debris, which are water-free. As a consequence, much higher volumes of centrifugation sludge are generated in the extracellular case, and if the concentration of solids in the sludge is relatively low, a large amount of enzyme is wasted, as explained above.

4.2.2.6 High-titer Scenarios

In Section 4.2.2.4, the decisive effect of volumetric productivity on the enzyme cost was clearly demonstrated; in fact, the production cost in the scenario of highest enzyme titer (12 g/L of enzyme) was found to be 57% lower than that in the baseline scenario (135 against 316 US$/kg). Therefore, in the interest of assessing the cost limits of recombinant enzyme production in *E. coli*, we simulated various scenarios based on the high-titer (12 g/L) case.

Firstly, given that the enzyme concentration in the cell lysate (11.5 g/L) is reasonably close to the target enzyme concentration in these scenarios, we decided to do away with the unit operation steps required to concentrate and stabilize the enzyme, which is not such a farfetched proposal in the context of on-site enzyme production within a biorefinery; in fact, other works in the literature that simulate on-site enzyme production make similar assumptions (9,10). This procedure alone reduces the enzyme cost by 32%, to 92 US$/kg. Then, we applied to the high-titer case several of the same process improvements applied to the baseline case, one at a time: tenfold reduction of IPTG use, tenfold reduction of glucose cost, use of inexpensive fermenter material (carbon steel), purchase of an extra main fermenter to be used in stagger mode, and use of the inoculum

recycle strategy for inoculum production. These measures led to cost reductions of 36-53% with respect to the previous case, as shown in Table 20.

Extracellular production can also be combined with a high production titer (12 g/L); in this case, protein concentration steps are not required, and a fairly high final titer is achieved (~18 g/L). Moreover, the enzyme cost is further reduced to 85 US$/kg, which corresponds to an 8% reduction with respect to the intracellular, high-titer, no concentration case. In addition, if the process improvements previously applied to the high-titer intracellular scenario are individually applied to high-titer extracellular production, further cost reductions can be obtained, in the range of 37-54%, as detailed in Table 20. Lastly, if all the factors presented above (with the exception of the extra main fermenter[**]) are combined with high-titer production, huge cost reductions are achieved: 21 US$/kg in the intracellular case and 20 US$/kg in the extracellular case. Although these figures do not match the values estimated for fungal cellulase mixtures in the literature, they do indicate that, with appropriate improvements on enzyme expression levels and intelligent choices in process design, enzyme costs within the range of 20-80 US$/kg could be realistically achieved using recombinant *E. coli*.

[**] When combined with the inoculum recycle scheme, the use of an extra main fermenter in stagger mode actually promotes a slight increase in enzyme cost. This happens because, when applied to the baseline scenario, the extra main fermenter halves the main fermentation duration, which is the longest step in the whole process, and consequently reduces the overall process cycle time; however, when combined with the inoculum recycle strategy, the extra fermenter does not reduce the overall process cycle time, given that the main fermentation and the seed fermentation, by design, have the same duration in this case. The additional facility-dependent costs associated with a new fermenter account for the modest cost increase in the latter case.

Table 20 – Enzyme cost for various simulated high-titer scenarios.

Production Type	Production Titer (g/L)	Fermenter Material	IPTG amount	Glucose Cost	Extra Fermenter	Cell recycle	Concentration Step	Final Titer (g/L)	Enzyme Cost (US$/kg)	Cost Reduction
Intracellular	12	SS316	1X	1X	no	no	yes	15.0	**135**	—
Intracellular	12	SS316	1X	1X	no	no	no	11.5	**92**	32%
Intracellular	12	CS	1X	1X	no	no	no	11.5	**78**	42%
Intracellular	12	SS316	0.1X	1X	no	no	no	11.5	**79**	41%
Intracellular	12	SS316	1X	0.1X	no	no	no	11.5	**75**	44%
Intracellular	12	SS316	1X	1X	yes	no	no	11.5	**86**	36%
Intracellular	12	SS316	1X	1X	no	yes	no	11.5	**64**	53%
Intracellular	12	CS	0.1X	0.1X	no	yes	no	11.5	**21**	84%
Extracellular	12	SS316	1X	1X	no	no	no	18.4	**85**	37%
Extracellular	12	CS	1X	1X	no	no	no	18.4	**72**	47%
Extracellular	12	SS316	0.1X	1X	no	no	no	18.4	**73**	46%
Extracellular	12	SS316	1X	0.1X	no	no	no	18.4	**68**	50%
Extracellular	12	SS316	1X	1X	yes	no	no	18.4	**82**	39%
Extracellular	12	SS316	1X	1X	no	yes	no	17.7	**62**	54%
Extracellular	12	CS	0.1X	0.1X	no	yes	no	17.7	**20**	85%

Caption: Unit production cost of the enzyme for various intracellular and extracellular scenarios, evaluating several process improvements Source: this work.

5 CONCLUSION

Despite the considerable technical and economic challenges that surround the production of fuels and chemicals from lignocellulosic biomass, as well as the industrial production of low-cost recombinant enzymes using *E. coli*, we believe that our work sheds light on several relevant questions.

Firstly, the experimental section shows that hyperthermophilic β-glucosidase TpBgl1 can be cloned and expressed into an *E. coli* system that dispenses with the use of antibiotics (*E. coli* SE1), and that high levels of enzyme activity on pNPG can be achieved on this system. These levels are similar (slightly higher) than those obtained with the BL21(DE3) strain. Moreover, the SE1 recombinant platform has displayed exceptional plasmid segregational stability (approximately 100%), and more so than the BL21(DE3) strain, when subjected to repeated passages in shake-flasks. This ability addresses a major problem in using multicopy plasmid recombinant platforms for efficient rProtein production on industrial scale, which is plasmid loss over prolonged cultures, and opens new possibilities for inoculum production on large scales.

Secondly, the techno-economic analysis of the production of the BGL enzyme for application on biomass hydrolysis, using recombinant *E. coli*, projects an enzyme cost of 316 US$/kg (in the baseline scenario), which is high in relation to the average cost of fungal cellulase cocktails. The facility-dependent cost, which is strongly associated with the cost of equipment, accounts for roughly half of the estimated cost, while the cost of raw materials, especially IPTG and glucose, and the cost of consumables (filtration cartridges and membranes) are all quite significant. However, the simulation of multiple scenarios and optimization measures indicate that the enzyme cost can be substantially reduced on many fronts, such as: substituting the (pure) carbon source for a similar, but cheaper, alternative, e. g. glucose-fructose syrup or glycerol; reducing the amount of IPTG used for induction or changing the induction strategy altogether; or replacing the diafiltration operation with cheaper alternatives. In fact, in the particular case of on-site enzyme use, concentration and stabilization of the enzyme may not be necessary, which could lead to large cost reductions. Developing *E. coli* strains or fermentation processes that generate high rEnzyme volumetric productivities can also significantly reduce the

cost of the enzyme, since the importance of this parameter for the process economics was corroborated by the simulations. In fact, the enzyme cost drops to approximately 135 US$/kg in the best productivity scenario (0.54 g L^{-1} h^{-1}), which corresponds to a 57% cost reduction with respect to the baseline scenario. With the *E. coli* SE1 system in mind, an alternative strategy to produce large inoculum volumes was also devised, in which a single seed fermenter is used: a fraction of the cell mass is preserved after each seed culture, so that the addition of fresh medium to the fermenter allows the immediate restart of the seed culture. This design avoids the need of an extensive seed train, which is costly and burdensome on a large scale. Furthermore, this inoculum production scheme can be carried out in such a manner that the cell cultures in the seed fermenter and in the main fermenter are synchronized, thereby reducing the overall process cycle time. As a consequence, higher annual production rates and lower enzyme production costs can be achieved. The extracellular production of the BGL enzyme using *E. coli*, reported in the literature, was simulated as well; the results reveal that secretion can indeed reduce the enzyme cost, provided that the clarification step is conducted efficiently. The combination of all these improvement measures could ultimately lead to a final enzyme cost near 20 US$/kg of protein, which is reasonably close to the cost of fungal cellulase mixtures reported in the literature.

Altogether, this work provides a comprehensive analysis of the factors affecting the production cost of low value-added enzymes using *E. coli* on a large scale, contributing to a better understanding of the actual potential use of this host in industry, particularly in the context of tailor-made enzymatic cocktails for biomass hydrolysis. Moreover, this study shows that the hyperthermophilic TpBgl1 enzyme can be successfully produced in *E. coli* SE1 with high activity levels, and that the toxin-antitoxin system provides nearly perfect plasmid segregational stability, even after repeated passages. These features could be exploited for the large-scale production of low-cost recombinant enzymes, especially if coupled with the novel inoculum production scheme proposed here.

5.1 Future works

The interdisciplinary nature of the present study opens many avenues for future investigation. On the experimental side, it would be useful to carry out fed-batch cultures using either *E. coli* BL21(DE3) or SE1, following the protocol considered in the techno-economic analysis, in order to ascertain the practical volumetric productivity of the process. Then, this productivity value could be employed as a parameter for further strain improvement through techniques of microbial engineering, such as evolutionary engineering, metabolic engineering, etc. In the particular case of the SE1 strain, the antibiotic resistance gene present in the plasmid could also be removed to mitigate the plasmid metabolic burden and, hopefully, increase the process productivity.

The reports of BGL extracellular production by Haq et al. (113) should be investigated as well; possibly, a signal peptide could be adjoined to the N-terminus of the enzyme, with the aim of sending the product to the extracellular medium, or at least accumulate it in the periplasm. The replacement of the *E. coli* host with a gram-positive bacterium such as *Bacillus subtilis* or even a yeast such as *Pichia pastoris* could be another potential strategy for secretory production of the enzyme.

Given the considerable contribution of (pure) glucose to the enzyme cost, it would also be interesting to experimentally evaluate cheaper carbon sources such as the C6-liquor from cellulose hydrolysis, glucose-fructose syrup or even glycerol from biodiesel production. Here, strain engineering would probably prove to be crucial. By the same token, the large contribution of IPTG to the enzyme cost encourages further research on the relationship between IPTG concentration and enzyme productivity. The use of lactose as a cheaper alternative to IPTG could also be assessed, and, more generally, alternative induction methods could be evaluated, such as thermal induction or xylose-based systems.

Since the BGL studied in this work displays a much lower activity on cellobiose than on pNPG (according to the literature), the enzyme could also be modified to hydrolyze cellobiose more efficiently, by means of protein engineering techniques. The stimulant effect of the enzyme product (glucose) on the enzyme activity also needs to be better understood; furthermore, it could be exploited to improve the process of hydrolysis. The

hyperthermophilic nature of the enzyme also provides a simple method of enzyme recovery based on thermal treatment, which should be experimentally evaluated as well. Likewise, enzyme recovery through precipitation with ethanol or similar substances is a potential alternative that has not been experimentally assessed yet.

Further research on the inoculum recycle scheme would also be vital to implement this promising strategy on industrial scale. Bioreactor assays similar to the ones carried out in shake-flasks in this work, but closer to the culture conditions on industrial scale, would probably be the first step. Moreover, not only plasmid segregational stability, but the genetic stability of the recombinant system in a broader sense (e. g. plasmid structural stability, allele segregation, host mutations, etc.) and plasmid copy number should be analyzed, as well as potentially deleterious effects of such strategy on enzyme productivity.

With respect to the techno-economic analysis, a more precise enzyme cost could be achieved if equipment prices were obtained for the particular context of the Brazilian sugarcane-ethanol industry, instead of using SuperPro Designer data. Research institutions such as the Brazilian Bioethanol Science and Technology Laboratory (CTBE) and the Institute of Applied Research (IPT) possess data banks containing such information. Furthermore, all experimental evaluations proposed above could be introduced into the simulator and validated as well.

Finally, a comprehensive analysis of the environmental impact of enzyme production would be interesting, especially taking into account that biomass hydrolysis is touted as an ecological process. In this respect, a Life Cycle Analysis (LCA) of the entire enzyme production process would be the ideal tool to assess the environmental footprint of the enzyme. Within this analysis, the impact of the antibiotic required for the BL21(DE3) platform should be given particular attention.

All in all, a great deal of experimental research has yet to be done to develop efficient microbial platforms for industrial enzyme production, and this work must be guided by ever more accurate technical, economic and environmental assessments. Nonetheless, we believe that our thesis demonstrates the great potential of microbial enzymes for a more sustainable production of chemicals and fuels in the future.

REFERENCES[††]

(1) CHUNDAWAT, S. P. S. et al. Deconstruction of Lignocellulosic Biomass to Fuels and Chemicals. **Annual Review of Chemical and Biomolecular Engineering**, v. 2, n. 1, p. 121–45, 2011.

(2) NANDA, S. et al. Pathways of Lignocellulosic Biomass Conversion to Renewable Fuels. **Biomass Conversion and Biorefinery**, v. 4, n. 2, p. 157–91, 26 Jun. 2014.

(3) BROWN, T. R.; BROWN, R. C. A Review of Cellulosic Biofuel Commercial-Scale Projects in the United States. **Biofuels, Bioproducts and Biorefining**, v. 7, n. 3, p. 235–45, May 2013.

(4) MENON, V.; RAO, M. Trends in Bioconversion of Lignocellulose: Biofuels, Platform Chemicals & Biorefinery Concept. **Progress in Energy and Combustion Science**, v. 38, n. 4, p. 522–50, Aug. 2012.

(5) BALAT, M. Production of Bioethanol from Lignocellulosic Materials via the Biochemical Pathway: A Review. **Energy Conversion and Management**, v. 52, n. 2, p. 858–75, Feb. 2011.

(6) GARVEY, M. et al. Cellulases for Biomass Degradation: Comparing Recombinant Cellulase Expression Platforms. **Trends in Biotechnology**, v. 31, n. 10, p. 581–93, Oct. 2013.

(7) LIU, G.; ZHANG, J.; BAO, J. Cost Evaluation of Cellulase Enzyme for Industrial-Scale Cellulosic Ethanol Production Based on Rigorous Aspen Plus Modeling. **Bioprocess and Biosystems Engineering**, v. 39, n. 1, p. 133–40, 5 Jan. 2016.

(8) ELLILÄ, S. et al. Development of a Low-Cost Cellulase Production Process Using Trichoderma Reesei for Brazilian Biorefineries. **Biotechnology for Biofuels**, v. 10, n. 1, p. 30, 2017.

[††] In accordance with the Brazilian Association of Technical Standards (ABNT NBR 6023).

(9) HUMBIRD, D. et al. **Process Design and Economics for Biochemical Conversion of Lignocellulosic Biomass to Ethanol**. Golden, CO, USA: National Renewable Energy Laboratory, May 2011. 136 pp. Technical Report, NREL/TP-5100-47764.

(10) KLEIN-MARCUSCHAMER, D. et al. The Challenge of Enzyme Cost in the Production of Lignocellulosic Biofuels. **Biotechnology and Bioengineering**, v. 109, n. 4, p. 1083–7, Apr. 2012.

(11) TUSÉ, D.; TU, T.; MCDONALD, K. A. Manufacturing Economics of Plant-Made Biologics: Case Studies in Therapeutic and Industrial Enzymes. **BioMed Research International**, v. 2014, p. 1–16, 2014.

(12) SCHUSTER, A.; SCHMOLL, M. Biology and Biotechnology of Trichoderma. **Applied Microbiology and Biotechnology**, v. 87, n. 3, p. 787–99, Jul. 2010.

(13) BISCHOF, R. H. et al. Cellulases and beyond: The First 70 Years of the Enzyme Producer Trichoderma Reesei. **Microbial Cell Factories**, v. 15, n. 1, p. 106, 10 Dec. 2016.

(14) NOVY, V. et al. The Micromorphology of Trichoderma Reesei Analyzed in Cultivations on Lactose and Solid Lignocellulosic Substrate, and Its Relationship with Cellulase Production. **Biotechnology for Biofuels**, v. 9, n. 1, p. 169, 9 Dec. 2016.

(15) MARTINEZ, D. et al. Genome Sequencing and Analysis of the Biomass-Degrading Fungus Trichoderma Reesei (Syn. Hypocrea Jecorina). **Nature Biotechnology**, v. 26, n. 5, p. 553–60, May 2008.

(16) SINGHANIA, R. R. et al. Role and Significance of Beta-Glucosidases in the Hydrolysis of Cellulose for Bioethanol Production. **Bioresource Technology**, v. 127, p. 500–7, Jan. 2013.

(17) PRYOR, S. W.; NAHAR, N. β-Glucosidase Supplementation during Biomass Hydrolysis: How Low Can We Go? **Biomass and Bioenergy**, v. 80, p. 298–302, 2015.

(18) GUSAKOV, A. V. et al. Design of Highly Efficient Cellulase Mixtures for Enzymatic Hydrolysis of Cellulose. **Biotechnology and Bioengineering**, v. 97, n. 5, p. 1028–38, 1 Aug. 2007.

(19) BUSSAMRA, B. C.; FREITAS, S.; COSTA, A. C. Da. Improvement on Sugar Cane Bagasse Hydrolysis Using Enzymatic Mixture Designed Cocktail. **Bioresource Technology**, v. 187, p. 173–81, Jul. 2015.

(20) JOHNSON, E. Integrated Enzyme Production Lowers the Cost of Cellulosic Ethanol. **Biofuels, Bioproducts and Biorefining**, v. 10, n. 2, p. 164–74, Mar. 2016.

(21) ADRIO, J. L.; DEMAIN, A. L. Microbial Enzymes: Tools for Biotechnological Processes. **Biomolecules**, v. 4, n. 1, p. 117–39, 16 Jan. 2014.

(22) NIELSEN, P. H.; OXENBØLL, K. M.; WENZEL, H. Cradle-to-Gate Environmental Assessment of Enzyme Products Produced Industrially in Denmark by Novozymes A/S. **The International Journal of Life Cycle Assessment**, v. 12, n. 6, p. 432–8, 12 Sep. 2007.

(23) CHEN, R. Bacterial Expression Systems for Recombinant Protein Production: E. Coli and Beyond. **Biotechnology Advances**, v. 30, n. 5, p. 1102–7, Sep. 2012.

(24) DEMAIN, A. L.; VAISHNAV, P. Production of Recombinant Proteins by Microbes and Higher Organisms. **Biotechnology Advances**, v. 27, n. 3, p. 297–306, May 2009.

(25) WAEGEMAN, H.; SOETAERT, W. Increasing Recombinant Protein Production in Escherichia Coli through Metabolic and Genetic Engineering. **Journal of industrial microbiology & biotechnology**, v. 38, n. 12, p. 1891–910, 8 Dec. 2011.

(26) CHOI, J. H.; KEUM, K. C.; LEE, S. Y. Production of Recombinant Proteins by High Cell Density Culture of Escherichia Coli. **Chemical Engineering Science**, v. 61, n. 3, p. 876–85, 2006.

(27) TRIPATHI, N. K. Production and Purification of Recombinant Proteins from Escherichia Coli. **ChemBioEng Reviews**, v. 3, n. 3, p. 116–33, Jun. 2016.

(28) TERPE, K. Overview of Bacterial Expression Systems for Heterologous Protein Production: From Molecular and Biochemical Fundamentals to Commercial Systems. **Applied Microbiology and Biotechnology**, v. 72, n. 2, p. 211–22, 22 Sep. 2006.

(29) GUPTA, S. K.; SHUKLA, P. Advanced Technologies for Improved Expression of Recombinant Proteins in Bacteria: Perspectives and Applications. **Critical Reviews in**

Biotechnology, v. 36, n. 6, p. 1089–98, 18 Nov. 2016.

(30) SILVA, M. R. et al. High Cell Density Co-Culture for Production of Recombinant Hydrolases. **Biochemical Engineering Journal**, v. 71, p. 38–46, Feb. 2013.

(31) CARNEIRO, S.; FERREIRA, E. C.; ROCHA, I. Metabolic Responses to Recombinant Bioprocesses in Escherichia Coli. **Journal of Biotechnology**, v. 164, n. 3, p. 396–408, 10 Apr. 2013.

(32) ROSANO, G. L.; CECCARELLI, E. A. Recombinant Protein Expression in Escherichia Coli: Advances and Challenges. **Frontiers in Microbiology**, v. 5, p. 172, 17 Apr. 2014.

(33) HUANG, C.-J.; LIN, H.; YANG, X. Industrial Production of Recombinant Therapeutics in Escherichia Coli and Its Recent Advancements. **Journal of Industrial Microbiology & Biotechnology**, v. 39, n. 3, p. 383–99, 18 Mar. 2012.

(34) WYRE, C.; OVERTON, T. W. Use of a Stress-Minimisation Paradigm in High Cell Density Fed-Batch Escherichia Coli Fermentations to Optimise Recombinant Protein Production. **Journal of Industrial Microbiology and Biotechnology**, v. 41, n. 9, p. 1391–404, 24 Sep. 2014.

(35) VAN MELDEREN, L. Toxin-Antitoxin Systems: Why so Many, What For? **Current Opinion in Microbiology**, v. 13, n. 6, p. 781–5, Dec. 2010.

(36) VANDERMEULEN, G. et al. New Generation of Plasmid Backbones Devoid of Antibiotic Resistance Marker for Gene Therapy Trials. **Molecular therapy : the journal of the American Society of Gene Therapy**, v. 19, n. 11, p. 1942–9, Nov. 2011.

(37) PEUBEZ, I. et al. Antibiotic-Free Selection in E. coli: New Considerations for Optimal Design and Improved Production. **Microbial Cell Factories**, v. 9, n. 1, p. 65, Jan. 2010.

(38) OLIVEIRA, P. H.; MAIRHOFER, J. Marker-Free Plasmids for Biotechnological Applications - Implications and Perspectives. **Trends in Biotechnology**, v. 31, n. 9, p. 539–47, Sep. 2013.

(39) SINGHA, T. K. et al. Efficient Genetic Approaches for Improvement of Plasmid Based

Expression of Recombinant Protein in Escherichia Coli : A Review. **Process Biochemistry**, v. 55, p. 17–31, Apr. 2017. .

(40) GUPTA, V. K. et al. Fungal Enzymes for Bio-Products from Sustainable and Waste Biomass. **Trends in Biochemical Sciences**, v. 41, n. 7, p. 633–45, Jul. 2016.

(41) GUPTA, P.; PARKHEY, P. A Two-Step Process for Efficient Enzymatic Saccharification of Rice Straw. **Bioresource Technology**, v. 173, p. 207–15, Dec. 2014.

(42) LIAO, J. C. et al. Fuelling the Future: Microbial Engineering for the Production of Sustainable Biofuels. **Nature Reviews Microbiology**, v. 14, n. 5, p. 288–304, 30 May 2016.

(43) KUBICEK, C. P.; KUBICEK, E. M. Enzymatic Deconstruction of Plant Biomass by Fungal Enzymes. **Current Opinion in Chemical Biology**, v. 35, p. 51–7, Dec. 2016.

(44) HASUNUMA, T. et al. A Review of Enzymes and Microbes for Lignocellulosic Biorefinery and the Possibility of Their Application to Consolidated Bioprocessing Technology. **Bioresource Technology**, v. 135, p. 513–22, May 2013.

(45) LIMAYEM, A.; RICKE, S. C. Lignocellulosic Biomass for Bioethanol Production: Current Perspectives, Potential Issues and Future Prospects. **Progress in Energy and Combustion Science**, v. 38, n. 4, p. 449–67, Aug. 2012.

(46) WYMAN, C. et al. Hydrolysis of Cellulose and Hemicellulose. In: Dumitriu, S. (Ed.). **Polysaccharides: Structural Diversity amd Functional Versatility**. New York, NY, USA: Marcel Dekker, 2005. ch. 43, p. 1023–62.

(47) CARDONA, C. a; QUINTERO, J. a; PAZ, I. C. Production of Bioethanol from Sugarcane Bagasse: Status and Perspectives. **Bioresource Technology**, v. 101, n. 13, p. 4754–66, Jul. 2010.

(48) GÍRIO, F. M. et al. Hemicelluloses for Fuel Ethanol: A Review. **Bioresource Technology**, v. 101, n. 13, p. 4775–800, Jul. 2010.

(49) HAGHIGHI MOOD, S. et al. Lignocellulosic Biomass to Bioethanol, a Comprehensive Review with a Focus on Pretreatment. **Renewable and Sustainable Energy Reviews**, v.

27, p. 77–93, Nov. 2013.

(50) RAHIKAINEN, J. L. et al. Inhibitory Effect of Lignin during Cellulose Bioconversion: The Effect of Lignin Chemistry on Non-Productive Enzyme Adsorption. **Bioresource Technology**, v. 133, p. 270–8, Apr. 2013.

(51) VAN DYK, J. S.; PLETSCHKE, B. I. A Review of Lignocellulose Bioconversion Using Enzymatic Hydrolysis and Synergistic Cooperation between Enzymes--Factors Affecting Enzymes, Conversion and Synergy. **Biotechnology Advances**, v. 30, n. 6, p. 1458–80, Nov. 2012.

(52) RUBIN, E. M. Genomics of Cellulosic Biofuels. **Nature**, v. 454, n. 7206, p. 841–5, 14 Aug. 2008.

(53) HO, D. P.; NGO, H. H.; GUO, W. A Mini Review on Renewable Sources for Biofuel. **Bioresource Technology**, v. 169, p. 742–9, 2014.

(54) MAITY, S. K. Opportunities, Recent Trends, and Challenges of Integrated Biorefinery: Part I. **Renewable and Sustainable Energy Reviews**, v. 43, p. 1446-66, Mar. 2015.

(55) SOCCOL, C. R. et al. Bioethanol from Lignocelluloses: Status and Perspectives in Brazil. **Bioresource Technology**, v. 101, n. 13, p. 4820–5, Jul. 2010.

(56) WILSON, D. B. Microbial Diversity of Cellulose Hydrolysis. **Current Opinion in Microbiology**, v. 14, n. 3, p. 259–63, Jun. 2011.

(57) BLANCH, H. W.; SIMMONS, B. A.; KLEIN-MARCUSCHAMER, D. Biomass Deconstruction to Sugars. **Biotechnology Journal**, v. 6, n. 9, p. 1086–102, Sep. 2011.

(58) SIMS, R. E. H. et al. An Overview of Second Generation Biofuel Technologies. **Bioresource Technology**, v. 101, n. 6, p. 1570–80, Mar. 2010.

(59) CHOVAU, S.; DEGRAUWE, D.; VAN DER BRUGGEN, B. Critical Analysis of Techno-Economic Estimates for the Production Cost of Lignocellulosic Bio-Ethanol. **Renewable and Sustainable Energy Reviews**, v. 26, p. 307–21, Oct. 2013.

(60) DEN HAAN, R. et al. Progress and Challenges in the Engineering of Non-Cellulolytic

Microorganisms for Consolidated Bioprocessing. **Current Opinion in Biotechnology**, v. 33, p. 32–8, Jun. 2015.

(61) SILVEIRA, M. H. L. et al. Current Pretreatment Technologies for the Development of Cellulosic Ethanol and Biorefineries. **ChemSusChem**, v. 8, n. 20, p. 3366–90, Oct. 2015.

(62) PAULOVA, L. et al. Lignocellulosic Ethanol: Technology Design and Its Impact on Process Efficiency. **Biotechnology Advances**, v. 33, n. 6, p. 1091–107, Nov. 2015.

(63) SINDHU, R.; PANDEY, A. Biological Pretreatment of Lignocellulosic Biomass – An Overview. **Bioresource Technology**, v. 199, p. 76–82, Jan. 2016.

(64) CHANDEL, A. K. et al. Biodelignification of Lignocellulose Substrates: An Intrinsic and Sustainable Pretreatment Strategy for Clean Energy Production. **Critical Reviews in Biotechnology**, v. 35, n. 3, p. 281–93, 3 Jul. 2015.

(65) LAURICHESSE, S.; AVÉROUS, L. Chemical Modification of Lignins: Towards Biobased Polymers. **Progress in Polymer Science**, v. 39, n. 7, p. 1266–90, Jul. 2014.

(66) RAGAUSKAS, A. J. et al. Lignin Valorization: Improving Lignin Processing in the Biorefinery. **Science**, v. 344, n. 6185, p. 1246843.1–10, 16 May 2014.

(67) DAVIS, R. et al. **Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbon Fuels and Coproducts: 2018 Biochemical Design Case Update: Biochemical Deconstruction and Conversion of Biomass to Fuels and Products via Integrated Biorefinery Path**. Golden, CO, USA: National Renewable Energy Laboratory, Nov. 2018. 137 pp. Technical Report, NREL/TP-5100-71949.

(68) CHEN, Y.; NIELSEN, J. Advances in Metabolic Pathway and Strain Engineering Paving the Way for Sustainable Production of Chemical Building Blocks. **Current Opinion in Biotechnology**, v. 24, n. 6, p. 965–72, Dec. 2013.

(69) YANG, B. et al. Enzymatic Hydrolysis of Cellulosic Biomass. **Biofuels**, v. 2, n. 4, p. 421–49, 9 Jul. 2011.

(70) ZHAO, X.; ZHANG, L.; LIU, D. Biomass Recalcitrance. Part II: Fundamentals of

Different Pre-Treatments to Increase the Enzymatic Digestibility of Lignocellulose. **Biofuels, Bioproducts and Biorefining**, v. 6, n. 5, p. 561–79, Sep. 2012.

(71) JÄGER, G.; BÜCHS, J. Biocatalytic Conversion of Lignocellulose to Platform Chemicals. **Biotechnology Journal**, v. 7, n. 9, p. 1122–36, 1 Sep. 2012.

(72) ALFENORE, S.; MOLINA-JOUVE, C. Current Status and Future Prospects of Conversion of Lignocellulosic Resources to Biofuels Using Yeasts and Bacteria. **Process Biochemistry**, v. 51, n. 11, p. 1747–56, Nov. 2016.

(73) BEHERA, S. et al. Importance of Chemical Pretreatment for Bioconversion of Lignocellulosic Biomass. **Renewable and Sustainable Energy Reviews**, v. 36, p. 91–106, Aug. 2014.

(74) ALVIRA, P. et al. Pretreatment Technologies for an Efficient Bioethanol Production Process Based on Enzymatic Hydrolysis: A Review. **Bioresource Technology**, v. 101, n. 13, p. 4851–61, Jul. 2010.

(75) ZABED, H. et al. Fuel Ethanol Production from Lignocellulosic Biomass: An Overview on Feedstocks and Technological Approaches. **Renewable and Sustainable Energy Reviews**, v. 66, p. 751–74, Dec. 2016.

(76) SUN, S. et al. The Role of Pretreatment in Improving the Enzymatic Hydrolysis of Lignocellulosic Materials. **Bioresource Technology**, v. 199, p. 49–58, Jan. 2016.

(77) BALAT, M.; BALAT, H. Recent Trends in Global Production and Utilization of Bio-Ethanol Fuel. **Applied Energy**, v. 86, n. 11, p. 2273–82, Nov. 2009.

(78) SHIRKAVAND, E. et al. Combination of Fungal and Physicochemical Processes for Lignocellulosic Biomass Pretreatment – A Review. **Renewable and Sustainable Energy Reviews**, v. 54, p. 217–34, Feb. 2016.

(79) CONDE-MEJÍA, C.; JIMÉNEZ-GUTIÉRREZ, A.; EL-HALWAGI, M. A Comparison of Pretreatment Methods for Bioethanol Production from Lignocellulosic Materials. **Process Safety and Environmental Protection**, v. 90, n. 3, p. 189–202, May 2012.

(80) JÖNSSON, L. J.; ALRIKSSON, B.; NILVEBRANT, N.-O. Bioconversion of

Lignocellulose: Inhibitors and Detoxification. **Biotechnology for Biofuels**, v. 6, n. 1, p. 16, 2013.

(81) NIGAM, P. S.; SINGH, A. Production of Liquid Biofuels from Renewable Resources. **Progress in Energy and Combustion Science**, v. 37, n. 1, p. 52–68, Feb. 2011.

(82) JÖNSSON, L. J.; MARTÍN, C. Pretreatment of Lignocellulose: Formation of Inhibitory by-Products and Strategies for Minimizing Their Effects. **Bioresource Technology**, v. 199, p. 103–12, 2016.

(83) CAPOLUPO, L.; FARACO, V. Green Methods of Lignocellulose Pretreatment for Biorefinery Development. **Applied Microbiology and Biotechnology**, v. 100, n. 22, p. 9451–67, 6 Nov. 2016.

(84) WAHLSTROM, R. M.; SUURNAKKI, A. Enzymatic Hydrolysis of Lignocellulosic Polysaccharides in the Presence of Ionic Liquids. **Green Chemistry**, v. 17, n. 2, p. 694–714, 2015.

(85) LEVINE, S. E. et al. A Mechanistic Model for Rational Design of Optimal Cellulase Mixtures. **Biotechnology and Bioengineering**, v. 108, n. 11, p. 2561–70, Nov. 2011.

(86) VIIKARI, L.; VEHMAANPERÄ, J.; KOIVULA, a. Lignocellulosic Ethanol: From Science to Industry. **Biomass and Bioenergy**, v. 46, p. 13–24, Nov. 2012.

(87) KOPPRAM, R. et al. Lignocellulosic Ethanol Production at High-Gravity: Challenges and Perspectives. **Trends in Biotechnology**, v. 32, n. 1, p. 46–53, Jan. 2014.

(88) RASMUSSEN, H.; SØRENSEN, H. R.; MEYER, A. S. Formation of Degradation Compounds from Lignocellulosic Biomass in the Biorefinery: Sugar Reaction Mechanisms. **Carbohydrate Research**, v. 385, p. 45–57, Feb. 2014.

(89) SEIDL, P. R.; GOULART, A. K. Pretreatment Processes for Lignocellulosic Biomass Conversion to Biofuels and Bioproducts. **Current Opinion in Green and Sustainable Chemistry**, v. 2, p. 48–53, Oct. 2016.

(90) MORENO, A. D. et al. A Review of Biological Delignification and Detoxification Methods for Lignocellulosic Bioethanol Production. **Critical Reviews in Biotechnology**,

v. 35, n. 3, p. 342–54, 3 Jul. 2015.

(91) BRAGA, C. M. P. et al. Addition of Feruloyl Esterase and Xylanase Produced On-Site Improves Sugarcane Bagasse Hydrolysis. **Bioresource Technology**, v. 170, p. 316–24, Oct. 2014.

(92) MASRAN, R. et al. Harnessing the Potential of Ligninolytic Enzymes for Lignocellulosic Biomass Pretreatment. **Applied Microbiology and Biotechnology**, v. 100, n. 12, p. 5231–46, 26 Jun. 2016.

(93) CHEN, S. et al. Biological Pretreatment of Lignocellulosics: Potential, Progress and Challenges. **Biofuels**, v. 1, n. 1, p. 177–99, 9 Jan. 2010.

(94) BUGG, T. D. H. et al. The Emerging Role for Bacteria in Lignin Degradation and Bio-Product Formation. **Current opinion in biotechnology**, v. 22, n. 3, p. 394–400, Jun. 2011.

(95) LIGUORI, R.; FARACO, V. Biological Processes for Advancing Lignocellulosic Waste Biorefinery by Advocating Circular Economy. **Bioresource Technology**, v. 215, p. 13–20, Sep. 2016.

(96) HARRIS, P. V. et al. New Enzyme Insights Drive Advances in Commercial Ethanol Production. **Current Opinion in Chemical Biology**, v. 19, n. 1, p. 162–70, Apr. 2014.

(97) BORNSCHEUER, U.; BUCHHOLZ, K.; SEIBEL, J. Enzymatic Degradation of (Ligno)Cellulose. **Angewandte Chemie International Edition**, v. 53, n. 41, p. 10876–93, 6 Oct. 2014.

(98) JALAK, J. et al. Endo-Exo Synergism in Cellulose Hydrolysis Revisited. **The Journal of Biological Chemistry**, v. 287, n. 34, p. 28802–15, 17 Aug. 2012.

(99) MURPHY, L. et al. Product Inhibition of Five Hypocrea Jecorina Cellulases. **Enzyme and Microbial Technology**, v. 52, n. 3, p. 163–9, 5 Mar. 2013.

(100) CARBOHYDRATE ACTIVE ENZYMES DATABASE. Available at: <http://www.cazy.org>. Accessed 15 Dec. 2018.

(101) HEMSWORTH, G. R. et al. Lytic Polysaccharide Monooxygenases in Biomass

Conversion. **Trends in Biotechnology**, v. 33, n. 12, p. 747–61, 2015.

(102) DUTTA, S.; WU, K. C.-W. Enzymatic Breakdown of Biomass: Enzyme Active Sites, Immobilization, and Biofuel Production. **Green Chemistry**., v. 16, n. 11, p. 4615–26, 26 Aug. 2014.

(103) BHATIA, Y.; MISHRA, S.; BISARIA, V. S. Microbial β-Glucosidases: Cloning, Properties, and Applications. **Critical Reviews in Biotechnology**, v. 22, n. 4, p. 375–407, 29 Jan. 2002.

(104) KETUDAT CAIRNS, J. R.; ESEN, A. β-Glucosidases. **Cellular and Molecular Life Sciences**, v. 67, n. 20, p. 3389–405, 20 Oct. 2010.

(105) HENRISSAT, B.; DAVIES, G. Structural and Sequence-Based Classification of Glycoside Hydrolases. **Current Opinion in Structural Biology**, v. 7, n. 5, p. 637–44, Oct. 1997.

(106) TEUGJAS, H.; VÄLJAMÄE, P. Selecting β-Glucosidases to Support Cellulases in Cellulose Saccharification. **Biotechnology for Biofuels**, v. 6, n. 1, p. 105, 24 Jul. 2013.

(107) SALGADO, J. C. S. et al. Glucose Tolerant and Glucose Stimulated β-Glucosidases – A Review. **Bioresource Technology**, v. 267, p. 704–13, 1 Nov. 2018.

(108) KUUSK, S.; VÄLJAMÄE, P. When Substrate Inhibits and Inhibitor Activates: Implications of β-Glucosidases. **Biotechnology for Biofuels**, v. 10, n. 1, p. 7, 3 Dec. 2017.

(109) ELLEUCHE, S. et al. Extremozymes-Biocatalysts with Unique Properties from Extremophilic Microorganisms. **Current Opinion in Biotechnology**, v. 29, p. 116–23, Oct. 2014.

(110) KRÜGER, A. et al. Towards a Sustainable Biobased Industry – Highlighting the Impact of Extremophiles. **New Biotechnology**, v. 40, p. 144–53, Jan. 2018.

(111) AKRAM, F.; HAQ, I. ul; MUKHTAR, H. Gene Cloning, Characterization and Thermodynamic Analysis of a Novel Multidomain Hyperthermophilic GH Family 3 β-Glucosidase (TnBglB) from Thermotoga Naphthophila RKU-10T. **Process Biochemistry**, v. 66, p. 70–81, 1 Mar. 2018.

(112) AKRAM, F. et al. Cloning with Kinetic and Thermodynamic Insight of a Novel Hyperthermostable β-Glucosidase from Thermotoga Naphthophila RKU-10T with Excellent Glucose Tolerance. **Journal of Molecular Catalysis B: Enzymatic**, v. 124, p. 92–104, 1 Feb. 2016.

(113) HAQ, I. U. et al. Cloning, Characterization and Molecular Docking of a Highly Thermostable β-1,4-Glucosidase from Thermotoga Petrophila. **Biotechnology Letters**, v. 34, n. 9, p. 1703–09, 20 Sep. 2012.

(114) COTA, J. et al. Comparative Analysis of Three Hyperthermophilic GH1 and GH3 Family Members with Industrial Potential. **New Biotechnology**, v. 32, n. 1, p. 13–20, Jan. 2015.

(115) SCHRÖDER, C. et al. Characterization of a Heat-Active Archaeal β-Glucosidase from a Hydrothermal Spring Metagenome. **Enzyme and Microbial Technology**, v. 57, p. 48–54, Apr. 2014.

(116) PARK, A.-R. et al. Biochemical Characterization of an Extracellular β-Glucosidase from the Fungus, Penicillium Italicum, Isolated from Rotten Citrus Peel. **Mycobiology**, v. 40, n. 3, p. 173–80, 19 Sep. 2012.

(117) FUSCO, F. A. et al. Biochemical Characterization of a Novel Thermostable β-Glucosidase from Dictyoglomus Turgidum. **International Journal of Biological Macromolecules**, v. 113, p. 783–91, 2018.

(118) MÉNDEZ-LÍTER, J. A. et al. A Novel, Highly Efficient β-Glucosidase with a Cellulose-Binding Domain: Characterization and Properties of Native and Recombinant Proteins. **Biotechnology for Biofuels**, v. 10, n. 1, p. 256, 6 Dec. 2017.

(119) XIA, W. et al. Engineering a Highly Active Thermophilic β-Glucosidase to Enhance Its PH Stability and Saccharification Performance. **Biotechnology for Biofuels**, v. 9, n. 1, p. 147, 20 Dec. 2016.

(120) HARNPICHARNCHAI, P. et al. A Thermotolerant β-Glucosidase Isolated from an Endophytic Fungi, Periconia Sp., with a Possible Use for Biomass Conversion to Sugars. **Protein Expression and Purification**, v. 67, n. 2, p. 61–9, 1 Oct. 2009.

(121) YANG, X. et al. Molecular Characterization of a Highly-Active Thermophilic β-Glucosidase from Neosartorya Fischeri P1 and Its Application in the Hydrolysis of Soybean Isoflavone Glycosides. **PLoS ONE**, v. 9, n. 9, p. e106785, 4 Sep. 2014.

(122) KRISCH, J. et al. Characterization of a β-Glucosidase with Transgalactosylation Capacity from the Zygomycete Rhizomucor Miehei. **Bioresource Technology**, v. 114, p. 555–60, 1 Jun. 2012.

(123) RAMANI, G. et al. Molecular Cloning and Expression of Thermostable Glucose-Tolerant β-Glucosidase of Penicillium Funiculosum NCL1 in Pichia Pastoris and Its Characterization. **Journal of Industrial Microbiology & Biotechnology**, v. 42, n. 4, p. 553–65, 28 Apr. 2015.

(124) KA, L. et al. The Purification and Characterization of an Extremely Thermostable Alpha- Amylase from the Hyperthermophilic Archaebacterium Pyrococcus Furiosus. **Journal of Biological Chemistry**, v. 268, n. 32, p. 24394–401, 1993.

(125) MELEIRO, L. P. et al. A Novel β-Glucosidase from Humicola Insolens with High Potential for Untreated Waste Paper Conversion to Sugars. **Applied Biochemistry and Biotechnology**, v. 173, n. 2, p. 391–408, 14 May 2014.

(126) CHAN, C. S. et al. Characterization of a Glucose-Tolerant β-Glucosidase from Anoxybacillus Sp. DT3-1. **Biotechnology for Biofuels**, v. 9, n. 1, p. 174, 22 Dec. 2016.

(127) PAINBENI, E. et al. Purification and Characterization of a Bacillus Polymyxa β-Glucosidase Expressed in Escherichia Coli. **Journal of Bacteriology**, v. 174, n. 9, p. 3087–91, 1992.

(128) PENG, X. et al. A Multifunctional Thermophilic Glycoside Hydrolase from Caldicellulosiruptor Owensensis with Potential Applications in Production of Biofuels and Biochemicals. **Biotechnology for Biofuels**, v. 9, n. 1, p. 98, 30 Dec. 2016.

(129) PARRY, N. J. et al. Biochemical Characterization and Mechanism of Action of a Thermostable β-Glucosidase Purified from Thermoascus Aurantiacus. **Biochemical Journal**, v. 353, n. 1, p. 117–27, 2001.

(130) POZZO, T. et al. Structural and Functional Analyses of β-Glucosidase 3B from

Thermotoga Neapolitana: A Thermostable Three-Domain Representative of Glycoside Hydrolase 3. **Journal of Molecular Biology**, v. 397, n. 3, p. 724–39, 2 Apr. 2010.

(131) NGUYEN, N.-P.-T. et al. One-Step Purification and Characterization of a β-1,4-Glucosidase from a Newly Isolated Strain of Stereum Hirsutum. **Applied Microbiology and Biotechnology**, v. 87, n. 6, p. 2107–16, 8 Aug. 2010.

(132) JOO, A.-R. et al. Purification and Characterization of a β-1,4-Glucosidase from a Newly Isolated Strain of Fomitopsis Pinicola. **Applied Microbiology and Biotechnology**, v. 83, n. 2, p. 285–94, 21 May 2009.

(133) DEOG, K. J.; JIN, S. H.; KIYOSHI, H. Construction and Characterization of Novel Chimeric β-Glucosidases with Cellvibrio Gilvus (CG) and Thermotoga Maritima (TM) by Overlapping PCR. **Biotechnology and Bioprocess Engineering**, v. 14, n. 3, p. 266–73, 1 Aug. 2009.

(134) LANGSTON, J.; SHEEHY, N.; XU, F. Substrate Specificity of Aspergillus Oryzae Family 3 β-Glucosidase. **Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics**, v. 1764, n. 5, p. 972–8, 1 May 2006.

(135) BRENDA**. Information on EC 3.2.1.21 - beta-glucosidase**. Available at: <https://www.brenda-enzymes.org/enzyme.php?ecno=3.2.1.21&Suchword=&reference=&UniProtAcc=&organism%5B%5D=&show_tm=0>. Accessed 4 Jan. 2019.

(136) YAN, T.-R.; LIAU, J.-C. Synthesis of Alkyl β-Glucosides from Cellobiose with Aspergillus Niger β-Glucosidase II. **Biotechnology Letters**, v. 20, n. 7, p. 653–7, 1998.

(137) GAO, G. et al. A Novel Metagenome-Derived Gene Cluster from Termite Hindgut: Encoding Phosphotransferase System Components and High Glucose Tolerant Glucosidase. **Enzyme and Microbial Technology**, v. 84, p. 24–31, 1 Mar. 2016.

(138) BHALLA, T. C.; ASIF, M.; SMITA, K. Purification and Characterization of Cyanogenic β-Glucosidase from Wild Apricot (Prunus Armeniaca L.). **Process Biochemistry**, v. 58, p. 320–5, 1 Jul. 2017.

(139) BOHLIN, C. et al. A Comparative Study of Activity and Apparent Inhibition of

Fungal β-Glucosidases. **Biotechnology and Bioengineering**, v. 107, n. 6, p. 943–52, 15 Dec. 2010.

(140) DODD, D.; CANN, I. K. O. Enzymatic Deconstruction of Xylan for Biofuel Production. **GCB Bioenergy**, v. 1, n. 1, p. 2–17, 18 Feb. 2009.

(141) JUTURU, V.; WU, J. C. Microbial Xylanases: Engineering, Production and Industrial Applications. **Biotechnology Advances**, v. 30, n. 6, p. 1219–27, Nov. 2012.

(142) LIU, G. et al. Development of Highly Efficient, Low-Cost Lignocellulolytic Enzyme Systems in the Post-Genomic Era. **Biotechnology Advances**, v. 31, n. 6, p. 962–75, Nov. 2013.

(143) RABEMANOLONTSOA, H.; SAKA, S. Various Pretreatments of Lignocellulosics. **Bioresource Technology**, v. 199, p. 83–91, Jan. 2016.

(144) WENG, J.-K. et al. Emerging Strategies of Lignin Engineering and Degradation for Cellulosic Biofuel Production. **Current Opinion in Biotechnology**, v. 19, n. 2, p. 166–72, Apr. 2008.

(145) BUGG, T. D. H. et al. Pathways for Degradation of Lignin in Bacteria and Fungi. **Natural Product Reports**, v. 28, n. 12, p. 1883–96, 2011.

(146) BUGG, T. D.; RAHMANPOUR, R. Enzymatic Conversion of Lignin into Renewable Chemicals. **Current Opinion in Chemical Biology**, v. 29, p. 10–17, Dec. 2015.

(147) CAMARERO, S.; MARTÍNEZ, M. J.; MARTÍNEZ, A. T. Understanding Lignin Biodegradation for the Improved Utilization of Plant Biomass in Modern Biorefineries. **Biofuels, Bioproducts and Biorefining**, v. 8, n. 5, p. 615–25, Sep. 2014.

(148) PARISUTHAM, V.; KIM, T. H.; LEE, S. K. Feasibilities of Consolidated Bioprocessing Microbes: From Pretreatment to Biofuel Production. **Bioresource Technology**, v. 161, p. 431–40, Jun. 2014.

(149) COSGROVE, D. J. Growth of the Plant Cell Wall. **Nature Reviews Molecular Cell Biology**, v. 6, n. 11, p. 850–61, Nov. 2005.

(150) GOURLAY, K. et al. Swollenin Aids in the Amorphogenesis Step during the

Enzymatic Hydrolysis of Pretreated Biomass. **Bioresource Technology**, v. 142, p. 498–503, Aug. 2013.

(151) KIM, E. S. et al. Functional Characterization of a Bacterial Expansin from Bacillus Subtilis for Enhanced Enzymatic Hydrolysis of Cellulose. **Biotechnology and Bioengineering**, v. 102, n. 5, p. 1342–53, 1 Apr. 2009.

(152) MENG, X.; RAGAUSKAS, A. J. Recent Advances in Understanding the Role of Cellulose Accessibility in Enzymatic Hydrolysis of Lignocellulosic Substrates. **Current Opinion in Biotechnology**, v. 27, p. 150–8, Jun. 2014.

(153) QUIROZ-CASTAÑEDA, R. E. et al. Loosenin, a Novel Protein with Cellulose-Disrupting Activity from Bjerkandera Adusta. **Microbial Cell Factories**, v. 10, n. 1, p. 8, 2011.

(154) KIM, S.; JIMÉNEZ-GONZÁLEZ, C.; DALE, B. E. Enzymes for Pharmaceutical Applications—a Cradle-to-Gate Life Cycle Assessment. **The International Journal of Life Cycle Assessment**, v. 14, n. 5, p. 392–400, 29 Jul. 2009.

(155) GUSAKOV, A. V. Alternatives to Trichoderma Reesei in Biofuel Production. **Trends in biotechnology**, v. 29, n. 9, p. 419–25, Sep. 2011.

(156) PETERSON, R.; NEVALAINEN, H. Trichoderma Reesei RUT-C30 - Thirty Years of Strain Improvement. **Microbiology**, v. 158, n. 1, p. 58–68, 1 Jan. 2012.

(157) BLANCH, H. W. Bioprocessing for Biofuels. **Current Opinion in Biotechnology**, v. 23, n. 3, p. 390–5, Jun 2012.

(158) KLEIN-MARCUSCHAMER, D.; BLANCH, H. W. Renewable Fuels from Biomass: Technical Hurdles and Economic Assessment of Biological Routes. **AIChE Journal**, v. 61, n. 9, p. 2689–701, Sep. 2015.

(159) JØRGENSEN, H.; PINELO, M. Enzyme Recycling in Lignocellulosic Biorefineries. **Biofuels, Bioproducts and Biorefining**, v. 11, n. 1, p. 150–67, 1 Jan. 2017.

(160) ZHANG, M. et al. Enhanced Enzymatic Hydrolysis of Lignocellulose by Optimizing Enzyme Complexes. **Applied Biochemistry and Biotechnology**, v. 160, n. 5, p. 1407–

14, 14 Mar. 2010.

(161) HU, J. et al. The Addition of Accessory Enzymes Enhances the Hydrolytic Performance of Cellulase Enzymes at High Solid Loadings. **Bioresource Technology**, v. 186, p. 149–53, Jun. 2015.

(162) XIN, D. et al. Role of Hemicellulases in Production of Fermentable Sugars from Corn Stover. **Industrial Crops and Products**, v. 74, p. 209–17, Nov. 2015.

(163) RANA, V. et al. On-Site Enzymes Produced from Trichoderma Reesei RUT-C30 and Aspergillus Saccharolyticus for Hydrolysis of Wet Exploded Corn Stover and Loblolly Pine. **Bioresource Technology**, v. 154, p. 282–9, Feb. 2014.

(164) MA, L. et al. Improvement of Cellulase Activity in Trichoderma Reesei by Heterologous Expression of a Beta-Glucosidase Gene from Penicillium Decumbens. **Enzyme and Microbial Technology**, v. 49, n. 4, p. 366–71, 10 Sep. 2011.

(165) GUPTA, V. K. et al. The Post-Genomic Era of Trichoderma Reesei: What's Next? **Trends in Biotechnology**, v. 34, n. 12, p. 970–82, Dec. 2016. .

(166) KUBICEK, C. P. et al. Metabolic Engineering Strategies for the Improvement of Cellulase Production by Hypocrea Jecorina. **Biotechnology for Biofuels**, v. 2, p. 19, 2009.

(167) CHANDEL, A. K. et al. The Realm of Cellulases in Biorefinery Development. **Critical Reviews in Biotechnology**, v. 32, n. 3, p. 187–202, Sep. 2012.

(168) OLSON, D. G. et al. Recent Progress in Consolidated Bioprocessing. **Current Opinion in Biotechnology**, v. 23, n. 3, p. 396–405, Jun. 2012.

(169) GEFEN, G. et al. Enhanced Cellulose Degradation by Targeted Integration of a Cohesin-Fused β-Glucosidase into the Clostridium Thermocellum Cellulosome. **Proceedings of the National Academy of Sciences of the United States of America**, v. 109, n. 26, p. 10298–303, 26 Jun. 2012.

(170) SAINI, J. K.; SAINI, R.; TEWARI, L. Lignocellulosic Agriculture Wastes as Biomass Feedstocks for Second-Generation Bioethanol Production: Concepts and Recent

Developments. **3 Biotech**, v. 5, n. 4, p. 337–53, 21 Aug. 2015.

(171) KAWAGUCHI, H. et al. Bioprocessing of Bio-Based Chemicals Produced from Lignocellulosic Feedstocks. **Current Opinion in Biotechnology**, v. 42, p. 30–9, Dec. 2016.

(172) ASSOCIATION OF MANUFACTURERS AND FORMULATORS OFE ENZYME PRODUCTS (AMFEP). **List of Commercial Enzymes**: Update May 2015. 2015. Available at: <http://www.amfep.org/sites/default/files/201505/Amfep List of Enzymes update May 2015.pdf>. Accessed 8 Aug. 2017.

(173) MIDDELBERG, A. P. J. Process-Scale Disruption of Microorganisms. **Biotechnology Advances**, v. 13, n. 3, p. 491–551, Jan. 1995.

(174) MEYER, H.-P.; SCHMIDHALTER R., D. Microbial Expression Systems and Manufacturing from a Market and Economic Perspective. In: AGBO, E. C. (Ed.). **Innovations in Biotechnology**. [s.l.] InTech, 2012. ch. 10, p. 211–50.

(175) EITEMAN, M. A.; ALTMAN, E. Overcoming Acetate in Escherichia Coli Recombinant Protein Fermentations. **Trends in Biotechnology**, v. 24, n. 11, p. 530–6, Nov. 2006.

(176) SHILOACH, J.; FASS, R. Growing E. Coli to High Cell Density—A Historical Perspective on Method Development. **Biotechnology Advances**, v. 23, n. 5, p. 345–57, Jul. 2005.

(177) HERMANN, T. Industrial Production of Amino Acids by Coryneform Bacteria. **Journal of Biotechnology**, v. 104, n. 1–3, p. 155–72, Sep. 2003.

(178) LIU, M. et al. Metabolic Engineering of Escherichia Coli to Improve Recombinant Protein Production. **Applied Microbiology and Biotechnology**, v. 99, n. 24, p. 10367–77, 24 Dec. 2015.

(179) VAN DIEN, S. From the First Drop to the First Truckload: Commercialization of Microbial Processes for Renewable Chemicals. **Current Opinion in Biotechnology**, v. 24, n. 6, p. 1061–8, 2013.

(180) VOSS, I.; STEINBÜCHEL, A. Application of a KDPG-Aldolase Gene-Dependent Addiction System for Enhanced Production of Cyanophycin in Ralstonia Eutropha Strain H16. **Metabolic Engineering**, v. 8, n. 1, p. 66–78, 1 Jan. 2006.

(181) IKEDA, M. et al. Fermentative Production of Tryptophan by a Stable Recombinant Strain of Coryne-Bacterium Glutamicum with a Modified Serine-Biosynthetic Pathway. **Bioscience, Biotechnology and Biochemistry**, v. 58, n. 4, p. 674–8, 1994. .

(182) SQUIRES, C. H. et al. Heterologous Protein Production in P. Fluorescens. **BioProcess International**, v. December, p. 54–9, 2004.

(183) VIDAL, L. et al. Development of an Antibiotic-Free Plasmid Selection System Based on Glycine Auxotrophy for Recombinant Protein Overproduction in Escherichia Coli. **Journal of Biotechnology**, v. 134, n. 1–2, p. 127–36, 20 Mar. 2008.

(184) KROLL, J.; KLINTER, S.; STEINBÜCHEL, A. A Novel Plasmid Addiction System for Large-Scale Production of Cyanophycin in Escherichia Coli Using Mineral Salts Medium. **Applied Microbiology and Biotechnology**, v. 89, n. 3, p. 593–604, 5 Feb. 2011.

(185) YAMAGUCHI, Y.; PARK, J.-H.; INOUYE, M. Toxin-Antitoxin Systems in Bacteria and Archaea. **Annual Review of Genetics**, v. 45, p. 61–79, Jan. 2011.

(186) HAYES, F.; VAN MELDEREN, L. Toxins-Antitoxins: Diversity, Evolution and Function. **Critical reviews in biochemistry and molecular biology**, v. 46, n. 5, p. 386–408, Oct. 2011.

(187) AMERICAN SOCIETY OF MICROBIOLOGY (ASM). **Fight, Flight, or Freeze: Bacterial Self-Preservation through Toxin-Antitoxin Systems**. Available at: <https://www.asm.org/index.php/general-science-blog/item/7080-fight-flight-or-freeze-bacterial-self-preservation-through-toxin-antitoxin-systems>. Accessed 7 Jan. 2019.

(188) PAL, D. et al. Antibiotic-free Expression System for the Production of Human Interferon-beta Protein. **3 Biotech**, v. 8, n. 1, p. 36, 1 Jan. 2018.

(189) BARROCA, M. et al. Antibiotic Free Selection for the High Level Biosynthesis of a Silk- Elastin-like Protein. **Scientific Reports**, v. 6, n. 1, p. 39329, 16 Dec. 2016.

(190) SHI, X. et al. High-Level Expression of Recombinant Thermostable β-Glucosidase in Escherichia Coli by Regulating Acetic Acid. **Bioresource Technology**, v. 241, p. 795–801, Oct. 2017.

(191) HORN, U. et al. High Volumetric Yields of Functional Dimeric Miniantibodies in Escherichia Coli, Using an Optimized Expression Vector and High-Cell-Density Fermentation under Non-Limited Growth Conditions. **Applied Microbiology and Biotechnology**, v. 46, p. 524–32, 1996.

(192) PARK, S. Y. et al. Metabolic Engineering of Escherichia Coli for High-Level Astaxanthin Production with High Productivity. **Metabolic Engineering**, v. 49, p. 105–15, 1 Sep. 2018.

(193) LEE, S. Y. et al. Construction of Plasmids, Estimation of Plasmid Stability, and Use of Stable Plasmids for the Production of Poly(3-Hydroxybutyric Acid) by Recombinant Escherichia Coli. **Journal of Biotechnology**, v. 32, n. 2, p. 203–11, 14 Feb. 1994.

(194) MILLION-WEAVER, S.; CAMPS, M. Mechanisms of Plasmid Segregation: Have Multicopy Plasmids Been Overlooked? **Plasmid**, v. 75, p. 27–36, 1 Sep. 2014.

(195) SCHUMACHER, M. A. Bacterial Plasmid Partition Machinery: A Minimalist Approach to Survival. **Current Opinion in Structural Biology**, v. 22, n. 1, p. 72–9, 1 Feb. 2012.

(196) SUMMERS, D. Timing, Self-Control and a Sense of Direction Are the Secrets of Multicopy Plasmid Stability. **Molecular Microbiology**, v. 29, n. 5, p. 1137–45, 1 Sep. 1998.

(197) WILMS, B. et al. High-Cell-Density Fermentation for Production of L-N-Carbamoylase Using an Expression System Based on the Escherichia Coli RhaBAD Promoter. **Biotechnology and Bioengineering**, v. 73, n. 2, p. 95–103, 2001.

(198) OU, B. et al. Techniques for Chromosomal Integration and Expression Optimization in Escherichia Coli. **Biotechnology and Bioengineering**, v. 115, n. 10, p. 2467–78, 2018.

(199) GU, P. et al. A Rapid and Reliable Strategy for Chromosomal Integration of Gene(s)

with Multiple Copies. **Scientific Reports**, v. 5, n. 1, p. 9684, 8 Sep. 2015.

(200) YIN, J. et al. Development of an Enhanced Chromosomal Expression System Based on Porin Synthesis Operon for Halophile Halomonas Sp. **Applied Microbiology and Biotechnology**, v. 98, n. 21, p. 8987–97, 29 Nov. 2014.

(201) SRINIVASAN, S.; BARNARD, G. C.; GERNGROSS, T. U. Production of Recombinant Proteins Using Multiple-Copy Gene Integration in High-Cell-Density Fermentations of Ralstonia Eutropha. **Biotechnology and Bioengineering**, v. 84, n. 1, p. 114–20, 5 Oct. 2003.

(202) BARNARD, G. C. et al. High Level Recombinant Protein Expression in Ralstonia Eutropha Using T7 RNA Polymerase Based Amplification. **Protein Expression and Purification**, v. 38, n. 2, p. 264–71, Dec. 2004.

(203) GAO, Y. et al. Development of Genetically Stable Escherichia Coli Strains for Poly(3-Hydroxypropionate) Production. **PLoS ONE**, v. 9, n. 5, p. e97845, 16 May 2014.

(204) TYO, K. E. J.; AJIKUMAR, P. K.; STEPHANOPOULOS, G. Stabilized Gene Duplication Enables Long-Term Selection-Free Heterologous Pathway Expression. **Nature Biotechnology**, v. 27, n. 8, p. 760–5, 26 Aug. 2009.

(205) CÁNOVAS, M. et al. Factors Affecting the Biotransformation of Trimethylammonium Compounds into L-Carnitine by Escherichia Coli. **Biochemical Engineering Journal**, v. 26, p. 145–54, 2005.

(206) KISS, A. A.; GRIEVINK, J.; RITO-PALOMARES, M. A Systems Engineering Perspective on Process Integration in Industrial Biotechnology. **Journal of Chemical Technology & Biotechnology**, v. 90, n. 3, p. 349–55, 1 Mar. 2015.

(207) JIMÉNEZ-GONZÁLEZ, C.; WOODLEY, J. M. Bioprocesses: Modeling Needs for Process Evaluation and Sustainability Assessment. **Computers & Chemical Engineering**, v. 34, n. 7, p. 1009–17, Jul. 2010.

(208) SHANKLIN, T. et al. Selection of Bioprocess Simulation Software for Industrial Applications. **Biotechnology and bioengineering**, v. 72, n. 4, p. 483–9, 20 Feb. 2001.

(209) HANGOS, K. M.; CAMERON, I. T. **Process modelling and model analysis**. San Diego, CA, USA: Academic Press, 2001. 543 p.

(210) HEINZLE, E.; BIWER, A. P.; COONEY, C. L. **Development of Sustainable Bioprocesses**: Modeling and Assessment. Chichester, England: Wiley, 2007. 294 p.

(211) DARKWAH, K.; KNUTSON, B. L.; SEAY, J. R. A Perspective on Challenges and Prospects for Applying Process Systems Engineering Tools to Fermentation-Based Biorefineries. **ACS Sustainable Chemistry & Engineering**, v. 6, n. 3, p. 2829–44, 5 Mar. 2018.

(212) BALDEA, M. From Process Integration to Process Intensification. **Computers & Chemical Engineering**, v. 81, p. 104–14, 4 Oct. 2015.

(213) STANKIEWICZ, A. I.; MOULIJN, J. A. Process Intensification: Transforming Chemical Engineering. **Chemical Engineering Progress**, v. 96, p. 22–31, Jan. 2000.

(214) HERMANN, B. G.; BLOK, K.; PATEL, M. K. Producing Bio-Based Bulk Chemicals Using Industrial Biotechnology Saves Energy and Combats Climate Change. **Environmental Science & Technology**, v. 41, n. 22, p. 7915–21, Nov. 2007.

(215) WOODLEY, J. M. Bioprocess Intensification for the Effective Production of Chemical Products. **Computers & Chemical Engineering**, v. 105, p. 297–307, 4 Oct. 2017.

(216) KOUTINAS, M. et al. Bioprocess Systems Engineering: Transferring Traditional Process Engineering Principles to Industrial Biotechnology. **Computational and Structural Biotechnology Journal**, v. 3, n. 4, p. e201210022, 1 Oct. 2012.

(217) BOGLE, I. D. L. et al. A Process Systems Engineering View of Biochemical Process Operations. **Computers & Chemical Engineering**, v. 20, n. 6–7, p. 943–9, Jun. 1996.

(218) GERNAEY, K. V et al. Application of Mechanistic Models to Fermentation and Biocatalysis for Next-Generation Processes. **Trends in biotechnology**, v. 28, n. 7, p. 346–54, Jul. 2010.

(219) ZHUANG, K.; BAKSHI, B. R.; HERRGÅRD, M. J. Multi-Scale Modeling for Sustainable

Chemical Production. **Biotechnology Journal**, v. 8, n. 9, p. 973–84, 1 Sep. 2013.

(220) SHEN, C. R. et al. Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in Escherichia Coli. **Applied and Environmental Microbiology**, v. 77, n. 9, p. 2905–15, 1 May 2011.

(221) VICKERS, C. E.; KLEIN-MARCUSCHAMER, D.; KRÖMER, J. O. Examining the Feasibility of Bulk Commodity Production in Escherichia Coli. **Biotechnology Letters**, v. 34, n. 4, p. 585–96, 10 Apr. 2011.

(222) BIDDY, M. J. et al. The Techno-Economic Basis for Coproduct Manufacturing to Enable Hydrocarbon Fuel Production from Lignocellulosic Biomass. **ACS Sustainable Chemistry and Engineering**, v. 4, n. 6, p. 3196–211, 2016.

(223) DA SILVA, T. L.; GOUVEIA, L.; REIS, A. Integrated Microbial Processes for Biofuels and High Value-Added Products: The Way to Improve the Cost Effectiveness of Biofuel Production. **Applied Microbiology and Biotechnology**, v. 98, n. 3, p. 1043–53, 13 Feb. 2014.

(224) BUDZIANOWSKI, W. M. High-Value Low-Volume Bioproducts Coupled to Bioenergies with Potential to Enhance Business Development of Sustainable Biorefineries. **Renewable and Sustainable Energy Reviews**, v. 70, p. 793–804, 1 Apr. 2017.

(225) RENGANATH RAO, R. et al. Alkaline Protease Production from Brevibacterium Luteolum (MTCC 5982) Under Solid-State Fermentation and Its Application for Sulfide-Free Unhairing of Cowhides. **Applied Biochemistry and Biotechnology**, v. 182, n. 2, p. 511–28, 2 Jun. 2017.

(226) DE CASTRO, A. M. et al. Techno-Economic Analysis of a Bioprocess for the Production of Multienzyme Solutions from the Cake of Babassu Industrial Processing: Evaluation of Five Different Inoculum Propagation Strategies. **Biomass Conversion and Biorefinery**, v. 4, n. 3, p. 237–47, 4 Sep. 2014.

(227) ZHUANG, J.; MARCHANT, M. Economic Analysis of Cellulase Production Methods for Bio-Ethanol. **Applied Engineering in Agriculture**, v. 23, n. 5, p. 679–87, 2007.

(228) OLOFSSON, J. et al. Integrating Enzyme Fermentation in Lignocellulosic Ethanol Production: Life-Cycle Assessment and Techno-Economic Analysis. **Biotechnology for Biofuels**, v. 10, n. 51, p. 51, 23 Dec. 2017.

(229) HONG, Y. et al. Impact of Cellulase Production on Environmental and Financial Metrics for Lignocellulosic Ethanol. **Biofuels, Bioproducts and Biorefining**, v. 7, n. 3, p. 303–13, May 2013.

(230) BARTA, Z. et al. Process Design and Economics of On-Site Cellulase Production on Various Carbon Sources in a Softwood-Based Ethanol Plant. **Enzyme Research**, v. 2010, p. 1–8, 2010. .

(231) MICHAILOS, S.; PARKER, D.; WEBB, C. Simulation Studies on Ethanol Production from Sugar Cane Residues. **Industrial and Engineering Chemistry Research**, v. 55, n. 18, p. 5173–9, 11 May 2016.

(232) DAVIS, R. et al. **Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbons: Dilute-Acid and Enzymatic Deconstruction of Biomass to Sugars and Biological Conversion of Sugars to Hydrocarbons**. Golden, CO, USA: National Renewable Energy Laboratory, Mar. 2015. 124 pp. Technical Report, NREL/TP-5100-62498.

(233) TAO, L. et al. **NREL 2012 Achievement of Ethanol Cost Targets: Biochemical Ethanol Fermentation via Dilute-Acid Pretreatment and Enzymatic Hydrolysis of Corn Stover**. Golden, CO, USA: National Renewable Energy Laboratory, Apr. 2014. 36 pp. Technical Report, NREL/TP-5100-61563.

(234) DAVIS, R. et al. **Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbons: Dilute-Acid and Enzymatic Deconstruction of Biomass to Sugars and Biological Conversion of Sugars to Hydrocarbons**. Golden, CO, USA: National Renewable Energy Laboratory, Oct. 2013. 139 pp. Technical Report, NREL/TP-5100-60223.

(235) LÚ-CHAU, T. A. et al. Scale-up and Economic Analysis of the Production of Ligninolytic Enzymes from a Side-Stream of the Organosolv Process. **Journal of Chemical Technology & Biotechnology**, v. 93, n. 11, p. 3125–34, 1 Nov. 2018.

(236) KHOOTAMA, A.; PUTRI, D. N.; HERMANSYAH, H. Techno-Economic Analysis of Lipase Enzyme Production from Aspergillus Niger Using Agro-Industrial Waste by Solid State Fermentation. **Energy Procedia**, v. 153, p. 143–8, 1 Oct. 2018.

(237) NAKKEERAN, E. et al. Techno-Economic Analysis of Processes for Aspergillus Carbonarius Polygalacturonase Production. **Journal of Bioscience and Bioengineering**, v. 113, n. 5, p. 634–40, 1 May 2012.

(238) OSMA, J. F.; TOCA-HERRERA, J. L.; RODRÍGUEZ-COUTO, S. Cost Analysis in Laccase Production. **Journal of Environmental Management**, v. 92, n. 11, p. 2907–12, 1 Nov. 2011.

(239) CASTILHO, L. R. et al. Economic Analysis of Lipase Production by Penicillium Restrictum in Solid-State and Submerged Fermentations. **Biochemical Engineering Journal**, v. 4, n. 3, p. 239–47, 1 Feb. 2000.

(240) CASTRO, A. M. de et al. Economic Analysis of the Production of Amylases and Other Hydrolases by Aspergillus Awamori in Solid-State Fermentation of Babassu Cake. **Enzyme Research**, v. 2010, p. 1–9, 23 May 2010.

(241) PEZZELLA, C. et al. A Step Forward in Laccase Exploitation: Recombinant Production and Evaluation of Techno-Economic Feasibility of the Process. **Journal of Biotechnology**, v. 259, p. 175–81, 10 Oct. 2017.

(242) SEIDMAN, C. E. et al. Introduction of Plasmid DNA into Cells. In: **Current Protocols in Molecular Biology**. Hoboken, NJ, USA: Wiley, 2001. v. 272, n. 8, p. 5342–7.

(243) COLD SPRING HARBOR LABORATORY. SOC Medium for E. Coli. **Cold Spring Harbor Protocols**, v. 2012, n. 6, p. pdb.rec069732, 1 Jun. 2012.

(244) COLD SPRING HARBOR LABORATORY. LB (Luria-Bertani) Liquid Medium. **Cold Spring Harbor Protocols**, v. 2006, n. 1, p. pdb.rec8141, Jun. 2006.

(245) COLD SPRING HARBOR LABORATORY. 2× YT Medium. **Cold Spring Harbor Protocols**, v. 2014, n. 12, p. pdb.rec082222, 1 Dec. 2014.

(246) QIAGEN. **QIAprep Spin Miniprep Kit High-Yield Protocol**. Available at:

<https://www.qiagen.com/br/resources/resourcedetail?id=ebb38a38-c61c-40a7-910d-9b10d1148022&lang=en>. Accessed 26 May 2016.

(247) THERMO SCIENTIFIC. **Double Digest Calculator**. Available at: <https://www.thermofisher.com/br/en/home/brands/thermo-scientific/molecular-biology/thermo-scientific-restriction-modifying-enzymes/restriction-enzymes-thermo-scientific/double-digest-calculator-thermo-scientific.html>. Accessed 9 Oct. 2018.

(248) ARMSTRONG, J. A.; SCHULZ, J. R. Agarose Gel Electrophoresis. In: **Current Protocols Essential Laboratory Techniques**. Hoboken, NJ, USA: Wiley, 2015. p. 7.2.1–22.

(249) LEDERBERG, J.; LEDERBERG, E. M. Replica Plating and Indirect Selection of Bacterial Mutants. **Journal of Bacteriology**, v. 63, n. 3, p. 399–406, Mar. 1952.

(250) COLD SPRING HARBOR LABORATORY. Phosphate-Buffered Saline (PBS): **Cold Spring Harbor Protocols**, v. 2006, n. 1, p. pdb.rec8247, Jun. 2006.

(251) EGUIA, F. A. P. **Desenvolvimento de uma tecnologia para produção e purificação da proteína rhG-CSF (Fator Estimulador de Colônia de Granulócito humano recombinante) produzida em Escherichia coli**. São Paulo: Universidade de São Paulo, 2018.

(252) COLD SPRING HARBOR LABORATORY. M9 Minimal Medium (Standard). **Cold Spring Harbor Protocols**, v. 2010, n. 8, p. pdb.rec12295, Aug. 2010.

(253) ANKE N. KAYSER. **Stoffflüsse in Escherichia coli TG1 unter aeroben, glucoselimitierten Bedingungen**. Braunschweig, Germany: Technische Universität Braunschweig, 1999. 134 p.

(254) KRUGER, N. J. The Bradford Method for Protein Quantitation. In: WALKER J. M. (Ed). **Basic Protein and Peptide Protocols**. New Jersey, USA: Humana Press, 1994. v. 32, p. 9–16.

(255) BIO-RAD. **Handcasting Polyacrylamide Gels**. Bulletin 6201. Available at: < http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6201.pdf>. Accessed 12 Jun. 2018.

(256) COLD SPRING HARBOR LABORATORY. 4X SDS-PAGE Loading Buffer. **Cold Spring Harbor Protocols**, v. 2006, n. 25, p. pdb.rec10588-pdb.rec10588, 1 Oct. 2006.

(257) GARFIN, D. E. One-Dimensional Gel Electrophoresis. **Methods in Enzymology**, v. 182, n. C, p. 425–41, 1 Jan. 1990.

(258) GALLAGHER, S. R. One-Dimensional SDS Gel Electrophoresis of Proteins. In: **Current Protocols in Protein Science**. Hoboken, NJ, USA: Wiley, 2012. v. 1., p. 10.1.1–44.

(259) GOLD ANALISA DIAGNÓSTICA. **Instruções de Uso - kit Glicose REF. 434**. May 2018. Available at: < http://www.goldanalisa.com.br/arquivos/%7BF1F784D0-7245-4832-BC2B-3C801BB82D82%7D_GLICOSE_REF_434.pdf >. Accessed 27 Jan. 2019.

(260) GLAZYRINA, J. et al. High Cell Density Cultivation and Recombinant Protein Production with Escherichia Coli in a Rocking-Motion-Type Bioreactor. **Microbial Cell Factories**, v. 9, n. 1, p. 42, 30 May 2010.

(261) OLSSON, L.; NIELSEN, J. On-Line and in Situ Monitoring of Biomass in Submerged Cultivations. **Trends in Biotechnology**, v. 15, n. 12, p. 517–22, 1 Dec. 1997.

(262) NOVICK, R. P. Plasmid Incompatibility. **Microbiological Reviews**, v. 51, n. 4, p. 381–95, Dec. 1987.

(263) ENJALBERT, B. et al. Acetate Exposure Determines the Diauxic Behavior of Escherichia Coli during the Glucose-Acetate Transition. **Journal of Bacteriology**, v. 197, n. 19, p. 3173–81, 2015.

(264) KATAYAMA, K. Y. **Avaliação do Sistema de Estabilização Plasmidial Toxina-Antitoxina para a Produção De Proteínas Recombinantes em Escherichia coli**. São Paulo: Universidade de São Paulo, 2016. 66 p.

(265) OPENWETWARE. **E. coli Genotypes**. Available at: <https://openwetware.org/wiki/E._coli_genotypes>. Accessed 12 Dec. 2018.

(266) SZPIRER, C. Y.; MILINKOVITCH, M. C. Separate-Component-Stabilization System for Protein and DNA Production without the Use of Antibiotics. **BioTechniques**, v. 38, n.

5, p. 775–81, 2005.

(267) TERRINONI, M. et al. A Novel Nonantibiotic, Lgt-Based Selection System for Stable Maintenance of Expression Vectors in Escherichia Coli and Vibrio Cholerae. **Applied and Environmental Microbiology**, v. 84, n. 4, p. e02143-17, Feb. 2018.

(268) FIGUEIREDO, D. B. et al. Production and Purification of an Untagged Recombinant Pneumococcal Surface Protein A (PspA4Pro) with High-Purity and Low Endotoxin Content. **Applied Microbiology and Biotechnology**, v. 101, n. 6, p. 2305–17, 26 Mar. 2017.

(269) SENN, H. et al. The Growth of Escherichia Coli in Glucose-Limited Chemostat Cultures: A Re-Examination of the Kinetics. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1201, n. 3, p. 424–36, 15 Dec. 1994.

(270) FERREIRA, R. G.; AZZONI, A. R.; FREITAS, S. Techno-Economic Analysis of the Industrial Production of a Low-Cost Enzyme Using E. Coli: The Case of Recombinant β-Glucosidase. **Biotechnology for Biofuels**, v. 11, n. 1, p. 81, 29 Dec. 2018.

(271) DIAS, M. O. S. et al. Integrated versus Stand-Alone Second Generation Ethanol Production from Sugarcane Bagasse and Trash. **Bioresource Technology**, v. 103, n. 1, p. 152–61, Jan. 2012.

(272) MERCK PATENT GMBH (Darmstadt, Germany). Wolfgang Strittmater et al. **Process for the preparation of recombinant proteins in *E. coli* by high cell density fermentation**. US 6410270 B1, 28 Nov. 1996, 25 Jun. 2002. Available at: <https://patents.google.com/patent/US6410270B1>. Accessed 4 Feb. 2015.

(273) PHUE, J.-N.; SHILOACH, J. Transcription Levels of Key Metabolic Genes Are the Cause for Different Glucose Utilization Pathways in E. coli B (BL21) and E. Coli K (JM109). **Journal of Biotechnology**, v. 109, n. 1–2, p. 21–30, Apr. 2004.

(274) KORZ, D. J. et al. Simple Fed-Batch Technique for High Cell Density Cultivation of Escherichia Coli. **Journal of Biotechnology**, v. 39, n. 1, p. 59–65, Feb. 1995.

(275) MILO, R.; PHILLIPS, R. **What is the macromolecular composition of the cell?** Available at: < http://book.bionumbers.org/what-is-the-macromolecular-composition-

of-the-cell/>. Accessed 15 May 2016.

(276) MACRELLI, S.; MOGENSEN, J.; ZACCHI, G. Techno-Economic Evaluation of 2nd Generation Bioethanol Production from Sugar Cane Bagasse and Leaves Integrated with the Sugar-Based Ethanol Process. **Biotechnology for Biofuels**, v. 5, n. 1, p. 22, 2012.

(277) MARIÑO, M. A.; FREITAS, S.; MIRANDA, E. A. Ethanol Precipitation of Glycosyl Hydrolases Produced by Trichoderma Harzianum P49P11. **Brazilian Journal of Chemical Engineering**, v. 32, n. 2, p. 325–33, Jun. 2015.

(278) SANTOS, M. et al. Study of the Storage Conditions of the Sugarcane Bagasse through Thermal Analysis. **Química Nova**, v. 34, n. 3, p. 507–11, 2011.

(279) SANFORD, K. et al. Scaling up of Renewable Chemicals. **Current Opinion in Biotechnology**, v. 38, p. 112–22, Apr. 2016.

(280) ATHMANATHAN, A. **An Analysis Of The Impact Of Storage Temperature, Moisture Content & Duration Upon The Chemical Components & Bioprocessing Of Lignocellulosic Biomass**. West Lafayette, IN, USA: Purdue University. 2013. 158 p.

(281) PURCHASE, B. S.; ROSETTENSTEIN, S.; BEZUIDENHOUDT, D. V. Challenges and Potential Solutions for Storage of Large Quantities of Bagasse for Power Generation. **Proceedings of the South African Sugar Technologists' Association**, v. 86, p. 495–513, 2013.

(282) MATTHEWS, G. Fermentation Equipment Selection: Laboratory Scale Bioreactor Design Considerations. In: MCNEIL, B.; HARVEY, L. (Eds). **Practical Fermentation Technology**. Chichester, England: Wiley, 2008. ch. 2, p. 3–36.

(283) VIEIRA, J. P. F. et al. Single Cell Oil Production Integrated to a Sugarcane-Mill: Conceptual Design, Process Specifications and Economic Analysis Using Molasses as Raw Material. **Industrial Crops & Products**, v. 89, p. 478–85, Oct. 2016.

(284) AXELSSON, H. Cell Separation, Centrifugation. In: FLICKINGER, M. (Ed). **Downstream Industrial Biotechnology**: Recovery and Purification. Hoboken, NJ, USA: Wiley, 2013. ch. 2, p. 27–48.

(285) SCHUBERT, P. F.; FINN, R. K. Alcohol Precipitation of Proteins: The Relationship of Denaturation and Precipitation for Catalase. **Biotechnology and Bioengineering**, v. 23, n. 11, p. 2569–90, Nov. 1981.

(286) ASENJO, J. A.; ANDREWS, B. A. Aqueous Two-Phase Systems for Protein Separation: Phase Separation and Applications. **Journal of Chromatography A**, v. 1238, p. 1–10, May 2012.

(287) SILVÉRIO, S. C. et al. Laccase Recovery with Aqueous Two-Phase Systems: Enzyme Partitioning and Stability. **Journal of Molecular Catalysis B: Enzymatic**, v. 87, p. 37–43, Mar. 2013.

(288) MOLBASE. Available at: <http://www.molbase.com>. Accessed 15 May 2016.

(289) MOSAIC CORPORATION. **Brazil NPK Statistical Update - May 2016**. Available at: <http://www.mosaicco.com/documents/Brazil_NPK_Update___May_2016_Statistics.pdf >. Accessed 15 May 2016.

(290) POTASH CORPORATION. **Market Data - Selected Fertilizer Prices**. Available at: <http://www.potashcorp.com/customers/markets/market_data/prices/>. Accessed 15 May 2016.

(291) MUSSATTO, S. I. et al. Techno-Economic Analysis for Brewer's Spent Grains Use on a Biorefinery Concept: The Brazilian Case. **Bioresource Technology**, v. 148, p. 302–10, Nov. 2013.

(292) PORTAL BRASIL (GOVERNO FEDERAL). **Energia renovável soma 99% da potência contratada em leilão da Aneel**. Available at: <http://www.brasil.gov.br/infraestrutura/2016/04/energia-renovavel-soma-99-da-potencia-contratada-em-leilao-da-aneel>. Accessed 2 May 2016.

(293) THEODORO, J. M. P. **Considerações sobre os Custos Ambientais Decorrentes do Gerenciamento dos Resíduos Sólidos e dos Efluentes Industriais Gerados no Setor Sucroalcooleiro: Um Estudo de Caso**. Araraquara: Centro Universitário de Araraquara, 2005. 130 p.

(294) INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). **ipeadata**. Available at:

<http://www.ipeadata.gov.br/>. Accessed 5 May 2016.

(295) BANCO NACIONAL DE DESENVOLVIMENTO ECONÔMICO E SOCIAL (BNDES). **Prorenova Industrial**. Available at: <http://www.bndes.gov.br/SiteBNDES/bndes/bndes_pt/Institucional/Apoio_Financeiro/Programas_e_Fundos/Prorenova/prorenova_industrial.html>. Accessed 15 May 2016.

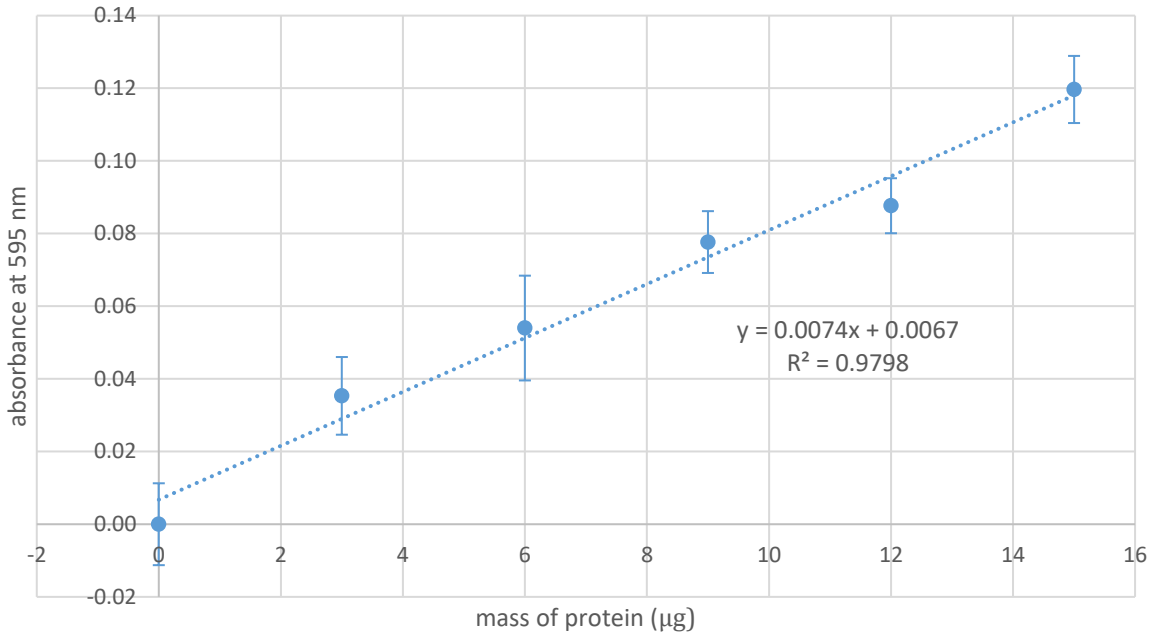
(296) CHEMICAL ENGINEERING MAGAZINE. **Chemical Engineering Plant Cost Index**. Available at: <http://www.chemengonline.com/pci-home>.

(297) BANCO CENTRAL DO BRASIL. **Taxas de Câmbio**. Available at: <http://www4.bcb.gov.br/pec/taxas/port/ptaxnpesq.asp>.

APPENDIX A – SUPPLEMENTARY CHARTS PRODUCED IN THE EXPERIMENTAL WORK

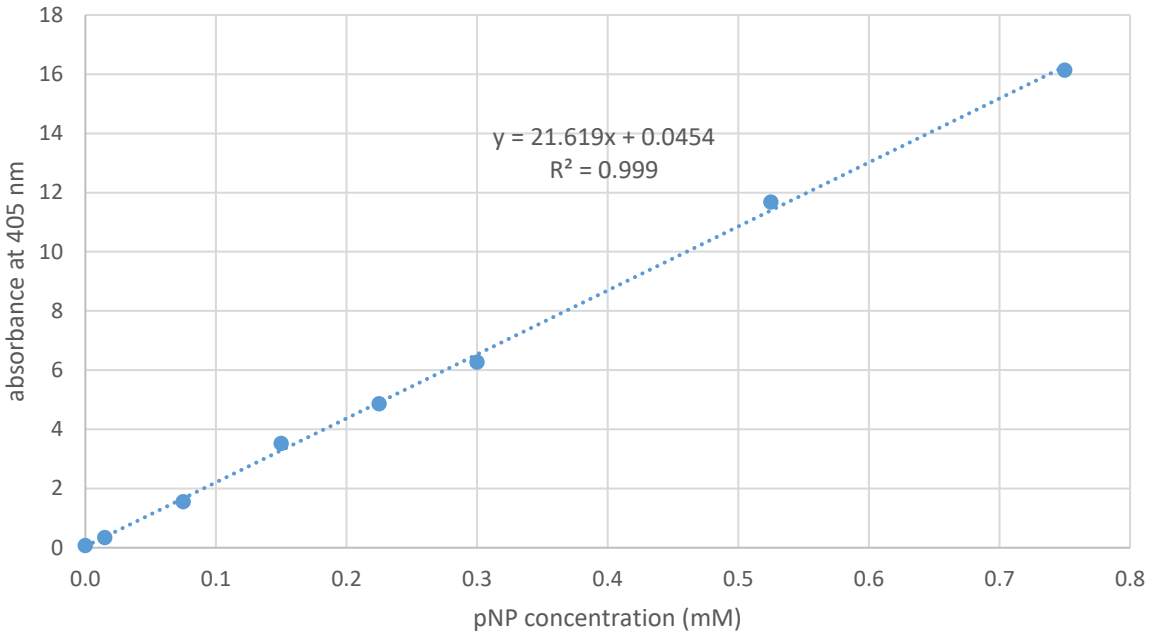
Here we present standard curves and other auxiliary plots generated during the experimental work described in Section 3.

Figure 25 – Standard curve of absorbance versus mass of protein



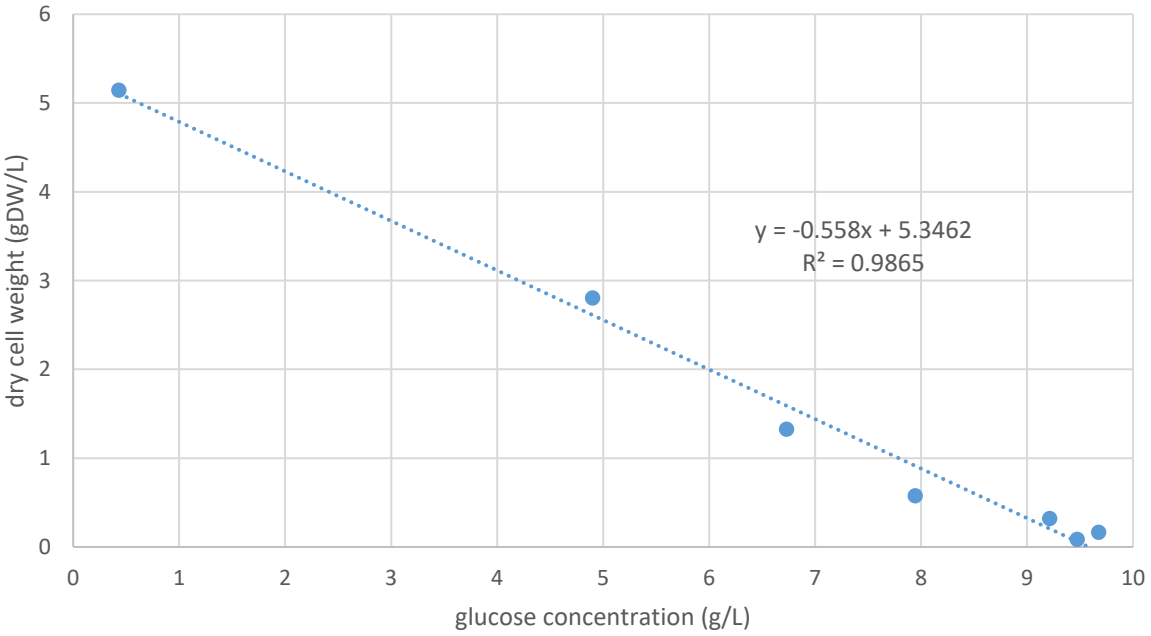
Caption: Standard curve relating mass of protein to absorbance at 405 nm in the Bradford assay. The protein chosen as the standard was hen egg lysozyme. The complete protocol employed is given in Section 3.1.4.7. Source: this work.

Figure 26 – Standard curve of absorbance versus pNP concentration



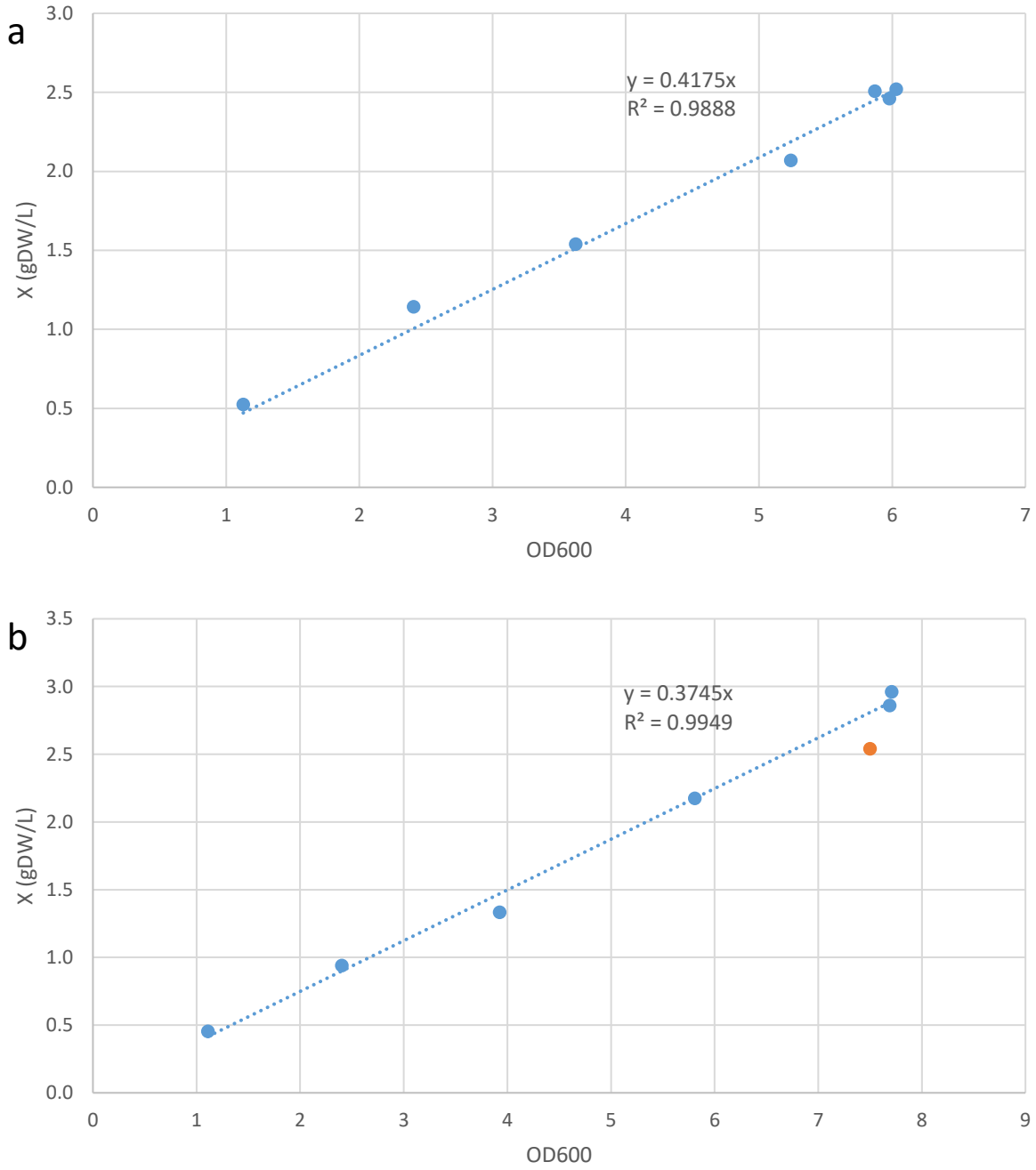
Caption: standard curve relating absorbance at 405 nm to *p*-nitrophenol (pNP) concentration. This plot was used to calculate β-glucosidase activity on *p*-nitrophenyl glucopyranoside (pNPG), given that pNPG hydrolysis releases pNP. The detailed protocol to build this curve is given in Section 3.1.4.6. Source: this work.

Figure 27 – Plot of dry cell weight versus glucose concentration in the bioreactor assay



Caption: Plot of cell concentration in terms of dry cell weight versus glucose concentration in a 2-L bioreactor culture using *E. coli* SE1/BGL. The bacteria were cultivated in HDF medium, at 37 °C, pH 7 and oxygen saturation set at 30%. The absolute value of the slope is numerically equal to the yield of biomass on glucose ($Y_{X/S}$). Source: this work.

Figure 28 – Standard curves of dry cell weight versus OD_{600} for *E. coli* BL21/BGL and SE1/BGL



Caption: standard curves relating cell concentration in terms of dry cell weight (X) and optical density at 600 nm (OD_{600}) for (a) *E. coli* BL21/BGL and (b) *E. coli* SE1/BGL. Note that the orange point in the SE1/BGL plot was considered as an outlier and disregarded in the linear regression. The complete protocol of dry cell weight determination is given in Section 3.1.4.12. Source: this work.

APPENDIX B – SUPPLEMENTARY INPUT DATA FOR THE SIMULATION

In this section, we provide supplementary data required to carry out the techno-economic analysis of recombinant β-glucosidase production, namely the main parameters of the downstream unit operations (Table 21), the costs of raw materials (Table 22), as well as the costs of utilities, labor, and financing (Table 23).

Table 21 – Main parameters of unit operations in the Downstream Section.

Parameter	**Value**
Homogenization	
Number of Passes	2
Pressure	1000 kPa
Cell Disruption	100%
rEnzyme Denaturation	5%
Centrifugation	
Sedimentation Efficiency	30%
Cell Debris Removal	70%
Solids Concentration in Heavy Stream	200 g/L
Microfiltration	
Whole-Cell Rejection Coefficient	1.00
Cell Debris Rejection Coefficient	0.95
rEnzyme Denaturation	3%
Concentration Factor	10
Dead-End Filtration	
Cell Debris Removal	100%
Contaminant Protein Removal	5%
Nucleic Acid Removal	5%
rEnzyme Removal	1%
Particle Concentration in Retentate (v/v)	50%
Diafiltration	
Volume Permeated	1
Contaminant-Protein Rejection Coefficient	0.80
rEnzyme Rejection Coefficient	1.00
rEnzyme Denaturation	3%
Concentration Factor	variable ζ

ζ Factor chosen so that the final rEnzyme concentration was equal to 15 g/L.
Caption: main parameters of downstream unit operations used in the production process of recombinant β-glucosidase. Save a few exceptions described in Section 4, all these parameters were held constant in every simulated scenario. The complete process description is provided in Section 4.1. Source: this work.

Table 22 – Costs of raw materials.

Raw Material	Price (US$/kg)	Reference
Glycerol	0.61	(288)
Glucose	0.66	(288)
$(NH_4)H_2PO_4$	0.35	(289)
Ammonia Gas	0.30	(290)
$(NH_4)_2HPO_4$	0.35	(290)
KH_2PO_4	1.29	(288)
Process Water	0.18×10^{-3}	(276)
Citric Acid	0.82	(288)
$CoCl_2$	9.52	(288)
$CuCl_2$	4.30	(288)
EDTA*Na_2	2.74	(288)
H_3BO_3	0.73	(288)
Iron III Citrate	7.63	(288)
$MgSO_4$	0.48	(288)
$MnCl_2$	1.71	(288)
Na_2MoO_4	13.48	(288)
$Zn(OAc)_2$	1.72	(288)
Kanamycin Sulfate	31.72	(288)
IPTG	601.00	(288)

Caption: Costs of raw materials used in the production process of recombinant β-glucosidase. Almost all prices of raw materials were obtained from the website molbase.com (288). It consists of an aggregator of international chemical suppliers, and provides a convenient "reference price" for each compound, which is an average price from multiple vendors automatically calculated by the website. Except for the cost of glucose, these values were kept constant in all simulated scenarios. Source: this work.

Table 23 – Costs of utilities, labor, and financing.

Parameter	Value	Unit	Source
Utilities			
Cooling Water (20 °C)	0.04	US$/t	(276)
Chilled Water (5 °C)	0.35	US$/t	Estimate based on SuperPro Designer default value
Steam	1.47	US$/t	(291)
Electricity	56	US$/(kW·h)	(292)
Waste Treatment			
Filter Cake	4.07	US$/t	(293)
Aqueous Waste	1.71	US$/$m^3$	(293)
Labor cost			
Basic Rate	7.93	US$/h	(294)
Operating Supplies	10%		SuperPro Designer default
Supervision	20%		SuperPro Designer default
Administration	60%		SuperPro Designer default
Adjusted Rate	15.07	US$/h	Calculated from above
Time Parameters			
Month/Year of Analysis and Construction	May 2016		
Construction Period	30	months	
Startup Period	6	months	
Project Lifetime	25	years	
Debt	50% of project		(295)
Loan Period	6	years	(295)
Loan Annual Interest Rate	10.25%		BNDES rate for medium/large ethanol plants (295)
Price Indices			
Chemical Engineering Plant Cost Index (CEPCI)	Inflation of equipment prices estimated by SuperPro Designer		(296)
Nominal Average Salary in Sao Paulo state	Inflation of labor cost		(294)
Índice de Preço ao Produtor (IPP) - Indústria de Transformação/ Petróleo e Biocombustíveis	Inflation of utilities and material costs		(294)
IPA-M	Inflation of waste treatment costs		(294)
Dollar/Real Exchange Rates	3.0		(297)

Caption: Costs of utilities, labor, and financing in the production process of recombinant β-glucosidase. All the values above were kept constant in every simulated scenario. Source: this work.

APPENDIX C – SUPPLEMENTARY OUTPUT DATA FROM THE PROCESS SIMULATION

In this appendix, we present additional data generated by the simulation of the enzyme production process in SuperPro Designer, considering the baseline scenario. Table 24 lists all the unit operations of the process, as well as their scheduling times; Table 25 shows major equipment specifications and purchase costs, as estimated by the simulator; and, lastly, Table 26 computes all the capital costs incurred in order to build the enzyme plant, including installation, building, engineering and startup costs.

Table 24 – Description and scheduling of each procedure performed during a single batch of enzyme production.

Task	Duration (h)	Start Time (h)	End Time (h)	Description
Complete Recipe	60.15	0.00	60.15	
P-24 in V-104	1.51	0.00	1.51	Blending / Storage
CHARGE-2	0.26	0.00	0.26	Charge 155 kg of Main Medium to V-104 (in P-24), using stream 'Medium (SFR-1)'.
TRANSFER-OUT-1	1.00	0.26	1.26	Transfer 100% from V-104 (in P-24) to ST-103 (in P-23) for 1.00 h.
CIP-1	0.25	1.26	1.51	Execute CIP steps: Cleaning Step #1. .
P-23 in ST-103	1.50	0.26	1.76	Heat Sterilization
STERILIZE-1	1.00	0.26	1.26	Sterilize at a holding T = 140 °C, for 1.0 h, the sterilization criterion is $\ln(N_0/N) = 13.82$.
CIP-1	0.25	1.26	1.51	Execute CIP steps: Cleaning Step #1. .
HOLD-1	0.25	1.51	1.76	Hold for 15 min.

Caption: Brief description and scheduling of each procedure performed during a single batch ("recipe") of the enzyme production process, in the baseline scenario. Each major piece of equipment is indicated by an acronym or abbreviation (e.g. SFR-1 stands for Seed Fermenter #1) and assigned a set of procedures (e.g. P-20). Note: table split into five pages (pp. 181-185). Source: this work.

Task	Duration (h)	Start Time (h)	End Time (h)	Description
P-20 in SFR-1	15.36	0.26	15.62	Seed Fermentation
TRANSFER-IN-1	1.00	0.26	1.26	Transfer in for 1.0 h from ST-103 (in P-23) to SFR-1 (in P-20).
REACT-1	1.00	1.26	2.26	React in batch mode for 60 min., under adiabatic conditions and pressure of 5.210 bar.
REACT-2	1.00	1.26	2.26	React in batch mode for 60 min., under adiabatic conditions and pressure of 5.210 bar.
CHARGE-1	0.08	1.26	1.34	Charge 4.66 g of Kanamycin Sulfa to SFR-1 (in P-20), using stream 'S-124'.
CHARGE-2	0.25	1.34	1.59	Charge 435 g of NH3 solution to SFR-1 (in P-20), using stream 'S-101'.
FERMENT-1	13.02	1.59	14.61	Ferment for 13.02 h, at 26 °C and pressure of 1.014 bar, with an aeration rate of 1.0 VVM.
TRANSFER-OUT-1	0.26	14.61	14.87	Transfer 100% from SFR-1 (in P-20) to SFR-2 (in P-14) at 600.000 kg/h.
CIP-1	0.25	14.87	15.12	Execute CIP steps: Cleaning Step #1. .
SIP-1	0.50	15.12	15.62	SIP for 30 min using Steam. The agent rate per volume contribution is 100 (kg/h)/m3.
P-21 in G-104	13.02	1.59	14.61	Gas Compression
COMPRESS-1	13.02	1.59	14.61	Compress air/gas with a pressure increase of 5.000 bar, for 13.02 h.
P-22 in AF-105	13.02	1.59	14.61	Air Filtration
FILTER-1	13.02	1.59	14.61	Filter gaseous stream, for 13.02 h.
P-25 in AF-107	13.02	1.59	14.61	Air Filtration
FILTER-1	13.02	1.59	14.61	Filter gaseous stream, for 13.02 h.
P-18 in V-103	3.23	11.63	14.86	Blending / Storage
CHARGE-2	0.98	11.63	12.61	Charge 2952.980 kg of mixture to V-103 (in P-18), using stream 'Medium (SFR-2)'.
TRANSFER-OUT-1	2.00	12.61	14.61	Transfer 100% from V-103 (in P-18) to ST-101 (in P-17) for 2.0 h.
CIP-1	0.25	14.61	14.86	Execute CIP steps: Cleaning Step #1. .

Caption: Brief description and scheduling of each procedure performed during a single batch (“recipe”) of the enzyme production process, in the baseline scenario. Each major piece of equipment is indicated by an acronym or abbreviation (e.g. SFR-1 stands for Seed Fermenter #1) and assigned a set of procedures (e.g. P-20). Note: table split into five pages (pp. 181-185). Source: this work.

Task	Duration (h)	Start Time (h)	End Time (h)	Description
P-17 in ST-101	2.50	12.61	15.11	Heat Sterilization
STERILIZE-1	2.00	12.61	14.61	Sterilize at a holding T = 140 °C, for 2.0 h, the sterilization criterion is $\ln(N_0/N)$ = 13.82.
CIP-1	0.25	14.61	14.86	Execute CIP steps: Cleaning Step #1. .
HOLD-1	0.25	14.86	15.11	Hold for 15 min.
P-14 in SFR-2	18.67	12.61	31.29	Seed Fermentation
TRANSFER-IN-1	2.00	12.61	14.61	Transfer in for 2.0 h from ST-101 (in P-17) to SFR-2 (in P-14).
CHARGE-1	0.25	14.36	14.61	Charge 93.25 g of Kanamycin Sulfa to SFR-2 (in P-14), using stream 'S-111'.
CHARGE-2	0.26	14.61	14.88	Charge 8.70 kg of NH3 solution to SFR-2 (in P-14), using stream 'S-102'.
TRANSFER-IN-2	0.26	14.61	14.87	Transfer in for 15.5 min from SFR-1 (in P-20) to SFR-2 (in P-14).
REACT-1	2.00	14.87	16.87	React in batch mode for 120 min., under adiabatic conditions and pressure of 5.258 bar. .
REACT-2	2.00	14.87	16.87	React in batch mode for 120 min., under adiabatic conditions and pressure of 5.258 bar. .
FERMENT-1	13.02	16.87	29.89	Ferment for 13.02 h, at 26 °C and pressure of 1.014 bar, with an aeration rate of 1.0 VVM.
TRANSFER-OUT-1	0.64	29.89	30.54	Transfer 100% from SFR-2 (in P-14) to FR-101 (in P-1) at 80 L/min.
CIP-1	0.25	30.54	30.79	Execute CIP steps: Cleaning Step #1. .
SIP-1	0.50	30.79	31.29	SIP for 30 min using Steam. The agent rate per volume contribution is 100 (kg/h)/m3. .
P-15 in G-103	13.02	16.87	29.89	Gas Compression
COMPRESS-1	13.02	16.87	29.89	Compress air/gas with a pressure increase of 5.000 bar, for 13.02 h.
P-16 in AF-104	13.02	16.87	29.89	Air Filtration
FILTER-1	13.02	16.87	29.89	Filter gaseous stream, for 13.02 h.

Caption: Brief description and scheduling of each procedure performed during a single batch ("recipe") of the enzyme production process, in the baseline scenario. Each major piece of equipment is indicated by an acronym or abbreviation (e.g. SFR-1 stands for Seed Fermenter #1) and assigned a set of procedures (e.g. P-20). Note: table split into five pages (pp. 181-185). Source: this work.

Task	Duration (h)	Start Time (h)	End Time (h)	Description
P-19 in AF-106	13.02	16.87	29.89	Air Filtration
FILTER-1	13.02	16.87	29.89	Filter gaseous stream, for 13.02 h.
P-13 in V-102	7.84	22.30	30.14	Blending / Storage
CHARGE-1	2.53	22.30	24.83	Charge 79.7 MT of mixture to V-102 (in P-13), using stream 'Medium'.
AGITATE-1	5.06	24.83	29.89	Agitate for 5.06 h at a specific power of 0.10 kW/m3.
TRANSFER-OUT-1	5.06	24.83	29.89	Transfer 100% from V-102 (in P-13) to ST-102 (in P-5) for 5.06 h.
CIP-1	0.25	29.89	30.14	Execute CIP steps: Cleaning Step #1. .
P-5 in ST-102	5.56	24.83	30.39	Heat Sterilization
STERILIZE-1	5.06	24.83	29.89	Sterilize at a holding T = 140 °C, for 5.1 h, the sterilization criterion is $\ln(N_0/N) = 13.82$.
CIP-1	0.25	29.89	30.14	Execute CIP steps: Cleaning Step #1. .
HOLD-1	0.25	30.14	30.39	Hold for 15 min.
P-1 in FR-101	29.82	24.83	54.65	Bioreactor
TRANSFER-IN-MEDIUM	5.06	24.83	29.89	Transfer in at 15 m3/h from ST-102 (in P-5) to FR-101 (in P-1).
REACT-1	5.06	24.83	29.89	React in batch mode for 304 min., under adiabatic conditions and pressure of 4.2 bar. .
REACT-2	5.06	24.83	29.89	React in batch mode for 304 min., under adiabatic conditions and pressure of 4.1 bar. .
REACT-3	5.06	24.83	29.89	React in batch mode for 304 min., under adiabatic conditions and pressure of 4.1 bar. .
CHARGE-AMMONIA	5.06	24.83	29.89	Charge 1.469 MT of NH3 solution to FR-101 (in P-1), using stream 'Ammonia 25%'.
CHARGE-ANTIBIOTICS	0.00	29.89	29.89	Charge 2331 g of Kanamycin Sulfa to FR-101 (in P-1), using stream 'Antibiotics'.
TRANSFER-IN-INOCULUM	0.64	29.89	30.54	Transfer in for 38.6 min from SFR-2 (in P-14) to FR-101 (in P-1).
FERMENT	22.07	30.54	52.61	Ferment for 22 h, at 26 °C and pressure of 150.5 kPa, with an aeration rate of 1 VVM.

Caption: Brief description and scheduling of each procedure performed during a single batch (“recipe”) of the enzyme production process, in the baseline scenario. Each major piece of equipment is indicated by an acronym or abbreviation (e.g. SFR-1 stands for Seed Fermenter #1) and assigned a set of procedures (e.g. P-20).. Note: table split into five pages (pp. 181-185). Source: this work.

Task	Duration (h)	Start Time (h)	End Time (h)	Description
CHARGE-INDUCER	0.03	37.54	37.57	Charge 18.52 mol of IPTG to FR-101 (in P-1), using stream 'Inducer'.
TRANSFER-OUT-BROTH	1.30	52.61	53.90	Transfer 100% from FR-101 (in P-1) to V-101 (in P-8) at 60 m3/h.
CIP-1	0.25	53.90	54.15	Execute CIP steps: Cleaning Step #1. .
SIP-1	0.50	54.15	54.65	SIP for 30 min using Steam. The agent rate per volume contribution is 100 (kg/h)/m3. .
P-2 in G-101	22.07	30.54	52.61	Gas Compression
COMPRESS-1	22.07	30.54	52.61	Compress air/gas with a pressure increase of 5.000 bar, for 1324 min.
P-3 in AF-101	22.07	30.54	52.61	Air Filtration
FILTER-1	22.07	30.54	52.61	Filter gaseous stream, for 1324 min.
P-6 in AF-102	22.07	30.54	52.61	Air Filtration
FILTER-1	22.07	30.54	52.61	Filter gaseous stream, for 1324 min.
P-8 in V-101	7.55	52.61	60.15	Storage
TRANSFER-IN-1	1.30	52.61	53.90	Transfer in for 77.8 min from FR-101 (in P-1) to V-101 (in P-8).
TRANSFER-OUT-1	6.00	53.90	59.90	Transfer 100% from V-101 (in P-8) to HG-101 (in P-10) for 6.00 h.
CIP-1	0.25	59.90	60.15	Execute CIP steps: Cleaning Step #1. .
P-10 in HG-101	6.25	53.90	60.15	Homogenization
HOMOGENIZE-1	6.00	53.90	59.90	Homogenize for 6.0 h, number of passes is 1, pressure drop is 1000 bar, and exit temp. is 18 °C.
CIP-1	0.25	59.90	60.15	Execute CIP steps: Cleaning Step #1. .
P-11 in DS-102	6.25	53.90	60.15	Centrifugation
CENTRIFUGE-1	6.00	53.90	59.90	Centrifuge for 6.0 h to remove solids.
CIP-1	0.25	59.90	60.15	Execute CIP steps: Cleaning Step #1. .

Caption: Brief description and scheduling of each procedure performed during a single batch ("recipe") of the enzyme production process, in the baseline scenario. Each major piece of equipment is indicated by an acronym or abbreviation (e.g. SFR-1 stands for Seed Fermenter #1) and assigned a set of procedures (e.g. P-20). Note: table split into five pages (pp. 181-185). Source: this work.

Table 25 – Major equipment specifications and purchase costs.

Quantity / Standby / / Staggered	Name	Description	Unit Cost (10^3 US$)	Cost (10^3 US$)
1 / 0 / 0	FR-101	Fermenter Vessel Volume = 100 m3	1460	1460
1 / 0 / 0	G-101	Centrifugal Compressor Compressor Power = 458 kW	493	493
1 / 0 / 0	ST-102	Heat Sterilizer Rated Throughput = 15.0 m3/h	487	487
1 / 0 / 0	V-101	Vertical-On-Legs Tank Vessel Volume = 86.4 m3	130	130
1 / 0 / 0	HG-101	Homogenizer Rated Throughput = 216 L/min	80	80
6 / 0 / 0	DS-102	Disk-Stack Centrifuge Throughput = 36.1 L/min	440	2640
1 / 0 / 0	AF-101	Air Filter Rated Throughput = 915 m3/h	7	7
1 / 0 / 0	AF-102	Air Filter Rated Throughput = 3518 m3/h	15	15
1 / 0 / 0	DE-101	Dead-End Filter Filter Area = 50.0 m2	145	145
1 / 0 / 0	V-102	Blending Tank Vessel Volume = 84.3 m3	393	393
1 / 0 / 0	SFR-2	Seed Fermenter Vessel Volume = 3.88 m3	681	681
1 / 0 / 0	G-103	Centrifugal Compressor Compressor Power = 17.8 kW	64	64
1 / 0 / 0	AF-104	Air Filter Rated Throughput = 35.5 /h	6	6
1 / 0 / 0	ST-101	Heat Sterilizer Rated Throughput = 1.46 m3/h	237	237
1 / 0 / 0	V-103	Blending Tank Vessel Volume = 3.25 m3	197	197
1 / 0 / 0	AF-106	Air Filter Rated Throughput = 202 m3/h	6	6
1 / 0 / 0	SFR-1	Seed Fermenter Vessel Volume = 194 L	447	447
1 / 0 / 0	G-104	Centrifugal Compressor Compressor Power = 0.89 kW	64	64

Quantity / Standby / / Staggered	Name	Description	Unit Cost (10^3 US$)	Cost (10^3 US$)
1 / 0 / 0	AF-105	Air Filter	6	6
		Rated Throughput = 1.78 m3/h		
1 / 0 / 0	ST-103	Heat Sterilizer	118	118
		Rated Throughput = 154 L/h		
1 / 0 / 0	V-104	Blending Tank	130	130
		Vessel Volume = 171 L		
1 / 0 / 0	AF-107	Air Filter	6	6
		Rated Throughput = 10.1 m3/h		
8 / 0 / 0	DF-101	Diafilter	115	920
		Membrane Area = 74.3 m2		
		Unlisted Equipment Cost (0.20×EPC)		2183
		Total Equipment Purchase Cost (EPC)		10914

Caption: Specification and free-on-board (FOB) cost of all the major pieces of equipment employed in the baseline scenario of enzyme production. All equipment prices were estimated using SuperPro Designer, adjusted to the year of 2016 with the CEPCI index (see Table 23 in Appendix B). Source: this work.

Table 26 – Complete list of capital costs.

Description	Cost (10^6 US$)	Factor	Calculation Basis
Total Plant Direct Cost (TPDC) (physical cost)			
Equipment Purchase Cost (EPC)	10.91		
Installation	4.78	custom factor for each piece of equipment; for unlisted equipment, 0.50×EPC	
A. Process Piping	3.82	0.35	EPC
B. Instrumentation	4.37	0.40	EPC
C. Insulation	0.33	0.03	EPC
D. Electrical	1.09	0.10	EPC
E. Buildings	4.91	0.45	EPC
F. Yard Improvement	1.64	0.15	EPC
G. Auxiliary Facilities	4.37	0.40	EPC
TPDC (EPC+Installation+A+B+C+D+E+F+G)	36.22		
Total Plant Indirect Cost (TPIC)			
Engineering	9.05	0.25	TPDC
Construction	12.68	0.35	TPDC
TPIC	21.73		
Total Plant Cost (TPC = TPDC+TPIC)			
TPC	57.95		
Contractor's Fee & Contingency (CFC)			
Contractor's Fee	2.90	0.05	TPC
Contingency	5.80	0.10	TPC
CFC	8.69		
Direct Fixed Capital Cost (DFC = TPC+CFC)			
DFC	66.64		
Working Capital (WC)			
WC	0.81	30 days of labor, raw materials, utilities and waste treatment	
Startup and Validation (SV)			
SV	3.33	0.05	DFC
Total Capital Expenses (CAPEX = DFC+WC+SV)			
CAPEX	70.78		

Caption: Full list of capital expenses required to build the enzyme production plant, taking the baseline process into consideration. The total plant direct cost is composed of the total equipment purchase cost (EPC) plus installation costs and other expenses, which are calculated by multiplying the EPC by suitable factors. Except for the working capital, all other capital costs ultimately derive from the plant direct cost, also by using multipliers. The multipliers used here are those proposed by SuperPro Designer, which are the same suggested by Heinzle, Biwer and Cooney (210). All costs are given in 2016 U. S. dollars. Source: this work.