






Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



**Universitat Autònoma
de Barcelona**

Departament d'Enginyeria Química, Biològica i Ambiental

**Cloning, expression and purification/immobilisation
of CBM–tagged enzymes involved in multienzymatic
production of lactic acid.**

Memòria per optar al grau de Doctor per la Universitat Autònoma de Barcelona, sota la direcció de la Dra. Marina Guillén Montalbán i el Dr. Ramón Román Roldán, i sota la tutela de la Dra. Gloria González Anadón

Per

Mario Benito Peinado

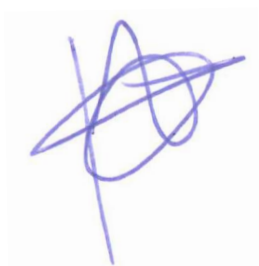
A Bellaterra, Juliol de 2022

La Dra. Marina Guillén Montalbán, en qualitat de professora del Departament d'Enginyeria Química, Biològica i Ambiental de la Universitat Autònoma de Barcelona i el Dr. Ramón Román Roldán, responsable de qualitat en laboratori a Laminar Pharma

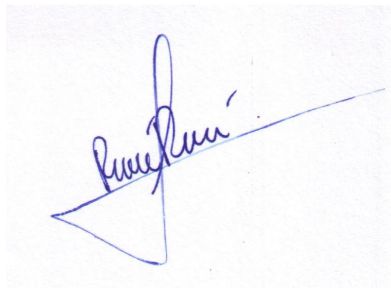
Certifiquem:

Que el graduat en Biotecnologia Mario Benito Peinado ha dut a terme sota la nostra direcció el treball titulat **Cloning, expression and purification/immobilisation of CBM-tagged enzymes involved in multienzymatic production of lactic acid**. El mateix, es presenta en aquesta memòria i constitueix el manuscrit de la tesi per optar al Grau de Doctor en Biotecnologia per la Universitat Autònoma de Barcelona.

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



Dra. Marina Guillén Montalbán



Dr. Ramón Román Roldán

A Bellaterra, 4 de Juliol de 2022

"A wizard is never late, Frodo Baggins. Nor is he early, he arrives precisely when he means to."

Peter Jackson, 2001

TABLE OF CONTENTS

RESUM	11
SUMMARY	13
ABREVIATIONS AND NOMENCLATURE	15
1. INTRODUCTION	18
1.1. Brief introduction to bioprocesses	20
1.2. The role of biocatalysts in bioprocesses	21
1.2.1. The use of enzymes	22
1.3. <i>Escherichia coli</i> as a platform for recombinant protein production	24
1.3.1. Regulation of gene expression in <i>Escherichia coli</i>	27
1.3.2. High cell-density cultures as a process intensification strategy	33
1.4. Enzyme purification and immobilisation as process intensification strategies	37
1.4.1. Enzyme purification	37
1.4.2. Enzyme immobilisation	39
1.5. Carbohydrate-Binding Modules	42
1.5.1. Carbohydrate-Binding Modules as immobilisation tags	44
1.5.2. Carbohydrate-Binding Modules as affinity purification tags	45
2. CONTEXT AND OBJECTIVES	47
2.1. Context of the thesis work	47
2.1.1. The Research Group	47
2.1.2. BIOCON-CO ₂ project	48
2.1.3. Enzymes involved in L-lactic acid production	51
2.2. Objectives of the thesis work	55
3. MATERIALS AND METHODS	56
3.1. Bacterial strains	56

3.1.1. <i>Escherichia coli</i> DH5 α	56
3.1.2. <i>Escherichia coli</i> TOP10	56
3.1.3. <i>Escherichia coli</i> NEB 10- β	57
3.1.4. <i>Escherichia coli</i> M15 Δ glyA	57
3.2. Plasmids	58
3.2.1. pBAD	58
3.2.2. pVEF	59
3.2.3. pUC57	62
3.3. Molecular biology techniques	63
3.3.1. DNA amplification	63
3.3.2. DNA isolation and manipulation	66
3.3.3. Enzyme cloning	67
3.3.4. DNA transformation	69
3.3.5. DNA sequencing	70
3.4. Growth medium composition	71
3.4.1. Complex mediums	71
3.4.2. Defined mediums	71
3.4.3. Medium supplements	76
3.5. Cultivation conditions	77
3.5.1. Shake-flask cultures	78
3.5.2. Bioreactor cultures	79
3.6. Enzyme recovery	82
3.7. Analytical methods	83
3.7.1. Biomass measurement	83
3.7.2. Substrates and by-products concentration measurement	83
3.7.3. Total protein content and enzyme titre determination	83
3.7.4. Enzymatic activity assessment	84
3.8 Biocatalysts immobilisation	86

3.8.1 Regenerated amorphous cellulose preparation	86
3.8.2. Immobilisation of fusion proteins to cellulosic supports	87
3.8.3. Immobilised derivatives storage stability	88
3.8.4. Eluent screening	88
3.9. Fast liquid protein chromatography	89
4. RESULTS I: HISTIDINE-TAGGED ENZYMES OVEREXPRESSION IN <i>E. COLI</i>	91
4.1. Overexpression studies with the NEB 10-β strain	92
4.1.1. Shake-flask overexpression experiments with complex medium	93
4.1.2. Batch production processes at bench-scale reactor with complex medium	96
4.1.3. Shake-flask overexpression experiments with defined medium	100
4.1.4 Fed-batch production processes at bench-scale reactor with defined medium	105
4.2. Changing <i>E. coli</i> host strain from NEB 10-β to M15ΔglyA	109
4.2.1. Molecular biology: cloning target genes into pVEF vector	110
4.2.2. <i>E. coli</i> M15 Δ glyA transformation and expression screening	114
4.3. Histidine-tagged enzymes production with the M15ΔglyA strain	116
4.3.1. Shake-flask overexpression experiments	119
4.3.2. Batch production processes at bench-scale reactor	122
4.3.3. Fed-batch production processes at bench-scale reactor	124
4.4. Comparison of production parameters between NEB 10-β and M15ΔglyA strains	129
4.5. Conclusions	132
5. RESULTS II: CARBOHYDRATE-BINDING MODULE-TAGGED ENZYMES PRODUCTION IN <i>E. COLI</i>	133
5.1. Generation of target CBM-fused proteins	134
5.1.1. Molecular biology: cloning target genes into pVEF vector	135

5.1.2. <i>E. coli</i> M15 Δ <i>glyA</i> transformation and expression screening	139
5.2. CBM–tagged enzymes production with the M15Δ<i>glyA</i> strain	142
5.2.1. Alcohol dehydrogenase from <i>Saccharomyces cerevisiae</i>	143
5.2.2. Pyruvate decarboxylase from <i>Saccharomyces cerevisiae</i>	146
5.2.3. Lactate dehydrogenase from <i>Thermotoga maritima</i>	148
5.3. CBM–tagged ScADH fed-bach production studies	152
5.3.1. CBM-ScADH production in fed-batch cultures	153
5.3.2. Comparison between histidine– and CBM–tagged ScADH variants	156
5.4. Conclusions	160
6. RESULTS III: ONE-STEP PURIFICATION/IMMOBILISATION OF CBM–FUSED ENZYMES	161
6.1. Immobilisation of CBM–fused enzymes into cellulosic support	162
6.1.1. CBM3–fused enzymes	163
6.1.2. CBM9–fused enzymes	172
6.2 Comparison between CBM3 and CBM9 domains as immobilisation tags	180
6.3. Use of Carbohydrate–Binding Modules as purification tags	181
6.3.1. Eluent screening	182
6.3.2. Purification of CBM9–fused enzymes	185
6.3.3. Purification of CBM3–fused enzymes	191
6.4. Comparison between CBM domains and histidine as purification tags	197
6.5. Conclusions	198
7. OVERALL CONCLUSIONS	200
8. APPENDIX	202

8.1. Results I	202
8.1.1. Adaptation of NEB 10- β cells to defined medium	202
8.1.2. Fed-batch production processes at bench-scale reactor and defined medium	203
8.1.3. Adaptation of M15 Δ <i>glyA</i> cells to defined medium	205
8.1.4. Batch production processes at bench-scale reactor	206
8.1.5. Fed-batch production processes at bench-scale reactor	208
8.2. Results II	210
8.2.1. Adaptation of M15 Δ <i>glyA</i> to defined medium	210
9. ACKNOWLEDGEMENTS	211
9.1. Funding and contributions	211
10. SCIENTIFIC CONTRIBUTIONS	212
11. REFERENCES	213
AGRAÏMENTS	229

RESUM

Aquesta tesi respon a la necessitat de reduir els costos operacionals associats a la producció, a la purificació i a la immobilització de biocatalitzadors emprats en bioprocessos. Per a desenvolupar i validar les diverses estratègies d'intensificació de procés plantejades en el present treball, s'han escollit tres biocatalitzadors —alcohol deshidrogenasa (ADH), piruvat descarboxilasa (PDC) i lactat deshidrogenasa (LDH)— els quals formen part d'un sistema biocatalític proposat dins del projecte europeu BIOCON-CO₂, per a sintetitzar àcid làctic.

En aquest context, per tal d'obtenir els biocatalitzadors d'interès de la manera més eficient possible, s'han integrat diversos procediments de diferents camps, incloent mètodes de clonació de gens i processos de producció de proteïnes recombinants amb el bacteri Gram negatiu *Escherichia coli*, així com el desenvolupament d'un procés d'una única etapa de purificació/immobilització de biocatalitzadors, basat en l'afinitat de dominis proteics d'unió a carbohidrats cap a suports cel·lulòsics de baix cost.

En la primera part d'aquesta tesi, es va emprar la soca d'*E. coli* NEB 10- β per a l'expressió dels tres enzims implicats en el sistema multienzimàtic proposat per a la síntesi d'àcid làctic amb una cua d'histidines fusionada al seu extrem N-terminal. El procés de producció, basat en el sistema d'expressió P_{BAD}, va demostrar ser exitós no només utilitzant medis de cultiu complexos, sinó també amb medis químicament definits, suplementats amb aminoàcids, seguint una estratègia de cultiu en fed-batch de dos etapes. Tot i així, per tal d'avaluar si es podien assolir millors rendiments de producció, els tres enzims d'interès es van clonar en un vector d'expressió desenvolupat en el Grup de recerca, basat en el promotor T5 *lac*, i es van produir posteriorment utilitzant una soca auxotròfica derivada de la soca M15 a escala de reactor de laboratori mitjançant cultius en fed-batch i amb un medi de cultiu definit lliure d'antibiòtics. D'aquesta manera, la soca M15 Δ *glyA* va demostrar ser una soca hoste més adequada i més robusta per a la producció en grans quantitats dels enzims necessaris, tot i que la soca NEB 10- β també va resultar ser un hoste microbià factible.

La segona part d'aquesta tesi es va centrar en la generació de noves variants dels enzims que contenien diferents mòduls d'unió a carbohidrats (en anglès CBM) fusionats als respectius extrems N-terminal,

amb l'objectiu de desenvolupar un mètode de purificació/immobilització d'una única etapa que permeti recuperar els enzims produïts amb una alta especificitat i eficiència, mitjançant l'ús de suports d'immobilització de cel·lulosa, que no només són econòmics sinó que també permeten evitar la presència de ions metàl·lics, presents habitualment en mostres processades amb resines d'afinitat que contenen metalls quelats.

Els tres enzims diana es van coexpressar amb èxit amb dos CBM diferents —un d'ells provinent de la xilanasa 10A de *T. maritima* (CBM9) i un altre originari de la proteïna estructural A del cel·lulosoma de *C. Thermocellum* (CBM3)— emprant la soca d'*E. coli* M15ΔglyA. Les proteïnes de fusió resultants es van produir amb èxit a altes concentracions en un reactor de laboratori seguint la mateixa estratègia aplicada per produir les variants amb cua d'histidina. Els resultats van mostrar que la soca M15ΔglyA era capaç de produir amb la mateixa eficiència els enzims fusionats amb els CBM com els fusionats a la cua d'histidines, demostrant-se també que els dominis CBM no tenen un impacte negatiu significatiu en la capacitat catalítica dels enzims.

Finalment, es va desenvolupar i caracteritzar el procés de purificació/immobilització d'un sol pas dels enzims fusionats amb els CBM en suports cel·lulòsics de baix cost i d'alta selectivitat, incloent l'estudi de l'estabilitat en condicions d'emmagatzematge dels derivats immobilitzats i la determinació de la capacitat de càrrega màxima dels suports. Els enzims fusionats amb CBM9 van demostrar que s'uneixen eficaçment a la cel·lulosa amorfa regenerada a altes càrregues d'enzim, mentre que els enzims fusionats amb CBM3 es van unir amb una afinitat més alta a la cel·lulosa microcristal·lina. A més, el suport RAC també es va testar per a la purificació de les proteïnes mitjançant cromatografia líquida de proteïnes, amb l'objectiu d'establir una alternativa a la cromatografia d'afinitat amb metalls quelats, mitjançant la qual només el domini CBM9 va demostrar ser adequat com a cua de purificació per als tres enzims avaluats.

SUMMARY

The present thesis responds to the need to reduce the operational costs associated with the production, purification and immobilisation of biocatalysts used in bioprocesses. The process intensification strategies described have been developed and validated for three different enzymes —an alcohol dehydrogenase (ADH), a pyruvate decarboxylase (PDC) and a lactate dehydrogenase (LDH)— which are involved in a biocatalytic system proposed by the European BIOCON-CO₂ project, which aims to synthesize lactic acid.

In this context, in order to obtain the biocatalysts of interest in the most efficient way, the integration of several procedures from different fields is attempted, including gene cloning methods and processes for the production of recombinant proteins with the Gram-negative bacterium *Escherichia coli*, and also with the development of a one-step biocatalyst purification/immobilisation process, based on the affinity of carbohydrate-active protein domains towards low-cost cellulosic supports.

In the first part of this thesis, the *E. coli* strain NEB 10- β was used for the expression of the three enzymes involved in the multienzyme system proposed to obtain lactic acid synthesis with a histidine tag fused to their N-terminal end. The production process, based on the P_{BAD} expression system, proved to be successful, not only by using complex culture mediums but also with chemically defined medium supplemented with amino acids, used in a two-step fed batch strategy. Nevertheless, aiming to assess if higher production yields could be achieved, the three target enzymes were cloned into an in-house developed expression vector, based on the T5 *lac* promoter, and were subsequently produced by using one auxotrophic M15-derived strain at bench-scale reactor through fed-batch cultures and antibiotic-free defined culture medium. Thus, M15 Δ *glyA* strain proved to be a more suitable and more robust *E. coli* host strain to produce the required enzymes at high titres, although the NEB 10- β strain also resulted to be a feasible microbial host.

The second part of this thesis was focused on the generation of new enzyme variants, containing different Carbohydrate-Binding Modules (CBM) fused to their N-terminal end, aiming to develop a one-step purification/immobilisation method which enables to recover the enzymes produced with high specificity and efficiency, by using cheap immobilisation cellulosic supports while avoiding the presence of metal ions,

usually present in samples processed with immobilized–metal affinity resins.

The three target enzymes were successfully co-expressed with two different CBMs —one from *T. maritima* Xylanase 10A (CBM9) and one from *C. Thermocellum* cellulosomal–scaffolding protein A (CBM3)— in *E. coli* M15ΔglyA strain. The resulting fusion proteins were successfully produced at high titres in a bench-scale reactor by following the same strategy applied to produce the histidine–tagged variants. Results showed that M15ΔglyA strain was able to produce CBM–fused enzymes as efficiently as histidine-tagged ones, and it was also proved that CBM domains do not have a significant negative impact in enzyme performance.

Finally, it was developed and characterised the one-step purification/immobilisation process of the different CBM–tagged variants onto low-cost and highly selective cellulosic supports, including the study of storage stability of the immobilised derivatives and the determination of maximum loading capacity of the supports. CBM9–fused enzymes proved to bind effectively to regenerated amorphous cellulose (RAC) at high loads, whereas CBM3–tagged enzymes bound with higher affinity onto microcrystalline cellulose (MC). Moreover, RAC support was also used for protein purification by fast liquid protein chromatography (FPLC), aiming to establish an alternative to metal affinity chromatography, by which only CBM9 proved to be a suitable purification tag for the three enzymes tested.

ABBREVIATIONS AND NOMENCLATURE

<i>ara</i>	Arabinose
AU	Enzyme activity unit
BSA	Bovine serum albumin
bp	DNA base pair
cAMP	Cyclic adenosine monophosphate
CAP	Catabolite activation protein
CBM	Carbohydrate-binding module
CoA	Coenzyme A
CV	Column volume
DCU	Digital control unit
DCW	Dry cell weight
DM	Defined culture medium
DTT	Dithiothreitol
EC	Enzyme commission number
EDTA	Ethylenediaminetetraacetic acid
EG	Ethylene glycol
Enz	Enzyme
EU	European Union
IBB	Institut de Biotecnologia i Biomedicina
IDA	Iminodiacetic acid
IMAC	Immobilized-metal affinity chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IY	Immobilisation yield
LA	Lithium acetate buffer
<i>lac</i>	Lactose
LB	Lysogeny broth culture medium
LIC	Ligation-independent cloning
LPS	Lipopolysaccharide
MC	Microcrystalline cellulose
MCF	Microbial cell factory
MCS	Multiple cloning site
NAD ⁺	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)
NADP	Nicotinamide adenine dinucleotide phosphate
OD ₆₀₀	Optical density at a wavelength (λ) of 600 nm
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction

PEG	Polyethylene glycol
PLA	Polylactic acid
PMSF	Phenylmethanesulphonyl fluoride
PO ₂	Dissolved oxygen
RA	Retained activity
RAC	Regenerated amorphous cellulose
RBS	Ribosome binding site
RPP	Recombinant protein production
SDS	Sodium dodecyl sulphate
ScADH	Alcohol dehydrogenase from <i>Saccharomyces cerevisiae</i>
ScPDC	Pyruvate decarboxylase from <i>Saccharomyces cerevisiae</i>
SGB	Servei de Genòmica i Bioinformàtica (UAB)
SHMT	Serine hydroxymethyl transferase
SLIC	Sequence- and ligation-independent cloning
SOP	Standard operating procedures
TB	Terrific Broth culture medium
TES	Trace elements solution
ThDP	Thiamine diphosphate
TmLDH	Lactate dehydrogenase from <i>Thermotoga maritima</i>
TPP	Thiamine pyrophosphate
Y _{X/S}	Biomass/substrate yield
ε	Molar extinction coefficient
μ	Specific growth rate
μ _{max}	Maximum specific growth rate

1. INTRODUCTION

According to the definition of the European Federation of Biotechnology (EFB,1989), biotechnology is an integrated application of natural and engineering sciences with the aim of using living organisms, cells and their component parts for products and services. In other words, biotechnology consists in the obtention of goods and services by using biological tools.

Nowadays, biotechnology represents one of the most important highly-developed technologies which enables the creation of novel products, services and industrial processes in various facets of society and yet, the origins of biotechnology date back to ancient times, being practised by agriculturalists who established better-quality species of plants and animals by methods of cross-pollination or cross-breeding, and the (unconscious) utilization of micro-organisms to produce products such as cheese, yogurt, bread, beer and wine (Bhatia and Goli, 2018).

Nevertheless, later developments in molecular biology — mainly since 1970s with the advent of recombinant DNA technologies — have led to the arisen of the term "modern biotechnology", which comprises applications far beyond breeding or fermentation, and involves genetic engineering and cell manipulation, among other techniques.

The definition of biotechnology can be further divided into different areas known as:

- Red biotechnology, which is related to medicine and human health. It studies discovery of new drugs, construction of artificial organs employing stem cells to replace/regenerate injured tissues, production of vaccines and antibiotics, regenerative therapies and new diagnostics (Sasson et al., 2005).
- Green biotechnology, which applies to agriculture and involves such processes as the development of pest-resistant grains and the accelerated evolution of disease-resistant animals. It also aims to progress in the discovery of new fertilizers and biopesticides (Silveira et al., 2005).
- Blue biotechnology, which encompasses processes in the marine and aquatic environments, such as control the proliferation of noxious water-borne organisms, explore and use marine biodiversity as a source of new products, bioprospecting the environment and using molecular biology and microbial ecology in marine organisms to obtain beneficial advances for humanity (Tramper et al., 2003).

- White biotechnology focuses on the production and processing of chemicals, materials and energy using living cells, such as yeast, fungi, bacteria, plants and enzymes for the industrial scale synthesis of products (Barcelos et al., 2018). It seeks to reduce the environmental impact moving from oil-based to sustainable processes.

In addition, another biotechnology colours such as brown (desert biotechnology), purple (patents and inventions), or yellow (insect biotechnology) can be found in some classifications (Kafarski, 2012).

Regarding **white biotechnology**, it is estimated that its implementation at industrial scale can lead to the reduction of carbon dioxide emissions, and water and energy consumption (Barcelos et al., 2018), and it can also lead to a reduction in capital and operating costs. Its global market by 2019 was estimated to represent USD 260.4 billion (IMARC Group, 2021), with a growth rate of 8.9%, including biofuels, production of chemicals such as organic acids, alcohols, amino acids, vitamins and bioplastics, some of which are produced entirely or almost entirely by white biotechnology processes.

Other products of great economic and industrial interest to white biotechnology are enzymes, among other recombinant proteins. In this context, the expression of heterologous proteins using genetic recombination is still the high point in the development and exploitation of modern biotechnology.

1.1. Brief introduction to bioprocesses

A bioprocess can be defined as a specific process that uses complete living cells or their components (e.g., enzymes) to obtain desired products. Bioprocess technology is therefore a vital part of biotechnology, that aims to produce target products at industrial scale with the desired yield and quality (purity) of the end products, as alternative to the chemical production process developed so far, taking into advantage that bioprocesses are considered to be energy saving processes, performed under mild conditions (ambient temperature and pressure), that can decrease emission of pollutants, employ carbon neutral raw materials and use organic waste as raw materials, among other characteristics.

In general, bioprocesses consist of an ordered sequence of several individual processes, which can be divided into i) upstream, ii) midstream and iii) downstream processes, that refer to preparation, production and purification, respectively.

The archetypical bioprocess is based on growth of a microorganism under conditions which encourage the production of a value-added product that can be recovered for its use. In this sort of bioprocesses, upstream stage can be defined as the entire series of steps that goes from early cell isolation and cultivation to cell banking (strain development), whilst midstream is defined as culture expansion of the cells at the desired scale until final harvest (fermentation), and the down stream part of a bioprocess refers to the steps where the cell mass from the midstream is processed to meet purity and quality requirements (separation and purification).

Bioprocess development involves identifying the robust design space for each bioproduct. The biological characteristics of cells and enzymes often impose constraints on bioprocessing; knowledge of them is therefore an important prerequisite for rational engineering design (Doran, 2013).

All areas of bioprocess development—the cell or enzyme used, the culture conditions provided, the fermentation equipment, and the operations used for product recovery—are interdependent, since improvement in one area can be disadvantageous to another. Therefore, ideally bioprocess development should proceed using an integrated approach, that

requires previous experiments to understand the interaction of parameters of the specific bioprocess.

In [Figure 1.1](#), some of the main considerations in bioprocess development are shown, including some of those related to the strain development and fermentation, as well as those for the separation and purification (recovery) of final product.

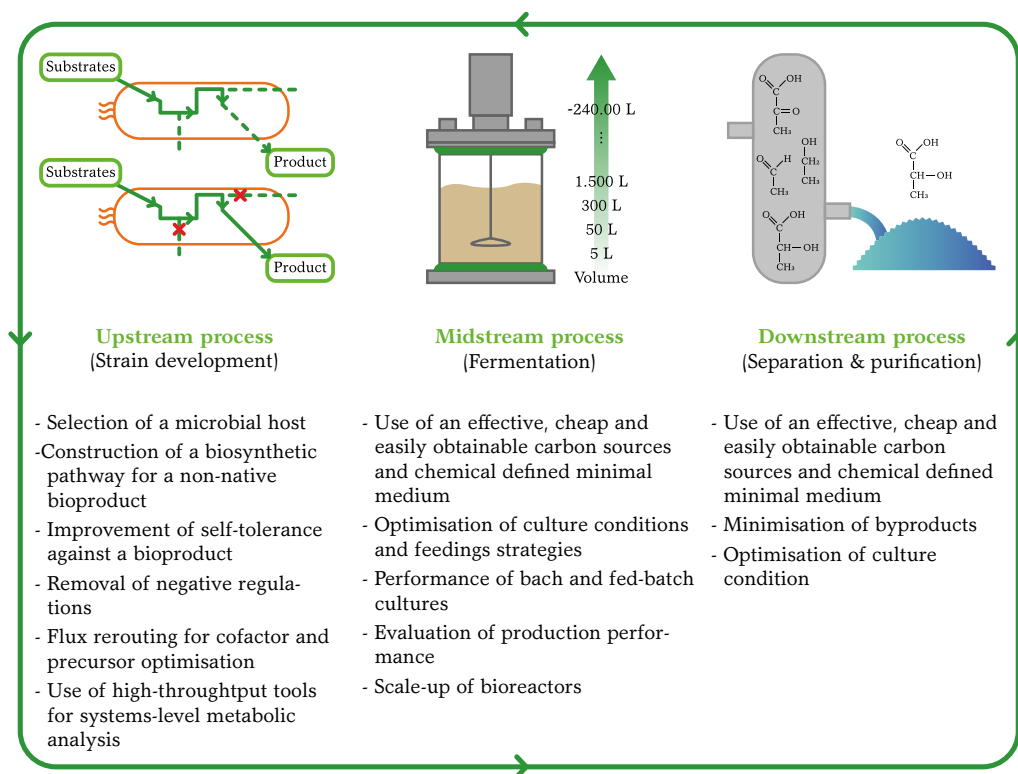


Figure 1.1. Schematic representation of the three major bioprocess stages (**top**), and some of the main variables that are taken into consideration to improve microbial strain's production performance and to reduce overall operation costs (**bottom**). Partially extracted from Nature Biotechnology (Lee and Kim, 2015).

1.2. The role of biocatalysts in bioprocesses

In chemical terms, catalysis is the acceleration of a chemical reaction by a catalyst, which reduces the necessary activation energy needed to convert a substrate to a product, without changing the extent nor

the equilibrium of the reaction. The catalyst is not consumed or altered during the reaction so, theoretically, it can be used indefinitely to convert the substrate into product.

Biocatalysis is referred to the use of a biological catalysts (enzymes, whole cells, etc.) for the transformation of natural and non-natural compounds. Over the years, biocatalysis has emerged as an alternative to the classic chemical synthesis established in the industry for decades. The environmental benefits associated to the use of biocatalysts includes lower amounts of generated wastes and more efficient use of resources and energy, among others. However, this is limited by the stability of the catalyst, that is, its capacity to retain its active structure through time at the conditions of reaction (Illanes, 2008).

1.2.1. The use of enzymes

Enzymes are a well-known kind of biocatalysts consisting in proteins, often conjugated with other molecules (metal ions, carbohydrates, lipids, etc.), which are crucial for life, as they are involved in all reactions that took place inside a living organism. Not only enzymes have an important role on life but also on most of bioprocesses performed to produce foods and beverages, improvement of detergents and textile bleaching, among others (Lobedanz et al., 2016).

The principal enzymes of industrial interest are used primarily for the conversion of starch and in processes such as baking, fruit processing, and brewing. Classes such as amylases, proteases, oxidases and lipases may be cited as examples. The global enzyme market was valued at USD 2.9 billion in 2008, rising to USD 4.4 billion in 2015, and reaching a value of USD 10.7 billion in 2020, which is expected to expand at a compound annual growth rate of 6.5% from 2021 to 2028 (Grand View Research, 2021).

The numerous advantages of enzymes led to its increased utilization in the synthesis of a vast scope of target molecules over the last decades, especially when processes with high regio- and enantio-selectivity are required (Robinson, 2015; Hedstrom, 2010). However, many industries are still reluctant to adopt enzyme biocatalysts in large-scale production processes since the conditions under which these enzymes are expected

to operate in industry are also very often far from those found in nature, further affecting their activity and stability (Smith et al., 2015).

Therefore, since the performance of biocatalyst and its production contribute to the final operating cost, it is advisable to explore diverse strategies in order to improve biocatalyst yield, especially in those bioprocesses focused on producing low value products, such as bulk chemicals, rather than complex molecules such as therapeutical drugs and fine chemicals, since operating costs must be minimised to grant the economic viability of the bioprocess.

In this context, process intensification is of key importance, since it can bring significant benefits in terms of process and chain efficiency, lower capital and operating expenses, higher quality of products, less wastes and improved process safety (Dimian et al., 2014). In the case of bioprocesses developed to obtain bulk chemicals by using enzymes, process intensification can be achieved through different approaches:

- Biocatalyst modification (protein engineering) via rational design or directed evolution, which is commonly used to create enzymes able to i) perform biotransformations under unnatural conditions, such as in organic solvents or reaction solutions with high concentrations of substrate or product, to ii) improve enzyme stability under the desired reaction conditions or even iii) to modify enzyme specificity towards substrates of interest (Cherry and Fidantsef, 2003; Wu et al., 2021).

- Improved heterologous production of enzymes in a suitable host, what can lead to the obtention of high quantities of enzymes with the use of inexpensive culture growth mediums. The production of the required enzymes is a bioprocess itself, that can also be optimised and intensified, as discussed below.

- Enzyme purification is also an advisable strategy, that contribute to prevent secondary undesired reactions that may be caused by other molecules and compounds present in crude cell extracts, and to reduce the implementation of excessive downstream steps. Besides, isolated enzymes can be thoroughly characterised in terms of its catalytic capabilities and in terms of the necessary components and conditions for its in vitro activity and suitable applications.

- Enzyme immobilisation into larger particles of inert support materials, which aims to promote a simple separation of the biocatalyst from the product stream, allowing not only to recover the enzymes for subse-

quent recycle but also to improve their operational stability and to control the microenvironment of the biocatalyst (Ramesh et al., 2016).

1.3. *Escherichia coli* as a platform for recombinant protein production

A wide range of expression systems exist for recombinant protein production (RPP), which are mainly based on prokaryotic (bacteria) cells, or eukaryotic systems—including fungi, yeast, insect and mammalian cells, among others—.

Nevertheless, all of them have inherent advantages and drawbacks (Waegeman and Soetaert, 2011), which are listed in [Table 1.1](#), so therefore, the selection of an appropriate expression system for RPP is probably one of the most important and challenging part of an entire bioprocess. An optimal expression system can be selected only if the productivity, bioactivity, purpose, and physicochemical characteristics of the interest protein are taken into consideration, together with the cost, convenience and safety of the system itself (Yin et al., 2007).

Table 1.1. Advantages and drawbacks of most frequently used recombinant protein production host systems. Based on paper review (Yin et al., 2007).

Host system	Advantages	Disadvantages
Bacteria	<ul style="list-style-type: none"> Large knowledge base Simple scale-up Low cost and time Easy operation High gene expression, tightly regulated and controlled 	<ul style="list-style-type: none"> Protein solubility issues, refolding usually required No post-translational modifications High endotoxin content in Gram-negative bacteria
Yeast	<ul style="list-style-type: none"> Eukaryotic protein processing, facilitates glycosylation and disulphide bonds Easy to scale up to fermentation Simple media requirements Simple recovery of secreted product 	<ul style="list-style-type: none"> Growth conditions may require optimization Gene expression less easily controlled Glycosylation not identical to mammalian systems

Fungi	<p>Low cost</p> <p>Moderate to high production rate</p> <p>Large quantities of product secreted</p> <p>Rather easy genetic manipulation</p>	<p>Refolding sometimes required</p> <p>Glycosylation not identical to mammalian systems</p>
Insect cells	<p>Near mammalian protein processing</p> <p>Greater growth yield than mammalian systems</p>	<p>More demanding culture conditions than bacteria or yeast</p> <p>Low growth rate</p>
Mammalian cells	<p>Highest level protein processing</p> <p>Same biological activity than native proteins</p>	<p>More demanding culture conditions than bacteria or yeast</p> <p>Low growth rate</p> <p>Manipulated cells can be genetically unstable</p> <p>Low expression levels</p>

Escherichia coli (*E. coli*) is a Gram-negative, facultative anaerobic and non-sporulating bacterium, commonly found in the lower intestine of warm-blooded organisms. *E. coli* cells are typically rod-shaped and measures about 2–6 μm long and 0.25–1 μm in diameter, with a cell volume of 0.6–0.7 μm^3 . It was first described and isolated by Theodor Escherich in 1885 (Allocati et al., 2013).

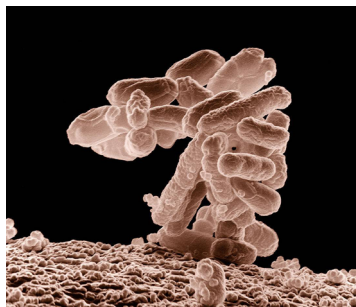


Figure 1.2. Low-temperature electron micrograph of a cluster of *E. coli* bacteria, magnified 10^4 times. Extracted from Agricultural Research Service, the research agency of the United States Department of Agriculture.

E. coli is probably the most widely used prokaryotic system for the synthesis of heterologous proteins thanks to their many advantages, such high growth rates and high production yields—recombinant protein can represent up to 50% of soluble proteins—achieved by using inexpensive culture media, the wide knowledge of its metabolism and genome and its easy transformation ability with exogenous DNA, among others (Chen, 2012; Schumann and Ferreira, 2004; Tripathi, 2016).

Several *E. coli* strains have been established in industrial bioprocesses, mainly *E. coli* BL21 and K-12 derived strains (Terpe, 2006). Additionally, thanks to its DNA structure and location, the bacterial genome can be easily modified aiming to optimize the protein overexpression yield, improve protein folding or minimize the side products formation.

Besides, along with chromosomal DNA, bacteria may also contain expression vectors—generally plasmids—which are extra-chromosomal and self-replicating DNA elements that contain the gene or genes of interest that encode for target products. Moreover, expression vectors used nowadays also contain multiple combinations of replicons, selection markers, transcriptional promoters and terminators and multiple cloning sites (Tripathi, 2016).

To date, the vast majority of recombinant protein expression is plasmid-based, primarily because higher gene dosages can be obtained compared with chromosomally integrated recombinant genes; multiple copies of the plasmid are present in the cell, e.g., between 500 to 700 copies in the pUC-derived plasmids (Waegeman and Soetaert, 2011). Furthermore, genetic manipulations are much more straightforward and less time consuming on plasmids than on the cell genome.

On the other hand, plasmid maintenance imposes a metabolic burden on the host cells, which results in a reduced growth rate and viability (Chen, 2012). Moreover, plasmid selection markers such antibiotics are commonly used to exert selective pressure to enhance stability of plasmid-bearing cultures. The potential risk of spreading antibiotics resistance genes make large scale use of antibiotics highly undesirable, what has prompted researchers to develop alternative methods to stabilise plasmid-bearing *E. coli* cultures (Vidal et al., 2008; Hägg et al., 2004).

Moreover, *E. coli* intracellular protein overexpression implies a multiple-step downstream process to recover target products, which usually

must be separated not only from the culture media mixture but also from the rest of proteins and other non-protein cell parts. Downstream steps could represent one of the major portions of total process cost (Labrou, 2014) and sample processing limitations could represent a process bottleneck.

1.3.1. Regulation of gene expression in *Escherichia coli*

A wide variety of expression vectors are commercially available to be used for RPP, which can be sorted into different groups depending on the promoter that contain; many promoters such as *lac*, *tac*, *trc*, T7, *rha*-BAD *ara*BAD, p_L , *PhyA*, and *PhoA* have been successfully used for the production of recombinant products in *E. coli* with relatively low to very high expression levels (Schleif, 2010; Terpe, 2006; Huang et al., 2012). Some of them are listed in Table 1.2.

Table 1.2. Some *E. coli* promoter systems commonly used for heterologous protein production.

Promoter system	Base of regulation / induction	Reference
T7 RNA polymerase	Engineered <i>lac</i> promoter that utilizes T7 RNA polymerase. High-level overexpression induced by adding IPTG (0.05–2.0 mM)	(Studier and Moffatt, 1986)
T5 <i>lac</i> promoter	Comprised by the fusion between the strong coliphage T5 promoter P_{N25} and the <i>lac</i> operator. Transcription from this promoter can be controlled by the <i>lac</i> repressor (tightly regulated). High-level overexpression induced by adding IPTG (0.05 - 2.0 mM)	(Gentz and Bujard, 1985) (Bujard et al., 1987)
Phage promoter p_L (λ)	p_L strong bacteriophage λ promoter is regulated by the temperature-sensitive <i>cI857</i> repressor. Controlled expression by shifting the temperature of cultures from 30 to 42°C.	(Elvin et al., 1990)

<i>araBAD</i> promoter (P _{BAD})	Tightly regulated by <i>araC</i> repressor, can fine-tune expression levels in a dose-dependent manner. Variable from low to high overexpression level induced by adding L-arabinose (0.001 - 1.0 %)	(Guzman et al., 1995a)
<i>tetA</i> promoter/operator	The <i>tetA</i> promoter is tightly regulated by the <i>tetR</i> gene, encoding the repressor, which regulates both its own synthesis and transcription of <i>tetA</i> . Variable from middle to high overexpression level induced by anhydrotetracycline (200 $\mu\text{g}\cdot\text{L}^{-1}$), independently on <i>E. coli</i> metabolic state.	(Skerra, 1994)

In general, a useful promoter system must be strong, has a low basal expression level (i.e., it is tightly regulated), must be easily transferable to other *E. coli* strains, and the induction must be simple and cost-effective, and should be independent on the commonly used ingredients of culturing media.

In this context, the ***lac* promoter**, comprised within the *E. coli* lactose operon, is one of the most extensively used for heterologous protein production. The lactose (*lac*) operon is required for the transport and metabolism of lactose in *E. coli* and many other enteric bacteria, allowing for the effective digestion of the disaccharide lactose when glucose is not available through the activity of a β -galactosidase enzyme (Juers et al., 2012). Gene regulation of the *lac* operon was the first genetic regulatory mechanism to be clearly understood and thus, it has become a foremost example of prokaryotic gene regulation.

The *lac* operon consists of three structural genes:

- *lacZ*, that encodes β -galactosidase. Apart from cleaving lactose to form glucose and galactose via hydrolysis, this enzyme catalyses the transgalactosylation of lactose to allolactose, and thus, the allolactose can also be cleaved to the monosaccharides.
- *lacY*, which encodes a β -galactoside permease, a transmembrane symporter that pumps β -galactosides (including lactose) into the cell using a proton gradient in the same direction.

- *lacA*, that encodes β -galactoside transacetylase, an enzyme that transfers an acetyl group from coenzyme A (CoA) to the hydroxyl group of the galactosides

In addition, *lac* operon is affected by the regulatory *lacI* gene, which encodes the repressor protein of the system, that inhibits the expression of the structural genes. The *lac* genes are organised into an operon, as shown in Figure 1.3, oriented in the same direction and co-transcribed into a single polycistronic mRNA molecule, and it is regulated by a two-part mechanism that ensures that gene transcription only occurs when necessary. Transcription of all genes starts with the binding of the RNA-polymerase enzyme to the promoter, immediately upstream of the genes.

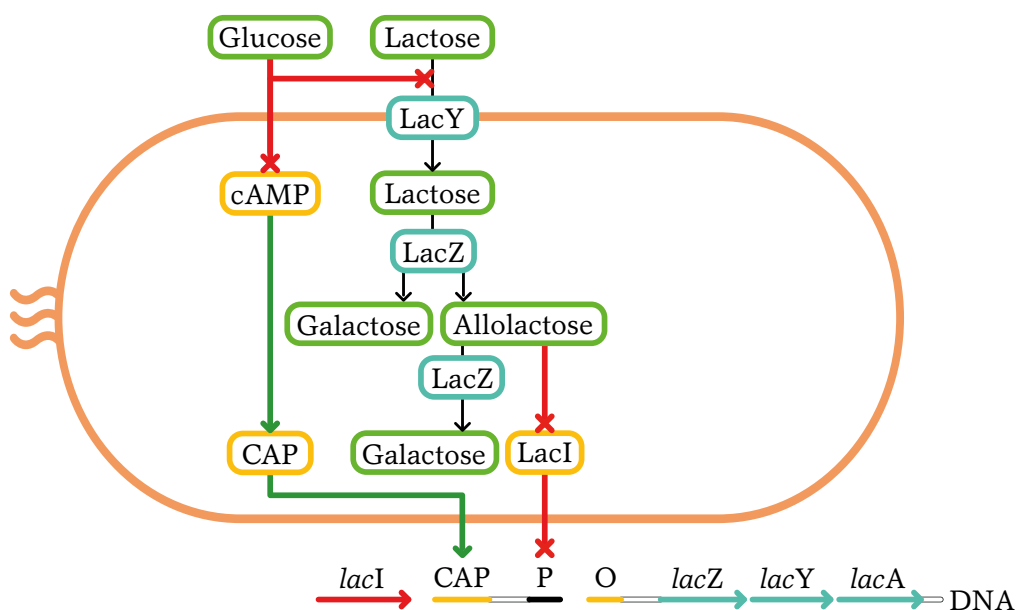


Figure 1.3. Schematic representation of the *lac* operon regulation mechanism. P, promoter region; O, operator region; cAMP, cyclic adenosine monophosphate; CAP, catabolite activation protein; *LacI*, repressor; *LacZ*, β -galactosidase; *LacY*, permease.

The first control mechanism is the regulatory response to lactose. The repressor protein encoded by *lacI* gene prevents RNA polymerase from binding to the operator of the operon in absence of lactose. The *lacI* gene coding for the repressor lies nearby the *lac* operon and is always expressed (constitutive).

When cells are grown in the presence of lactose, however, the *lac* repressor is inactivated when it is bound by allolactose, causing an allosteric shift. Thus altered, the repressor is unable to bind to the operator, allowing RNA polymerase to transcribe the *lac* genes and thereby leading to higher levels of the encoded proteins.

The second control mechanism is a response to glucose, which uses the catabolite activator protein (CAP) to greatly increase production of β -galactosidase in the absence of glucose. Cyclic adenosine monophosphate (cAMP) is a signal molecule whose prevalence is inversely proportional to that of glucose. It binds to the CAP, which in turn allows the CAP to bind to the CAP binding site, assisting the RNA polymerase in binding to the *lac* promoter.

In the absence of glucose, the cAMP concentration is high and binding of CAP-cAMP to the DNA significantly increases the production of β -galactosidase, enabling the cell to hydrolyse lactose and release galactose and glucose. The effect of glucose and lactose concentrations on the transcriptional levels is described in [Figure 1.4](#):

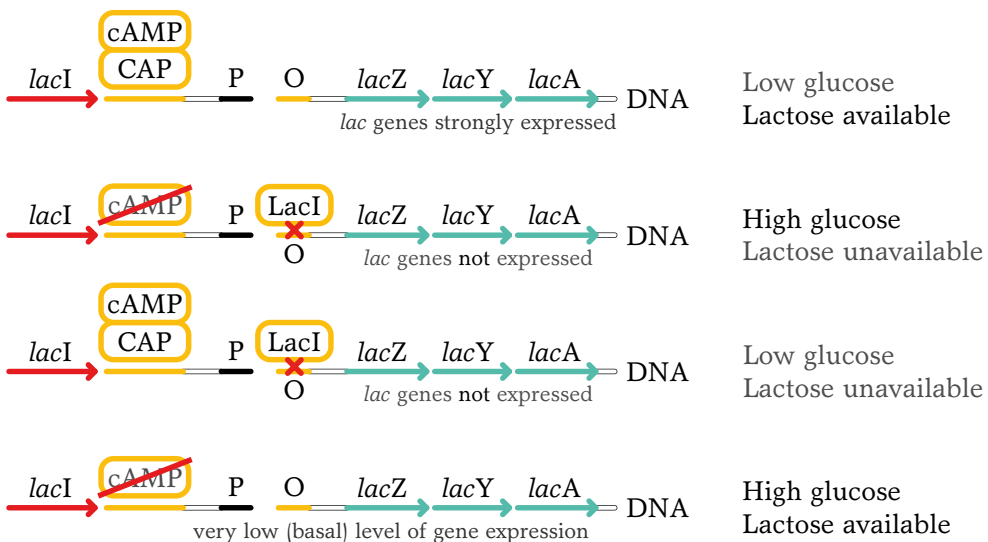


Figure 1.4. Schematic representation of the effect derived from lactose and glucose concentration on the transcriptional levels of the *lac* promoter of *E. coli*. P, promoter region; O, operator region; cAMP, cyclic adenosine monophosphate; CAP, catabolite activation protein; *LacI*, repressor.

Nevertheless, as previously mentioned, β -galactosidase is also able to transform allolactose to galactose. Hence, if allolactose was used in any *lac* operon-based expression system to induce the transcription of the genes responsible for the overexpression of the heterologous target proteins, it would be necessary to maintain allolactose (or lactose) levels during all the process. Fortunately, induction of *lac* promoter can also be achieved by adding non-hydrolysable lactose analogues, such as isopropyl- β -D-1-thiogalactopyranoside (IPTG), which only needs to be added in a single pulse.

Even if the *lac* promoter is rather weak and rarely used for high level production of recombinant proteins, many promoters were constructed from *lac*-derived regulatory element, improving the strength of the expression system by fusing the three structural genes of *lac* operon to a stronger promoter, obtained from T7 and T5 coliphages (Polisky et al., 1976; Studier and Moffatt, 1986; Bujard et al., 1987), or by generating synthetic promoters such *tac* (de Boer et al., 1983) and *trc* promoters (Amann et al., 1983).

The other promoter used in this work is the ***araBAD* (P_{BAD}) promoter**, which is present in the pBAD vectors and is a useful alternative for heterologous protein production in *E. coli*. When a gene is cloned downstream of the P_{BAD} promoter, its expression is controlled by the AraC protein, which has the dual role of repressor/activator. Expression is induced to high levels on media containing L-arabinose whilst expression is tightly shut off on media containing glucose but lacking arabinose (Terpe, 2006). In general, the *araC*- P_{BAD} promoter system allows high-level expression, tightly regulated protein expression, and very inexpensive induction.

The promoter is a part of the arabinose operon, which contains three structural genes:

- *araA*, that encodes L-arabinose isomerase enzyme, that converts intracellular L-arabinose to L-ribulose.
- *araB*, which encodes a kinase enzyme that phosphorylates L-ribulose.
- *araD*, that encodes an epimerase enzyme that that converts L-ribulose-phosphate to D-xylulose-phosphate, that goes to pentose phosphate metabolic pathway.

In *E. coli*, the P_{BAD} promoter is adjacent to the *araCp* (P_C) promoter, which transcribes the *araC* gene in the opposite direction. The cyclic

AMP receptor protein CAP binds between the P_{BAD} and P_C promoters, stimulating transcription of both when bound by cAMP (Schleif, 1992).

The schematic representation of arabinose operon and its regulation is provided in Figure 1.5:

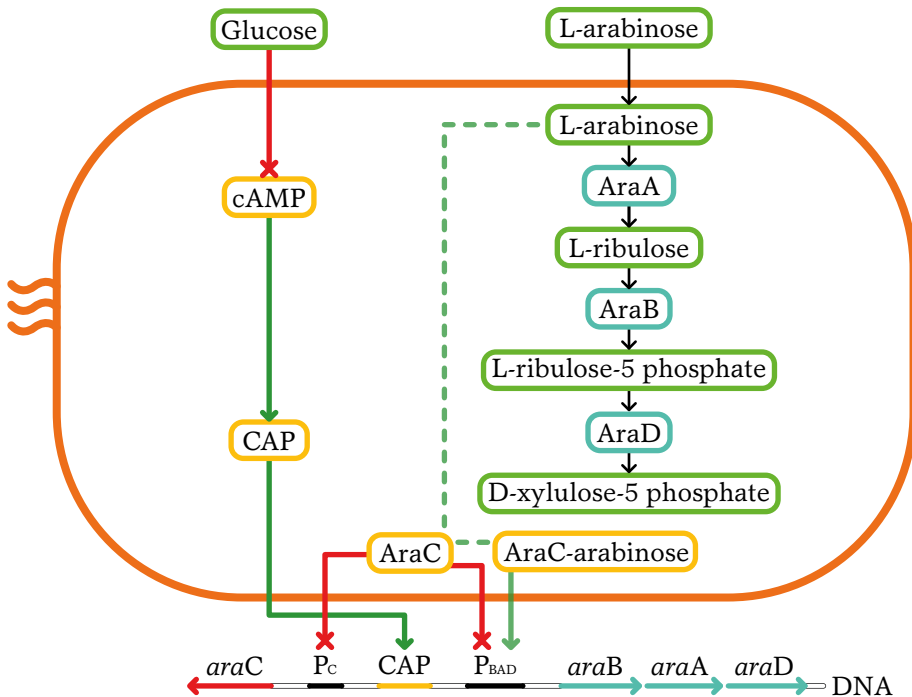


Figure 1.5. Schematic representation of the arabinose operon regulation mechanism. P_C and P_{BAD} , promoter regions; cAMP, cyclic adenosine monophosphate; CAP, catabolite activation protein; AraC, repressor; AraB, kinase; AraA, isomerase; AraD, epimerase.

Transcription initiation at the P_{BAD} promoter occurs in the presence of high arabinose, that binds to to AraC. The AraC-arabinose dimer contributes to activation of the P_{BAD} promoter. Additionally, CAP binds to two CAP binding sites and helps activate the P_{BAD} promoter, and thus, low cAMP levels (high glucose concentrations) prevent CAP from activating the P_{BAD} promoter (Guzman et al., 1995a).

Without arabinose, and regardless of glucose concentration, the P_{BAD} and P_C promoters are repressed by AraC. The N-terminal arm of AraC inte-

acts with its DNA binding domain, allowing two AraC proteins to bind to the O₂ and I₁ operator sites. The two bound AraC proteins dimerize and cause looping of the DNA, that prevents binding of CAP and RNA polymerase, avoiding the transcription of the genes (Schleif, 1992). The effect of glucose and arabinose concentrations on the transcriptional levels is depicted in Figure 1.6:

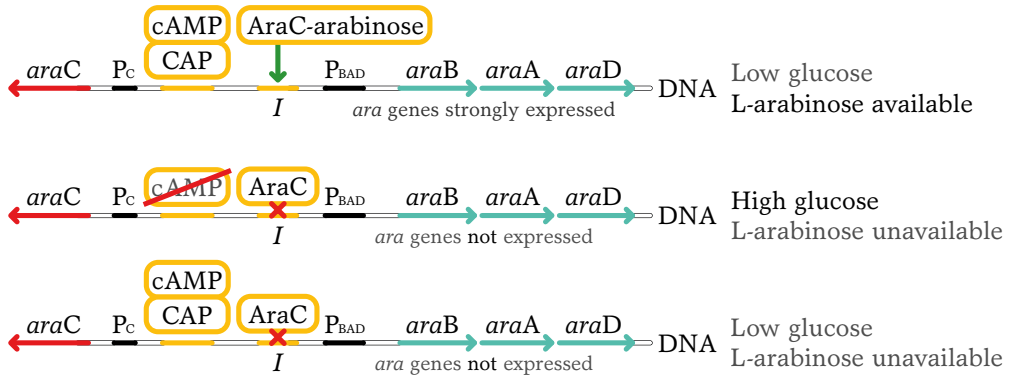


Figure 1.6. Schematic representation of the effect derived from L-arabinose and glucose concentration on the transcriptional levels of the arabinose promoter of *E. coli*. P_C and P_{BAD} , promoter regions; I , operator site; cAMP, cyclic adenosine monophosphate; CAP, catabolite activation protein; AraC, repressor.

In bacterial strains which are deleted for *ara* genes, expression of the cloned gene reaches maximal induction upon adding 0.001% (w/v) L-arabinose. However, in strains that can metabolize L-arabinose, 1% (w/v) is necessary for full induction (Mayer, 1995). Moreover, it has been proven to function under high cell density fermentation (Román et al., 2020), but the protein quality was shown to be lower than in low densities (DeLisa et al., 1999).

1.3.2. High cell–density cultures as a process intensification strategy

Once an optimal expression system has been constructed, protein production can be enhanced by increasing the production of protein per cell per unit time (specific productivity), or by increasing the cell con-

centration per unit time (cell productivity). In this context, various culture techniques have been developed for growing recombinant and non-recombinant *E. coli* strains in bioreactors, in which environmental conditions that influence bacterial cell growth —pH, concentration of dissolved oxygen (PO_2), temperature, etc.— can be controlled. The three classical operational strategies that have been followed most frequently for RPP are: i) continuous, ii) discontinuous (batch) and iii) semi-continuous (fed-batch) fermentations.

When using *E. coli* as expression system for the heterologous production of proteins, high cell-density cultures have been developed not only to improve productivity, but also to provide advantages such as reduced culture volume, enhanced downstream processing, reduced wastewater, lower production costs and reduced investment in equipment (Lee, 1996). Nevertheless, the main reason that justifies the pursuit of high cell-density cultures relies on the fact that *E. coli* does not excrete proteins into extracellular medium but tends to keep them in the cytoplasm instead, having then a limited accumulation capacity. Thus, it is necessary to reach high numbers of *E. coli* cells to obtain as much target protein as possible.

In addition, the fusion of signal peptides to target proteins might affect the mRNA stability and may cause protein degradation if targeted proteins accumulate at the translocation site of the cell wall (Owji et al., 2018). Plus, obtaining diluted products in large final volumes may also represent an important drawback for some volume-limited downstream equipment, even if avoiding cell lysis could be seen a priori as an advantage.

Regarding high cell-density cultures, fed-batch operating mode allows to reach high cell concentration levels while helps to avoid the inhibitory effect of high substrate levels —what may occur if operating in discontinuous, for example—. Cell concentrations of greater than 70 grams dry cell weight per culture litre can be routinely obtained by fed-batch cultures if a balanced nutrient medium that contains all the necessary components for supporting cell growth is used (Lee, 1996). However, this cultivation approach has several drawbacks, including limited oxygen transfer capacity, the formation of growth inhibitory by-products, and limited heat dissipation, most of them caused because of the high cell densities reached.

Probably, the main characteristic of a fed-batch process is the control of cell growth by limiting the amount of carbon source in the culture medium; the specific growth rate (μ) of the bacterial culture can be easily regulated by controlling the addition flux of a feeding medium that contains the substrate—normally glucose or glycerol—at a certain concentration. Thus, one of the main advantages that fed-batch strategy brings to *E. coli* cultivation is the avoid of by products formation, that may be produced due to an unrestricted cell growth.

E. coli cells produce the by-product acetate under anaerobic or oxygen-limiting conditions. Moreover, acetate is produced by *E. coli* cultures growing in the presence of excess glucose even under aerobic conditions (Lee, 1996; Han and Eiteman, 2019; Pinhal et al., 2019). Acetate is produced when carbon flux into the central metabolic pathway exceeds the biosynthetic demands and the capacity for energy generation within the cell (overflow metabolism). An accumulation at high levels of acetate reduces growth rate, biomass yield and maximum attainable cell densities (Phue and Shiloach, 2005; Han and Eiteman, 2019). Therefore, controlling the growth rate of the culture via substrate addition is essential.

Though, there are many strategies than can be applied to reduce the formation of acetate in high cell-density cultures, including temperature decrease, culture composition design and even microbial engineering, among others (Lee, 1996; Lee and Kim, 2015; Pinhal et al., 2019).

When fed-batch operation mode is considered for growing *E. coli* cultures at bench-scale reactors, one must consider that reaching high cell densities may take excessively long periods, given that reactors are normally inoculated at low cell densities, with the addition of volume limited shake-flask inoculum cultures. In this context, one of the simplest and easiest ways to proceed is to perform a two-step process, consisting in a first batch step followed by the fed-batch phase, allowing then to grow cell inoculum at the maximum specific growth rate (μ_{\max}) during the batch step with the nutrients initially present in the reactor, and subsequently sustaining cell growth—and therefore modulating it—during the second step by using various regimes of nutrient feed, until fermentation is complete.

Besides, substrate feeding can be added to the culture by following different approaches, such as constant flux addition, or increasing addition.

In the second case, the flux increase can be gradual, stepped, lineal or exponential. The main advantage of exponential feeding is that a certain specific growth rate can be maintained along the process.

As any other step within the bioprocess, the production yield can be enhanced by **optimising the fed-batch culture process**. Regarding *E. coli* cultures, one of the key variables that should be considered is the maximum biomass concentration that can be achieved with a certain reactor setup, since it is desirable to reach the highest number of cells, as explained above. Furthermore, another variable that should be taken into consideration, especially when using strong expression promoters, is for how long *E. coli* cells can be overexpressing the target protein before reaching a metabolic collapse caused by the addition of the inducer (whose concentration can also be optimised).

Hence, it is advisable to carry out two preliminary fed-batch cultures; the first one should be performed without inducing protein overexpression, to assess the maximum biomass levels than could be reached, and the second one should be performed by inducing protein overexpression at low cell densities, aiming to determine the effective process time until the metabolic collapse of the culture, and the maximum specific production yield —i.e., the maximum amount of target protein produced per unit of biomass— of the *E. coli* strain used. Next, the optimisation of the production process would consist in adjusting the induction time aiming to reach both maximum cell growth and metabolic collapse simultaneously.

As mentioned before, temperature is also an important variable that can be used to control cell metabolism in a simple way. By lowering the culture temperature from 37°C (considered optimum for *E. coli*) to 25–30°C, nutrient uptake and growth rate can be reduced, thus reducing the formation of toxic by-products and the generation of metabolic heat. Lowering culture temperature also reduces cellular oxygen demand, which enables higher cell-densities to be obtained without the need for pure oxygen (Lee, 1996). Furthermore, it is possible to reduce the formation of inclusion bodies for some proteins by growing recombinant cells at lower temperatures (Miret et al., 2020).

1.4. Enzyme purification and immobilisation as process intensification strategies

Biotechnology seeks to develop new protein-based applications and their commercial exploitation, in which the quality and the purity grade of final products is usually tightly related to their market value. The purity grade of synthesised proteins is normally more important in fine products rather than bulk ones, especially depending on their final use. As mentioned before, those bioprocesses focused on producing low-valued products at large scale seek to maximise biocatalyst yield to bring biotransformations closer to industrial feasibility. The use of purified enzymes and the possibility to recycle them through biocatalyst immobilisation are therefore two interesting process intensification strategies that might be considered.

1.4.1. Enzyme purification

Regarding the production of enzymes that are going to be used in biocatalytic processes, obtaining purified and isolated biocatalysts instead of using crude cell broths is an advisable strategy, helping to prevent secondary undesired reactions and reducing the implementation of excessive downstream steps in the subsequent processes. In that context, the development of efficient and effective purification methods and materials is then a requirement to bring these bioprocesses to techno-economic viability.

The various steps in the downstream processing protocol are normally divided into those that separate the protein and nonprotein parts of the mixture and those that separate the desired protein from all other proteins while retaining the biological activity and chemical integrity of the polypeptide. These last steps are typically the most laborious and difficult aspect of protein purification (Labrou, 2014). Downstream steps exploit differences in chemical, structural and functional properties between the target protein and other proteins in the crude mixture. A wide variety of protein purification methods that can be combined to generate a suitable purification scheme are available. Usually, one executes a series of purification steps, and only rarely proteins can be purified in a single step. Therefore, developing an efficient and a cost-effective single-step purification process is one of the most challenging tasks of bioprocess engineering.

Chromatography is certainly the principal operation in downstream processing, especially affinity chromatography, which is a quite effective purification method thanks to the highly selective biological interactions established between one molecule (the ligand) immobilised to a surface and another molecule (the target) present in a mobile phase as part of a complex mixture, providing a fast and often single step purification process (Urh et al., 2009; Labrou, 2014). In addition, the interaction is typically reversible, and the target molecule captured can be desorbed from the adsorbent support, recovering a highly purified product.

Many affinity interactions are used to purify enzymes, but probably one of the most typically used is immobilised-metal affinity chromatography (IMAC), in which the ligand is a metal chelator to which a metal ion has been bound (Regnier et al., 2020). E. g., supports functionalized with Co^{2+} or Ni^{2+} ions are capable to bind histidine-bearing peptides, especially those proteins with a poly-histidine tag fused to one of their ends (Crowe et al., 1996). Bound proteins are eluted by decreasing the buffer pH or increasing imidazole concentrations (Tripathi, 2016).

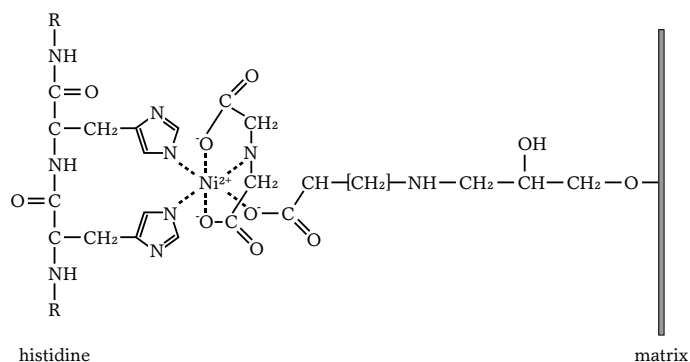


Figure 1.7. Purification of proteins using polyhistidine affinity tags and nickel-nitriloacetic acid (Ni^{2+} -NTA) functionalised matrix.

However, these kinds of resins used to purify histidine-tagged proteins contribute significantly to the final downstream cost, what brings to contemplate other alternatives, especially for industrial-scale processes in which large cell lysate volumes must be processed. Additionally, the presence of metal ions in final product, that can be released to the mobile phase, is not compatible with processes focused on pharma or food

industries, whose commercialised products must overcome strict purity requirements.

Downstream processing has thus been challenged with demands of high yields, resolving power, and cost efficiency, what has triggered remarkable developments in improvising process tools and innovative strategies for protein separation and yet, there is room for improvement, and advances in downstream techniques are of key importance, since they can have a significant economic impact in the whole bioprocess development, together with evolved upstream and midstream techniques.

1.4.2. Enzyme immobilisation

In 1971, during the first Enzyme Engineering Conference, immobilised enzymes were defined as "enzymes physically confined or localised in a certain defined region of space with retention of their catalytic activity, and which can be used repeatedly and continuously" (Khan, 2021).

Immobilised enzymes are currently the object of considerable interest, due to their benefits over free enzymes in solution. There are several reasons for the preparation and use of immobilised enzymes, in addition to a more convenient handling of enzyme preparations. Immobilised enzymes are more robust and more resistant to environmental changes. More importantly, the heterogeneity of the immobilised enzyme systems allows an easy recovery of both enzymes and products, multiple reuses of enzymes, continuous operation of enzymatic processes, rapid termination of reactions, improved operational and storage stability, and greater variety of bioreactor designs (Cao, 2011; Homaei et al., 2013).

The properties of immobilised enzyme preparations are governed by the properties of both the enzyme and the carrier material. The specific interaction between the latter provides an immobilised enzyme with distinct chemical, biochemical, mechanical and kinetic properties (Tischer and Wedekind, 1999). It stabilises structure of the enzymes, thereby allowing their applications under harsh environmental conditions (pH, temperature, organic solvents), and thus enable their uses in non-aqueous enzymology, and in the fabrication of biosensors (Dwevedi and Kayastha, 2011).

Alternatively, biocatalyst immobilisation might also refer to the whole cells with enzymes inside or outside to constitute a whole cell system.

Here, contrary to enzyme immobilisation system in which an enzyme is attached to a solid support, the target living cells are immobilised in a whole cell system with preservation of their metabolic activity (Lin and Tao, 2017).

Biocatalyst immobilisation is especially appropriate in those biotransformations whose downstream processes require less unit operations or a free-allergen product (especially useful in the food and pharmaceutical industries), and when the enzyme production represents a significant cost to the process. Besides, immobilisation facilitates to perform processes in other reactor configurations than discontinuous (batch) stirred tank reactors, enabling the operation in continuous reactors, packed bed reactors, plug-flow reactors or fluidized bed reactors, among others.

On the downsides of immobilisation, there is the loss of total or partial enzymatic activity during the attachment of the biocatalyst and the added cost that the carrier and the immobilisation process represents. Furthermore, placing the enzymes in confined regions of space, as it is a small pore, entails that the substrate/s transport is governed by mass transfer effects.

These various phenomena, referred to as diffusional limitations, can lead to a reduced reaction rate, i.e., to a reduced efficiency as compared to the soluble enzyme. In stirred tanks, external diffusion plays a minor role as long as the reaction liquid is stirred sufficiently. Further, partition effects can lead to different solubilities inside and outside the carriers. Partition must be taken into account when ionic or adsorptive forces of low concentrated solutes interact with carrier materials. The most crucial effects are observed in porous particles due to internal diffusion. Reaction rate is then limited by the concentration gradient and catalytic activity can get significantly reduced (Tischer and Wedekind, 1999).

There exist **different enzyme immobilisation methods** that can be sorted into those which use a solid pre-existent support or those based on carrier-free immobilisation. On the one hand, immobilisation methods performed without prefabricated carriers tend to become cheaper, since the enzyme molecules are immobilised to themselves with a reagent (cross-linking) or confined by different techniques (encapsulation, entrapment or use of ultrafiltration membranes). However, carrier-free techniques usually present some important drawbacks at industrial-scale processes, such as poor mechanical properties (cross-linking and encap-

sulation), high mass transfer limitations (entrapment and cross-linking), difficult recovery of the derivatives compared to enzymes immobilised onto solid supports (cross-linking, encapsulation) or non-improvement of stability (containment by ultrafiltration membranes) (Tischer and Wedekind, 1999; Illanes, 2008).

On the other hand, immobilisation techniques based on the attachment of enzymes to preexistent supports usually allow easiest reuse of enzymes in conventional reactors, since their mechanical properties are more robust, and the size of the particles is more uniform. The enzyme binding can be achieved by covalent bonds by promoting chemical reaction between the superficial amino acid residues of the enzymes and the functional groups of the support. This immobilisation method may produce multipoint strong bonds which usually stabilises the enzyme structure, resulting in an improvement of stability (dos Santos et al., 2015). However, the irreversible immobilisation also presents some drawbacks, including the loss of catalytic activity caused when the covalent bonds interact with the active site of the enzymes, and the finite use of the support once the enzyme is deactivated, among others.

Alternatively, enzyme binding can be achieved by non-covalent interactions, based on Van der Waals and hydrophobic forces, or ionic bonds. Nevertheless, non-covalently immobilised enzymes can be released from the support depending on the reaction conditions (Illanes, 2008).

Regarding the supports used to immobilise biocatalysts, knowledge of the structural characteristics is helpful for achieving a high performance biocatalyst; the affordability, the chemical composition, the mechanical strength, the hydrophilicity, the particle size, the porosity and the maximum enzyme loading are some of the crucial features to be taken into consideration, although the development of a suitable immobilisation procedure often follows empirical guidelines (trial and error strategy) (Dos Santos et al., 2015).

In addition, even if the available set of materials for enzyme immobilisation is vast and new carriers with modern functionalities continuously appear, one of the major drawbacks that still presents biocatalyst immobilisation relies on the support costs, which can increase the price of the biocatalyst more than 4-fold. (Datta et al., 2013; Basso and Serban, 2019). Therefore, the final choice to use immobilised enzymes depends on the economic evaluation of costs associated with their use versus benefits

obtained in the process. The final implementation strongly depends on the use of newly developed immobilisation supports, derived from low-cost and high-available materials, such as cellulose or activated carbon, among others.

A brief overview of the existing immobilisation methods is provided below in [Figure 1.8](#):

Immobilisation Methods

Attachment to (prefabricated) carriers:

- Covalent binding
- Ionic binding
- Adsorptive binding
- Metal binding
- (bio)Affinity binding

Crosslinking:

- Direct crosslinking
- Crystallisation + crosslinking
- Co-crosslinking
- Copolymerisation after chemical modification

Encapsulation / inclusion:

- Membrane devices
- Microcapsules
- Liposomes / reversed micelles
- Organic solvents

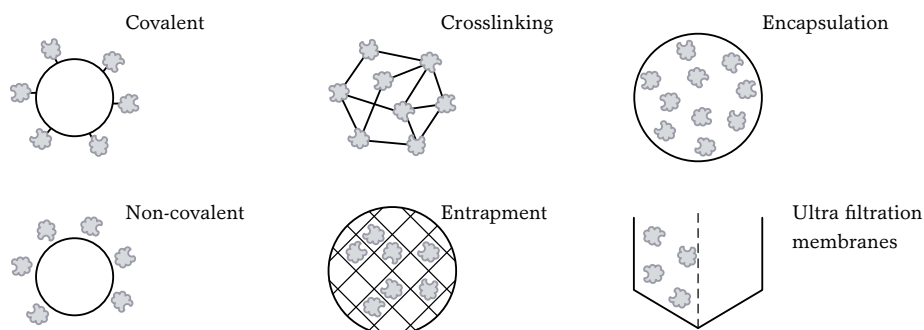


Figure 1.8. Overview of the different immobilisation methods.

1.5. Carbohydrate–Binding Modules

Carbohydrate–Binding Modules (CBM) are well–described as non-catalytic domains involved in substrate binding of many carbohydrate–active enzymes (Levy and Shoseyov, 2002), which have shown high affinity for a wide range of polysaccharides. The requirement of CBMs existing as modules within larger enzymes sets this class of carbohydrate–binding protein apart from other non-catalytic sugar binding proteins such as lectins and sugar transport proteins.

CBMs were previously defined as cellulose-binding domains (CBDs), based on the initial discovery of several modules that bound cellulose, and were classified depending on their amino acid similarity (Tomme et al., 1988; Gilkess et al., 1988). However, additional modules in carbohydrate-active enzymes are continually being found that bind other carbohydrates apart from cellulose, hence the need to reclassify these polypeptides using more inclusive terminology (Boraston et al., 2004).

Currently, thousands of CBMs have been divided into 89 different families based on amino acid sequence, binding specificity, and structure (Oliveira et al., 2015). To date, more than $3 \cdot 10^5$ CBM domains are included in these families, plus other 2240 non-classified modules.

In recent years the practical use of CBMs has been established in different fields of biotechnology, and the number of published articles and patents is constantly on the rise. Three basic properties have contributed to CBMs being perfect candidates for many applications: i) CBMs are usually independently folding units and therefore can function autonomously in chimeric proteins; ii) the attachment matrices are abundant and inexpensive and have excellent chemical and physical properties; and iii) the binding specificities can be controlled, and therefore the right solution can be adapted to an existing problem (Shoseyov et al., 2006).

CBMs have considerable capacity for hydrogen-bond formation between their polar residues and the hydroxy groups of carbohydrates. Indeed, an extensive network of direct and indirect hydrogen bonds is formed between the protein and the carbohydrate.

Nevertheless, not all CBM domains can bound any carbohydrate compound. In this context, another useful classification of CBMs was proposed, in which these protein modules have been grouped into three types (Figure 1.9) (Boraston et al., 2004; Armenta et al., 2017):

- Type A (surface-binding) includes members of CBM families 1, 2a, 3, 5 and 10 that bind to insoluble, highly crystalline cellulose and/or chitin. Their binding sites are planar and rich in aromatic amino acid residues creating a flat platform to bind to the surfaces of crystalline polysaccharides and show little or no affinity for soluble carbohydrates. The interaction of Type A modules with crystalline cellulose is associated with positive entropy, demonstrating that the thermodynamic forces that drive the binding of CBMs to crystalline ligands are relatively unique among carbohydrate-binding proteins.

- Type B (glycan–chain–binding) are the most abundant form of CBMs reported to date and include examples from families 2b, 4, 6, 15, 17, 20, 22, 27, 28, 29, 34 and 36. Type B binding sites appear as extended grooves or clefts comprised of binding subsites, which typically interact with individual glycan chains rather than crystalline surfaces, showing increased affinities for penta- and hexa-saccharides and negligible interaction with oligosaccharides with a degree of polymerization of three or less.

- Type C (small–sugar–binding) currently include examples from families 9, 13, 14, 18 and 32. Type C binding sites are short pockets that bind optimally to mono-, di- or tri-saccharides, as well as the reducing ends of small sugars. Type C CBMs can bind to several cellulose allomorphs and the insoluble fraction of oat spelt xylan.

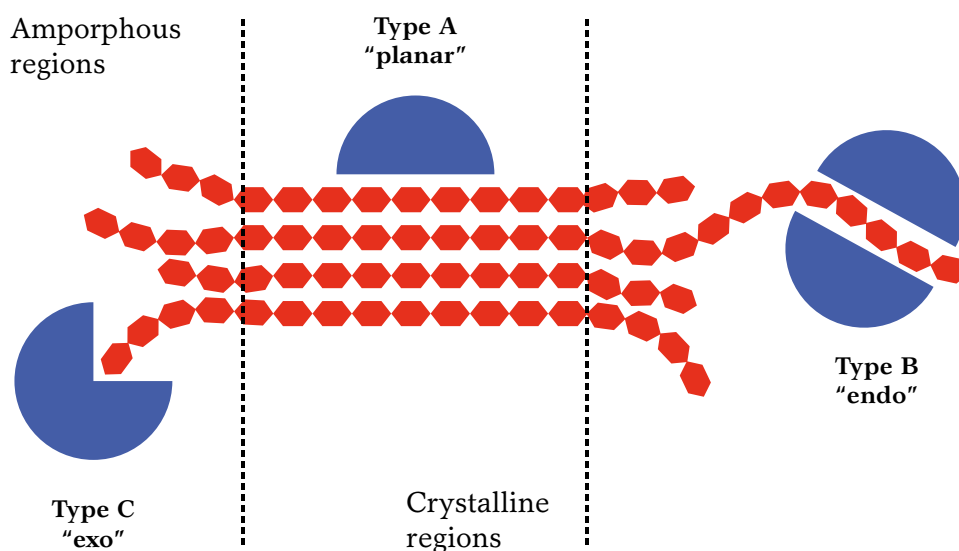


Figure 1.9. Schematic representation of different CBM types binding with different regions of a polysaccharide substrate. Extracted from CAZypedia (Lammerts van Bueren and Ficko-Blean, 2021).

1.5.1. Carbohydrate–Binding Modules as immobilisation tags

Some reports have affirmed the feasibility of employing a CBM as an affinity tag for enzyme immobilisation and processing. In those studies carbohydrates were used as an affinity support for enzyme immobili-

sation, with high capacity, while retaining enzymatic activity; in some instances, increased enzymatic activity was reported (Kauffman et al., 2000; Levy and Shoseyov, 2002; Oliveira et al., 2015). Other studies have shown that a CBM serving as a fusion partner may have additional values. In the expression of CBM–lipase fusion protein in yeast, for example, it was shown that CBM also enhanced secretion (Ahn et al., 2004).

Moreover, surface-exposed CBMs can be an efficient means of whole-cell immobilisation. Whole-cell immobilisation by cellulosic material was first demonstrated when an *E. coli* surface-anchored CBM, was attached to cellulose. The cells bound tightly to cellulose at a wide range of pHs, and the extent of immobilisation was dependent on the amount of surface anchored CBM (Levy and Shoseyov, 2002).

The dominant impediment to the use of current commercially available generic affinity tags in large-scale bioprocessing is cost. Therefore, those affinity tags that bind to an inexpensive, chemically and hydrodynamically robust chromatographic resins are highly desirable. Hence, the main advantage of using CBMs for enzyme immobilisation is the substitution of expensive functionalised carriers for a more affordable immobilisation supports such as cellulose, what can bring biotransformations closer to economically feasibility, specially at industrial scale, where large amounts or immobilisation resins are required. Moreover, cellulosic supports are usually inert and show good mechanical properties.

1.5.2. Carbohydrate–Binding Modules as affinity purification tags

As mentioned, bioprocessing is the major application for CBMs, given that large-scale recovery and purification of biologically active molecules continue to be challenges for many biotechnological products.

Biospecific affinity purification (affinity chromatography) has become one of the most rapidly developing divisions of immobilised affinity ligand technology. Affinity chromatography is a quite effective purification method, thanks to the highly selective biological interactions established between one molecule (the ligand) immobilised to a surface and another molecule (the target) present in a mobile phase as part of

a complex mixture, providing a fast and often single step purification process (Urh et al., 2009).

Many protein entities have been expressed fused to CBMs, establishing CBMs as high-capacity purification tags for the isolation of biologically active target peptides at relatively low cost. The interaction established between CBMs and carbohydrates are typically reversible, and the target CBM–fused molecules captured can usually be desorbed from the adsorbent support, recovering a highly purified product (Shoseyov et al., 2006).

For example, the affinity of a family–9 CBM (CBM9, [Figure 1.10](#)) for mono- and disaccharides has important consequences for its practical application as an affinity tag. Glucose and cellobiose can be used to desorb CBM9 from cellulose and, thus, the low cost and availability of cellulose and these competitive eluents make CBM9 an attractive candidate for use as a purification tag. Many CBMs have been used as purification tags thanks to their highly specificity. The most frequently used CBMs are those that bind to cellulose, especially CBM3 (Morag et al., 1995; Wan et al., 2011; Myung et al., 2011) and CBM9 (Boraston et al., 2001; Kavosi et al., 2004; Ye et al., 2011), but many others have been studied (Sugimoto et al., 2012; Wang et al., 2012).

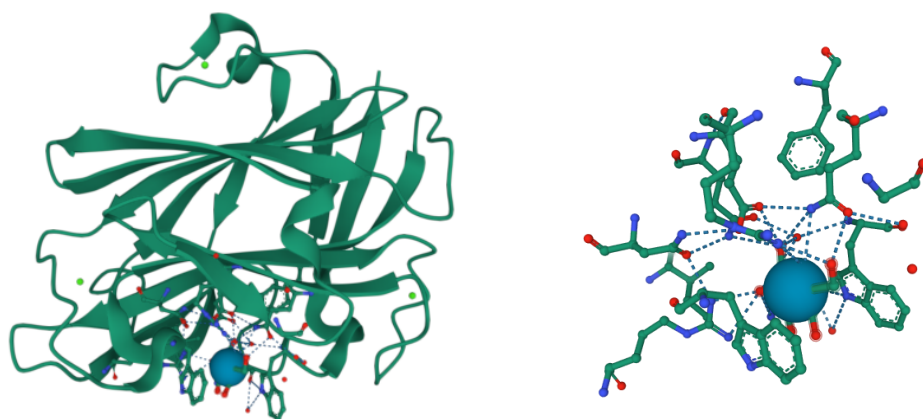


Figure 1.10. Schematic representation of a family–9 CBM (CBM9) of an endo- β -1-4-xylanase A from *T. maritima* (*TmXyn10A*) (**left**), and the interactions established between the amino acid residues and carbohydrate (**right**). Extracted from Protein Data Bank (PDB 1I8A) through Mol* Viewer (Sehnal et al., 2021).

2. CONTEXT AND OBJECTIVES

2.1. Context of the thesis work

The research work performed along this PhD thesis has been included in the BIOCON-CO₂ project (grant agreement number 761042), which is a four-year project financed by the European Commission under the program Horizon 2020. BIOCON-CO₂ stands for **BIO**technological processes based on microbial platforms for the **CON**version of **CO**₂ from iron steel industry into commodities for chemicals and plastics.

The thesis work described in the following sections has been performed at the Universitat Autònoma de Barcelona (UAB, Bellaterra, Spain), into the Bioprocess Engineering and Applied Biocatalysis Research Group, which has an extended background in biochemical engineering.

2.1.1. The Research Group

Bioprocess Engineering and Applied Biocatalysis Research Group (ENG-4BIO) is located at Department of Chemical, Biological and Environmental Engineering of the Escola d'Enginyeria from UAB. The research work of the group is mainly focused on process development for recombinant protein production in *E. coli* and their subsequent use in biocatalysis. Moreover, the group has acquired an increasing know-how in the development of enzymatic synthesis processes to produce biologically active products and in biocatalyst immobilisation.

Previous works within the group related to recombinant protein production consisted in:

- Media formulation for aldolase production (Durany et al., 2004) (Vidal et al., 2008).
- Implementation of glucose-limited fed-batch operation processes (Pinsach et al., 2008).
- Development of a model of the whole recombinant production process with high cell density *E. coli* cultures (Calleja et al., 2014) (Calleja et al., 2016).
- Development of an expression system based on glycine auxotrophy to ensure plasmid stability avoiding antibiotic supplementation (Vidal et al., 2008) (Pasini et al., 2016).

Likewise, most recent research related to enzyme biocatalysis and biocatalyst immobilisation consisted in:

- Development of a hybrid enzyme–nanoparticle complex obtained by enzyme immobilisation onto magnetic particles (Masdeu et al., 2018).
- Trimethyl- ϵ -caprolactone synthesis with an immobilised glucose dehydrogenase and an immobilised cyclohexanone monooxygenase (Solé et al., 2019, a).
- Co-immobilisation of a monooxygenase and a glucose dehydrogenase on different supports for application as a self-sufficient oxidative biocatalyst (Solé et al., 2019, b).
- Enzymatic synthesis of a statin precursor by immobilised alcohol dehydrogenase with a coupled cofactor regeneration system (García-Bofill et al., 2021, a).
- Biocatalytic synthesis of vanillin by an immobilised eugenol oxidase (García-Bofill et al., 2021, b).

2.1.2. BIOCON-CO₂ project

BIOCON-CO₂ has a consortium of recognized industry experts and leading academic organisations, comprised of eighteen industrial and academic partners based in seven countries all over Europe, as well as an industry partner from Israel. The consortium comprises five SMEs (small and medium enterprises), five large industries, four research and technology organisations and four universities (Figure 2.1), being the UAB one of them.



Figure 2.1. BIOCON-CO₂ consortium partners. UAB pointed out in red.

The main objective of BIOCON-CO₂ project is to develop and validate in industrially relevant environment a flexible platform to biologically transform CO₂ into added-value chemicals and plastics. The project aims to capture at least 4% of the total market share at medium term (1.4 Megatonnes CO₂ per year) and 10% at long term (3.5 Mtonnes) contributing to reduce EU dependency from fuel oils and support the EU leadership in CO₂ reuse technologies. The versatility and flexibility of the platform, based on three main stages—CO₂ solubilisation, bioprocess and downstream—, has been assessed by developing several technologies and strategies for each stage that will be combined as puzzle pieces. The specific objectives of the project have been divided into ten work packages (WP).

BIOCON-CO₂ has attempted to develop four microbial cell factories (MCFs) based on low-energy biotechnological processes using CO₂ from iron and steel industry as a direct feedstock to produce commodities with application in chemicals and plastics sectors using different biological systems:

- Clostridia fermentative process for C3-C6 alcohols production based on anaerobic metabolism.
- *Oligotropha carboxidovorans* fermentative process for 3-hydroxypropionic acid (3-HP) production based on aerobic metabolism.
- One-pot biocatalytic production of formic (methanoic) acid by using recombinant resting *E. coli* cells.
- One-pot biocatalytic synthesis of L-lactic (2-hydroxypropanoic) acid by using a multi-enzymatic system.

The technologic, socio-economic and environmental feasibility of the processes has been assessed to ensure their future industrial implementation, replicability and transfer to other CO₂ sources, such as gas streams from cement and electricity generation industries.

BIOCON-CO₂ aims to overcome the current challenges of the industrial scale implementation of the biotechnologies routes for CO₂ reuse by developing engineered enzymes, immobilisation in nanomaterials, genetic and metabolic approaches, strain acclimatization, engineered carbonic anhydrases, pressurized fermentation, trickle bed reactor using advanced materials and electrofermentation.

The main role of ENG4BIO Group within BIOCON-CO₂ project is to develop a **one-pot biocatalytic process to obtain L-lactic acid** from

CO₂ and ethanol using a multienzymatic system (Figure 2.2), which involves the action of three different enzymes in a one-pot reaction:

- Alcohol dehydrogenase from *Saccharomyces cerevisiae* (ScADH), responsible for converting ethanol into acetaldehyde by an oxidative reaction
- Pyruvate decarboxylase from *S. cerevisiae* (ScPDC), used to transform acetaldehyde into pyruvate by adding a CO₂ molecule.
- Lactate dehydrogenase from *Thermotoga maritima* (TmLDH), which reduces pyruvate to lactic acid.

The system includes the recycling of nicotinamide adenine dinucleotide (NAD) cofactor, which works as electron donor/acceptor in reduction-oxidation (redox) reactions. At first, one molecule of NAD⁺ is reduced to NADH to allow the oxidation of ethanol and afterwards, one NADH molecule is required to reduce one molecule of pyruvate into lactate (Figure 2.2). This way, the cofactor is recycled, avoiding the need of adding it continuously along the reaction time.

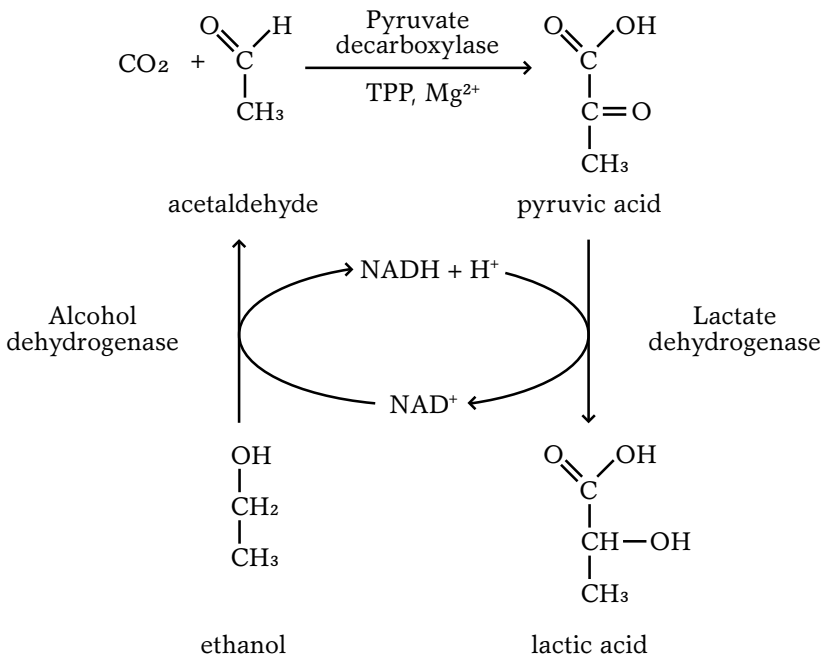


Figure 2.2. Schematic representation of the multienzymatic system designed for L-lactic acid production by using CO₂ and ethanol as substrates.

The development of the system described is included inside BIOCON-CO₂ WP5. However, the role of UAB did not only include the development of this system itself but it also included the production and immobilisation at up-scale of the enzymes involved in, aiming to obtain high quantities of robust biocatalysts (reusable and stable) in a self-sufficient way.

Finally, UAB was also involved in the process of lactic acid polymerisation to obtain polylactic acid (PLA), based on esterification reaction by using a lipase. This part of the research was included in WP6.

2.1.3. Enzymes involved in L-lactic acid production

Alcohol dehydrogenases (ADHs, EC 1.1.1.1) constitute a large family of oxidoreductase enzymes responsible for the reversible oxidation of primary alcohols into their corresponding aldehydes and the oxidation of secondary alcohols into ketones, with the concomitant reduction of NAD or NADP cofactors (Figure 2.3). These enzymes have been identified not only in yeasts, but also in several other eukaryotes and even prokaryotes (De Smidt et al., 2008).

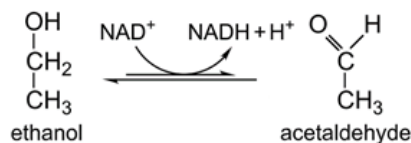


Figure 2.3. Schematic representation of alcohol dehydrogenase natural reaction.

The alcohol dehydrogenase from *Saccharomyces cerevisiae* (*ScADH*) have been studied intensively for over half a century; most of the work on yeast ADH genetics is due to the analysis of Christel Drewke and Michael Ciriacy, who revealed the existence of four different isoenzymes in yeast (ADH 1, 2, 3 and 4) (Drewke and Ciriacy, 1988). Later studies confirmed the existence of three more.

Yeast ADH1 is a tetramer of four identical subunits —homotetramer— with 347 amino acid residues each and a total calculated mass of 147396 Da (almost 150 kDa, 37 kDa per monomer).

Zinc atoms are essential for maintaining the quaternary structure of the enzyme and both zinc and the coenzyme are bound at, or near to, each of the four reactive cysteines (Figure 2.4) (Leskovac et al., 2002). The important role of the zinc atom in alcohol oxidation is to stabilise the alcoholate ion for the hydride transfer step in the reverse direction, working as an electron attractor, which gives rise to an increased electrophilic character of the aldehyde, consequently facilitating the transfer of a hydride ion to the aldehyde (Leskovac et al., 2002; Baker et al., 2009).

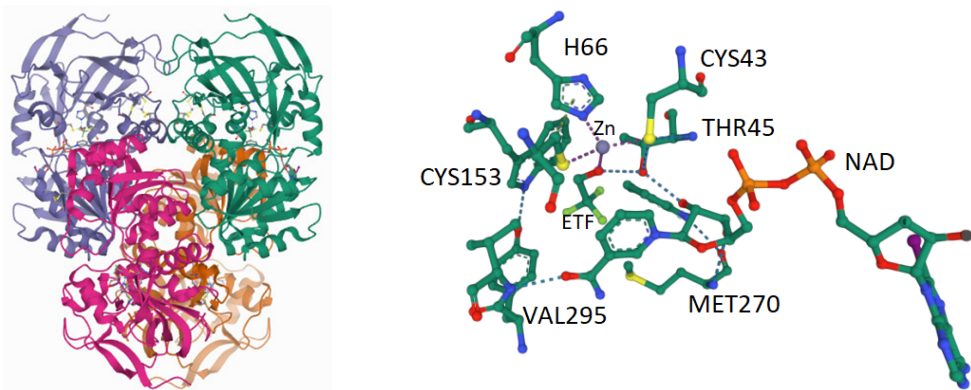


Figure 2.4. X-ray diffraction image (2.40 Å) of alcohol dehydrogenase 1 from *S. cerevisiae* (ScADH) in its homotetrameric structure (left), and the interactions of one of the active sites (right), using trifluoroethanol (ETF) as substrate, ligated to the catalytic zinc (Zn). Extracted from Protein Data Bank (PDB 4W6Z) through Mol* Viewer (Sehnal et al., 2021).

Pyruvate decarboxylase (PDC, EC 4.1.1.1) catalyse the non-oxidative decarboxylation of pyruvate yielding acetaldehyde and carbon dioxide during alcoholic fermentation (Figure 2.5) (König, 1998). Most of the produced acetaldehyde is subsequently reduced to ethanol, but some is required for cytosolic acetyl-CoA production for biosynthetic pathways. Nevertheless, some PDC isoenzymes —PDC1 included— are also able to decarboxylate more complex 2-oxo acids that comes from transaminated aminoacids, during amino acid catabolism (Lehmann et al., 1973) (Dickinson et al., 2003). Apart from yeasts, PDC enzymes have also been found in plant seeds and a few bacteria.

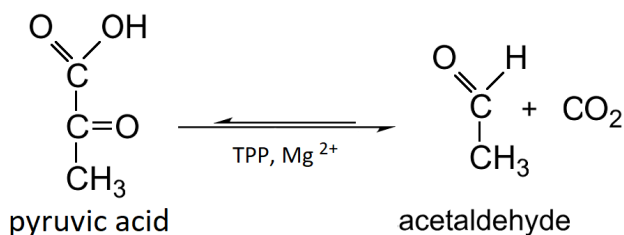


Figure 2.5. Schematic representation of pyruvate decarboxylase natural reaction.

All the PDCs from yeast and bacteria studied so far are tetramers in the native state and built up of identical or almost identical subunits with molecular masses of about 60 kDa. Each PDC1 monomer from *S. cerevisiae* (ScPDC) weights specifically 61495 Da (61.5 kDa, 563 amino acid residues), constituting a 246 kDa homotetramer, and was discovered by Neuberger and Karczaga and subsequently isolated and partially purified (Boiteux and Hess, 1970; Arjunan et al., 1996).

The catalytic activity of PDC depends on the presence of the cofactor thiamine diphosphate (ThDP), which is bound mainly via a divalent metal ion (magnesium in most cases) to the protein moiety at the interface of two subunits (Figure 2.6). PDC is a very powerful tool in the enzymatic synthesis of chiral amines by combining it with transaminases when alanine is used as amine donor (Alcover et al., 2019).

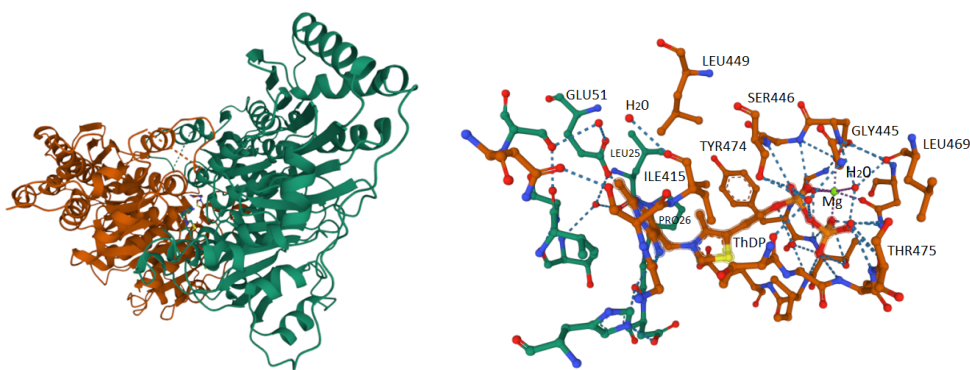


Figure 2.6. X-ray diffraction image (2.30 Å) of pyruvate decarboxylase 1 from *S. cerevisiae* (ScPDC) dimeric structure (left), and the substrate-binding region of one subunit (right), based on the interactions established between magnesium (Mg), water (H₂O), thiamine diphosphate (ThDP) and the amino acid residues. The space channel generated by the interaction of the two binding regions allows the entry of substrate. Extracted from Protein Data Bank (PDB 1PVD) through Mol* Viewer (Sehnal et al., 2021).

Lactate dehydrogenase (LDH, EC 1.1.1.27) catalyses the last step in anaerobic glycolysis, consisting in the conversion of pyruvate to lactate, with the concomitant oxidation of NADH to NAD⁺ (Figure 2.7) (Ostendorp et al., 1996).

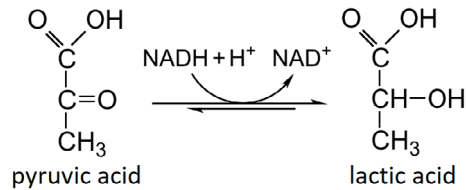


Figure 2.7. Schematic representation of lactate dehydrogenase natural reaction.

In general, with few exceptions, the protein has been found to be a homotetramer with a subunit molecular mass of approximately 35 kDa. LDH from *T. maritima* (*Tm*LDH) represents a 144 kDa homotetramer, composed of 4 identical subunits of 319 amino acid residues (34994 Da), which is found to be the most thermostable LDH isolated so far. NADH coenzyme and the effector fructose 1,6-bisphosphate contribute to confer a long-term stability of the enzyme at temperatures up to 80°C (Hecht et al., 1989). Surprisingly, apart from the major tetrameric fraction, about 10% of the recombinant protein occurs as a stable homo-octamer (Ostendorp et al., 1996).

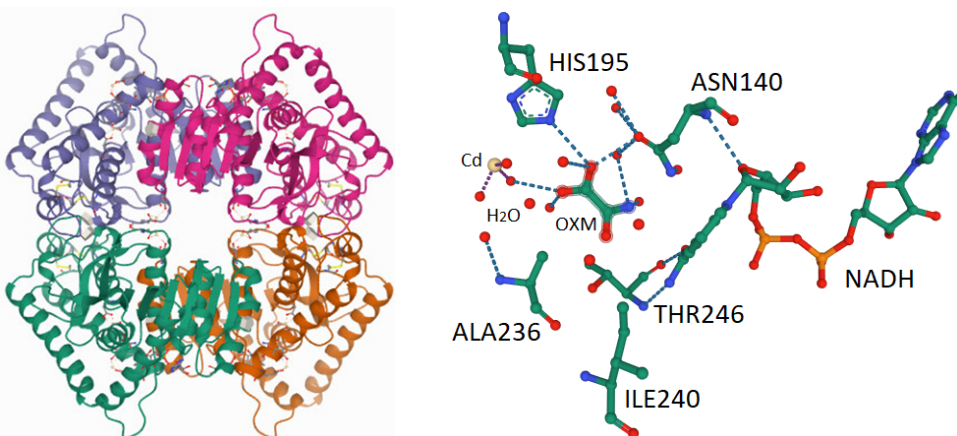


Figure 2.8. X-ray diffraction image (2.10 Å) of L-lactate dehydrogenase from *T. maritima* (*Tm*LDH) in its homotetrameric structure (left), and the substrate-binding region of one subunit (right), using oxamic acid (OXM) as substrate, ligated to the catalytic cadmium (Cd). Extracted from Protein Data Bank (PDB 1A5Z) through Mol* Viewer (Sehnal et al., 2021).

2.2. Objectives of the thesis work

The main objective of the present work is to develop a process for the production of purified recombinant enzymes, which has been assessed by producing the three enzymes required to perform the one-pot biocatalytic system to obtain L-lactic acid. This main goal can be translated into several specific objectives, which should be divided between project requirements and further research work. Broadly speaking, this thesis is entirely focused on recombinant protein production (RPP) process intensification, starting with heterologous overexpression in *E. coli* and moving on to enzyme purification/immobilisation. The specific objectives are:

- The evaluation of the production of the three enzymes involved in L-lactic acid synthesis at bench-scale reactor, with an N-terminal histidine tag and with the *E. coli* NEB 10- β strain, by using a defined culture medium.
- The production of these three enzymes with the *E. coli* M15 Δ glyA strain at bench-scale reactor, by cloning the genes of interest into an in-house developed expression vector and using an antibiotic-free defined culture medium.
- The comparison of the production yields achieved with both bacterial strains, among the different production strategies followed to overexpress the three histidine-tagged enzymes.
- The generation of new enzyme variants by fusing different carbohydrate-binding modules (CBM) to the N-terminal end of the three enzymes of interest, and the assessment of their correct expression with *E. coli* M15 Δ glyA strain.
- The production of the CBM-tagged at reactor scale enzymes by using the antibiotic-free defined culture medium.
- The development and characterisation of a one-step purification/immobilisation process of the different CBM-tagged variants onto low-cost cellulosic supports, including the study of storage stability of the immobilised derivatives and the determination of maximum loading capacity, depending on the support and the CBM used.
- The development of a fast liquid protein chromatography (FPLC) purification method by using the same cellulosic supports, and the comparison of the purification yields achieved between cellulose- and nickel-based purification supports.

3. MATERIALS AND METHODS

3.1. Bacterial strains

Several *Escherichia coli* strains have been used along the thesis work, for vector propagation, cloning and for the overexpression of the recombinant proteins designed. All bacterial strains used are presented below.

3.1.1. *Escherichia coli* DH5 α

E. coli DH5 α is a commercial strain from InvitrogenTM (Thermo Fisher Scientific, Waltham, MA, USA) widely used for routine cloning applications thanks to its high transformation efficiency. Like many cloning strains, DH5 α has several features that make it useful for recombinant DNA methods, being some of them (Grant et al., 1990):

- The *endA1* mutation inactivates an intracellular endonuclease that degrades plasmid DNA, achieving cleaner preparations of DNA.
- The *hsdR17* mutation eliminates the restriction endonuclease of the EcoKI restriction-modification system, so DNA lacking the EcoKI methylation will not be degraded, preventing cleavage of heterologous DNA by an endogenous endonuclease.
- $\Delta(lacZ)M15$ is the alpha acceptor allele needed for blue-white screening with many *lacZ* based vectors.
- The *recA1* deficiency eliminates undesired homologous recombination.

Genotype: F $\cdot\phi$ 80*lacZ* Δ M15 $\Delta(lacZYA-argF)$ U169 *recA1 endA1 hsdR17*(r_K⁻, m_K⁺) *phoA supE44 λ thi-1 gyrA96 relA1*.

3.1.2. *Escherichia coli* TOP10

E. coli TOP10 (InvitrogenTM) cells were used as an alternative to *E. coli* DH5 α cells in first cloning experiments, even if both genotype characteristics are similar. The *mcrA* mutation supports efficient transfection of methylated DNA.

Genotype: F $\cdot mcrA$ $\Delta(mrr-hsdRMS-mcrBC)$ ϕ 80*lacZ* Δ M15 $\Delta lacX74$ *recA1 araD139 $\Delta(ara-leu)$ 7697 galU galK λ rpsL(Str^R) endA1 nupG*.

3.1.3. *Escherichia coli* NEB 10- β

E. coli NEB 10- β was intended to be used as a polyvalent and transversal strain, assessing its capability to be effectively employed for multiple purposes, from molecular biology to protein production. NEB 10- β is a commercial strain from New England BioLabs[®] (Ipswich, MA, USA) derived from DH10B cells (Thermo Fisher Scientific). Despite that its genotype characteristics make it useful for recombinant DNA methods, especially plasmid cloning, it was used for protein overexpression instead, taking advantage of its *araD139* mutation, which makes it unable to metabolize arabinose. This characteristic can be useful for applying arabinose as an inducer for protein expression, through the inducible *araBAD* promoter (P_{BAD}). NEB 10- β strains used in this work were kindly sheared by Dr. Marco W. Fraaije, from Groningen University (RUG, Netherlands).

Genotype: $\Delta(ara-leu)$ 7697 *araD139 fhuA $\Delta lacX74$ galK16 galE15 e14- ϕ 80dlacZ Δ M15 recA1 relA1 endA1 nupG rpsL (StrR) rph spoT1 $\Delta(mrr-hsdRMS-mcrBC)$.*

3.1.4. *Escherichia coli* M15 $\Delta glyA$

E. coli M15 $\Delta glyA$ strain (Vidal et al., 2008) was used for protein overexpression at high cell density cultures. This thiamine-dependent strain is derived from *E. coli* M15 (Qiagen, Hilden, Germany), which contains a M15 deletion in the *lacZ* gene (Beckwith, 1964; Villarejo and Zabin, 1974). M15 $\Delta glyA$ includes the deletion of the *glyA* gene, which encodes for the enzyme serine hydroxymethyl transferase (SHMT). This enzyme is responsible for the reversible interconversion between L-threonine and glycine plus acetaldehyde and the reversible interconversion between serine and glycine (Stover et al., 1992).

Genotype: F⁻ ϕ 80 Δlac M15 *thi- lac- mtl- recA+* KmR $\Delta glyA$.

The use of this auxotrophic strain with a complementary vector which contains *glyA* gene allows to avoid the use of antibiotics as selection markers.

3.2. Plasmids

The expression systems used during the thesis are presented below.

3.2.1. pBAD

The pBAD vector system is based on the *araBAD* operon, which controls *E. coli* L-arabinose metabolism (see Section 1.3, *Escherichia coli* as a platform for recombinant protein production). The gene of interest is placed into the pBAD vector downstream of the *araBAD* promoter (Figure 3.1), which drives expression of the gene of interest in response to L-arabinose and is inhibited by glucose. Precise control of expression levels makes this system ideal for producing problematic proteins, such as proteins with toxicity or insolubility issues (Guzman et al., 1995b).

pBAD vector contains the following key features:

- The *araBAD* promoter drives transcription of the gene of interest when L-arabinose is present and glucose is absent. This promoter also controls *araC* expression.
- The ribosome-binding site (RBS) and translation initiation element from T7 bacteriophage. This allows for efficient production of the protein of interest.
- ORF: The open reading frame of the gene of interest.
- The *rrnB* terminator signal sequence stops the transcript made from the gene of interest, preventing run-on transcription.
- Ampicillin resistance gene allows the plasmid to be maintained by ampicillin selection in *E. coli*.
- pBR322 origin of replication. Plasmids carrying this origin exist in medium copy numbers in *E. coli*.
- *araC* encodes the regulatory protein of the *E. coli* *araBAD* operon. *araC* inhibits expression from the *araBAD* promoter in the absence of L-arabinose or the presence of glucose and activates transcription in the presence of L-arabinose and the absence of glucose (see Section 1.3, *Escherichia coli* as a platform for recombinant protein production).

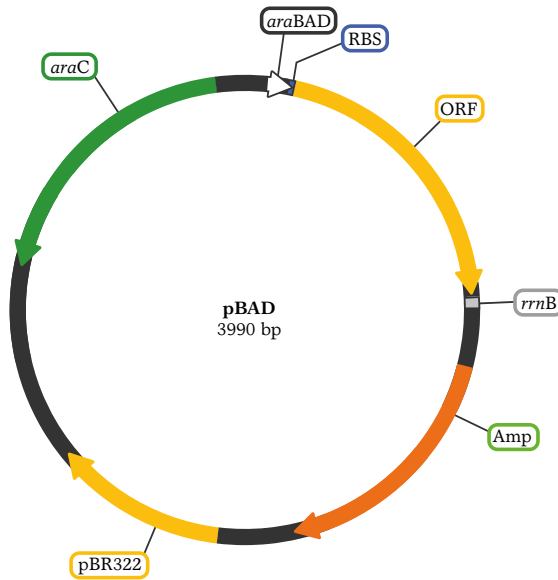


Figure 3.1. Schematic representation of pBAD expression vector (3990 bp). Features: *araBAD* promoter; RBS, ribosome binding site; ORF, open reading frame; *rrnB* transcriptional termination region; Amp, ampicillin resistance; pBR322 replication origin; *araC* gene.

3.2.2. pVEF

The pVEF plasmid is an expression vector derived from an in-house developed plasmid (Pasini et al., 2016), which was obtained from the commercial pQE-40 vector (Qiagen). In this plasmid, the gene of interest is placed after an IPTG-inducible T5 promoter, derived from T5 phage, which has a double a *lac* operator (*lacO*) repression module that provides a tightly regulated expression of recombinant proteins (Figure 3.2), since *lacI* repressor is expressed constitutively (see Section 1.3, *Escherichia coli* as a platform for recombinant protein production).

Besides, noteworthy among its features:

- T5 promoter allows for a strong overexpression of the protein of interest.
- The RBS with a high transcription level, followed by a *SmaI* restriction site that enables to clone the genes of interest right after the RBS, by using SLIC technique.

- *lacI-glyA* cassette, regulated by a constitutive promoter (J23110), provides a constant level of *lacI* repressor as well as a constitutive over-expression of SHMT enzyme.
- Two transcriptional termination regions: the t_0 from phage lambda and the *rrnB* terminator signal from T1 phage.
- ColE1 replication origin: medium copy number (15–60 copies per cell).
- Ampicillin resistance is conferred by a β -lactamase coding sequence (*bla* gene).

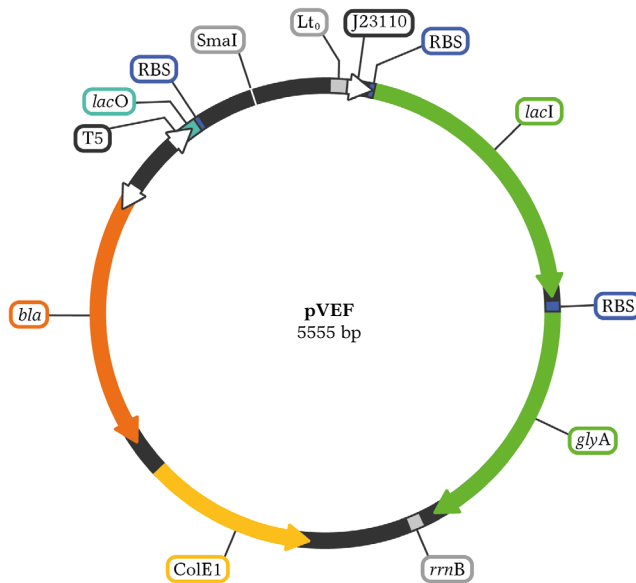


Figure 3.2. Schematic representation of pVEF expression vector (5555 bp). Features: T5 promoter; *lacO*, two *lac* operator regions; RBS, ribosome binding site; SmaI restriction site; *Lto*, lambda to transcriptional termination; J23110 constitutive promoter; *lacI* gene; *glyA* gene; *rrnB* transcriptional termination; *bla*, ampicillin resistance gene; ColE1 replication origin.

In the present work, target enzymes were cloned into the pVEF vector. Below, it is shown a schematic representation corresponding to the **two different variants of newly cloned** vectors (Figure 3.3) In all cases, enzyme-coding DNA sequences were cloned right after the SmaI restriction site. On the one hand, a 6x histidine-tag (18 bp) was fused to the 5' region of target genes (Figure 3.3 top) and, on the other hand, different DNA sequences coding for four carbohydrate-binding modules (CBM) were fused instead (Figure 3.3 bottom), spacing these tags from target enzymes with a linker (36 bp).

The enzyme-coding DNA sequences corresponded to ScADH (1047 bp), ScPDC (1692 bp) and TmLDH (960 bp), whilst CBM-coding sequences corresponded to CBM2 (315 bp), CBM2.2 (336 bp), CBM3 (462 bp), CBM9 (567 bp)

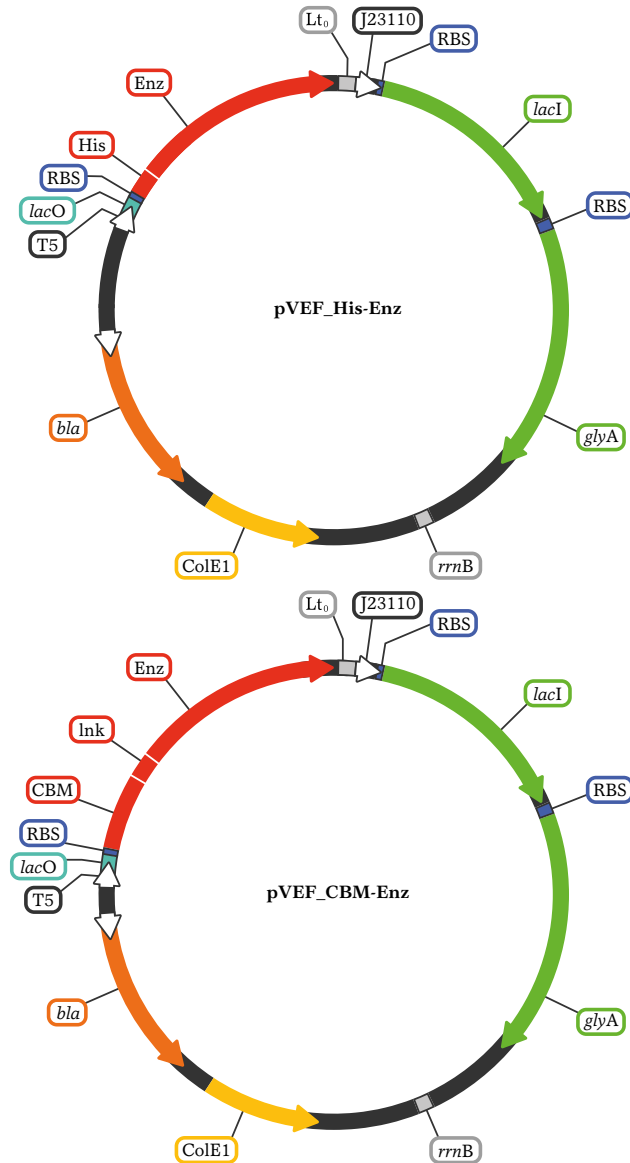


Figure 3.3. Schematic representation of pVEF expression vector containing a histidine-tagged enzyme coding DNA sequence (top) or a CBM-tagged enzyme coding sequence (bottom). Features: T5 promoter; *lacO*, two *lac* operator regions; RBS, ribosome binding site; *Lto*, lambda to transcriptional termination; J23110 constitutive promoter; *lacI* gene; *glyA* gene; *rrnB* transcriptional termination; *bla*, ampicillin resistance gene; *ColE1* replication origin; His, 6x histidine sequence; CBM, carbohydrate-binding module sequence; Lnk, linker sequence; Enz, enzyme sequence.

3.2.3. pUC57

The pUC57 is a common used plasmid cloning vector isolated from *E. coli* DH5 α strain, in which the DNA fragments corresponding to CBM3 from CtCipA (*Clostridium thermocellum* cellulosomal–scaffolding protein A), to CBM2 from CcEngD (*Clostridium cellulovorans* endoglucanase D), to CBM2 from AcEng5 (*Actinomyces* sp. 40 glycosyl hydrolase 5, from now on CBM2.2) and CBM9 from TmXyn10A (*Thermotoga maritima* endo-1,4- β -xylanase 10A) were cloned by GenScript[®] Biotech (Piscataway, NJ, USA) considering codon usage optimization for *E. coli*.

pUC57 vector (Figure 3.4) contains the following key features:

- *lac* promoter drives transcription of the gene of interest when lactose is present and *lacZ* gene codifies for N-terminal fragment of the β -galactosidase.
- MCS: multiple cloning site, containing restriction sites for several enzymes, is split into the *lacZ* gene.
- Ampicillin resistance is conferred by a β -lactamase coding sequence (*bla* gene).
- pMB1 high-copy-number replication origin (100–300 copies per cell).

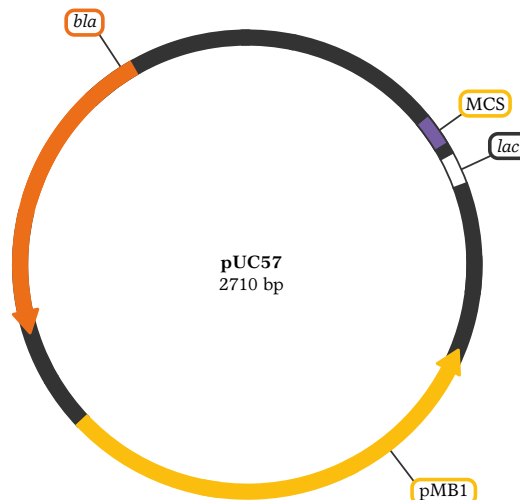


Figure 3.4. Schematic representation of pUC57 expression vector (2710 bp). Features: *lac* promoter; MCS, multiple cloning site; *bla* ampicillin resistance gene; pMB1 replication origin.

3.3. Molecular biology techniques

Basic molecular biology techniques related to recombinant DNA technology have been used in this work. All molecular biology reagents, FastDigest restriction enzymes and DNA purification kits were purchased from Thermo Fisher Scientific™ (Waltham, MA, USA), unless otherwise stated, and were used by following standard operating procedures (SOP) described in laboratory and in the manufacturer's instructions. Below, different methods and proceedings are exposed.

3.3.1. DNA amplification

The polymerase chain reaction (PCR) has been used to amplify specific DNA regions. PCR was invented in 1983 by the American biochemist Kary Mullis, who stated that "lets you pick the piece of DNA you're interested in and have as much of it as you want" (Mullis, 1990). The PCR method relies on thermal cycling, exposing reactants to repeated cycles of heating and cooling to permit different temperature-dependent reactions —specifically, DNA melting and enzyme-driven DNA replication —(Garibyan and Avashia, 2013).

PCR employs several components and reagents that include:

- DNA template, containing the target region to be amplified.
- Two primers, which are short single strand DNA fragments (oligonucleotides) that are a complementary sequence to the target DNA region. Each primer binds to the 3' end of the sense and the anti-sense strands of the DNA, serving as an extension point for the DNA polymerase to build on.
- The DNA polymerase is the enzyme that links individual nucleotides together to form the PCR product.
- Deoxynucleoside triphosphates (dNTPs), the building blocks from which the DNA polymerase synthesizes a new DNA strand.

In this work, all DNA sequences were amplified by PCR using a Phusion Flash High-Fidelity PCR Master Mix, which has a low mutation rate and creates blunt ends in DNA sequences that are suitable for the SLIC technique. PCR reactions were performed in an Applied Biosystems™ MiniAmp™ Thermal Cycler (Thermo Fisher Scientific), and amplified DNA was purified with a GeneJET PCR purification Kit.

Altogether, the genes corresponding to the three enzymes involved in L-lactic acid production—which were obtained from a pBAD vector shared by Dr. Fraaije (RUG)—were cloned into a pVEF expression vector, firstly with a six-histidine coding sequence fused to N-terminal and afterwards, with four different CBM sequences. Common PCR was performed for poly-histidine-tagged enzymes amplification, whereas it was required to perform a two-step overlap-extension PCR for cloning CBM-tagged enzymes, including a 36-nucleotide linker fragment between CBMs and enzyme sequences, whose DNA sequence was 5' AGCGCGGGCAGCAGCGCGGGCAGCGGCAGCGGC 3'. All the primers designed to amplify the genes of interest by PCR are listed in Tables 3.1 and 3.2.

Table 3.1. List of all oligonucleotides used as primers in polymerase chain reactions for DNA amplification of histidine-fused enzymes. T_m, melting temperature in °C, GC, guanine–cytosine content in%. Codon corresponding to first histidine residue pointed out in bold (blue), first codifying codon pointed out in bold and stop codon underlined. Fw, forward, Rev, reverse.

Name (sense)	5' to 3' DNA sequence	Size (bp)	T _m (°C)	GC (%)
Frag.For. (Fw)	AGGAGAAATTAACCC AT GGGCAGCAGC CAT CAT	33	68	48
ScADH (Rev)	CTAATTAAGCTTCC CTA TTTAGAAGTGCAACAACG	37	61	35
ScPDC (Rev)	CAGCTAATTAAGCTTCC CTA CTGCTTGCCA	31	64	45
TmLDH (Rev)	CTAATTAAGCTTCC CTA GCCGCTGGTGTTTGG	34	65	47

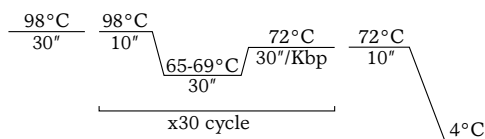
Table 3.2. List of all oligonucleotides used as primers in polymerase chain reactions for DNA amplification of CBM-fused enzymes. T_m, melting temperature in °C, GC, guanine–cytosine content in%. Overlapping region between the two DNA fragments of each construct pointed out in bold (green), first codifying codon pointed out in bold and stop codon underlined. Fw, forward, Rev, reverse.

Name (sense)	5' to 3' DNA sequence	Size (bp)	T _m (°C)	GC (%)
CBM3-linker (Fw)	AGGAGAAATTAACCC ATG AACCTGAAAGTGGAA	33	63	39
CBM2-linker (Fw)	TAAAGAGGAGAAATTAACCC ATG CCGAGC	29	62	45
CMB2.2-linker (Fw)	AGGAGAAATTAACCC ATG CCGACCACCC	28	66	54
CBM9-linker (Fw)	AGGAGAAATTAACCC ATG GTGGCGACCG	28	66	54

Linker-ScADH (Rev)	GGATAGACATGCCGCTGCCGCT	22	66	64
Linker-ScADH (Fw)	CGGCAGCGGCATGTCTATCCAGAAACTCA	30	68	57
ScADH (Rev)	CTAATTAAGCTTCCCTT <u>TA</u> TTTAGAAGTGTCAACAACG	37	61	35
Linker-ScPDC (Rev)	TTTCAGACATGCCGCTGCCGCT	22	65	59
Linker-ScPDC (Fw)	CGGCAGCGGCATGTCTGAAATTA	23	60	52
ScPDC (Rev)	CAGCTAATTAAGCTTCCCTT <u>ACT</u> GCTTGGCA	31	64	45
Linker-TmLDH (Rev)	CAATCTTCATGCCGCTGCCGCTG	23	65	61
Linker-TmLDH (Fw)	GCAGCGGCATGAAGATTGGTATC	23	60	52
TmLDH (Rev)	CTAATTAAGCTTCCCTT <u>AG</u> CCGCTGGTGTTTTGG	34	65	47

Annealing step lasted 30 seconds and the temperature depended on the design of the primers, while elongation step temperature was fixed on 72°C and time varied depending on the length of the fragment (30 second per each 1000 bp). Temperature profiles of PCR methods are represented in [Figure 3.5 A](#) for common PCR and in [Figure 3.5 B](#) for overlap–extension PCR.

A



B

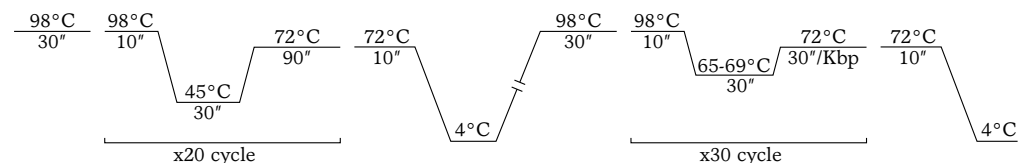


Figure 3.5. Temperature profile and number of cycles of polymerase chain reaction methods employed for common PCR (**A**) and for overlap-extension PCR (**B**).

3.3.2 DNA isolation and manipulation

Plasmid DNA isolation is an essential technique used in molecular cloning in order not only to construct new clones but to analyse the ones obtained. For the extraction of plasmid DNA from recombinant strains, commercial GeneJET Plasmid Miniprep Kit was used. The DNA extraction is based on the alkaline lysis method developed by H.C. Birnboim and J. Doly (Birnboim and Doly, 1979), which consists in a denaturation of chromosomal DNA and proteins under alkaline conditions (pH 12.0-12.5), while plasmid DNA is bound to a resin and eluted afterwards using a low-ionic-strength buffer.

The enzymatic restriction of DNA was employed for linearizing the pVEF vector prior to enzyme cloning, and for checking the correct cloning of the new constructs obtained. The digestion of DNA was carried out using restriction enzymes. Enzyme and DNA concentrations depended on the specific application, and so did temperature and reaction time. Nevertheless, the restriction process was usually performed at 37°C for 30 minutes, and a deactivation step of 5 minutes at 85°C was applied when necessary.

In general, a restriction enzyme unit is defined as the amount of enzyme required to completely digest 1 mg of DNA in 60 minutes under optimum conditions. The endonucleases used on that purpose, are listed below in [Table 3.3](#):

Table 3.3. List of all restriction enzymes used along this thesis work. Cleavage sites are indicated with arrows (↓ ↑).

Enzyme	Cleavage site
SmaI	5'...CCC↓GGG...3' 3'...GGG↑CCC...5'
XbaI	5'...T↓CTAGA...3' 3'...AGATC↑A...5'
XhoI	5'...C↓TCGAG...3' 3'...GAGCT↑C...5'

The length of all DNA fragments resulting from PCR reactions were checked by agarose gel **DNA electrophoresis**, which promotes DNA separation according to the length of the DNA fragments by applying a constant voltage, by which DNA fragments travel from the positive cell electrode to the negative one during the time of exposure.

1% agarose gels were prepared by mixing 0.5 g of commercial TopVision agarose tablets in 50 mL 5 mM Lithium acetate buffer (LA) and SYBR™ Safe as DNA staining reagent (5 µL). Samples were prepared with 6x DNA loading dye and loaded into the gel. Electrophoresis gels were run at 200 V in LA buffer for 15 minutes. GeneRuler DNA Ladder Mix was loaded as standard molecular weight marker.

3.3.3 Enzyme cloning

Fusion proteins were cloned into a previously SmaI linearized pVEF vector using a variation of the sequence- and ligation-independent cloning (SLIC) method (Jeong et al., 2012; Islam et al., 2017) (Figure 3.6 A). SLIC is a molecular biology technique derived from previously developed ligation-independent cloning (LIC) method (Aslanidis and Jong, 1990), which was based on the 3'-to-5' exonuclease activity of T4 DNA polymerase and has been used as a high-throughput method due to its uniformity and cost-effectiveness.

However, LIC requires a specifically designed vector containing a long stretch of sequence that lacks a particular deoxynucleoside triphosphate (Jeong et al., 2012), whereas SLIC overcomes the sequence restraint of LIC and allows the assembly of multiple overlapping fragments simultaneously as a simple, cost-effective, time-saving, and versatile cloning method.

One-step SLIC allows inserting a gene of interest into any vector at any position regardless of sequence variations of the genes of interest since there is no need to digest inserts with restriction endonucleases. By modulating the junction between the homologous and gene-specific regions of PCR primers for the insert, one can delete or add any additional sequences, such as a restriction enzyme site or a tag sequence, providing a robust basis for high-throughput gene cloning (Jeong et al., 2012; Islam et al., 2017).

Briefly, two separate T4 DNA polymerase reactions for linearized vector and insert were carried out containing in 10 μ L final volume, 40–50 ng of either linearized vector DNA or the corresponding insert DNA in 1:2 molar ratio (Figure 3.6 B). Reaction buffer was composed of 200 mM Urea, 20 mM DTT, 33 mM Tris-acetate (pH 7.9), 10 mM Magnesium acetate, 66 mM potassium acetate, 0.1 mg·mL⁻¹ BSA and 2.5 U of T4 DNA polymerase.

Both were incubated at 12°C for 10 minutes. Ethylenediaminetetraacetic acid (EDTA) (45.5 mM) was added in each tube to stop the reaction and the tubes were incubated at 75°C for 10 minutes afterwards. The two reactions were then mixed and annealed progressively using a temperature decrease ramp of 2.4°C·min⁻¹ for 20 minutes, from 72 to 24°C (Figure 3.6 B).

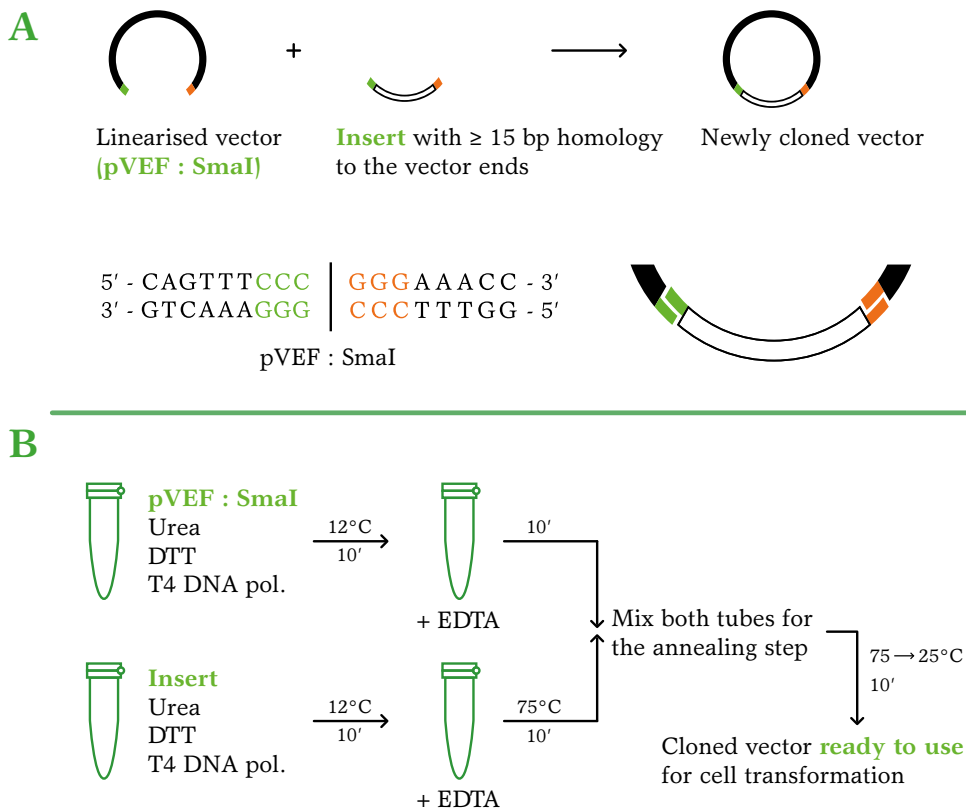


Figure 3.6. Schematic diagrams of one-step SLIC. General overview (A). Homologous regions are in the same color. The SmaI restriction site is indicated. (B) Step-by-step SLIC procedure.

3.3.4. DNA transformation

The conversion of one genotype into another by the introduction of exogenous DNA is termed transformation, which was first reported in *Streptococcus pneumoniae* in 1928 by Frederick Griffith. The mechanism is marked by two phases, the first phase involves the uptake of the DNA across the cellular envelope and the second phase involves the setting up of the DNA in the cell as a stable genetic material.

Transformation of heterologous DNA into *E. coli* was achieved by a physicochemical method, based on the capability of calcium ions (Ca^{2+}) to create pores in the cell membrane; the repulsion between foreign DNA and the bacterial cell, owing to negative charges on them both, are overcome by these divalent cations, which do not only mask the negative charge of DNA but also cause changes in the membrane permeability. It has been reported that DNA binds to the lipopolysaccharide (LPS) receptor molecules on the competent cell surface, and the divalent cations generate coordination complexes with the negatively charged DNA molecules and LPS (Bergmans et al., 1981, Asif et al., 2017). Additionally, the heat shock step strongly depolarizes the cell membrane of CaCl_2 -treated cells, helping to deal with the barriers to DNA uptake.

Chemical competent cells were prepared as follows: 50 μL of cell stock of the desired bacterial strain were grown in 250 mL shake-flask cultures with 10 mL Lysogeny broth medium (LB) (37°C , 200 rpm, overnight). 1 mL of first cultures was used to inoculate 500 mL Erlenmeyer flasks with 50 mL LB, under same culture conditions. When cell growth reached and OD_{600} of approximately 0.4 (two hours after inoculation), cells were harvested by centrifugation (4000 rpm, 4°C , 10 minutes) and were resuspended in 30 mL of ice-cold 0.1 M CaCl_2 solution. Afterwards, a second centrifugation step was applied, and the resulting cell pellet was resuspended in 2 mL calcium solution. Once homogenized, 2 mL (1:1 v/v ratio) of ice-cold 0.1 M CaCl_2 solution containing glycerol (40% w/w) were added. Finally, competent cells were aliquoted into pre-chilled eppendorf tubes (100 μL each) and were kept under ultra-refrigeration conditions (-80°C).

Competent cells were checked by seeding 50 μL of each cell stock prepared in LB-agar plates supplemented with $100 \text{ mg}\cdot\text{L}^{-1}$ ampicillin (incubated in a Sanyo MIR-154 incubator, 37°C , overnight), expecting to observe no colonies in any case.

For the heat–shock transformation, the following steps were followed: 100 μL of *E. coli* competent cells were transformed by with 10 μL of SLIC product as follows: Competent cells were taken out of -80°C and thawed on ice (approximately 20 minutes). 10 μL of SLIC product were added to the aliquots and the cell-DNA mixture was incubated on ice for 10 minutes. Tubes were then placed into a 42°C water bath for 60 seconds and were transferred to the ice for 2 minutes afterwards. Finally, 390 μL of LB medium without antibiotic were added to the bacteria and tubes were incubated at 37°C for 45 minutes.

Both positive and negative controls of the transformation were performed by adding 10 μL of circular pVEF (at $10\text{ ng}\cdot\mu\text{L}^{-1}$) and 10 μL of sterile deionized water, respectively. Additionally, 10 μL of the SLIC negative control (linearized plasmid without insert) were added to a third control tube.

Finally, transformed clones were selected using LB-agar plates with $100\text{ mg}\cdot\text{L}^{-1}$ ampicillin (37°C , overnight) and transformations were confirmed by colony-PCR and by plasmid restriction pattern (Sections 3.3.1 and 3.3.2, respectively).

3.3.5. DNA sequencing

Eventually, all the genes obtained by PCR and all the different constructs generated along this thesis work were sequenced and verified using an ABI 3130XL DNA sequencer device from Applied Biosystems™ (Thermo Fisher Scientific), at IBB facilities (Institute of Biotechnology and Biomedicine, UAB, Spain), via Genomic and Bioinformatics Service (SGB).

The Sanger method was applied for DNA sequencing, based on the use of a thermostable polymerase that allows a cyclic sequencing reaction (Sanger et al., 1977). The primers designed for PCR amplification (Tables 3.1 and 3.2) were also used for DNA sequencing, with the addition of one forward–sensed primer based on pVEF vector, whose homologous region is placed 166 bp prior to RBS. Its corresponding sequence was 5' CGAAAAGTGCCACCTGACGT 3' and it was used aiming to determine with higher accuracy the DNA region corresponding to the start of transcription.

3.4. Growth medium composition

Along this thesis work, different culture mediums have been used for cell growth. Complex mediums were used at early stages of the research, for molecular biology experiments, first cell stocks generation and for preliminary production studies in well plates and in shake flasks. Afterwards, a defined and antibiotic-free minimum medium (DM) with glucose as carbon source was used for final cryostock generation and for protein overexpression at bioreactor scale.

All reagents used for medium preparation were purchased from Sigma-Aldrich (St. Louis, Mi, USA), unless otherwise stated.

3.4.1. Complex mediums

Lysogeny broth (LB) —which has come to colloquially mean Luria–Bertani medium—, containing $10\text{ g}\cdot\text{L}^{-1}$ peptone, $5\text{ g}\cdot\text{L}^{-1}$ yeast extract and $10\text{ g}\cdot\text{L}^{-1}$ sodium chloride (NaCl), was used for molecular biology experiments, for first cell stocks generation and for preliminary production screenings. LB was also used for Petri dish cultures, for which bacteriological grade agar was added to a final concentration of $15\text{ g}\cdot\text{L}^{-1}$. In all cases LB medium was sterilised by autoclaving (121°C , 30 minutes).

Terrific broth (TB) medium, containing $23.6\text{ g}\cdot\text{L}^{-1}$ tryptone, $11.8\text{ g}\cdot\text{L}^{-1}$ yeast extract, $4\text{ g}\cdot\text{L}^{-1}$ glycerol (by adding $8\text{ mL}\cdot\text{L}^{-1}$ of a 50% w/w glycerol solution), $9.4\text{ g}\cdot\text{L}^{-1}$ K_2HPO_4 , $2.2\text{ g}\cdot\text{L}^{-1}$ KH_2PO_4 , was used for preliminary protein overexpression experiments performed in Erlenmeyer flasks. TB medium components were sterilised by autoclaving (121°C , 30 minutes).

3.4.2. Defined mediums

The defined minimum medium (DM) is a glucose mineral medium used for final cryostock generation and for protein overexpression at shake-flask and bioreactor scale. This medium has been widely used inside the research Group (Durany et al., 2004, Vidal et al., 2008, Pasini et al., 2016). The diverse components of the DM were prepared separately as stock solutions as listed below (Tables 3.4 and 3.5):

Table 3.4. List of all stock solutions used for DM preparation.

Component	Concentration (g·L ⁻¹)
Glucose	400
MgSO ₄ · 7H ₂ O	400
(NH ₄) ₂ SO ₄	300
K ₂ HPO ₄	52.8
KH ₂ PO ₄	10.4
NaCl	8
FeCl ₃	50
Thiamine	50
CaCl ₂ · 2H ₂ O	100

Additionally, a trace elements solution (TES) was prepared:

Table 3.5. List of all TES components and their final concentration in stock preparation.

Component	Concentration (g·L ⁻¹)
AlCl ₃	0.04
ZnSO ₄ · 7H ₂ O	1.74
CoCl ₂ · 6H ₂ O	0.16
CuSO ₄ · H ₂ O	1.55
H ₃ BO ₃	0.01
MnCl ₂ · 4H ₂ O	1.42
NiCl ₂ · 6H ₂ O	0.01
Na ₂ MoO ₄	0.02

Glucose, phosphates (macro-elements) and ammonium sulphate stock solutions were sterilised by autoclaving (121°C, 30 minutes) while TES and the rest of the solutions were sterilised separately by filtration using a 0.2 µm syringe filter. All components were mixed once the heat of autoclaved solutions decreased to room temperature (24°C).

Nevertheless, it was also required to prepare an additional stock solution of glycerol (500 g·L⁻¹, sterilised by autoclaving at 121°C for 30 minutes) that was used as carbon source in cultures where *E. coli* NEB 10-β containing a pBAD vector was grown, mainly because the *araBAD* operon is inhibited by glucose (see section 1.1.1.2). Moreover, NEB 10-β cells required the addition of three amino acids when DM was used, as previously described (Miret et al., 2020). For that reason, three concentrated amino acid stock solutions were prepared as described below (Table 3.6), which were sterilised by filtration using a 0.2 µm syringe filter.

Table 3.6. List of the three amino acid stock solutions prepared.

Component	Concentration (g·L ⁻¹)
L-valine (L-val)	7.04
L-leucine (L-leu)	14
L-isoleucine (L-ile)	10.4

Defined medium used for shake-flask cultures —including cultures for protein overexpression experiments and the preparation of the inoculums for the bioreactor processes— was prepared from sterile concentrated stock solutions. The list of all components and their final concentration are listed below (Table 3.7):

Table 3.7. DM composition for shake-flask cultures.

Component	Concentration (g·L ⁻¹)
Glucose	10
MgSO ₄ · 7H ₂ O	0.2
(NH ₄) ₂ SO ₄	4
K ₂ HPO ₄	13.2
KH ₂ PO ₄	2.6
NaCl	2
FeCl ₃	0.025
Thiamine	0.1
CaCl ₂ · 2H ₂ O	0.005
TES	3 mL·L ⁻¹ / 0.72 mL·L ⁻¹ *
L-val **	0.07
L-leu **	0.21
L-ile **	0.104

* TES proportion decreased in flask cultures where NEB 10-β strain was grown.

**DM had to be supplemented with amino acids in cultures where NEB 10-β strain was grown.

For *E. coli* NEB 10-β overexpression experiments in shake-flask cultures, a second DM that contained glycerol instead of glucose was prepared at the same final concentration. The use of the two DM prepared for this specific case is described in section 3.5.1.

Defined medium used for batch stages performed during bioreactor processes was prepared as follows (Table 3.8):

Table 3.8. DM composition for bioreactor cultures at batch phase.

Component	Concentration (g·L ⁻¹)
Glucose	20
MgSO ₄ · 7H ₂ O	0.5
(NH ₄) ₂ SO ₄	6
K ₂ HPO ₄	13.2
KH ₂ PO ₄	2.6
NaCl	2
FeCl ₃	0.025
Thiamine	0.2
CaCl ₂ · 2H ₂ O	0.005
TES	3 mL·L ⁻¹ / 0.72 mL·L ⁻¹ *
L-val **	0.28
L-leu **	0.84
L-ile **	0.416

* TES proportion also decreased in bioreactor cultures where NEB 10-β strain was grown.

**DM had to be supplemented with amino acids in cultures where NEB 10-β strain was grown.

For the bioreactor production processes in which a **fed-batch strategy** was followed, a concentrated feeding solution was added after the end of batch phase to reach high cell densities. Its composition is described in [Table 3.9](#). Once again, For *E. coli* NEB 10-β cultures, a second feeding medium that contained glycerol instead of glucose was prepared at the same final concentration. The use of the two DM prepared for this specific case is described in Section 3.5, Cultivation conditions.

Table 3.9. Concentrated feeding solution used in fed-batch processes.

Component	Concentration (g·L ⁻¹)
Glucose	500
MgSO ₄ · 7H ₂ O	9.5
FeCl ₃	0.5
Thiamine	0.02
CaCl ₂ · 2H ₂ O	0.1
TES	60 mL·L ⁻¹
L-val *	1.68
L-leu *	5.02
L-ile *	2.56

*DM had to be supplemented with amino acids in cultures where NEB 10-β strain was grown.

3.4.3. Medium supplements

The antibiotic used and the inducers employed for protein overexpression were prepared, sterilised by filtration (0.2 μm) and stored in aliquots at -20°C as stock solutions:

Table 3.10. Medium supplements.

Component	Concentration (g·L ⁻¹)
Ampicillin	100
IPTG	23.83
L-arabinose	200

3.5. Cultivation conditions

Prior to perform all the overexpression experiments that have been carried out along this work, different ready-to-use cell bank cryostocks were established for all the *E. coli* strains employed for that purpose. In addition, stocks were prepared with both complex and defined mediums. This way, bacterial cells were stored at -80°C in a mixture of culture media with glycerol (25% v/v final concentration).

For **stock generation**, a single colony was picked from plate cultures, and cells were grown in 50 mL Erlenmeyer flasks containing 5 mL of the desired medium —LB or DM— (37°C , 140 rpm, overnight). Afterwards, 5 mL of previous culture were transferred to a new flask containing 30 mL of culture medium, starting at an OD_{600} of approximately 0.5. Same culture conditions were kept and after six hours the biomass was separated from the culture media by centrifugation at 5000 rpm for 10 minutes at 4°C in an Avanti™ J20 centrifuge (Beckman Coulter, Brea, CA, USA) in 50 mL falcon tubes. Culture medium was discarded, and the resulting pellet was resuspended with 5 mL of fresh medium. Once homogenized, 5 mL of an autoclaved glycerol solution 50% (v/v) were added. The final mixture was aliquoted in cryovials containing each 1 mL of the final cell bank stock.

In all cases where complex medium was used, cultures were supplemented with ampicillin ($0.1\text{ g}\cdot\text{L}^{-1}$ final concentration) as selection marker, aiming to avoid the loose of the plasmid transformed in each case.

After molecular biology experiments by which new transformants were obtained, it was required to perform a previous **overexpression screening experiment** with the aim to establish a ready-to-use cell bank cryostock with the most productive *E. coli* M15 Δ glyA colonies.

Well plate cultures were seeded with transformed colonies from LB-agar plates. Protein production screenings were carried out (2 mL LB medium with $0.1\text{ g}\cdot\text{L}^{-1}$ ampicillin, 24°C , 24h, 140 rpm, 0.4 mM IPTG) and more productive candidates were used to generate the final stock.

Before generating cell bank stocks, M15 Δ glyA colonies were adapted to DM by performing a **three-step adaptation cultures** in 100 mL Erlenmeyer flasks with 30 mL of DM (overnight at 37°C and 140 rpm of agitation).

3.5.1. Shake-flask cultures

For flask-scale cultures, inoculums were prepared by adding 100 μL of the corresponding cryostock into a 50 mL Erlenmeyer flask with 10 mL of culture medium. Growth was performed overnight at 37°C with 140 rpm of agitation.

For protein overexpression experiments in flasks, all the cultivations were performed for triplicate at 24°C and 140 rpm. 250 mL Erlenmeyer flasks containing 50 mL of culture medium were used and inoculum was added to start the production process at an OD_{600} of approximately 0.5 in all cases.

Induction step: For the protein overexpression experiments performed with complex medium the corresponding inducer was added with a single pulse at a final concentration of 0.06 $\text{g}\cdot\text{L}^{-1}$ for the expression systems induced with IPTG and at a concentration of 0.1 $\text{g}\cdot\text{L}^{-1}$ for the expression vectors induced with L-arabinose, and all cultures were supplemented with ampicillin (0.1 $\text{g}\cdot\text{L}^{-1}$).

However, when DM was employed the induction step was performed in a different way according to the expression vector used; on the one hand, for *E. coli* NEB 10- β strain transformed with pBAD vectors, a medium replacement was needed. Cultures started with the glucose-based DM and after 24 hours of growth, the biomass was centrifuged (5000 rpm, 10 minutes, 4°C) and the resulting pellets were resuspended with 10 mL of fresh glycerol-based DM. A second centrifugation process was carried out under the same conditions and pellets were resuspended in 50 mL of glycerol-based DM. Finally, a pulse of L-arabinose was added at a final concentration of 0.1 $\text{g}\cdot\text{L}^{-1}$. Overexpression experiments ended 24 hours after the medium replacement.

On the other hand, cultures performed with *E. coli* M15 ΔglyA containing a pVEF vector did not need any medium change. For that, a pulse of IPTG was added after 6 to 8 hours of growth at a final concentration of 0.25 mM (0.06 $\text{g}\cdot\text{L}^{-1}$), and the overexpression experiments ended at 24 hours of growth.

In all cases, inductor was added in two of the three flasks, whereas the third culture was not induced (negative control) and samples before and after the induction step were taken from all flasks.

3.5.2. Bioreactor cultures

Bioreactor cultivations were carried out by following two different process approaches; for some overexpression experiments a batch strategy was chosen whereas for some other experiments fed-batch processes were carried out. Nevertheless, in all bioreactor cultivations an Applikon ez-Control (Applikon Biotechnology[®], Delft, Netherlands) was used, equipped with pH, dissolved oxygen (PO₂) and temperature probes. The digital control unit (DCU) controlled the pH, stirring, temperature and dissolved oxygen of the culture.

For batch fermentations, a 5 L vessel was used. In all cases the temperature was maintained at 24°C along the process, regardless of the culture medium used, and the pH was kept at 7.00 ± 0.05 by adding NH₄OH 15% (v/v) and 2 M H₂SO₄ solutions; NH₄⁺ from base addition did also serve as the nitrogen source for biomass growth. Airflow of 1 vvm was applied and oxygen saturation level was controlled through the stirring (450–1100 rpm); it was set to a PO₂ of 60% for *E. coli* M15ΔglyA strain and 30% for NEB 10-β strain. In the latter case, ampicillin (0.1 mg·mL⁻¹) had to be added in all cases.

For fermentations with TB medium, pre-inoculum cultures were prepared in 50 mL Erlenmeyer flasks with 15 mL of LB plus 100 μL of the corresponding cryostock, (37°C, 140 rpm, overnight and ampicillin 0.1 g·L⁻¹). Inoculum cultures started at an OD₆₀₀ of 0.2 in duplicate 500 mL Erlenmeyer flasks, using 100 mL of LB per flask and the corresponding volume of pre-inoculum supplemented with the antibiotic (24°C, 140 rpm). Cultures were kept under incubation for 4 to 6 h, until the biomass concentration overtook an OD₆₀₀ of 1.

The whole volume of inoculum cultures (200 mL) was used to inoculate the reactor, that contained 3.5 L of TB culture medium (also with antibiotic at 0.1 g·L⁻¹). The induction was performed 4 hours after the inoculation, and the process ended at 22–24 hours, when all substrates were consumed and cellular metabolism decreased notably, reaching a final culture volume of around 3.7 and 3.8 L. Culture was induced by adding a single pulse of 1 mL·L⁻¹ of a 20% (w/w) arabinose solution (final concentration of 0.2 g·L⁻¹) for NEB 10-β strains, and a single pulse of 2.5 mL·L⁻¹ of 100 mM IPTG solution (0.25 mM final concentration) for M15ΔglyA strains.

For batch processes with DM, pre-inoculum cultures were prepared in 50 mL Erlenmeyer flasks with 15 mL of DM plus 100 μ L of the corresponding cryostock, (37°C, 140 rpm, overnight). Inoculum cultures started at an OD₆₀₀ of 0.2 in duplicate 500 mL Erlenmeyer flasks, using 100 mL of DM per flask and the corresponding volume of pre-inoculum (24°C, 140 rpm). Cultures were kept under incubation for 4 to 6 h, until the biomass concentration overtook an OD₆₀₀ of 1.

Inoculum volume was added to the reactor with 4 L of DM, starting at an OD₆₀₀ of approximately 0.1. Culture growth was kept overnight, and the induction phase occurred afterwards (at 16–17 hours of process) by adding a single pulse of IPTG (0.25 mM final concentration). Batch processes stopped when all glucose was consumed, and cell metabolism decreased, what happened after 26–28 hours process.

In all cases, samples before and after the induction were taken and protein content and enzyme activity were measured afterwards.

Fed-batch experiments were performed in the same bioreactor equip than for batch cultures, and culture conditions were similarly set and controlled, with the modification of pure oxygen addition after the maximum stirring was achieved in the cascade control for PO₂ maintenance. A 2 or a 5 L vessel were used, and temperature was maintained at 37°C and pH at 7.0 ± 0.05 by adding NH₄OH 15% v/v and 2 M H₂SO₄ solutions. Airflow of 1vvm was applied, and oxygen saturation level was set to a PO₂ of 60% for M15 Δ glyA strains and of 30% for NEB 10- β , controlled through a cascade of stirring (450–1100 rpm) and pure oxygen addition after maximum stirring was reached (setup images in [Figure 3.7](#)). Initial batch phase was started by transferring 200 mL of inoculum to 800 mL of DM, with 20 g·L⁻¹ of glucose. Afterwards, once the initial glucose was totally consumed, the substrate limiting fed-batch phase started by adding exponentially the feeding medium through a preprogrammed exponential addition, using the following equation (Eq. 1):

$$F = \frac{\mu \cdot X \cdot V_0 \cdot e^{(\mu \cdot \Delta t)}}{Y_{X/S} \cdot S_0} \quad (1)$$

Where:

- F is the feeding flux (mL·min⁻¹).
- μ is the fixed specific growth rate (0.2 h⁻¹)
- Δt is the time interval in which the feeding flux is applied (1 h).
- X is the predicted biomass concentration (g·L⁻¹) in the bioreactor at the end of the time interval.
- V_0 is the culture volume (L) at the beginning of the time interval.
- $Y_{X/S}$ is the biomass/substrate yield (set at 0.3 g·g⁻¹).
- S_0 is the concentration of substrate —glucose or glycerol— in the feeding medium (~500 g·L⁻¹).

For processes with M15 Δ glyA, **induction phase** started by adding a pulse of 100 mM IPTG (2.5 mL·L⁻¹, 0.25 mM final concentration) when the culture surpassed a biomass concentration OD₆₀₀ \approx 100, and the whole process ended when cell growth stopped and glucose accumulation was detected, reaching in all cases a final culture volume of approximately 2 L.

On the other hand, for processes with NEB 10- β cells, two different addition media were used: one for a first fed-batch growth phase and another for the induction phase, being the first one the same glucose-based medium used for M15 Δ glyA and the second one a glycerol-based and amino acid-supplemented DM (see Table 3.9) changing at this point the specific growth rate (μ) from 0.2 to 0.05 h⁻¹. This approach was developed and described by J. Miret (Miret et al., 2020), who after the first fed-batch phase applied a a step of 1 h without feeding and with a linear decreasing of temperature to 24 °C, that was followed by a second fed-batch phase, in which the induction was carried out by adding arabinose at a final concentration of 0.2 g·L⁻¹ in the bioreactor.

Once the induction phase started, the above mentioned NH₄OH solution used as a corrector agent was switched by a NH₄OH solution containing the amino acids L-leu, L-ile and L-val at a concentration of 90, 45 and 30 g·L⁻¹, respectively. That fact was required since the proportion amino acid vs carbon source in the base medium could not be met in the feeding medium due to solubility issues.

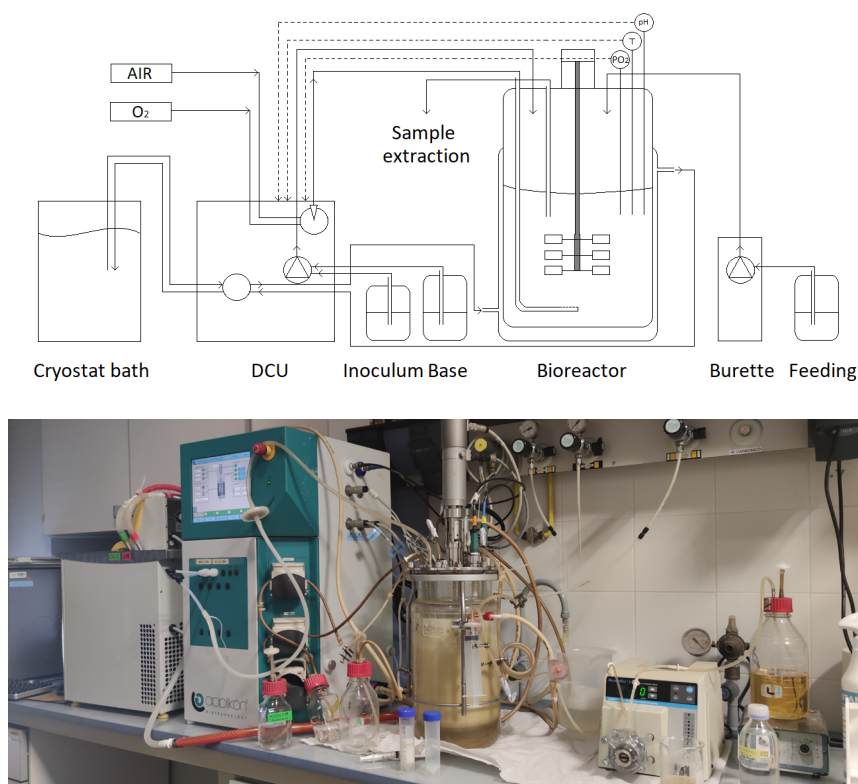


Figure 3.7. Schematic representation of the bioreactor setup (top) and a picture of the actual bioreactor setup (bottom).

3.6. Enzyme recovery

Biomass was separated from the culture media by centrifugation at 7000 rpm for 20 min at 4°C in an Avanti™ J20 centrifuge (Beckman Coulter, Brea, CA, USA). For shake-flask overexpression experiments 50 mL falcon tubes were used whereas for bioreactor production 500 mL centrifugation tubes were employed. The pellet was divided into several aliquots and kept frozen at -20°C. When needed, the aliquot of interest was resuspended in 50 mM Tris-HCl (pH 7.50) buffer and lysed using an OneShot cell disruptor (Constant Systems Ltd, Daventry, UK), for 2 cycles at 1.47 kbar pressure. Phenylmethylsulphonyl fluoride (PMSF) was added as protease inhibitor, unless otherwise stated at final concentration of 0.1 mM (17.42 mg·L⁻¹). Cell debris and insoluble fraction was removed by centrifugation at 13000 rpm for 30 min at 4°C. Cell lysate for further usages was stored at -20°C in 50 mL falcon tubes.

3.7. Analytical methods

3.7.1. Biomass measurement

The biomass concentration was measured in terms of absorbance at 600 nm of wavelength (OD_{600}) using a HACH[®] D3900 (Hach, Loveland, CO, USA) spectrophotometer, diluting the samples in an absorbance lower than 0.9 with distilled water. Biomass concentration expressed as dry cell weight (DCW) was calculated considering that 1 OD_{600} unit is equivalent to 0.3 g DCW·L⁻¹ (Pinsach et al., 2008; Miret et al., 2020). All samples were measured for duplicate.

3.7.2. Substrates and by-products concentration measurement

In all cases, 1 mL of culture sample was centrifuged (13000 rpm, 3 minutes, 24°C) and filtered (0.45 µm) to remove biomass prior to the analytic procedures.

The resulting supernatant was then used for **glucose and lactate concentration measurement**, using an YSI 20170 system (YSI Inc., Yellow Springs, OH, USA). All samples were diluted with distilled water to a glucose concentration lower than 10 g·L⁻¹ and were analyzed for duplicate.

Supernatant was also used for **glycerol and acetate concentration determination**, by high-performance liquid chromatography (HPLC) Ultimate 3000 (Thermo Scientific) equipped with a Dionex UltiMate 3000 variable wavelength detector. An ICsep COREGEL 87H3 (Transgenomic[®], Omaha, NE, USA) 7.8 x 300 mm column was employed, using a 6 mM H₂SO₄ solution (pH 2.0) as mobile phase (0.3 mL·min⁻¹, 40°C).

3.7.3. Total protein content and enzyme titre determination

Total intracellular protein content present in cell lysates were determined with the Bradford method and relative amount of target enzymes was determined by protein electrophoresis.

The **Bradford method** was carried out with Coomassie Protein Assay Reagent Kit (Thermo Scientific) and bovine serum albumin (BSA) as protein standard, following manufacturer's instructions. The assays were

performed in 96-microwell plates, by mixing 200 μL of Coomassie reagent with 7 μL of a sample (diluted with distilled water if needed). A Multiskan™ FC equipment (Thermo Scientific) was used for the absorbance reading (595 nm) 15 minutes after the mixture of the components.

BSA serial dilutions between 0.4 and 0.1 $\text{mg}\cdot\text{mL}^{-1}$ were used to generate a regression line that allowed to determine the concentration of the analyzed samples as a function of their absorbance, and all samples were analyzed for duplicate (BSA included).

For **protein electrophoresis**, samples were subjected to sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis to determine the percentage of target enzymes among the rest of intracellular soluble proteins present in the lysates obtained.

For that, NuPAGE® 12% Bis Tris gels MES-SDS as running buffer were used, following the manufacturer's manual (Invitrogen™). Samples were prepared by mixing 5 μL of sample buffer (4X) with 3 μL of miliQ water, 2 μL of NuPAGE® sample reducing agent (10X) and with 10 μL of the sample (with a protein concentration of approximately 0.4 $\text{mg}\cdot\text{mL}^{-1}$). After 10 minutes of incubation at 70°C, 15 μL of each vial were loaded into the gel. Blue Prestained Precision Plus Protein standard (Bio-Rad Laboratories, Hercules, CA, USA) was used as molecular weight marker (5 μL), and protein migration was promoted by the electric current generated (200 V, 40 minutes) in a Mini-PROTEAN II apparatus (Bio-Rad). Afterwards, gels were rinsed with distilled water and then covered with a 40% (v/v) methanol and 10% (v/v) acetic acid aqueous solution, used for gel fixation for one hour.

After fixation, gels were washed again and then covered with Bio-Safe™ Coomassie (Bio-Rad) for at least 2 hours, and the excess of dye was removed from gel with distilled water for 2 more hours. Finally, the relative amounts of target proteins could be quantified by densitometry using the Bio-Rad Gel Doc™ EZ System and the specific software Image Lab™.

3.7.4. Enzymatic activity assessment

The production yield of all the target proteins overexpressed was not only measured in terms of protein concentration but also in terms of enzymatic activity. All measurements were carried out by using a Cary-50Bio UV-visible spectrophotometer (Agilent Technologies, Santa Clara,

CA, USA). For soluble enzyme samples, a 2 mL and 1 cm-depth polystyrene BRAND® UV cuvettes (Sigma Aldrich) were used at final assay volume of 1.5 mL; for immobilised enzyme samples 3.5 mL quartz cuvettes HELMA® 100-QS (Hellma Analytics, Mülheim, Germany) with magnetic stirring were used, at final assay volume of 3 mL. All activity values samples were analysed for triplicate, and the temperature was controlled and set at 25°C for all measurements.

One unit of enzyme activity (AU) is defined as the amount of enzyme required to catalyse the conversion of 1 µmol of substrate per minute at 25°C.

Alcohol dehydrogenase (ADH) activity present in lysates and in immobilised derivatives was determined by following spectrophotometrically the formation of NADH at 340 nm of wavelength —being the molar extinction coefficient (ϵ) of β -NADH of $6.22 \text{ mM}^{-1}\cdot\text{cm}^{-1}$ — as a consequence of the enzyme conversion of ethanol to acetaldehyde that requires the presence of NAD^+ as cofactor. The reaction mixture contained ethanol at 543.6 mM, β - NAD⁺ at 7.5 mM, 20 mM of phosphate buffer (pH 8.80), and 50 µL of enzyme sample.

Pyruvate decarboxylase (PDC) activity was measured by following spectrophotometrically the depletion of NADH at 340 nm of wavelength ($\epsilon = 6.22 \text{ mM}^{-1}\cdot\text{cm}^{-1}$) caused by the conversion of acetaldehyde to ethanol (which is a coupled reaction performed by ADH) that requires the presence of NADH as cofactor. The reaction mixture contained 33 mM sodium pyruvate, 0.11 mM β - NADH, $3.5 \text{ AU}\cdot\text{mL}^{-1}$ ScADH (commercial) and 0.1 mM TPP and 0.1 mM MgCl_2 in citrate buffer 200 mM (pH 6.0), plus 50 µL of enzyme sample.

Lactate dehydrogenase (LDH) activity was also measured by following spectrophotometrically the depletion of NADH at 340 nm of wavelength ($\epsilon = 6.22 \text{ mM}^{-1}\cdot\text{cm}^{-1}$) but due to the conversion of pyruvate to lactate. In this case, reaction mixture contained 33 mM sodium pyruvate, 0.11 mM β - NADH, 50 mM of phosphate buffer (pH 7.6) and 50 µL of the sample.

In all cases, enzyme activity was determined once the difference of absorbance at 340 nm per minute was measured, using the following equation (Eq. 2):

$$\text{Activity (UA} \cdot \text{mL}^{-1}) = \frac{\Delta \text{Abs} \cdot \Delta T^{-1} \cdot V_t \cdot L}{\varepsilon \cdot V_{enz}} \quad (2)$$

Where:

- $\Delta \text{Abs} \cdot \Delta t^{-1}$ is the difference in absorbance per minute.
- V_t is the total assay volume (mL).
- L is the depth of assay cuvette (1 cm).
- ε is the molar extinction coefficient ($\text{mM}^{-1} \cdot \text{cm}^{-1}$).
- V_{enz} is the volume of enzyme sample added to the assay.

3.8 Biocatalysts immobilisation

Immobilisation experiments performed in this work were carried out by immobilising the three enzymes fused whether to the CBM3 domain from *C. thermocellum* or the CBM9 domain from *T. maritima*. Immobilisation experiments were performed aiming to establish a one-step purification/immobilisation process for the CBM–fused enzymes, based on affinity interactions between CBMs and cellulosic supports. Microcrystalline cellulose (MC) and regenerated amorphous cellulose (RAC) were used.

3.8.1 Regenerated amorphous cellulose preparation

Regenerated amorphous cellulose (RAC) cellulose was prepared from Avicel[®] PH-101 (Sigma) and Avicel[®] PH-200—which was kindly donated by DuPont[™] N&B (New York, NY, USA)—by following a well-described procedure (YH et al., 2006), consisting in acid treatment of microcrystalline cellulose at low temperature.

Firstly, 20 g of cellulose was suspended in 200 mL of cold distilled water (4°C), 10 mL aliquots of the suspension were mixed with 40 mL of 85% (v/v) orthophosphoric acid (H₃PO₄) in 50 mL falcon tubes, that were kept under refrigeration (4°C) and roller agitation. Afterwards, each falcon tube was diluted and mixed in 500 mL of ice–cold distilled water and the whole volume was vacuum filtered. RAC was rinsed out for a second time with the same volume of water, and 2 mL of 2 M sodium carbonate (Na₂CO₃) were added to neutralise any rest of acid. A last wash step with 500 mL of water was performed and the support was filtered again.

RAC was suspended in distilled water at a ratio of 10 grams of recovered support per litre and was stored under refrigeration (4°C). 1.5 mg of 2% (w/v) sodium azide (NaN₃) solution was added to avoid microbial contamination.

3.8.2. Immobilisation of fusion proteins to cellulosic supports

Immobilisation experiments were performed with 9 mL of cell lysate (see Section 3.6, Enzyme recovery) containing CBM3- or CBM9-fused enzymes and 1 mL of filtered support. Immobilisation processes were carried out at room temperature (24°C) with roller bottles and all experiments were done for triplicate. Supernatant and suspension activities were analysed over time until a steady state was reached and in all cases the activity of a blank (no support) was also monitored to ensure that the enzyme activity was not affected by protocol's conditions.

Several samples were taken along the experiment time and enzyme activity was measured for triplicate. Error bars correspond to standard error of all activity measurements.

Immobilisation yield (IY) and retained activity (RA) were calculated using the following equations (Eq. (3) and (4), respectively):

$$IY (\%) = \frac{A_i - A_{sn}}{A_i} \cdot 100 \quad (3)$$

$$RA (\%) = \frac{A_{sus} - A_{sn}}{A_i} \cdot 100 \quad (4)$$

Where:

- A_i is the initial enzyme activity offered.
- A_{sn} is the remaining enzyme activity measured in supernatant at the end of experiment .
- A_{sus} is the enzyme activity of the suspension (supernatant plus support) measured at the end of experiment.

In order to determine the immobilisation parameters, small amounts of enzyme were loaded ($\leq 50 \text{ AU}\cdot\text{mL}^{-1}$ of support) so that diffusional limitations were minimized. Nevertheless, for maximum load capacity assessment, cell lysate volumes used were higher than 9 mL since gra-

ter enzyme amounts were required to determine the maximum activity units that can be immobilised in 1 mL of cellulosic support, accepting this way the presence of diffusional limitations.

3.8.3. Immobilised derivatives storage stability

Immobilised derivatives were recovered by vacuum filtration and were resuspended in fresh 50 mM Tris-HCl pH 7.50 buffer. Samples were kept under refrigeration and activity was measured after 1 ,2 ,3 ,4, 7, 15 and 30 days for blank samples and immobilised derivatives, aiming to determine the half-life time of both soluble and immobilised forms under storage conditions. Before measuring enzyme activity, samples were tempered (24°C) and homogenised at roller bottles.

3.8.4. Eluent screening

CBM-tagged *ScADH* soluble enzymes obtained from cell lysates (2 mL) were incubated at room temperature (24°C) under roller agitation with different Tris-HCl pH 7.50 buffer solutions containing 3.6 M D-sorbitol, 10.2 M glycerol, 13.4 M ethylene glycol (EG), 1 M D-glucose, 0.9 M D-mannitol, 1 M Polyethylene glycol (PEG), and 3.2 M Xylitol at final volume of 10 mL. Enzyme activity of each vial was analysed for 2 h to assess enzyme stability towards these compounds, that were considered as possible eluent candidates. All experiments were performed by triplicate, including a negative control (blank) containing 2 mL of CBM-*ScADH* soluble enzymes and 8 mL of Tris-HCl buffer.

Same compounds were subsequently used in an elution screening experiment of CBM9-*ScADH* fusion protein from RAC support, unless for D-sorbitol, which showed unsuccessful results at stability tests. 0.4 mg (0.375 mL) of a CBM9-*ScADH* immobilised derivative were incubated under the same conditions reported for stability test, at a final volume of 10 mL. The enzyme activity of the immobilised derivative suspended in Tris-HCl buffer was measured prior to start the eluent screening experiment, to check that samples kept ADH activity, and after mixing 1896 with the different solutions, supernatant activity was measured for 1 h, to assess the elution capability of each compound. Once again, all candidates were tested by triplicate.

3.9. Fast liquid protein chromatography

Fast liquid protein chromatography (FPLC) method was used for protein purification (Figure 3.8). An ÄKTA™ Pure 150 equip from Cytiva (Uppsala, Sweden) with a Pharmacia XK-16/20 column (30 mL) was used in all purification processes, unless otherwise stated. Process was carried out at room temperature (24°C) and all fractions were collected and subsequently analysed by Bradford assay, by SDS-PAGE electrophoresis and by activity assay.

Histidine-tagged enzymes were purified by **immobilised-metal affinity chromatography** (IMAC) with FPLC by using 10 mL of a crosslinked agarose derivatized with iminodiacetic acid (IDA) from ABT (Madrid, Spain) and functionalised with Ni²⁺ ions.

The column bead was conditioned with 2.5 column volumes (CV) of 50 mM imidazole and 500 mM NaCl in 50 mM Tris-HCl buffer (pH 7.50) and 20 mL of cell lysate containing the target histidine-tagged enzymes were loaded onto the column at 1 mL·min⁻¹—establishing a residence time of 20 minutes—. Column was washed with 4 CV of the same solution used for equilibration, and finally target enzymes were eluted with 2 CV of 500 mM imidazole in 50 mM phosphate buffer (pH 7.50).

A second step consisting in a size-exclusion chromatography was required to remove imidazole (desalting) by using a Pharmacia XK-16/70 column packed with 120 mL of Sephadex® G-10 support (Merck Life Science, Darmstadt, Germany). Column was conditioned with 2 CV of 10 mM phosphate buffer (pH 7.50) and sample was loaded onto the column at 3.5 mL·min⁻¹.

For purification of CBM-fused enzymes, a **cellulose-based affinity chromatography** process based on previous reports (Boraston et al., 2001) was carried out, using 15 mL of Avicel® PH-200-derived RAC. 10 mL of the clarified cell lysate containing the target CBM-fused protein were loaded at 0.5 mL·min⁻¹ onto the column—previously conditioned with 2 CV of 50 mM Tris-HCl buffer, pH 7.50—. The column was first washed with 5 CV of 200 mM NaCl in 50 mM Tris-HCl buffer, pH 7.50 (5 mL·min⁻¹) and with 3 CV of 50 mM Tris-HCl buffer, pH 7.50 (5 mL·min⁻¹) afterwards. The bounds between CBMs and RAC were desorbed with 3 CV of 2 M glucose in 50 mM Tris-HCl buffer, pH 7.50 (2 mL·min⁻¹).

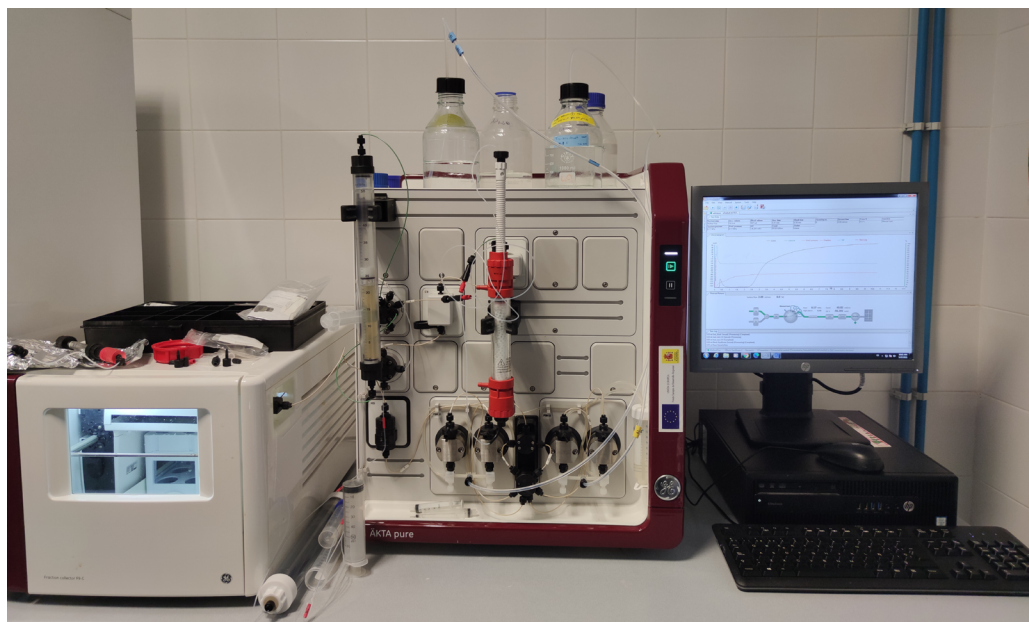


Figure 3.8. Picture of the actual FPLC setup.

4. RESULTS I: HISTIDINE-TAGGED ENZYMES OVEREXPRESSION IN *E. COLI*

Abstract

The bacteria *E. coli* is the main workhorse used for cloning, molecular biology and recombinant protein expression purposes. Hence, using one single *E. coli* strain to develop all these tasks could result in a simplification of the production process of a newly cloned protein of interest. In this work, the *E. coli* strain NEB 10- β was intended to be used as a polyvalent and transversal strain, assessing its capability to be effectively used for the expression of the histidine-tagged enzymes involved in L-lactic acid synthesis, obtaining between 20 and 54 mg of target enzymes per gram of biomass by following a two-step fed-batch strategy with defined culture medium. Nevertheless, aiming to assess if higher production yields could be achieved with a typically-used *E. coli* strain, the three target enzymes were cloned into an in-house developed expression vector by following a sequence- and ligation-independent cloning method (SLIC). Afterwards, the resulting fusion proteins were produced by using the auxotrophic M15 Δ glyA strain, with an antibiotic-free minimum culture medium. Specific production range varied from 45 and 41 mg of His-ScADH and TmLDH per gram of biomass, up to 85 mg·g⁻¹ DCW of His-ScPDC, corresponding to 7300, 2400 and 1900 AU·g⁻¹DCW, respectively. Finally, production parameters were compared not only between the two *E. coli* strains but between the different production strategies followed to produce histidine-tagged enzymes.

4. RESULTS I: HISTIDINE-TAGGED ENZYMES OVEREXPRESSION IN *E. COLI*

4.1. Overexpression studies with the NEB 10- β strain

Given the multi enzymatic system designed to produce lactic acid (described in Section 2.1, Context of the thesis work), it was firstly necessary to obtain enough amount of the three enzymes involved in it before trying to produce the target molecule. In that context, the easiest and fastest way to proceed with the biocatalysts production was to use the existing recombinant protein production (RPP) platform, already developed by the project partner (*Rijksuniversiteit Groningen*, RUG) responsible for enzyme engineering.

This specific platform in which *ScADH*, *ScPDC* and *TmLDH* were overexpressed consisted in the bacterial *E. coli* NEB 10- β strain, with the three corresponding genes cloned into the pBAD vector (each one for separate), which expression is regulated by the L-arabinose-inducible PBAD promoter (see Section 3.2, Plasmids).

Although NEB 10- β strain is widely used for molecular biology tasks thanks to its high cloning efficiency, it is not usually employed to obtain high titres of proteins at reactor scale —where other *E. coli* strains such as BL21 and K-12 derived strains are preferred (Terpe, 2006)—, since NEB 10- β strain requires an extra amino acid supply. However, using molecular biology strains to obtain high titres of a newly cloned protein of interest without needing to switch to a classical production strain would definitely result in a simpler development of the RPP process, especially if the main objective relies on using the overexpressed proteins rather than optimising an industrial production process.

Therefore, the first objective of the experiments compiled in this chapter of results was to produce successfully the three enzymes, all of them containing an N-terminal hexa-histidine tag in NEB 10- β . Firstly, preliminary experiments were performed with Terrific broth (TB) complex medium at shake-flask scale, aiming to assess the correct overexpression of the target enzymes and to generate a cell stock to start working with. Afterwards, enzyme production was scaled-up to bench reactor, and finally it was tested the viability to produce the three enzymes by following a two-step fed-bach approach, using glucose and glycerol as

main carbon sources of a defined minimum medium (DM) developed specifically to fulfil NEB 10- β requirements (Miret et al., 2020).

Besides, it was considered interesting to evaluate the production capability of *E. coli* NEB 10- β strain depending on the culture medium used (complex or defined) and the production scale, aiming to assess if the use of defined culture mediums is not only a cost-effective but a more productive option, in contrast to complex mediums.

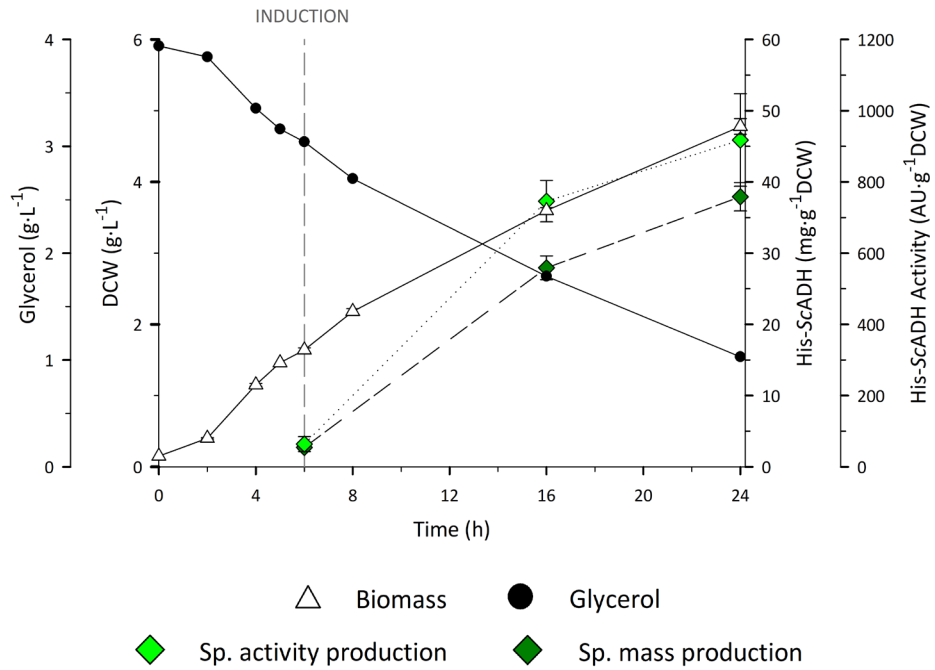
Thus, in addition to the fermentation processes, most relevant production parameters calculated for each of the three enzymes of interest are presented, trying to draw general conclusions that may help to elucidate the behaviour of *E. coli* NEB 10- β strain in RPP processes.

As can be noticed in the following sections, production processes were carried out at a temperature of 24 or 37°C depending on the culture strategy followed. Shake-flask and reactor-scale batch production processes were performed at 24°C, as a strategy to reduce not only the formation of toxic by-products but also of inclusion bodies, by lowering the nutrient uptake and growth rate, since induction took place with high concentrations of substrate (see Section 1.3, *Escherichia coli* as a platform for recombinant protein production). On the other hand, fed-batch cultures were always performed at 37°C, trying to avoid excessive duration of the first batch phase (where protein overexpression was not induced), and using instead the limitation of feeding influx to regulate cell metabolism during the fed-batch phase of the process.

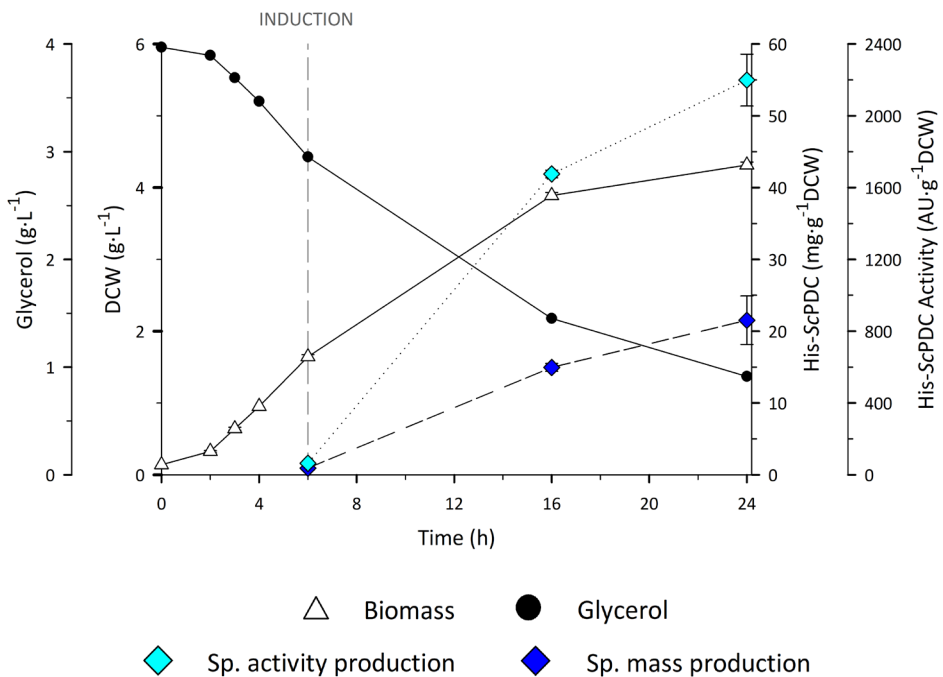
4.1.1. Shake-flask overexpression experiments with complex medium

As mentioned, first preliminary experiments were carried out at flask scale with complex TB medium, as described in Section 3.5, Cultivation conditions. The evolution of overexpression processes is depicted in [Figure 4.1](#):

A: His-ScADH



B: His-ScPDC



C: His-*Tm*LDH

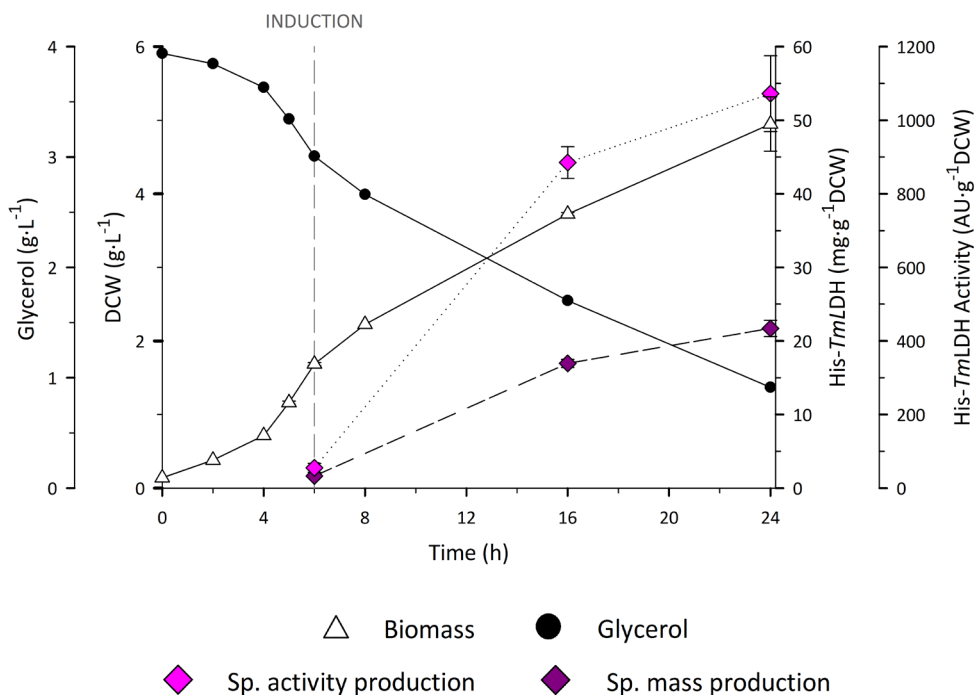


Figure 4.1. *E. coli* NEB 10- β flask cultures performed to produce histidine-tagged ScADH (A), ScPDC (B) and *Tm*LDH (C) enzymes with TB medium at flask scale. Profiles along time of biomass DCW (Δ) and glycerol (\bullet) (g·L⁻¹), specific activity (\blacklozenge) (AU·g⁻¹DCW) and specific mass production (\blacklozenge) (mg·g⁻¹DCW). Induction phase indicated. Culture conditions: 24°C, pH 7.0, 140 rpm, 0.1 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹ ampicillin. Error bars correspond to standard error of culture duplicates.

Regarding the development of the processes, similar timings were observed among the three cases, reaching in all cases between 4 and 5 grams of biomass per culture litre, with maximum specific growth rates (μ_{\max}) between 0.4 and 0.5 h⁻¹. Even if glycerol was not totally exhausted in none of the cases, it was decided to stop cultures after 24 hours.

Preliminary overexpression experiments at shake-flask scale showed promising results, obtaining more than 4 mg of active target enzymes in all cases. According to protein electrophoresis (results not shown), target enzymes represented up to a 13% of total protein content. These results suggest that protein overexpression was correctly repressed in absence of the inductor L-arabinose.

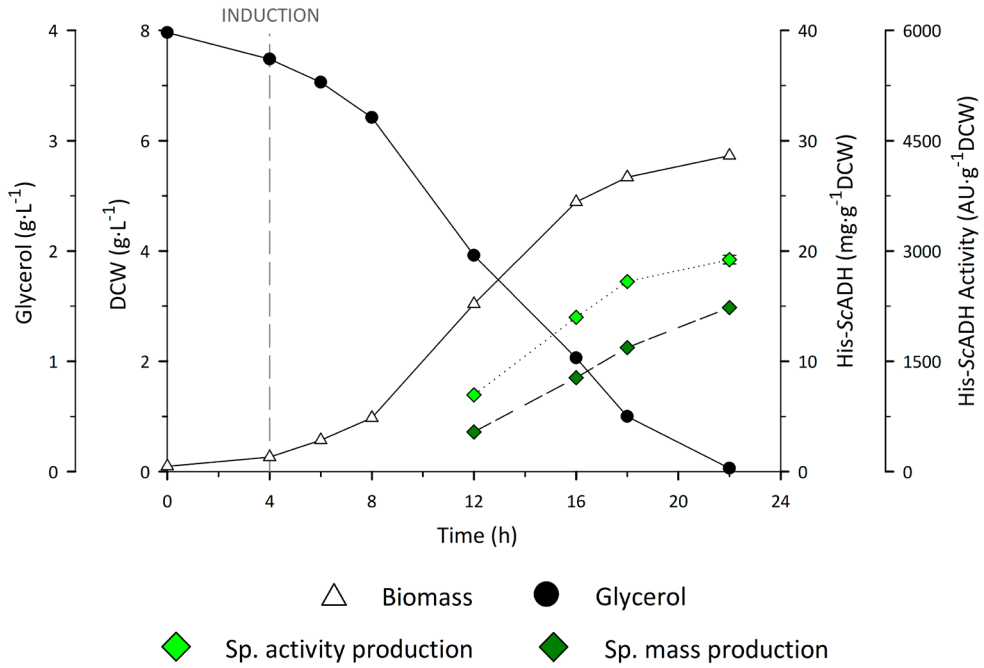
4.1.2. Batch production processes at bench–scale reactor with complex medium

After achieving the overexpression of the three candidate enzymes in NEB 10- β strain, the following experiments were carried out, aiming to assess the protein overexpression at reactor scale, to determine if production yields could be improved with the scale-up from flasks to reactor, with controlled PO₂ and pH conditions TB medium.

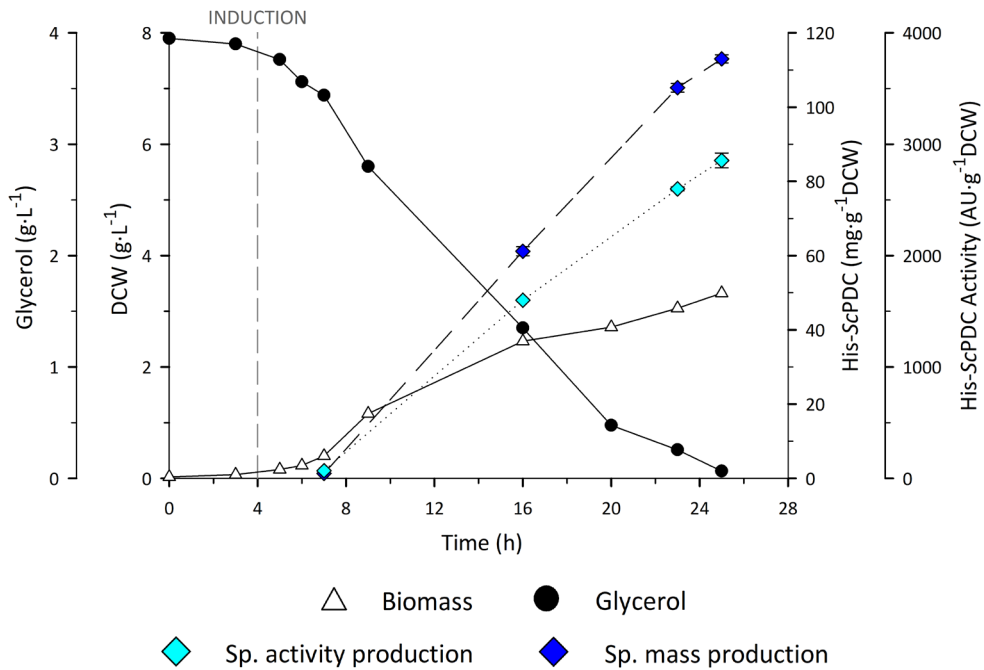
Fermentation parameters followed along the batch cultures, and most relevant production parameters calculated for each fermentation are depicted in [Figure 4.2](#).

Regarding the development of the processes, similar timings were observed among the three cases, and initial glycerol was totally exhausted in all fermentations. However, profile curves corresponding to specific mass and specific activity productions showed different tendencies among the three cultures; A positive shift was observed at the end of His-ScPDC and His-TmLDH producing processes ([Figures 4.2 B and C](#), respectively), whereas it seemed that His-ScADH production started to slow at the end of the culture ([Figure 4.2 A](#)).

A: His-ScADH



B: His-ScPDC



C: His-*Tm*LDH

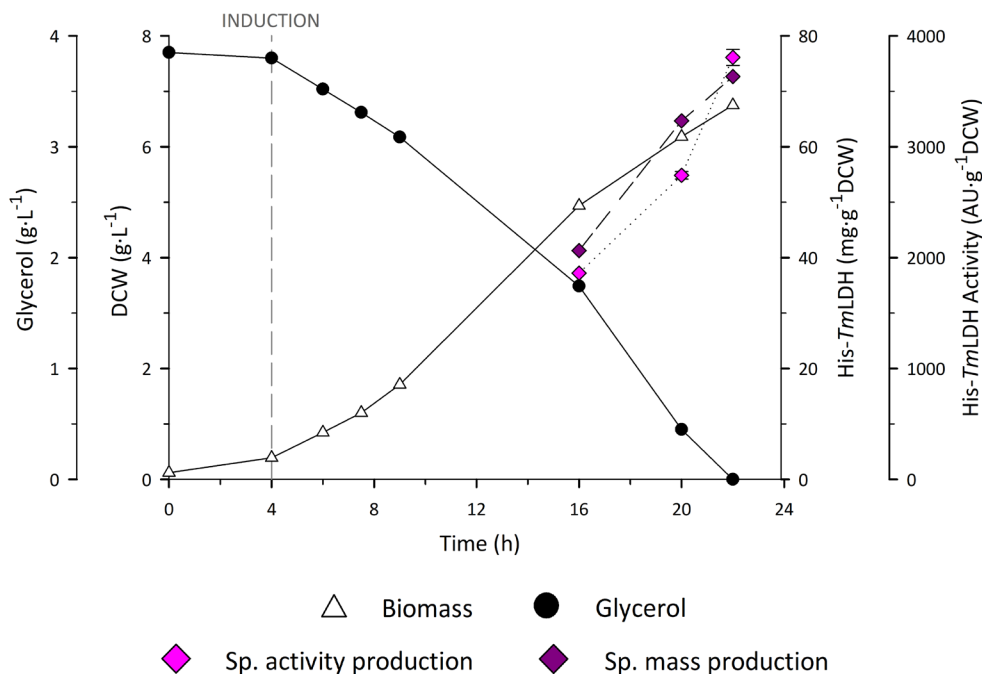


Figure 4.2. *E. coli* NEB 10- β batch cultures performed to produce histidine-tagged *Sc*ADH (A), *Sc*PDC (B) and *Tm*LDH (C) enzymes with TB medium. Profiles along time of biomass DCW (Δ) and glycerol (\bullet) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Induction phase indicated. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 30% PO_2 , 0.2 $\text{g}\cdot\text{L}^{-1}$ L-arabinose, 0.1 $\text{g}\cdot\text{L}^{-1}$ ampicillin. Error bars correspond to standard error of three sample measurements.

Below, most relevant production parameters calculated for each enzyme are shown in [Table 4.1](#). Nevertheless, the intention of the author is not to establish a comparison among the different enzymes overexpressed, but to discuss the differences observed among the two culture scales (flasks and reactor) used for protein production.

Table 4.1. Production parameters calculated for histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes using *E. coli* NEB 10- β strain and TB medium in 50 mL shake flask and 5 L batch cultures. Shake-flask culture conditions: 24°C, pH 7.0, 140 rpm, 0.1 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹ ampicillin; \pm correspond to the standard error of culture duplicates. Reactor conditions: 24°C, pH 7.0, 450-1200 rpm, 30% PO₂, 0.2 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹ ampicillin; \pm correspond to the standard error of three sample measurements.

Production parameter	Scale	His- <i>ScADH</i>	His- <i>ScPDC</i>	His- <i>TmLDH</i>
Specific production (mg Enz·g ⁻¹ DCW)	Flask	38 \pm 2	22 \pm 3	22 \pm 1
	Reactor	15 \pm 0	113 \pm 3	73 \pm 0
Specific production (AU·g ⁻¹ DCW)	Flask	918 \pm 130	2199 \pm 144	1072 \pm 103
	Reactor	2883 \pm 59	2854 \pm 50	3860 \pm 72
Specific activity (AU·mg ⁻¹ Enz)	Flask	24 \pm 2	102 \pm 20	61 \pm 6
	Reactor	194 \pm 4	25 \pm 0	52 \pm 1

Results showed that, in general, specific production increased from flask to reactor cultures. In mass terms, histidine-tagged *ScPDC* and *TmLDH* specific productions were 5- to 3-fold higher than shake-flask values, although His-*ScADH* yield decreased. Besides, in terms of enzyme specific production, the three processes showed improved yields in contrast to flask cultures, possibly due to the better control of culture variables that bioreactor setup can offer.

Though, it was observed a significant increase in enzyme specific activity for *ScADH* and a decrease for *ScPDC* enzyme when scaling up their production. These results can somehow point out a low reproducibility of production (and protein folding) efficiency of this microbial platform in TB medium cultures. Besides, it is widely reported that in some cases the heterologous production of some eukaryotic enzymes in bacterial strains can lead to lower production yields and lower productivities, especially for enzymes such as *ScPDC*, to the best knowledge of the author.

4.1.3. Shake-flask overexpression experiments with defined medium

As previously stated, it was considered interesting to assess the viability to produce the three enzymes by using minimum defined medium (DM) instead of using complex medium. This change was promoted mainly because the use of chemically defined minimum mediums is one alternative that grants a better reproducibility and control of performance in cell culture, better productivity and product quality as well as improved communication and transfer of the formulation, fulfilling new quality and regulatory requirements of biotechnology industry (Whitford et al., 2018). Moreover, using DM allows to control cell growth by limiting the carbon source, what enables to conduct effectively fed-batch cultures for process intensification.

The defined medium used contained glucose as a carbon source, and it also had to be supplemented with leucine, isoleucine and valine amino acids, since the NEB 10- β strain is auxotroph for the two first acids and, in addition, this auxotrophy also affects the genes involved in the biosynthesis pathway of valine. Thus, first growth phase with glucose as carbon source was followed by a medium switch to glycerol-based medium —supplemented with arabinose— to avoid glucose-mediated carbon catabolite repression of the *araBAD* promoter.

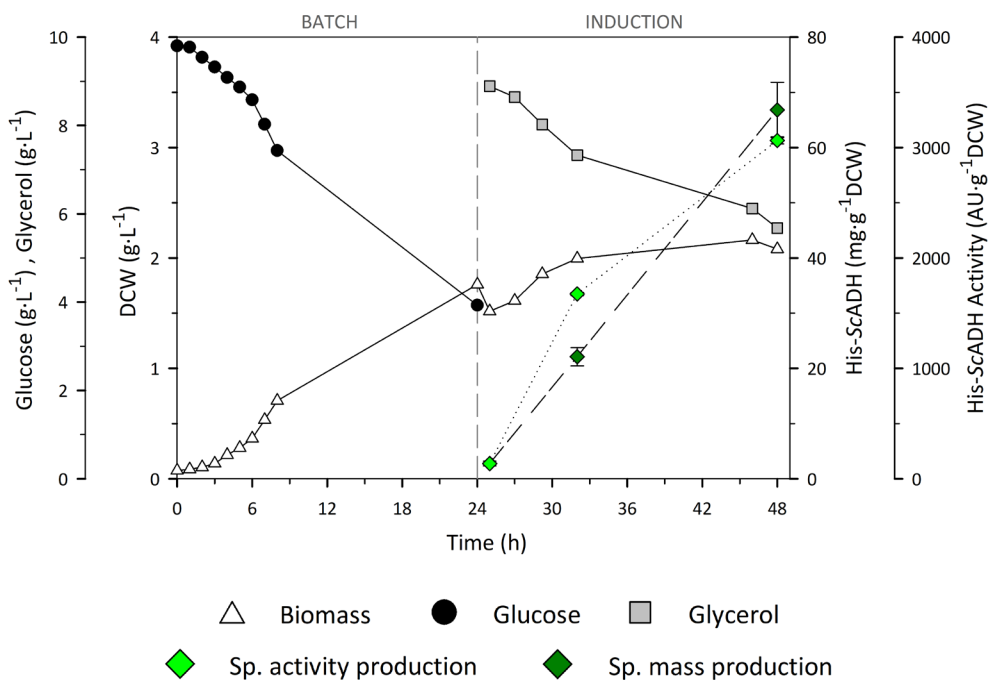
Nevertheless, prior to produce the target enzymes with *E. coli* NEB 10- β , cells were adapted to defined medium through a three-step cell growth at flask-scale (see Section 3.5, Cultivation conditions). This process was useful to i) check that *E. coli* cells grew correctly with DM, reaching an exponential phase of growth, and ii) estimate the maximum specific growth rate (μ_{\max}) of the strain under the setup conditions, by following the cell growth along the three cycles of the adaptation process for each variant.

The adaptation resulted in the generation of the cell cryostocks that were used as pre-inoculums in subsequent production experiments with DM. Biomass evolution throughout the culture cycles showed that cell growth with DM occurred as expected, presenting similar profiles from cycle to cycle (see Figure 8.1, from the Appendix section). Therefore, cultures derived from the third step were used to generate the corresponding cell banks.

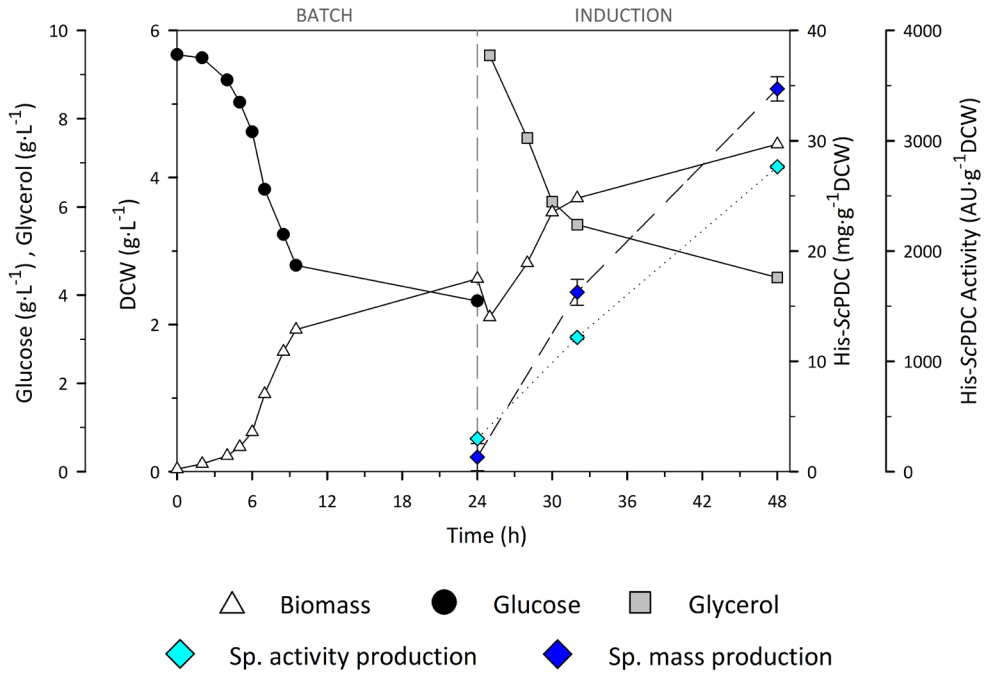
First overexpression experiments with DM were carried out at flask scale, aiming to assess the enzyme production with the new culture medium. As previously stated, it was required to perform a medium exchange (see Section 3.5, Cultivation conditions) from glucose- to glycerol-based culture medium, necessary to induce protein overexpression.

Three flasks were inoculated with the corresponding variants, and two of them were induced afterwards, whilst the third culture remained as a negative control. Process parameters followed along the cultures and calculated specific productions of each enzyme are shown below in [Figure 4.3](#) and in [Table 4.2](#).

A: His-ScADH



B: His-ScPDC



C: His-TmLDH

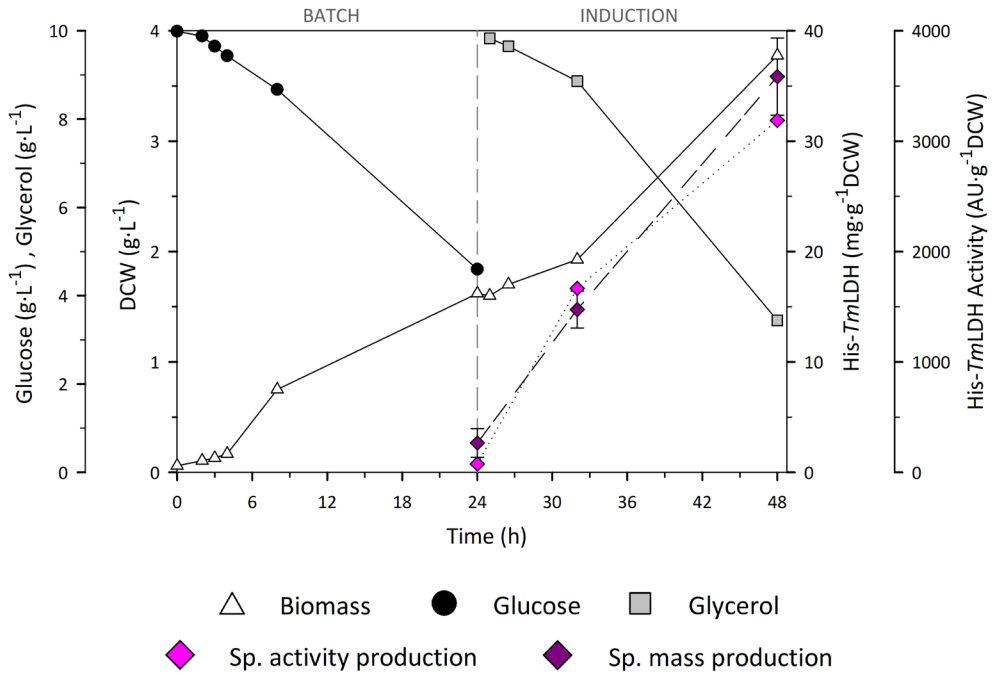


Figure 4.3. *E. coli* NEB 10- β flask cultures performed to produce histidine-tagged ScADH (A), ScPDC (B) and TmLDH (C) enzymes with DM at flask scale. Profiles along time of biomass DCW (Δ), glucose (\bullet) and glycerol (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge \blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (\blacklozenge \blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Induction phase (and medium replacement) indicated. Culture conditions: 24°C, pH 7.0, 140 rpm, 0.1 $\text{g}\cdot\text{L}^{-1}$ L-arabinose, 0.1 $\text{g}\cdot\text{L}^{-1}$ ampicillin. Error bars correspond to standard error of culture duplicates.

Table 4.2. Production parameters calculated for histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes using *E. coli* NEB 10- β strain and DM in 50 mL shake flask cultures. Culture conditions: 24°C, pH 7.0, 140 rpm, 0.1 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹ ampicillin. \pm correspond to the standard error of culture duplicates.

Production parameter	His- <i>ScADH</i>	His- <i>ScPDC</i>	His- <i>TmLDH</i>
Specific production (mg Enz·g ⁻¹ DCW)	67 \pm 5	35 \pm 1	34 \pm 3
Specific production (AU·g ⁻¹ DCW)	3065 \pm 32	2764 \pm 16	3190 \pm 92
Specific activity (AU·mg ⁻¹ Enz)	46 \pm 5	80 \pm 9	90 \pm 3

Promising results were obtained, since NEB 10- β cells proved to overexpress correctly target enzymes by using DM and at higher production yields than with complex medium, apparently. Differences observed between the two media and between induced and control shake-flask cultures are shown in [Figure 4.4](#):

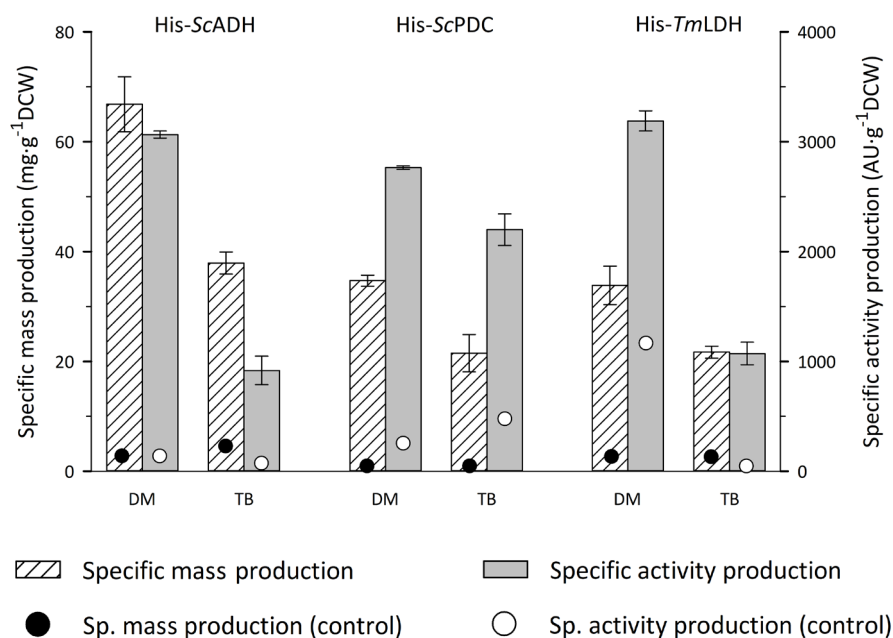


Figure 4.4. *E. coli* NEB10- β specific activity (AU·g⁻¹DCW) and specific mass (mgEnz·g⁻¹DCW) production values obtained in 50 mL flask cultures performed to produce histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes with DM and TB mediums. Culture conditions: 24°C, pH 7.0, 140 rpm, 0.1 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹ ampicillin. Error bars correspond to standard error of culture duplicates.

As can be seen in [Figure 4.4](#), higher yields were reported for flask cultures with defined medium rather than for flask cultures with complex medium in all cases. Moreover, remarkable differences were shown in contrast to negative controls.

According to SDS-PAGE (not shown), His-*ScADH* enzyme corresponded to a 11% of total protein content in the supernatant of induced cultures with DM. Similar values were reported for His-*TmLDH* enzyme (16% of relative quantity), and even higher were observed among His-*ScPDC* producing cultures (30% of relative quantity). Similar differences between induced cultures and negative controls than those observed in flasks with complex medium were reported, what meant that overexpression system was correctly repressed in absence of arabinose.

4.1.4 Fed-batch production processes at bench-scale reactor with defined medium

Finally, after having achieved positive results in flask-scale production tests with the use of defined medium, following experiments aimed to assess the production of target enzymes at 2 L reactor scale. The fermentation process consisted of a batch phase with glucose as carbon source and a subsequent fed-batch phase, in which a first exponential glucose-based feeding was combined with a later glycerol-based feeding, as described in Section 3.5, Cultivation conditions. The shift of the culture feeding solution meant the beginning of the induction phase, as happened with shake-flask cultures.

His-*ScADH* was the first enzyme produced at reactor scale and was used as a model enzyme, whilst histidine-tagged *ScPDC* and *TmLDH* enzymes productions were useful to validate the reproducibility of the process platform.

Fermentation parameters followed along the fed-batch culture, and enzyme specific activity and specific mass production analysed afterwards are depicted in [Figure 4.5](#), and the evolution of the fed-batch cultures performed to obtain the other two histidine-tagged enzymes are provided in [Figures 8.2](#) and [8.3](#) of the Appendix section.

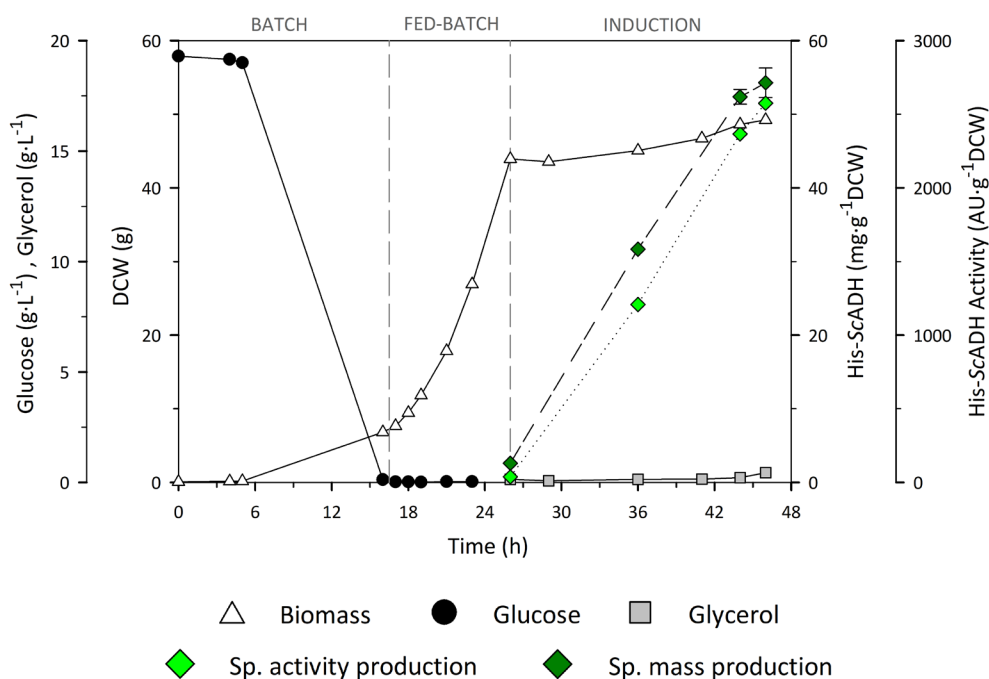


Figure 4.5. *E. coli* NEB 10- β fed-batch culture performed to produce His-ScADH enzyme with DM. Profiles along time of biomass DCW (Δ) (g), glucose (\bullet) and glycerol (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Batch, fed-batch and induction phases indicated. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 30% PO_2 , 0.2 $\text{g}\cdot\text{L}^{-1}$ L-arabinose, 0.1 $\text{g}\cdot\text{L}^{-1}$ ampicillin. Error bars correspond to standard error of three sample measurements.

The fed-batch culture was developed as expected; NEB 10- β cells grew during batch phase at a biomass/substrate yield ($Y_{X/S}$) of $0.35 \text{ gDCW}\cdot\text{g}^{-1}$ of glucose, and afterwards, the feeding with glucose was added and biomass grew at a μ of 0.2 h^{-1} , until the desired biomass concentration was reached. At that point, glucose was shifted for glycerol, and the culture was induced by adding arabinose to the culture. In Figure 4.5, biomass values are provided as total grams of DCW instead of $\text{gDCW}\cdot\text{L}^{-1}$, since it was considered more accurate considering that volume was not constant during fed-batch phase.

As can be seen, specific production profiles still showed an increasing trend at the end of the process, meaning that probably, production yields could be further optimised. As example, Miret et al. achieved significantly higher His-ScADH production yields when a linear decreasing of temperature from 37 to 24°C was applied, and the results obtained

were published together with the production of a fusion protein phosphite dehydrogenase-cyclohexanone monooxygenase, as a validation of the developed production platform (Miret et al., 2020).

Specific production values calculated for the three fed-batch fermentations with DM are listed in Table 4.3. Results showed that the three target enzymes were successfully produced, although different specific production values were obtained depending on the enzyme.

Table 4.3. Production parameters calculated for histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes using *E. coli* NEB 10- β strain and DM in 2 L fed-batch cultures. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 30% PO₂, 0.2 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹ ampicillin. Error bars correspond to standard error of three sample measurements.

Production parameter	His- <i>ScADH</i>	His- <i>ScPDC</i>	His- <i>TmLDH</i>
Specific production (mg Enz·g ⁻¹ DCW)	54 ± 2	25 ± 1	19 ± 2
Specific production (AU·g ⁻¹ DCW)	2575 ± 5	226 ± 13	1240 ± 9
Specific activity (AU·mg ⁻¹ Enz)	47 ± 0	9 ± 1	64 ± 0

Finally, it was considered interesting to compare the production yields obtained for one of the enzymes (His-*ScADH*), depending on the fermentation approach followed, aiming to assess if the **change of the production strategy** from TB batch to DM fed-batch cultures was as effective as expected, even if none of the processes were optimised. Therefore, the parameters compared among the different strategies depended on the production process, i.e., on final volume and total biomass reached.

As can be seen in Table 4.4, calculated values suggest that a fed-batch strategy with DM should be considered the most effective process to obtain His-*ScADH* enzyme. Regarding the enzyme concentration (mg·L⁻¹), almost a 20-fold increase was reported, corresponding to a 11-fold higher volumetric productivity. In terms of activity units, around a 5-fold increase was measured in final concentration (AU·L⁻¹), what meant a 2.3-fold increase in volumetric productivity.

Table 4.4. Productivity parameters calculated for His-ScADH enzyme using *E. coli* NEB 10- β strain at bench-scale reactor, depending on the culture medium used. \pm correspond to the standard error of three sample measurements.

Production parameter	Batch TB	Fed-batch DM
Total biomass (g DCW)	15.7 \pm 0	49.2 \pm 0
Final volume (L)	3.75	1.65
Enzyme titre (mg Enz·L ⁻¹)	85 \pm 0	1616 \pm 52
Enzyme activity (AU·L ⁻¹)	1.6·10 ⁴ \pm 340	7.7·10⁴ \pm 147
Volumetric productivity (mg Enz·L ⁻¹ ·h ⁻¹)	4 \pm 0	45 \pm 1
Volumetric productivity (AU·L ⁻¹ ·h ⁻¹)	751 \pm 15	1670 \pm 3

However, this productivity increase from batch to fed-batch cultures did not happen for the other two enzymes produced, especially in terms of enzyme activity. For example, 380 AU·L⁻¹·h⁻¹ of ScPDC were produced with TB medium, whereas 7 AU·L⁻¹·h⁻¹ were produced in fed-batch culture with DM. Moreover, 1070 AU·L⁻¹·h⁻¹ of *Tm*LDH were produced with TB medium, whilst 372 AU·L⁻¹·h⁻¹ were obtained with DM (Figures 8.2 and 8.3, respectively).

The production downshift observed between the two operational modes could be caused by some differences, including i) the effect of temperature, which may affect to protein folding capability, and ii) the effect of a possible amino acid imbalance in DM, which may hinder protein synthesis pathway.

These results suggest that probably RPP with *E. coli* NEB 10- β stain is a more robust expression system with the use of complex culture medium and at lower culture temperatures, rather than with DM, where high amounts of amino acids must be added as medium supplement, with all that implies (low solubilisation, increased final costs, etc.).

Thus, considering that the aim of this work is to produce recombinant enzymes with defined mediums and additionally avoiding the use of antibiotics, it would be interesting to explore other alternatives to produce the three enzymes, which are requested at high loads to perform biocatalytic processes.

4.2. Changing *E. coli* host strain from NEB 10- β to M15 Δ glyA

Although the three enzymes of interest were successfully overexpressed with the *E. coli* NEB 10- β strain, the vicissitudes associated RPP based on fed-batch fermentations with DM such as the need to change the carbon source in the middle of the process and the amino acid supplementation requirements, led to the decision to switch the microbial cell factory (MCF) for another *E. coli* strain more suitable for heterologous protein production by using the desired operational setup (fed-batch cultures with DM).

In that sense, a well-characterized and previously developed in our laboratory *E. coli* M15-derived strain (Vidal et al., 2008, Pasini et al., 2016) was chosen to overexpress target proteins. This enabled us to introduce a key feature; this mutant strain contains a deletion of the *glyA* gene, which encodes for serine hydroxymethyl transferase (SHMT), a 46 kDa enzyme involved in the main glycine biosynthesis pathway in *E. coli* when glucose is used as carbon source (Plamann and Stauffer, 1983). SHMT enzyme catalyses the reversible interconversion between serine and glycine, plus, it also catalyses the reversible conversion between L-threonine and glycine (Florio et al., 2010). This strain was successfully used in the past to overproduce DHAP-dependent aldolase FucA (Pasini et al., 2016), among other recombinant proteins.

The decision to change the MCF to produce target enzymes did not only affect to the *E. coli* strain selected but also to the plasmid into which genes of interest were cloned. The perspective of a fresh start was given as an opportunity to upgrade certain aspects of the production process, being one of the most interesting the possibility to induce protein overexpression with IPTG instead of L-arabinose by changing the regulatory operon, what led to a significant simplification of the process, since it was avoided the use of glycerol as carbon source.

Moreover, there was already developed one pQE-40-derived vector with the *lacI-glyA* cassette included, regulated by a constitutive promoter (J23110), which also contained an ampicillin-resistance gene, as previously described (see Section 3.2, Plasmids). The use of this vector in combination with the auxotrophic strain allowed us to apply selective pressure with two different mechanisms (ampicillin resistance and/or auxotrophy), depending on the need. Thus, ampicillin was used for

molecular biology experiments in which complex medium was used, but was avoided in subsequent overexpression experiments with DM.

The aim of the experiments described below was to establish the RPP platform based on the auxotrophic *E. coli* M15 Δ glyA strain and the use of DM and test the viability of the new transformants to express the three histidine-tagged enzymes in an active and soluble form.

Therefore, DNA fragments corresponding to histidine-tagged enzymes were cloned into the pVEF vector designed for protein overexpression by using a variation of the SLIC method (Jeong et al., 2012, Islam et al., 2017), as previously explained (Section 3.3, Molecular biology techniques). Once the cloning was achieved for all the target proteins, well-plate overexpression screening experiment with the resulting M15 Δ glyA transformant colonies were conducted to corroborate the capability of the cells to produce the target enzymes. In addition, colonies who showed higher overexpression levels were picked to generate the cell banks used for the subsequent production processes.

4.2.1. Molecular biology: cloning target genes into pVEF vector

Firstly, DNA fragments corresponding to the three histidine-tagged enzymes included in pBAD vectors were amplified by PCR (Figure 4.6) to obtain enough DNA to use in subsequent cloning experiments.

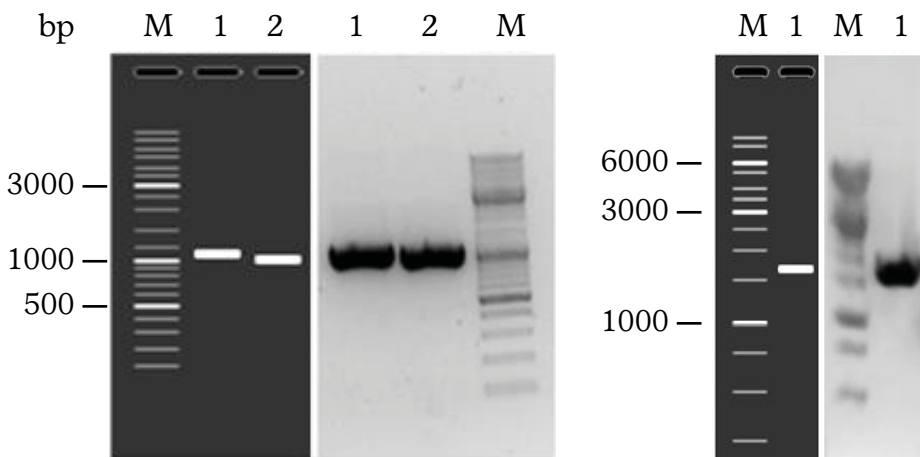


Figure 4.6. Agarose 1% DNA electrophoresis gel with amplified DNA fragments corresponding to PCR products of His-ScADH (1047 bp), His-TmLDH (960 bp) (left) and His-ScPDC (1692 bp) (right). Simulation of expected bands included. Lane M: molecular weight standard [bp].

Initial cloning experiments were performed with DNA fragments corresponding to encoding sequences for *ScADH* and *TmLDH* enzymes, which were the first to be obtained by PCR. Aiming to compare the transformation efficiency of two different *E. coli* strains by the heat-shock method, DH5 α and TOP10 competent cells were transformed with the resulting SLIC products, and in all cases several colonies were grown in agar plates, shown in [Figure 4.7](#):

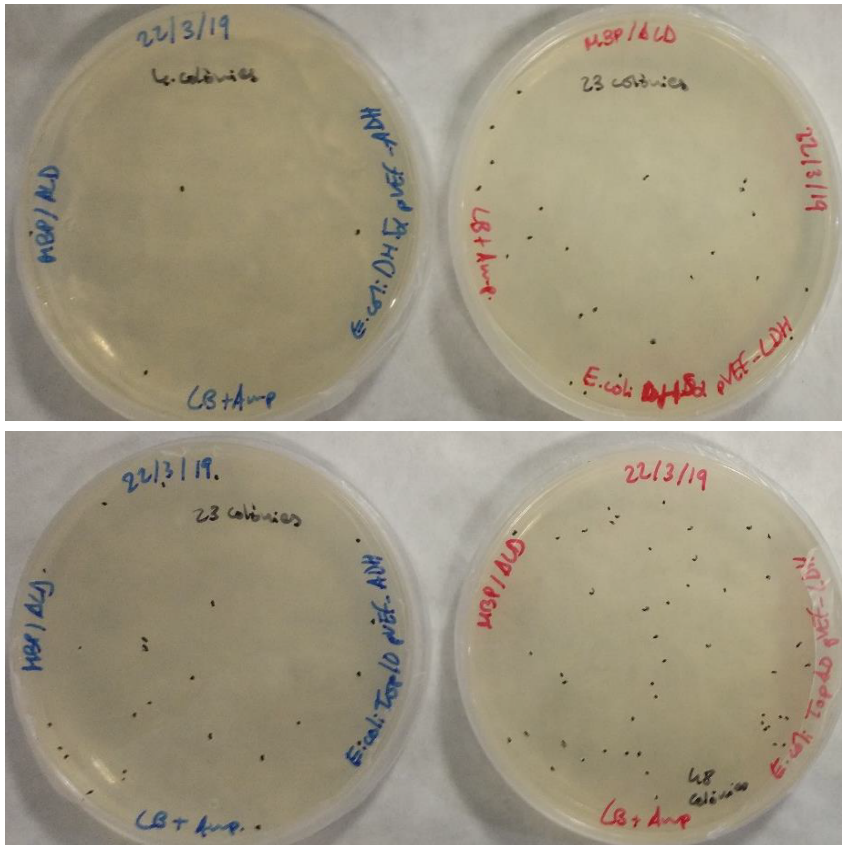


Figure 4.7. Petri dishes with LB-agar medium containing *E. coli* DH5 α (top) and TOP10 (bottom) competent cells transformed with pVEF_His-ScADH (blue) and pVEF_His-TmLDH (red) vectors obtained via SLIC method. Culture conditions: 37°C, 0.1 g·L⁻¹ ampicillin.

In order to assess the efficiency of the transformation, some colonies of each petri dish were tested via PCR, being all of them positive, as can be observed in DNA electrophoresis gels shown in [Figure 4.8](#). However, an amplified DNA band in negative control of PCR was detected, that corresponded to His-ScADH fragment, most probably due to a mani-

pulation mistake. Afterwards, one positive colony of each *E. coli* strain transformed with the two pVEF vectors were checked again via XhoI restriction digest (Figure 4.9).

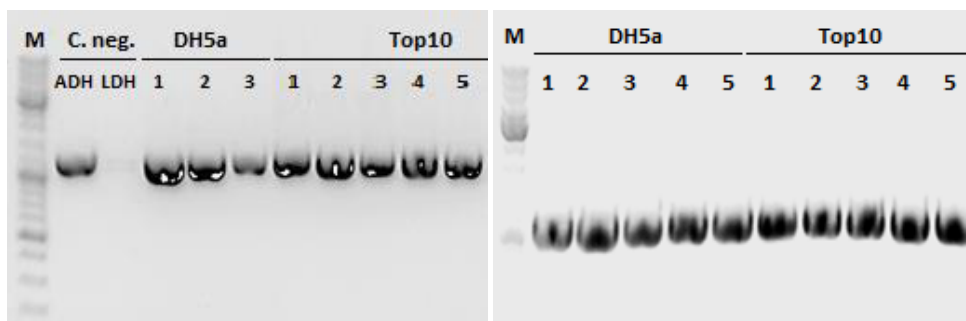


Figure 4.8. Agarose 1% DNA electrophoresis gels with colony PCR products corresponding to His-ScADH (1047 bp) (left) and to His-TmLDH (960 bp) (right) amplified fragments, from picked *E. coli* DH5a and TOP10 transformants. Negative control of PCR performed with the two pairs of oligonucleotides. Lane M: molecular weight standard (bp).

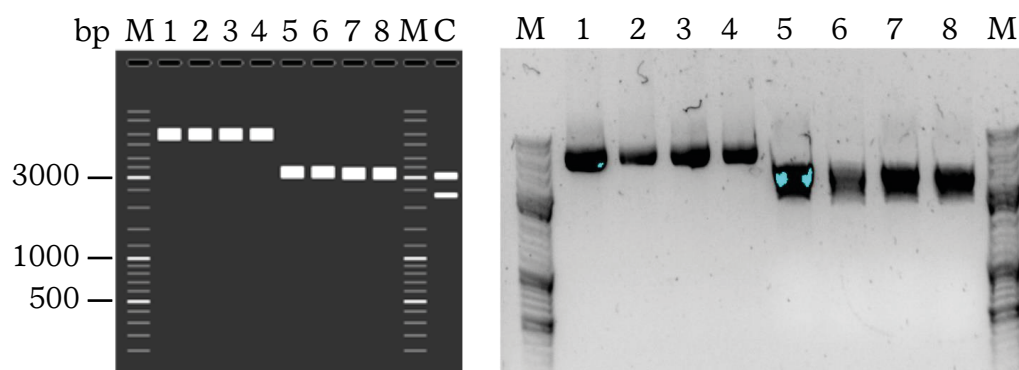


Figure 4.9. Simulation of expected restriction pattern, including the simulation of pVEF (without insert) pattern re-restriction cut by XhoI (left). Agarose 1% DNA electrophoresis gel with pVEF_His-ScADH and pVEF_His-TmLDH pattern restrictions cut by XhoI (right). Lanes M: molecular weight standard (bp); lanes 1 to 4: negative controls (uncut plasmids); lanes 5 and 6: pVEF_His-ScADH obtained from DH5a and TOP10 cells, respectively, digested with XhoI; lanes 7 and 8: pVEF_His-TmLDH obtained from DH5a and TOP10 cells, respectively, digested with XhoI.

Results obtained did not only suggest that both strains were suitable for plasmid DNA transformation, but it could also be assessed that SLIC technique was highly efficient, since all colonies picked from different plates incorporated the pVEF vector, with the corresponding inserts correctly cloned. Thus, these were considered key results, as they proved that this methodology was useful to establish a fast and versatile cloning procedure, enabling to transform DNA inserts to *E. coli* cells with a

high grade of success. In addition, positive results were also achieved for His-ScPDC coding DNA sequence (Figure 4.10), which was cloned into pVEF vector and subsequently transformed into *E. coli* DH5 α competent cells.

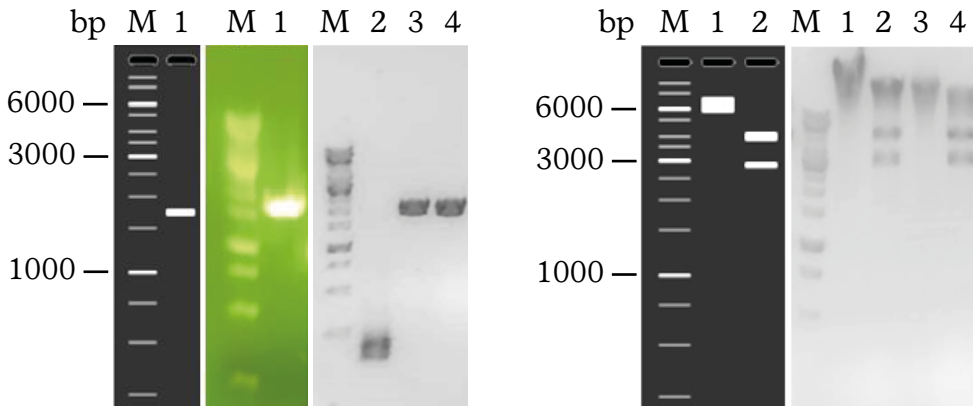


Figure 4.10. Agarose 1% DNA electrophoresis gel with amplified DNA fragments corresponding to PCR product of His-ScPDC (1692 bp) (left). Lanes M: molecular weight standard (bp); lane 1: amplified DNA fragment prior to SLIC; Lane 2: negative colony PCR product after SLIC; Lanes 3 and 4: positive colony PCR products after SLIC. And agarose 1% DNA electrophoresis gel with pVEF_His-ScPDC pattern restrictions cut by XbaI (right). Lane M: molecular weight standard (bp); lanes 1 and 3: negative controls (uncut plasmids); lanes 2 and 4: pVEF_His-ScPDC digested with XbaI. Simulation of expected bands included for both experiments.

As can be seen in Figure 4.10, two positive colonies —among the three picked— derived from *E. coli* DH5 α transformation with SLIC product showed correct DNA amplification and correct pattern restriction. Thus, results obtained in His-ScPDC cloning were useful to prove the robustness of SLIC method once again, that definitely turned out to be an effective and cost-saving cloning method.

Finally, the region of interest of the three newly cloned pVEF vectors were analysed by Sanger sequencing, showing no mutations.

The schematic structure of the new plasmid constructs obtained via SLIC that incorporated histidine-tagged ScADH, ScPDC and TmLDH enzymes is shown below in Figure 4.11, and more detailed information is provided in Section 3.2, Plasmids.

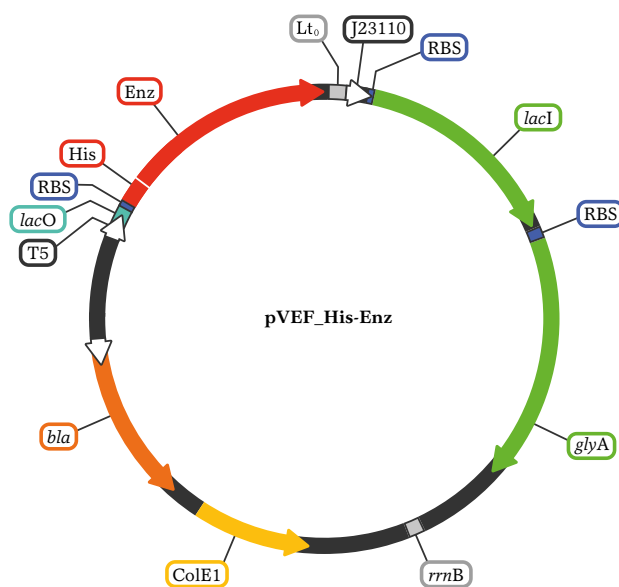


Figure 4.11. Schematic representation of pVEF expression vector containing a histidine-tagged enzyme coding DNA sequence. Features: T5 promoter; *lacO*, two *lac* operator regions; RBS, ribosome binding site; *Lto*, lambda to transcriptional termination; J23110 constitutive promoter; *lacI* gene; *glyA* gene; *rrnB* transcriptional termination; *bla*, ampicillin resistance gene; ColE1 replication origin; His, 6x histidine sequence; Enz, enzyme sequence.

4.2.2. *E. coli* M15 Δ *glyA* transformation and expression screening

Once the molecular biology experiments were satisfactorily accomplished, proving that *E. coli* DH5 α cells were able to produce the recombinant enzymes in induced cultures, it was required to change the host strain to overcome the subsequent overexpression experiments. Therefore, M15 Δ *glyA* competent cells were transformed with the three vectors since production processes at reactor scale were meant to be performed with M15 Δ *glyA* strain.

Firstly, an overexpression screening was performed to determine that M15 Δ *glyA* clones resulting from transformation were able to produce the target proteins, and to pick the most productive ones among several colonies randomly chosen for each of the variants. The corresponding SDS-PAGE gels are provided below:

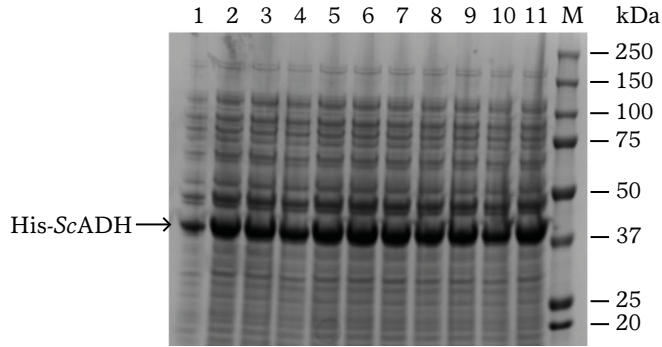


Figure 4.12. SDS-PAGE of *E. coli* M15ΔglyA final samples of the screening experiment for His-ScADH overexpression. Lane M: molecular weight standard (kDa); lanes 1 to 11: induced cultures. Culture conditions: 2 mL LB, 24°C, 24 h, 140 rpm, 0.4 mM IPTG. His-ScADH (37 kDa) corresponding band indicated with arrow.

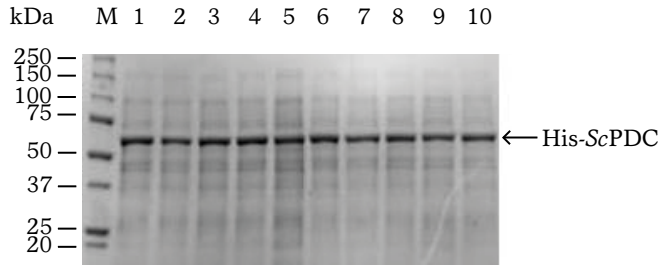


Figure 4.13. SDS-PAGE of *E. coli* M15ΔglyA final samples of the screening experiment for His-ScPDC overexpression. Lane M: molecular weight standard (kDa); lanes 1 to 10: induced cultures. Culture conditions: 2 mL LB, 24°C, 24 h, 140 rpm, 0.4 mM IPTG. His-ScPDC (63 kDa) corresponding band indicated with arrow.

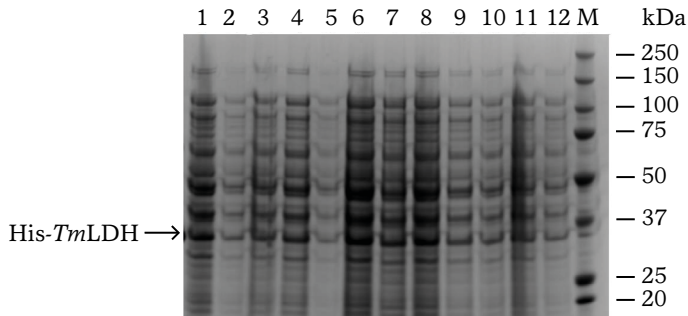


Figure 4.14. SDS-PAGE of *E. coli* M15ΔglyA final samples of the screening experiment for His-TmLDH overexpression. Lane M: molecular weight standard (kDa); lanes 1 to 12: induced cultures. Culture conditions: 2 mL LB, 24°C, 24 h, 140 rpm, 0.4 mM IPTG. His-TmLDH (34 kDa) corresponding band indicated with arrow.

Overexpression screening experiments were useful to determine that minor differences were observed among the tested clones of the same variant. However, the relative amount of target enzymes with respect to total protein content of the cell lysates varied significantly between the different constructs, according to protein electrophoresis.

Regarding His-ScADH relative content (Figure 4.12), between 14 and 16% of relative band intensity was measured in all samples, whilst around 20% of relative intensity was measured for His-ScPDC producing cultures (Figure 4.13). Finally, slightly lower values were shown for His-TmLDH (Figure 4.14), with almost a 10% of intensity. However, significant variances in total protein content can be seen among the His-TmLDH producing colonies, even if it was tried to standardise the biomass concentration of the samples loaded onto the electrophoresis gel. Thus, to determine the most productive clone, the intensity of the bands corresponding to target protein were normalised by the total intensity of each lane.

In conclusion, the change of the *E. coli* host strain was successfully achieved, since all M15 Δ glyA variants transformed with the newly cloned pVEF vectors were able to overexpress the proteins of interest. To conclude the molecular biology work, the colonies that showed higher overexpression yields for each of the three histidine-tagged target enzymes were chosen to establish the working cell cryostocks, as mentioned before.

4.3. Histidine-tagged enzymes production with the M15 Δ glyA strain

As previously stated, the decision to change the MCF to produce target enzymes came with substantial modifications in production procedures respect to previous processes with NEB 10- β strain, which were expected to lead to simpler and more efficient RPP processes.

The use of the auxotrophic M15-derived strain transformed with the generated pVEF vectors brought i) the use of glucose as the sole carbon source, avoiding the medium shift, and ii) the end of antibiotic addition as selection marker, which were considered the two main improvements respect to previous methodology. Besides, M15 Δ glyA strain was already

reported as an efficient recombinant protein producer *E. coli* strain (Vidal et al., 2008, Pasini et al., 2016).

In this section, the results obtained for histidine-tagged enzymes production with the M15 Δ *glyA* strain and antibiotic-free defined medium (DM) are presented. First preliminary experiments consisted in protein overexpression assessment at flask scale, whose corresponding results are presented for the three enzymes of interest (*ScADH*, *ScPDC* and *TmLDH*). The aim of these lower-scaled cell cultures was to check that the new M15 Δ *glyA* transformants were able to grow in DM without using ampicillin as selection marker, meaning that correct constitutive expression of SHMT enzyme was occurring —encoded by *glyA* gene of pVEF—, and to corroborate the correct overexpression of the three enzymes in a soluble and active form, what was only tested previously in well-plate screenings with LB medium.

Considering that considerable amounts of target enzymes were required by the whole project to attempt the obtention of lactic acid through the designed multi-enzymatic system, the three enzymes shall be produced at reactor scale with high-density cell cultures. Therefore, production processes were scaled up to bench-scale reactor and general protocols developed in the group were applied to obtain the enzymes.

On the one hand, batch fermentations were performed at 24°C (alike shake-flask cultures) aiming to slow cell metabolism and to promote the correct solubilisation of target enzymes, since insoluble inclusion bodies could be formed under accelerated cell growth conditions at higher temperatures (37°C), as previously reported for these type of proteins in other *E. coli* expression systems (Miret et al., 2020).

Batch processes were delimited by the total consumption of the initial amount of glucose contained in DM, what occurred approximately around 24 hours after the inoculation of reactor, and results obtained were useful to compare production yields between M15 Δ *glyA* and NEB 10- β strains.

On the other hand, fed-batch fermentations were carried out at 37°C, and exponential growth phase was limited by controlling the addition of concentrated feeding solution with a pre-set specific growth rate (μ) lower than the calculated at 24°C in batch cultures, aiming to slow down cell metabolism, as an alternative to temperature decrease, what could

have been led to excessive process duration of batch phase. Fed batch processes were delimited by the decrease of cell growth, possibly caused by limiting phenomena such as oxygen transfer limitations, side-products accumulation, etcetera.

Fed-batch processes were also scaled-up from 2 to 5 litre vessels, what allowed us to obtain significantly higher enzyme titres than with batch fermentations. Finally, results provided in this section have been compared according to their dependence; specific production values — dependent on the construction of the expression system— are provided as average values obtained from all bench-scale processes, whilst productivity values —dependent on culture volumes and cell densities, i.e., the production process— are compared among the different approaches. Specific enzyme activity of recovered proteins was also evaluated for each of the enzymes.

Nevertheless, prior to produce the enzymes of interest with *E. coli* M15 Δ glyA strain, transformed cells were adapted from complex medium (LB), used in molecular biology experiments, to DM, used to produce the recombinant proteins at reactor scale. This **adaptation process to defined medium** consisted in the same three-step flask-scaled cell cultures carried out previously for NEB 10- β cells, ending up with the generation of the working cell banks required to establish the fully antibiotic-free-based production process with DM.

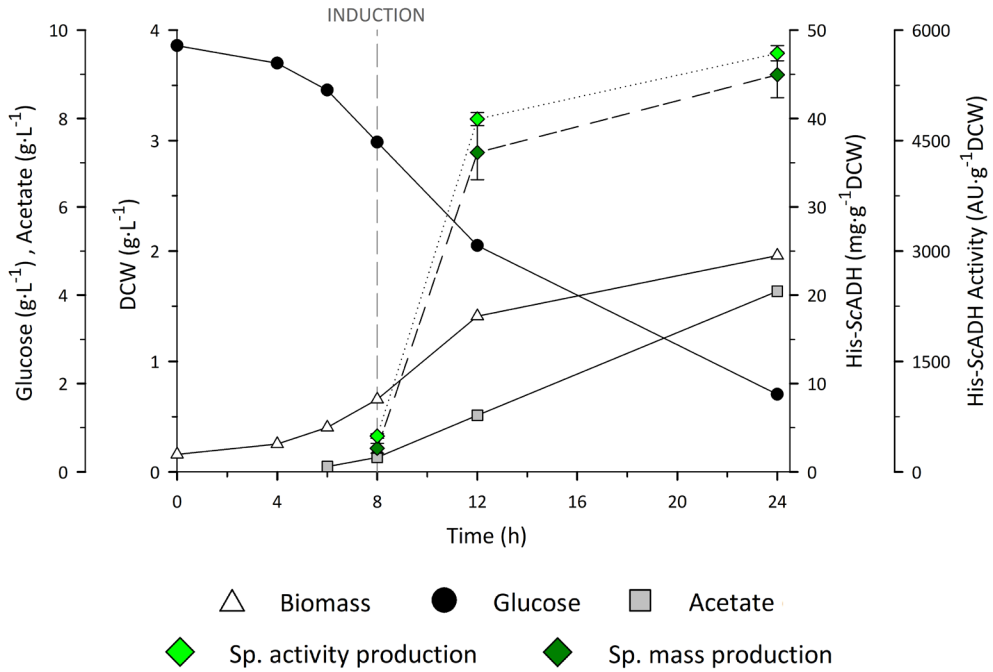
Biomass concentration levels were measured to check if *E. coli* cells grew as expected with defined medium throughout the culture cycles, and the resulting profiles are shown in [Figure 8.4](#) of the Appendix section. Results showed that cell growth in DM was correct, presenting similar profiles of biomass from cycle to cycle. In addition, maximum specific growth rates (μ_{\max}) reached during exponential growth phase were calculated in the last step of the adaptation process for each variant, obtaining μ_{\max} values of around 0.3 h^{-1} for the three variants. Therefore, cultures derived from the third step were used to generate the cell cryostocks.

4.3.1. Shake-flask overexpression experiments

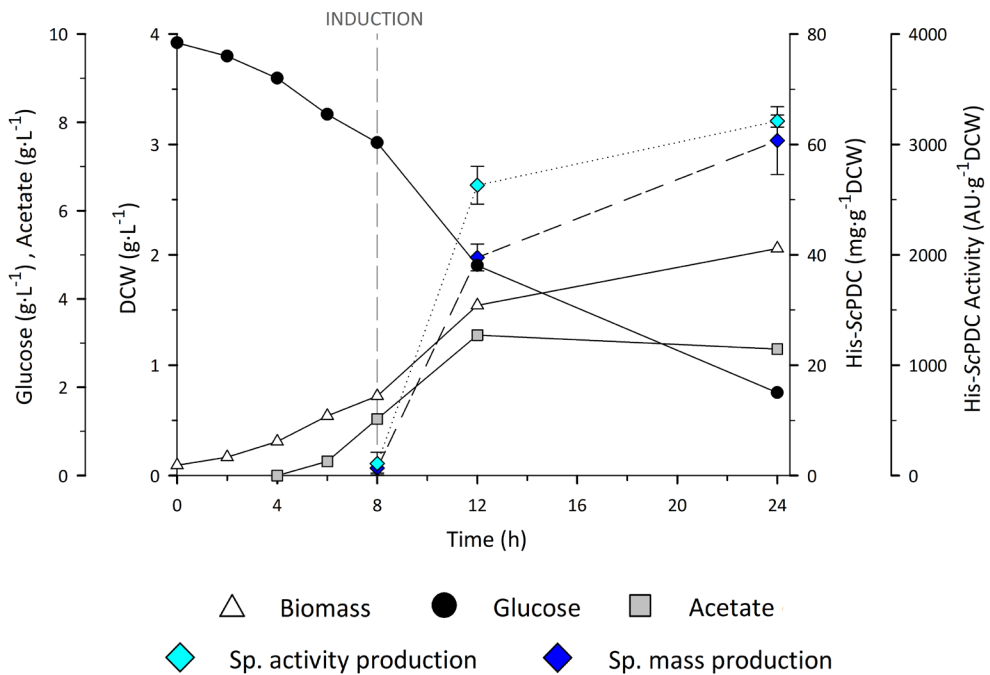
Once the cell stocks were available for RPP with DM, and before to carry out any production process at reactor scale, it was performed an overexpression experiment at flask scale with DM to assess the enzyme production, considering that first overexpression screening at well-plate scale was performed prior to the adaptation from LB to DM.

Same approach applied for NEB 10- β strain was followed; three flasks were inoculated with the corresponding variants, and two of them were induced afterwards, keeping the third as a negative control. Process parameters followed along the cultures and calculated specific productions (before and after induction) are shown below in [Figure 4.15](#):

A: His-ScADH



B: His-ScPDC



C: His-*Tm*LDH

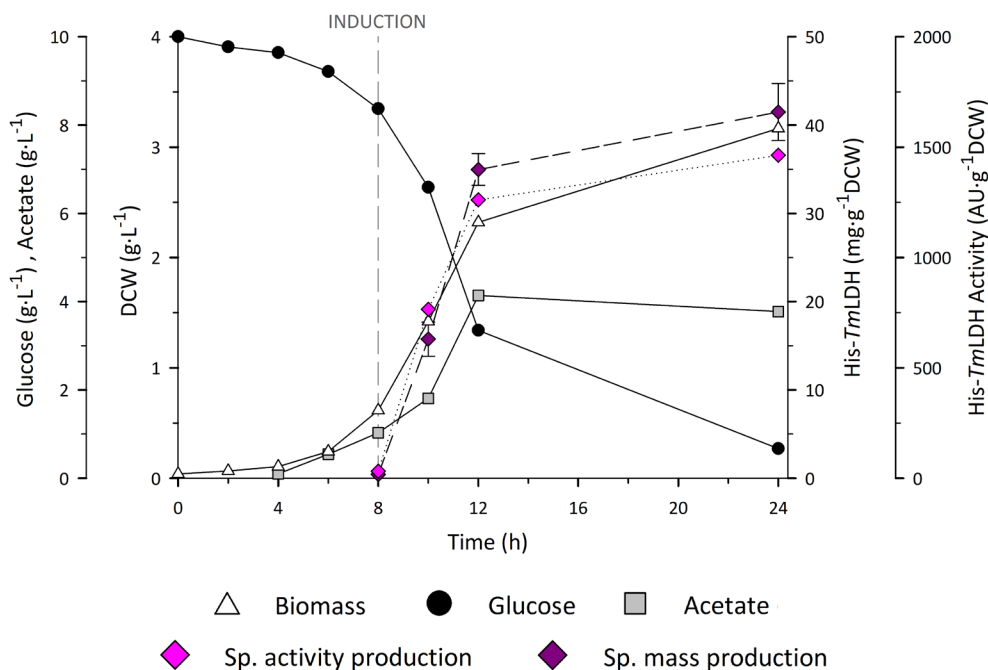


Figure 4.15. *E. coli* M15 Δ *glyA* flask cultures performed to produce histidine-tagged ScADH (A), ScPDC (B) and *Tm*LDH (C) enzymes with DM. Profiles along time of biomass DCW (Δ), glucose (\bullet) and acetate (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Induction phase indicated. Culture conditions: 24°C, pH 7.0, 140 rpm, 0.25 mM IPTG. Error bars correspond to standard error of culture duplicates.

Preliminary overexpression experiments at shake-flask scale showed promising results, obtaining between 4 to 6 milligrams of target enzymes, depending on the case. Moreover, it could be checked the strong repression of expression system before the addition of IPTG, which was also corroborated with the comparison of specific production parameters between induced cultures and negative controls.

However, as can be observed in Figure 4.15, glucose was not totally consumed in none of the cases and acetate accumulation reached almost inhibitory levels. Thus, it was expected to obtain higher yields with reactor-scale fermentations, in which parameters such as pH and dissolved oxygen are controlled, what could avoid excessive deviations from optimum culture conditions.

Below, most relevant production parameters calculated for each enzyme are listed in Table 4.5. As previously stated for protein overexpression with NEB 10- β strain, it was not intended to compare these values among the three enzymes, but to obtain reference production values for each of the variants. Therefore, the intention is to present the results obtained, which will be compared to other production scales later on.

Table 4.5. Production parameters calculated for histidine-tagged ScADH, ScPDC and TmLDH enzymes using *E. coli* M15 Δ glyA strain and DM in 50 mL shake flask cultures. Culture conditions: 24°C, pH 7.0, 140 rpm, 0.25 mM IPTG. \pm correspond to the standard error of culture duplicates.

Production parameter	His-ScADH	His-ScPDC	His-TmLDH
Specific production (mg Enz·g ⁻¹ DCW)	45 \pm 3	61 \pm 6	42 \pm 3
Specific production (AU·g ⁻¹ DCW)	5700 \pm 100	3200 \pm 55	1470 \pm 14
Specific activity (AU·mg ⁻¹ Enz)	24 \pm 2	53 \pm 1	35 \pm 0

Results obtained were useful to conclude that target enzymes overexpression was successfully achieved in all cases, proving that auxotrophic *E. coli* M15 Δ glyA strain was capable to produce recombinant proteins by using antibiotic-free DM, strengthen previous results (Pasini et al., 2016). In addition, specific production yields of induced cultures were up to one order of magnitude higher than negative controls, validating the regulation of the expression vector.

4.3.2. Batch production processes at bench-scale reactor

The next step taken was to scale up the production of the three histidine-tagged enzymes to bench scale by following a batch strategy as described in Section 3.5, Cultivation conditions. ScADH was the first enzyme produced at reactor scale and was used as a model enzyme to analyse the behaviour of M15 Δ glyA cells in batch and fed-batch cultures, aiming to establish the final operating conditions of the production platform, whilst ScPDC and TmLDH enzymes were useful to validate the reproducibility of the process platform.

However, considering the large number of fermentations performed, only ScADH producing cultures will be shown hereinafter, although results obtained in all production processes are provided in the following

tables. Thus, the rest of the figures corresponding to the production of ScPDC and TmLDH enzymes are provided in Figures 8.5 and 8.6 of the Appendix, respectively. Fermentation parameters followed along the batch culture, and enzyme specific activity and specific mass production analysed afterwards are depicted in Figure 4.16:

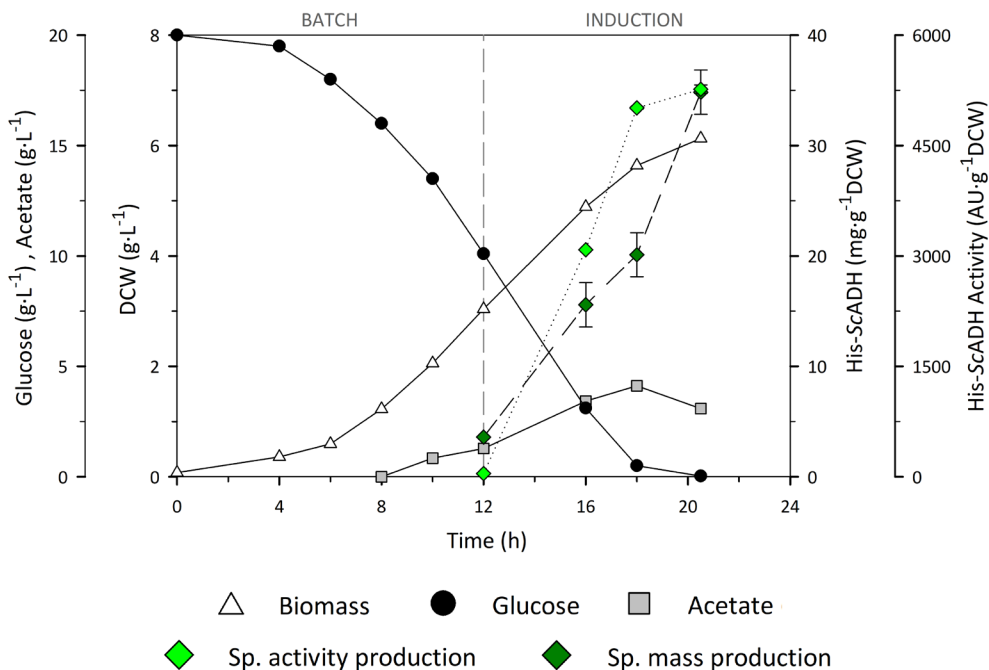


Figure 4.16. *E. coli* M15ΔglyA batch culture performed to produce His-ScADH enzyme with DM. Profiles along time of biomass DCW (Δ), glucose (●) and acetate (□) (g·L⁻¹), specific activity (◆) (AU·g⁻¹DCW) and specific mass production (◆) (mg·g⁻¹DCW). Batch and induction phases indicated. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

More than 1 gram of target enzyme was produced (230 mg·L⁻¹), corresponding to more than 1.5·10⁵ alcohol dehydrogenase activity units (3.5·10⁴ AU·L⁻¹). Regarding the development of the process, exponential growth phase occurred at a μ_{\max} of 0.28 h⁻¹ until induction moment (12 h), being this value slightly higher than μ_{\max} measured in flask cultures (0.24 h⁻¹), probably due to the improvement in oxygen transfer from flasks to reactor. Although acetate accumulation was detected during exponential cell growth, it was partially consumed during the last stage of the fermentation, when glucose was almost totally exhausted.

Finally, regarding the evolution of specific production profiles, it could be considered that maximum values were not reached during the batch process, since profile curves did not seem to be flattened, which was also reported for ScPDC and TmLDH enzymes (see Section 8.1 of the Appendix). In their corresponding batch fermentations, almost 2 grams of ScPDC were obtained ($450 \text{ mg}\cdot\text{L}^{-1}$), whilst 1 gram of TmLDH was produced ($205 \text{ mg}\cdot\text{L}^{-1}$), corresponding to $8.8\cdot 10^4$ and $6\cdot 10^4$ total AU, respectively ($2\cdot 10^4$ and $1.4\cdot 10^4 \text{ AU}\cdot\text{L}^{-1}$).

Below, specific production results obtained for the three batch fermentations are listed in Table 4.6:

Table 4.6. Production parameters calculated for histidine-tagged ScADH, ScPDC and TmLDH enzymes using *E. coli* M15 Δ glyA strain and DM in 5 L batch cultures. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. \pm correspond to the standard error three sample measurements.

Production parameter	His-ScADH	His-ScPDC	His-TmLDH
Specific production ($\text{mg Enz}\cdot\text{g}^{-1}\text{DCW}$)	35 ± 2	78 ± 2	31 ± 3
Specific production ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$)	5266 ± 58	3509 ± 67	2175 ± 52
Specific activity ($\text{AU}\cdot\text{mg}^{-1}\text{Enz}$)	151 ± 2	45 ± 1	69 ± 2

4.3.3. Fed-batch production processes at bench-scale reactor

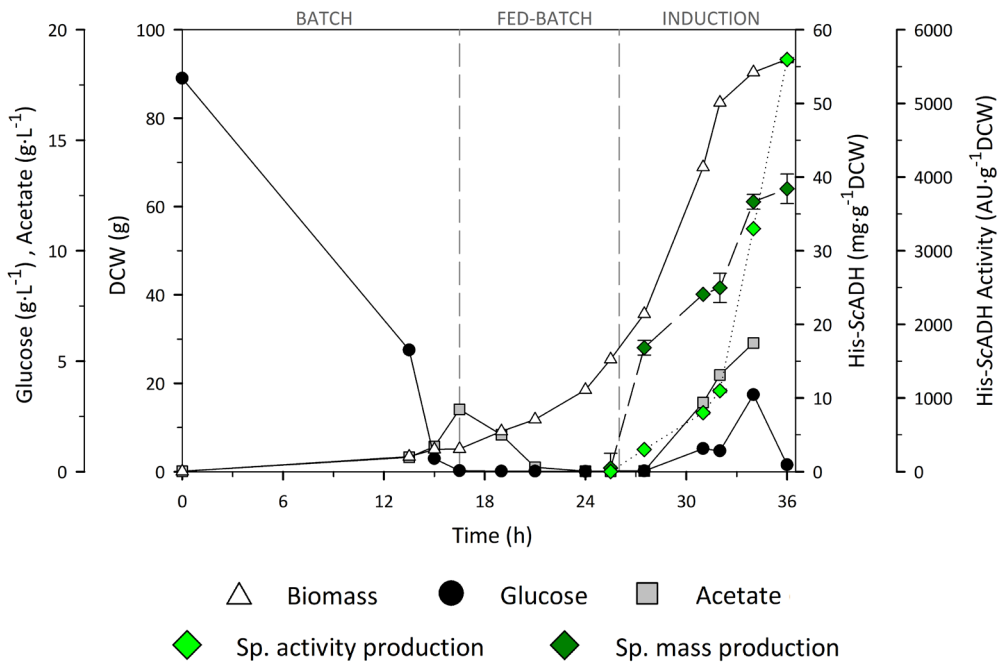
Finally, histidine-tagged enzymes were produced at bench scale by following a fed-batch strategy, whose main objective was to improve process productivities in comparison with batch cultures. As previously mentioned, fed-batch cultures were carried out at 2 and also at 5 L scale. This increase in culture volume was considered interesting to assess if similar productivity values could be obtained, regardless of the final volume. However, considering that the main objective was to obtain the three enzymes to work with, fed-batch processes were not optimised to try to obtain even higher enzyme titres.

His-ScADH enzyme was successfully produced in both fed-batch cultures, resulting in the culmination of the whole production process, which was also validated with the production of His-ScPDC and His-TmLDH enzymes. For that reason, after this first approximation to the production in fed-batch cultures with antibiotic-free DM, it can be concluded that

the three histidine-tagged enzymes were successfully produced with the platform developed.

Production parameters followed along the two His-ScADH producing processes are shown below in Figure 4.17, while the processes conducted to obtain the rest of enzymes (ScPDC and TmLDH) are provided in the Appendix section (Figures 8.7 and 8.8, respectively).

A: 2 L



B: 5 L

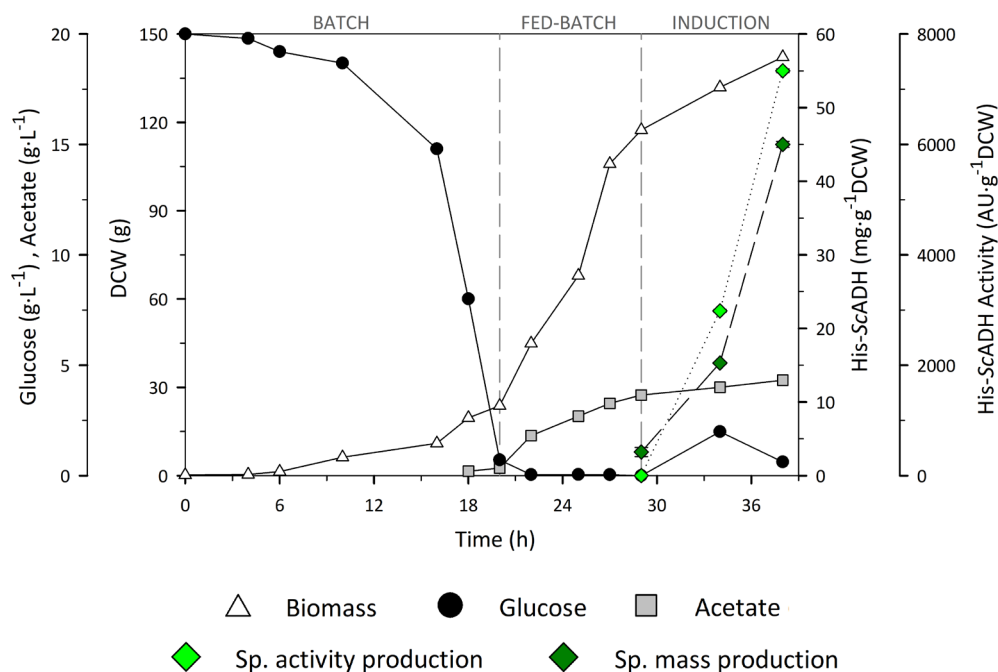


Figure 4.17. *E. coli* M15Δ*glyA* fed-batch cultures performed to produce His-ScADH enzyme with DM with a 2 L vessel (A) and a 5 L vessel (B). Profiles along time of biomass DCW (Δ) (g), glucose (●) and acetate (□) (g·L⁻¹), specific activity (◆) (AU·g⁻¹DCW) and specific mass production (◆) (mg·g⁻¹DCW). Batch, fed-batch and induction phases indicated. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

Similar behaviours were observed between the two processes, showing a calculated $Y_{X/S}$ of 0.29 g·g⁻¹ and of 0.34 g·g⁻¹ during batch stages of the 2 L (Figure 4.17 A) and the 5 L (Figure 4.17 B) processes, respectively. This slight difference between the two processes could be explained by the accumulation of acetate observed at the end of batch phase in the 2 L bench-scale reactor.

On the other hand, regarding the development of fed-batch phases, specific growth rates (μ) of 0.18 and 0.20 h⁻¹ were determined for the 2 L and the 5 L bench-scale reactors, respectively. In both cases, glucose started to accumulate approximately 5 hours after the induction of protein overexpression—and so did acetate—as a consequence of metabolic collapse caused by induction, even if none of the two processes reached

the peak of maximum production yield, since the corresponding profiles still showed an increasing tendency at the end of the fermentations.

Specific production results obtained for the three enzymes in the different fed-batch cultures are listed in [Tables 4.7 \(2 L\)](#) and [4.8 \(5 L\)](#):

Table 4.7. Production parameters calculated for histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes using *E. coli* M15 Δ *glyA* strain and DM in 2 L fed-batch cultures. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. \pm correspond to the standard error three sample measurements.

Production parameter	His- <i>ScADH</i>	His- <i>ScPDC</i>	His- <i>TmLDH</i>
Specific production (mg Enz·g ⁻¹ DCW)	38 ± 1	87 ± 3	36 ± 1
Specific production (AU·g ⁻¹ DCW)	5600 ± 26	1760 ± 6	2362 ± 28
Specific activity (AU·mg ⁻¹ Enz)	146 ± 1	20 ± 0	66 ± 1

Table 4.8. Production parameters calculated for histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes using *E. coli* M15 Δ *glyA* strain and DM in 5 L fed-batch cultures. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. \pm correspond to the standard error three sample measurements.

Production parameter	His- <i>ScADH</i>	His- <i>ScPDC</i>	His- <i>TmLDH</i>
Specific production (mg Enz·g ⁻¹ DCW)	45 ± 0	85 ± 3	41 ± 1
Specific production (AU·g ⁻¹ DCW)	7336 ± 34	2381 ± 13	1930 ± 7
Specific activity (AU·mg ⁻¹ Enz)	163 ± 1	28 ± 0	47 ± 0

As explained above, the change of the production strategy from batch to fed-batch cultures, was not only useful to obtain higher amounts of target enzymes but to compare the productivities reported, even if none of the processes was optimised. In general, same specific yields were calculated for fed-batch fermentations compared with batch processes, since all these parameters are not dependent on the culture strategy applied but on other considerations such the *E. coli* strain and the expression system, and similar specific activities were achieved for each of the enzymes produced among all the cultures, even if significant differences were observed in some cases. Thus, **specific production yields were consistent enough to be considered as an average value** for each of the enzymes of interest ([Table 4.9](#)), regardless of the culture strategy.

Table 4.9. Average production parameters calculated for histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes using *E. coli* M15 Δ *glyA* strain and DM. \pm correspond to the standard error of the average value calculated among the different production processes.

Production parameter	His- <i>ScADH</i>	His- <i>ScPDC</i>	His- <i>TmLDH</i>
Specific production (mg Enz·g ⁻¹ DCW)	39 \pm 3	84 \pm 3	36 \pm 3
Specific production (AU·g ⁻¹ DCW)	6070 \pm 640	2550 \pm 510	2160 \pm 125
Specific activity (AU·mg ⁻¹ Enz)	153 \pm 5	31 \pm 7	61 \pm 7

As can be observed, specific productions varied from one enzyme to another, meaning that probably the optimisation of the process should actually be aborred individually for each target enzyme, even if values within the same range were obtained for all recombinant proteins. Regarding the specific mass productions obtained, around 40 milligrams of *ScADH* and *TmLDH* were produced per biomass gram, whilst more than 80 milligrams of *ScPDC* were obtained. Nonetheless, in terms of moles of enzyme, specific production was near to 1 micromole of enzyme per biomass gram, yielding from 1.1 to 1.3 μ mol·g⁻¹DCW.

Regarding the operational mode followed, volumetric productivity values increased notably when production processes were scaled up from batch to fed-batch cultures, which is basically one of the main reasons why fed-batch processer are developed. In this case, results are provided in separate tables for each enzyme, aiming to lighten the data burden and to emphasise that comparison shall be done between the different processes but not between the three enzymes.

Productivity values calculated for His-*ScADH*, His-*ScPDC* and His-*TmLDH* enzymes are listed in [Tables 4.10](#), [4.11](#) and [4.12](#), respectively, in which higher values are highlighted in bold:

Table 4.10. Productivity parameters calculated for His-*ScADH* enzyme using *E. coli* M15 Δ *glyA* and DM in bench-scale reactor cultures. \pm correspond to the standard error of three sample measurements.

Production parameter	Batch	Fed-batch	
		2 L	5 L
Volumetric productivity (mg Enz·L ⁻¹ ·h ⁻¹)	11 \pm 1	54 \pm 2	48 \pm 0
Volumetric productivity (AU·L ⁻¹ ·h ⁻¹)	1700 \pm 19	7870 \pm 36	7760 \pm 35

Table 4.11. Productivity parameters calculated for His-ScPDC enzyme using *E. coli* M15ΔglyA and DM in bench-scale reactor cultures. ± correspond to the standard error of three sample measurements.

Production parameter	Batch	Fed-batch	
		2 L	5 L
Volumetric productivity (mg Enz·L ⁻¹ ·h ⁻¹)	17 ± 0	104 ± 3	116 ± 3
Volumetric productivity (AU·L ⁻¹ ·h ⁻¹)	781 ± 8	2100 ± 7	3230 ± 18

Table 4.12. Productivity parameters calculated for His-TmLDH enzyme using *E. coli* M15ΔglyA and DM in bench-scale reactor cultures. ± correspond to the standard error of three sample measurements.

Production parameter	Batch	Fed-batch	
		2 L	5 L
Volumetric productivity (mg Enz·L ⁻¹ ·h ⁻¹)	7 ± 1	35 ± 1	60 ± 1
Volumetric productivity (AU·L ⁻¹ ·h ⁻¹)	505 ± 12	2303 ± 30	2822 ± 10

In general, fed-batch cultures led to an increase of one order of magnitude in enzyme productivities, and so was observed for enzyme titre (mg·L⁻¹) and enzyme activity (AU·L⁻¹), which increased in the same proportion. Besides, it can be observed that higher productivities were calculated for fed-batch processes at 5 L scale, rather than at 2 L scale, unless for His-ScADH production, in which no significant differences were obtained. Thus, fed-batch cultures proved to be the most effective strategy tested, since remarkably higher enzyme titres were produced, even if production processes were still far from being optimised, what would be interesting to address in further experimental work.

4.4. Comparison of production parameters between NEB 10-β and M15ΔglyA strains

Finally, it was considered interesting to compare the production parameters obtained between the two *E. coli* strains used in this chapter of results. Once again, His-ScADH enzyme was used as a reference protein, and those results obtained in fed-batch cultures performed with DM were compared, since this was the desired production approach that must be established to fulfil the project requirements.

Even if NEB 10- β specific production yield was slightly higher, higher amounts of His-ScADH enzyme were produced with M15 Δ glyA strain (2.7 and 3.6 grams, respectively), obtaining then higher enzyme titres and higher volumetric productivities. In addition, differences between the two strains were even higher in terms of activity units, obtaining four times more alcohol dehydrogenase AU with the auxotrophic *E. coli* strain, what led to an almost 5-fold increase in volumetric productivity.

Regarding the specific production values obtained for the other two enzymes of interest, similar differences between the NEB 10- β and the M15 Δ glyA strains were shown, as can be seen in [Figure 4.18](#).

As mentioned, results suggest that the auxotrophic strain proved to be a more robust MCF, although production yields obtained also varied from one enzyme to another. Nevertheless, it has also been proved that the initially proposed expression platform based on NEB 10- β strain could be a useful alternative for some specific processes in which coexists molecular biology experiments with protein production, resulting in a simplification of the production pathway of a newly cloned protein, from the plasmid construction to the obtention of grams of product.

According to the results obtained, it can be concluded that the use of *E. coli* M15 Δ glyA strain as MCF together with the IPTG-inducible pVEF expression vector resulted in a better option for RPP, rather than NEB 10- β strain with the arabinose-inducible pBAD vector.

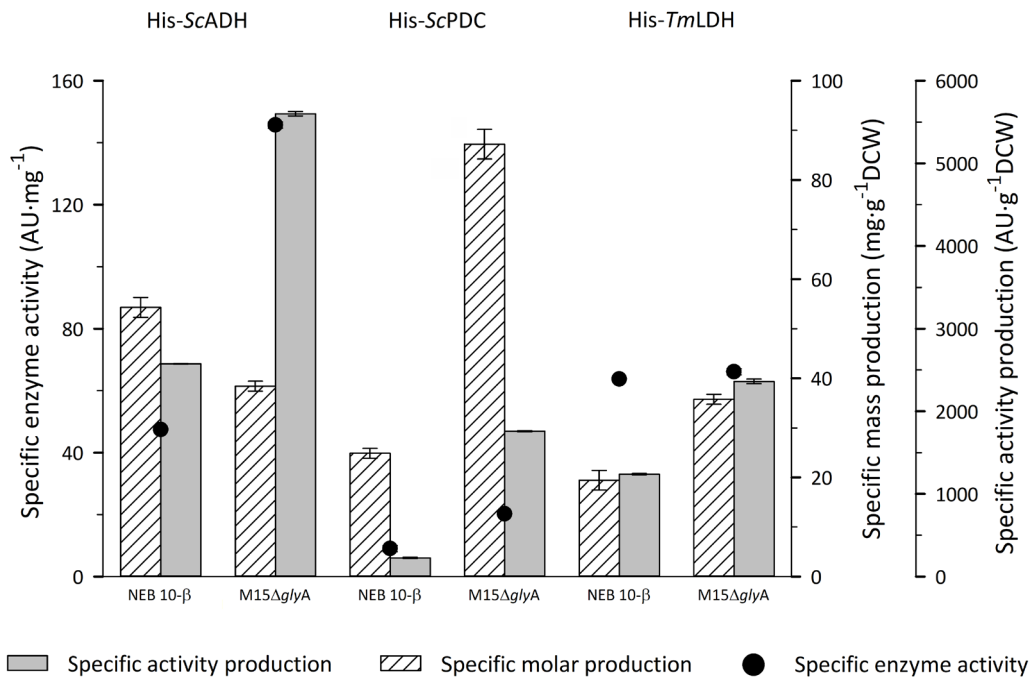


Figure 4.18. Comparison of specific production values obtained in fed-batch cultures performed to produce histidine-tagged ScADH, ScPDC and TmLDH enzymes with DM, between *E. coli* NEB10-β and M15ΔglyA strains. Specific activity (AU·g⁻¹DCW), specific mass (mg·g⁻¹DCW) production values and specific enzyme activity (AU·mg⁻¹Enz) of the three enzymes. Culture conditions for NEB 10-β: 37°C, pH 7.0, 450–1100 rpm, 30% PO₂, 0.2 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹ ampicillin. Culture conditions for M15ΔglyA: 37°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

4.5. Conclusions

Overexpression experiments of the three enzymes requested for L-lactic acid synthesis with the NEB 10- β strain showed promising results, obtaining higher production yields for cultures with defined medium rather than for cultures with complex medium, not only at flask scale but also at reactor scale in all cases. Therefore, a two-step fed-batch strategy that combined the use of glucose and glycerol as the two main carbon sources was successfully followed to obtain high loads of target enzymes.

However, aiming to avoid the medium shift and the amino acid supplementation, the microbial host strain was changed for a typically-used *E. coli* strain for RPP. The auxotrophic *E. coli* M15 Δ *glyA* strain, transformed with a pQE-40-derived vector that contained the *lacI-glyA* cassette, was chosen to overexpress the three histidine-tagged enzymes.

Molecular biology experiments led to the correct cloning of the three corresponding DNA sequences into the pVEF plasmid by following an optimised SLIC method, which turned out to be an effective and cost-saving technique. The use of the newly coned vectors in combination with the auxotrophic strain allowed us to apply selective pressure without the need to use antibiotic, and the use of IPTG instead of L-arabinose led to simpler fermentations, since it was avoided the use of glycerol as carbon source.

The three enzymes of interest were successfully overexpressed at bench-scale reactor with the auxotrophic strain and antibiotic-free DM through batch and also fed-batch cultures. Although similar specific production yields were obtained, volumetric productivity values increased notably when production processes were scaled up from batch to fed-batch cultures, obtaining in all cases high amounts of the histidine-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes, both in mass and activity units. Yet, there is space for improvement through process optimisation.

Finally, although the NEB 10- β strain proved to be viable to produce newly cloned enzymes, it can be concluded that M15 Δ *glyA* strain is a more suitable and more robust MCF to produce the required enzymes at high titres.

5. RESULTS II: CARBOHYDRATE–BINDING MODULE–TAGGED ENZYMES PRODUCTION IN *E. COLI*

Abstract

Two different carbohydrate–binding modules (CBM) were successfully fused to the three enzymes of interest by following the same method used to clone the histidine–tagged enzymes onto the expression vector. The resulting fusion proteins were produced at high titres in a bench–scale reactor using the auxotrophic M15 Δ *glyA* strain, with the antibiotic-free minimum culture medium. The specific production range varied from 50 to 70 mg of target protein per gram of biomass, obtaining around 45 mg·g⁻¹ of CBM-*ScADH* variants, with a specific activity of >65 AU·mg⁻¹, and up to 70 and 77 mg·g⁻¹ of CBM–fused *ScPDC* and *TmLDH* enzymes, respectively, with a specific activity between 20 to 28 AU·mg⁻¹. In this chapter, production parameters of the overexpressed enzymes were also compared with the histidine–tagged variants, aiming to evaluate the effect of the fused CBM tags on the catalytic capacity of the enzymes. Results showed that M15 Δ *glyA* strain was able to produce CBM–fused enzymes as efficiently as histidine–tagged ones, and that CBM domains do not have a significant negative impact in enzyme performance.

5. RESULTS II: CARBOHYDRATE–BINDING MODULE–TAGGED ENZYMES PRODUCTION IN *E. COLI*

5.1. Generation of target CBM–fused proteins

Once the three target enzymes were successfully overexpressed with the auxotrophic *E. coli* M15 Δ *glyA* strain, the next objective was addressed; the obtention of an expression system in M15 Δ *glyA* strain able to produce the three enzymes of interest with carbohydrate–binding module (CBM) tags fused to N-terminal end of each enzyme. This decision was taken considering all the possible applications that CBMs can bring in enzyme immobilisation and purification (see Section 1.5, Carbohydrate–Binding Modules), compared to the traditional histidine tags. I.e., the fusion of a CBM domain may represent a significant step forward in economic and environmental viability of a hypothetical industrial–scale production process.

Some CBM domains have been used as fusion tags, showing not only their highly specificity towards carbohydrates, but other properties such as enhanced protein folding and solubility, as well as increased protein overexpression yields (Oliveira et al., 2015 Shoseyov et al., 2006, Ye et al., 2011). Therefore, a CBM3 domain from *C. thermocellum*, two CBM2 domains —one from *C. cellulovorans* and another from *Actinomyces* (CBM2 and CBM2.2, respectively)—, and a CBM9 domain from *T. maritima* have been tried to fuse to each of the enzymes of interest.

The new constructs were cloned into the pVEF vector designed for protein overexpression by using the same methodology employed to clone the histidine–tagged enzymes, with minor differences, as explained in Section 3.3, Molecular biology techniques. The aim of the experiments described below was to demonstrate that the proposed RPP system, based on the auxotrophic *E. coli* M15 Δ *glyA* strain and the use of defined medium (DM), can be used to successfully express the designed CBM–fused enzymes in an active and soluble form, which are meant to contribute to a cheaper and more environmentally friendly downstream process.

Once the cloning was achieved for all the CBM–fused proteins, the well–plate overexpression screening experiment with the resulting M15 Δ *glyA* transformant colonies was carried out to corroborate the capability of

the cells to produce the target enzymes. In addition, the colony of each construct who showed the greater over-expression levels was picked to generate the cell bank for the subsequent production processes.

5.1.1. Molecular biology: cloning target genes into pVEF vector

Firstly, DNA fragments corresponding to the four chosen CBM domains were amplified by PCR (Figure 5.1), prior to use them in subsequent cloning experiments.

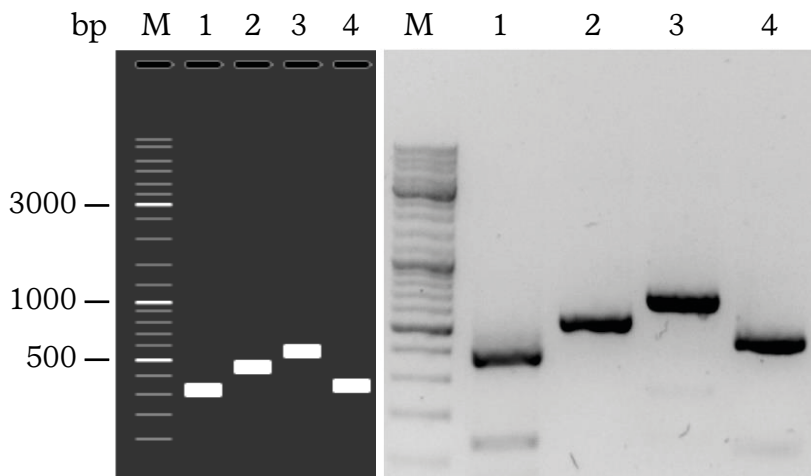


Figure 5.1. Simulation of expected amplified DNA fragments by PCR (left) and 1% agarose DNA electrophoresis gel with actual PCR products (right). Lane M: molecular weight standard (bp); lanes 1 to 4: PCR products corresponding to CBM2 (351 bp), CBM3 (492 bp), CBM9 (603 bp) and CBM2.2 (372 bp), respectively.

ScADH was used for the initial fusion experiments with all the CBM candidates. *E. coli* DH5 α competent cells were transformed with the resulting SLIC products, and in all cases several colonies were grown in agar plates, e. g., 101 colonies with the CBM3-ScADH construct cloned were grown in LB-agar plates, and 9 randomly picked colonies were checked by colony PCR (Figure 5.2), obtaining 7 positive colonies, which were correctly transformed.

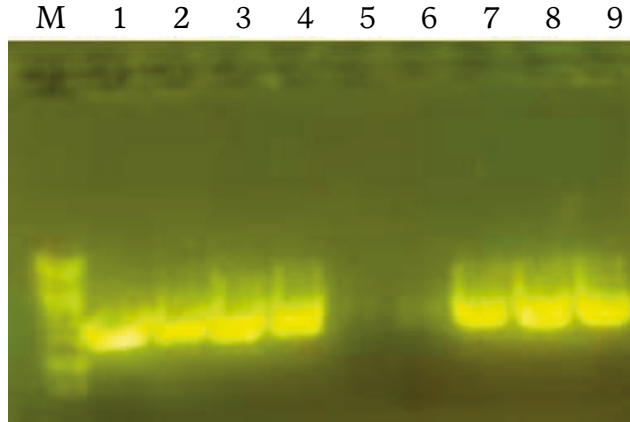


Figure 5.2. Agarose 1% DNA electrophoresis gel with colony PCR products corresponding to CBM3-ScADH amplified fragment (1570 bp). Lane M: molecular weight standard (bp); Positive colonies detected in lanes 1 to 4 and 7 to 9.

Afterwards, three positive colonies (Figure 5.2, lanes 7 to 9) were checked again via XbaI restriction digest (Figure 5.3). The transformed plasmid of the two positive cases was extracted and sequenced, showing no mutations.

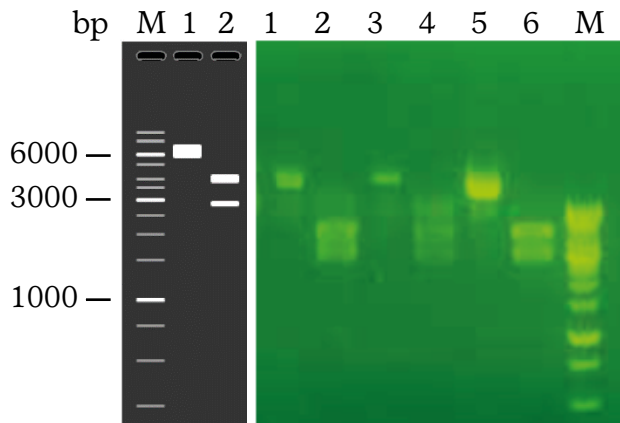


Figure 5.3. Simulation of pVEF_CBM3-ScADH pattern restriction cut by XbaI (left). Lane 1: negative control (uncut plasmid); lane 2: expected restriction pattern. Agarose 1% DNA electrophoresis gel with actual re-restriction patterns (right). Lanes 1,3 and 5: negative controls; lanes 2 and 6: correct restriction patterns; lane 4: incorrect restriction pattern; lane M: Molecular weight standard (bp).

Regarding the rest of the generated constructs (with CBM2, 2.2 and 9), several positive colonies were obtained in all cases, e.g, 56 DH5 α colo-

nies were grown with the CBM2.2-ScADH construct, and 12 of them were checked by PCR (Figure 5.4), obtaining 10 positive cases. Thus, molecular biology experiments led to the correct cloning of the four DNA constructs into the expression vector, what reinforced the robustness and the effectiveness of the cloning method, since it proved to be useful not only with relative short DNA fragments but with almost 2000 bp-long fragments.

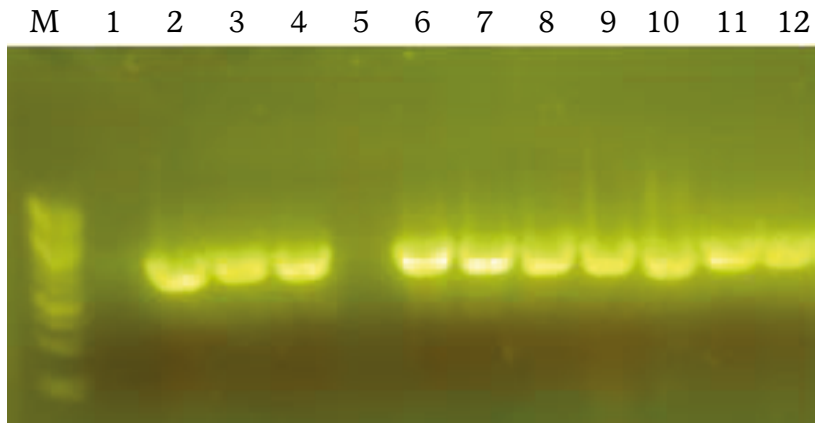


Figure 5.4. Agarose 1% DNA electrophoresis gel with colony PCR products corresponding to CBM2.2-ScADH amplified fragment (1452 bp). Lane M: molecular weight standard (bp); Positive colonies detected in lanes 2 to 4 and 6 to 12.

A fraction of the transformed *E. coli* DH5 α cultures used for plasmid extraction were induced with IPTG aiming to assess the viability of new transformants, in terms of protein expression and enzyme activity levels. Alcohol dehydrogenase enzyme activity and the corresponding overexpressed fusion proteins were measured at satisfactory levels for the CBM3- and CBM9-fused ScADH constructs, whereas CBM2- and CBM2.2-ScADH could not be correctly overexpressed, even if pVEF vector contained the correct DNA sequences cloned onto it. Thus, the results obtained in the expression test performed with ScADH variants led to discard two CBM domains for the following enzymes to be cloned.

Therefore, ScPDC and *Tm*LDH enzymes were fused only to CBM3 and CBM9 domains, following the same procedure that was used previously. On the one hand, regarding **ScPDC enzyme**, more than 30 colonies of DH5 α cells transformed with CBM-ScPDC constructs were grown in plates. In both cases, plasmid DNA of one of the positive colonies was

extracted, obtaining correct restriction patterns with XbaI digestion, and finding no mutations in any of the DNA sequences analysed. In addition, PDC activity was detected for both variants after adding a pulse of IPTG into the flask cultures.

On the other hand, **CBM-*Tm*LDH** cloning also resulted in several transformed colonies, obtaining several positive colonies as well, for each of the variants. Finally, the cultures derived from positives colonies were induced with IPTG showing LDH activity in all cases.

The schematic structure of the new plasmid constructs obtained via SLIC that incorporated CBM–fused *Sc*ADH, *Sc*PDC and *Tm*LDH enzymes is shown below (Figure 5.5), and more de-tails about them are provided in Section 3.2, Plasmids.

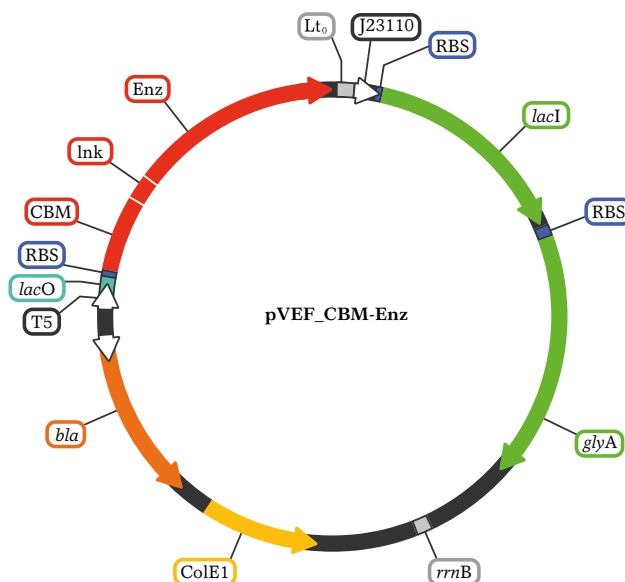


Figure 5.5. Schematic representation of pVEF expression vector containing a CBM–tagged enzyme coding sequence. Features: T5 promoter; *lacO*, two *lac* operator regions; RBS, ribosome binding site; *Lto*, lambda to transcriptional termination; J23110 constitutive promoter; *lacI* gene; *glyA* gene; *rrnB* transcriptional termination; *bla*, ampicillin resistance gene; ColE1 replication origin; CBM, carbohydrate–binding module sequence; Lnk, linker sequence; Enz, enzyme sequence.

5.1.2. *E. coli* M15ΔglyA transformation and expression screening

Once *E. coli* DH5α cells were able to produce the recombinant fusion proteins in induced cultures, the production process at reactor scale was meant to be performed with M15ΔglyA cells, which is the RPP platform chosen to produce the target enzymes at larger scale. Thus, competent cells of the later strain were transformed with the new expression vectors, aiming to overexpress the corresponding proteins of interest.

E. coli M15ΔglyA overexpression screening was performed to determine that all clones were able to produce the target proteins, and to pick the most productive one among several colonies randomly chosen for each of the CBM-fused enzymes, as was performed previously with the histidine-fused versions. The corresponding SDS-PAGE gels are provided below:

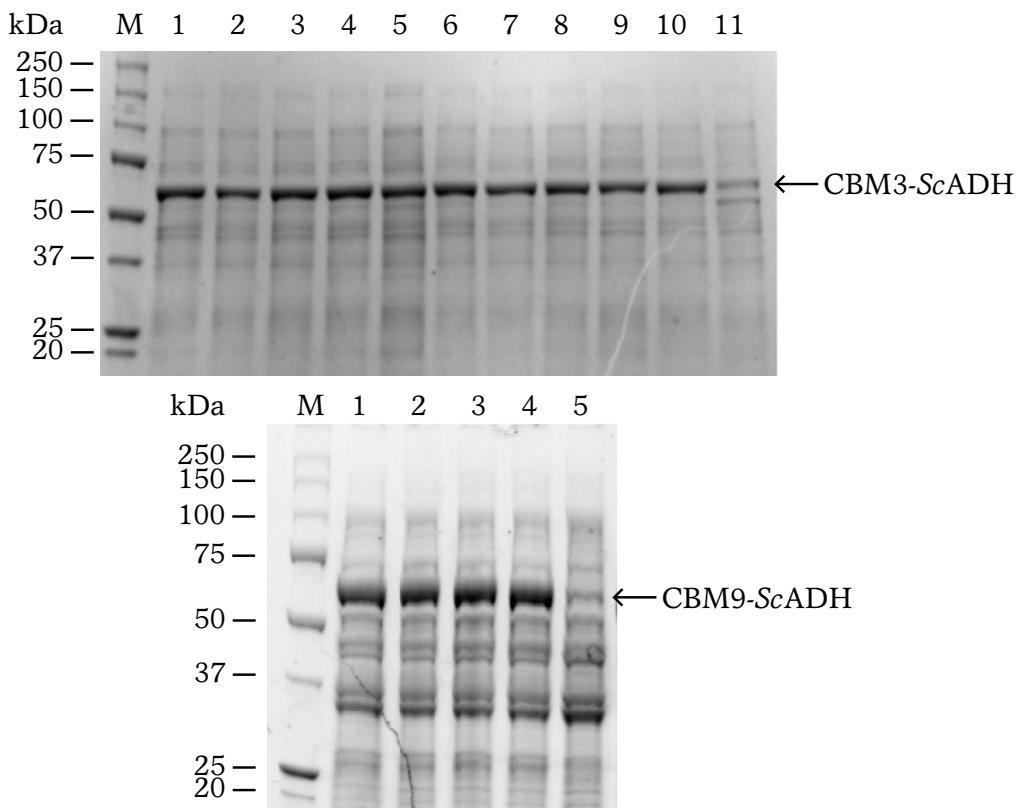


Figure 5.6. SDS-PAGE of *E. coli* M15ΔglyA final samples of the screening experiment for CBM3-ScADH overexpression (**top**). Lane M: molecular weight standard (kDa); lanes 1 to 10: induced cultures; lane 11: negative control. And for CBM9-ScADH (**bottom**). Lane M: molecular weight standard (kDa); lanes 1 to 4: induced cultures; lane 5: negative control. Culture conditions: 2 mL LB, 24°C, 24 h, 140 rpm, 0.4 mM IPTG. CBM3-ScADH (55 kDa) and CBM9-ScADH (64 kDa) corresponding bands indicated with arrows.

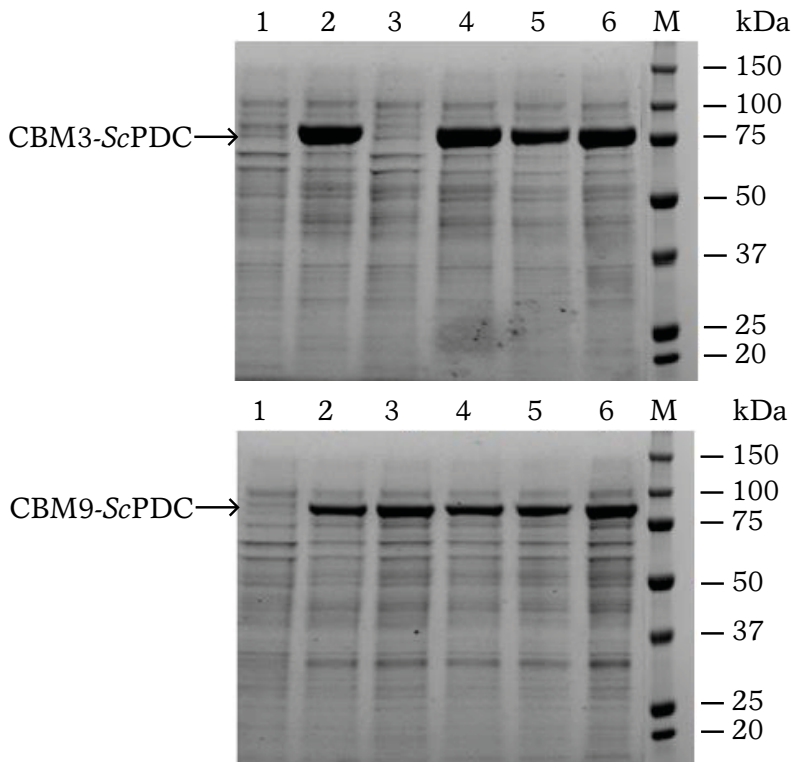


Figure 5.7. SDS-PAGE of *E. coli* M15 Δ glyA final samples of the screening experiment for CBM3-ScPDC overexpression (**top**). Lane 1: negative control; lanes 2 to 6: induced cultures; lane M: molecular weight standard (kDa). And for CBM9-ScPDC (**bottom**). Lane 1: negative control; lanes 2 to 6: induced cultures; lane M: molecular weight standard (kDa). Culture conditions: 2 mL LB, 24°C, 24 h, 140 rpm, 0.4 mM IPTG. CBM3-ScPDC (80 kDa) and CBM9-ScPDC (88 kDa) corresponding bands indicated with arrows.

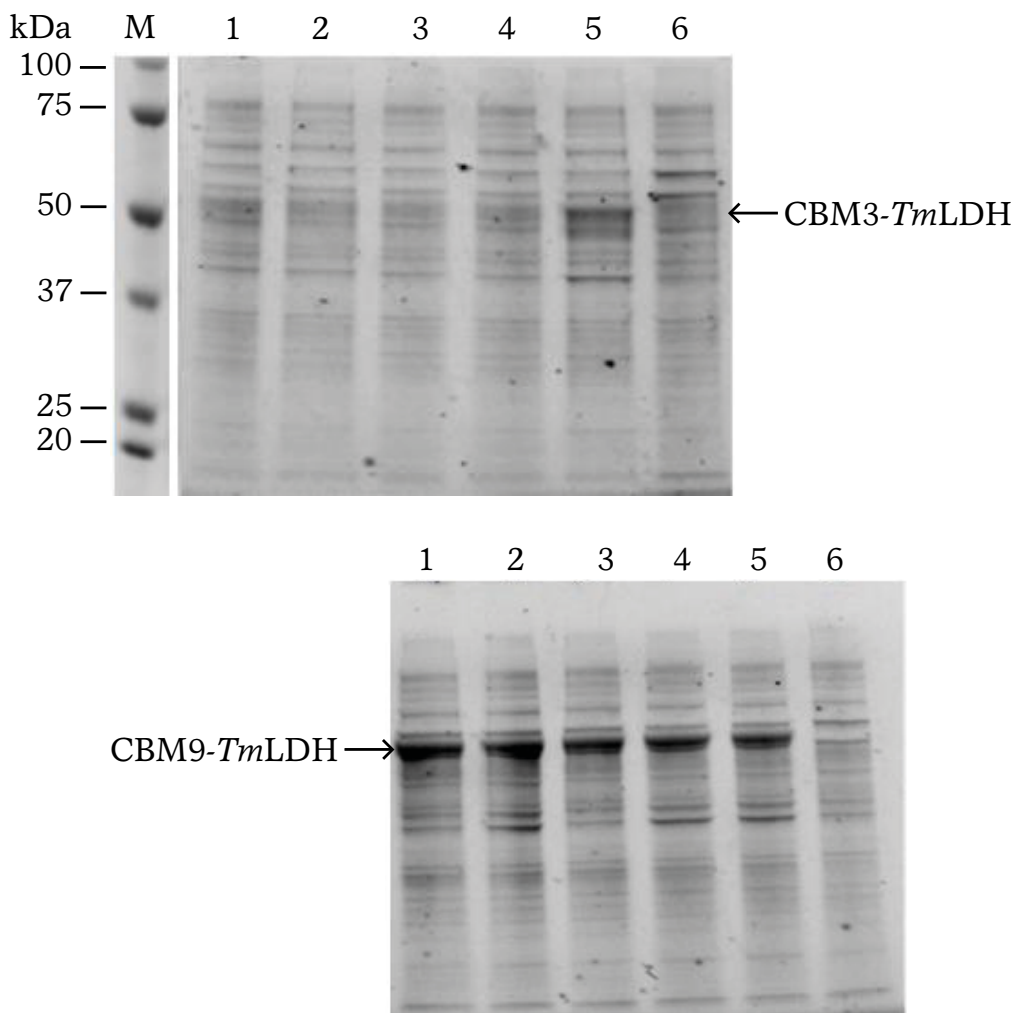


Figure 5.8. SDS-PAGE of *E. coli* M15AglyA final samples of the screening experiment for CBM3-*TmLDH* overexpression (**top**). Lanes 1 to 5: induced cultures; lane 6: negative control; lane M: molecular weight standard (kDa). And for CBM9-*TmLDH* (**bottom**). Lanes 1 to 5: induced cultures; lane 6: negative control. Culture conditions: 2 mL LB, 24°C, 24 h, 140 rpm, 0.4 mM IPTG. CBM3-*TmLDH* (53 kDa) and CBM9-*TmLDH* (57 kDa) corresponding bands indicated with arrows.

The screening experiment was useful to determine that minor differences were observed between clones of the same variant, but the percentage of recombinant protein with respect to total protein content (determined by SDS-PAGE) varied significantly between the different constructs; for **CBM-fused ScADH**, a 34% of relative band intensity was observed for CBM3-*ScADH* (Figure 5.6 top) and a 40% was achieved for CBM9-*ScADH* (Figure 5.6 bottom). Similar values were obtained for **ScPDC enzyme**,

reaching a 35–37% of relative intensity for CBM3–fused version (Figure 5.7 top) and a 50 to 60% for CBM9-ScPDC (Figure 5.7 bottom). However, lower values were shown for *Tm*LDH protein, with a 15% of intensity for the best CBM3-*Tm*LDH producer candidate —the rest of samples represented a 6%— (Figure 5.8 top) and with a 22% of relative content for CBM9–tagged enzyme in the best case (Figure 5.8 bottom).

5.2. CBM–tagged enzymes production with the M15ΔglyA strain

As previously stated, one of the goals of this thesis is related to the obtention of the three enzymes involved in acid lactic synthesis —ADH, PDC and LDH—, with different CBM tags fused to them, aiming to evaluate and characterize their purification/immobilisation in a single-step process, to assess if these domains can be used as fusion tags for that purpose.

Therefore, to carry out the subsequent experiments, it was required to produce all fusion proteins (six variants in total) through fermentation processes with *E. coli*. However, considering the large number of different productions and that this work is part of a European project, which also includes the supply of enzymes for the subsequent development of the proposed MCF, it was not able to study each production case in depth and optimize them. Instead, general protocols of the group were applied to obtain the enzymes.

It was also considered interesting, given the large volume of data obtained, to assess whether the fusion proteins may affect to the behaviour of *E. coli* M15ΔglyA cultures. Thus, not only the fermentations are presented, but an attempt has been made to draw the most general conclusions possible and try to find any tendence or pattern, although it would be necessary to perform more cultures to obtain more statistically reliable data, which would be possible to do in coming studies.

Prior to produce the six fusion proteins of interest with *E. coli* M15ΔglyA cells at reactor scale, *E. coli* cells were adapted from complex (LB) to defined medium (DM), ending up with the generation of all the working cell banks required to establish the fully antibiotic-free–based production process. Cell growth measured during the adaptation process to defined medium is shown in Figure 8.9 of the Appendix section.

All cultures showed similar growth rates along the whole adaptation process. The specific growth rate values obtained for CBM–fused variants were also similar to the ones measured during the adaptation of histidine–tagged ones. Thus, adaptation to DM was considered satisfactory.

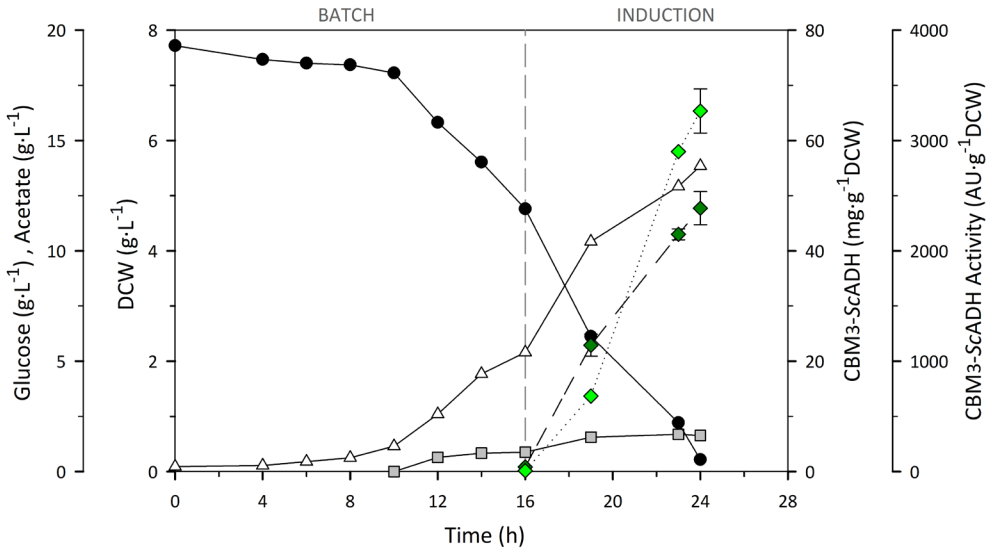
Afterwards, CBM–fused enzymes were produced in the bench–scale reactor by following a batch strategy as described in Section 3.5, Cultivation conditions. Once again, **batch cultures** were set up to produce the six CBM–tagged enzymes at 24°C. The main reason was that during batch phase *E. coli* cells grew at maximum μ , given that glucose was not at a limiting concentration, but the opposite. Hence, the decrease of culture temperature was used to slow cell metabolism and try to avoid the formation of inclusion bodies, which has been already reported for these proteins in other *E. coli* expression systems (Miret et al., 2020).

5.2.1. Alcohol dehydrogenase from *Saccharomyces cerevisiae*

In this context, **CBM-ScADH variants** were the first pair of fusion proteins that were produced, given that ScADH was established as model enzyme along this work. Process parameters were followed along the two processes, whose end was determined by the total consumption of glucose. Enzyme specific activity and specific mass production were analysed afterwards.

As can be seen in [Figure 5.10](#), both CBM-ScADH variants were successfully produced. Similar total amounts of target protein were produced for both cases (1.26 grams of CBM3-ScADH and 1.32 grams of CBM9-ScADH), even if slightly higher total enzyme activity units were obtained for CBM3–tagged enzyme ($8.6 \cdot 10^4$ and $7.8 \cdot 10^4$ AU, respectively). These results were of key importance since they validate the viability of the proposed RPP system, proving to be able to produce such complex proteins in a soluble and active form at high titres, by using the selected antibiotic-free defined medium.

A: CBM3-ScADH



B: CBM9-ScADH

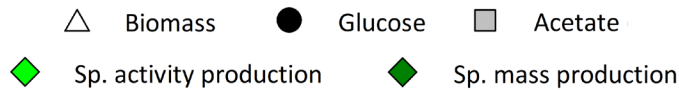
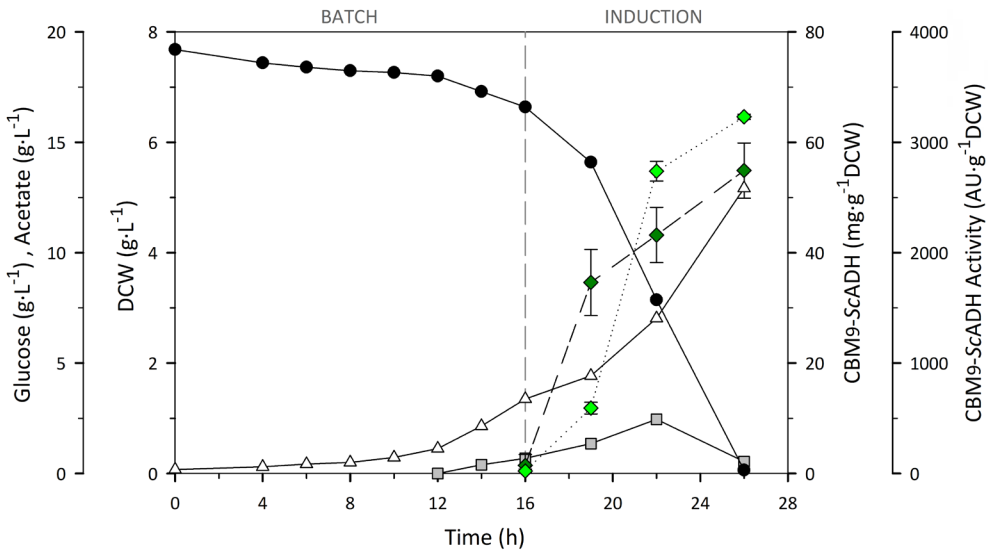


Figure 5.10. *E. coli* M15ΔglyA batch cultures performed to produce CBM3-ScADH (A) and CBM9-ScADH (B) fusion proteins with DM. Profiles along time of biomass DCW (Δ), glucose (●) and acetate (□) ($\text{g}\cdot\text{L}^{-1}$), specific activity (◆) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (◆) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Batch and induction phases indicated. Culture conditions: 24°C , pH 7.0, 450–1100 rpm, 60% PO_2 , 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

Regarding the development of the processes, maximum specific growth rates (μ_{\max}) of 0.34 and 0.25 h⁻¹ were reached during exponential growth phase, for CBM3-ScADH (Figure 5.10 A) and for CBM9-fused variant (Figure 5.10 B), respectively.

Below, main production parameters calculated for each variant are shown in Table 5.1. Nevertheless, some of these values, which are given in milligrams of enzyme, cannot be compared between the two variants, even if they are relevant results form a process-development point of view, since the resulting recombinant proteins have different molecular weight —CBM3-ScADH weights 55 kDa whereas CBM9-ScADH weights 64 kDa—. Otherwise, values expressed in activity units (AU) or micro-moles (μmol) can be compared, since the same activity test is performed, and mass values were normalised to the number of mols of enzyme.

Table 5.1. Production parameters calculated for CBM-fused ScADH enzymes using *E. coli* M15 Δ glyA and DM in batch cultures at 5 L-reactor scale. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. \pm correspond to the standard error of three sample measurements. Comparable parameters are presented in bold.

Production parameter	CBM3-ScADH	CBM9-ScADH
Total biomass (g DCW)	26.3 \pm 0	24.1 \pm 0
Final volume (L)	4.76	4.67
Enzyme activity (AU·L ⁻¹)	1.81·10 ⁴ \pm 1500	1.67·10 ⁴ \pm 120
Enzyme titre (mg Enz·L ⁻¹)	264 \pm 19	284 \pm 25
Specific production (AU·g⁻¹DCW)	3270 \pm 200	3232 \pm 23
Specific production (mg Enz·g ⁻¹ DCW)	48 \pm 3	55 \pm 5
Specific production ($\mu\text{mol Enz}\cdot\text{g}^{-1}\text{DCW}$) *	0.87 \pm 0	0.86 \pm 0
Specific activity (AU·mg ⁻¹ Enz)	69 \pm 6	59 \pm 0
Specific activity (AU·μmol^{-1}Enz) *	3749 \pm 36	3742 \pm 27

* Considering a molar weight of 55 kDa for CBM3-ScADH and of 64 kDa for CBM9-ScADH.

As shown in [Table 5.1](#), similar values were obtained for the two variants in all parameters calculated. In addition, slight differences are not significant if the \pm error is considered. Regarding the specific production values, almost 0.9 micromoles of both variants were produced per gram of biomass, and same values of alcohol dehydrogenase activity were produced, yielding more than 3200 AU·g⁻¹DCW.

Finally, regarding the specific activity of the proteins produced, it can be stated that none of the two CBM domains fused to ScADH enzyme seemed to affect to its correct catalytic activity, since one mol of each fusion protein corresponded to the same ADH activity units. Thus, both candidates were considered suitable to be produced at bench-scale reactor with the auxotrophic *E. coli* strain by using the antibiotic-free DM.

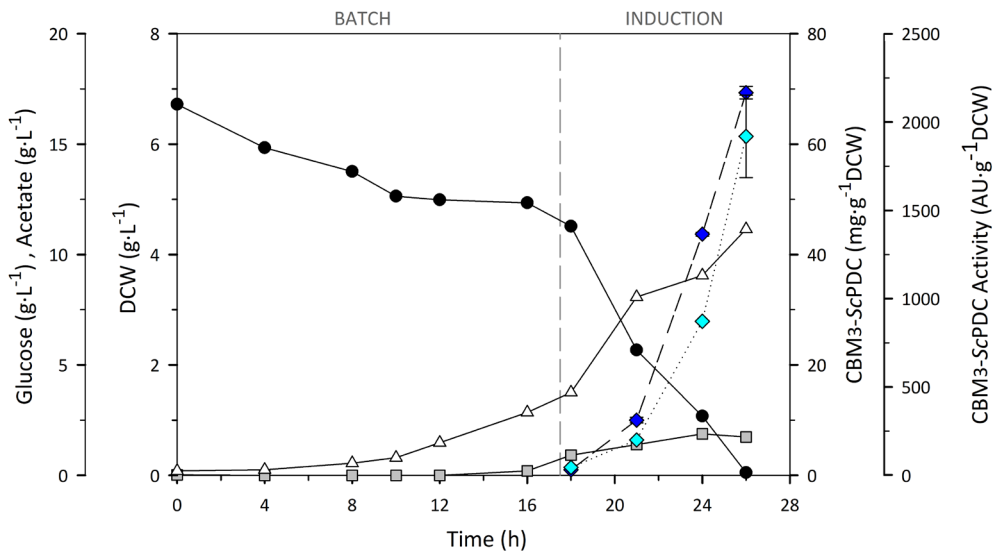
5.2.2. Pyruvate decarboxylase from *Saccharomyces cerevisiae*

CBM-fused ScPDC variants were the following pair of fusion proteins tested for production at reactor scale. Biomass, glucose and acetate profiles along the fermentation time are shown in [Figure 5.11](#), and so are enzyme specific activity and specific mass production. The rest of the production parameters that were calculated for both variants are listed in [Table 5.2](#).

Once again, both candidates were successfully produced at bench-scale reactor (5 L) with the auxotrophic *E. coli* strain by using the antibiotic-free DM; 1.46 g of CBM3-ScPDC and 1.1 g of CBM9-ScPDC were obtained, corresponding to more than 4·10⁴ and 2.2·10⁴ total AU, respectively.

Regarding the two fermentations, similar final biomass levels were reached for both variants at the end of the processes ([Table 5.2](#)). In addition, a μ_{\max} of 0.22 h⁻¹ was calculated in both cases during the exponential growth phase ([Figure 5.11](#)).

A: CBM3-ScPDC



B: CBM9-ScPDC

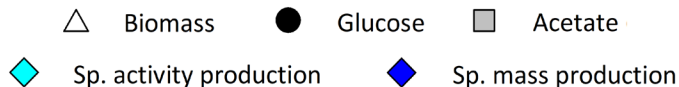
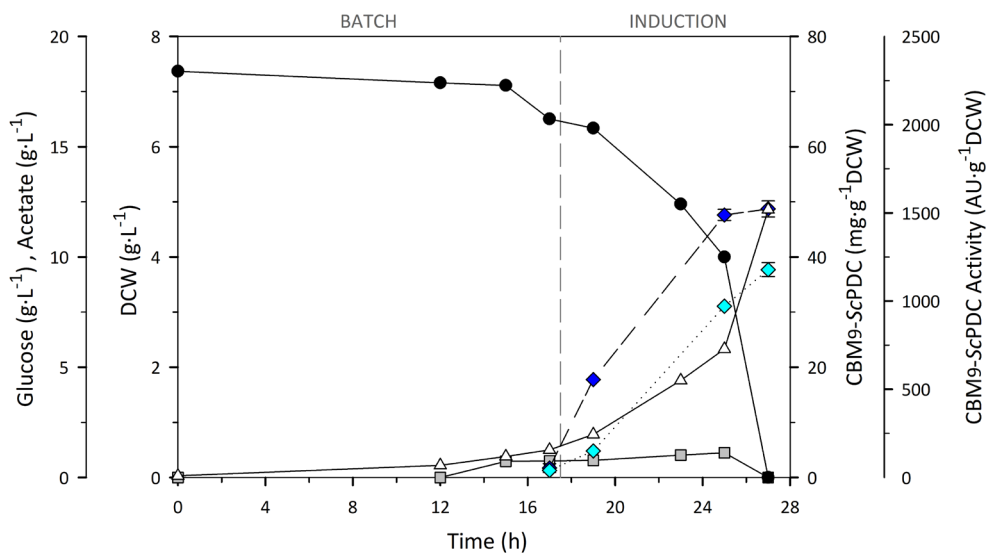


Figure 5.11. *E. coli* M15ΔglyA batch cultures performed to produce CBM3-ScPDC (A) and CBM9-ScPDC (B) fusion proteins with DM. Profiles along time of biomass DCW (Δ), glucose (●) and acetate (□) (g·L⁻¹), specific activity (◆) (AU·g⁻¹DCW) and specific mass production (◆) (mg·g⁻¹DCW). Batch and induction phases indicated. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

Table 5.2. Production parameters calculated for CBM–fused ScPDC enzymes using *E. coli* M15ΔglyA and DM in batch cultures at 5 L–reactor scale. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. ± correspond to the standard error of three sample measurements. Comparable parameters are presented in bold.

Production parameter	CBM3-ScPDC	CBM9-ScPDC
Total biomass (g DCW)	21 ± 0	22.7 ± 1
Final volume (L)	4.71	4.68
Enzyme activity (AU·L ⁻¹)	8.6·10 ³ ± 734	4.72·10 ³ ± 166
Enzyme titre (mg Enz·L ⁻¹)	309 ± 17	175 ± 6
Specific production (AU·g⁻¹DCW)	1920 ± 165	971 ± 34
Specific production (mg Enz·g ⁻¹ DCW)	69 ± 1	49 ± 1
Specific production (μmol Enz·g⁻¹DCW) *	0.87 ± 0	0.55 ± 0
Specific activity (AU·mg ⁻¹ Enz)	27.7 ± 2	20 ± 1
Specific activity (AU·μmol⁻¹Enz) *	2214 ± 190	1750 ± 62

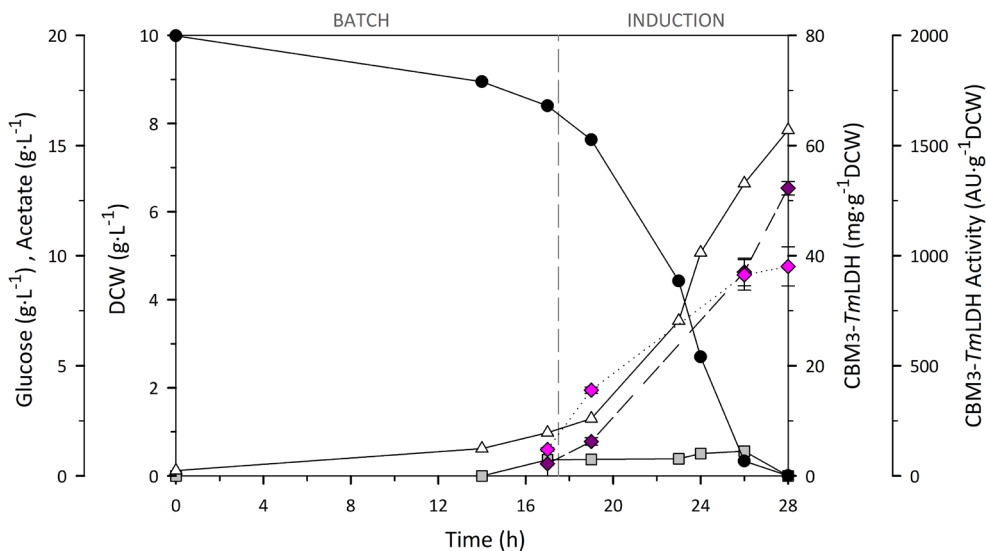
* Considering a molar weight of 80 kDa for CBM3-ScPDC and of 88 kDa for CBM9-ScPDC.

As shown in [Table 5.2](#), higher values were obtained for the CBM3–fused variant in all parameters calculated, even if CBM9–fused variant also showed satisfactory recombinant protein production (RPP) yields. Most remarkable differences were observed in terms of specific production, since activity production (AU·g⁻¹DCW) of CBM3–fused variant was almost 2-fold higher than CBM9, and molar production (μmol Enz·g⁻¹DCW) was 1.6-fold higher. Though, considering the difference in specific mass between the two variants, minor differences were observed. CBM3-ScPDC enzyme showed a 26% increase in specific activity, respect to the CBM9 variant.

5.2.3. Lactate dehydrogenase from *Thermotoga maritima*

Parameters measured along the fermentation of CBM-*Tm*LDH variants are shown below in [Figure 5.12](#):

A: CBM3-*Tm*LDH



B: CBM9-*Tm*LDH

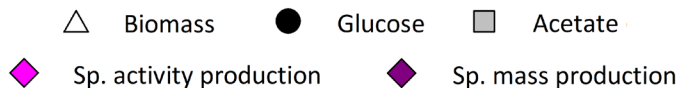
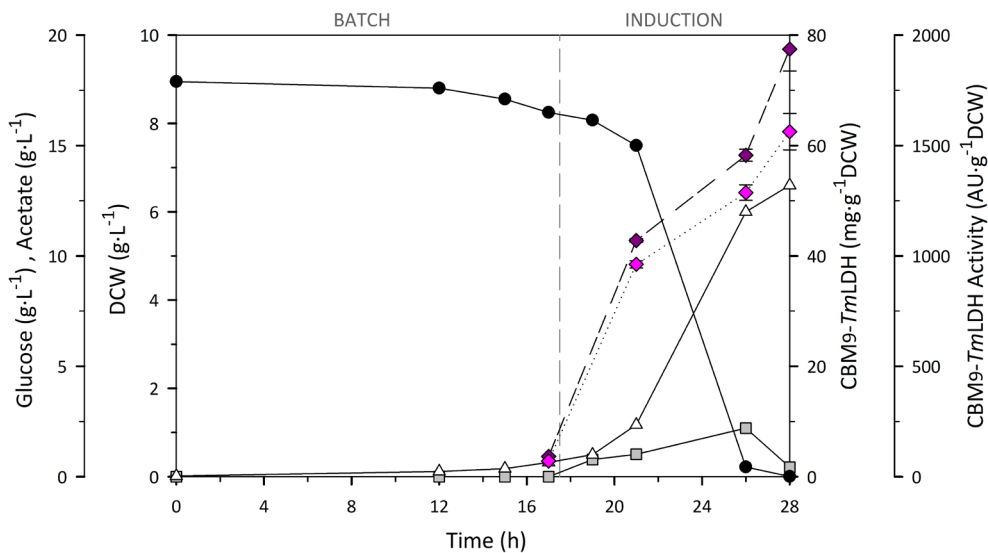


Figure 5.12. *E. coli* M15ΔglyA batch cultures performed to produce CBM3-*Tm*LDH (A) and CBM9-*Tm*LDH (B) fusion proteins with DM. Profiles along time of biomass DCW (△), glucose (●) and acetate (□) (g·L⁻¹), specific activity (◆) (AU·g⁻¹DCW) and specific mass production (◆) (mg·g⁻¹DCW). Batch and induction phases indicated. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

Once again, high loads of target proteins were produced. The two fermentations showed similar profiles along time and the total consumption of glucose was also achieved at same process time, and acetate produced was totally depleted from culture medium in the two cases.

Table 5.3. Production parameters calculated for CBM-fused *Tm*LDH enzymes using *E. coli* M15ΔglyA and DM in batch cultures at 5 L-reactor scale. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. ± correspond to the standard error of three sample measurements. Comparable parameters are presented in bold.

Production parameter	CBM3- <i>Tm</i> LDH	CBM9- <i>Tm</i> LDH
Total biomass (g DCW)	36.7 ± 1	30 ± 0
Final volume (L)	4.68	4.60
Enzyme activity (AU·L ⁻¹)	7.5·10 ³ ± 700	1.03·10 ⁴ ± 544
Enzyme titre (mg Enz·L ⁻¹)	410 ± 9	511 ± 26
Specific production (AU·g⁻¹DCW)	951 ± 89	1563 ± 82
Specific production (mg Enz·g ⁻¹ DCW)	52 ± 1	77 ± 4
Specific production (μmol Enz·g⁻¹DCW) *	0.99 ± 0	1.36 ± 0.1
Specific activity (AU·mg ⁻¹ Enz)	18 ± 2	20 ± 1
Specific activity (AU·μmol⁻¹Enz) *	963 ± 90	1154 ± 61

* Considering a molar weight of 53 kDa for CBM3-*Tm*LDH and of 57 kDa for CBM9-*Tm*LDH.

The results obtained showed that CBM9-fused enzyme was produced more efficiently than CBM3 variant. A 37% increase in specific molar production (μmol Enz·g⁻¹DCW) and a 64% increase in specific activity production (AU·g⁻¹DCW) were reported for CBM9-*Tm*LDH variant, respect to CBM3-*Tm*LDH (Table 5.3).

Nevertheless, regarding the specific activity of both fusion proteins, similar values were obtained in both cases, even if slightly higher values were reported for CBM9-tagged enzyme (differences are not significant if the ± error is considered). Still, both processes were considered satisfactory since there was produced a significant amount of active enzyme

to work with and none of the two CBM domains fused to *Tm*LDH enzyme seemed to have any impact on catalytic activity.

Below, it is briefly presented a [general discussion of production tests](#) based on batch fermentations, whose main results are provided in [Figure 5.13](#).

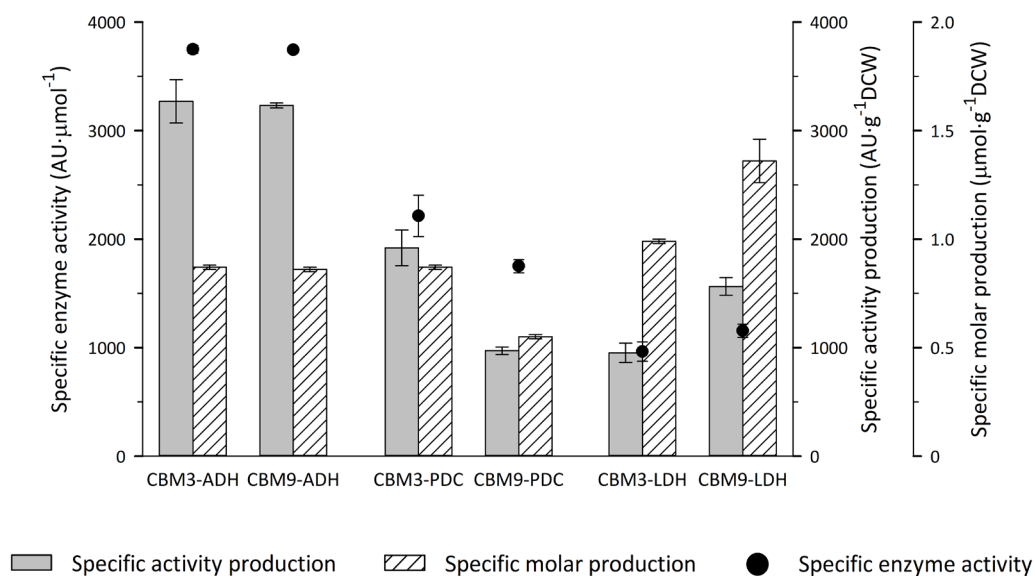


Figure 5.13. *E. coli* M15Δ*glyA* specific activity (AU·g⁻¹DCW) and specific molar (μmol·g⁻¹DCW) production values obtained in batch cultures performed to produce CBM-tagged *Sc*ADH, *Sc*PDC and *Tm*LDH enzymes with DM, and specific enzyme activity (AU·μmol⁻¹Enz) of the six fusion proteins overexpressed. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

In general, *E. coli* M15Δ*glyA* strain variants were able to produce more than one gram of active and soluble target product in all cases, corresponding to approximately 0.9–1 micromole of enzyme, fulfilling the obtention of enough amount of enzyme for the subsequent immobilisation experiments. Focusing on *E. coli* production among the two variants of each enzyme, there is not a clear relation between the CBM tag used and the production yields obtained, meaning that both CBM domains can be considered suitable candidates to be used in RPP as fusion tags.

However, the suitability of the two CBM used in this study should be evaluated for each enzyme individually, since different results were

obtained in each case; *ScADH* variants showed no significant differences among them, whereas *ScPDC* and *TmLDH* production values varied depending on which CBM was fused to them (Figure 5.13).

In addition, it can be deduced through the specific production profiles (Figures 5.10 to 5.12), that *E. coli* cells did not reach metabolic collapse in none of the fermentations, since the corresponding curves do not bend at the end of the cultures but show an ascendant tendency instead, except maybe for CBM9-*ScPDC* variant (Figure 5.11 B). Therefore, maximum production yields may not have been reached. Nevertheless, the objective of these experiments was far from process optimization, as mentioned above.

Finally, one common fact that was observed in all fermentations is that target protein overexpression mechanism was strongly repressed, since cell lysate samples corresponding to pre-induction showed minor enzyme activity values (basal activity), which increased notably along the induction phase of the fermentations, as expected.

5.3. CBM–tagged *ScADH* fed-batch production studies

Aiming to implement the production of CBM–tagged enzymes at industrial scale, the following studies at bench–scale reactor that were carried out consisted in the preliminary study of their production in fed-batch fermentations, with high cell densities cultures. The scale-up of the CBM–tagged enzymes production was useful to assess if higher productivities could be achieved compared to batch processes, since fed-batch duration was not limited by the depletion of initial glucose but could be extended until the metabolic collapse of the culture or until an insufficient oxygen transfer capacity in the reactor.

In this context, *ScADH* was used again as model enzyme, thanks to the high activities produced in batch cultures and the similarity of production yields obtained previously for the two CBM tested. Therefore, two fed-batch fermentations were performed to produce the CBM–tagged *ScADH* variants. Nonetheless, given that the intention was not to find the maximum production, fed-batch processes were not deeply analysed and subsequently optimised, and so, same general protocol used to produce histidine–tagged enzyme was followed (see Section 3.5, Cultivation conditions).

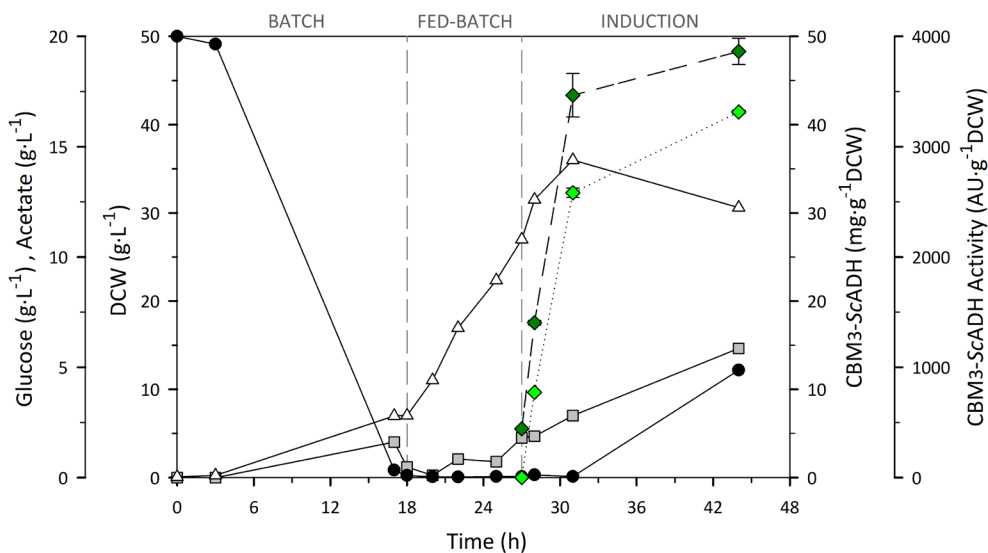
Fed-batch fermentations were carried out at 37°C, aiming to avoid an excessive process duration and considering that *E. coli* growth is favoured at this temperature, rather than at 24°C. However, it was set a specific growth rate of 0.2 h⁻¹ by controlling the addition of concentrated feeding solution to prevent an unrestrained cell growth. Thus, setting a μ lower than the calculated at 24°C in batch cultures, may allow to emulate the metabolic state of batch cultures.

Afterwards, production yields calculated in fed-batch processes were compared to the ones obtained in batch cultures to check if production yields were improved. Furthermore, specific productions and specific activities were also compared with histidine-tagged variant to determine if the *E. coli* M15 Δ glyA strain could produce CBM-fused proteins as efficiently as histidine-tagged and to evaluate the effect of the CBM fused tags to catalytic capacity of ScADH enzyme.

5.3.1. CBM-ScADH production in fed-batch cultures

As mentioned above, a fed-batch strategy was followed to overexpress the two CBM-fused ScADH variants. Fermentation parameters followed along the two processes are shown below:

A: CBM3-ScADH



B: CBM9-ScADH

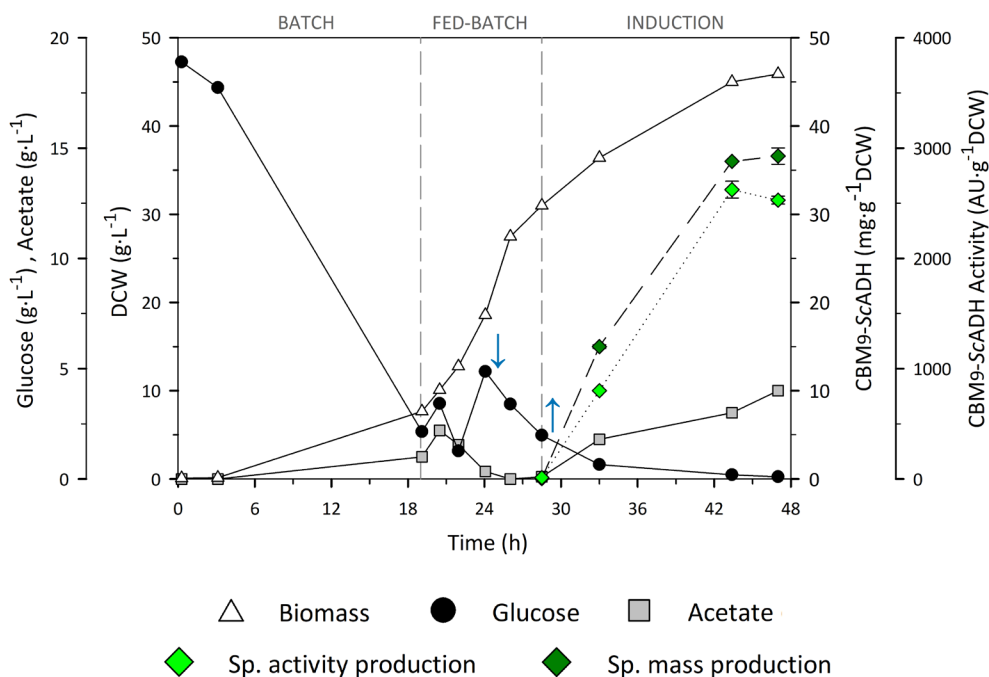


Figure 5.14. *E. coli* M15 Δ glyA fed-batch cultures performed to produce CBM3-ScADH (A) and CBM9-ScADH (B) fusion proteins with DM. Profiles along time of biomass DCW (Δ), glucose (\bullet) and acetate (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\cdot\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\cdot\text{DCW}$). Batch, fed-batch and induction phases indicated. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Arrows indicate the stop (\downarrow) and the resume of the feeding (\uparrow) when needed. Error bars correspond to standard error of three sample measurements.

Regarding the development of the processes, batch phase lasted for about 18 hours. Fed-batch phase duration was also similar in both cases, reaching the desired biomass concentration for culture induction in about 8 to 10 hours of exponential growth. Specific growth rate (μ) of 0.15 h^{-1} was calculated during exponential feeding addition for CBM3-ScADH variant (Figure 5.14 A), whereas a μ of 0.18 h^{-1} was determined for CBM9-fused variant (Figure 5.14 B). The differences between the pre-set μ (0.2 h^{-1}) in feeding addition and the experimentally measured μ could be due to the coefficient of cell maintenance (Vidal Conde, 2007), which was not considered in the design of the feeding addition profile.

Besides, CBM9-ScADH culture (Fig. 5.14 B) showed an accumulation of glucose at the early stage of fed-batch phase, probably caused by an acetate accumulation at the end of batch phase. Thus, feeding was sto-

pped until glucose concentration decreased below 5 g·L⁻¹ and acetate was completely exhausted.

Main production parameters calculated for each variant are shown in [Table 5.4](#). As previously explained, some of the results provided which are considered relevant cannot be compared between the two variants, since the resulting recombinant proteins have different specific mass, whereas some other results, expressed in enzyme AU or µmol, can be compared.

Table 5.4. Production parameters calculated for CBM-fused ScADH enzymes using *E. coli* M15ΔglyA and DM in fed-batch cultures at 2 L-reactor scale. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. ± correspond to the standard error of three sample measurements. Comparable parameters are presented in bold.

Production parameter	CBM3-ScADH	CBM9-ScADH
Total biomass (g DCW)	61.8 ± 0	90 ± 0
Final volume (L)	2.0	1.96
Enzyme activity (AU·L ⁻¹)	1.02·10 ⁵ ± 1130	1.16·10 ⁵ ± 1640
Enzyme titre (mg Enz·L ⁻¹)	1550 ± 107	1780 ± 21
Specific production (AU·g⁻¹DCW)	3316 ± 42	2528 ± 36
Specific production (mg Enz·g ⁻¹ DCW)	50.7 ± 3	38.7 ± 0
Specific production (µmol Enz·g⁻¹DCW) *	0.93 ± 0.1	0.61 ± 0
Specific activity (AU·mg ⁻¹ Enz)	65.4 ± 1	65.3 ± 1
Specific activity (AU·µmol⁻¹Enz) *	3580 ± 53	3860 ± 59

* Considering a molar weight of 55 kDa for CBM3-ScADH and of 64 kDa for CBM9-ScADH.

As shown in [Table 5.4](#), similar values were obtained for the two variants in most of the parameters calculated, as was reported for batch cultures ([Table 5.1](#)). 3.13 grams of CBM3-ScADH and 3.48 grams of CBM9-ScADH were produced, and similar total enzyme activity units were also achieved (2.05·10⁵ and 2.27·10⁵ AU, respectively). However, a 30% increase of specific activity production and a 52% increase of specific molar production were calculated for CBM3-tagged variant, compared

to CBM9-ScADH, even if the second variant showed slightly higher specific activity.

After this first checking of the production in fed-batch cultures with DM, it can be concluded that both CBM domains fused to ScADH enzyme were successfully produced, and that results obtained were very similar between them, even if some operational issues were reported in the two processes. Yet, better results could certainly be obtained in future experiments, by optimizing some operational parameters such as the moment of induction.

As explained, the **change of the production approach from batch to fed-batch** cultures, was useful to assess if improved production yields could be reached, even if both processes were not previously optimised. In general, same specific yields were calculated for fed-batch fermentations compared with batch processes, and similar specific activity of the fusion proteins produced were achieved (Tables 5.1 and 5.4).

However, fed-batch strategy led to an increase in enzyme titre ($\text{mg}\cdot\text{L}^{-1}$) of one order of magnitude, and a 5.9- to 6.3-fold increase was reported in alcohol dehydrogenase activity produced per culture litre ($\text{AU}\cdot\text{L}^{-1}$), for CBM3- and CBM9-ScADH enzymes, respectively. Moreover, volumetric productivity of enzyme activity was triplicated for CBM3-tagged variant and was almost quadruplicated for CBM9-ScADH. Finally, volumetric productivity expressed in mass terms ($\text{mg}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$) was also triplicated in both cases.

Thus, fed-batch cultures were proved to be the most effective strategy tested, since remarkably higher titres were produced, meaning that this approach brings a more efficient use of time and defined culture medium.

5.3.2. Comparison between histidine- and CBM-tagged ScADH variants

Finally, the comparison between the three variants of ScADH enzyme was focused on *E. coli* M15 Δ glyA specific production yields ($\mu\text{mol}\cdot\text{g}^{-1}\cdot\text{DCW}$ and $\text{AU}\cdot\text{g}^{-1}\cdot\text{DCW}$) and the specific molar activity ($\text{AU}\cdot\mu\text{mol}^{-1}$) of each variant, which are the three comparable parameters, as previously explained. The corresponding values obtained in 2 L fed-batch fermentations are provided in Figure 5.15.

On the one hand, similar specific production values were obtained among the three variants; M15 Δ glyA strains produced 1.04 μmol of His-ScADH, 0.87 μmol of CBM3-ScADH and 0.86 CBM9-ScADH per gram of DCW. Besides, similar total enzyme amounts were produced of each ScADH variant, obtaining 3.58, 3.13 and 3.48 grams of enzyme, respectively. These are promising results, since they proved that CBM-tagged enzymes can be produced at high tires with the chosen RPP system, at similar yields than histidine-tagged enzymes.

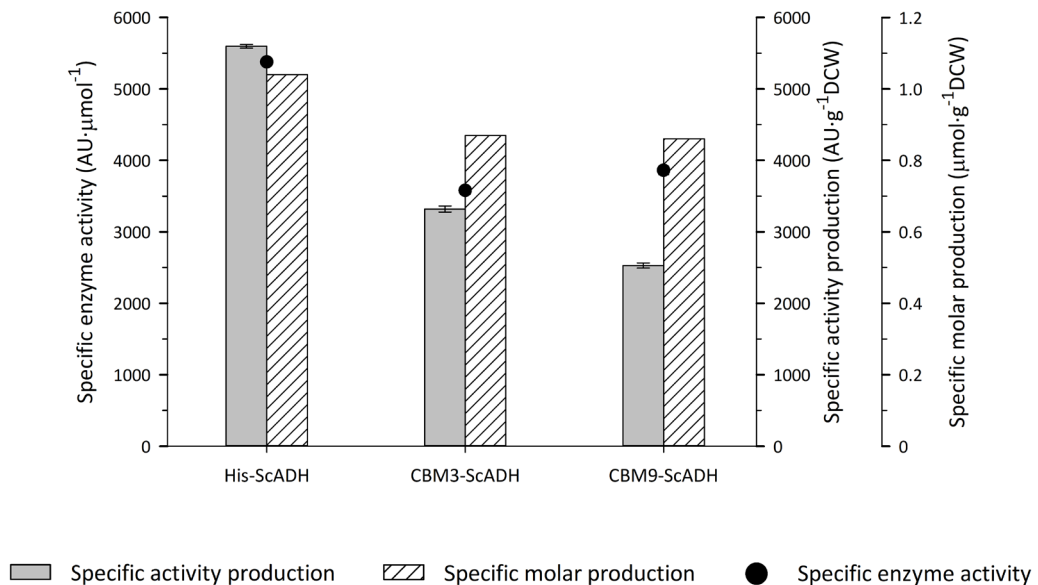


Figure 5.15. *E. coli* M15 Δ glyA specific activity (AU·g⁻¹DCW) and specific mass (mg·g⁻¹DCW) production values obtained in fed-batch cultures performed to produce histidine- and CBM-tagged ScADH enzymes with DM, and specific enzyme activity (AU·μmol⁻¹Enz) of the three fusion proteins. Culture conditions: 37°C, pH 7.0, 450-1200 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

On the other hand, specific activity production varied among the three variants, since more than 5000 AU of the histidine-tagged ScADH variant were produced per gram of biomass, whilst approximately 3300 and 2500 AU·g⁻¹DCW were obtained for CBM3- and CBM9-tagged variants, respectively. In addition, specific enzyme activity of His-ScADH variant was 1.4-fold higher than CBM-fused ones, showing up to 5400 AU per micromole of enzyme.

Even if these results could suggest that CBM domains negatively affect the catalytic capability of *ScADH* enzyme, the resulting CBM-tagged enzymes could be produced at high concentrations in a soluble and functional form as well, and total produced activity values still fluctuated within the same magnitude order. Besides, CBM tags can provide significant advantages in subsequent purification and immobilisation processes. It must be considered that the cost of cellulosic supports is between 3 and 50-fold lower than functionalised agarose beads (depending on the supplier). Hence, the decrease in enzyme specific activity shall be considered an affordable compromise, aiming to develop a more sustainable and cost-effective downstream process which would also expand the range of possible applications, since metal ions would be avoided from final products.

Finally, the **specific enzyme activities of *ScPDC* and *TmLDH* enzymes** were compared among their three variants. However, as previously stated, this comparison was done to try to find general conclusions or a possible tendency in regard to the correlation between the enzyme activity and the fusion tags, and thus, this comparison was not addressed to the production processes but was focused only on the specific enzyme activities measured in final products.

In this context, results do not show a clear relation between the N-terminal tag used and the specific activities obtained, as happened when production yields were analysed. On the one hand, *ScADH* and *TmLDH* enzymes were more active in their histidine-tagged versions, rather than with CBM domains fused to them, e. g., 2200 AU were measured per μmol of His-*TmLDH* enzyme, whilst approximately 1200 and 1000 AU were measured for CBM9- and CBM3- *TmLDH* variants, respectively.

On the other hand, *ScPDC* enzyme showed higher specific activity in its CBM3-tagged version ($2200 \text{ AU}\cdot\mu\text{mol}^{-1}$), which corresponded to 1.3-fold increase in specific activity, compared to His-*ScPDC* variant. Hence, the differences between the three enzymes and the lack of a clear behaviour hinder the conception of a general conclusion, meaning that the adequacy to use any of the two CBM domains should probably be evaluated for each enzyme individually.

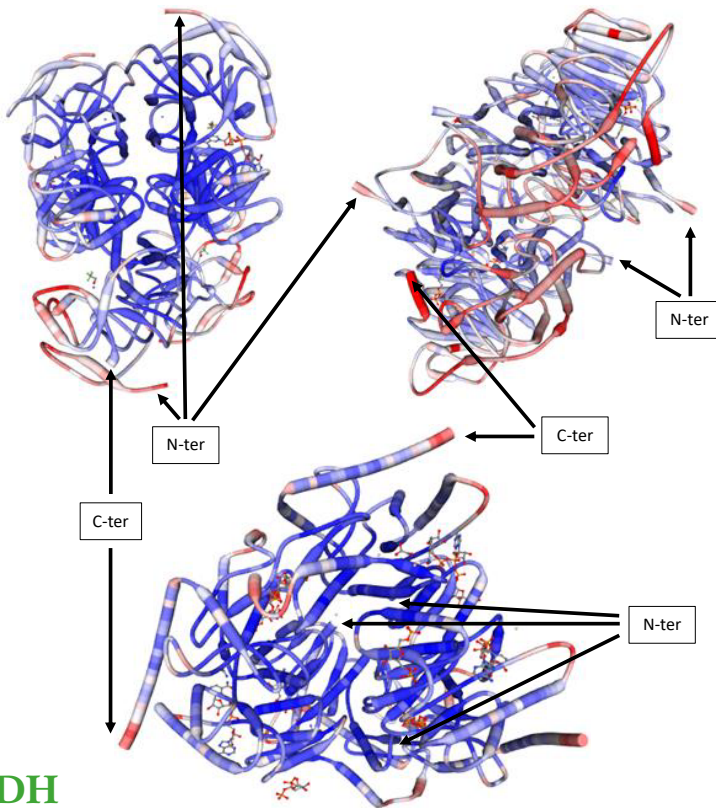
One possible hypothesis that would explain the differences observed in specific activity values between the three variants of each enzyme is that the fusion of CBM domains at N-terminal end probably may

alter, in a major grade than histidine tag, key space regions such as the catalytic centre or the substrate-binding region, or may affect to the interaction between the monomers —*ScADH*, *ScPDC* and *TmLDH* are actually tetrameric enzymes, as previously described— that shape the quaternary structure of the proteins.

Three-dimensional images of the enzymes tested (Figure 5.16) help to comprehend that the relative position of the amino- and carboxyl-terminal ends vary among each protein, and that the orientation of the N-terminus (where CBM domains are fused) vary among the three enzymes. E.g., *TmLDH* amino-terminal ends are oriented towards the inner region (Figure 5.16 C), and probably the fusion of a heavy-weighted tag such as CBM domains should be tried at C-terminal end instead.

A: *ScADH*

B: *ScPDC*



C: *TmLDH*

Figure 5.16. X-ray diffraction images of *ScADH* (2.40 Å) (A), *ScPDC* (2.40 Å) (B) and *TmLDH* (2.10 Å) (C) enzymes obtained from SWISS-MODEL (Bienert et al., 2017), in their homotetrameric quaternary structure. N- and C-terminal ends indicated with arrows.

Unfortunately, as said before, the work described in this document is part of a time-limited and well-delimited European project. Therefore, it was not possible to deeply address the study of CBM tags effects on enzyme structures. Nevertheless, it would be an interesting question to explore in future experiments.

5.4. Conclusions

Molecular biology experiments led to the correct cloning of the four DNA constructs into the pVEF plasmid vector, consisting in the fusion of the alcohol dehydrogenase from *S. cerevisiae* with four different CBM domains. Once again, SLIC technique turned out to be an effective and cost-saving method, since DH5 α cell colonies were grown in LB-agar plates in all cases after cell transformation process, what reinforced the robustness and the effectiveness of the cloning method.

Two of the newly *E. coli* DH5 α transformed variants were able to overexpress the target enzyme, corresponding to CBM3- and CBM9-tagged versions. These two CBM domains were also successfully fused to pyruvate decarboxylase and lactate dehydrogenase enzymes, and *E. coli* M15DglyA cells were transformed with the resulting plasmids.

The six fusion proteins could be also overexpressed at high titres in a bench-scale reactor with the auxotrophic strain and DM via batch fermentations. ScADH variants showed no significant differences between the two variants, whereas ScPDC and TmLDH production values varied depending on the CBM. More than one gram of target product was produced in all cases, fulfilling the obtention of enough amount of enzyme for the subsequent immobilisation experiments.

Finally, it was compared the specific enzyme activity of the CBM-tagged enzymes respect to their histidine-tagged versions, and different results were obtained for each of the enzymes tested, even if in all cases total produced activity fluctuate within the same magnitude order.

It can be theorised that the relative position of the fused CBM domains at N-terminal end may affect to tertiary and quaternary structure of the peptides, and thus, to their catalytic activity. Nonetheless, definitely satisfactory results were obtained, since there were obtained six different CBM-tagged enzymes in a soluble and active form, meaning that the objective was achieved.

6. RESULTS III: ONE-STEP PURIFICATION/IMMOBILISATION OF CBM-FUSED ENZYMES

Abstract

CBM-fused proteins obtained by recombinant protein production at bench-scale processes were bound to a low-cost and highly selective cellulosic support by one-step purification/immobilisation process. CBM9-fused enzymes proved to bind effectively to regenerated amorphous cellulose (RAC) at high loads, whereas CBM3-tagged enzymes bound with higher affinity onto microcrystalline cellulose (MC). The best results with both tags were obtained with *ScADH* enzyme —more than 100 mg·mL⁻¹ support were immobilised, retaining almost total activity offered—. In addition, RAC support was also used for protein purification, aiming to establish an alternative to metal affinity chromatography. CBM9 domain proved to be a suitable tag for enzyme purification in all tested biocatalysts using glucose as eluent. However, CBM3-tagged enzymes were not efficiently desorbed using the methodology developed for CBM9-tagged variants. Once again, the best results were achieved for *ScADH* enzyme, with a recovery yield of >92% and almost a 10-fold higher purity grade.

6. RESULTS III: ONE-STEP PURIFICATION/IMMOBILISATION OF CBM–FUSED ENZYMES

6.1. Immobilisation of CBM–fused enzymes into cellulosic support

As described in the introduction of this work (Section 1.4, Enzyme purification and immobilisation as process intensification strategies), one approach commonly used to bring biotransformations closer to industrial feasibility is biocatalyst immobilisation. The use of immobilised derivatives containing biocatalysts does not only allow to reuse them but can also improve their stability towards several reaction conditions such temperature, pH, or solvents, thereby increasing the biocatalyst yield (Tischer and Wedekind, 1999).

However, one of the main drawbacks that characterises enzyme immobilisation is the cost of commercial supports used, which contribute significantly to the final downstream cost. Therefore, the use of cheap immobilisation supports such as cellulose stands as a promising alternative, especially for industrial–scale processes focused on the production of low–valued final products.

The experiments described below were carried out aiming to establish a one-step purification/immobilisation process for the CBM–fused enzymes that have been successfully overexpressed and produced in *E. coli* (see Section 5.2, CBM–tagged enzymes production with the M15 Δ glyA strain). The immobilisation method is then based on affinity interactions between CBMs and cellulosic supports. The characterization of the CBM3– and CBM9–fused enzymes affinity towards cellulose was performed by immobilising the enzymes produced at bench–scale reactor onto a regenerated amorphous cellulose (RAC) support.

Firstly, immobilisation experiments were carried out loading low amounts of enzyme aiming to avoid diffusional limitations, so that immobilisation parameters such as retained activity could be determined without the effect of this sort of hindrances. Afterwards, the stability of both immobilised and soluble enzymes under storage conditions (50 mM Tris-HCl buffer pH 7.50, 4°C) was studied. Finally, the maximum enzyme loading capacity of RAC support was analysed for both CBMs by increasing the offered enzyme quantities.

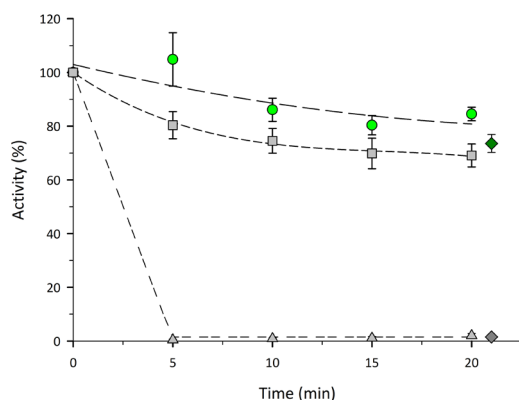
Though, considering that diverse CBM families can be established according to the type of compounds by which these domains present a greater binding affinity (Section 1.5, Carbohydrate–Binding Modules), CBM3–tagged enzymes were also immobilised to non–treated microcrystalline cellulose (MC), since CBM3 domains are characterized to bind with higher affinity towards crystalline cellulose supports, whereas CBM9 domains rather bind to amorphous cellulose instead (Oliveira et al., 2015).

6.1.1. CBM3–fused enzymes

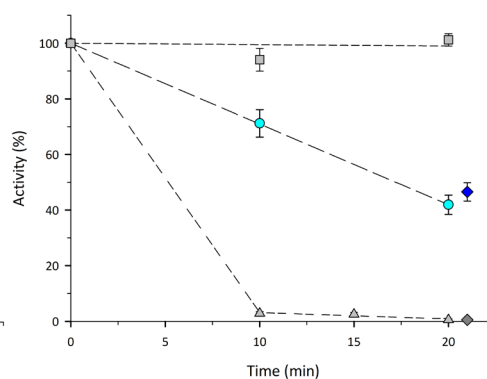
First batch experiments were carried out to **characterise the immobilisation process** of CBM3–tagged enzymes towards the amorphous cellulose. The enzyme activities measured along the immobilisation processes of the three enzymes are depicted in [Figure 6.1](#), whereas the immobilisation parameters calculated are listed in [Table 6.1](#).

The experiments performed revealed the high affinity of the carbohydrate–binding module towards the cellulosic support, as well as the short time needed to establish the union between them, since the equilibrium was reached after five to ten minutes of incubation in all cases. Even if almost total binding of target protein was achieved for two of the three enzymes tested, the retained activity of the immobilised derivatives varied significantly among the three cases ([Figure 6.1](#) and [Table 6.1](#)). In all cases, blank samples (lysate without support) kept a high percentage of initial activity (> 75–80%), what meant that experiment conditions were suitable, in terms of enzyme stability.

A: CBM3-ScADH



B: CBM3-ScPDC



C: CBM3-TmLDH

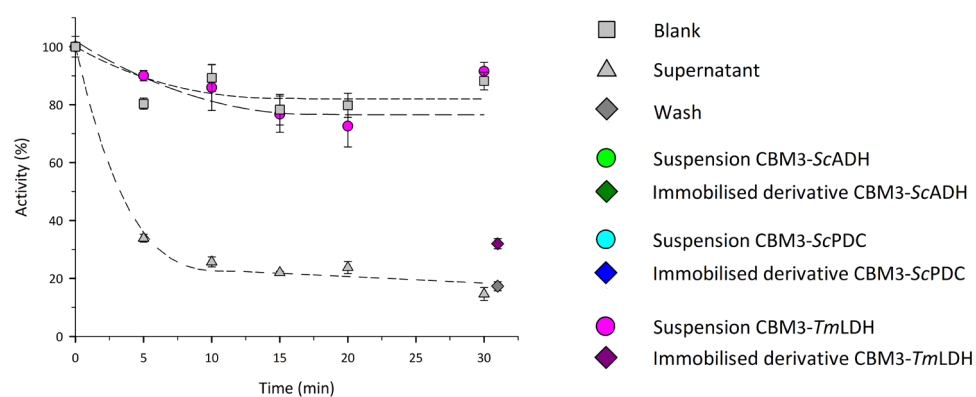


Figure 6.1. Enzyme activities measured along the immobilisation processes of CBM3-ScADH (A), CBM3-ScPDC (B) and CBM3-TmLDH (C) with 1 mL of RAC support. Blank (□), suspension (●●), supernatant (Δ), retained activity of immobilised derivatives (◆◆) and supernatant activity after wash (◇). Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, roller agitation. Error bands correspond to the standard error of three replicates.

Table 6.1. Immobilisation parameters calculated for CBM3-fused enzymes onto RAC support. Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, 1 mL RAC support, roller agitation. ± correspond to the standard error of three replicates.

Enzyme	Total protein offered (mg)	Enzyme offered (mg·mL ⁻¹ RAC)	Enzyme offered (AU·mL ⁻¹ RAC)	IY (%)	RA (%)	Recovered activity (%) *
ScADH	2.5 ± 0	0.4 ± 0	29 ± 1	98 ± 1	86 ± 2	74 ± 3
ScPDC	1.5 ± 0	0.2 ± 0	49 ± 1	98 ± 1	43 ± 3	46 ± 3
TmLDH	11.4 ± 1	3.8 ± 0	50 ± 2	85 ± 3	77 ± 7	32 ± 1

IY: immobilisation yield; RA: retained activity. * Activity recovered after sample filtration.

Results showed a slight deactivation of **CBM3–fused ScADH** (Figure 6.1 A) due to the immobilisation process, but all activity was maintained in the final immobilised derivative (86.1%). On the contrary, a progressive activity loss was observed during the immobilisation of **ScPDC enzyme** (Figure 6.1 B) —which also showed no activity in supernatant—, whose derivative retained a 45.6% of total offered activity. Although this fact may indicate the presence of diffusional limitations, it could not be checked since the experiment was carried out at activity test threshold detection. Finally, **TmLDH** suspension (Figure 6.1 C) kept the whole activity (91.5%) during the experiment, but the recovered support after filtration showed low retained activity (32%). Moreover, a 14.7% of total LDH activity remained in the supernatant fraction, probably due to an excessive enzyme loading. This fact occurred because all the experiments were set up according to enzyme activity of the samples, and not to the total amount of target protein offered. I.e., it is possible that the amount of enzyme offered (3.8 mg) exceeded the maximum loading capacity (Table 6.1).

Protein analysis was estimated for the three experiments and the results obtained are listed in Table 6.2. Briefly, total protein content of blank and supernatant was calculated by measuring their protein concentrations, and the amount of enzyme immobilised to RAC was estimated by measuring the difference of relative intensity of the band corresponding to target enzyme (in SDS-PAGE) between the blank samples and the supernatant samples (Figures 6.2, 6.3 and 6.4).

Table 6.2. Protein analysis of immobilisation experiments of CBM3–fused enzymes onto RAC support. Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, 1 mL RAC support, roller agitation. ± correspond to the standard error of three replicates.

Enzyme	Total protein offered (mg)	Initial enzyme content (%)	Total protein in final supernatant (mg)	Total protein immobilised (mg)	Total enzyme immobilised* (mg)	Enzyme content in the immobilised derivative (%)
ScADH	2.5 ± 0	15 ± 1	2.1 ± 0	0.4 ± 0	0.34 ± 0	85
ScPDC	1.5 ± 0	13 ± 0	1.2 ± 0	0.3 ± 0	0.2 ± 0	67
TmLDH	11.4 ± 1	33 ± 1	8 ± 1	3.4 ± 0	3.1 ± 0	91

* Calculated from SDS-PAGE differences between blank and final supernatant samples.

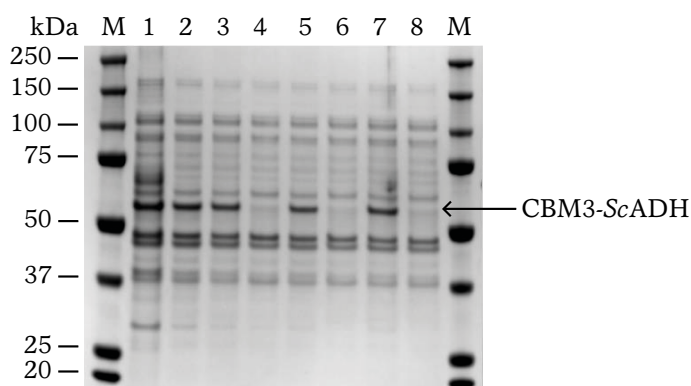


Figure 6.2. SDS-PAGE of the immobilisation experiments of CBM3-ScADH fusion protein onto cellulosic RAC support. Lanes M: molecular weight standard (kDa); lanes 1 and 2: clarified cell extract before and after diluting, respectively; lanes 3, 5 and 7: blank samples of each replicate; lanes 4, 6 and 8: final supernatant samples of each replicate. CBM3-ScADH (57 kDa) corresponding band indicated with arrow.

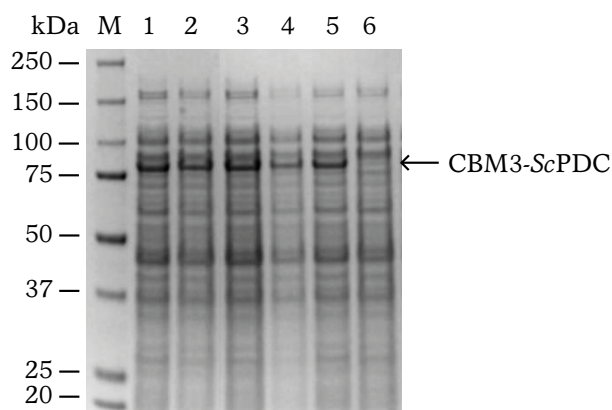


Figure 6.3. SDS-PAGE of the immobilisation experiments of CBM3-ScPDC fusion protein onto cellulosic RAC support. Lane M: molecular weight standard (kDa); lanes 1 and 2: clarified cell extract before and after diluting, respectively; lanes 3 and 5: blank samples of two replicates; lanes 4 and 6: final supernatant samples of two replicates. CBM3-ScPDC (80 kDa) corresponding band indicated with arrow.

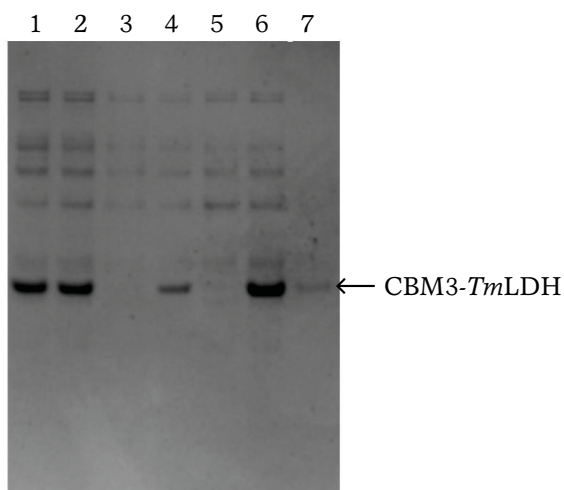


Figure 6.4. SDS-PAGE of the immobilisation experiments of CBM3-*TmLDH* fusion protein onto cellulosic RAC support. Lane 1: clarified cell extract; lanes 2, 4 and 6: blank samples of each replicate; lanes 3, 5 and 7: final supernatant samples of each replicate. CBM3-*TmLDH* (53 kDa) corresponding band indicated with arrow.

Regarding the protein content analysis (Table 6.2), in all cases the high specificity of the binding between CBM3 domain and the cellulosic support was demonstrated, since total protein content difference between initial and final supernatant samples were close to the amount of target protein offered to support. This fact can be clearly observed in protein electrophoresis (Figures 6.2, 6.3 and 6.4), where a remarkable decrease of the bands corresponding to target enzymes can be seen in all three cases.

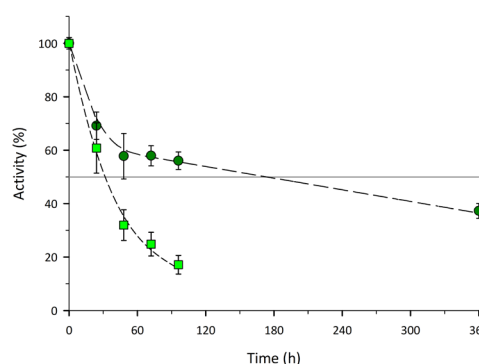
For example, it was estimated that 0.37 mg of CBM3-*ScADH* enzyme were offered (Table 6.1), and according to SDS-PAGE (Figure 6.2), and relative content in supernatant decreased from 14.6% to 0.7%, what means that approximately 0.34 mg of target protein were bound to RAC (Table 6.2), which corresponded to the 86% of total protein immobilised (0.4 mg). Moreover, this value coincides with what was measured by activity assay, since mass balance showed a IY of 92% (0.34 of 0.37 mg) whereas a IY of 98% was estimated by activity assay (Table 6.1).

Protein analysis also revealed a high specificity for the rest of the enzymes towards RAC; almost all CBM3-*ScADH* enzyme offered —0.17 mg of 0.2 mg— was attached to RAC, increasing the enzyme content from 13 to 67% after immobilisation (Table 6.2). Regarding CBM3-*TmLDH*,

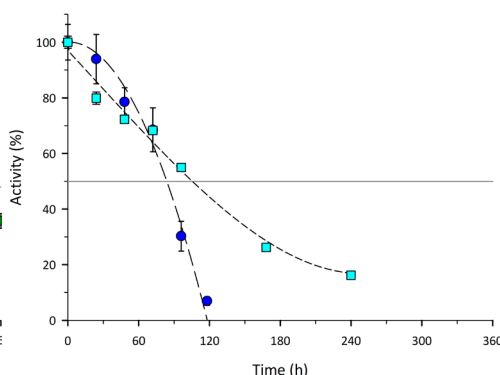
3.11 mg of enzyme from the 3.8 mg offered (82%) were bound to RAC, reaching more than 91% of enzyme content in the final immobilised derivative (Table 6.2). Moreover, the IY obtained by activity assay was close to the value obtained by protein analysis (85%).

The enzyme storage stability assessment of both soluble enzymes and immobilised derivatives is depicted in Figure 6.5:

A: CBM3-ScADH



B: CBM3-ScPDC



C: CBM3-TmLDH

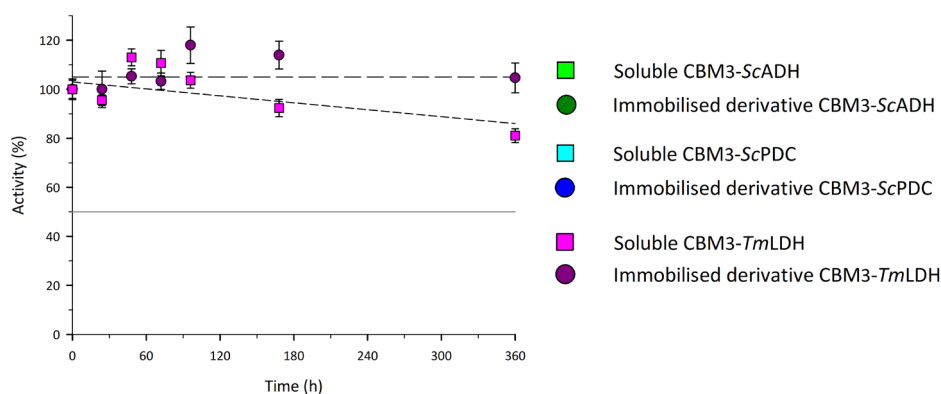


Figure 6.5. Stability of CBM3-ScADH (A), CBM3-ScPDC (B) and CBM3-TmLDH (C) fusion proteins, soluble (□) and immobilised onto RAC support (○), under storage conditions (50 mM Tris-HCl buffer pH 7.50, 4°C). Error bars correspond to standard error of three replicates.

On the one hand, the experiment revealed the notorious high stability of TmLDH enzyme (Figure 6.5 C), whose immobilised derivative kept the whole activity a fortnight after the beginning of the experiment and

the soluble enzyme only suffered a 20% activity loss. Nevertheless, it should be mentioned that this fact was somehow expected, considering that *Tm*LDH enzyme was isolated from an extremophile organism. On the other hand, results showed that the immobilisation process allowed a 5.5-fold increase of half-life time for *Sc*ADH enzyme—from 32 to 174 hours— (Figure 6.5 A), which turned out to be the less stable protein among all soluble ones, under the tested conditions.

On the contrary, immobilisation process did not result in an increased stability of *Sc*PDC enzyme (Figure 6.5 B), being the half-life of immobilised derivative a 20% lower than soluble one (84 and 105 hours, respectively).

The **maximum enzyme loading capacity** of CBM3–fused enzymes on RAC support was determined by increasing the offered enzyme quantities to the support (Table 6.3). In this case, due to mass transfer limitations, steric hindrances and other possible phenomena commonly associated to highly loaded immobilisation supports, retained activity values were underestimated (Wingard et al., 1976). Thus, retained activity coefficient previously assessed with no-limiting conditions (Table 6.1) was used to calculate the theoretical final retained activity of the high-loaded derivatives by assuming to be equal in all cases, since it is not dependent on enzyme amount (García-Bofill et al., 2021, b). For experiments with MC, retained activity coefficient was also calculated previously, obtaining a 95% RA for CBM3-*Sc*ADH, an 86% for CBM3-*Sc*PDC and a 92% for CBM3-*Tm*LDH (Table 6.3).

Additionally, the maximum loading capacity of CBM3–tagged enzymes on MC support was also determined, by performing the same experiments concurrently with the two types of cellulose, aiming to explore other alternatives with CBM3 domains. As mentioned, maximum capacity was assessed by increasing the amount of enzyme offered to 1 mL of support, and experiments stopped once IY decreased significantly from 100% (< 90%), meaning that a remarkable percentage of enzyme offered remained unbound in the supernatant.

Table 6.3. Results obtained of maximum enzyme loading capacity of CBM3–fused enzymes onto amorphous (RAC) and microcrystalline (MC) cellulosic supports. Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, 1 mL RAC support, roller agitation. \pm correspond to the standard error.

Enzyme	Support	Enzyme offered (AU·mL ⁻¹ support)	IY (%)	Enzyme immobilised (AU·mL ⁻¹ support)	Theoretical RA (AU·mL ⁻¹ support)	Measured RA (AU·mL ⁻¹ support)
ScADH	RAC (RA: 86 %)	501 \pm 4	100 \pm 0	501 \pm 0	431 \pm 11	74 \pm 3
		1002 \pm 9	100 \pm 0	1001 \pm 0	861 \pm 22	400 \pm 6
		5012 \pm 33	86 \pm 5	4311 \pm 263	3710 \pm 110	962 \pm 90
	MC (RA: 95 %)	9963 \pm 50	50 \pm 1	4976 \pm 84	4284 \pm 219	686 \pm 80
		501 \pm 3	100 \pm 0	501 \pm 0	476 \pm 2	239 \pm 1
		1003 \pm 4	100 \pm 0	1002 \pm 0	952 \pm 4	537 \pm 7
5012 \pm 44		100 \pm 0	5001 \pm 1	4751 \pm 20	1741 \pm 75	
		10025 \pm 5	94 \pm 0	9468 \pm 2	8995 \pm 40	1282 \pm 33
ScPDC	RAC (RA: 43 %)	49 \pm 1	98 \pm 1	48 \pm 0	21 \pm 0	23 \pm 2
		200 \pm 10	79 \pm 1	159 \pm 2	68 \pm 5	42 \pm 2
		546 \pm 17	42 \pm 2	228 \pm 12	98 \pm 15	108 \pm 13
	MC (RA: 86 %)	49 \pm 1	100 \pm 0	49 \pm 0	42 \pm 0	44 \pm 1
		250 \pm 24	88 \pm 1	220 \pm 3	189 \pm 0	174 \pm 3
		501 \pm 47	72 \pm 8	363 \pm 42	311 \pm 0	219 \pm 9
TmLDH	RAC (RA: 77 %)	50 \pm 2	85 \pm 3	43 \pm 2	33 \pm 3	16 \pm 1
		100 \pm 2	72 \pm 1	73 \pm 1	56 \pm 7	54 \pm 0
		50 \pm 1	92 \pm 1	46 \pm 1	42 \pm 0	39 \pm 1
	MC (RA: 92 %)	100 \pm 2	76 \pm 1	76 \pm 1	69 \pm 3	26 \pm 1

IY: immobilisation yield; RA: retained activity. Theoretical RA coefficients (2nd column) were determined under no diffusional-limiting conditions.

In most of the cases, it was observed that maximum loading value increased among two consecutive experiments (Table 6.3), even if both were performed at high loads (IY lower than 100%), what could reveal that multi-layer immobilisation is occurring. For example, $3710 \pm 110 \text{ AU}\cdot\text{mL}^{-1}$ of CBM3-ScADH were bound to RAC when $5000 \text{ AU}\cdot\text{mL}^{-1}$ were offered, reaching a IY lower than 100% (86%), what would indicate that maximum loading capacity was already reached. However, when higher amounts were offered ($10000 \text{ AU}\cdot\text{mL}^{-1}$) the final immobilised activity obtained was 15% higher ($4284 \pm 220 \text{ AU}\cdot\text{mL}^{-1}$).

This phenomenon has already been reported in other immobilisation processes (Solé et al., 2019, García-Bofill et al., 2021, a). Besides, it was proved that these supposed multi-layered derivatives could keep bound all the immobilised protein after their recovery, since the activity measurement of the filtered volume—which comprised the filtered supernatant and the subsequent wash—did not vary significantly from the last measure of supernatant prior to filtration step. For that, the maximum loading values obtained inside the range of activities tested are presented in bold in Table 6.3 and are shown in Figure 6.6, together with their corresponding mass values:

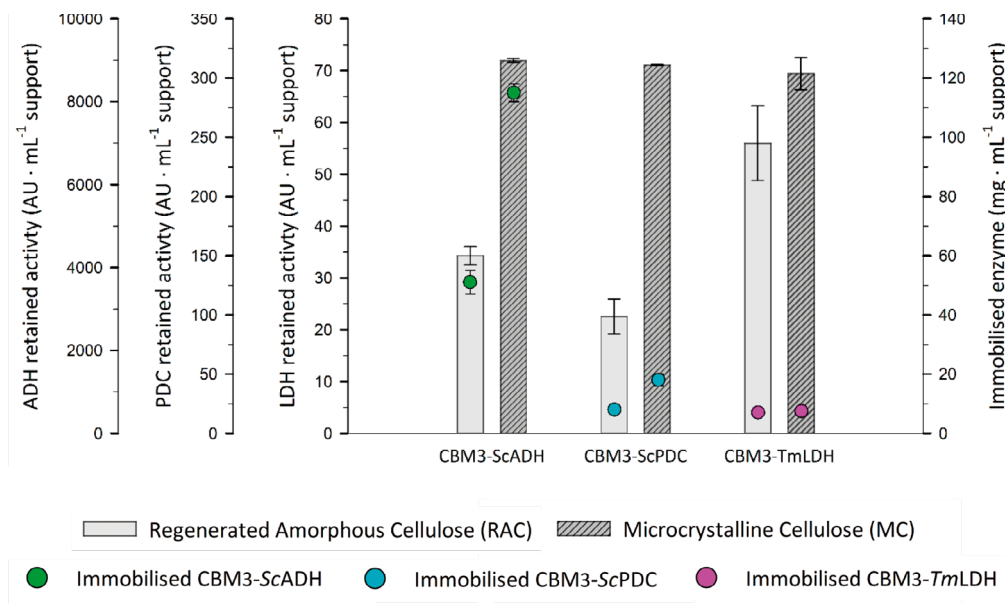


Figure 6.6. Maximum enzyme loading capacity ($\text{AU}\cdot\text{mL}^{-1}$) of CBM3-fused enzymes onto amorphous (RAC) and microcrystalline (MC) cellulosic supports (bars), and their corresponding mass values (mg) (\circ). Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C , 1 mL RAC support, roller agitation. Error bars correspond to the standard error.

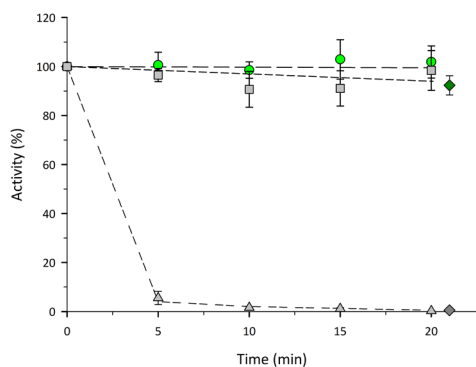
Regarding the experiments with MC support, higher maximum loadings were obtained compared with the ones obtained with RAC, even if differences between the two supports varied among the enzymes tested. For *ScADH* enzyme, maximum retained activity was 2.2-fold higher for MC support in contrast to RAC, which also corresponded to higher amounts of enzyme immobilised ($\text{mg}\cdot\text{mL}^{-1}$). Similar results were obtained for *ScP-DC* enzyme, showing a 3-fold increase in activity immobilised compared with RAC, that corresponded to a 2.3-fold increase in total amount of immobilised enzyme (21.5 vs $9.2 \text{ mg}\cdot\text{mL}^{-1}$). However, minor differences were obtained for *TmLDH*, with which a 20% increase in maximum loading was reached with MC support, corresponding with an increase of the same proportion in terms of enzyme amount (6.3 vs $7.5 \text{ mg}\cdot\text{mL}^{-1}$ for RAC and for MC, respectively).

In summary, results obtained for CBM3-tagged enzymes immobilisation experiments demonstrated that the most suitable strategy with CBM3 fusion tag would be using microcrystalline cellulose instead of amorphous.

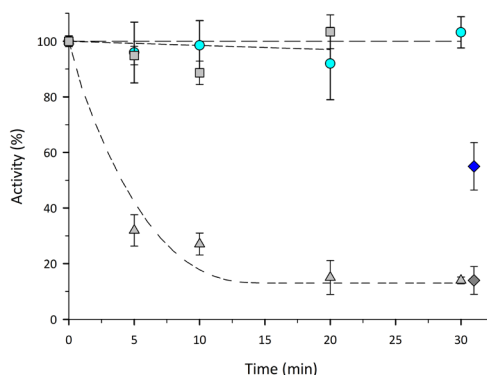
6.1.2. CBM9-fused enzymes

The **characterisation of the immobilisation process** of CBM9-tagged enzymes onto the amorphous cellulose (Figure 6.7) confirmed once again the high affinity of CBMs towards cellulose. As happened with CBM3 domain, the equilibrium was reached after just five to ten minutes of incubation for the three enzymes.

A: CBM9-*ScADH*



B: CBM9-*ScPDC*



C: CBM9-*TmLDH*

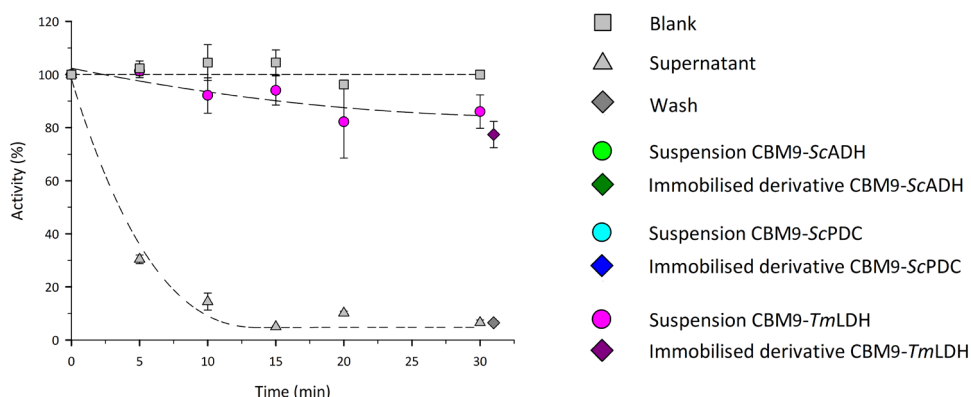


Figure 6.7. Enzyme activities measured along the immobilisation processes of CBM9-*ScADH* (A), CBM9-*ScPDC* (B) and CBM9-*TmLDH* (C) with 1 mL of RAC support. Blank (\square), suspension (\bullet), supernatant (Δ), retained activity of immobilised derivatives (\blacklozenge) and supernatant activity after wash (\blacklozenge). Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, roller agitation. Error bands correspond to the standard error of three replicates.

The totality of blanks and almost all suspension samples kept the whole initial activity offered, having only observed a slight deactivation of CBM9-fused *TmLDH* protein (14%) in suspension measurements (Figure 6.7 C), meaning that experiment conditions were suitable, as it was already assessed for CBM3-tagged proteins.

Total binding of target protein was achieved for *ScADH* enzyme (Figure 6.7 A), and so occurred for *TmLDH*, since activity measured in supernatant could be considered residual (6.5%). However, it was measured a 14% of initial *ScPDC* activity in final supernatant samples (Figure 6.7 B), what meant that probably the maximum loading capacity of RAC

support was exceeded, as occurred with CBM3-*Tm*LDH enzyme (Figure 6.1 C).

Immobilisation yield and retained activity coefficients (Table 6.4) were calculated for the three enzymes, and protein analysis (Table 6.5) was checked by following the same method used with CBM3-tagged proteins. In this case, electrophoresis gels with which relative intensity bands of target proteins were calculated are shown in Figures 6.8, 6.9 and 6.10.

Table 6.4. Immobilisation parameters calculated for CBM9-fused enzymes onto RAC support. Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, 1 mL RAC support, roller agitation. \pm correspond to the standard error of three replicates.

Enzyme	Total protein offered (mg)	Enzyme offered (mg·mL ⁻¹ RAC)	Enzyme offered (AU·mL ⁻¹ RAC)	IY (%)	RA (%)	Recovered activity (%) *
<i>Sc</i> ADH	2.4 \pm 0	0.2 \pm 0	16 \pm 3	100 \pm 0	98 \pm 2	92 \pm 4
<i>Sc</i> PDC	15.4 \pm 1	2 \pm 0	46 \pm 6	84 \pm 2	85 \pm 6	55 \pm 8
<i>Tm</i> LDH	8.3 \pm 0	1 \pm 0	43 \pm 1	94 \pm 2	79 \pm 4	77 \pm 5

IY: immobilisation yield; RA: retained activity. *Activity recovered after sample filtration.

Table 6.5. Protein analysis of immobilisation experiments of CBM9-fused enzymes onto RAC support. Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, 1 mL RAC support, roller agitation. \pm correspond to the standard error of three replicates.

Enzyme	Total protein offered (mg)	Initial enzyme content (%)	Total protein in final supernatant (mg)	Total protein immobilised (mg)	Total enzyme immobilised* (mg)	Enzyme content in the immobilised derivative (%)
<i>Sc</i> ADH	2 \pm 0	12 \pm 1	1.8 \pm 0	0.21 \pm 0	0.21 \pm 0	100
<i>Sc</i> PDC	15.4 \pm 1	13 \pm 0	14.2 \pm 0	1.2 \pm 0	0.9 \pm 0	75
<i>Tm</i> LDH	8.3 \pm 0	12 \pm 1	7.5 \pm 0	0.8 \pm 0	0.8 \pm 0	100

* Calculated from SDS-PAGE differences between Blank and final supernatant samples.

Results showed that high retained activities were achieved in all cases, but not all the recovered derivatives maintained all the activity (Table 6.4). On the one hand, the immobilised derivatives of *Sc*ADH enzyme

kept a 92.4% of the activity after the recovery process, and a 75% of activity recovery was achieved for CBM9-*Tm*LDH. However, *Sc*PDC enzyme derivatives only conserved a 55% of total activity offered, even if suspension samples kept the totality of activity.

Protein analysis strengthened what was measured via enzyme activity test; **CBM9-*Sc*ADH** presence in lysate decreased from 12.2% to 1.1% of relative content, according to protein electrophoresis (Figure 6.8), and 0.21 mg of target protein from 0.24 mg offered were calculated to have bound to RAC (Table 6.5), meaning that almost a 90% of target enzyme was recovered in immobilised derivative. *Sc*PDC analysis revealed that approximately 0.9 mg of target enzyme were immobilised, which corresponded to the 45% of total offered (2 mg). Finally, 0.8 mg of **CBM9-*Tm*LDH** were bound to support (Table 6.5), which corresponded to the 80% of total enzyme offered (1 mg·mL⁻¹), matching also with the relative intensity differences observed between blank and final supernatant samples, measured via electrophoresis (Figure 6.10).

In all cases the high specificity of CBM9–tagged enzymes towards RAC was proved, increasing the relative enzyme content from 12–13% in initial lysates to 100, 75 and 100% in the recovered derivatives, for *Sc*ADH, *Sc*PDC and *Tm*LDH, respectively.

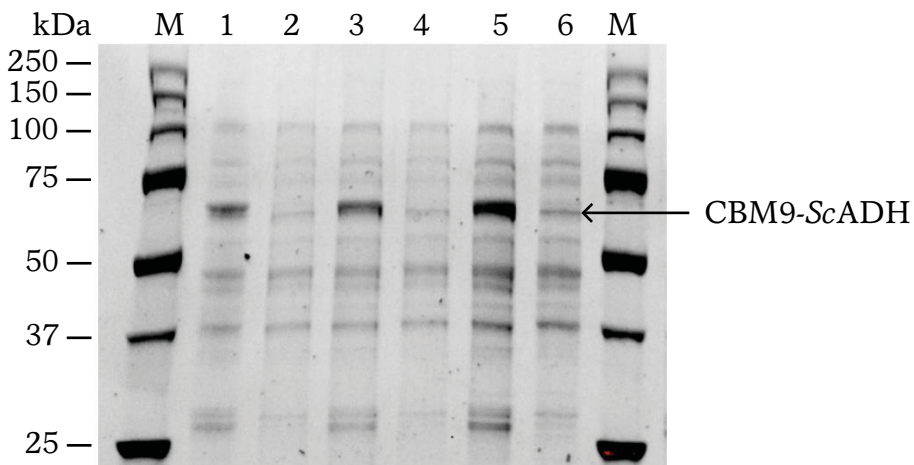


Figure 6.8. SDS–PAGE of the immobilisation experiments of CBM9-*Sc*ADH fusion protein onto cellulosic RAC support. Lanes M: molecular weight standard (kDa); lanes 1, 3 and 5: blank samples of each replicate; lanes 2, 4 and 6: final supernatant samples of each replicate. CBM9-*Sc*ADH (63 kDa) corresponding band indicated with arrow.

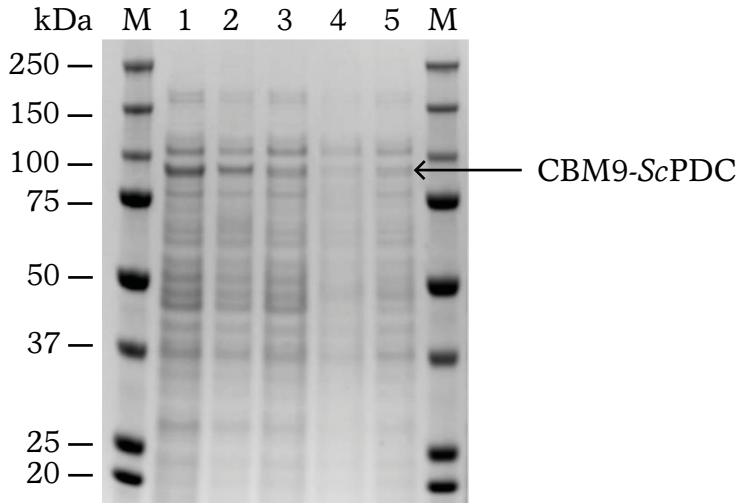


Figure 6.9. SDS-PAGE of the immobilisation experiments of CBM9-ScPDC fusion protein onto cellulosic RAC support. Lanes M: molecular weight standard (kDa); lane 1: clarified cell extract; lanes 2 and 3: blank samples of two replicates; lanes 4 and 5: final supernatant samples of two replicates. CBM9-ScPDC (88 kDa) corresponding band indicated with arrow.

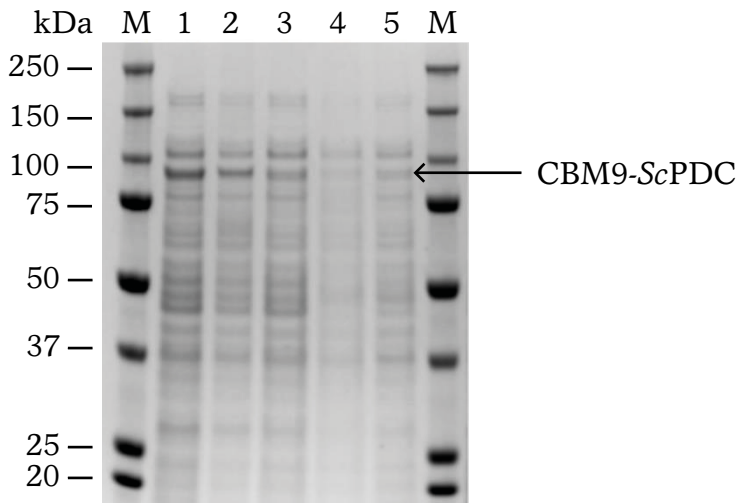
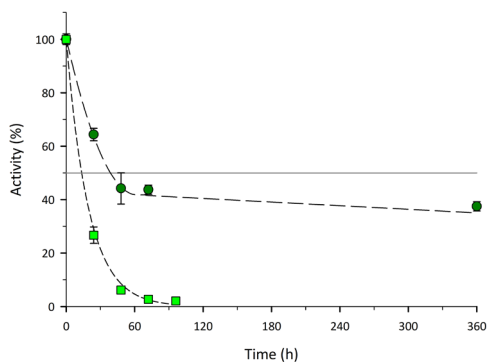


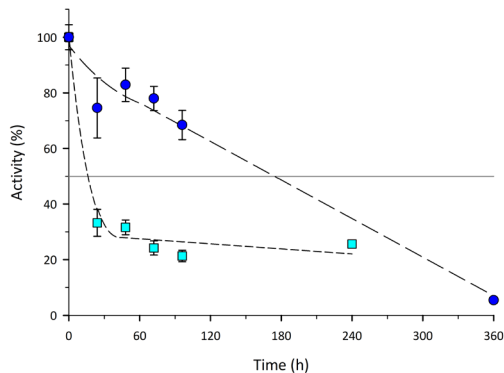
Figure 6.10. SDS-PAGE of the immobilisation experiments of CBM9-TmLDH fusion protein onto cellulosic RAC support. Lane 1: clarified cell extract; lanes 2, 3 and 4: final supernatant samples of each replicate. CBM9-TmLDH (55 kDa) corresponding band indicated with arrow.

The **enzyme storage stability assessment** of both soluble enzymes and immobilised derivatives is depicted in **Figure 6.11**:

A: CBM9-*ScADH*



B: CBM9-*ScPDC*



C: CBM9-*TmLDH*

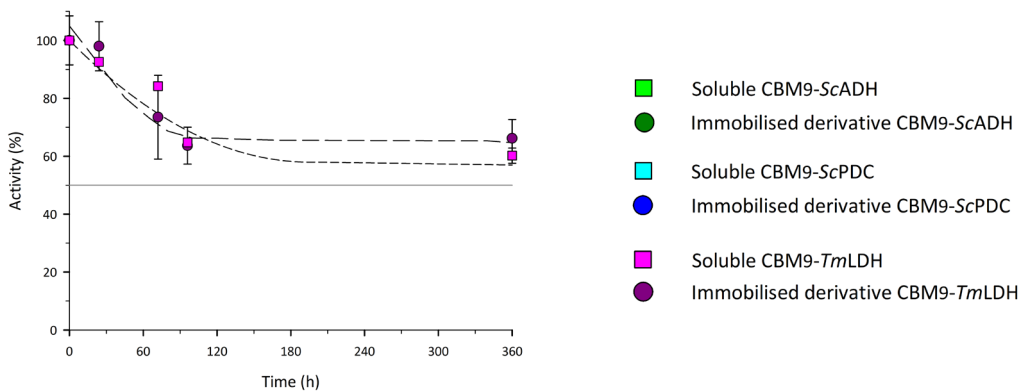


Figure 6.11. Stability of CBM9-*ScADH* (A), CBM9-*ScPDC* (B) and CBM9-*TmLDH* (C) fusion proteins, soluble (□) and immobilised onto RAC support (○), under storage conditions (50 mM Tris-HCl buffer pH 7.50, 4°C). Error bars correspond to standard error of three replicates.

Results showed that the immobilisation process allowed the enhancement of enzyme stability in all cases. On the one hand, *ScADH* and *ScPDC* soluble forms were quickly inactivated, being their half-life times of 14 and 16 hours, respectively. Nevertheless, the stability increase achieved with immobilisation was different among the two of them, leading to a stability improvement of 2.9 folds for **CBM9-*ScADH*** (Figure 6.11 A), and of 11 folds for ***ScPDC*** enzyme (Figure 6.11 B). *ScADH* derivative, which seemed to be inactivated faster than *ScPDC* one, finally retained a 40%

of initial activity after 15 days, whereas the second one, which retained a 70% of activity after 96 hours, ended up completely inactivated.

On the other hand, *Tm*LDH soluble enzyme activity slowly decreased and was finally stabilised at around a 60% of initial activity at the end of the experiment (Figure 6.11 C). In this case, enzyme stability of immobilised derivative was similar to soluble form, even if the final activity measured was slightly higher than soluble (66%). Once again, the obtained results consolidate CBM domains as feasible one-step purification/immobilisation tags thanks to the improvement of enzyme stability once the fusion peptides are bound to the support.

The **maximum enzyme loading capacity** of CBM9–fused enzymes on RAC support was also assessed (Table 6.6). However, the maximum loading capacity on MC support is not compiled in this work, since it has already been reported that CBM9 domains—and more precisely the same CBM9 domain used in this work— exhibit greater affinity towards amorphous cellulose rather than crystalline (Morag et al., 1995). The maximum loading values obtained—inside the range of activities tested—are presented in bold in Table 6.6.

Results obtained revealed that 1 mL of RAC support was able to bind up to 8300 AU of CBM9-ScADH enzyme from cell lysate, which would correspond to approximately to 115 mg of target protein. Additionally, 310 UA·mL⁻¹ of CBM9-ScPDC and 151 AU·mL⁻¹ of CBM9-*Tm*LDH fusion proteins, corresponding both to 10 mg, could also be bound to RAC. As happened with CBM3–tagged variants, multi-layer immobilisation could have occurred during the experiments, since it was observed an increase in maximum loading capacity between consecutive experiments that showed an IY lower than 90%. This phenomenon can be clearly seen within the range of AU tested for ScADH enzyme and, additionally, results obtained for CBM9-ScPDC also goes in line with this hypothesis.

Table 6.6. Results obtained of maximum enzyme loading capacity of CBM9–fused enzymes onto RAC support. Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, 1 mL RAC support, roller agitation. ± correspond to the standard error.

Enzyme	Enzyme offered (AU·mL ⁻¹ support)	IY (%)	Enzyme immobilised (AU·mL ⁻¹ support)	Theoretical RA (AU·mL ⁻¹ support)	Measured RA (AU·mL ⁻¹ support)
	16 ± 3	100 ± 0	16 ± 0	16 ± 0	15 ± 1
	54 ± 0	100 ± 0	54 ± 0	53 ± 1	53 ± 0
	3797 ± 142	99 ± 0	3751 ± 7	3667 ± 64	1656 ± 9
ScADH (RA: 98 %)	5907 ± 101	100 ± 1	5885 ± 1	5750 ± 100	3738 ± 45
	8017 ± 316	97 ± 0	7790 ± 8	7613 ± 136	4155 ± 103
	10062 ± 84	74 ± 0	7494 ± 652	7326 ± 171	4151 ± 649
	11112 ± 189	63 ± 0	7028 ± 34	6871 ± 117	2585 ± 93
	16051 ± 273	53 ± 1	8472 ± 136	8282 ± 141	3554 ± 440
	154 ± 4	92 ± 1	142 ± 1	120 ± 3	88 ± 1
ScPDC (RA: 85 %)	363 ± 10	87 ± 0	317 ± 1	267 ± 7	282 ± 6
	501 ± 15	73 ± 1	368 ± 6	310 ± 8	204 ± 56
	43 ± 1	94 ± 2	40 ± 1	32 ± 2	31 ± 2
TmLDH (RA: 79 %)	246 ± 7	72 ± 2	177 ± 5	140 ± 7	42 ± 4
	346 ± 8	55 ± 1	190 ± 8	151 ± 8	60 ± 14

IY: immobilisation yield; RA: retained activity. Theoretical RA coefficients (1st column) were determined under no diffusional-limiting conditions.

6.2 Comparison between CBM3 and CBM9 domains as immobilisation tags

The results obtained with the immobilisation experiments validate the strategy of using CBM-tagged enzymes as a promising system for one step purification/immobilisation process thanks to i) the high specificity of CBM domains towards cellulosic supports compared to the other proteins present in *E. coli* lysates and ii) the high retained activities obtained in the final immobilised derivatives. However, each domain has proved to be more suitable for a different sort of cellulose; CBM3-fused enzymes proved to bind more effectively towards microcrystalline cellulose, whereas CBM9-tagged enzymes bound with greater affinity to amorphous cellulose.

Below, it is presented a brief comparison between the two CBM domains used in this work as immobilisation tags. Maximum enzyme loadings of CBM-fused enzymes obtained in this study are shown in [Figure 6.12](#):

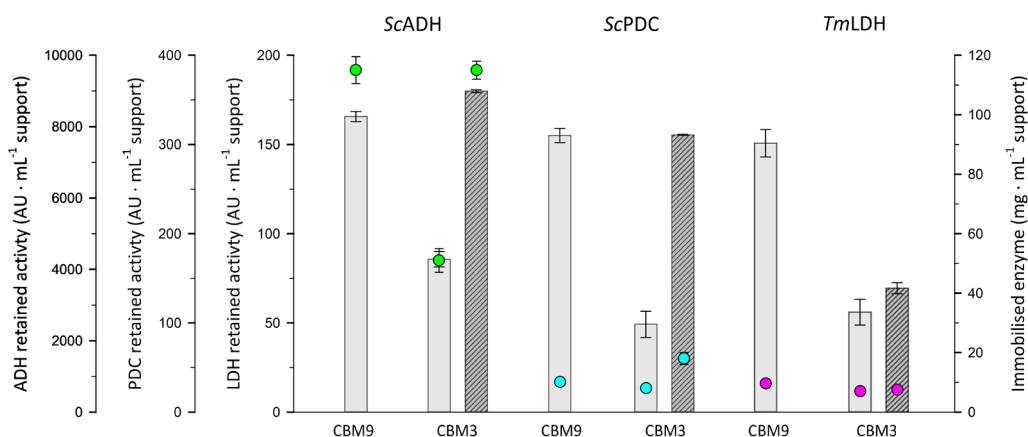


Figure 6.12. Maximum enzyme loading capacity (AU·mL⁻¹) of CBM-fused enzymes onto amorphous (RAC) and microcrystalline (MC) cellulose supports (bars), and their corresponding mass values (mg·mL⁻¹) (○). Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, 1 mL RAC support, roller agitation. Error bars correspond to the standard error.

Results depicted in [Figure 6.12](#) makes graphically evident what was already stated above; maximum enzyme loads vary significantly depending on the cellulose support used and the CBM domain fused to target enzymes. In all cases best results were obtained by using RAC with CBM9 domain and using MC with CBM3-fused enzymes. Thus, in terms of

AU, similar loadings were achieved with CBM9 tag onto RAC and with CBM3 domain onto MC support, except for *Tm*LDH enzyme, with which a 2-fold increase in maximum loading was observed for the immobilisation of the CBM9–fused version onto RAC. Nevertheless, in terms of milligrams, minor differences were shown to bound to each of the supports tested. Regarding *Sc*PDC enzyme, a 1.8-fold increase in total amount of protein immobilised was calculated for the immobilisation of CBM3-*Sc*PDC enzyme onto MC, in contrast to CBM9-*Sc*PDC onto RAC.

In conclusion, both CBM–tagged variants were proved to be suitable for one-step purification/immobilisation process, retaining in almost all cases a high percentage of the activity offered, and improving in most of them the enzyme storage stability.

6.3. Use of Carbohydrate–Binding Modules as purification tags

As previously mentioned, CBM domains can be used as an alternative of traditional metal–chelated resins used in protein purification processes. Their usage would enable us to recover highly purified soluble target proteins and overcome the hindrances of IMAC methods such as high-cost resins and metal ions presence in recovered product. The feasibility of CBM–fused enzymes purification by cellulose affinity chromatography would also bring to light the versatility of the cellulose as a support for both i) one-step immobilisation/purification and ii) enzyme purification.

For that purpose, it was decided to set up a fast protein liquid chromatography (FPLC) method to recover CBM–fused enzymes in soluble form after removing the undesired proteins from cell lysates, based on the reversibility of the bound between cellulose and the protein. However, it was necessary to perform a preliminary screening study prior to FPLC experiments, aiming to determine which compound would be the most suitable eluent. In this context, a wide range of compounds have been reported to elute CBM domains from cellulosic supports; CBM3 domain has been successfully eluted with ethylene glycol (Wan et al., 2011, Myung et al., 2011), EDTA (Hong et al., 2008) or trimethylamine (Morag et al., 1995), whereas CBM9 domain has been eluted with glucose (Boraston et al., 2001, Kavoosi et al., 2004). Nonetheless, some of these compounds could not be used in our case of study since they would significantly compromise enzyme activity due to their denaturing effect.

Prior to compare the efficiency of purification process depending on which CBM is fused to target enzyme, brief stability and elution screenings were performed with some of the compounds mentioned above; some polyalcohols, whose hydroxyl (OH) groups could be useful to unbound CBM domains from cellulose (see Section 1.5, Carbohydrate-Binding Modules), and glucose, which has already been used as eluent in previous reports (Boraston et al., 2001).

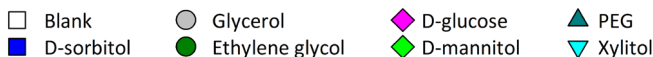
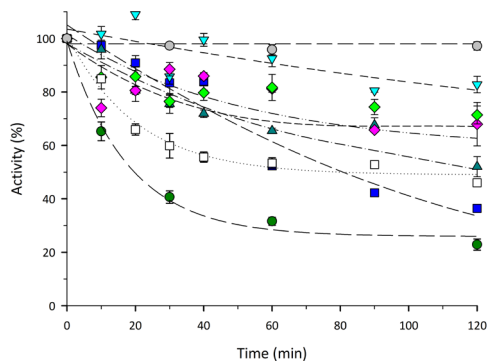
Related to the support used in FPLC purifications, RAC obtained from Avicel[®] PH-101—which has a particle size of approximately 50 μm —was unviable for FPLC performance, because the cellulose bread ended up compacting and column flowthrough collapsed. For that reason, the same cellulose support but with higher particle size (Avicel[®] PH-200, ~ 180 μm , see Section 3.9, Fast Liquid Protein Chromatography) was employed. For that, subsequent experiments were performed with RAC derived from Avicel[®] PH-200 cellulose.

6.3.1. Eluent screening

As mentioned, first experiments were focused on the screening of several compounds to be used as eluent before starting with FPLC processes. Therefore, prior to assess the eluting capability of each of the compounds picked, an enzyme stability study was carried out. The two CBM-fused versions of *ScADH* enzyme were incubated at room temperature (24°C) with different solutions that contained the selected compounds at 75% of their maximum solubility in water (see Section 3.9, Fast Liquid Protein Chromatography). Though, the screening was only performed with *ScADH* because of two reasons: i) it proved to be the less stable enzyme (under the tested conditions) and it was probable that any of the compounds could have more impact on its stability rather than on the other two and ii) since protein elution is dependent on the interaction between the support and the CBM, it was only required to evaluate the desorption of each CBM domain from RAC support once and thus, iii) *ScADH* has been used as model enzyme along this work.

The evolution of ADH activity of the two CBM-fused variants along the time experiment is shown in [Figure 6.13](#):

A: CBM3-ScADH



B: CBM9-ScADH

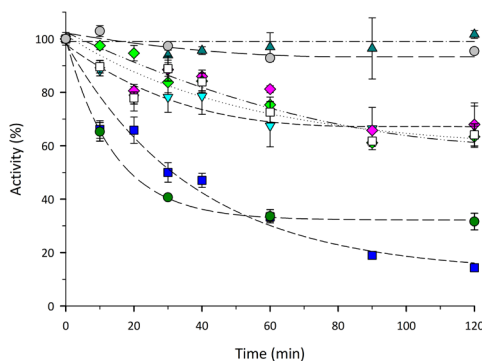


Figure 6.13. Stability of CBM3-ScADH (A) and CBM9-ScADH (B) fusion proteins (in soluble form). Enzymes were incubated in 50 mM Tris-HCl pH 7.50 buffer —blank (□)—, containing 3.6 M D-sorbitol (■); 10.2 M glycerol (●); 13.4 M ethylene glycol (●); 1 M D-glucose (◆); 0.9 M D-mannitol (◆); 1 M Polyethylene glycol (PEG) (▲) and 3.2 M Xylitol (▼) solutions. Experiments conditions: 24°C, roller agitation. Error bars correspond to standard error of three replicates.

As can be seen in Figure 6.13 A, stability of CBM3-ScADH was significantly affected by the presence of ethylene glycol (EG), and samples incubated with sorbitol also showed low stability after long incubation periods, decreasing up to the 36% of the initial value. On the contrary, enzyme stability was improved for the rest of compounds, in comparison with the control sample, especially for samples incubated with glycerol.

On the other hand, CBM9-ScADH stability was improved in samples incubated with glycerol and with polyethylene glycol (PEG), retaining almost the totality of initial activity after 2 h. However, enzyme activity was clearly affected not only by the presence of EG but also sorbitol, whilst the rest of compounds tested showed similar stability than control samples.

Considering the results obtained in stability tests, sorbitol and ethylene glycol (EG) were discarded, and the rest of the compounds tested were used to carry out the desorption batch experiment, as described in Section 3.8, Biocatalysts immobilisation). However, it should be mentioned that EG was the only compound found in literature that has been efficiently used to desorb CBM3 domains from cellulose (Wan et al., 2011,

Myung et al., 2011), to the best knowledge of the author, apart from denaturing compounds. Nevertheless, considering the results obtained in stability tests (Figure 6.13), EG could not be considered to recover CBM3–fused ScADH enzyme, since the objective of the present study is not only to purify the enzymes but to obtain them in their active form, as already mentioned.

Considering this, desorption studies were carried out with the CBM9–tagged variant, and the suitability of the subsequent developed purification process for CBM9-ScADH was evaluated afterwards to purify CBM3–tagged enzymes. Therefore, a first elution study was performed with CBM9-ScADH fusion protein and the results obtained are shown in Table 6.7, in which activity values of initial derivatives are listed, as well as the activity measured in the supernatants and the remaining activity immobilised onto the support after the experiment.

Table 6.7. Enzyme recovery parameters obtained for CBM9-ScADH enzyme elution from RAC support by using different compounds as eluent. Experiment conditions: 50 mM Tris-HCl pH 7.50 buffer, 24°C, roller agitation. \pm correspond to the standard error of three replicates.

Eluent	Immobilised derivative activity		Final supernatant activity		Support activity		Total balance (%)
	(AU)	(%)	(AU)	(%)	(AU)	(%)	
Glycerol	302 \pm 6	100 \pm 2	289 \pm 9	96 \pm 3	39 \pm 1	13 \pm 0	109
D-glucose	197 \pm 12	100 \pm 6	165 \pm 2	84 \pm 1	6 \pm 0	3 \pm 0	87
PEG	81 \pm 6	100 \pm 8	1 \pm 0	1 \pm 0	78 \pm 4	96 \pm 5	98
Xylitol	378 \pm 3	100 \pm 1	22 \pm 1	6 \pm 0	251 \pm 28	66 \pm 7	72
D-mannitol	243 \pm 1	100 \pm 0	27 \pm 0	11 \pm 0	148 \pm 3	61 \pm 1	72

Regarding the percentage of activity recovered in the supernatant after one hour of incubation, only glycerol and glucose proved to be suitable candidates for protein desorption from RAC. However, the high viscosity of the elution buffer with minimal concentration of glycerol needed to desorb CBM9–fused enzyme from RAC support (75% v/v) was incompatible with the FPLC system performance and, when glycerol was diluted up to 50% to reduce the viscosity, it was only achieved a 11% of enzyme recovery.

Therefore, glycerol was discarded for the following experiments, being glucose the sole compound that was tested as eluent in FPLC experiments, which turned out to be satisfactory results, since glucose is a relatively cheap and highly available chemical compound, and it would be easy to remove from final product. Thus, it was decided to develop a FPLC process for CBM9–tagged enzyme purification by using glucose as eluent compound.

6.3.2. Purification of CBM9–fused enzymes

Once it was determined that glucose would be used as eluent in FPLC purification processes, first experiments were carried out with the same protein used in the previous screening. CBM9-ScADH affinity purification process was performed for triplicate, and the resulting chromatogram and the corresponding SDS–PAGE gel documentation is shown in Figures 6.14 and 6.15, respectively, whereas the purification metrics are provided in Table 6.8.

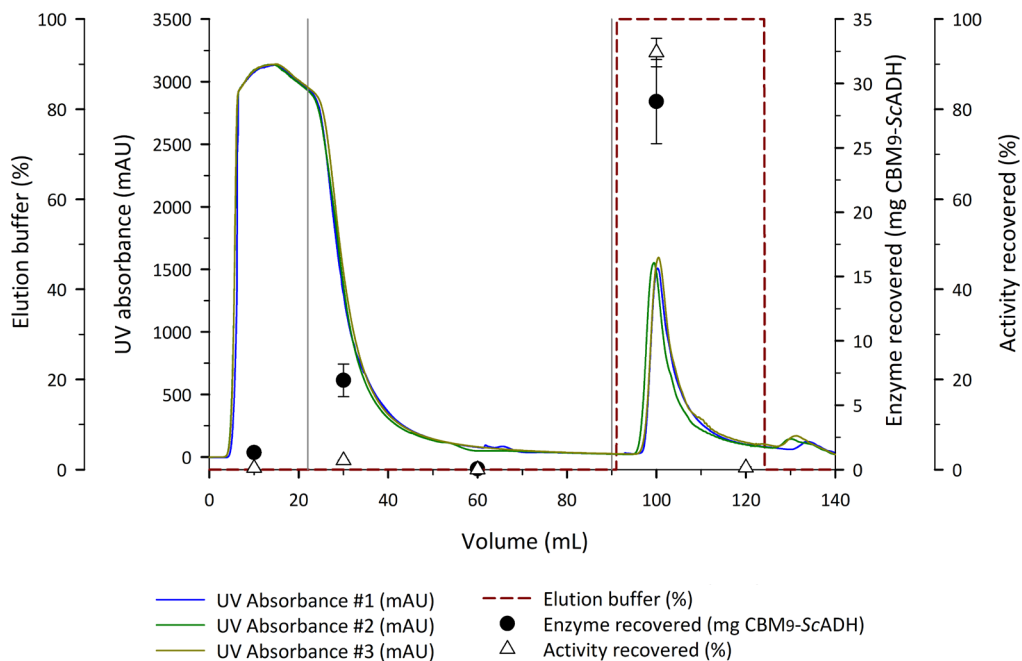


Figure 6.14. Chromatogram of three consecutive CBM9–ScADH FPLC purifications on pre-treated Avicel® PH-200 support (RAC). Enzyme recovered in each fraction (●) and the corresponding relative activity (Δ) plotted. Experiment conditions: 15 mL column volume, 10 mL sample volume, 24°C, pH 7.50. Loading buffer: 50 mM Tris-HCl; wash: 200 mM NaCl in 50 mM Tris-HCl; elution: 2 M glucose in 50 mM Tris-HCl.

FPLC purification resulted in a $93 \pm 3\%$ recovery of activity in average. A trivial *ScADH* fraction within the clarified lysate load was lost in the column flow-through (2.4%) and no activity was measured at any of the column washes. As expected, glucose solution was effective in desorbing all specifically bound target enzyme, which eluted from the column in a single and clear peak (Figure 6.14). In addition, it was proved that RAC support can be reused in consecutive purification batches, since process efficiency and product recovery did not vary significantly among the three experiments performed.

Table 6.8. Results obtained for FPLC-based CBM9-*ScADH* purification processes. \pm correspond to the standard error of three replicates.

Fraction	Total activity		Total protein		Specific activity (AU·mg ⁻¹ protein)	Purification factor
	(AU)	(%)	(mg)	(%)		
Lysate	4117 ± 99	100 ± 5	381 ± 9	100 ± 2	10.8 ± 0	-
Flow-through	98 ± 4	2 ± 0	341 ± 7	90 ± 2	0.3 ± 0.1	-
Wash	0.4 ± 0	0 ± 0	2.1 ± 0	0 ± 0	0.2 ± 0	-
Elution	3813 ± 129	93 ± 3	36 ± 2	9 ± 0	106.5 ± 9	10 ± 1

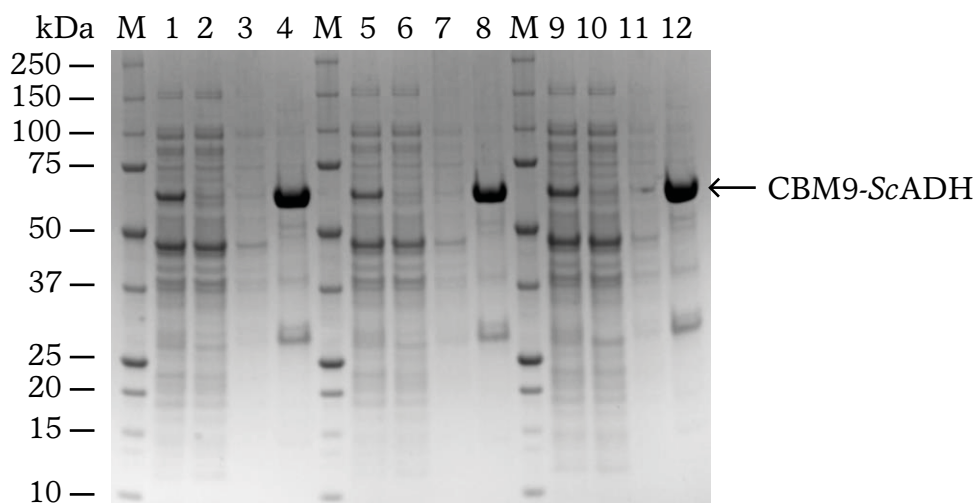


Figure 6.15. SDS-PAGE of the FPLC purifications of CBM9-*ScADH* fusion protein onto cellulosic RAC support. Lanes M: molecular weight standard (kDa); lanes 1, 5 and 9: clarified cell extract prior to column loading; lanes 2, 6 and 10: column flow through; lanes 3, 7 and 11: column wash; lanes 4, 8 and 12: purified CBM9-*ScADH* eluted with 50 mM Tris-HCl buffer containing 2M glucose. CBM9-*ScADH* (64 kDa) corresponding band indicated with arrow.

In that case, both mass and activity balances were correct, since the sum of all fractions corresponds to 379 mg of total protein (99.5% of initial offered) and to 3930 AU, that means a 95.4% of total (Table 6.8). Regarding the target enzyme, it has been estimated that 37 mg were recovered among all the fractions —29 mg in elution— from the 47 mg that were offered (78.5%). According to SDS-PAGE, CBM9-*ScADH* purity in elution fractions is about 80% (Figure 6.15, lanes 4, 8 and 12).

Once it was checked that glucose was suitable for CBM9-tagged protein elution in FPLC —it had already been demonstrated in batch experiments during eluent screening actually—, same procedure was followed for the rest of the enzymes.

CBM9-*ScPDC* purification experiment was carried out by duplicated (Figure 6.16) to prove that the method was suitable for a wide range of CBM9-fused enzymes, apart from *ScADH*. Purification metrics are given in Table 6.9, and the corresponding SDS-PAGE gel is shown in Figure 6.17.

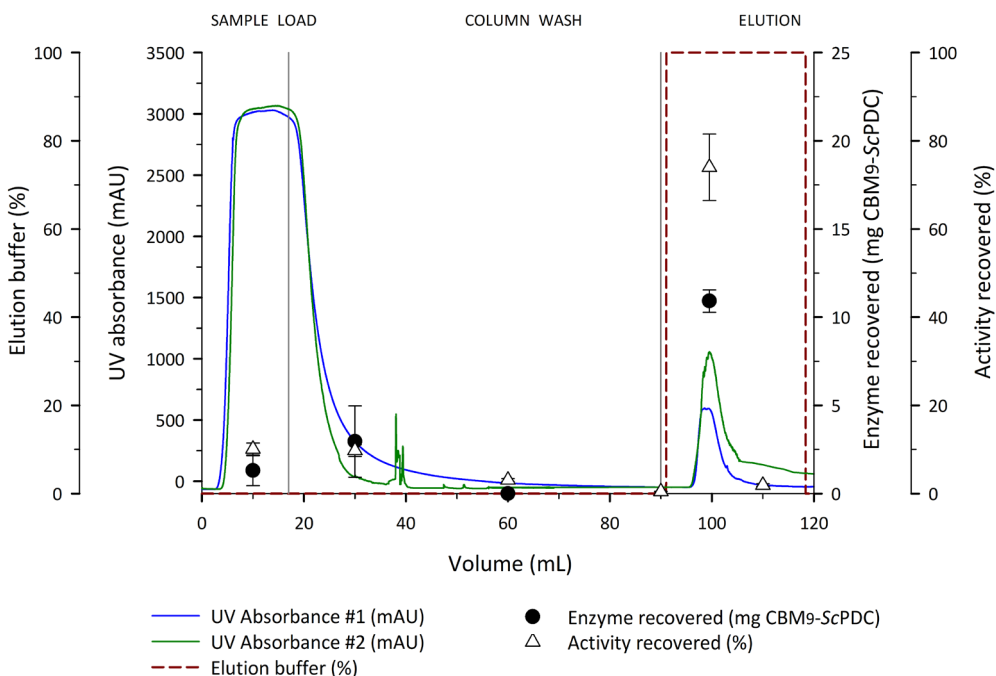


Figure 6.16. Chromatogram of two consecutive CBM9-*ScPDC* FPLC purifications on pre-treated Avicel® PH-200 support (RAC). Enzyme recovered in each fraction (●) and the corresponding relative activity (Δ) plotted. Experiment conditions: 15 mL column volume, 10 mL sample volume, 24°C, pH 7.50. Loading buffer: 50 mM Tris-HCl; wash: 200 mM NaCl in 50 mM Tris-HCl; elution: 2 M glucose in 50 mM Tris-HCl.

Table 6.9. Results obtained for FPLC-based CBM9-ScPDC purification processes. \pm correspond to the standard error of two replicates.

Fraction	Total activity		Total protein		Specific activity (AU·mg ⁻¹ protein)	Purification factor
	(AU)	(%)	(mg)	(%)		
Lysate	108 ± 6	100 ± 5	118 ± 2	100 ± 1	0.9 ± 0	-
Flow-through	21 ± 3	20 ± 3	97.4 ± 2	82 ± 2	0.3 ± 0	-
Wash	3.3 ± 0	3 ± 0	1.1 ± 0	1 ± 0	2.9 ± 0	-
Elution	80 ± 8	74 ± 7	12.5 ± 0	11 ± 0	6.4 ± 1	7 ± 0

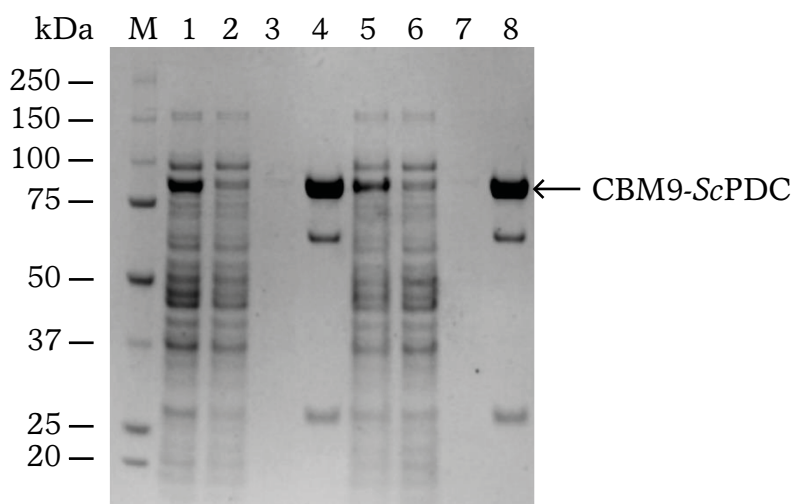


Figure 6.17. SDS-PAGE of the FPLC purifications of CBM9-ScPDC fusion protein onto cellulosic RAC support. Lane M: molecular weight standard (kDa); lanes 1 and 5: clarified cell extract prior to column loading; lanes 2 and 6: column flow through; lanes 3 and 7: column wash; lanes 4 and 8: purified CBM9-ScPDC eluted with 50 mM Tris-HCl buffer containing 2M glucose. CBM9-ScPDC (83 kDa) corresponding band indicated with arrow.

CBM9-ScPDC purification was satisfactorily performed, achieving a $74 \pm 7\%$ recovery, corresponding to 80 ± 8 AU from 108 ± 6 AU offered (Table 6.9). The rest of the activity was detected mainly in flowthrough fraction (21 ± 3 AU), meaning that this percentage of target protein was not correctly immobilised, even if maximum load capacity of RAC support was not exceeded. Taking this into account, it can be stated that almost all activity immobilised to the column could be recovered afterwards in

elution step, and the sum of the activity measured in all fractions (106.8 AU) coincided with total activity offered (Table 6.9).

Regarding mass balances, total protein loaded onto the column also matched with the sum of the fractions —118 and 111.7 mg, respectively — (Table 6.9). However, 11 mg of target protein were calculated to be present in elution fraction, corresponding to 44% of total offered (25 mg), being this percentage notably lower than measured via activity test, possibly due to measurement errors during protein analysis procedures. CBM9-ScPDC enzyme was recovered with an 85% purity grade according to SDS-PAGE (Figure 6.17).

Last affinity purification processes were performed by using the **CBM9-fused *Tm*LDH** enzyme, which was the last one to be tested. 10 mL of cell lysate containing 491 ± 19 AU (180 ± 23 mg) of CBM9-*Tm*LDH fusion protein were loaded onto the FPLC column (Figure 6.18), yielding an $86 \pm 7\%$ of activity recovery (Table 6.10), with a purity grade of 76% (Figure 6.19).

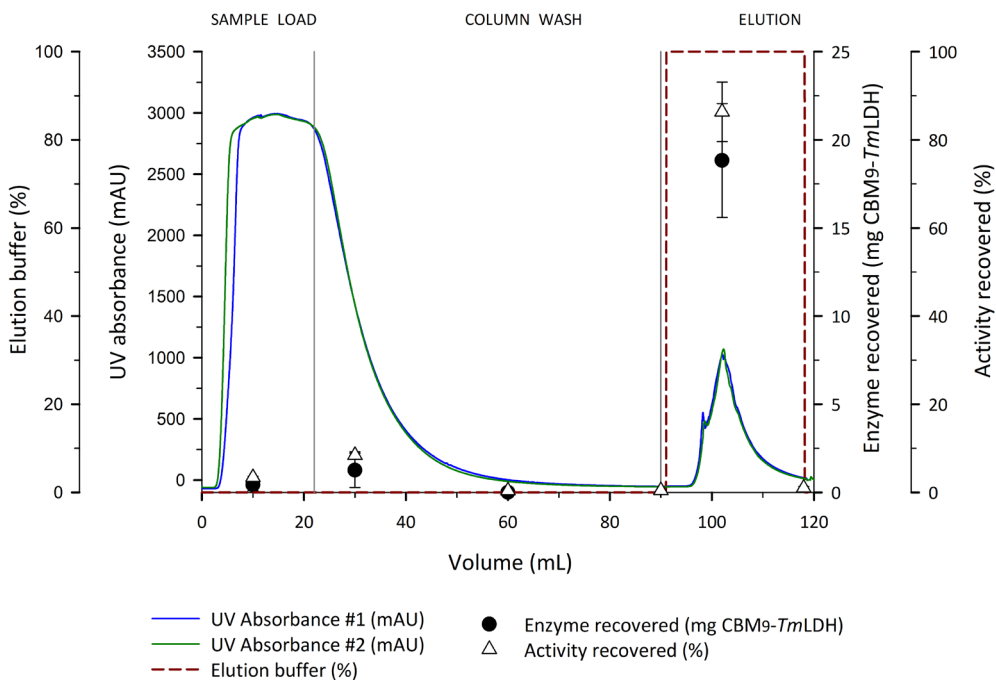


Figure 6.18. Chromatogram of two consecutive CBM9-*Tm*LDH FPLC purifications on pre-treated Avicel® PH-200 support (RAC). Enzyme recovered in each fraction (●) and the corresponding relative activity (Δ) plotted. Experiment conditions: 15 mL column volume, 10 mL sample volume, 24°C, pH 7.50. Loading buffer: 50 mM Tris-HCl; wash: 200 mM NaCl in 50 mM Tris-HCl; elution: 2 M glucose in 50 mM Tris-HCl.

Table 6.10. Results obtained for FPLC-based CBM9-*Tm*LDH purification processes. \pm correspond to the standard error of two replicates.

Fraction	Total activity		Total protein		Specific activity (AU·mg ⁻¹ protein)	Purification factor
	(AU)	(%)	(mg)	(%)		
Lysate	491 ± 19	100 ± 4	180 ± 23	100 ± 13	2.7 ± 0	-
Flow-through	58 ± 3	12 ± 0	168 ± 5	94 ± 3	0.4 ± 0	-
Wash	2 ± 1	0.4 ± 0	1.2 ± 0	1 ± 0	1.8 ± 1	-
Elution	424 ± 33	86 ± 7	24 ± 4	13 ± 2	18 ± 0	6.6 ± 1

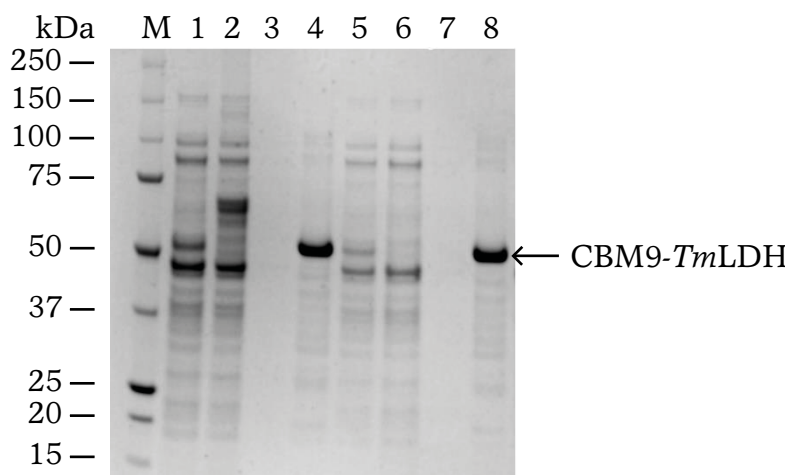


Figure 6.19. SDS-PAGE of the FPLC purifications of CBM9-*Tm*LDH fusion protein onto cellulosic RAC support. Lane M: molecular weight standard (kDa); lanes 1 and 5: clarified cell extract prior to column loading; lanes 2 and 6: column flow through; lanes 3 and 7: column wash; lanes 4 and 8: purified CBM9-*Tm*LDH eluted with 50 mM Tris-HCl buffer containing 2M glucose. CBM9-*Tm*LDH (51 kDa) corresponding band indicated with arrow.

On the one hand, activity balance was correct, since total activity units of cell lysate coincided with the sum of all fractions (490.8 and 490.7 AU, respectively). Elution fraction contained 423.7 AU, that corresponded to 86.4% of total activity loaded (Table 6.10). On the other hand, protein balance was also satisfactory; 195 mg were estimated to be recovered, whereas 180 mg were calculated to be offered to the column (Table 6.10). Finally, concerning the target enzyme, 20.5 mg were recovered, which represented the 81% of total estimated (25.3 mg), being this relative percentage similar enough to the obtained via activity analysis.

Even if satisfactory results were obtained for the purification of the three CBM9–tagged enzymes, there were detected some impurities in elution samples, especially for *ScADH* and *ScPDC* purification processes (Figures 6.15 and 6.17, respectively), which could be broken fractions of the fusion proteins that contained the CBM9 domain and remained attached to the support until the elution step, although it was not determined to what corresponded. E.g., in Figure 6.17 there were detected two more bands in elution samples, with a relative intensity of 10% approximately. The sizes of those bands are similar to molecular weights of both CBM9 domain and *ScPDC* enzyme (26 and 62 kDa in gel, respect to 21 and 61 kDa in theory), what could be showing that a fraction of the fusion protein was split in two parts. This fact would also explain the unbalance mentioned above and coincides with what was observed for CBM9–fused *ScADH* enzyme (Figure 6.15). However, all purification processes were performed by adding protease inhibitor (PMSF) to cell extract to precisely prevent the breaking of the fusion protein.

6.3.3. Purification of CBM3–fused enzymes

Considering the high yields achieved in FPLC purification processes for CBM9–fused enzymes, and the satisfactory performance of RAC support used as column bead, same purification procedure was tested to purify CBM3–tagged enzymes using glucose as eluent. The studies were carried out not only with RAC but also with MC, given the positive results obtained during immobilisation experiments of CBM3–tagged enzymes. First experiments were carried out by loading 10 mL of a clarified cell lysate containing 5850 ± 316 AU of *CBM3-ScADH* enzyme, corresponding to 440 ± 23 of total protein (Figure 6.20, Table 6.11). The experiment was performed by duplicate and similar results were obtained in both processes; a 6% of total loaded activity was recovered in elution fraction (350 ± 1 AU) and neither the flow-through nor the wash fractions presented any enzyme activity (Table 6.11), disregarding then a loss of target protein in previous fractions. Elution step resulted in the appearance of a peak in the chromatogram, but notably smaller than expected, meaning that glucose solution was unable to unbind the majority of CBM3–fused enzyme from cellulose, at least for that concentration and under the tested conditions (Figure 6.20).

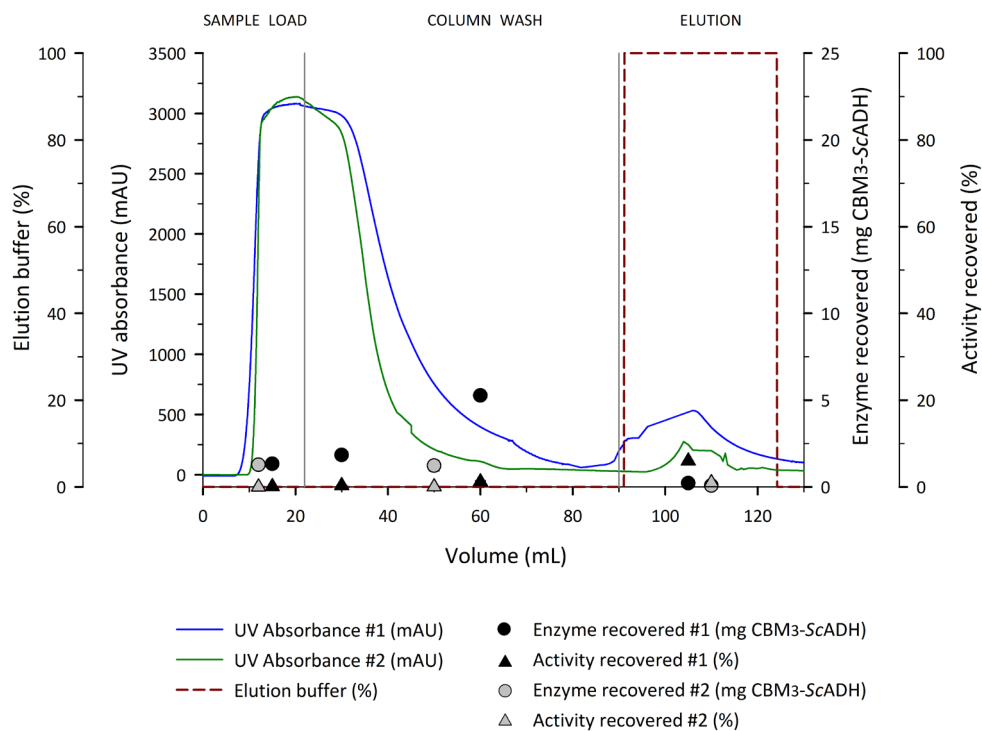


Figure 6.20. Chromatogram of two consecutive CBM3-ScADH FPLC purifications on pre-treated Avicel[®] PH-200 support (RAC). Enzyme recovered in each fraction (\circ) and the corresponding relative activity (Δ) plotted. Experiment conditions: 15 mL column volume, 10 mL sample volume, 24°C, pH 7.50. Loading buffer: 50 mM Tris-HCl; wash: 200 mM NaCl in 50 mM Tris-HCl; elution: 2 M glucose in 50 mM Tris-HCl.

Table 6.11. Results obtained for FPLC-based CBM3-ScADH purification processes. \pm correspond to the standard error of two replicates.

Fraction	Total activity		Total protein		Specific activity (AU·mg ⁻¹ protein)	Purification factor
	(AU)	(%)	(mg)	(%)		
Lysate	5854 ± 316	100 ± 5	440 ± 23	100 ± 5	13.3 ± 0.7	-
Flow-through	4 ± 0.1	0.1 ± 0	83 ± 7	19 ± 2	0.05 ± 0	-
Wash	68 ± 15	1.2 ± 0	94 ± 1	21 ± 0	0.73 ± 0	-
Elution	350 ± 1	6 ± 0	8.8 ± 0	2 ± 0	39.7 ± 0	3 ± 0

Protein analysis confirmed what was observed by chromatography, given that elution fraction only contained the 2% of total protein, when target overexpressed protein usually represents the 15–20% of total protein content. Moreover, when microcrystalline cellulose was used instead of RAC, only a 1% of initial activity (61 of 6404 AU loaded) was recovered, which is in accordance with results obtained from immobilisation process, i.e., the more affinity towards substrate, the stronger bound is established, and the harder to desorb CBM3–fused proteins.

Regarding that protein balance could not be correctly calculated for CBM3-ScADH enzyme, the following experiment was carried out by using CBM3-ScPDC fusion protein, but this time it was added to the procedure a final step that promoted protein denaturalization by loading sodium hydroxide (NaOH) onto the chromatography column (Figure 6.21), aiming to determine if CBM3–fused proteins remained bound to the support after elution step. Thus, protein balance and process parameters are listed in Table 6.12, and the corresponding electrophoresis gel is shown in Figure 6.22.

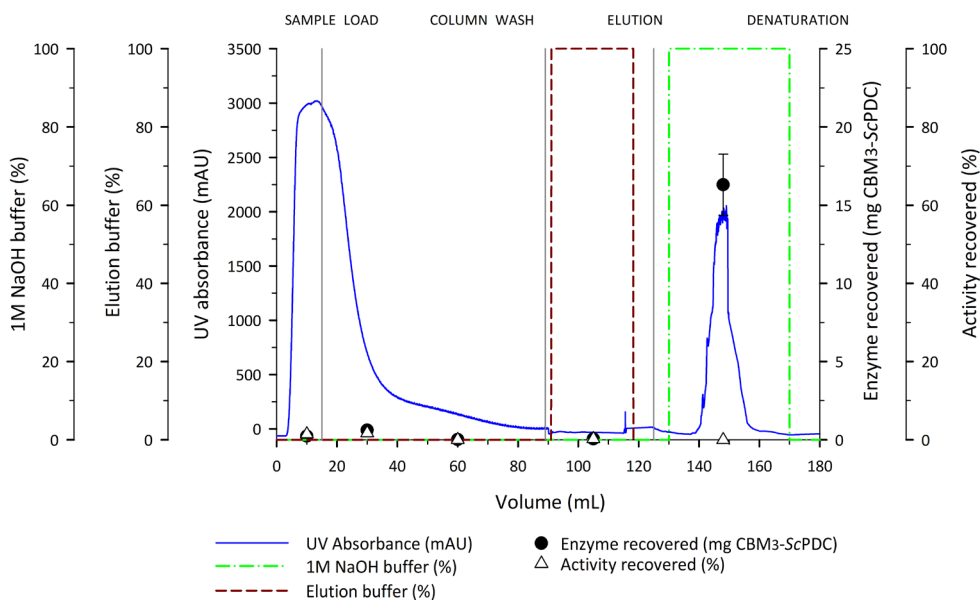


Figure 6.21. Chromatogram of CBM3–ScPDC FPLC purification on pre-treated Avicel® PH-200 support (RAC). Enzyme recovered in each fraction (●) and the corresponding relative activity (Δ) plotted. Experiment conditions: 15 mL column volume, 10 mL sample volume, 24°C, pH 7.50. Loading buffer: 50 mM Tris-HCl; wash: 200 mM NaCl in 50 mM Tris-HCl; elution: 2 M glucose in 50 mM Tris-HCl; denaturation: 1 M NaOH in distilled water. Error bars correspond to the standard error of three measurement replicates.

Once again, no activity was measured in any of the fractions collected and a clear and single peak was obtained when NaOH was added to RAC (Figure 6.21), which corresponded to CBM3-ScPDC fusion protein, as was confirmed by SDS-PAGE (Figure 6.22).

Table 6.12. Results obtained for FPLC-based CBM3-ScPDC purification process. \pm correspond to the standard error of three measurement replicates.

Fraction	Total activity		Total protein		Specific activity (AU·mg ⁻¹ protein)	Purification factor
	(AU)	(%)	(mg)	(%)		
Lysate	2924 ± 100	100 ± 3	120 ± 6	100 ± 5	24 ± 1	-
Flow-through	53 ± 3	2 ± 0	107 ± 2	89 ± 2	0.5 ± 0	-
Wash	2.6 ± 1	0 ± 0	1 ± 0	1 ± 0	2.6 ± 0	-
Elution	1 ± 0	0 ± 0	0 ± 0	0 ± 0	2.64 ± 0.5	0.1 ± 0

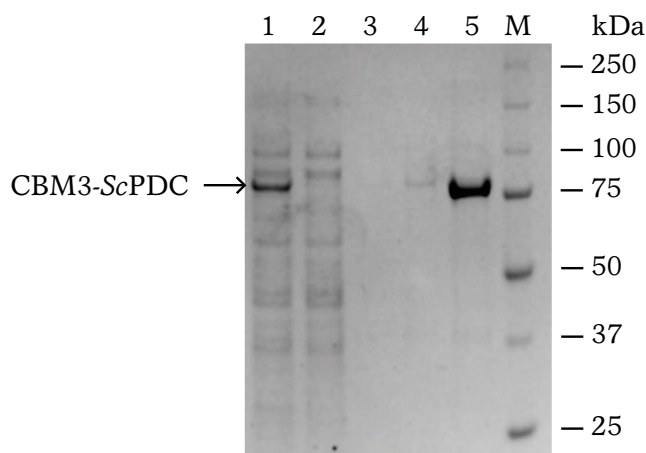


Figure 6.22. SDS-PAGE of the FPLC purification of CBM3-ScPDC fusion protein onto cellulosic RAC support. Lane M: molecular weight standard (kDa); lane 1: clarified cell extract loaded to column; lane 2: fraction corresponding to column flowthrough; lane 3: column wash; lane 4: elution fraction; lane 5: purified and denatured CBM3-ScPDC enzyme desorbed with 1 M NaOH solution. CBM3-ScPDC (79 kDa) corresponding band indicated with arrow.

For this experiment, having presumably recovered all proteins that were loaded onto the column, it made sense to check the protein balance (Table 6.12). It was estimated that 127 mg of total protein were recovered among all the fractions, corresponding to the 106% of initial value (120

mg), which can be considered a coherent estimation. 18.8 mg of total protein were measured in denatured fraction, and it was estimated that 16.3 mg corresponded to target protein, what represented the 68% of total offered (24 mg). Differently from the previous case, no activity was measured in elution fraction, and no peak was neither detected.

Finally, same procedure followed for *ScPDC* enzyme was used for *CBM3-TmLDH* purification (Figure 6.23). In this case, 429 ± 24 AU (135 ± 6 mg) were loaded onto the column (Table 6.13) and as expected, no activity was recovered in elution fraction (0.25% of total), whereas the protein peak appeared when sodium hydroxide was added, as happened with *CBM3-ScPDC* enzyme. It was also detected the presence of target protein in column flowthrough, even if no activity was measured in that fraction.

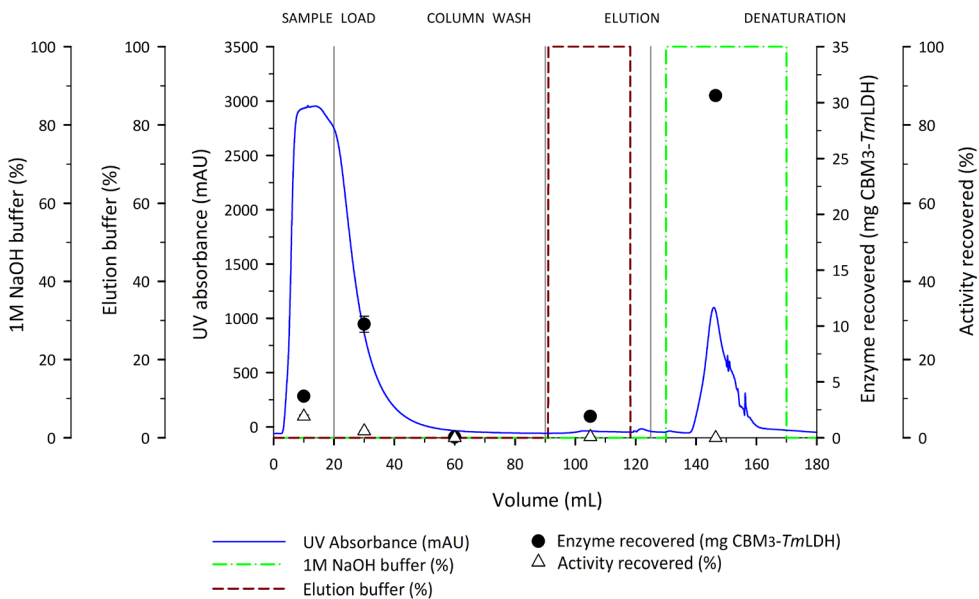


Figure 6.23. Chromatogram of *CBM3-TmLDH* FPLC purification on pre-treated Avicel® PH-200 support (RAC). Enzyme recovered in each fraction (●) and the corresponding relative activity (Δ) plotted. Experiment conditions: 15 mL column volume, 10 mL sample volume, 24°C, pH 7.50. Loading buffer: 50 mM Tris-HCl; wash: 200 mM NaCl in 50 mM Tris-HCl; elution: 2 M glucose in 50 mM Tris-HCl; denaturation: 1 M NaOH in distilled water. Error bars correspond to the standard error of three measurement replicates.

Table 6.13. Results obtained for FPLC-based CBM3-*Tm*LDH purification process. \pm correspond to the standard error of three measurement replicates.

Fraction	Total activity		Total protein		Specific activity (AU·mg ⁻¹ protein)	Purification factor
	(AU)	(%)	(mg)	(%)		
Lysate	429 ± 24	100 ± 5	135 ± 6	100 ± 4	3.2 ± 0	-
Flow-through	31 ± 1	7 ± 0	114 ± 7	84 ± 5	0.4 ± 0	-
Wash	0 ± 0	0 ± 0	1.5 ± 0	1 ± 0	0.1 ± 0	-
Elution	1 ± 0	0.3 ± 0	12 ± 1	9 ± 1	0.1 ± 0	0 ± 0

According to mass balance, almost 31 mg of target protein were desorbed from RAC when NaOH was added, corresponding to the 60% of total CBM3-*Tm*LDH offered (53 mg), and the sum of all fractions corresponded to 86% of total protein loaded. In this case, SDS-PAGE also proved that the protein peak of denaturation fraction corresponded to the target enzyme (Figure 6.24).

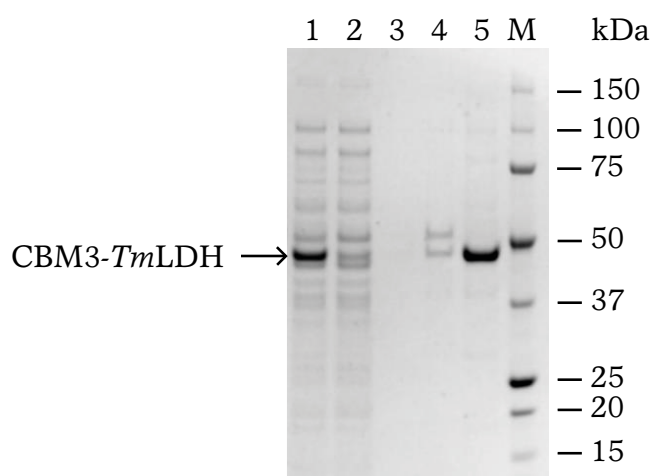


Figure 6.24. SDS-PAGE of the FPLC purification of CBM3-*Tm*LDH fusion protein onto cellulosic RAC support. Lane M: molecular weight standard (kDa); lane 1: clarified cell extract loaded to column; lane 2: fraction corresponding to column flowthrough; lane 3: column wash; lane 4: elution fraction; lane 5: purified and denatured CBM3-*Tm*LDH enzyme desorbed with 1 M NaOH solution. CBM3-*Tm*LDH (47 kDa) corresponding band indicated with arrow.

6.4. Comparison between CBM domains and histidine as purification tags

As mentioned at the early beginning of this chapter, the final goal of all the experiments described in the previous sections was to establish a one-step purification/immobilisation process for the CBM–fused enzymes, whose versatility and reversibility could also be exploited to substitute the use of IMAC purification process with histidine–tagged enzymes, for the reasons provided. Hence, once all the information obtained in these experiments was processed and analysed, there was nothing left to do but compare purification yields between the two methods.

Histidine–based purification experiments were performed as described in Section 3.9, Fast liquid protein chromatography, by using a cross-linked agarose support derivatized with iminodiacetic acid (IDA). Yet, aiming to be as much concise and brief as possible, only final yields will be provided below (Table 6.14).

Table 6.14. Process parameters obtained for the FPLC–based purification experiments of *ScADH*, *ScPDC* and *TmLDH* enzymes, by using CBM domains and histidine as purification tags. \pm correspond to the standard error of replicates.

Enzyme	Tag	Purification yield (%)	Purification factor	Purity grade (%)
<i>ScADH</i>	CBM3	6 \pm 0	3 \pm 0	-
	CBM9	92 \pm 3	10 \pm 1	80 \pm 2
	histidine	71 \pm 0	7 \pm 2	90 \pm 1
<i>ScPDC</i>	CBM3	0 \pm 0	0 \pm 0	87* \pm 0
	CBM9	74 \pm 8	7 \pm 1	87 \pm 1
	histidine	89 \pm 1	4 \pm 0	92 \pm 0
<i>TmLDH</i>	CBM3	0 \pm 0	0 \pm 0	83* \pm 0
	CBM9	86 \pm 7	7 \pm 1	80 \pm 3
	histidine	79 \pm 1	5 \pm 1	88 \pm 0

* Purity grade according to SDS-PAGE bands corresponding to denatured protein fraction.

According to the results obtained, it can be concluded that the methodology developed to recover purified CBM-tagged enzymes is only applicable for CBM9 domain, since it was not possible to desorb CBM3 domain neither from RAC nor MC supports with glucose after its immobilisation, although it could be recovered with denaturing conditions. Thus, prior to use CBM3 domain as purification tag it would be necessary to carry out additional studies to explore other compounds different from glucose to desorb CBM3 domain from cellulose, apart from EG, which was not suitable in this case due to enzyme inactivation.

Therefore, CBM9 domain proved to be a suitable option for protein purification with cellulosic supports, since high purification yields and purification factors were achieved with this tag for the enzymes tested. Nevertheless, elution samples of histidine-tagged enzymes also retained a remarkable percentage of initial activity and showed higher purity grades for the three enzymes evaluated. In summary, considering the list of drawbacks that can be overcome by using RAC instead of IMAC purification resins, CBM9 domain is a real and feasible alternative to polyhistidine tag, since the overall performance of both purification approaches can be equated, with same order of magnitude yields.

6.5. Conclusions

CBM domains tested in this work were proved to be suitable for one-step purification/immobilisation process, thanks to the high specificity of these domains towards cellulosic supports. Even if *ScADH* enzyme was the most efficiently immobilised protein onto both RAC and MC supports —not only with high retained activities obtained in the final immobilised derivatives but also with high enzyme loadings—, *ScPDC* and *TmLDH* enzymes showed promising results too. However, maximum load capacity of cellulose support was strongly dependent on the CBM fused; CBM3-tagged enzymes proved to bind more effectively towards microcrystalline cellulose, whereas CBM9-tagged enzymes bound with greater affinity to RAC.

Concerning the feasibility of CBM domains to be used in FPLC-based purification processes, the obtained results revealed that CBM3-fused enzymes are not applicable for purification process based on the affinity interaction, under the developed conditions for CBM9-tagged enzymes, since fusion proteins could not be easily desorbed from RAC, under the

tested conditions. For that purpose, a CBM9–fused strategy would be a better option, allowing both a single purification process and a one-step purification/immobilisation process. Additionally, it can be also concluded that CBM9 tags are more suitable and cheaper alternative to traditional polyhistidine tag used to purify proteins by IMAC chromatography.

7. OVERALL CONCLUSIONS

In this thesis several process intensification strategies have been explored, aiming to integrate them transversely for improving the bioprocess metrics of the production of the three enzymes of interest. In this context, the studied methodologies and the most significant results obtained consisted in:

- The heterologous production of the three enzymes in *E. coli* cells, with the main goal to develop a fed-batch production process based on the use of chemically defined minimum culture mediums. NEB 10- β strain showed promising results; enzyme overexpression was successfully achieved when following a two-step fed-batch strategy that combined the use of glucose and glycerol as the two main carbon sources. Higher volumetric productivities were achieved for ScADH enzyme, in contrast to batch cultures with complex medium, in mass but also in activity terms (11- and 2.3-fold increase, respectively). Nevertheless, this increased productivity did not happen for the rest of target enzymes. The results suggest that production downshift observed between the two operational modes could be caused by the effect of temperature and/or by the effect of a possible aminoacid imbalance in defined medium during the fed-batch stage.

Therefore, M15 Δ *glyA* strain was used as microbial host aiming to produce the three enzymes in a more straightforward way, avoiding the medium shift and the amino acid supplementation. Target enzymes were cloned into the in-house-developed pVEF plasmid, what allowed the use of antibiotic-free defined culture medium. The three enzymes were successfully produced at bench-scale reactor with the auxotrophic strain through batch and fed-batch cultures. In general, around 40 mg·g⁻¹ DCW of ScADH and *Tm*LDH were produced, whilst more than 80 mg·g⁻¹ DCW of ScPDC were obtained, yielding from 1.1 to 1.3 μ mol·g⁻¹DCW. Fed-batch processes led to an increase of one order of magnitude in volume productivity.

- The fusion of four CBM domains to ScADH enzyme via rational design aiming to develop a cost- and time-effective one-step purification/immobilisation method. Two of the variants (CBM3- and CBM9-ScADH) could be correctly overexpressed in *E. coli* and were subsequently fused to ScPDC and *Tm*LDH enzymes. The six fusion proteins were successfully produced in a soluble and active form in the M15 Δ *glyA* strain by following a batch strategy with the antibiotic-free defined medium.

In general, more than one gram of active and soluble target product was produced in all cases, corresponding to approximately 0.9–1 μmol of enzyme. Regarding the production yields achieved, similar results were obtained for the two *ScADH* variants, whereas *ScPDC* and *TmLDH* production values varied depending on the CBM. In all cases, specific activity values ($\text{AU}\cdot\mu\text{mol}^{-1}$) among the variants of each enzyme showed minor differences.

Finally, it was attempted to produce the two *ScADH* variants in fed-batch cultures, to compare the yields obtained with the histidine-tagged version, with which slightly better results were shown. Specific enzyme activity values of the CBM-tagged enzymes were compared respect to their histidine-tagged versions, and different results were obtained for each of the enzymes tested, even if in all cases total produced activity fluctuate within the same magnitude order.

- The development of a one-step purification/immobilisation process for the two CBM domains fused to the enzymes, which proved to be suitable candidates, thanks to their high specificity towards cellulosic supports. Promising values were obtained in all cases, since most of the fusion proteins tested retained high percentages of activity after their immobilisation, especially for *ScADH* variants, which showed the highest retained activity values. Results also showed that immobilised derivatives were more stable than soluble enzymes under storage conditions, in most of the cases. Maximum enzyme load capacity was strongly related on the CBM fused; CBM3-tagged enzymes proved to bind more effectively towards microcrystalline cellulose, whereas CBM9-tagged enzymes bound with greater affinity to RAC.

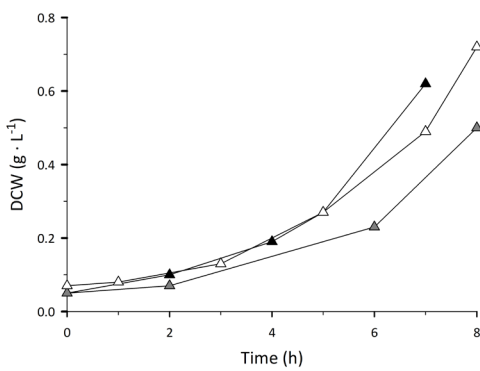
- The development of a FPLC-based purification process as another possibility different from immobilisation. The results obtained revealed that CBM3-fused enzymes are not applicable for protein purification, under the tested conditions, since fusion proteins could not be easily desorbed from neither RAC and MC supports. On the other hand, CBM9-tagged enzymes were recovered by using glucose as eluent. Purification factors of 10, 7 and 6.6 were obtained for CBM9-tagged *ScADH*, *ScPDC* and *TmLDH* enzymes, respectively, with a final purity grade between 80 and 88%. These values are found within the same range than the obtained for histidine-tagged variants, meaning therefore that the developed process stands as a suitable and cheap alternative to traditional IMAC chromatography methods.

8. APPENDIX

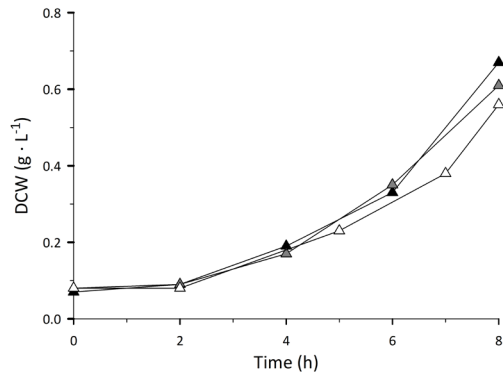
8.1. Results I

8.1.1. Adaptation of NEB 10- β cells to defined medium

A: His-ScADH



B: His-ScPDC



C: His-TmLDH

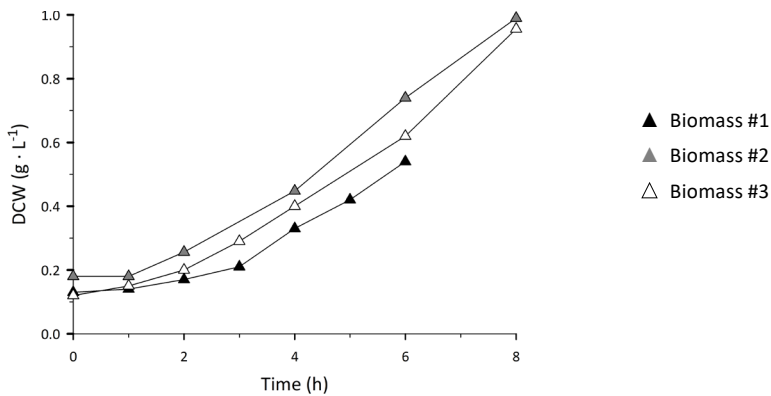


Figure 8.1. Shake-flask cultures performed to adapt *E. coli* NEB 10- β cells to DM medium. Biomass DCW ($\text{g} \cdot \text{L}^{-1}$) measured along time for histidine-tagged ScADH (A), ScPDC (B) and TmLDH (C) variants, during the first (black), second (grey) and third (white) growth cycles. Culture conditions: 37°C, pH 7.0, 140 rpm.

8.1.2. Fed-batch production processes at bench-scale reactor and defined medium

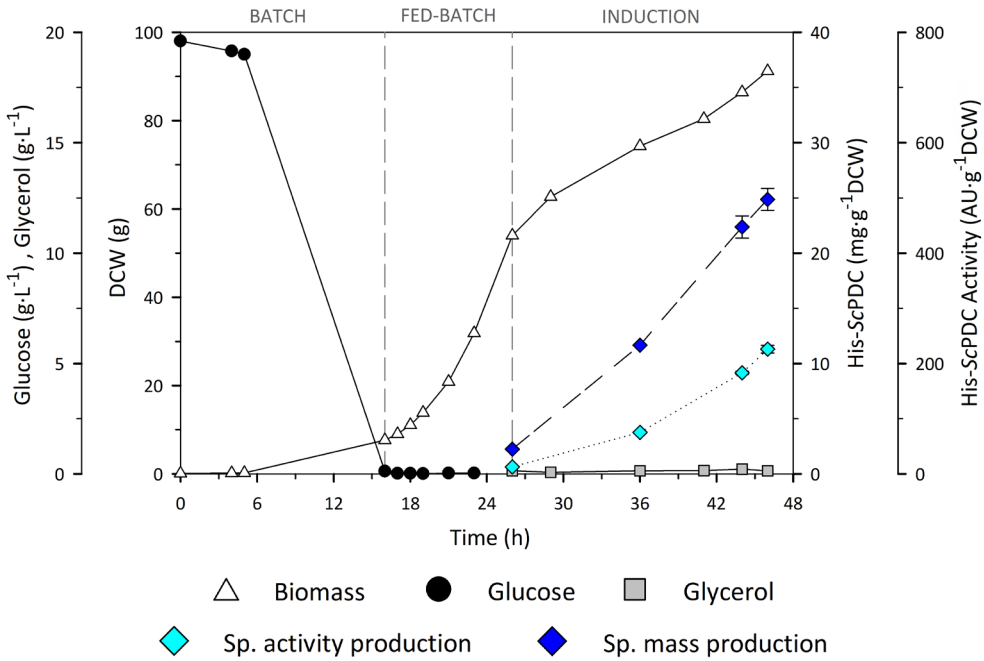


Figure 8.2. *E. coli* NEB 10-β fed-batch culture performed to produce His-ScPDC enzyme with DM. Profiles along time of biomass DCW (Δ) (g), glucose (●) and glycerol (◻) (g·L⁻¹), specific activity (◊) (AU·g⁻¹·DCW) and specific mass production (◆) (mg·g⁻¹·DCW). Batch, fed-batch and induction phases indicated. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 30% PO₂, 0.2 g·L⁻¹ L-arabinose, 0.1 g·L⁻¹. Error bars correspond to standard error of three sample measurements.

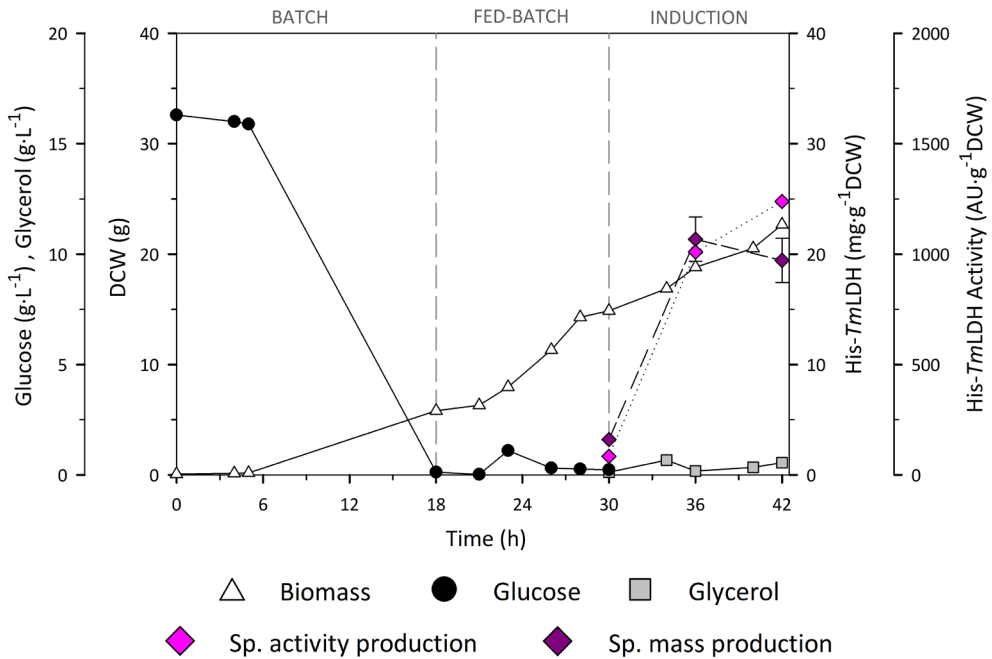
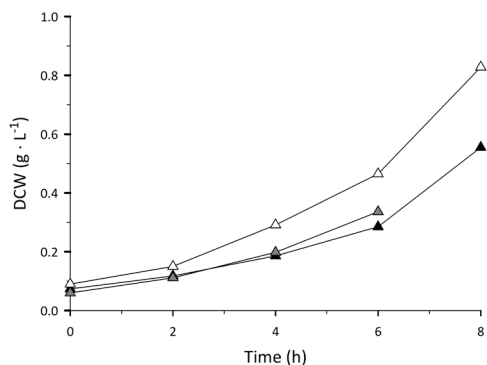


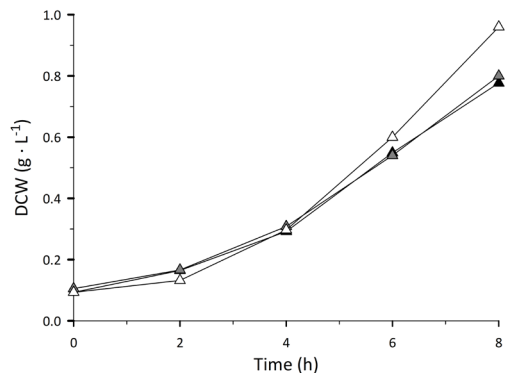
Figure 8.3. *E. coli* NEB 10- β fed-batch culture performed to produce His-TmLDH enzyme with DM. Profiles along time of biomass DCW (Δ) (g), glucose (\bullet) and glycerol (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Batch, fed-batch and induction phases indicated. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 30% PO_2 , 0.2 $\text{g}\cdot\text{L}^{-1}$ L-arabinose, 0.1 $\text{g}\cdot\text{L}^{-1}$. Error bars correspond to standard error of three sample measurements.

8.1.3. Adaptation of M15 Δ glyA cells to defined medium

A: His-ScADH



B: His-ScPDC



C: His-TmLDH

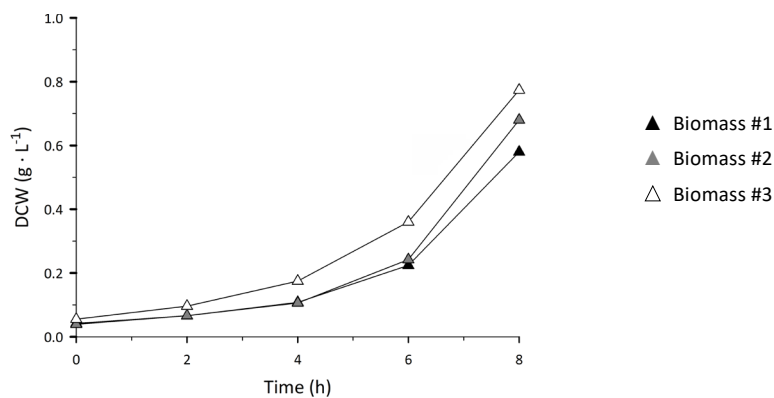


Figure 8.4. Shake-flask cultures performed to adapt *E. coli* M15 Δ glyA cells to DM. Biomass DCW ($\text{g}\cdot\text{L}^{-1}$) measured along time for histidine-tagged ScADH (A), ScPDC (B) and TmLDH (C) variants, during the first (black), second (grey) and third (white) growth cycles. Culture conditions: 37°C, pH 7.0, 140 rpm.

8.1.4. Batch production processes at bench-scale reactor

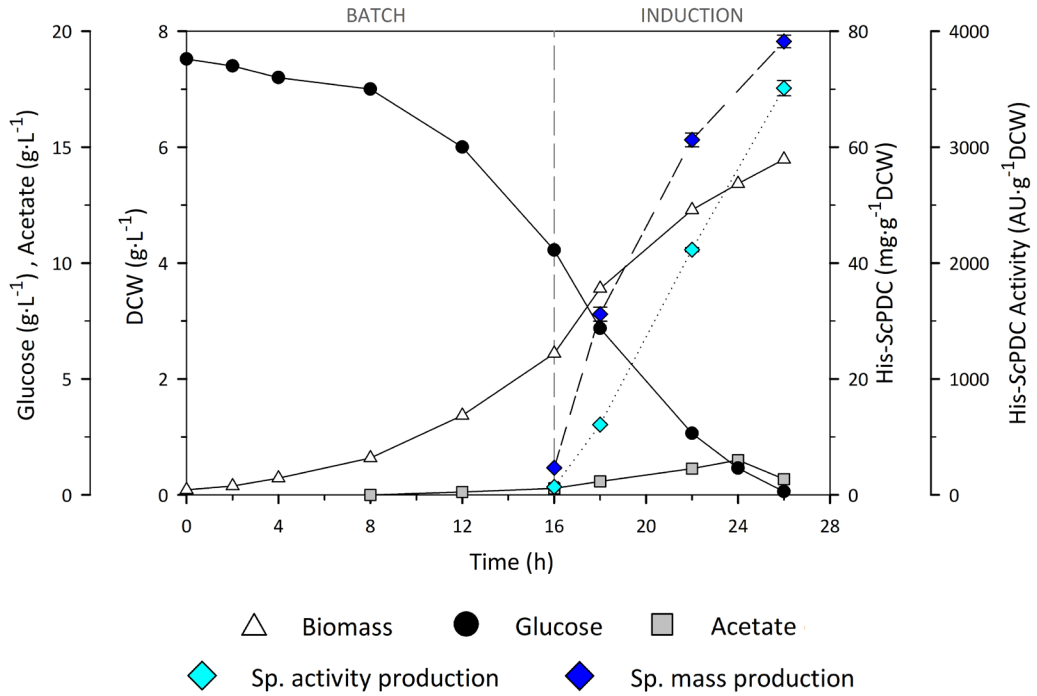


Figure 8.5. *E. coli* M15ΔglyA batch culture performed to produce His-ScPDC enzyme with DM. Profiles along time of biomass DCW (Δ), glucose (●) and acetate (◻) (g·L⁻¹), specific activity (◊) (AU·g⁻¹DCW) and specific mass production (◆) (mg·g⁻¹DCW). Batch and induction phases indicated. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

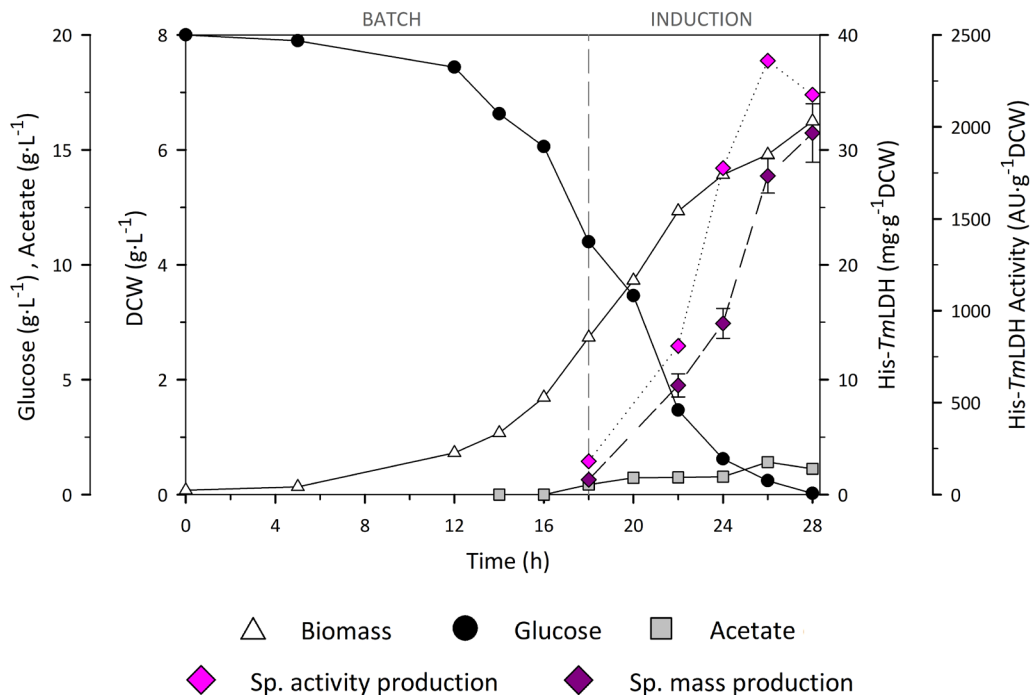


Figure 8.6. *E. coli* M15ΔglyA batch culture performed to produce His-TmLDH enzyme with DM. Profiles along time of biomass DCW (Δ), glucose (\bullet) and acetate (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Batch and induction phases indicated. Culture conditions: 24°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

8.1.5. Fed-batch production processes at bench-scale reactor

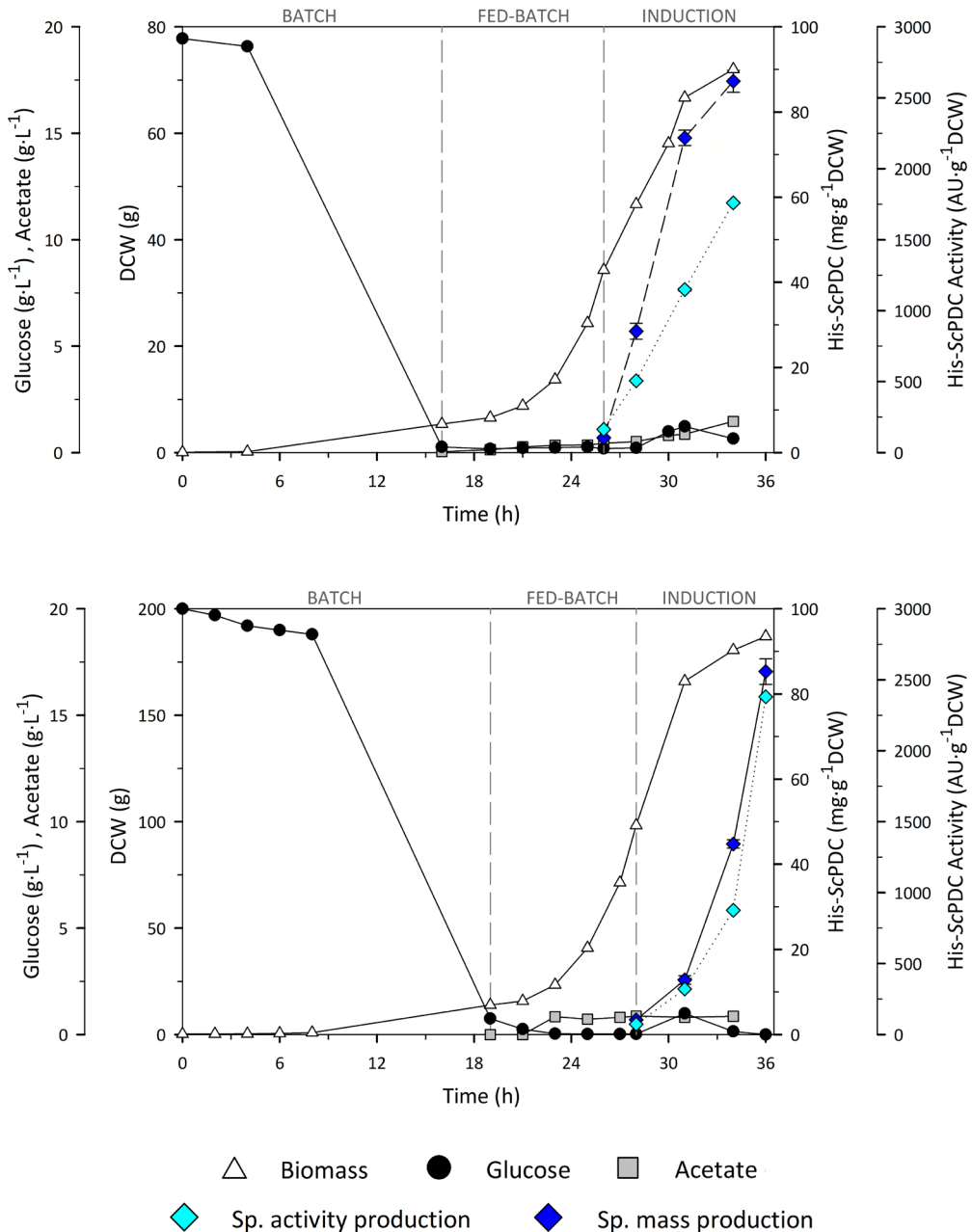


Figure 8.7. *E. coli* M15 Δ glyA fed-batch cultures performed to produce His-ScPDC enzyme with DM with a 2 L vessel (top) and a 5 L vessel (bottom). Profiles along time of biomass DCW (\triangle) (g), glucose (\bullet) and acetate (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\diamond) ($\text{AU}\cdot\text{g}^{-1}\cdot\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\cdot\text{DCW}$). Batch, fed-batch and induction phases indicated. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 60% PO_2 , 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

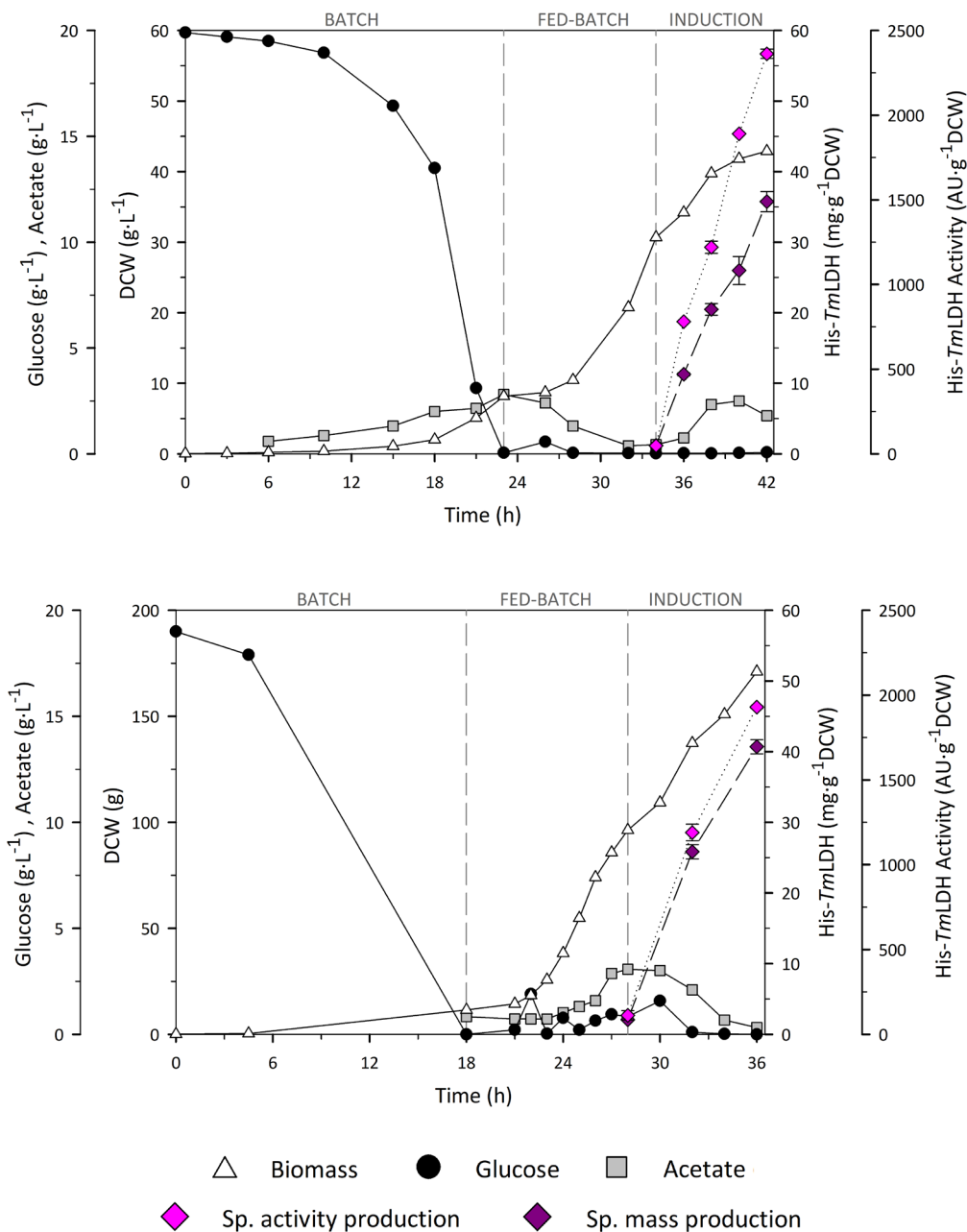
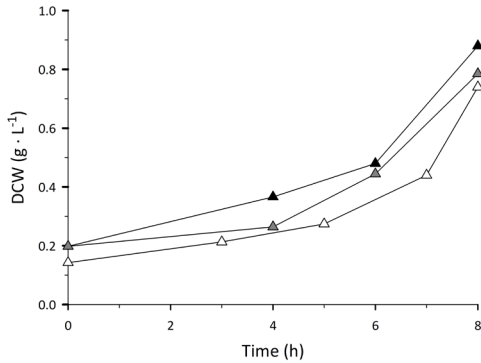


Figure 8.8. *E. coli* M15 Δ glyA fed-batch cultures performed to produce His-TmLDH enzyme with DM with a 2 L vessel (top) and a 5 L vessel (bottom). Profiles along time of biomass DCW (Δ) (g), glucose (\bullet) and acetate (\square) ($\text{g}\cdot\text{L}^{-1}$), specific activity (\blacklozenge) ($\text{AU}\cdot\text{g}^{-1}\text{DCW}$) and specific mass production (\blacklozenge) ($\text{mg}\cdot\text{g}^{-1}\text{DCW}$). Batch, fed-batch and induction phases indicated. Culture conditions: 37°C, pH 7.0, 450–1100 rpm, 60% PO₂, 0.25 mM IPTG. Error bars correspond to standard error of three sample measurements.

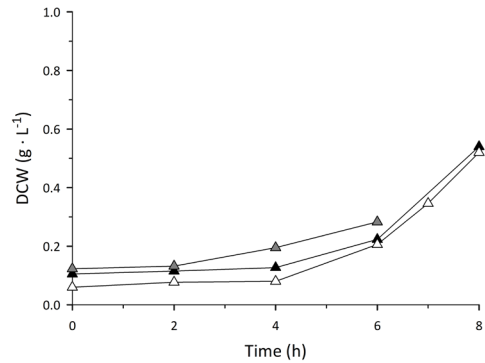
8.2. Results II

8.2.1. Adaptation of M15 Δ glyA to defined medium

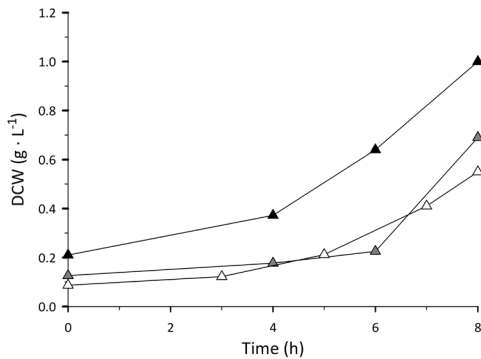
A: CBM3-ScADH



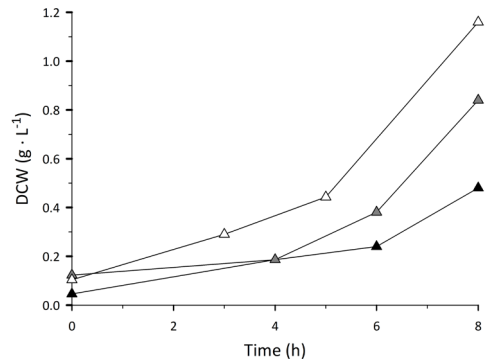
CBM9-ScADH



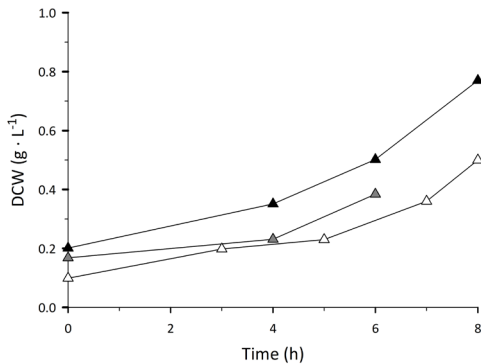
B: CBM3-ScPDC



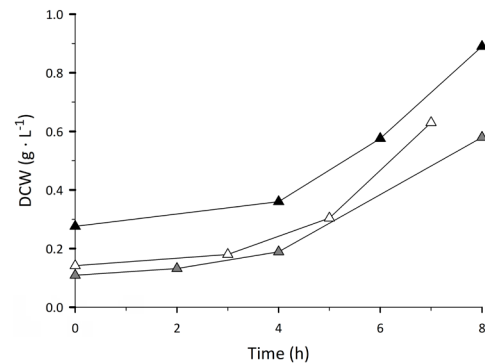
CBM9-ScPDC



C: CBM3-TmLDH



CBM9-TmLDH



▲ Biomass #1 ▲ Biomass #2 △ Biomass #3

Figure 8.9. *E. coli* M15 Δ glyA shake-flask cultures performed to adapt the new transformants to DM. Biomass DCW (g·L⁻¹) measured along time for CBM-ScADH (A), CBM-ScPDC (B) and CBM-TmLDH (C) variants, during the first (black), second (grey) and third (white) growth cycles. Culture conditions: 37°C, pH 7.0, 140 rpm.

9. ACKNOWLEDGEMENTS

9.1. Funding and contributions

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement n°761042 (BIOCON-CO₂). This output reflects the views only of the author and the European Commission cannot be held responsible for any use which may be made of the information contained therein. The research group is recognized by Generalitat de Catalunya as 2017 SGR 1462.



The author thanks Prof. Marco Fraaije, from Rijksuniversiteit Groningen, for sharing the pBAD vectors and Pharmaceutical Solutions, from DuPont Nutrition & Biosciences for kindly donating a sample of Avicel[®] PH-200 cellulose. Author also thanks COST Action CM 1303–Systems Biocatalysis for financial support.

Finally, the author acknowledges also UAB for funding his PhD PIF grant.

10. SCIENTIFIC CONTRIBUTIONS

- Miret, J., Román, R., Benito, M., Casablanca, A., Guillén, M., Álvaro, G., González, G., "Development of a highly efficient production process for recombinant protein expression in *Escherichia coli* NEB10 β ", *Biochemical Engineering Journal*, 159, 107612, 1-8, 2020.
<https://doi.org/10.1016/J.BEJ.2020.107612>

Production of His-ScADH enzyme with NEB 10- β strain as a validation of a fed-batch production process developed to obtain PTDH-CHMO fusion protein.

- Benito, M., Román, R., Ortiz, G., Casablanca, A., Álvaro, G., Caminal, G., González, G., Guillén, M., "Cloning, expression, and one-step purification/immobilisation of two carbohydrate-binding module-tagged alcohol dehydrogenases", *Journal of Biological Engineering*, 16, 16, 2022.
<https://doi.org/10.1186/s.13036-022-00295-8>

Fusion of CBM3 and CBM9 domains to ScADH enzyme and subsequent development of a one-step purification/immobilisation method as a process intensification strategy.

11. REFERENCES

- Ahn, J.O., Choi, E.S., Lee, H.W., Hwang, S.H., Kim, C.S., Jang, H.W., Haam, S.J., Jung, J.K., "Enhanced secretion of *Bacillus stearothermophilus* L1 lipase in *Saccharomyces cerevisiae* by translational fusion to cellulose-binding domain", *Applied Microbiology and Biotechnology*, 64, 833–839, 2004.
<https://doi.org/10.1007/S00253-003-1547-5>
- Alcover, N., Carceller, A., Álvaro, G., Guillén, M., "Zymobacter palmae pyruvate decarboxylase production process development: Cloning in *Escherichia coli*, fed-batch culture and purification", *Engineering in Life Sciences*, 19, 502–512, 2019.
<https://doi.org/10.1002/elsc.201900010>
- Allocati, N., Masulli, M., Alexeyev, M.F., di Ilio, C., "Escherichia coli in Europe: An Overview", *International Journal of Environmental Research and Public Health*, 10, 6235-6254, 2013.
<https://doi.org/10.3390/IJERPH10126235>
- Amann, E., Brosius, J., Ptashne, M., "Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli*", *Gene*, 25 (2), 167–178, 1983.
[https://doi.org/10.1016/0378-1119\(83\)90222-6](https://doi.org/10.1016/0378-1119(83)90222-6)
- Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D., Jordan, F., "Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast *Saccharomyces cerevisiae* at 2.3 Å resolution", *Journal of Molecular Biology*, 256, 590–600, 1996.
<https://doi.org/10.1006/jmbi.1996.0111>
- Armenta, S., Moreno-Mendieta, S., Sánchez-Cuapio, Z., Sánchez, S., Rodríguez-Sanoja, R., "Advances in molecular engineering of carbohydrate-binding modules", *Proteins: Structure, Function, and Bioinformatics*, 85 (9), 1602–1617, 2017.
<https://doi.org/10.1002/PROT.25327>
- Asif, A., Mohsin, H., Tanvir, R., Rehman, Y., "Revisiting the Mechanisms Involved in Calcium Chloride Induced Bacterial Transformation", *Frontiers in Microbiology*, 8, 2169, 1-5, 2017.
<https://doi.org/10.3389/FMICB.2017.02169>

- Aslanidis, C., Jong, P.J. de, "Ligation-independent cloning of PCR products (LIC-PCR)", *Nucleic Acids Research*, 18, 6069–6074, 1990.
<https://doi.org/10.1093/NAR/18.20.6069>
- Baker, P.J., Britton, K.L., Fisher, M., Esclapez, J., Pire, C., Bonete, M.J., Ferrer, J., Rice, D.W., "Active site dynamics in the zinc-dependent medium chain alcohol dehydrogenase superfamily", *Proc. Natl. Acad. Sci. USA*, 106 (3), 779-784, 2009.
<https://doi.org/10.1073/PNAS.0807529106>
- Barcelos, M.C.S., Lupki, F.B., Campolina, G.A., Nelson, D.L., Molina, G., "The colors of biotechnology: general overview and developments of white, green and blue areas", *FEMS Microbiology Letters*, 365 (21), 1-11, 2018.
<https://doi.org/10.1093/femsle/fny239>
- Basso, A., Serban, S., "Industrial applications of immobilized enzymes—A review", *Molecular Catalysis*, 479, 110607, 2019.
<https://doi.org/10.1016/J.MCAT.2019.110607>
- Beckwith, J.R., "A deletion analysis of the lac operator region in *Escherichia coli*", *Journal of Molecular Biology*, 8 (3), 427–430, 1964.
[https://doi.org/10.1016/S0022-2836\(64\)80206-0](https://doi.org/10.1016/S0022-2836(64)80206-0)
- Bergmans, H.E., Die, I.M. van, Hoekstra, W.P., "Transformation in *Escherichia coli*: stages in the process", *Journal of Bacteriology* 146 (2), 564-570, 1981.
<https://doi.org/10.1128/jb.146.2.564-570.1981>
- Bhatia, S., Goli, D., "Introduction to Pharmaceutical Biotechnology: Basic techniques and concepts", Vol.1, IOP Publishing Ltd., 2018. ISBN: 978-0-7503-1300-1.
<https://doi.org/10.1088/978-0-7503-1299-8>
- Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., Schwede, T., "The SWISS-MODEL Repository—new features and functionality", *Nucleic Acids Research*, 45, 313-319, 2017.
<https://doi.org/10.1093/NAR/GKW1132>
- Birnboim, H.C., Doly, J., "A rapid alkaline extraction procedure for screening recombinant plasmid DNA", *Nucleic Acids Research* 7 (6), 1513-1523, 1979.
<https://doi.org/10.1093/NAR/7.6.1513>

- Boiteux, A., Hess, B., "Allosteric properties of yeast pyruvate decarboxylase", *FEBS Letters* 9 (5), 293–296, 1970.
[https://doi.org/10.1016/0014-5793\(70\)80381-7](https://doi.org/10.1016/0014-5793(70)80381-7)
- Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J., "Carbohydrate-binding modules: Finetuning polysaccharide recognition", *Biochemical Journal*, 382 (3), 769–781, 2004.
<https://doi.org/10.1042/BJ20040892>
- Boraston, A.B., Creagh, A.L., Alam, M.M., Kormos, J.M., Tomme, P., Haynes, C.A., Warren, R.A.J., Kilburn, D.G., "Binding specificity and thermodynamics of a family 9 carbohydratebinding module from *Thermotoga maritima* xylanase 10A", *Biochemistry*, 40, 6240–6247, 2001.
<https://doi.org/10.1021/bi0101695>
- Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.T., Dobberstein, B., "A T5 Promoter-Based Transcription-Translation System for the Analysis of Proteins in Vitro and in Vivo", *Methods in Enzymology*, 155 (C), 416–433, 1987.
[https://doi.org/10.1016/0076-6879\(87\)55028-5](https://doi.org/10.1016/0076-6879(87)55028-5)
- Calleja, D., Fernández-Castañé, A., Pasini, M., de Mas, C., López-Santín, J., "Quantitative modeling of inducer transport in fed-batch cultures of *Escherichia coli*", *Biochemical Engineering Journal*, 91, 210–219, 2014.
<https://doi.org/10.1016/J.BEJ.2014.08.017>
- Calleja, D., Kavanagh, J., Mas, C. de, López-Santín, J., "Simulation and prediction of protein production in fed-batch *E. coli* cultures: An engineering approach", *Biotechnology and Bioengineering*, 113 (4), 772–782, 2016.
<https://doi.org/10.1002/BIT.25842>
- Cao, L., "Immobilized Enzymes: science or art?". *Current Opinion in Chemical Biology*, 9 (2), 217-226, 2005.
<https://doi.org/10.1016/j.cbpa.2005.02.014>
- Cao, L., "Immobilized Enzymes". *Comprehensive Biotechnology*, Second Edition, Elsevier B.V., 2, 461-476, 2011. ISBN: 978-0-08-088504-9.
<https://doi.org/10.1016/B978-0-08-088504-9.00168-9>

- Chen, R., "Bacterial expression systems for recombinant protein production: E. coli and beyond", *Biotechnology Advances*, 30, 1102–1107, 2012.
<https://doi.org/10.1016/J.BIOTECHADV.2011.09.013>
- Cherry, J.R., Fidantsef, A.L., "Directed evolution of industrial enzymes: an Update", *Current Opinion in Biotechnology*, 14, 438–443, 2003.
[https://doi.org/10.1016/S0958-1669\(03\)00099-5](https://doi.org/10.1016/S0958-1669(03)00099-5)
- Crowe, J., Masone, B.S., Ribbe, J., "One-Step Purification of Recombinant Proteins with the 6xHis Tag and Ni-NTA Resin", *Methods Mol. Biol.*, 58, 491–510, 1996.
<https://doi.org/10.1385/0-89603-402-X:491>
- Datta, S., Christena, L.R., Rajaram, Y.R.S., "Enzyme immobilisation: an overview on techniques and support materials", *3 Biotech*, 3, 1-9, 2013.
<https://doi.org/10.1007/S13205-012-0071-7>
- de Boer, H.A., Comstock, L.J., Vassert, M., "The tac promoter: A functional hybrid derived from the trp and lac promoters", *Proc. Natl. Acad. Sci. USA*, 80, 21-25, 1983.
- de Smidt, O., du Preez, J.C., Albertyn, J., "The alcohol dehydrogenases of *Saccharomyces cerevisiae*: A comprehensive review", *FEMS Yeast Research*, 8 (7), 967–978, 2008.
<https://doi.org/10.1111/j.1567-1364.2008.00387.x>
- DeLisa, M.P., Li, J., Rao, G., Weigand, W.A., Bentley, W.E., "Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor", *Biotechnology and Bioengineering*, 65 (1), 54–65.,1999.
- Dickinson, J.R., Salgado, L.E.J., Hewlins, M.J.E., "The Catabolism of Amino Acids to Long Chain and Complex Alcohols in *Saccharomyces cerevisiae*", *Journal of Biological Chemistry*, 278 (10), 8028–8034, 2003.
<https://doi.org/10.1074/JBC.M211914200>
- Dimian, A.C., Bildea, C.S., Kiss, A.A., "Process Intensification", *Computer Aided Chemical Engineering*, 35, 397–448, 2014.
<https://doi.org/10.1016/B978-0-444-62700-1.00010-3>

- Doran, P.M., "Bioprocess Development", Bioprocess Engineering Principles, Second Edition, Elsevier Ltd., 3-11, 2013. ISBN: 978-0-12-220851-53–11.
<https://doi.org/10.1016/B978-0-12-220851-5.00001-0>
- dos Santos, J.C.S., Barbosa, O., Ortiz, C., Berenguer-Murcia, A., Rodrigues, R.C., FernandezLafuente, R., "Importance of the Support Properties for Immobilisation or Purification of Enzymes", ChemCatChem, 7 (16), 2413–2432, 2015.
<https://doi.org/10.1002/CCTC.201500310>
- Drewke, C., Ciriacy, M., "Overexpression, purification and properties of alcohol dehydrogenase IV from *Saccharomyces cerevisiae*", Biochimica et Biophysica Acta, 950 (1), 54–60, 1988.
[https://doi.org/10.1016/0167-4781\(88\)90072-3](https://doi.org/10.1016/0167-4781(88)90072-3)
- Durany, O., Caminal, G., de Mas, C., López-Santín, J., "Studies on the expression of recombinant fuculose-1-phosphate aldolase in *E. Coli*", Process Biochemistry, 39 (11), 1677–1684, 2004.
[https://doi.org/10.1016/S0032-9592\(03\)00302-9](https://doi.org/10.1016/S0032-9592(03)00302-9)
- Dwevedi, A., Kayastha, A.M., "Enzyme Immobilisation: A Breakthrough in Enzyme Technology and Boon to Enzyme Based Industries", Protein Structure, Nova Science Publishers, 31–50, 2011. ISBN: 978-1-61209-656-8
- Elvin, C.M., Thompson, P.R., Argall, M.E., Philip Hendr, N., Stamford, P.J., Lilley, P.E., Dixon, N.E., "Modified bacteriophage lambda promoter vectors for overproduction of proteins in *Escherichia coli*", Gene, 87 (1), 123–126, 1990.
[https://doi.org/10.1016/0378-1119\(90\)90503-J](https://doi.org/10.1016/0378-1119(90)90503-J)
- "Enzymes Market Size & Share Report, 2021-2028", Market Analysis Report, Grand View Research, 1-153, 2021. ISBN: 978-1-68038-022-4
- Florio, R., di Salvo, M.L., Vivoli, M., Contestabile, R., "Serine hydroxymethyltransferase: A model enzyme for mechanistic, structural, and evolutionary studies", Biochimica et Biophysica Acta, 1814 (11), 1489–1496, 2010.
<https://doi.org/10.1016/J.BBAPAP.2010.10.010>

- García-Bofill, M., Sutton, P.W., Guillén, M., Álvaro, G., "Enzymatic synthesis of a statin precursor by immobilised alcohol dehydrogenase with NADPH oxidase as cofactor regeneration System", *Applied Catalysis A, General*, 609, 117909, 2021 (a).
<https://doi.org/10.1016/J.APCATA.2020.117909>
- García-Bofill, M., Sutton, P.W., Straatman, H., Brummund, J., Schürmann, M., Guillén, M., Álvaro, G., "Biocatalytic synthesis of vanillin by an immobilised eugenol oxidase: High biocatalyst yield by enzyme recycling", *Applied Catalysis A, General*, 610, 117934, 2021 (b).
<https://doi.org/10.1016/J.APCATA.2020.117934>
- Garibyan, L., Avashia, N., "Polymerase Chain Reaction", *Journal of Investigative Dermatology*, 133, 1–4, 2013.
<https://doi.org/10.1038/JID.2013.1>
- Gentz, R., Bujard, H., "Promoters recognized by *Escherichia coli* RNA polymerase selected by function: highly efficient promoters from bacteriophage T5", *Journal of Bacteriology*, 164 (1), 70-77, 1985.
<https://doi.org/10.1128/jb.164.1.70-77.1985>
- Gilkess, N.R., Antony, R., Warren, J., Miller, R.C., Kilburn, D.G., "Precise Excision of the Cellulose Binding Domains from Two *Cellulomonas fimi* Cellulases by a Homologous Protease and the Effect on Catalysis*", *The Journal of Biological Chemistry*, 263 (21), 10401–10407, 1988.
[https://doi.org/10.1016/S0021-9258\(19\)81530-2](https://doi.org/10.1016/S0021-9258(19)81530-2)
- Grant, S.G.N., Jesseet, J., Bloomt, F.R., Hanahan, D., "Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants", *Proc. Natl. Acad. Sci. USA*, 87, 4645–4649, 1990.
- Guzman, L.-M., Belin, D., Carson, M.J., Beckwith, J., "Tight Regulation, Modulation, and High Level Expression by Vectors Containing the Arabinose P BAD Promoter", *Journal of Bacteriology*, 177 (14), 4121–4130, 1995.
- Hägg, P., Wa De Pohl, J., Abdulkarim, F., Isaksson, L.A., "A host/plasmid system that is not dependent on antibiotics and antibiotic resistance genes for stable plasmid maintenance in *Escherichia coli*", *Journal of Biotechnology*, 111, 17–30, 2004.
<https://doi.org/10.1016/j.jbiotec.2004.03.010>

- Han, Q., Eiteman, M.A., "Acetate formation during recombinant protein production in *Escherichia coli* K-12 with an elevated NAD(H) pool", *Engineering in Life Sciences*, 19, 770-780, 2019.
<https://doi.org/10.1002/elsc.201900045>
- Hecht, K., Wrba, A., Jaenicke, R., "Catalytic properties of thermophilic lactate dehydrogenase and halophilic malate dehydrogenase at high temperature and low water activity", *European Journal of Biochemistry*, 183, 69–74, 1989.
<https://doi.org/10.1111/j.1432-1033.1989.tb14897.x>
- Hedstrom, L., "Enzyme Specificity and Selectivity", *Encyclopedia of Life Sciences*, John Wiley & Sons, Ltd, 1-8, 2010. ISBN: 9780470015902.
<https://doi.org/10.1002/9780470015902.a0000716.pub2>
- Homaei, A.A., Sariri, R., Vianello, F., Stevanato, R., "Enzyme immobilisation: an Update", *J. Chem. Biol.*, 6 (8), 185–205, 2013.
<https://doi.org/10.1007/s12154-013-0102-9>
- Hong, J., Ye, X., Wang, Y., Zhang, Y.H.P., "Bioseparation of recombinant cellulose-binding module-proteins by affinity adsorption on an ultra-high-capacity cellulosic adsorbent", *Analytica Chimica Acta*, 621 (2), 193–199, 2008.
<https://doi.org/10.1016/j.aca.2008.05.041>
- Huang, C.-J., Lin, H., Yang, X., "Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements", *J. Ind. Microbiol. Biotechnol.*, 39, 383–399, 2012.
<https://doi.org/10.1007/s10295-011-1082-9>
- Illanes, A., "Enzyme biocatalysis: Principles and Applications", 1st ed., Springer-Verlag New York Inc., 1-56, 2008. ISBN: 9780470015902.
<https://doi.org/10.1007/978-1-4020-8361-7>
- Islam, M.N., Lee, K.W., Yim, H.S., Lee, S.H., Jung, H.C., Lee, J.H., Jeong, J.Y., "Optimizing T4 DNA polymerase conditions enhances the efficiency of one-step sequence- and ligationindependent cloning", *BioTechniques*. 63 (3), 125–130, 2017.
<https://doi.org/10.2144/000114588>

- Jeong, J.Y., Yim, H.S., Ryu, J.Y., Lee, H.S., Lee, J.H., Seen, D.S., Kang, S.G., "One-step sequence-and ligation-independent cloning as a rapid and versatile cloning method for functional genomics Studies", *Applied and Environmental Microbiology*, 78 (15), 5440–5443, 2012.
<https://doi.org/10.1128/AEM.00844-12>
- Juers, D.H., Matthews, B.W., Huber, R.E., "LacZ β -galactosidase: Structure and function of an enzyme of historical and molecular biological importance", *Protein Science*, 21, 1792–1807, 2012.
<https://doi.org/10.1002/pro.2165>
- Kafarski, P., "Rainbow code of biotechnology", *CHEMIK*, 66 (8), 811–816, 2012.
- Kauffman, C., Shoseyov, O., Shpigel, E., Bayer, E.A., Lamed, R., Shoham, Y., Mandelbaum, R.T., "Novel Methodology for Enzymatic Removal of Atrazine from Water by CBD-Fusion Protein Immobilized on Cellulose", *Environmental Science & Technology*, 34 (7), 1292–1296, 2000.
<https://doi.org/10.1021/es990754h>
- Kavooosi, M., Meijer, J., Kwan, E., Creagh, A.L., Kilburn, D.G., Haynes, C.A., "Inexpensive onestep purification of polypeptides expressed in *Escherichia coli* as fusions with the family 9 carbohydrate-binding module of xylanase 10A from *T. Marítima*", *Journal of Chromatography B*, 807 (1), 87–94, 2004.
<https://doi.org/10.1016/j.jchromb.2004.03.031>
- Khan, M.R., "Immobilized enzymes: a comprehensive review", *Bulletin of the National Research Centre*, 207 (45), 1-13, 2021.
<https://doi.org/10.1186/s42269-021-00649-0>
- König, S., "Subunit structure, function and organisation of pyruvate decarboxylases from various organisms", *Biochimica et Biophysica Acta*, 1385 (2), 271–286, 1998.
[https://doi.org/10.1016/S0167-4838\(98\)00074-0](https://doi.org/10.1016/S0167-4838(98)00074-0)
- Labrou, N.E., "Protein Purification: An Overview", *Methods in Molecular Biology*, 1129, 3–10, 2014.
https://doi.org/10.1007/978-1-62703-977-2_1

- Lammerts van Bueren, A., Ficko-Blean, E., "Carbohydrate-binding modules", CAZypedia, Carbohydrate-active Enzymes, Accessed 11-04-2022.
https://www.cazypedia.org/index.php/Carbohydrate-binding_modules
- Lee, S.Y., "High cell-density culture of *Escherichia coli*", *Trends in Biotechnology*, 14 (3), 98–105, 1996.
[https://doi.org/10.1016/0167-7799\(96\)80930-9](https://doi.org/10.1016/0167-7799(96)80930-9)
- Lee, S.Y., Kim, H.U., "Systems strategies for developing industrial microbial strains", *Nature Biotechnology*, 33 (10), 1061–1072, 2015.
<https://doi.org/10.1038/nbt.3365>
- Lehmann, H., FISCHER, G., Hübner, G., KOHNEBT, K. -D, Schellenberger, A., "The Influence of Steric and Electronic Parameters on the Substrate Behaviour of α -Oxo Acids to Yeast Pyruvate Decarboxylase", *European Journal of Biochemistry*, 32 (1), 83–87, 1973.
<https://doi.org/10.1111/J.1432-1033.1973.TB02582.X>
- Leskovac, V., Trivić, S., Peričin, D., "The three zinc-containing alcohol dehydrogenases from baker's yeast, *Saccharomyces cerevisiae*", *FEMS Yeast Research*, 2 (4), 481–494, 2002.
<https://doi.org/10.1111/J.1567-1364.2002.TB00116.X>
- Levy, I., Shoseyov, O., "Cellulose-binding domains: Biotechnological Applications", *Biotechnology Advances*, 20 (3-4), 191–213, 2002.
[https://doi.org/10.1016/S0734-9750\(02\)00006-X](https://doi.org/10.1016/S0734-9750(02)00006-X)
- Lin, B., Tao, Y., "Whole-cell biocatalysts by design", *Microbial Cell Factories*, 16 (1), 1–12, 2017. <https://doi.org/10.1186/S12934-017-0724-7/FIGURES/2>
- Lobedanz, S., Damhus, T., Borchert, T. v., Hansen, T.T., Lund, H., Lai, W., Lin, M., Leclerc, M., Kirk, O., "Enzymes in Industrial Biotechnology", *Kirk-Othmer Encyclopedia of Chemical Technology*, John Wiley & Sons Inc., 1–73, 2016. ISBN: 9780471484943.
<https://doi.org/10.1002/0471238961.0914042114090512.A01.PUB3>
- Masdeu, G., Kralj, S., Pajk, S., López-Santín, J., Makovec, D., Álvaro, G., "Hybrid chloroperoxidase-magnetic nanoparticle clusters: effect of functionalization on biocatalyst performance", *Journal of Chemical Technology & Biotechnology*, 93 (1), 233–245, 2018.
<https://doi.org/10.1002/JCTB.5345>

- Mayer, M.P., "A new set of useful cloning and expression vectors derived from pBlueScript", *Gene*, 163 (1), 41–46, 1995.
[https://doi.org/10.1016/0378-1119\(95\)00389-N](https://doi.org/10.1016/0378-1119(95)00389-N)
- Mergulhão, F.J.M., Summers, D.K., Monteiro, G.A., "Recombinant protein secretion in *Escherichia coli*", *Biotechnology Advances*, 23 (3), 177–202, 2005.
<https://doi.org/10.1016/J.BIOTECHADV.2004.11.003>
- Miret, J., Román, R., Benito, M., Casablancas, A., Guillén, M., Álvaro, G., González, G., "Development of a highly efficient production process for recombinant protein expression in *Escherichia coli* NEB10 β ", *Biochemical Engineering Journal*, 159, 107612, 1-8, 2020.
<https://doi.org/10.1016/J.BEJ.2020.107612>
- Morag, E., Lapidot, A., Govorko, D., Lamed, R., Wilchek, M., Bayer, E.A., Shoham, Y., "Expression, purification, and characterization of the cellulose-binding domain of the scaffoldin subunit from the cellulosome of *Clostridium thermocellum*", *Applied and Environmental Microbiology*, 61 (5), 1980–1986, 1995.
<https://doi.org/10.1128/aem.61.5.1980-1986.1995>
- Mullis, K.B., "The Unusual Origin of the Polymerase Chain Reaction", *Scientific American*, 262 (4), 56–65, 1990.
<https://doi.org/10.2307/24996713>
- Myung, S., Zhang, X.Z., Percival Zhang, Y.H., "Ultra-stable phosphoglucose isomerase through immobilisation of cellulose-binding module-tagged thermophilic enzyme on low-cost highcapacity cellulosic adsorbent", *Biotechnology Progress*, 27 (4), 969–975, 2011.
<https://doi.org/10.1002/BTPR.606>
- Oliveira, C., Carvalho, V., Domingues, L., Gama, F.M., "Recombinant CBM-fusion technology - Applications overview", *Biotechnology Advances*, 33 (3-4), 358–369, 2015.
<https://doi.org/10.1016/j.biotechadv.2015.02.006>
- Ostendorp, R., Auerbach, G., Jaenicke, R., "Extremely thermostable L(+)-lactate dehydrogenase from *Thermotoga maritima*: Cloning, characterization, and crystallization of the recombinant enzyme in its tetrameric and octameric State", *Protein Science*, 5 (5), 862–873, 1996.
<https://doi.org/10.1002/pro.5560050508>

- Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A., Ghasemi, Y., "A comprehensive review of signal peptides: Structure, roles, and Applications", *European Journal of Cell Biology*, 97 (6), 422–441, 2018.
<https://doi.org/10.1016/J.EJCB.2018.06.003>
- Pasini, M., Fernández-Castané, A., Jaramillo, A., de Mas, C., Caminal, G., Ferrer, P., "Using promoter libraries to reduce metabolic burden due to plasmid-encoded proteins in recombinant *Escherichia coli*", *New Biotechnology*, 33 (1), 78–90, 2016.
<https://doi.org/10.1016/j.nbt.2015.08.003>
- Phue, J.N., Shiloach, J., "Impact of dissolved oxygen concentration on acetate accumulation and physiology of *E. coli* BL21, evaluating transcription levels of key genes at different dissolved oxygen conditions", *Metabolic Engineering*, 7 (5-6), 353–363, 2005.
<https://doi.org/10.1016/J.YMBEN.2005.06.003>
- Pinhal, S., Ropers, D., Geiselmann, J., de Jong, H., Metcalf, W.W., "Acetate Metabolism and the Inhibition of Bacterial Growth by Acetate", *Journal of Bacteriology*, 201 (13), 147–166, 2019.
- Pinsach, J., de Mas, C., López-Santín, J., Striedner, G., Bayer, K., "Influence of process temperature on recombinant enzyme activity in *Escherichia coli* fed-batch cultures", *Enzyme and Microbial Technology*, 43 (7), 507–512, 2008.
<https://doi.org/10.1016/j.enzmictec.2008.08.007>
- Plamann, M.D., Stauffer, G. v., "Characterization of the *Escherichia coli* gene for serine hydroxymethyltransferase", *Gene*, 22 (1), 9–18, 1983.
[https://doi.org/10.1016/0378-1119\(83\)90059-8](https://doi.org/10.1016/0378-1119(83)90059-8)
- Polisky, B., Bishop, R.J., Gelfandt, D.H., "A plasmid cloning vehicle allowing regulated expression of eukaryotic DNA in bacteria", *Proc. Natl. Acad. Sci. USA*, 73 (11), 3900–3904, 1976.
- Ramesh, H., Nordblad, M., Whittall, J., Woodley, J.M., "Considerations for the Application Of Process Technologies in Laboratory- and Pilot-Scale Biocatalysis for Chemical Synthesis", *Practical Methods for Biocatalysis and Biotransformations 3*, First ed., John Wiley & Sons Ltd., 1–30, 2016. ISBN: 9781118605257.
<https://doi.org/10.1002/9781118697856.ch01>

- Regnier, F.E., Kim, J.H., Narasimhan, M.L., Cho, W., "Affinity-targeting schemes for protein biomarkers", *Proteomic and Metabolomic Approaches to Biomarker Discovery*, Elsevier Inc., 215–245, 2020.
<https://doi.org/10.1016/B978-0-12-818607-7.00013-X>
- Robinson, P.K., "Enzymes: principles and biotechnological Applications", *Essays in Biochemistry*, 59, 1–41, 2015. <https://doi.org/10.1042/BSE0590001>
- Román, R., Lončar, N., Casablanco, A., Fraaije, M.W., Gonzalez, G., "High-level production of industrially relevant oxidases by a two-stage fed-batch approach: overcoming catabolite repression in arabinose-inducible *Escherichia coli* Systems", *Applied Microbiology and Biotechnology*, 104, 5337–5345, 2020.
<https://doi.org/10.1007/s00253-020-10622-y>
- Sanger, F., Nicklen, S., Coulson, A.R., "DNA sequencing with chain-terminating inhibitors", *Proc. Natl. Acad. Sci. USA*, 74 (12), 5463–5467, 1977.
- Sasson, A., "Medical biotechnology: Achievements, prospects and perceptions", *United Nations University*, 1-17, 2005. 154. ISBN: 92-808-1114-2
- Schleif, R., "AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action", *FEMS Microbiol Rev.*, 34, 779-796, 2010.
<https://doi.org/10.1111/j.1574-6976.2010.00226.x>
- Schumann, W., Ferreira, L.C.S., "Production of recombinant proteins in *Escherichia coli*", *Genetics and Molecular Biology*, 27 (3), 442–453, 2004.
<https://doi.org/10.1590/S1415-47572004000300022>
- Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Svobodová, S., Berka, K., Bazgier, V., Velankar, S., Burley, S.K., Koča, J., Koča, K., Rose, A.S., "Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures", *Nucleic Acids Research*, 49, 432-437, 2021.
<https://doi.org/10.1093/nar/gkab314>

- Shoseyov, O., Shani, Z., Levy, I., "Carbohydrate Binding Modules: Biochemical Properties and Novel Applications", *Microbiology and Molecular Biology Reviews*, 70 (2), 283–295, 2006.
<https://doi.org/10.1128/MMBR.00028-05>
- Silveira, J.M.F.J. da, Borges, I. de C., Buainain, A.M., "Biotecnologia e agricultura: da ciência e tecnologia aos impactos da inovação", *São Paulo em Perspectiva*, 19 (2), 101–114, 2005.
<https://doi.org/10.1590/S0102-88392005000200009>
- Skerra, A., "Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*", *Gene*, 151 (1-2), 131–135, 1994.
[https://doi.org/10.1016/0378-1119\(94\)90643-2](https://doi.org/10.1016/0378-1119(94)90643-2)
- Smith, M.R., Khera, E., Wen, F., "Engineering Novel and Improved Biocatalysts by Cell Surface Display", *Industrial & Engineering Chemistry Research*, 54 (16), 4021–4032, 2015.
<https://doi.org/10.1021/IE504071F>
- Solé, J., Brummund, J., Caminal, G., Schürman, M., Álvaro, G., Guillén, M., "Trimethyl- ϵ -caprolactone synthesis with a novel immobilized glucose dehydrogenase and an immobilized thermostable cyclohexanone monooxygenase", *Applied Catalysis A, General*, 585, 117187, 2019 (a).
<https://doi.org/10.1016/J.APCATA.2019.117187>
- Solé, J., Brummund, J., Caminal, G., Schürman, M., Álvaro, G., Guillén, M., "Ketoisophorone Synthesis with an Immobilized Alcohol Dehydrogenase", *ChemCatChem*, 11 (19), 4862–4870, 2019 (b).
<https://doi.org/10.1002/CCTC.201901090>
- Solé, J., Caminal, G., Schürmann, M., Álvaro, G., Guillén, M., "Co-immobilisation of P450 BM3 and glucose dehydrogenase on different supports for application as a self-sufficient oxidative biocatalyst", *Journal of Chemical Technology & Biotechnology*, 94 (1), 244–255, 2019 (c).
<https://doi.org/10.1002/JCTB.5770>
- Stover, P., Zamora, M., Shostak, K., Gautam-Basak, M., Schirchs, V., "Escherichia coli Serine Hydroxymethyltransferase: The role of histidine 228 in determining reaction specificity", *The Journal of Biological Chemistry*, 267 (25), 17679–11687, 1992.
[https://doi.org/10.1016/S0021-9258\(19\)37096-6](https://doi.org/10.1016/S0021-9258(19)37096-6)

- Studier, F.W., Moffatt, B.A., "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes", *Journal of Molecular Biology*, 189 (1), 113–130, 1986.
[https://doi.org/10.1016/0022-2836\(86\)90385-2](https://doi.org/10.1016/0022-2836(86)90385-2)
- Sugimoto, N., Igarashi, K., Samejima, M., "Cellulose affinity purification of fusion proteins tagged with fungal family 1 cellulose-binding domain", *Protein Expression and Purification*, 82, 290–296, 2012.
<https://doi.org/10.1016/j.pep.2012.01.007>
- Terpe, K., "Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial Systems", *Applied Microbiology and Biotechnology*, 72 (2), 211–222, 2006.
<https://doi.org/10.1007/s00253-006-0465-8>
- Tischer, W., Wedekind, F., "Immobilized Enzymes: Methods and Applications", *Biocatalysis - From Discovery to Application*, Topics in Current Chemistry, Springer, 95-126, 1999. ISBN: 978-3-540-68116-8.
https://doi.org/10.1007/3-540-68116-7_4
- Tomme, P., van Tilbeurgh, H., Pettersson, G., van Damme, J., Vandekerckhove, J., Knowles, J., Teer, T., Claeysens, M., "Studies of the cellulolytic system of *Trichoderma reesei* QM 9414 Analysis of domain function in two cellobiohydrolases by limited proteolysis", *Eur. J. Biochem.*, 170, 575–581, 1988.
- Tramper, J., Battershill, C., Brandenburg, W., Burgess, G., Hill, R., Luiten, E., Müller, W., Osinga, R., Rorrer, G., Tredici, M., Uriz, M., Wright, P., Wijffels, R., "What to do in marine biotechnology?", *Biomolecular Engineering*, 20 (4-6), 467–471, 2003.
[https://doi.org/10.1016/S1389-0344\(03\)00077-7](https://doi.org/10.1016/S1389-0344(03)00077-7)
- Tripathi, N.K., "Production and Purification of Recombinant Proteins from *Escherichia coli*", *ChemBioEng Reviews*, 3 (3), 116–133, 2016.
<https://doi.org/10.1002/CBEN.201600002>
- Urh, M., Simpson, D., Zhao, K., "Affinity Chromatography: General Methods", *Methods in Enzymology*, 463, 417–438, 2009.
[https://doi.org/10.1016/S0076-6879\(09\)63026-3](https://doi.org/10.1016/S0076-6879(09)63026-3)

- Vidal Conde, Luis., "Producción de aldolasas recombinantes de la biología molecular al desarrollo de procesos", 2007. TDX (Tesis Doctorals en Xarxa). Universitat Autònoma de Barcelona.
- Vidal, L., Pinsach, J., Striedner, G., Caminal, G., Ferrer, P., "Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*", *Journal of Biotechnology*, 134 (1-2), 127–136, 2008.
<https://doi.org/10.1016/j.jbiotec.2008.01.011>
- Villarejo, M.R., Zabin, I., " β Galactosidase from termination and deletion mutant strains", *Journal of Bacteriology*, 120 (1), 466–474, 1974.
<https://doi.org/10.1128/JB.120.1.466-474.1974>
- Waegeman, H., Soetaert, W., "Increasing recombinant protein production in *Escherichia coli* through metabolic and genetic engineering", *J. Ind. Microbiol. Biotechnol.*, 38 (8), 1891–1910, 2011.
<https://doi.org/10.1007/s10295-011-1034-4>
- Wan, W., Wang, D., Gao, X., Hong, J., "Expression of family 3 cellulose-binding module (CBM3) as an affinity tag for recombinant proteins in yeast", *Applied Microbiology and Biotechnology*, 91 (3), 789–798, 2011.
<https://doi.org/10.1007/s00253-011-3373-5>
- Wang, S., Cui, G.-Z., Song, X.-F., Feng, Y., Cui, Q., "Efficiency and stability enhancement of cisepoxysuccinic acid hydrolase by fusion with a carbohydrate binding module and immobilisation onto cellulose", *Appl. Biochem. Biotechnol.*, 168 (3), 708–717, 2012.
<https://doi.org/10.1007/S12010-012-9811-8>
- "White Biotechnology Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2021-2026", IMARC Group, 2021.
- Whitford, W.G., Lundgren, M., Fairbank, A., "Cell Culture Media in Bioprocessing", *Biopharmaceutical Processing*, 147–162, 2018.
<https://doi.org/10.1016/B978-0-08-100623-8.00008-6>
- Wingard, L.B., Katzir, E., Goldstein, L., "Immobilized enzyme principles", *Applied Biochemistry and Bioengineering*, 1st ed., Academic Press Inc., 1, 1976. ISBN: 9781483215846

- Wu, S., Snajdrova, R., Moore, J.C., Baldenius, K., Bornscheuer, U.T., "Biocatalysis: Enzymatic Synthesis for Industrial Applications", *Angew. Chem. Int. Ed.*, 60, 88–119, 2021.
<https://doi.org/10.1002/anie.202006648>
- Ye, X., Zhu, Z., Zhang, C., Zhang, Y.H.P., "Fusion of a family 9 cellulose-binding module improves catalytic potential of *Clostridium thermocellum* cellodextrin phosphorylase on insoluble cellulose", *Applied Microbiology and Biotechnology*, 92 (3), 551–560, 2011.
<https://doi.org/10.1007/s00253-011-3346-8>
- YH, Z., J, C., LR, L., LR, K., "A transition from cellulose swelling to cellulose dissolution by phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure", *Biomacromolecules*, 7 (2), 644–648, 2006.
<https://doi.org/10.1021/BM050799C>
- Yin, J., Li, G., Ren, X., Herrler, G., "Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes", *Journal of Biotechnology*, 127, 335–347, 2007.
<https://doi.org/10.1016/j.jbiotec.2006.07.01>

AGRAÏMENTS

Heus aquí la conclusió d'aquest treball, en què he intentat englobar bona part de la feina feta els darrers anys de la meua vida. Amb ell poso punt i final —o potser punt i apart, mai se sap— a una de les etapes més extenses de la meua curta existència; la carrera universitària. Han sigut deu anys increïbles, dels quals en faig un balanç descaradament positiu i que sempre recordaré amb un somriure sincer. Aquesta etapa es tanca per donar pas a una de nova i de moment desconeguda, que pràcticament coincideix amb l'inici d'un altre gran canvi a nivell personal que portava temps trucant a la porta i que a dia d'avui ja és tota una realitat. Acabo el doctorat just després d'haver volat del niu, com se sol dir. D'encetar la vida en parella i començar a llaurar la terra on espero que germinin algun dia les llavors d'un nou nucli familiar, juntament amb la meua estimada Clàudia i les nostres "ratetes", altrament conegudes com Otto i Lila.

Desitjo que aquestes properes —i últimes— línies serveixin doncs d'agraïment a totes aquelles persones amb qui m'he topat durant aquests darrers anys i que, en major o menor grau, han contribuït al fet que assolixi el màxim grau universitari amb aquest treball. Tot i que m'agradaria ser breu, tampoc em vull deixar a ningú —és realment just i necessari—, i si algú que ho llegeix no s'hi troba explícitament anomenat, si us plau, se senti inclòs. Fent mostra del meu "frikisme", diré que no us ho podré agrair ni la meitat del que jo voldria, i el que jo voldria és menys de la meitat del que la meitat de vosaltres mereixeríeu. Poc més cal dir a qui ho hagi entès.

En primer lloc, he de donar les gràcies als meus directors/tutors, que m'han acompanyat i guiat tot aquests anys, i que són font inesgotable d'experiència i bons consells. A la Dra. Gloria González, modesta i humil com ella sola, però que val or. A la Dra. Marina Guillén, que no només m'ha guiat durant la tesi sinó en tots els anys en què he format part del grup, i que és la millor directora que un pot tenir; si la Gloria és or, ella també, com a mínim. També a la Dra. Glòria Caminal, que tot i no constar en aquest elenc de directores ha exercit com a tal, aportant una perspectiva més allunyada i potser més directa, molt necessària. Sento una admiració immensa per vosaltres, que sou tres estandards com la copa d'un pi, exemple clar del pes que les dones tenen (i han de tenir sempre) en el món acadèmic i de la investigació.

He de donar les gràcies també al Dr. Ramón Román, tot i que encara no he acabat de treure en clar què carai ha aportat, per bé que firma com a director de la tesi. Imagino que alguna cosa haurà fet, sinó de què... Bromas aparte, Ramón, me siento realmente afortunado de que nos hayamos cruzado en el camino y he haber compartido contigo varios años en la planta. Muchas gracias por todo lo que me has ayudado tirando adelante los experimentos más peliagudos de la tesis, en los que al principio —y al final también— iba más perdido que un pulpo en un garaje. Gracias por el buen rollo y por lo poco que se ha notado lo muy por encima que, de momento, estás de mí. Supongo que el roce hace el cariño y el hecho de que tu primer hijo lleve mi nombre me alegra a la par que me alivia; será que ni soy tan malo ni soy tan feo. Que podamos reírnos por muchos años más del maravilloso circo de tristes payasos llamado F.C.B. que tantas alegrías y tantas risas a su costa nos ha regalado estos últimos tiempos.

Vull donar les gràcies al meravellós grup d'Aldolases i en general a tota la família "asera", sumant la bona gent de Lipases. Gràcies al Dr. Gregorio Álvaro per obrir-me les portes fa tants anys i per haver-me deixat quedar fins avui, i sobretot als companys de laboratori que he tingut. Gràcies als veterans Gerard, Luismi, Jordi i Natàlia, pels bons consells i l'acollida al grup, i sobretot gràcies al Miquel García, amb qui probablement més temps han coexistit les nostres tesis, merci de tot cor per tot el que m'has ajudat i pels bons moments que hem passat, has sigut un excel·lent company, tan de lab com de despatx. Eskerrik asko a mi otro compañero de despacho Josu, felizmente doctorado hace poco. Espero que als dos us vagi genial i que la vida us tracti tan bé com tracteu vosaltres a qui us envolten.

Vull fer menció expressa a la Garazi i a la Nicole, alumnes de pràctiques que m'han donat la vida. Eskerrik asko Garazi y mil gracias Nicole, me habéis ayudado más de lo que os podéis imaginar a tirar adelante muchos de los experimentos que se recogen en este trabajo. Por no hablar de lo simpatiquísimas y trabajadoras que sois. Me alegro mucho de que hayamos coincidido y espero que hayáis aprendido de mí lo que no se debe hacer en un laboratorio. També vull donar les gràcies a la Raquel per l'ajuda rebuda durant la seva estança al grup.

Com se sol dir, moltíssims ànims i força a tots aquells que esteu a mitges: Albert(s), Arnau, Núria, Sílvia, Quique... sou molt bona gent. Muchísi-

mos ànimos també a Carolina, Diego, Yerko, Roberto, espero que todo os vaja genial. Pot quedar lleig que sigui jo qui ho digui, però Lipases-Al-dolases som el millor grup del Departament amb diferència. Gràcies a tots pel bon rotllo i la vostra simpatia, també de tots aquells que hi eren quan vaig arribar: Kírian, Javi, Juanjo, Xavi, Sergi...

Gràcies també a la resta de persones del grup i de tot el Departament amb les que he compartit laboratori i feina, com l'Oscar, la Jessica, l'Arnau Boix, el Gerard Guerra, etc., així com a la gent de la Planta Pilot, en especial al Dr. Antoni Casablanca per a deixar-m'hi entrar, tot i ser jo perico i ell culé. Gràcies a la gent que treballa al Departament; als tècnics Pili, Manuel i Rosi, a les noies de secretaria Rosa, Montse, Laura, Alba, Marta... I sobretot gràcies a les dones de la neteja que es presenten cada dia a les set del matí per deixar nets despatxos i laboratoris. En especial gràcies a la Mariluz, que em feia de despertador amb la seva ràdio els dies que "pringava" fent fermentacions, i que ara gaudeix d'una merescudíssima jubilació. I gràcies a la Luisa, por todos esos ratitos en los que parabas en nuestro laboratorio a descansar un poco y a hablar de cualquier cosa, por siempre interesarte por nosotros y por ser una autentica jabata infatigable, espero que pronto te toque descansar y disfrutar de la vida como te mereces.

Ara ja sí que acabo definitivament, donant les gràcies a les dues persones que més m'estimen i sense les quals seria inconcebible estar escrivint aquests paràgrafs d'agraïments; els meus pares Encarna i José Antonio, que s'han sacrificat des que em van portar al món per tal de procurar-me una bona educació i la millor formació acadèmica possible. Tot i que mai a la vida un fill li pot arribar a agrair a uns pares tot el que han fet per ell, us he intentat honorar arribant fins al final. Ho he fet tan bé com he sabut i us torno a donar les gràcies per ser excel·lents en la vostra tasca parental. Us estimo molt, també a tu Pau, a qui apart he de donar les gràcies per dedicar-te a maquetar aquest treball i aportar el teu criteri gràfic/artístic.

A tothom,

Moltes gràcies.