UNIVERSITAT AUTÒNOMA DE BARCELONA

ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA

DEPARTAMENT D'ENGINYERIA QUÍMICA

DEVELOPMENT OF RECOMBINANT ALDOLASE PRODUCTION PROCESS IN *Escherichia coli*.

Memòria que per optar al grau de Doctor

per la Universitat Autònoma de Barcelona presenta

Jaume Pinsach i Boada

Bellaterra, 2009

JOSEP LÓPEZ SANTÍN, Catedràtic del Departament d'Enginyeria Química de la Universitat Autònoma de Barcelona, i CARLES DE MAS ROCABAYERA, Professor Titular del Departament d'Enginyeria Química de la Universitat Autònoma de Barcelona,

CERTIFIQUEM:

Que el titulat en Enginyeria Química Jaume Pinsach i Boada ha dut a terme sota la nostra direcció el treball amb el nom "Development of recombinant aldolase production process in *Escherichia coli*" que presenta en aquesta memòria i que constitueix la seva Tesi per optar al grau de Doctor per la Universitat Autònoma de Barcelona.

I perquè es prengui coneixement i consti als efectes oportuns, presentem a l'Escola Tècnica Superior d'Enginyeria de la Universitat Autònoma de Barcelona l'esmentada Tesi, signant el present certificat a

Bellaterra, febrer del 2009

Dr. Josep López Santín   Dr. Carles de Mas Rocabayera

Els homes sovint oblidem que la felicitat
és una disposició de la ment,
no una condició de les circumstàncies.

John Locke

A l'Anna

Vull donar les gràcies a en Josep i a en Carles no només per donar-me l'oportunitat de realitzar aquest treball, guiar-me i donar-me suport al llarg de tot aquest temps, sinó també pel seu tracte, la seva manera de fer i la seva confiança. Abans de començar amb prou feines ens coneixíem, i ara puc dir que em sento afortunat pel fet d'haver compartit tot això amb vosaltres.

Agraïments també a tota la resta de companys i companyes "aldolaseros" per la vostra ajuda, per la vostra paciència i per tot el que hem fet i viscut plegats: a la Trinitat, a en Luis, a en Goyo, a la Dolors, a la Inés, a les Glòries, a en JuanMi, a en Sergio, a en Jordi, a en Pau, a la Cristina, a la Lorena, a la Noelia, a l'Engràcia... ens ho hem passat molt bé! Agraïments especials també per l'Oriol i en Carles Casas, que tot i no tenir relació amb el treball m'han ajudat molt...

A la resta de gent del Departament, perquè penso que aquesta etapa la recordaré per com de bé ens ho hem passat, per l'ambient que hem viscut, per la quantitat de coses que hem fet (a part de la feina, es clar!). Tot això ha fet que companys i companyes passin a ser amics i amigues...

I'd like to thank Prof. Modak, Kakasaheb, Saurabh, Tukaram, Shryia, Girija, Vishal, Vijay, Vishu, JP, Ganesh, Rachna, Indraneel and all the rest of the people in IISc Bangalore not only for your time and help, but also for your friendship. I learned many things from you: some of them about work, but also others still more valuable... it was a great experience for me! Unluckily, I had not enough information about the process to calibrate and optimise a robust production model, but I learned many things about modelling and optimisation strategies over there. My apologies because I have not been able to translate all your teachings into results!

I'd like to thank the austrian crew as well for their hospitality and help: Prof. Bayer, Gerald, Markus, Gerd, Sabine, Xu-Bin, Florian, Thomas, Karo, Sven, Norbert and all the rest of the people in BOKU's Biotechnology Department.

Als amics, tant de Colomers com d'Olot, per la seva comprensió. Finalment la tesi s'ha acabat, o sigui que espero poder dedicar-vos més temps a tots i a totes. Com diuen, la veritable amistat no depèn ni de l'espai ni del temps...

Vull donar les gràcies de manera especial a la meva família, que sempre m'ha recolzat i animat a tirar endavant qualsevol projecte que m'ha passat pel cap. La vostra confiança cega en mi i en les meves il·lusions és quelcom que no té preu. Em sento realment afortunat... I a la família d'Olot, que també ens ha animat i ajudat durant tota aquesta etapa tant a l'Anna com a mi.

Finalment, vull donar les gràcies a l'Anna. Per tot el que hem viscut, per tot el que hem compartit, pel que ens hem estimat. Podria agrair-te moltes altres coses, però el més important és que amb tu he après què és realment important a la vida. Això tot just acaba de començar i ens queda un llarg trajecte per recórrer, moltes il·lusions per compartir...

TABLE OF CONTENTS.

SUMMARY.

This work has been focused on the development of recombinant aldolase production process in *Escherichia coli*.

On the one hand, the production process of recombinant fuculose 1-phosphate aldolase (FucA) was automated when using an expression system based on an inducible weak promoter. Since pulse induction did not significantly affect the host cell metabolism until high product yields were reached, a robust process control could be implemented by automating an exponential substrate feed. High cell density cultures were obtained at limiting substrate levels by controlling the specific growth rate. The implemented control algorithm used a simple feedback loop based on indirect biomass estimation (using exhaust gas analysis) and mass balances.

On the other hand, the production process of recombinant rhamnulose 1-phosphate aldolase (RhuA) using an expression system based on an inducible strong promoter was developed. Results obtained from preliminary expression studies performed in shake flask cultures indicated that the use of a strong promoter presented important advantages. When compared to the expression system based on a weak promoter, high product yields were obtained in shorter times, and lower amounts of inducer were required. However, when going to high cell density cultures, the system was found to be extremely sensitive to induction. Thus, special attention was paid on the development of suitable process strategies.

First, process development was focused on the implementation of alternative induction strategies in high cell density cultures. While high intracellular specific activities were achieved in shake flasks, lower specific values were measured in standard pulse-induced fed-batch cultures. The conditions under which high metabolic load on host cells compromised the productivity of the process were identified and avoided by implementing a continuous inducer addition strategy. When tuning protein expression rates, the length of the production phase was extended, the requirements of inducer were decreased and high product yields were achieved. However, results still showed lower specific enzyme activity than in the case of preliminary expression studies.

Since decreased enzyme activity has been often attributed to temperature-driven stress (translated into either protein misfolding and/or proteolysis itself), optimisation of process temperature was carried out. Lower specific production rates were measured at a reduced process temperature, but significant improvement was achieved in terms of bioactive RhuA yields.

Besides, preliminary studies on process modelling were carried out. Even though it was not possible to calibrate a production model robust enough to describe all situations which could potentially occur (and, thus, mathematical optimisation of the process strategy would not be reliable), insight of the process was gained, and some of the process bottlenecks and unknown key features were identified.

From these learnings, an alternative growth strategy was then developed to minimise the starvation induced-stress on host cells. High cell density cultures were obtained at non-limiting substrate levels by controlling the glucose concentration in the culture at inhibiting values to avoid acetate accumulation. After optimising the induction strategy, high product yields as well as high specific enzyme activities were obtained.

Finally, an analysis of the global production process was performed. Since the aim of the whole process was to obtain an immobilised enzyme ready to be used as biocatalyst, all different production strategies were compared taking into account their impact on downstream yields. When doing so, results showed that the alternative process strategies which had been developed in this work significantly increased the immobilised specific yields of active RhuA with respect to the previous strategy. These results reinforced the need to optimise recombinant protein production processes considering both the production and downstream stages as a whole.

ABBREVIATIONS.

Ac	acetate concentration ($g \cdot L^{-1}$).
AC	activity concentration ($AU \cdot L^{-1}$).
ATP	adenosine-5'-triphosphate.
AU	activity units ($\mu mol \cdot min^{-1}$ at 25ºC).
CAP	catabolite activator protein.
cAMP	cyclic adenosine monophosphate.
CER	carbon dioxide evolution rate (mmol $CO_2 \cdot L^{-1} \cdot h^{-1}$).
CFU	colony forming units.
CoASH	coenzyme A.
D	dilution rate (h^{-1}).
DC	dead cells.
DCU	digital control unit.
DHAP	dihydroxyacetone phosphate.
DNA	deoxyribonucleic acid.
EDTA	ethylenediaminetetraacetic acid.
$f_{gDCW/OD600}$	optical density at 600 nm to dry cell weight conversion factor.
$f_{I/X}$	IPTG to biomass ratio ($g \cdot g^{-1}$).
F_I	inducer feed rate ($L \cdot h^{-1}$).
F_S	substrate feed rate ($L \cdot h^{-1}$).
FSC	forward side scatter.
fucA	gene coding for FucA.
FucA	fuculose 1-phosphate aldolase.
F_V	flow rates other than substrate or inducer feeds (base addition, phosphate addition, sample taking,...) ($L \cdot h^{-1}$).
Gal	galactose concentration ($g \cdot L^{-1}$).
Gal_0	galactose concentration at inoculation time ($g \cdot L^{-1}$).
gDCW	dry cell weight (g).
G_e	inlet molar gas flow rate ($mol \cdot h^{-1}$).
glyA	gene coding for GlyA/SHMT.
GlyA	glycine hydroxymethyltransferase.
G_s	outlet molar gas flow rate ($mol \cdot h^{-1}$).
HPLC	high performance liquid cromatography.
I	inducer concentration ($g \cdot L^{-1}$ or $\mu mol \cdot L^{-1}$).
I_0	inducer concentration at inoculation time ($g \cdot L^{-1}$ or $\mu mol \cdot L^{-1}$).
I_F	IPTG concentration in inducer feed ($g \cdot L^{-1}$ or $\mu mol \cdot L^{-1}$).
IPTG	isopropyl-β-D-thiogalactopyranoside.

K_S	saturation constant for the growth kinetics on glucose ($g \cdot L^{-1}$).
K_{SI}	substrate inhibition constant ($g \cdot L^{-1}$).
m_1	maintenance coefficient for X_1 ($g \cdot gDCW^{-1} \cdot h^{-1}$).
m_2	maintenance coefficient for X_2 ($g \cdot gDCW^{-1} \cdot h^{-1}$).
MB	microburette.
m_{CO2X}	CO_2 maintenance coefficient ($mmol\ CO_2 \cdot g\ DCW^{-1} \cdot h^{-1}$).
MCS	multiple cloning site.
MFM	mass flow meter.
mRNA	messenger ribonucleic acid.
MS	mass spectrometer.
m_{SX}	maintenance coefficient ($g \cdot gDCW^{-1} \cdot h^{-1}$).
N	strirring speed (rpm).
NAD^+	nicotinamide adenine dinucleotide.
OD600	optical density at 600 nm.
OUR	oxygen uptake rate ($mmol\ O_2 \cdot L^{-1} \cdot h^{-1}$).
P	product concentration ($mg \cdot L^{-1}$).
P_0	product concentration at inoculation time ($mg \cdot L^{-1}$).
PCN	plasmid copy number ($plasmid \cdot cell^{-1}$).
pO2	oxygen partial pressure (%).
PTS	phosphotransferase system.
q_{Ac}	specific acetate accumulation rate ($g \cdot gDCW^{-1} \cdot h^{-1}$).
q_{CO2}	specific CO_2 production rate ($mmol\ CO_2 \cdot g\ DCW^{-1} \cdot h^{-1}$).
Q_e	continuous liquid inlet flow rate ($L \cdot h^{-1}$).
q_P	specific production rate ($AU \cdot gDCW^{-1} \cdot h^{-1}$).
q_{P0}	basal specific production rate ($AU \cdot gDCW^{-1} \cdot h^{-1}$).
q_{Pmax}	maximum specific production rate ($AU \cdot gDCW^{-1} \cdot h^{-1}$).
Q_s	continuous liquid outlet flow rate ($L \cdot h^{-1}$).
q_S	specific substrate consumption rate ($g \cdot gDCW^{-1} \cdot h^{-1}$).
rhaD	gene coding for RhuA.
RhuA	rhamnulose 1-phosphate aldolase.
RNA	ribonucleic acid.
S	substrate concentration ($g \cdot L^{-1}$).
S_0	substrate concentration at inoculation time ($g \cdot L^{-1}$).
S_F	substrate feed concentration ($g \cdot L^{-1}$).
SHMT	serine hydroxymethyltransferase.
SSC	side scatter signal.
t	time (h).
t_0	inoculation time (h).
T	temperature (ºC).

TCA	tricarboxylic acid cycle.
TCN	total cell number.
V	culture volume (L).
V_0	culture volume at inoculation time (L).
V_a	volume of acid added into the bioreactor (L).
V_{af}	volume of antifoam added into the bioreactor (L).
V_{ad}	feed volume to be added to the fermentor in a time interval (L).
V_b	volume of base added into the bioreactor (L).
VBNC	viable but not culturable.
X	total cell concentration (gDCW·L^{-1}).
X_0	total cell concentration at inoculation time (gDCW·L^{-1}).
X_1	viable and culturable cell concentration (gDCW·L^{-1}).
X_2	viable but not culturable cell concentration (gDCW·L^{-1}).
$x_{CO2,e}$	carbon dioxide molar fraction in inlet gas.
$x_{CO2,s}$	carbon dioxide molar fraction in exhaust gas.
$x_{O2,e}$	oxygen molar fraction in inlet gas.
$x_{O2,s}$	oxygen molar fraction in exhaust gas.
Y_{XCO2}	CO_2 yield (g DCW·mmol CO_2^{-1}).
Y_{XP}	phosphorus to biomass yield (gDCW·g^{-1}).
Y_{XS}	substrate to biomass yield (gDCW·g^{-1}).

Greek letters:

α	segregation rate from X_1 to X_2 (h^{-1}).
Δt	time interval (h).
μ	specific growth rate (h^{-1}).
μ_{fix}	specific growth rate set point value (h^{-1}).
μ_S	specific growth rate due to substrate consumption (h^{-1}).
μ_1	specific growth rate of X_1 (h^{-1}).

PREFACE.

Several aspects of the production process of recombinant proteins in *Escherichia coli* were studied in this work. Each one of them has been presented in an independent chapter, but there is an obvious common aim throughout them. Therefore, the reader might find some redundant information (including some topics of introduction, figures and references). However, this information has been specifically included to facilitate the reading of each chapter independently.

In chapter 1, general introduction, an overview of recombinant protein production in *E. coli* and aldolases is presented.

In chapter 2, context and objectives, the aims of the present work are derived from previous work done within the research group.

In chapter 3, common materials and methods used throughout the work are described. However, specific materials and methods can be also found in the correspondent chapters.

In chapter 4, the control and automation of fuculose 1-phosphate aldolase (FucA) production process using an expression system based on an inducible "weak" promoter is presented.

In chapter 5, alternative induction strategies to overexpress rhamnulose 1-phosphate aldolase (RhuA) using an expression system based on an inducible "strong" promoter are implemented and compared.

In chapter 6, the effects of process temperature on RhuA yields and bioactivity are studied.

In chapter 7, preliminary work on RhuA production process modelling is presented. A model describing cell growth is calibrated and, although a robust model for protein production is not presented, the most significant features of the production process are studied.

Finally, in chapter 8, an analysis of the global yields of alternative process strategies taking into account both production as well as downstream stages is presented.

1. GENERAL INTRODUCTION.

Biotechnology is defined by the United Nations as "any technological application that uses biological systems, living organisms or derivatives thereof, to make or modify products or processes for specific use". Based on its application, three different branches are distinguished: red (medical) biotechnology, green (agricultural) biotechnology and white (industrial) biotechnology.

The production of enzymes as are aldolases using microorganisms can be included into the white branch of biotechnology. This branch is currently promoted for the sustainable economic future of modern industrialised societies, but development of novel enzymes, processes, products and applications is required to fulfil these objectives [1]. White biotechnology is expected to have a positive impact not only on economy (especially on the chemical industry), but also on environment and on society, because the driving force behind it is sustainability. It drives to clean processes that improve manufacture of chemical products by cutting down waste production and lowering energy consumption. The main applications of white biotechnology are focused on replacing fossil fuels with renewable resources, replacing conventional processes with bioprocesses and creating new high-value bioproducts. Global biotechnology market reports indicate that the biotech market was worth over 171 billion $ in 2007 (http://www.fiercebiotech.com/research/biotechnology-global-industry-guide), and it is expected to be one important pillar of european economy, increasingly used for industrial processing, pharmaceuticals, agriculture and food. The products of white biotechnology and bioenergy are expected to represent one-third share of industrial production by 2030, reaching an estimated value of 300 billion € (http://www.europabio.org/articles/cologne_paper.pdf).

Amongst bioproducts, enzymes are used in a wide range of applications and industries [2]. They are required to synthesize kilograms of stereochemically chiral synthons that are used as building blocks to produce highly active pharmaceuticals, and at a bigger scale as active ingredients for bulk products such as high performance laundry detergents. Their versatility allows their use in many applications, including processes to degrade natural polymers (such as starch, cellulose and proteins) as well as for the regioselective or enantioselective synthesis of asymmetric chemicals. The aldolases produced in this work can be included into the latter group.

In the future, novel biotechnological applications should boost the market for industrial enzymes. Owing to a wealth of potentially useful biocatalysts, biotechnological applications of microbial resources are increasingly popular with the chemical industries and are viewed as indispensable for the modern organic chemist [3]. In the face of increasing energy costs,

scarcity of fossil resources, environmental pollution and a globalised economy, the large-scale use of biotechnology instead of, or to complement, traditional industrial production processes is viewed as both an opportunity and a necessity. With many successfully implemented processes operating worldwide and the number of industrial biotransformation having doubled every decade since 1960 [4], unavailability of an appropriate biocatalyst is thought to be a limiting factor for any biotransformation process [5]. Therefore, the development of robust processes for their production becomes of major importance.

1.1. **Recombinant protein production in *E. coli*.**

Demands of the expanding biotechnology industry have driven to different improvements in protein expression technology, which have been translated into the production of a spectrum of recombinant proteins in different systems for a wide variety of purposes. Most of the recombinant proteins are nowadays produced either in bacteria, yeasts, engineered animal cell lines, hybridoma cells or even human cells. However, research continues on the development of alternative production systems, particularly in the use of transgenic animals or plants [6].

In the recent years, baculovirus and mammalian cell cultures have gained importance for the production of biopharmaceuticals due to the increasing needs of complex proteins and antibodies. An alternative to baculovirus or mammalian expression systems are yeasts, especially when large amounts of secreted protein are required [7]. If the product contains post-translational modifications, *Saccharomyces cerevisiae* or *Pichia pastoris* may offer an economic alternative because they can grow to high cell densities using minimal media [8], but the use of yeasts is limited due to their inability to modify proteins with human glycosylation structures for the production of therapeutic glycoproteins [9]. The use of transgenic plants and animals as production vehicles may also play a role in applications requiring exceptional product volumes, but regulatory issues still remain to be addressed.

Even though the choice of expression system is progressively widening, *E. coli* is still the dominant host for recombinant protein production. It is used in many industry fields to produce high value intermediates, detergents, nutraceuticals and pharmaceuticals, amongst others. It is the most popular choice when simple proteins are required, and significant advances have been also made to overexpress complex proteins, hormones, interferons and interleukins in it [6]. The production of heterologous proteins or parts thereof in different extra-cytoplasmic compartments (in the periplasm, outer membrane or extracellularly) of *Escherichia coli* offers multiple new applications, for example, in vaccine development, immobilised enzymes and bioremediation. Nowadays, not only surface display of short peptides, but also cell-surface anchoring or secretion of functional proteins is possible.

The past 20 years have seen enormous progress in the understanding of the mechanisms used by this enteric bacterium, and factors influencing folding, stability and export of extra-cytoplasmic proteins are also better understood [10]. All this knowledge has been exploited to overcome problems like inclusion body formation, proteolytic degradation and disulfide bond generation that have long impeded the production of complex proteins in a properly folded and biologically active form. The application of this information to industrial processes, together with emerging strategies for creating folding modulators and performing glycosylation guarantee

that *E. coli* will remain as an important host for the production of both commodity and high value added proteins in the future.

1.1.1.   *E. coli* as a host.

Escherichia coli (Figure 1.1) is a Gram-negative, facultative anaerobe and non-sporulating bacterium that is commonly found in the lower intestine of warm-blooded animals. The morphology of cells is rod-shaped, about 2 micrometers long and 0.5 micrometers in diameter. Most *E. coli* strains are harmless, but others can cause serious damage to humans. The harmless strains are part of the normal flora of the gastrointestinal tract, and can benefit their hosts by producing vitamin K2 and by preventing the establishment of pathogenic bacteria within the intestine [11]. However, *E. coli* can be easily grown outside the intestine as well, and it has become a model organism in biotechnology because of the simplicity of its genetics and manipulation.

Figure 1.1. *Escherichia coli* cells electron micrograph (magnified 10.000 times).

Unlike eukaryotes, the bacterial chromosome is not enclosed inside of a membrane-bound nucleus but instead resides inside the bacterial cytoplasm. In there, the chromosome exists as a compact (usually circular) supercoiled structure which can be easily modified. To date, the most popular target for genetic manipulation in *E. coli* has been the modification of host cell metabolism to reduce acetate formation. Acetate is a by-product of *E. coli* metabolism which can be toxic for cells and can inhibit their growth. Strain-level modifications range from acetate synthesis decrease [12, 13] to increase acetate utilisation [14], divert glycolysis to other pathways [15-17] or modify the PTS glucose transport system [18]. Other genetic manipulations have been focused on improved protein folding (which may be achieved by overexpression of intracellular chaperones) [19], or efficient disulfide bond formation and

isomerisation in *E. coli* cytoplasm, allowing the expression of complicated proteins in an active form [20]. All these modifications have driven to the existence of various strains of *Escherichia coli* with different genotypes (genetic constitution) which are used as expression systems. In particular, *E. coli* K12 and *E. coli* B and their many derivatives are the most commonly used hosts [21].

All these features allow *E. coli* to be one of the most competitive hosts for rapid and economical production of simple recombinant proteins, aminoacid and metabolite production when compared to other hosts. These advantages, coupled to the wealth of biochemical and genetic knowledge, have enabled the production of economically sensitive products as insulin and bovine growth hormone. Current commercial products mainly include recombinant proteins which are considered low-volume high-value products [22]. However, recent advancements in metabolic engineering made it possible to use *E. coli* as a platform to produce high-volume low-value products [23]. In all these processes, high cell density and high volumetric yields are essential for economical feasibility. Capital investment and operation costs are reduced due to size reduction of fermentation equipment, upstream utilities, downstream units, etc. In this sense, process development plays an important role to maximise target product yields and to minimise production costs.

1.1.2.   Cell growth.

When using *E. coli* as expression system, high cell density cultures are usually carried out to increase target protein yields. Main factors which affect high density growth of *E. coli* are solubility of solid and gaseous substrates in water (limitation/inhibition of substrates with respect to growth may occur), accumulation of products or metabolic by-products to inhibitory levels, high evolution rates of CO_2 and heat as well as increasing viscosity of the medium [24, 25].

Since this work was developed at bench scale, neither heat transfer or high power consumption due to high viscosity of the medium nor the utilisation of pure oxygen were problems, but *E. coli* growth can be also limited/inhibited by nutritional requirements including carbon, nitrogen, phosphorus, sulphur, magnesium, potassium, iron, manganese, zinc, copper and other growth factors [26]. For high cell density cultures, the nutritional requirements cannot be added initially to the basal media, since most media ingredients become inhibitory to *E. coli* when added at high concentrations. Another problem is the precipitation of media components that can hamper adequate supply, interfere with the fermentation process and monitoring devices and

affect downstream recovery and purification processes. Therefore, a well designed media and feeding strategy are essential to achieve high cell density cultures.

Another motivation to develop different feeding strategies was to decrease acetate accumulation. As explained before, high acetate concentration can inhibit growth and recombinant protein production [27]. The excretion of acetate results from the need to regenerate NAD^+ consumed by glycolysis and to recycle the coenzyme A (CoASH) required to convert pyruvate to acetyl-CoA (Figure 1.2). Since the tricarboxylic acid cycle (TCA) completes the oxidation of acetyl-CoA to carbon dioxide, acetogenesis occurs whenever the full TCA cycle does not operate or when the carbon flux into cells exceeds its capacity [14, 28]. Thus, it occurs anaerobically during mixed-acid fermentation [29], and it is directly related to the rate at which cells grow or the rate at which they consume the usual substrate, glucose. It also occurs aerobically [30] when unlimited growth on excess of glucose inhibits respiration [31, 32], a behaviour called the Crabtree effect [33]. As a consequence of the latter, 15% of the glucose can be excreted as acetate [32]. However, as will be described in chapter 8, acetate accumulation can be also avoided by inhibiting growth when growing on excess of glucose.

Nevertheless, acetate is only partially oxidised. Thus, it is a potential source of both carbon and energy which can be assimilated through the glyoxylate shunt [34]. Therefore, the ability of *E. coli* to remove this potential toxin from its environment by consuming it permits to solve this problem under specific conditions [35, 36].

Total acetate concentration in the medium is the net balance of production and consumption by microorganisms, being able to influence growth, cell concentration and the level of protein produced. Various genetic and biochemical engineering or combined strategies have been tried to avoid or reduce acetate formation [14, 37-39]. As explained before, molecular biology is a powerful technique to overcome acetate accumulation, but acetate accumulation can be also avoided in stirred tank reactors by means of culture medium design and methods for carbon source feeding. From a process engineering point of view, there are two principal strategies for the control of nutrient feed: the open-loop and the closed-loop (feedback) control.

Open-loop fed-batch methods are based on mathematical models that describe growth patterns and the expected demand for nutrients. Using such models, an optimal concentration of all components and feed rate can be calculated. In particular, limiting carbon source is one of the most popular strategies to control cell growth under aerobic conditions. The specific growth rate [27] and specific glucose uptake [40] can be controlled under critical values, avoiding acetate accumulation. These equations are used to pre-determine linear or exponential addition rates.
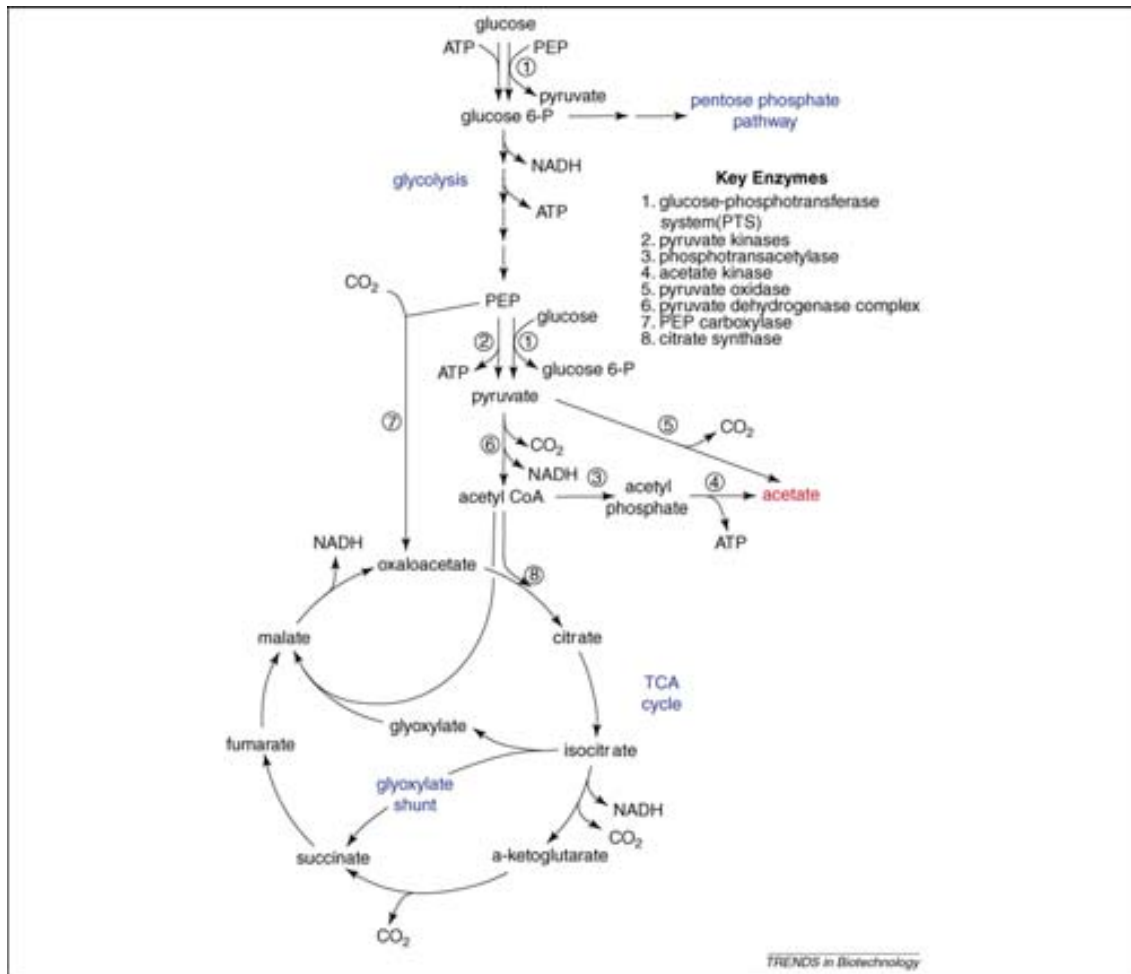
Figure 1.2. Key biochemical pathways in *E. coli* involved in the aerobic consumption of glucose and the synthesis of acetate.

The second feeding approach is to use direct or indirect feedback control techniques. Direct feedback is based on monitoring the concentration of the substrate for controlled addition of nutrients. Indirect feedback control is based on monitoring parameters such as pH, dissolved oxygen, carbon dioxide exhaust rate and cell concentration (optical density, NAD^+, ATP, ...). Special attention must be paid to high cell density cultures because of the lack of sensors for fast, reliable and robust measurement of essential control variables. Moreover, the time-variant dynamics of fed-batch processes and the complexity and non-linearity of the intracellular and extracellular dynamics complicates the process control.

Both open-loop as well as feedback control techniques have been used in this work. The choice of the nutrient feeding strategy has been done depending on the expression system, since not any strategy is suitable in any case. Process optimisation performed in present work required the development of different nutrient feeding strategies depending on both the expression system and target protein which were considered.

1.1.3.   Induction of recombinant protein expression.

Along with chromosomal DNA, most bacteria also contain small independent pieces of DNA called plasmids which are capable of replicating independently of the chromosome (Figure 1.3). Plasmids (or expression vectors) often encode for genes that are advantageous but not essential to their bacterial host, and can be easily gained, lost or transferred by cells. These properties make them basic tools in biotechnology: they are used to multiply particular genes or to make large amounts of proteins. In the latter case, bacteria containing a plasmid harbouring the gene of interest are grown and, then, transcription of the inserted gene can be induced to produce the target protein.
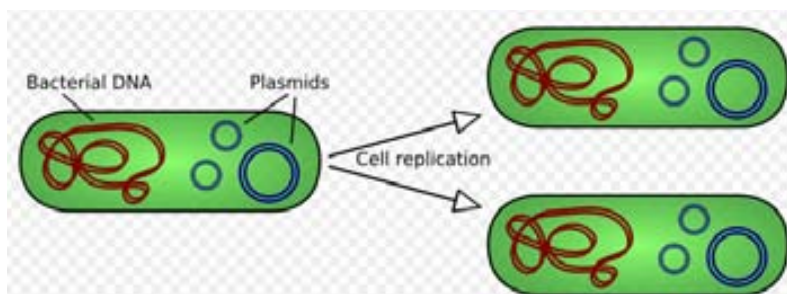

Figure 1.3. Plasmid replication scheme.

In order to achieve these goals, several expression vectors have been designed. Amongst other features, they usually contain an origin of replication that allows replication of the plasmid in the host, a multiple cloning site where target DNA can be inserted, a promoter used to drive transcription of the target gene, and antibiotic resistance that allows for identification of which cells have uptaken the vector through selection pressure. Gene dosage (plasmid copy number), promoter strength, mRNA stability and the efficiency of the translation initiation are some of the plasmid features which may affect target protein yields.

A wide variety of plasmids are now available commercially, most of them using moderate-to-high copy numbers. These can drive rapid expression, but usually require selection pressure, for example, by antibiotics. As alternatives, low copy number plasmids as well as integration of the expression cassette into the chromosome have been successfully used. The first group is useful when inclusion body formation is desired, but the second one is preferable when the product is desired either soluble and active or secreted. In the latter case, high product yields will depend upon sustaining the production period.

In order to control recombinant protein expression in *E. coli*, several inducible promoters have been developed. When using them, induction of the target gene transcription is controlled by a signal which will depend on the type of promoter. Chemical induction is the most common

strategy (being the signal the presence or absence of a particular molecule), but other induction strategies like the thermal one have been also successfully applied. Idealistically, the promoter should not allow expression before induction, and should allow adjustable induction of protein expression. The arabinose promoter system (*araB*) comes closest to fulfil these objectives, but in large scale production, inexpensive and automatic induction is preferred. Then, the *phoA* (alkaline phosphatase) promoter regulated by phosphate levels in the medium can be used. Cold induced promoters might be also considered when improved folding in the cytoplasm is required [41]. However, the most common promoters are the *lac*-derived ones, since the expression of the recombinant protein can be easily induced by either lactose or its synthetic analogue isopropyl-β-D-thiogalactopyranoside (IPTG).

The *lac operon* is required for the transport and metabolism of lactose in *E. coli* and some other bacteria. Regulation of the *E. coli* lactose operon (Figure 1.4) is one of the best studied areas in molecular biology and this knowledge has been used to construct several different *lac*-type promoters [42, 43]. Most of these systems contain:

- a *lacI* gene encoding the *lac* repressor, which prevents transcription of the three structural genes by binding to the *lac* operator region situated next to the promoter.
- a catabolite activator protein (CAP), which assists transcription in the absence of glucose. Two cAMP molecules might bind CAP functioning as allosteric effectors by increasing the protein's affinity for DNA, only when glucose levels are low.
- a promoter (P), region of DNA which contains specific sequence providing a binding site for RNA polymerase, a protein which transcribes the downstream genes into mRNA.
- the operator region (O). mRNA transcription starts when the repressor is removed from the operator by an inducer.
- three structural genes:
    o *lac Z* encodes β-galactosidase (lacZ), an enzyme which cleaves lactose into glucose and galactose.
    o *lac Y* encodes β-galactoside permease (lacY), a membrane-bound transport protein which pumps lactose into the cell.
    o *lac A* encodes β-galactoside transacetylase (lacA), an enzyme that transfers an acetyl group from acetyl-CoA to β-galactosides.

Transcription of the *lac* operon genes is controlled by two mechanisms. The first control mechanism depends on lactose availability. When lactose is not present, the *lac* repressor protein tightly binds to the operator keeping transcription of the downstream genes by RNA polymerase at very low levels (the repressor binding to the operator interferes to the binding of RNA polymerase to the operator). By contrast, when lactose is present, a lactose metabolite

called allolactose binds to the repressor avoiding binding of the latter to the operator. In this case, RNA polymerase is able to transcribe the downstream genes at high levels.

The second control mechanism depends on glucose levels. In the absence of glucose, two cAMP molecules bind to CAP (cAMP levels are high), enhancing transcription of the downstream genes by RNA polymerase. By contrast, in the presence of glucose, cAMP levels are low and do not bind to CAP. This latter behaviour is known as catabolite repression, explaining the differences of transcription levels when inducer is available and glucose is present or not. As it will be explained in chapter 8, transcription is also achieved in the presence of glucose, although higher levels of inducer are required.
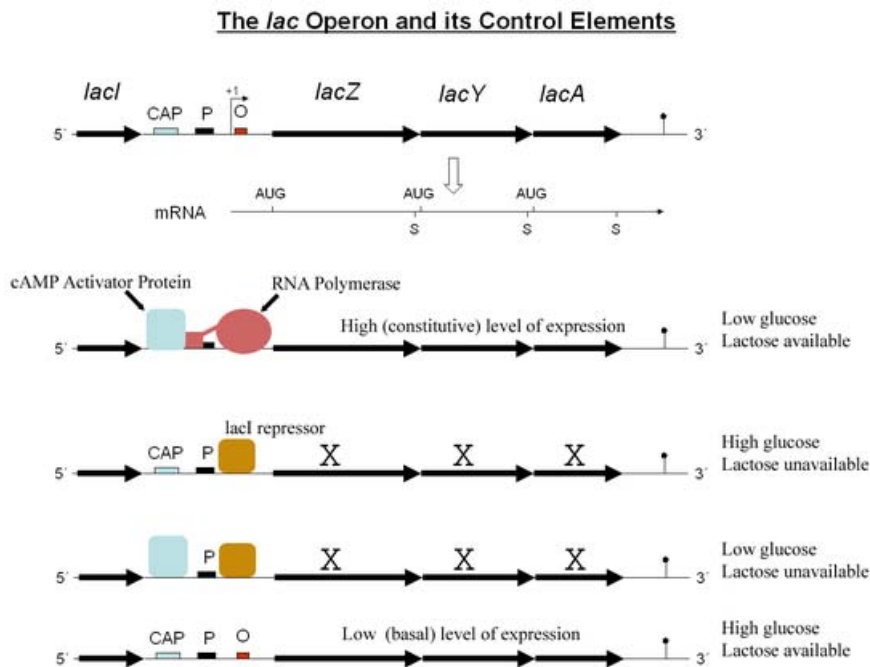
Figure 1.4. The *lac* operon and its control elements.

Usually, in recombinant protein expression plasmids based on *lac* promoters, the gene coding for the target protein is cloned downstream the *lac* operator, substituting the three structural genes of the *lac* operon.

Lactose is the natural inducer of the *lac* operon, but it has to be metabolised for an effective induction of protein expression. Since lactose metabolism is regulated by glucose levels, the quality of induction depends on careful monitoring of both glucose and lactose levels [44]. As a consequence, the most commonly used inducer for these *lac*-type promoter systems is isopropyl-β-D-thiogalactopyranoside (IPTG) (Figure 1.5). In contrast to lactose (natural inducer

of these systems if altered by β-galactosidase), IPTG is not metabolised, contributing to simplify the control of induction.
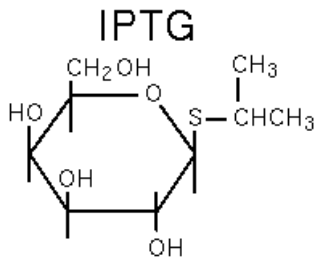
IPTG

Figure 1.5. Isopropyl-β-D-thiogalactopyranoside molecule.

For efficient transcription from *lac*-derived promoters, IPTG must enter the cells efficiently. This synthetic inducer is expensive and may interfere with cell growth. Owing the cost and potential toxic effects of IPTG, it would be advantageous to use less inducer to obtain the same level of transcription [45]. Lactose permease (the *lacY* gene product) takes part in transporting IPTG across the cell membrane [46]. Inducer transport by lactose permease is subject to carbon catabolite repression (decrease in activation level of the lactose operon due to the presence of glucose) and inducer exclusion (increase in the external inducer concentration threshold value at which the system shifts from an uninduced to an induced state) [47]. However, owing to other mechanisms of transport (including diffusion), induction can easily be obtained without the lactose permease [48, 49]. Once inside the cell, IPTG induces the *lac* operon as an all-or-none phenomenon. The production of permeases in a single cell can be either switched (induced) or shut off (uninduced) and, therefore, intermediate levels of enzyme production are a consequence of the coexistence of these two populations of cells [50].

In *Escherichia coli*, strong induction of recombinant protein expression has been shown to be deleterious due to a heavy metabolic burden on the host cell, which may completely cease cell growth before maximum product accumulation has occurred. Metabolic burden can be explained by a rapid increase in the demand of metabolites not usually produced in sufficient amounts by the metabolic pathways active under normal growth conditions [51]: resources (raw material and energy) are withdrawn from the host metabolism for maintenance and expression of the foreign DNA [52]. This undesired behaviour will depend on the system which is used (strain, plasmid copy number and construction, product, etc.) as well as on the operational strategy.

Both expression systems used in this work contained vectors with *lac*-based promoters. Hence, IPTG was used as inducer (although other inducers were also tested, see section *10.2. Galactose as an alternative inducer*). However, the behaviour of host cells was found to be completely different when working with "weak" or "strong" promoters and, therefore, most of the efforts of this work were focused on defining the best process strategy in each case to increase recombinant aldolase yields.

1.2. The aldolases.

The pharmaceutical and fine chemical industries often seek entire sets of multiple diverse biocatalysts to build toolboxes for biotransformations. These toolboxes need to be rapidly accessible to meet the strict time-lines of a biosynthetic feasibility evaluation in competition with traditional synthetic chemistry [1].

Aldolases are a group of C-C bond forming enzymes with widespread applications [53, 54]. Stereoselective C-C bond formation catalysed by aldolases has been reported as an attractive alternative to conventional chiral organic chemistry methods (Figure 1.6) for chemical and pharmaceutical industries [55-57]. Aldolases catalyse C-C bond formation with defined stereochemistry yielding enantiomerically pure products, even when the starting materials are non-chiral substrates [58].
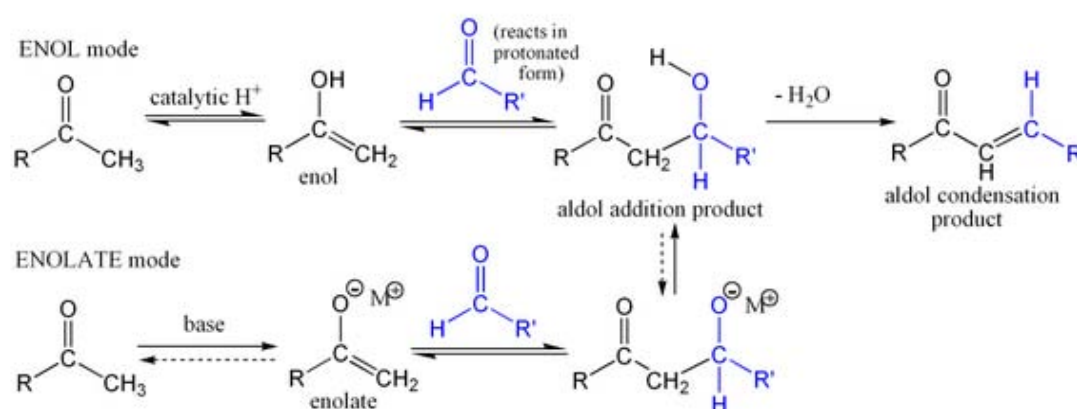
Figure 1.6. Aldol reactions.

The four enzymes of the dihydroxyacetone phosphate (DHAP) dependent aldolases group (Figure 1.7) have been proposed as a battery of biocatalysts able to catalyze the synthesis of diols of complementary stereochemistry [59]. DHAP dependent aldolases have been used as catalysts for the synthesis of enzyme inhibitors such as iminocyclitols [60, 61], which have been largely investigated as therapeutic targets for the design of new antibiotics, antimetastatic, antihyperglycemic, or immunostimulatory agents [62].

Aldolases isolated from different sources have been cloned in *E. coli* for their overexpression [63-65].

Figure 1.7. DHAP-dependent aldolases and their complementary stereochemistry based on their catalysed reactions.

One of the aldolases which has been produced in this work is the DHAP dependent fuculose-1-phosphate aldolase (FucA, EC 4.1.2.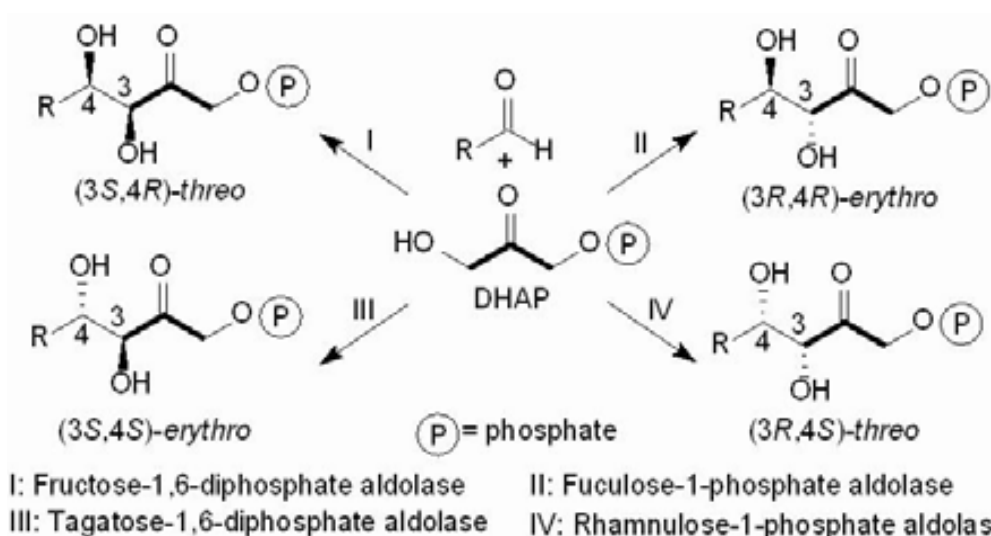17). FucA catalyzes the selective formation of R-R diols (D-*eritro* compounds) by aldol addition of DHAP to a wide variety of aldehydes [66]. It had been previously cloned in *E. coli* XL1 Blue and expressed as fusion protein with a six-histidine tag under the control of isopropyl-β-D-thiogalactopyranoside (IPTG) inducible weak promoter *trc* [67, 68] in pTrcfuc plasmid and under λ promoter in a pORTlaLfuc plasmid [69].

Rhamnulose 1-phosphate aldolase (RhuA, EC 4.1.2.19) is the other DHAP-dependent aldolase which has been produced in this work. RhuA catalyses the reversible asymmetrical aldol addition of DHAP to L-lactaldehyde generating (3*R*,4*S*) L-rhamnulose 1-phosphate *in vivo*. The possibility to use a wide range of acceptor substrates together with the predictability of the stereochemistry of the product (controlled by the enzyme) has developed an increasing interest for this enzyme for chiral synthesis [54, 61]. RhuA had been previously cloned and expressed as a fusion protein with a six-histidine tag in *E. coli* XL1 Blue under the control of IPTG inducible weak promoter *trc* [64] in pTrcrham plasmid as well as in *E. coli* M15 and under the control of the strong phage T5 promoter in pQErham and pQEαβrham plasmids [70, 71].

Production of both FucA and RhuA in an active form is challenging partly due to their complex tertiary structure defined by homotetramers (Figure 1.8). The four active sites are placed between the different subunits, where a Zn molecule acts as prosthetic group. Thus, the production process was not only developed to maximise the amount of protein, but also to ensure proper protein structure conformation to maximise enzyme activity.
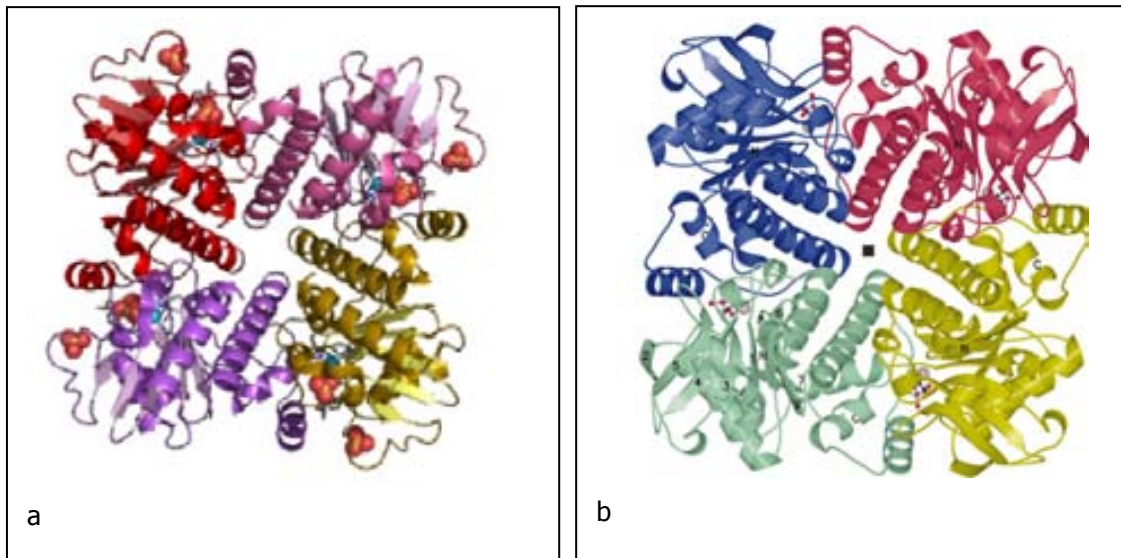
Figure 1.8. DHAP dependent aldolase crystallographic structures. (a) FucA (Protein Data Bank, European Molecular Biology Laboratory) and (b) RhuA [72] tetramers.

References.

[1] P.Lorenz, J.Eck, Metagenomics and industrial applications, Nature Reviews Microbiology 3 (2005) 510-516.

[2] O.Kirk, T.V.Borchert, C.C.Fuglsang, Industrial enzyme applications, Current Opinion in Biotechnology 13 (2002) 345-351.

[3] H.E.Schoemaker, D.Mink, M.G.Wubbolts, Dispelling the myths - Biocatalysis in industrial synthesis, Science 299 (2003) 1694-1697.

[4] A.J.J.Straathof, S.Panke, A.Schmid, The production of fine chemicals by biotransformations, Current Opinion in Biotechnology 13 (2002) 548-556.

[5] A.Schmid, et al., Industrial biocatalysis today and tomorrow, Nature 409 (2001) 258-268.

[6] G.Walsh, Biopharmaceuticals: recent approvals and likely directions, Trends Biotechnology 23 (2005) 553-558.

[7] I.Hunt, From gene to protein: a review of new and enabling technologies for multi-parallel protein expression, Protein Expression and Purification 40 (2005) 1-22.

[8] S.Macauley-Patrick, M.L.Fazenda, B.McNeil, L.M.Harvey, Heterologous protein production using the *Pichia pastoris* expression system, Yeast 22 (2005) 249-270.

[9] T.U.Gerngross, Advances in the production of human therapeutic proteins in yeasts and filamentous fungi, Nature Biotechnology 22 (2004) 1409-1414.

[10] P.Cornelis, Expressing genes in different *Escherichia coli* compartments, Current Opinion in Biotechnology 11 (2000) 450-454.

[11] R.Bentley, R.Meganathan, Biosynthesis of vitamin K (menaquinone) in bacteria, Microbiological Reviews 46 (1982) 241-280.

[12] J.Contiero, et al., Effects of mutations in acetate metabolism on high-cell-density growth of *Escherichia coli*, Journal of Industrial Microbiology & Biotechnology 24 (2000) 421-430.

[13] J.C.Diazricci, L.Regan, J.E.Bailey, Effect of Alteration of the Acetic-Acid Synthesis Pathway on the Fermentation Pattern of *Escherichia coli*, Biotechnology and Bioengineering 38 (1991) 1318-1324.

[14] W.R.Farmer, J.C.Liao, Reduction of aerobic acetate production by *Escherichia coli*, Applied Environmental Microbiology 63 (1997) 3205-3210.

[15] A.A.Aristidou, K.Y.San, G.N.Bennett, Modification of Central Metabolic Pathway in *Escherichia coli* to Reduce Acetate Accumulation by Heterologous Expression of the

Bacillus subtilis Acetolactate Synthase Gene, Biotechnology and Bioengineering 44 (1994) 944-951.

[16] L.L.Bermejo, N.E.Welker, E.T.Papoutsakis, Expression of *Clostridium acetobutylicum* ATCC 824 genes in *Escherichia coli* for acetone production and acetate detoxification, Applied and Environmental Microbiology 64 (1998) 1079-1085.

[17] N.N.Dedhia, W.E.Chen, J.E.Bailey, Managing Carbon Flow in *Escherichia coli* by Controlled Synthesis and Degradation of Glycogen, Abstracts of Papers of the American Chemical Society 207 (1994) 184-BIOT.

[18] G.Gosset, Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system, Microbial Cell Factories 4 (2005) 14.

[19] J.G.Thomas, A.Ayling, F.Baneyx, Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from *E. coli*: To fold or to refold, Applied Biochemistry and Biotechnology 66 (1997) 197-238.

[20] P.H.Bessette, F.Aslund, J.Beckwith, G.Georgiou, Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm, Proceedings of the National Academy of Sciences of the United States of America 96 (1999) 13703-13708.

[21] S.B.Noronha, H.J.Yeh, T.F.Spande, J.Shiloach, Investigation of the TCA cycle and the glyoxylate shunt in *Escherichia coli* BL21 and JM109 using (13)C-NMR/MS, Biotechnology and Bioengineering 68 (2000) 316-327.

[22] D.Riesenberg, et al., High cell density fermentation of recombinant *Escherichia coli* expressing human interferon alpha 1, Applied Microbiology and Biotechnology 34 (1990) 77-82.

[23] S.Y.Lee, H.N.Chang, Characteristics of poly(3-hydroxybutyric acid) synthesis by recombinant *Escherichia coli*, Annals of the New York Academy of Sciences 782 (1996) 133-142.

[24] J.Shiloach, R.Fass, Growing *E. coli* to high cell density--a historical perspective on method development, Biotechnology Advances 23 (2005) 345-357.

[25] D.Riesenberg, High-Cell-Density Cultivation of *Escherichia coli*, Current Opinion in Biotechnology 2 (1991) 380-384.

[26] I.C.Gunsalus, R.Y.Stanier, The bacteria, Academic Press, Inc., New York, 1960.

[27] R.A.Majewski, M.M.Domach, Simple Constrained-Optimization View of Acetate Overflow in *Escherichia coli*, Biotechnology and Bioengineering 35 (1990) 732-738.

[28] M.El-Mansi, Flux to acetate and lactate excretions in industrial fermentations: physiological and biochemical implications, Journal of Industrial Microbiology and Biotechnology 31 (2004) 295-300.

[29] A.Bock, G.Sawers, Fermentation, in: H.E.Umbarger (Ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, ASM Press, Washington D.C. (1996) 262-282.

[30] R.Britten, Extracellular Metabolic Products of *Escherichia coli* During Rapid Growth, Science 119 (1954) 578.

[31] H.Holms, Flux analysis and control of the central metabolic pathways in *Escherichia coli*, FEMS Microbiology Reviews 19 (1996) 85-116.

[32] W.H.Holms, The central metabolic pathways of *Escherichia coli*: relationship between flux and control at a branch point, efficiency of conversion to biomass, and excretion of acetate, Current Topics in Cell Regulation 28 (1986) 69-105.

[33] H.G.Crabtree, Observations on the carbohydrate metabolism of tumours, Biochemical Journal 23 (1929) 536-545.

[34] H.L.Kornberg, Role and Control of Glyoxylate Cycle in *Escherichia coli* - First Colworth Medal Lecture, Biochemical Journal 99 (1966) 1-&.

[35] T.D.Brown, M.C.Jones-Mortimer, H.L.Kornberg, The enzymic interconversion of acetate and acetyl-coenzyme A in *Escherichia coli*, Journal of General Microbiology 102 (1977) 327-336.

[36] G.L.Kleman, W.R.Strohl, Acetate metabolism by *Escherichia coli* in high-cell-density fermentation, Applied and Environmental Microbiology 60 (1994) 3952-3958.

[37] C.H.Chou, G.N.Bennett, K.Y.San, Effect of Modified Glucose-Uptake Using Genetic-Engineering Techniques on High-Level Recombinant Protein-Production in *Escherichia coli* Dense Cultures, Biotechnology and Bioengineering 44 (1994) 952-960.

[38] M.El-Mansi, W.H.Holms, Control of Carbon Flux to Acetate Excretion During Growth of *Escherichia coli* in Batch and Continuous Cultures, Journal of General Microbiology 135 (1989) 2875-2883.

[39] K.Konstantinov, M.Kishimoto, T.Seki, T.Yoshida, A Balanced Do-Stat and Its Application to the Control of Acetic-Acid Excretion by Recombinant *Escherichia coli*, Biotechnology and Bioengineering 36 (1990) 750-758.

[40] K.Han, H.C.Lim, J.Hong, Acetic-Acid Formation in *Escherichia coli* Fermentation, Biotechnology and Bioengineering 39 (1992) 663-671.

[41] M.Mujacic, K.W.Cooper, F.Baneyx, Cold-inducible cloning vectors for low-temperature protein expression in *Escherichia coli*: application to the production of a toxic and proteolytically sensitive fusion protein, Gene 238 (1999) 325-332.

[42] E.Amann, J.Brosius, "ATG vectors' for regulated high-level expression of cloned genes in *Escherichia coli*, Gene 40 (1985) 183-190.

[43] M.Lanzer, H.Bujard, Promoters Largely Determine the Efficiency of Repressor Action, Proceedings of the National Academy of Sciences of the United States of America 85 (1988) 8973-8977.

[44] P.Neubauer, K.Hofmann, O.Holst, B.Mattiasson, P.Kruschke, Maximizing the Expression of A Recombinant Gene in *Escherichia coli* by Manipulation of Induction Time Using Lactose As Inducer, Applied Microbiology and Biotechnology 36 (1992) 739-744.

[45] L.H.Hansen, S.Knudsen, S.J.Sorensen, The effect of the *lacY* gene on the induction of IPTG inducible promoters, studied in *Escherichia coli* and *Pseudomonas fluorescens*, Current Microbiology 36 (1998) 341-347.

[46] P.R.Jensen, H.V.Westerhoff, O.Michelsen, The Use of *Lac*-Type Promoters in Control Analysis, European Journal of Biochemistry 211 (1993) 181-191.

[47] M.Santillan, M.C.Mackey, Influence of catabolite repression and inducer exclusion on the bistable behavior of the *lac* operon, Biophysical Journal 86 (2004) 1282-1292.

[48] J.Beckwith, The lactose operon, in: F.C.Neidhart (Ed.), *Escherichia coli* and *Salmonella typhimurium*, American Society of Microbiology, Washington D.C. (1987) 1444-1452.

[49] A.Khlebnikov, J.D.Keasling, Effect of *lacY* expression on homogeneity of induction from the P-*tac* and P-*trc* promoters by natural and synthetic inducers, Biotechnology Progress 18 (2002) 672-674.

[50] J.M.G.Vilar, C.C.Guet, S.Leibler, Modeling network dynamics: the *lac* operon, a case study, Journal of Cell Biology 161 (2003) 471-476.

[51] B.R.Glick, Metabolic Load and Heterologous Gene-Expression, Biotechnology Advances 13 (1995) 247-261.

[52] W.E.Bentley, N.Mirjalili, D.C.Andersen, R.H.Davis, D.S.Kompala, Plasmid-Encoded Protein - the Principal Factor in the Metabolic Burden Associated with Recombinant Bacteria, Biotechnology and Bioengineering 35 (1990) 668-681.

[53] W.D.Fessner, V.Helaine, Biocatalytic synthesis of hydroxylated natural products using aldolases and related enzymes, Current Opinion in Biotechnology 12 (2001) 574-586.

[54] C.J.Hamilton, Enzymes in preparative mono- and oligo-saccharide synthesis, Natural Product Reports 21 (2004) 365-385.

[55] T.D.Machajewski, C.H.Wong, The catalytic asymmetric aldol reaction, Angewandte Chemie-International Edition 39 (2000) 1352-1374.

[56] C.H.Wong, L.H.Randall, Enzymes in Organic Synthesis: Application to the Problems of Carbohydrate Recognition (Part 1), Angewandte Chemie International Edition in English 34 (1995) 412-432.

[57] N.Wymer, E.J.Toone, Enzyme-catalyzed synthesis of carbohydrates, Current Opinion in Chemical Biology 4 (2000) 110-119.

[58] H.J.M.Gijsen, L.Qiao, W.Fitz, C.H.Wong, Recent advances in the chemoenzymatic synthesis of carbohydrates and carbohydrate mimetics, Chemical Reviews 96 (1996) 443-473.

[59] K.K.C.Liu, et al., Use of Dihydroxyacetone Phosphate Dependent Aldolases in the Synthesis of Deoxyazasugars, Journal of Organic Chemistry 56 (1991) 6280-6289.

[60] L.Espelt, et al., Stereoselective aldol additions catalyzed by dihydroxyacetone phosphate-dependent aldolases in emulsion systems: Preparation and structural characterization of linear and cyclic iminopolyols from aminoaldehydes, Chemistry-A European Journal 9 (2003) 4887-4899.

[61] S.Takayama, G.J.McGarvey, C.H.Wong, Microbial aldolases and transketolases: New biocatalytic approaches to simple and complex sugars, Annual Review of Microbiology 51 (1997) 285-310.

[62] V.H.Lillelund, H.H.Jensen, X.Liang, M.Bols, Recent developments of transition-state analogue glycosidase inhibitors of non-natural product origin, Chemical Reviews 102 (2002) 515-553.

[63] W.D.Fessner, et al., Enzymes in Organic-Synthesis .1. Diastereoselective Enzymatic Aldol Additions - L-Rhamnulose and L-Fuculose 1-Phosphate Aldolases from *Escherichia coli*, Angewandte Chemie-International Edition in English 30 (1991) 555-558.

[64] E.García-Junceda, G.J.Shen, T.Sugai, C.H.Wong, A New Strategy for the Cloning, Overexpression and One-Step Purification of 3 DHAP-Dependent Aldolases - Rhamnulose-1-Phosphate Aldolase, Fuculose-1-Phosphate Aldolase and Tagatose-1,6-Diphosphate Aldolase, Bioorganic & Medicinal Chemistry 3 (1995) 945-953.

[65] L.Vidal, J.Calveras, P.Clapes, P.Ferrer, G.Caminal, Recombinant production of serine hydroxymethyl transferase from *Streptococcus thermophilus* and its preliminary evaluation as a biocatalyst, Applied Microbiology and Biotechnology 68 (2005) 489-497.

[66] R.Alajarin, et al., Synthesis, structure, and pharmacological evaluation of the stereoisomers of furnidipine, Journal of Medical Chemistry 38 (1995) 2830-2841.

[67] O.Durany, C.de Mas, J.Lopez-Santin, Fed-batch production of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 40 (2005) 707-716.

[68] E.Garcia-Junceda, G.J.Shen, R.Alajarin, C.H.Wong, Cloning and overexpression of rhamnose isomerase and fucose isomerase, Bioorganic and Medicinal Chemistry 3 (1995) 1349-1355.

[69] O.Durany, et al., Production of fuculose-1-phosphate aldolase using operator-repressor titration for plasmid maintenance in high cell density *Escherichia coli* fermentations, Biotechnology and Bioengineering 91 (2005) 460-467.

[70] L.Vidal, et al., High-level production of recombinant His-tagged rhamnulose 1-phosphate aldolase in *Escherichia coli*, Journal of Chemical Technology and Biotechnology 78 (2003) 1171-1179.

[71] L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

[72] M.Kroemer, G.E.Schulz, The structure of L-rhamnulose-1-phosphate aldolase (class II) solved by low-resolution SIR phasing and 20-fold NCS averaging, Acta Crystallographica section D: Biological Crystallography 58 (2002) 824-832.

2. CONTEXT AND OBJECTIVES.

This work was carried out within a research group which is involved in the production of aldolases and their utilisation in biocatalysis. Specifically, the present work was focused on the development of enzyme production stage.

Previous works within the group related to recombinant protein production consisted in media formulation for recombinant aldolase production [1], implementation of substrate-limited fed-batch strategies to obtain recombinant aldolases in high cell density cultures [2, 3], production of recombinant aldolases using alternative expression systems [4-6] as well as production of different aldolases using the same expression system [7].

Similar operational strategies had been previously applied to all systems. Predefined exponential feeding had been used in fed-batch cultures using minimal media and glucose as a sole carbon source to increase biomass concentration. Constant specific growth rate to avoid acetate accumulation was maintained by limiting glucose and, when high cell densities were reached, an IPTG pulse was added to the cultures, inducing the overexpression of the correspondent recombinant aldolase.

The main objective of the present was to improve the production process of different recombinant aldolases in *Escherichia coli*: fuculose 1-phosphate aldolase (FucA) and rhamnulose 1-phosphate aldolase (RhuA) were produced in *E. coli* XL1 Blue MRF' pTrcfuc and *E. coli* M15 [pREP4] ΔglyA pQE$\alpha\beta$rham, respectively.

In the first case the process had been successfully optimised and the aim of this work was its control and automation.

In the second case, the productivity of the process had been increased in high cell density cultures, but results showed that the host cells were not exploited at their maximum capacity and that the activity of the enzyme was not the expected one. Therefore, part of the present work was also focused on finding out the reasons for these problems and on fixing them by means of process strategies.

The main objective was then translated into several secondary objectives:

- To automate FucA production process using *E. coli* XL1 Blue MRF' pTrcfuc by implementing simple and robust specific growth rate control.

- To optimise RhuA production in high cell density cultures using *E. coli* M15 ΔglyA [pREP4] pQEαβrham. This included:

    o The development and implementation of alternative substrate feeding strategies to minimise the starvation-induced stress in host cells.

    o The development and implementation of alternative inducer feeding strategies to minimise the metabolic burden in host cells by tuning the transcription rate.

    o The study of process temperature reduction to decrease temperature-driven stress in host cells and improve active protein yields.

    o An analysis of the impact of the production process strategy on the global immobilised biocatalyst yields, considering both production as well as the downstream stages.

    o The development and calibration of a robust production model for process optimisation.

References.

[1] O.Durany, G.Caminal, C.de Mas, J.Lopez-Santin, Studies on the expression of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 39 (2004) 1677-1684.

[2] O.Durany, C.de Mas, J.Lopez-Santin, Fed-batch production of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 40 (2005) 707-716.

[3] L.Vidal, P.Ferrer, G.Alvaro, M.D.Benaiges, G.Caminal, Influence of induction and operation mode on recombinant rhamnulose 1-phosphate aldolase production by *Escherichia coli* using the T5 promoter, Journal of Biotechnology 118 (2005) 75-87.

[4] O.Durany, et al., Production of fuculose-1-phosphate aldolase using operator-repressor titration for plasmid maintenance in high cell density *Escherichia coli* fermentations, Biotechnology and Bioengineering 91 (2005) 460-467.

[5] L.Vidal, et al., High-level production of recombinant His-tagged rhamnulose 1-phosphate aldolase in *Escherichia coli*, Journal of Chemical Technology and Biotechnology 78 (2003) 1171-1179.

[6] L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

[7] L.Vidal, J.Calveras, P.Clapes, P.Ferrer, G.Caminal, Recombinant production of serine hydroxymethyl transferase from *Streptococcus thermophilus* and its preliminary evaluation as a biocatalyst, Applied Microbiology and Biotechnology 68 (2005) 489-497.

3. MATERIALS AND METHODS.

The materials and methods described in the following section apply to all experiments except when specifically indicated. These are general materials and methods which have been used throughout the whole work, although specific materials and methods used in particular parts of it can be also found in the corresponding chapters.

3.1. Strains and vectors.

Two different *E. coli*-based systems were used to overexpress recombinant aldolases. Both strains were stored at -80ºC in glycerol stocks using commercial Cryobilles (AES Laboratoire, France) from exponential phase cultures in Luria-Bertani medium supplemented with antibiotic (LB-A).

The first one, *E. coli* XL1 Blue MRF' pTrcfuc, was used to overexpress fuculose 1-phosphate aldolase (FucA) [1]. The isolated gene *fucA* from *E. coli* had been previously cloned in the commercial strain XL1 Blue MRF' from Invitrogen [2], which is auxotrophic for thiamine (vitamin B1). This strain harbours pTrcHis (Figure 3.1) plasmids, which are pBR322-derived expression vectors designed for efficient recombinant protein expression and purification in *E. coli*. High levels of expression are possible using the IPTG inducible *trc* (*trp-lac*) promoter. The pTrcHis vectors also contain a copy of the *lacI* gene coding for the *lac* repressor protein. This allows for efficient repression of transcription of the cloned insert in *E. coli* regardless of whether the strain is *lacI^{q+}* or *lacI^{q-}*. pTrcHis allows the use of ampicillin as selection marker due to the constitutive overexpression of β-lactamase and includes the sequence to overexpress a hexahistidine tag (6xHis tag) at the N-terminus of the target protein. The metal binding domain of the fusion peptide allows simple, one step purification of recombinant proteins by Immobilized Metal Affinity Chromatography (IMAC).

The second expression system was used to overexpress rhamnulose 1-phosphate aldolase (RhuA) with a hexahistidine tag at its N-terminus [2, 3]. The K-12 derived strain *E. coli* M15 ΔglyA [pREP4] harbouring the vector pQEαβrham was used for its overexpression [4]. It was derived from the commercial *E. coli* M15 [pREP4] pQE40 system from Qiagen.
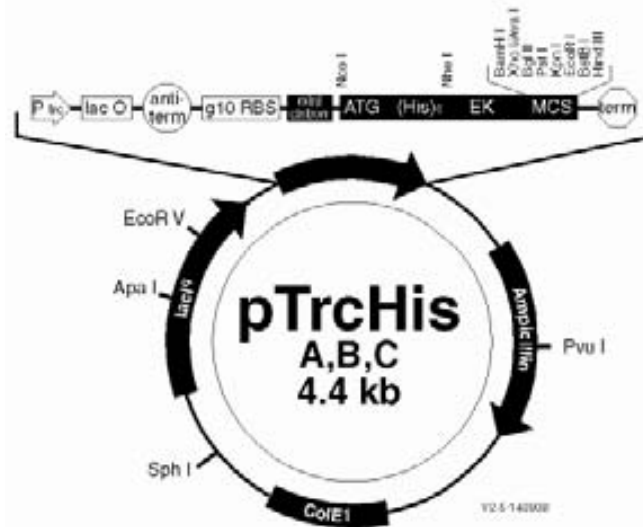
Figure 3.1. pTrcHis plasmid (www.invitrogen.com).

The high-level expression of 6xHis-tagged proteins in *E. coli* using pQE vectors (Figure 3.2) is based on the T5 promoter transcription–translation system. pQE plasmids belong to the pDS family of plasmids. These are low-copy plasmids with optimized promoter–operator element consisting of phage T5 promoter (recognized by the *E. coli* RNA polymerase) and two *lac* operator sequences which increase *lac* repressor binding to ensure efficient repression of the powerful T5 promoter. pQE plasmids also contain 6xHis-tag coding sequence at the N-terminus as well as β-lactamase gene (*bla*) conferring resistance to ampicillin. The extremely high transcription rate initiated at the T5 promoter can only be efficiently regulated and repressed by the presence of high levels of the *lac* repressor protein. Expression of recombinant proteins encoded by pQE vectors is rapidly induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG).
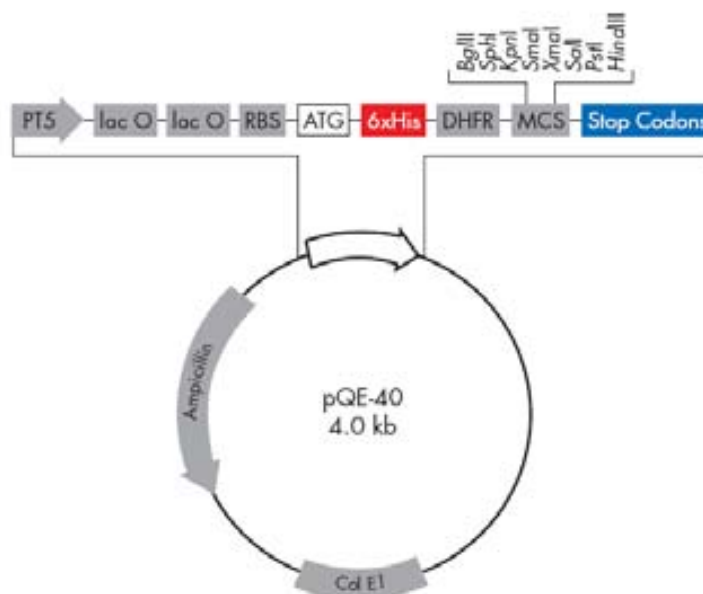

Figure 3.2. pQE vector (www.qiagen.com).

The host strain also contains the low-copy plasmid pREP4 (Figure 3.3) which confers kanamycin resistance and constitutively expresses the *lac* repressor protein encoded by the *lacI* gene. The pREP4 plasmid is derived from pACYC and contains the p15A replicon. Multiple copies of pREP4 are present in the host cells that ensure the production of high levels of the *lac* repressor protein which binds to the operator sequences and tightly regulates recombinant protein expression. The pREP4 plasmid is compatible with all plasmids carrying the ColE1 origin of replication, and is maintained in *E. coli* in the presence of kanamycin.
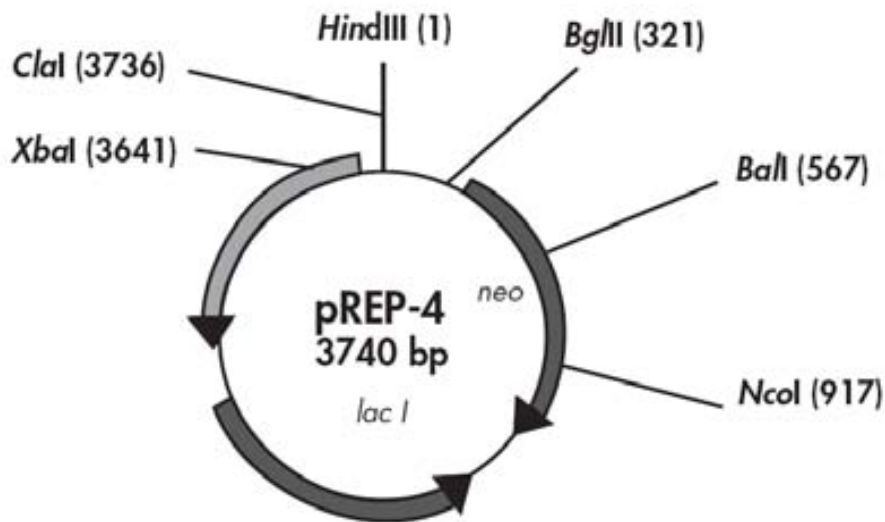

Figure 3.3. Restriction map of the pREP4 helper plasmid (www.qiagen.com).

In previous work carried out in the research group, *E. coli* M15 [pREP4] pQErham was modified to obtain *E. coli* M15 ΔglyA [pREP4] pQEαβrham [4]. This expression system was designed as an antibiotic-free plasmid selection system based on glycine auxotrophy to ensure plasmid stability without the need of antibiotic supplementation. The *glyA* gene was deleted from the chromosome of *E. coli* M15 to obtain *E. coli* M15 ΔglyA, and it was inserted in the pQErham plasmid under the control of a constitutive promoter (P3) to obtain pQEαβrham (Figure 3.4). Using this system, only bearing-plasmid cells are able to grow in synthetic medium without glycine supplementation thanks to constitutive SHMT expression.
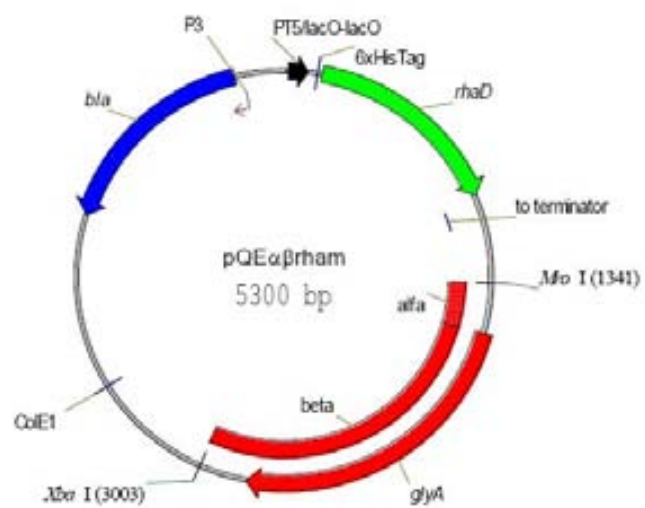
Figure 3.4. pQEαβrham plasmid.

3.2. Media composition.

LB medium, with 10 g L^{-1} peptone, 5 g L^{-1} yeast extract and 10 g L^{-1} NaCl, was used for all preinoculum preparations.

LB medium supplemented with antibiotic was used to prepare cell banks. In the case of *E. coli* XL1 Blue MRF' pTrcfuc it was supplemented with ampicillin to 0.1 g·L^{-1}, while in the case of *E. coli* M15 ΔglyA [pREP4] pQEαβrham was supplemented with ampicillin and kanamycin to 0.1 and 0.025 g·L^{-1}, respectively (contamination was avoided and it ensured the cell bank contained only plasmid-bearing cells, since the auxotrophy-based plasmid maintenance system is only effective in defined media without glycine supplementation).

A defined mineral medium, utilising glucose as the sole carbon source, was used for inocula, for shake flask experiments and for all bioreactor cultivations.

The medium for shake flask cultures was composed of 5 g·L^{-1} glucose, 11.9 g·L^{-1} K$_2$HPO$_4$, 2.4 g·L^{-1} KH$_2$PO$_4$, 1.8 g·L^{-1} NaCl, 3 g·L^{-1} (NH$_4$)$_2$SO$_4$, 0.11 g·L^{-1} MgSO$_4$·7H$_2$O, 0.01 g·L^{-1} FeCl$_3$, 0.03 g·L^{-1} thiamine and 0.72 mL·L^{-1} of trace elements solution.

The batch phase for bioreactor cultivations was composed of 20 g·L^{-1} glucose, 11.9 g·L^{-1} K$_2$HPO$_4$, 2.4 g·L^{-1} KH$_2$PO$_4$, 1.8 g·L^{-1} NaCl, 3 g·L^{-1} (NH$_4$)$_2$SO$_4$, 0.45 g·L^{-1} MgSO$_4$·7H$_2$O, 0.02 g·L^{-1} FeCl$_3$, 0.1 g·L^{-1} thiamine and 2.86 mL·L^{-1} of trace elements solution.

The feed medium for high-cell-density fed-batch fermentations consisted of 487 g·L^{-1} glucose, 9.7 g·L^{-1} MgSO$_4$·7H$_2$O, 0.5 g·L^{-1} FeCl$_3$, 0.34 g·L^{-1} thiamine, 64 mL·L^{-1} trace elements solution and 0.5 mL·L^{-1} of antifoam (Sigma).

The trace elements solution composition contained (g·L^{-1}): 1.44 CaCl$_2$·2H$_2$O, 0.042 AlCl$_3$·6H$_2$O, 0.87 ZnSO$_4$·7H$_2$O, 0.16 CoCl$_2$·6H$_2$O, 1.6 CuSO$_4$, 0.01 H$_3$BO$_3$, 1.42 MnCl$_2$·4H$_2$O, 0.01 NiCl$_2$·6H$_2$O, 0.02 Na$_2$MoO$_4$·H$_2$O [5].

Phosphates were not included in the feeding solution for fed-batch cultures in order to avoid co-precipitation with magnesium salts. Instead, a concentrated phosphate solution containing 500 g·L^{-1} K$_2$HPO$_4$ and 100 g·L^{-1} KH$_2$PO$_4$ was pulsed during the fed-batch phase to avoid their depletion when necessary (calculated yield Y$_{XP}$= 18 gDCW·g P^{-1}).

3.3.   Experimental setup. Hardware and software.

Experiments were carried out using a Biostat® B bioreactor (Braun Biotech Int.) equipped with a 2 L fermentation vessel. Its digital control unit (DCU) controlled the pH, stirring, temperature and dissolved oxygen of bioreactor. Two gas mass flow meters (MFM, Bronkhorst EL-FLOW) were used to control and to measure inlet and outlet gas mass flows, respectively. Exhaust gas flow was continuously sampled and analysed by a mass spectrometer (MS, Balzers Quadstar 422). A microburette (MB, Crison Instruments MICRO BU 2030) with a 2.5 mL syringe (Hamilton) was used in fed-batch cultures (Figure 3.5) for discrete feed addition.
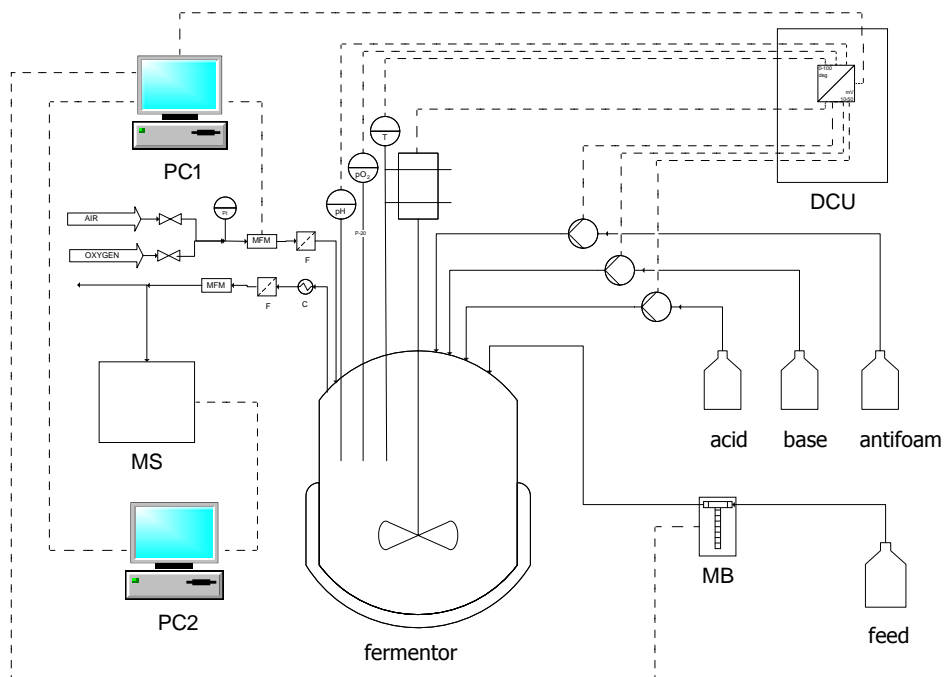

Figure 3.5. Schematic representation of the experimental setup used to perform fed-batch cultures.

Specific software was developed (using Microsoft Visual Basic 6.0.) to configure communication between devices, store acquired data in a database and monitor process variables along the fermentation (Figure 3.6). It also allowed on-line calculation of different process variables such as culture volume, carbon exhaust rate, oxygen uptake rate and volume of feed to supply along time. Substrate mass balance was used as control law to determine feed rate, and the total cell concentration was either calculated by biomass mass balances or estimated by exhaust gas measurements. Data acquisition, treatment, storage and actuation frequency was user defined, usually to 30 seconds.
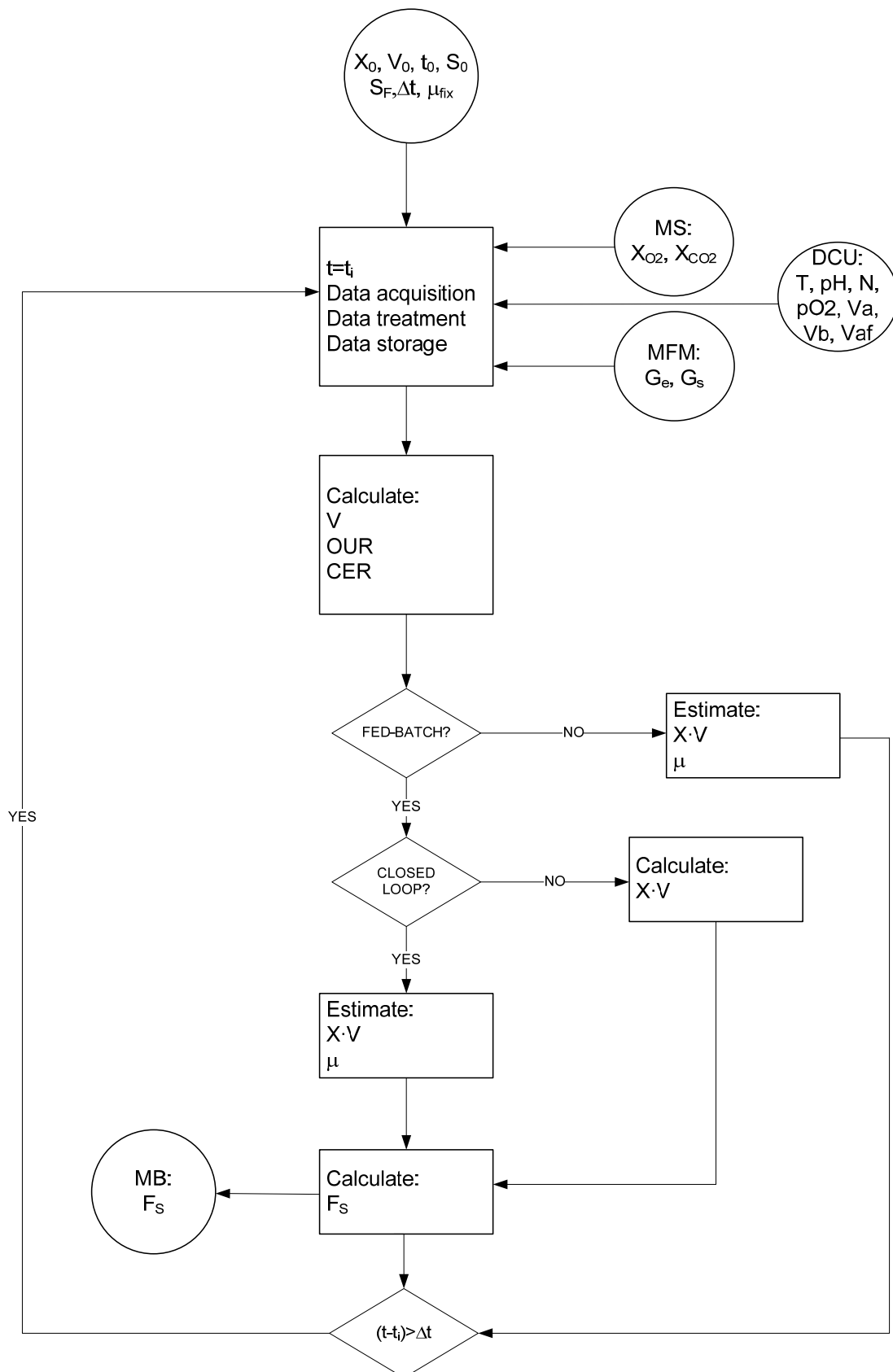
Figure 3.6. Flowchart of the developed software.

Communication of all devices with the computer where the main software was installed (PC1 in Figure 3.5) was implemented using different strategies depending on the case:

- The DCU was connected to PC1 using the respective serial ports and a RS422/RS232 conversion card. The communication protocol was implemented as a subroutine in the Visual Basic house-made software.
- Data from the MS was acquired by connecting the computer which controls it (PC2) with PC1 by using the respective Ethernet cards directly connected by a twisted cable with RJ45 connectors. Then, a private Microsoft Network was created so that all files of PC2 were accessible from PC1.
- Both mass flow meters were connected to PC1 by means of an A/D D/A converter, since their output goes from 0 to 5 VDC. The communication protocol was provided by the converter manufacturer and implemented as a subroutine in the Visual Basic house-made software.
- The microburette was controlled by the main software. In order to send the instructions to it, a serial port of PC1 and RS232 communication protocol was used. The communication protocol was provided by the MB manufacturer.

3.4. Cultivation conditions.

Preinoculum cultures were grown from glycerol stocks in a 100 mL shake flask containing 15 mL LB media and incubated overnight at 37ºC on a rotary shaker at 250 rpm.

To prepare inoculum, 5 mL of preinoculum cultures were transferred to a 0.5 L shake flask containing 100 mL of defined medium and it was incubated at 37ºC for 5 h with shaking at 250 rpm.

For fed-batch bioreactor experiments, 80 mL of inoculum culture were transferred aseptically to the bioreactor containing 720 mL of defined medium. All growth experiments were carried out using a Biostat® B bioreactor (Braun Biotech Int.) equipped with a 2 L fermentation vessel. The pH was maintained at 7.00±0.05 by adding 25% (w/v) NH$_4$OH solution to the reactor. The temperature was kept at 37ºC throughout the growth when not stated differently. The pO2 value was maintained at 50% of air saturation by adapting the stirrer speed between 450 and 1120 rpm and supplying air and/or pure oxygen at a flow rate of 2 vvm. The end of the batch phase was identified by a reduction in the oxygen consumption rate and an increase in pH. A simple mathematical model based on mass balances and substrate consumption kinetics was used to control the specific growth rate at a constant value by limiting the carbon source to the cultures [6] according to Eq. 3.1:

$$V_{ad}(t) = \frac{1}{S_F} \cdot (\frac{m_{sx}}{\mu_{fix}} + \frac{1}{Y_{xs}}) \cdot X(t) \cdot V(t) \cdot (\exp(\mu_{fix} \cdot \Delta t) - 1) \qquad \text{(Eq. 3.1)}$$

where $V_{ad}(t)$ is the feed volume (L) to be added at a time t, S_F is the substrate feed concentration (g·L^{-1}), m_{SX} is the maintenance coefficient (g·gDCW^{-1}·h^{-1}), μ_{fix} is the predefined specific growth rate (h^{-1}), Y_{XS} is the substrate to biomass yield (gDCW·g^{-1}), Δt is the time interval between additions (h) and X(t)·V(t) is the total biomass (gDCW), defined by Eq. 3.2:

$$X(t) \cdot V(t) = X_0 \cdot V_0 \cdot \exp(\mu_{fix} \cdot t) \qquad \text{(Eq. 3.2)}$$

where $X_0 \cdot V_0$ is the total biomass when the feed is started (gDCW).

In order to keep the specific growth rate to the desired set point, the following values were set in the feeding model: 0.31 gDCW·g^{-1} for biomass-substrate yield and 0 g·gDCW^{-1}·h^{-1} for maintenance coefficient in the case of *E. coli* XL1 Blue MRF' pTrcfuc, while for *E. coli* M15 ΔglyA [pREP4] pQEαβrham, 0.35 gDCW·g^{-1} was set for biomass-substrate yield and 0.10 g·gDCW^{-1}·h^{-1} for maintenance coefficient.

3.5.   **Analytical methods**.

Dry cell weight (DCW) concentrations were estimated using a calibration curve between DCW and optical density at 600 nm (OD600). In the case of *E. coli* XL1 Blue MRF' pTrcfuc, 1 unit of optical density at 600 nm was equivalent to approximately 0.25 gDCW·L^{-1} for non-induced fed-batch cultures (Figure 3.7) –see section *4.2. Analytical methods* for values obtained in induced fed-batch cultures as well as for non-induced continuous cultures-. For *E. coli* M15 ΔglyA [pREP4] pQEαβrham, this value was approximately 0.31 in all cases (Figure 3.8).
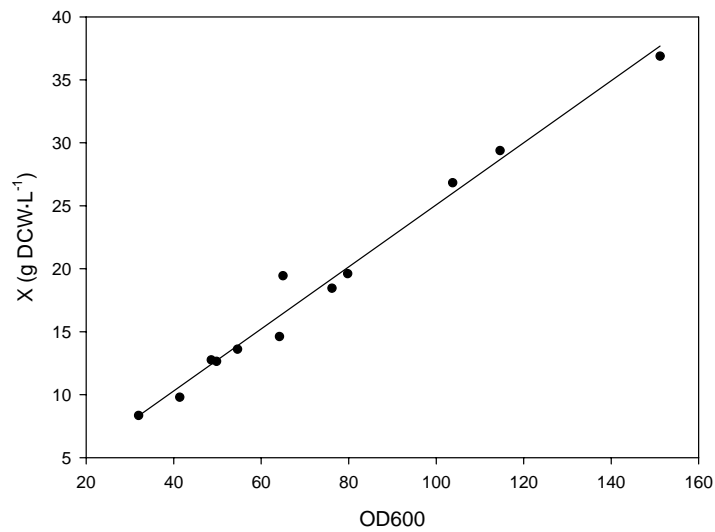

Figure 3.7. Dry cell weight concentration versus optical density at 600 nm in fed-batch cultures of *E. coli* XL1 Blue MRF' pTrcfuc.
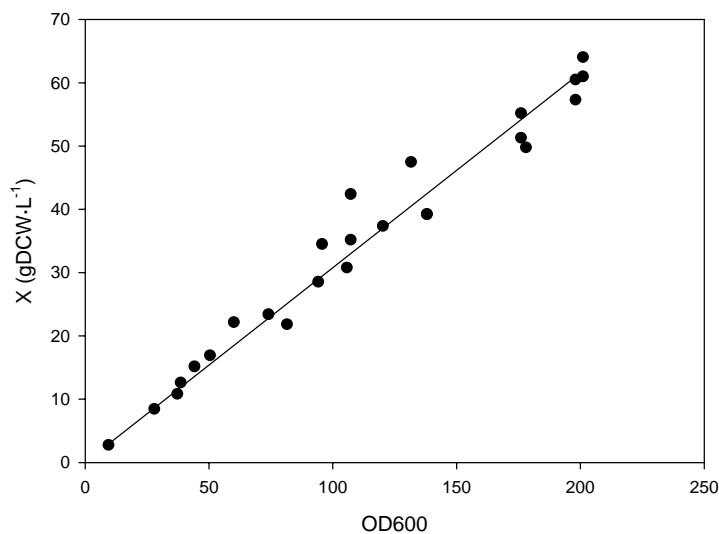

Figure 3.8. Dry cell weight concentration versus optical density at 600 nm in fed-batch cultures of *E. coli* M15 ΔglyA [pREP4] pQEαβrham.

The DCW values for the calibration curve calculation were obtained from culture broth samples centrifuged and washed twice with 0.9% (w/v) NaCl isotonic solution, filtered and dried at 105ºC until constant weight.

OD600 was obtained from fermentation broth samples, measuring absorbance at 600 nm (Philips PU8620) in the linear range 0.1-0.9. Samples were diluted with distilled water to keep absorbance in the linear region when required.

For other determinations, one millilitre of culture medium was centrifuged, and the supernatant was then filtered and used for glucose, organic acids, ammonium and phosphate measurements. Glucose and organic acids were analyzed by HPLC (Helwett Packard 1050) on an Aminex HPX-87H (Biorad) ion exchange column at 25ºC with refractive index detector (Helwett Packard 1047) using 15 mM H_2SO_4 (pH=3.0) as eluent at a flow rate of 0.6 mL·min^{-1}. Ammonium and phosphates were determined by colorimetric kit assays (Dr. Lange) following the supplier's instructions.

To quantify the product, broth samples were withdrawn and centrifuged. After rejecting the supernatant, the pellet was resuspended in 100 mM Tris·HCl (pH=7.5) to a final OD600=3 for enzyme determination. Cell suspensions were placed in ice and sonicated with four 15 s pulses at 50 W with 2 min intervals in ice between each pulse using a Vibracell$^®$ model VC50 (Sonics & Materials) equipped with a microtip probe. Cellular debris was removed by centrifugation and the clear supernatant was collected for product analysis. Two different methods were used to quantify the product in each sample. The first one was used to quantify the amount of protein, while the second one was used to quantify its activity:

a. Total intracellular soluble protein was determined using a Comassie® Protein Assay Reagent Kit (Pierce) as detailed in the supplier's instructions. To determine the percentage of RhuA amongst the rest of the intracellular soluble proteins, 12% polyacrylamide SDS-PAGE gels were performed in a Miniprotean® II instrument (Bio-Rad) according to the manufacturer's instructions and quantified by Kodak Digital Science® densitometry software.

b. Fuculose 1-phosphate aldolase (FucA) activity was measured using a coupled enzymatic assay [7]. In the first step, fuculose-1-phosphate is cleaved to L-lactaldehyde and dihydroxyacetone phosphate (DHAP) by FucA; in the second step, DHAP is reduced using rabbit muscle glycerol 3-phosphate dehydrogenase (GDH) and reduced nicotine adenine dinucleotide (NADH). 2 mM of fuculose-1-phosphate, 0.14 mM NADH and 1.7 activity units of GDH were incubated in 100 mM Tris·HCl buffer 150 mM KCl pH 7.5 at 25ºC. The assay was initiated upon addition of the sample to be titrated for this aldolase activity. One activity unit (AU) of FucA is defined as the

amount of enzyme that catalyzes the hydrolysis of 1 μmol of fuculose-1-phosphate to DHAP and L-lactaldehyde per minute at 25ºC and pH 7.5. The time course of the reaction was followed measuring the decrease of NADH monitoring the absorbance at 340 nm in a UV-VIS Cary (Varian) spectrophotometer (ε = 6.22 mM^{-1}cm^{-1}).

Determination of RhuA activity was carried out with a coupled continuous assay [3]. In the first step, rhamnulose 1-phosphate is cleaved to L-lactaldehyde and DHAP. In the second step, DHAP is reduced using rabbit muscle GDH and NADH. In the assay, the sample to be titrated was mixed with a solution containing 0.15 mM NADH, 2.0 mM rhamnulose 1-phosphate, 100 mM KCl, 50 mM Tris·HCl, and 2.5 AU·mL^{-1} GDH at pH 7.5. The reaction was incubated at 25ºC and the absorbance at 340 nm monitored in a UV-VIS Cary (Varian) spectrophotometer (ε = 6.22 mM^{-1}cm^{-1}). One unit of RhuA activity was defined as the amount of enzyme required to convert 1 µmol of rhamnulose 1-phosphate in DHAP per minute at 25ºC and pH=7.5.

References.

[1] O.Durany, C.de Mas, J.Lopez-Santin, Fed-batch production of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 40 (2005) 707-716.

[2] E.García-Junceda, G.J.Shen, T.Sugai, C.H.Wong, A New Strategy for the Cloning, Overexpression and One-Step Purification of 3 DHAP-Dependent Aldolases - Rhamnulose-1-Phosphate Aldolase, Fuculose-1-Phosphate Aldolase and Tagatose-1,6-Diphosphate Aldolase, Bioorganic & Medicinal Chemistry 3 (1995) 945-953.

[3] L.Vidal, et al., High-level production of recombinant His-tagged rhamnulose 1-phosphate aldolase in *Escherichia coli*, Journal of Chemical Technology and Biotechnology 78 (2003) 1171-1179.

[4] L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

[5] L.Vidal, P.Ferrer, G.Alvaro, M.D.Benaiges, G.Caminal, Influence of induction and operation mode on recombinant rhamnulose 1-phosphate aldolase production by *Escherichia coli* using the T5 promoter, Journal of Biotechnology 118 (2005) 75-87.

[6] J.Pinsach, C.de Mas, J.Lopez-Santin, A simple feedback control of *Escherichia coli* growth for recombinant aldolase production in fed-batch mode, Biochemical Engineering Journal 29 (2006) 235-242.

[7] O.Durany, G.Caminal, C.de Mas, J.Lopez-Santin, Studies on the expression of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 39 (2004) 1677-1684.

4. A SIMPLE FEEDBACK CONTROL OF *E. coli* GROWTH FOR RECOMBINANT FUCULOSE 1-PHOSPHATE ALDOLASE PRODUCTION IN FED-BATCH MODE.

Abstract.

Fed-batch production of recombinant fuculose 1-phosphate aldolase (FucA) by *E. coli* XL1 Blue MRF' pTrcfuc was automated by using a simple feedback specific growth rate control strategy. Non-induced continuous cultures were conducted in order to characterize substrate consumption and carbon dioxide production yields and rates. In fed-batch cultures, substrate feeding rate was adjusted using on-line biomass estimation based on exhaust gas analysis and macroscopic mass balances. Overexpression of recombinant protein induced by isopropyl-β-D-thiogalactopyranoside (IPTG) under *trc* promoter did not affect significantly the specific growth rate during 7 hours after induction. Growth and protein production curves were parallel until high level of protein expression started to inhibit cell growth. The proposed specific growth rate control strategy was successfully applied to both non-induced and induced fed-batch cultures.

4.1. Introduction and objectives.

In previous work [1], FucA was produced in substrate-limited fed-batch cultures at constant specific growth rate (μ). Feeding was added following a predefined exponential function calculated assuming a constant apparent biomass-substrate yield. Optimal inducer pulse concentration and induction time were determined in order to obtain maximum productivity of recombinant protein. The metabolic burden associated to recombinant FucA overexpression progressively inhibited culture growth after 7 hours of induction, but growth was not apparently influenced by the expression of the protein during one doubling time in which protein expression and growth curves were parallel (Figure 4.1). However, specific growth rate control using this simple exponential feeding strategy required considerable operator effort to address process deviations from the expected evolution of the culture.
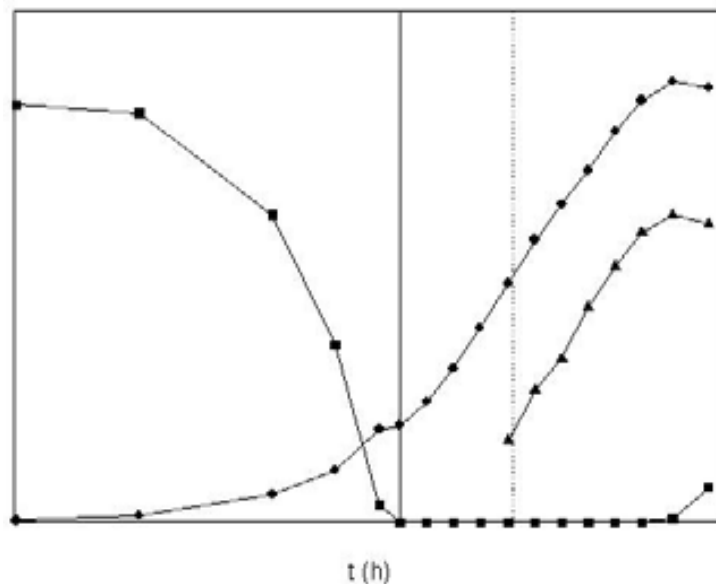

Figure 4.1. Schematic representation of the substrate (■), biomass (●) and product (▲) evolution in a standard pulse-induced fed-batch culture. The start of the substrate feed (—) and an inducer pulse (⋯) are depicted by vertical lines.

Inferential methods applied to feedback control are frequently used to provide a corrective scheme to compensate process disturbances and model inaccuracies [2], although other strategies including fuzzy control and neural networks have also been applied to control fed batch fermentations [3]. The implementation of an inferential control scheme requires the knowledge of the substrate demand, which depends on biomass concentration, specific growth rate and biomass-substrate stoichiometry. Due to the lack of reliable and cheap on-line sensors, indirect specific growth rate [4, 5] and biomass [6, 7] estimation techniques using mass balances based on exhaust gas analysis have been applied. Some of these approaches

considered time-varying model parameters, since cell metabolism changed during the course of the fermentation, especially when metabolism shifted from pure cell growth to product synthesis. As a result, parameter estimation required complex algorithms for nonlinear systems as are fed-batch fermentations.

However, considering an expression system that does not overburden metabolism and a recombinant protein which is not toxic to host cell, a simple model able to describe non-induced growth might also describe induced growth [8, 9]. Since FucA overexpression under promoter *trc* apparently did not affect specific growth rate for 7 hours during the induction period, model parameters obtained from non-induced continuous cultures could be potentially used during this first part of induction phase to control our system. From this moment on, metabolic burden strongly affected cell growth. The lack of knowledge of this second part of induction phase limited any kind of control.

As recombinant protein production at industrial scale requires reproducible processes, the objective of this chapter was to develop a simple feedback growth control method for automated fed-batch production of recombinant FucA based on macroscopic mass balances and exhaust gas analysis. The same strategy could be applied to any other system in which recombinant protein production is growth-associated and does not overburden host cell metabolism too quickly.

4.2. Materials and methods.

Bacterial strain.

E. coli XL1 Blue MRF' pTrcfuc [10] was the recombinant strain used for both continuous and fed-batch cultures. See section *3.1. Strains and vectors* for further details on the expression system. The strain was stored at -80ºC in glycerol stocks using commercial Cryobilles (AES Laboratoire, France) from exponential phase cultures in Luria-Bertani medium supplemented with antibiotic (LB-A).

Media.

Luria-Bertani plus antibiotic was used as basic complete medium. LB-A medium has been described in section *3.2. Media composition*.

A defined mineral medium (MD), with glucose as the sole carbon source was employed in all batch growths. MD medium contained ($g \cdot L^{-1}$): glucose, 20; K_2HPO_4, 13.23; KH_2PO_4, 2.65; NaCl, 2.04; $(NH_4)_2SO_4$, 4.10; $MgSO_4 \cdot 7H_2O$, 0.5; $FeCl_3$, 0.026; ampicillin, 0.1; thiamine, 0.01 and trace element solution, 2.86 $mL \cdot L^{-1}$. See section *3.2. Media composition* for trace element solution composition.

Defined medium composition used to feed continuous cultures (MD-C) was adapted from MD medium. Two different MD-C medium compositions were used. When working at low specific growth rate values, higher biomass concentrations were required to obtain non-noisy measurements. For steady states obtained at 0.05 h^{-1} and 0.10 h^{-1}, MD-C contained ($g \cdot L^{-1}$): glucose, 20; K_2HPO_4, 2.70; KH_2PO_4, 0.54; NaCl, 1.63; $(NH_4)_2SO_4$, 3.28; $MgSO_4 \cdot 7H_2O$, 0.40; $FeCl_3$, 0.02; ampicillin, 0.037; thiamine, 0.008 and trace element solution, 2.26 $mL \cdot L^{-1}$. MD-C used to work from 0.15 h^{-1} to 0.25 h^{-1} contained ($g \cdot L^{-1}$): glucose, 7; K_2HPO_4, 0.94; KH_2PO_4, 0.19; NaCl, 0.57; $(NH_4)_2SO_4$, 1.15; $MgSO_4 \cdot 7H_2O$, 0.14; $FeCl_3$, 0.01; ampicillin, 0.037; thiamine, 0.003 and trace element solution, 0.80 $mL \cdot L^{-1}$.

Feeding medium in fed-batch experiments contained ($g \cdot L^{-1}$): glucose, 479; $(NH_4)_2SO_4$, 8.8; $MgSO_4 \cdot 7H_2O$, 9.50; $FeCl_3$, 0.49; ampicillin, 0.03; thiamine, 0.29 and trace element solution, 63 $mL \cdot L^{-1}$. Phosphates were not included in feeding solution to avoid solubility problems and a separated concentrated solution containing 500 $g \cdot L^{-1}$ K_2HPO_4 and 100 $g \cdot L^{-1}$ KH_2PO_4 was used when necessary.

Analytical methods.

Biomass concentration was determined either as optical density at 600 nm (OD600) or as dry cell weight concentration as described in *3.5. Analytical methods*. One unit of optical density was found to equal 0.36 g DCW·L^{-1} for continuous cultures conditions, 0.25 g DCW·L^{-1} for non-induced fed-batch cultures and 0.29 g DCW·L^{-1} for induced fed-batch cultures.

Concentrations of glucose, acetate and other volatile organic products as lactate and ethanol were estimated as indicated in *3.5. Analytical methods*.

Ammonium was determined by continuous flow analysis [11]. Samples were centrifuged and filtered before diluting with water in order to keep ammonium concentration in the lineal range for this method (from 0.5 to 50 mg·L^{-1}).

A quadrupole mass spectrometer (Balzers Quadstar 422) was used for on-line exhaust gas analysis. Exhaust gas humidity was reduced by using an exhaust gas cooler and a silica column. N_2, O_2, CO_2 and Ar molar fractions were determined with Faraday Cup [12]. Inlet CO_2 air composition was obtained from a 12 h measurement average before inoculation.

FucA activity concentration was estimated by a specific assay [13] and one unit of enzyme activity (UA) was defined as the quantity of enzyme that catalyses the formation of 1 μmol DHAP·min^{-1} at pH=7.5 and 25ºC. Further details on the enzymatic assay can be found in section *3.5. Analytical methods*.

Cultivation conditions.

A single glycerinate unit was inoculated in 15 mL of LB-A medium in a 100 mL shake flask and incubated overnight at 37ºC. This preculture was used to inoculate 200 mL of MD medium in a 1 L shake flask and incubated overnight at 37ºC. After this period, the volume of the culture used to inoculate the fermentor was 10% (v/v) of the starting batch volume, which was 0.8 L in fed-batch experiments and 1 L in continuous cultures.

Fermentations were performed in a Biostat B (Braun Biotech Int.) with 2 L working volume vessel at 37ºC. Ammonium was supplemented as a concentrated base solution (25% w/v NH$_4$OH) employed to control pH and a small amount was included in feed solution to keep ammonium concentration around 1 g·L^{-1}. 2 M H$_2$SO$_4$ and this 25% w/v NH$_4$OH solutions were used to control pH at 7±0.05. Air or pure oxygen was supplied at a controlled flow rate and

agitation speed was varied from 400 to 1000 rpm to keep the dissolved oxygen concentration at 20% of the saturation value.

The initial amount of phosphates in fed-batch experiments supported at least 30-35 g DCW without phosphate limitation. A 10 mL pulse of concentrated phosphates solution was added to bioreactor only when necessary. A discrete addition of 1.8 mmol IPTG was used for induction of FucA overexpression.

Experimental setup and software.

Braun Biostat B digital control unit (DCU) controlled pH, agitation, temperature and dissolved oxygen of bioreactor. Two gas mass flow meters (Bronkhorst, EL-FLOW) were used to control and to measure inlet and outlet gas mass flow, respectively. Exhaust gas flow was continuously sampled and analysed by a mass spectrometer (MS). Two peristaltic pumps (Cole Parmer Instrument Company, Masterflex) were used for feeding and broth withdrawing during continuous cultures (Figure 4.2), and a microburette (Crison Instruments, MICRO BU 2030) with a 2.5 mL syringe (Hamilton) was used in fed-batch cultures for discrete feed addition (Figure 3.5, see section *3.3. Experimental setup. Hardware and software.*).
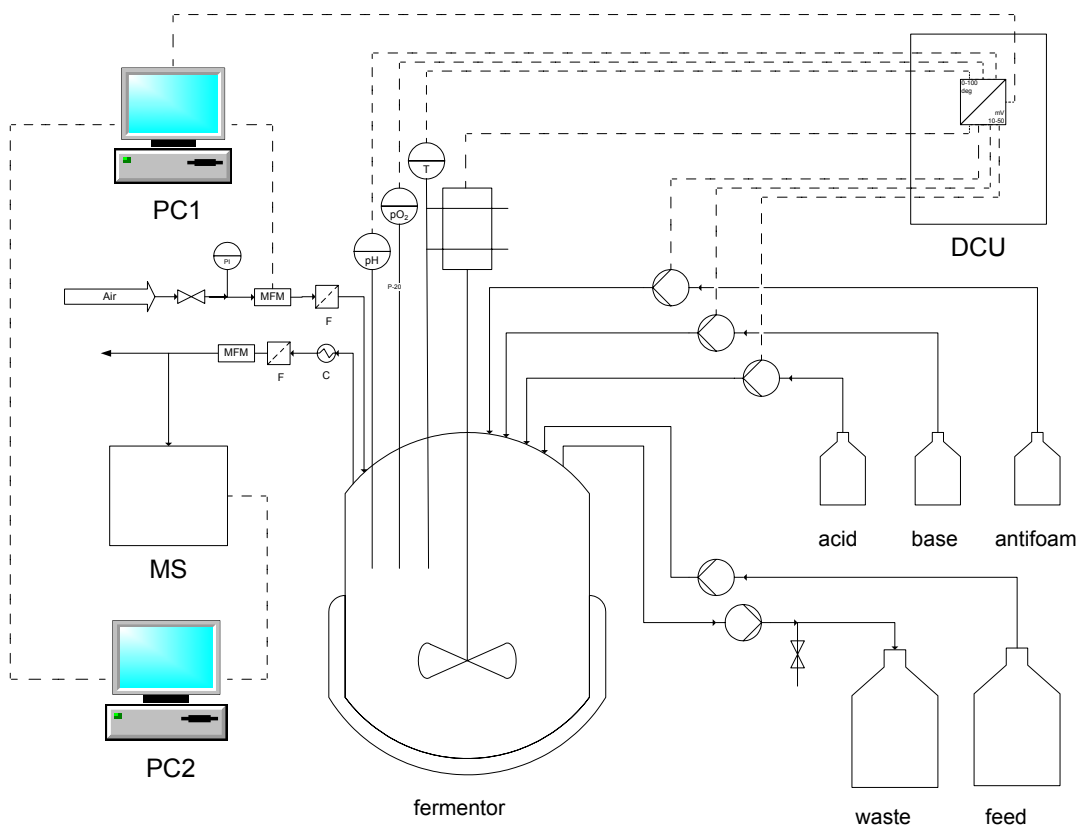

Figure 4.2. Schematic representation of the experimental setup used to perform continuous cultures.

Specific software was developed (using Microsoft Visual Basic 6.0.) to configure communication between devices, store acquired data in a database and monitor process variables along the fermentation (further details can be found in section *3.3. Experimental setup. Hardware and software.*). It also allowed on-line calculation of different process variables such as culture volume, carbon exhaust rate and volume of feed to supply. Data acquisition, treatment, storage and actuation frequency was 1 minute.

These experimental values were combined with substrate consumption and carbon dioxide production models to implement the specific growth rate control. Substrate mass balance was used as control law to determine feed rate, and CO_2 production model combined with exhaust gas analysis allowed on-line biomass estimation for feedback corrections.

4.3. Control scheme and model considerations.

4.3.1. Control scheme.

Exponential feeding based on substrate consumption model with feedback biomass correction was used for specific growth rate control. Biomass estimation was determined on-line by mass spectrometry exhaust gas analysis and carbon dioxide production model as shown in Figure 4.3.
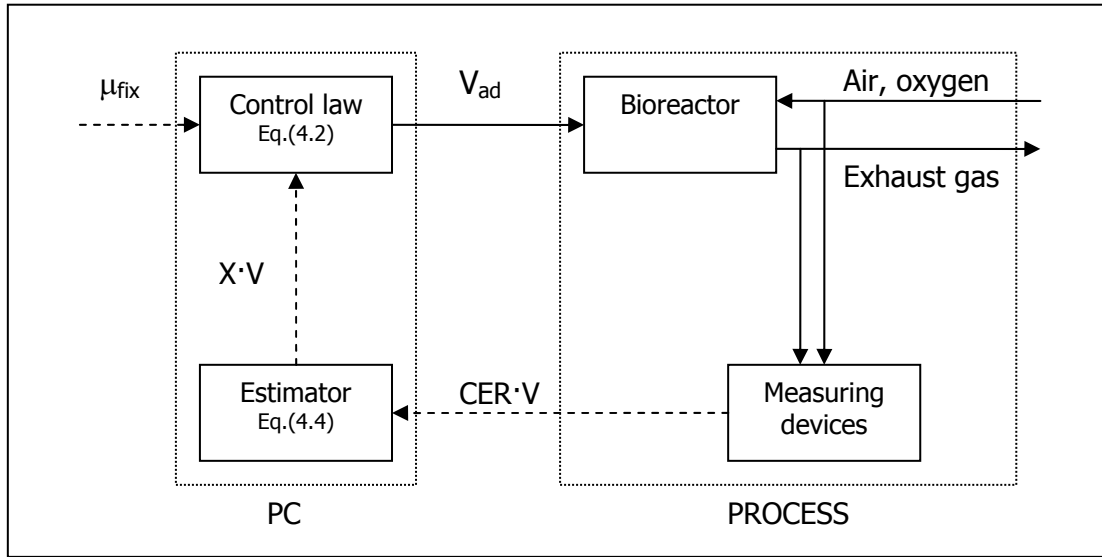

Figure 4.3. Control configuration scheme.

4.3.2. Process model description.

Considering substrate-limited fed-batch operation, combination of biomass and substrate mass balances led to:

$$\frac{F_S \cdot S_F}{X \cdot V} = \frac{1}{Y_{xs}} \cdot \frac{1}{X \cdot V} \cdot \frac{d(X \cdot V)}{dt} + m_{sx} \qquad \text{(Eq. 4.1)}$$

Since constant biomass-substrate yield and maintenance coefficient values were assumed, feeding rate profile could be obtained by integration of Eq. 4.1. Considering discrete additions, the feed volume to be added to bioreactor in a time interval was given by the following expression:

$$V_{ad}(t) = \frac{1}{S_F} \cdot \left(\frac{m_{sx}}{\mu_{fix}} + \frac{1}{Y_{xs}}\right) \cdot X(t) \cdot V(t) \cdot (\exp(\mu_{fix} \cdot \Delta t) - 1) \qquad \text{(Eq. 4.2)}$$

Utilisation of Eq. 4.2 as control law to maintain specific growth rate at the set point value (μ_{fix}) required the knowledge of total biomass along the fed-batch phase (X·V) as well as the biomass-substrate yield and maintenance coefficient.

Typical inferential methods based on exhaust gas analysis use oxygen consumption or carbon dioxide production models as estimators for biomass. However, oxygen consumption was not as accurate as carbon dioxide production model because differences between outlet gas and inlet air oxygen concentration measurements were small and highly affected by noise, especially when low biomass concentrations were present in bioreactor (data not shown). Moreover, the use of pure oxygen when high cell densities were achieved required different model parameters. Therefore, carbon dioxide production model was used because of its better biomass estimation performance.

Neglecting variations on dissolved carbon dioxide concentration [14], the following relationship was obtained by combination of biomass and CO_2 mass balances for fed-batch operation:

$$\frac{CER{\cdot}V}{X{\cdot}V} = \frac{1}{Y_{XCO_2}}{\cdot}\frac{1}{X{\cdot}V}{\cdot}\frac{d(X{\cdot}V)}{dt} + m_{CO_2 X} \qquad \text{(Eq. 4.3)}$$

Assuming constant biomass-carbon dioxide yield and maintenance coefficient, total biomass could be estimated by rearranging and integrating Eq. 4.3:

$$X(t){\cdot}V(t) = \exp(-m_{CO_2 X}{\cdot}Y_{XCO_2}{\cdot}t){\cdot}\left[X_0{\cdot}V_0 + \int_{t_0}^{t} Y_{XCO_2}{\cdot}(CER{\cdot}V){\cdot}\exp(m_{CO_2 X}{\cdot}Y_{XCO_2}{\cdot}t){\cdot}dt \right]$$

$$\text{(Eq. 4.4)}$$

Eq. 4.4 would allow on-line biomass estimation, provided that model parameters are known and on-line exhaust gas analysis is available. Total biomass estimation could then be used as a feedback correction in Eq. 4.2.

This simple control scheme (Figure 4.3) did not require data filtering. Eq. 4.4 eliminated (CER·V) noise by integrating all experimental values along the fermentation, obtaining a non-noisy biomass estimation. Furthermore, the use of calculated (X·V) from experimentally measured (CER·V) made the feeding rate independent of calculated culture volume in Eq. 4.2, which enhanced the robustness of this method.

4.3.3. Determination of model parameters.

In order to estimate the parameters relating substrate consumption and CO_2 production to specific growth rate, continuous cultures were performed, as this information is most effectively obtained through the use of this operational strategy [15].

Considering steady state operation (μ=D), substrate mass balance was applied to bioreactor to obtain:

$$q_s = m_{sx} + \frac{1}{Y_{xs}} \cdot \mu$$
(Eq. 4.5)

Specific substrate consumption rate could be calculated from experimental measurements by using the following expression:

$$q_s = \frac{Q_e \cdot S_F - Q_s \cdot S}{X \cdot V}$$
(Eq. 4.6)

Neglecting carbon dioxide accumulation in the culture broth, carbon dioxide production in continuous culture mass balance led to:

$$q_{CO_2} = m_{CO_2X} + \frac{1}{Y_{XCO_2}} \cdot \mu$$
(Eq. 4.7)

Specific CO_2 production rate could be calculated from experimental measurements using Eq. 4.8:

$$q_{CO_2} = \frac{G_s \cdot x_{CO_2,s} - G_e \cdot x_{CO_2,e}}{X \cdot V} = \frac{CER \cdot V}{X \cdot V}$$
(Eq. 4.8)

4.4. Results and discussion.

4.4.1. Determination of model parameters.

Non-induced continuous cultures were conducted to determine parameters for both substrate consumption and carbon dioxide production models in metabolically stable and substrate-limited growth conditions. Once continuous cultures reached the steady state, stable values of measured variables were obtained. Three residence times after setting a constant flow rate, averages for different measured variables were calculated for each one of the steady states. Five different steady states were reached at specific growth rates from 0.05 to 0.25 h^{-1}.

Eq. 4.5 was used to obtain substrate consumption model parameters (Figure 4.4) and Eq. 4.7 was used to obtain CO_2 production model parameters (Figure 4.5). Considering 95% confidence [16], the obtained inverse yield for glucose ($1/Y_{XS}$) was 3.19±0.20 g substrate·g DCW^{-1}, glucose maintenance coefficient (m_{SX}) was 0.01±0.01 g substrate·g $DCW^{-1}·h^{-1}$, inverse yield for CO_2 ($1/Y_{XCO2}$) was 48.10±6.36 mmol CO_2·g DCW^{-1} and CO_2 maintenance coefficient (m_{CO2X}) was -0.11±0.45 mmol CO_2·g $DCW^{-1}·h^{-1}$. Maintenance coefficients were meaningless and were consequently neglected without significant repercussion on both model predictions. Confidence interval values were 6% for glucose consumption and 13% for carbon dioxide production inverse yields.
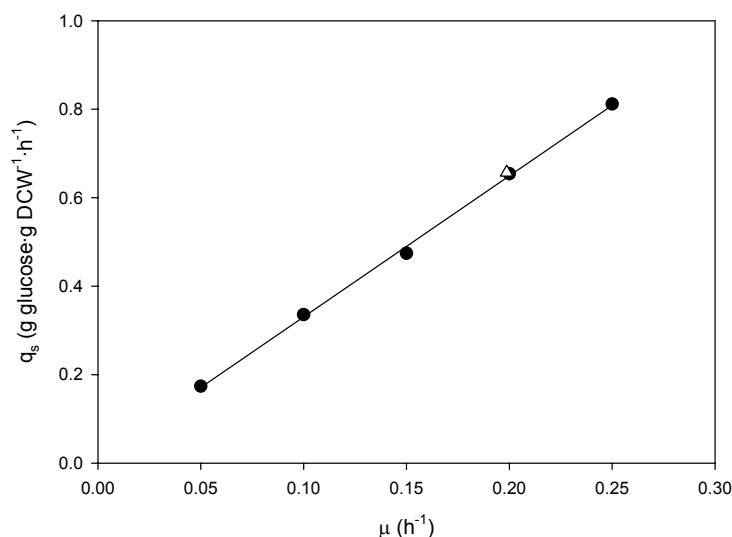

Figure 4.4. Substrate consumption model (—) obtained from non-induced continuous cultures (●) –present section-. Time averaged values for specific substrate consumption and specific growth rates from the open-loop fed-batch culture (Δ) –see section 4.4.2-.
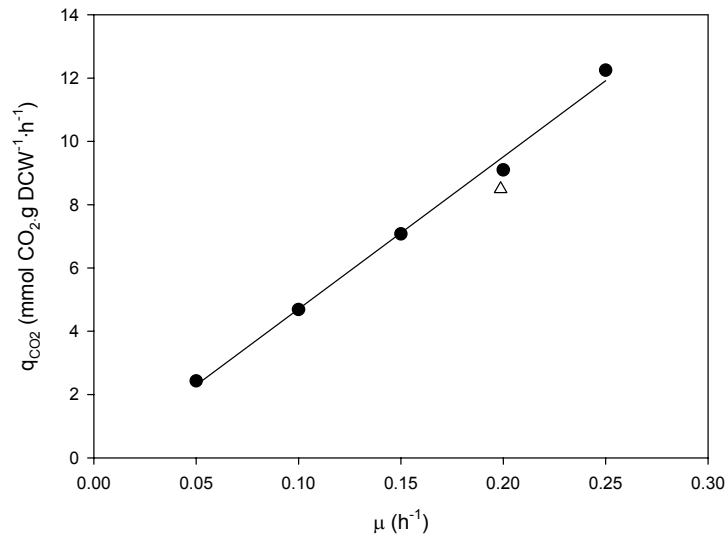
Figure 4.5. Carbon dioxide production model (—) obtained from non-induced continuous cultures (●) –present section-. Time averaged values for specific carbon dioxide production and specific growth rates from the open-loop fed-batch culture (Δ) –see section 4.4.2-.

4.4.2.   Model validation. Open-loop fed-batch cultures.

A non-induced fed-batch culture was performed to validate parameters for both models obtained from continuous cultures in high cell density culture conditions. A predefined exponential growth curve considering a constant specific growth rate of 0.20 h^{-1} allowed the calculation of an exponential addition profile according to Eq. 4.2 and the biomass-glucose yield previously determined. This predefined feeding policy was successfully implemented, and specific growth rate value obtained from OD600 measurements was 0.20 h^{-1} during the fed period (Figure 4.6).
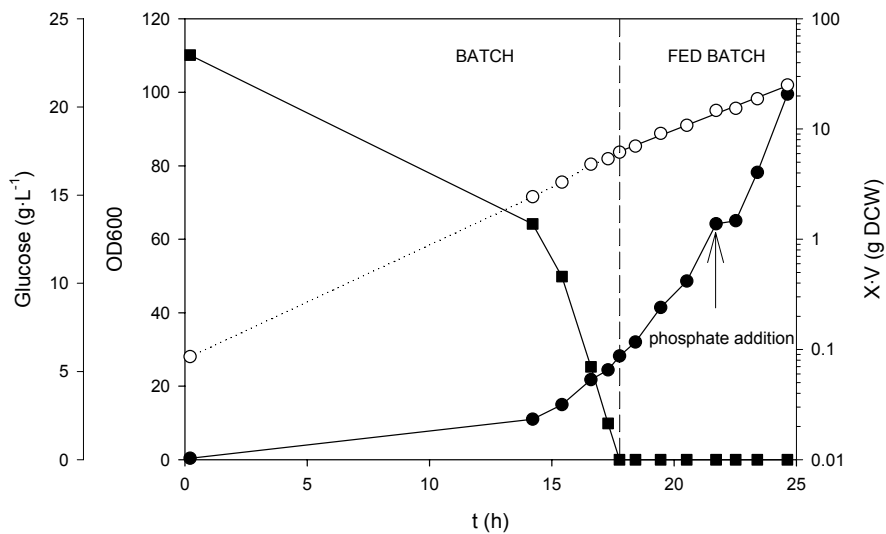

Figure 4.6. Open-loop fed-batch culture evolution of total biomass (o), optical density at 600 nm (●) and substrate concentration (■). μ=0.20 h^{-1}.

Using exhaust gas measurements and Eq. 4.4, off-line model prediction of biomass evolution could be calculated and compared to experimentally measured biomass (Figure 4.7). Average error from model predictions to experimental biomass measurements was below 5% during the fed-phase. Total biomass was estimated from (CER·V) values in a robust manner because it was only slightly affected by disturbances and noise in experimental measurements. (CER·V) evolution from exhaust gas analysis is presented in Figure 4.8, where the end of the batch phase at 17.8 h, the change from air to pure oxygen at 19.2 h and a discrete phosphate addition at 21.6 h can be easily seen.
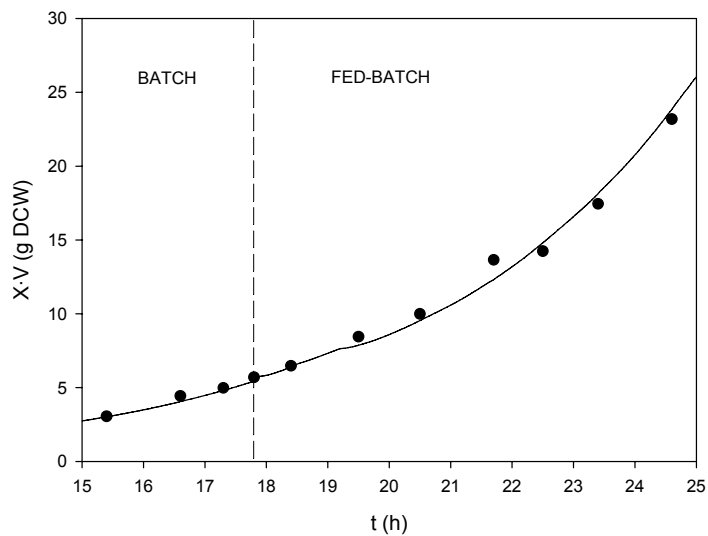

Figure 4.7. Total biomass during the end of batch phase and during the open-loop fed-batch phase: model prediction (—) and experimental biomass evolution (●).
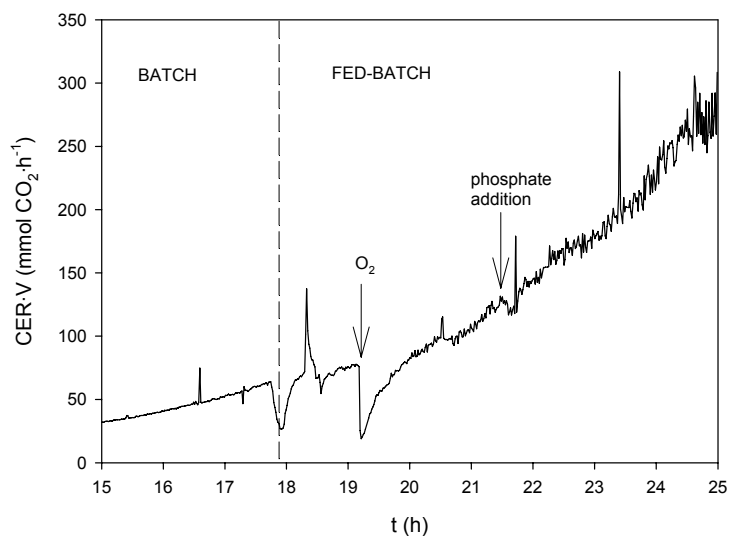

Figure 4.8. (CER·V) evolution during the open-loop fed-batch culture.

In order to validate parameters estimation from continuous to high cell density culture conditions, Eq. 4.1 was integrated obtaining:

$$\frac{\int_{t=0}^{t}\frac{F_{s}\cdot S_{0}}{X\cdot V}\cdot dt}{\int_{t=0}^{t}dt}=\frac{1}{Y_{xs}}\cdot\frac{\int_{X_0\cdot V_0}^{X\cdot V}\frac{d(X\cdot V)}{X\cdot V}}{\int_{t=0}^{t}dt}+m_{sx}$$

(Eq. 4.9)

where the left side was the time average for specific glucose consumption rate and the first term on the right is the time average for specific growth rate divided by Y_{XS}. In fed-batch experiments at a constant specific growth rate, Eq. 4.9 allows the calculation of Y_{XS} and m_{SX}.

Eq. 4.3 was also integrated obtaining an expression which allowed CO_2 production model parameters validation in fed-batch cultures:

$$\frac{\int_{t=0}^{t}\frac{CER\cdot V}{X\cdot V}\cdot dt}{\int_{t=0}^{t}dt}=\frac{1}{Y_{XCO_2}}\cdot\frac{\int_{X_0\cdot V_0}^{X\cdot V}\frac{d(X\cdot V)}{X\cdot V}}{\int_{t=0}^{t}dt}+m_{CO_2X}$$

(Eq. 4.10)

where left side was the time average for specific carbon dioxide production rate and the first term on the right represents the time average for specific growth rate divided by Y_{XCO2}. In fed-batch experiments at a constant specific growth rate, Eq. 4.10 allows the calculation of Y_{XCO2} and m_{CO2X}.

All integrals were numerically calculated using experimental data along the fermentation in order to reduce data noise.

The use of Eq. 4.9 and Eq. 4.10 allowed direct comparison to parameters obtained from continuous cultures. The obtained values have been used to plot the specific substrate consumption rate in Figure 4.4 and the specific carbon dioxide production rate in Figure 4.5, reasonably fitting the previously determined models.

Therefore, Eq. 4.2 using substrate-biomass yield obtained from continuous cultures was validated as control law and Eq. 4.4 using carbon dioxide-biomass yield obtained from continuous cultures was validated as biomass estimator.

4.4.3. Closed-loop fed-batch cultures.

Two closed-loop fed-batch cultures were performed. The first one was a non-induced culture to test the specific growth rate control using on-line biomass estimation for feedback control. This strategy was applied to a second induced fed-batch to obtain a controlled recombinant FucA production process.

4.4.3.1. Feedback control of non-induced growth.

Once parameters had been validated in high cell density cultures, on-line biomass estimation allowed the specific growth rate feedback control implementation. A specific growth rate set point value of 0.15 h^{-1} was chosen to avoid acetate accumulation. Oxygen supply limited the final dry cell weight concentration to approximately 40 g $DCW \cdot L^{-1}$ ($f_{gDCW/OD600}$=0.25 for non-induced cells). The fermentation ran automatically under the control of Eq. 4.2 where (X·V) was estimated from exhaust gas analysis and Eq. 4.4. Experimental value of μ was 0.16 h^{-1}, quite close to the set point (Figure 4.9).
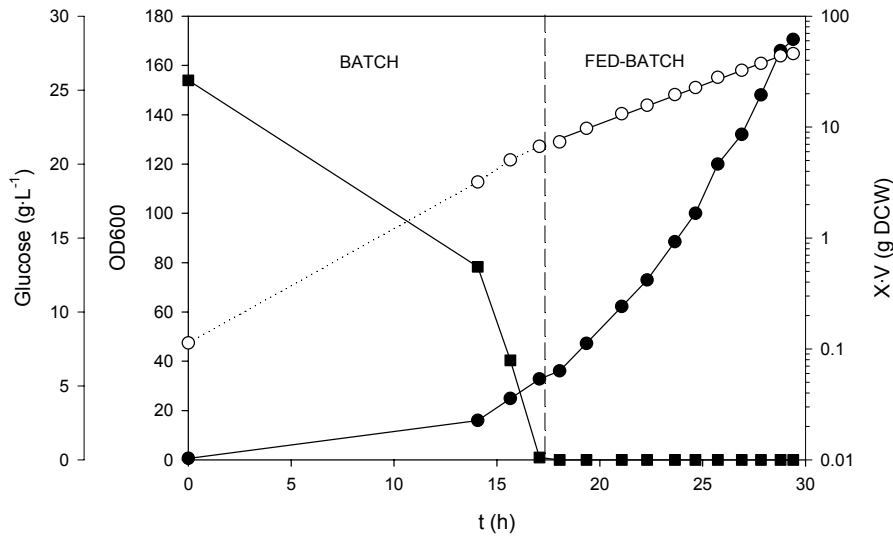

Figure 4.9. Non-induced closed-loop fed-batch evolution of of total biomass (o), optical density at 600 nm (●) and substrate concentration (■). μ=0.16 h^{-1} (μ_{fix}=0.15 h^{-1}).

Disturbances in CO_2 measurements due to change from air to pure oxygen at 21.5 h and a pulse addition of 10 mL of concentrated solution of phosphates at 25 h were absorbed by integration of measured (CER·V) values in Eq. 4.4 without any need of data filtering (data not shown). Exponential growth was controlled along the fed phase proving the robustness of the control strategy.

4.4.3.2. Feedback control of induced growth. Recombinant aldolase production.

Specific growth rate set point for the non-induced fed-batch phase was 0.15 h^{-1} and it was reduced to 0.10 h^{-1} for the induction phase. These values were low enough to avoid high oxygen transfer rates, to keep low acetate concentrations and to allow proper folding of the recombinant protein [1]. The proposed control system also worked well, and the experimental specific growth rate value was 0.16 h^{-1} during the non-induced fed-batch phase and 0.09 h^{-1} during the induction phase (Figure 4.10). Final dry cell weight concentration was, as in non-induced experiments, around 40 g DCW·L^{-1} ($f_{gDCW/OD600}$=0.29 for induced cells), proving that induction time was optimal (oxygen supply becomes limiting 7 hours after induction, when metabolic burden starts to inhibit cell growth). In spite of noise and disturbances in (CER·V) measurements (as the change from air to pure oxygen at 24 h, the change of specific growth rate setpoint at 27 h, or induction at 28 h), model estimation of total biomass was similar to experimentally measured values.
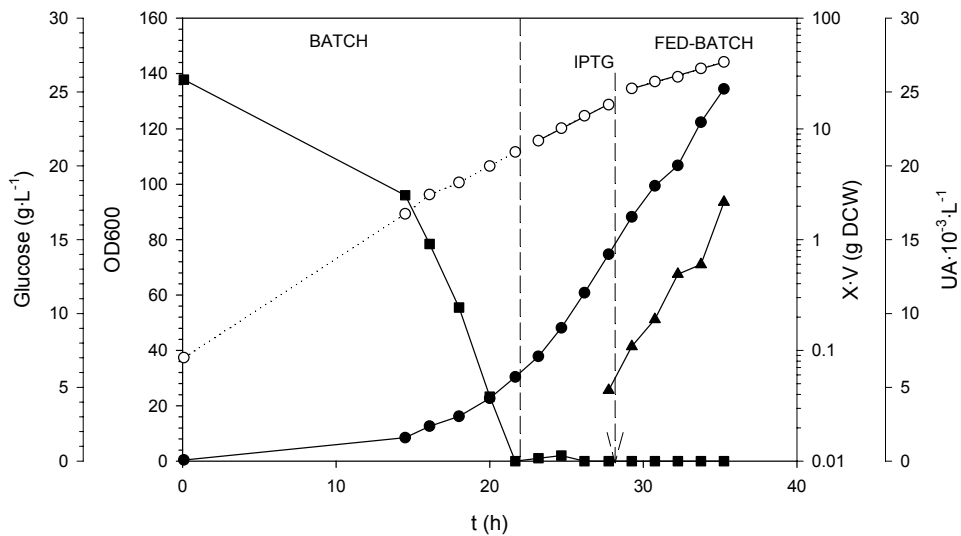

Figure 4.10. Induced closed-loop fed-batch evolution of total biomass (o), optical density at 600 nm (●), substrate concentration (■) and FucA concentration in units of activity (UA) per litre (▲). Before induction: μ=0.16 h^{-1} (μ_{fix}=0.15 h^{-1}). After induction: μ=0.09 h^{-1} (μ_{fix}=0.10 h^{-1}).

Growth control was not apparently affected by recombinant protein overexpression during one doubling time. During this period production and growth curves were parallel (Figure 4.10) as previously reported. Thus, the hypothesis that models obtained from non-induced cultures were able to describe the first part of the induction phase has been validated for our expression system and recombinant protein.

4.5. Conclusions.

Recombinant FucA production in fed-batch mode was automated by a simple feedback specific growth rate control method. Models for substrate consumption and carbon dioxide production obtained from non-induced continuous cultures were employed and modelling and sensitivity problems were avoided. On-line carbon dioxide production rate data provided the required biomass estimation, which was used to adjust feeding rate to control specific growth rate during the fed-phase, using substrate consumption model as a control law.

Recombinant FucA production under the control of *trc* promoter in *E. coli* XL1 Blue MRF' pTrcfuc was growth-associated and did not affect significantly the proposed specific growth rate control system for 7 hours. Significant amount of recombinant protein was obtained during this period. This confirmed the proposed strategy as a robust method to control growth-associated recombinant protein production in fed-batch mode.

This approach to fermentation control might be applicable to other similar systems as well, particularly to those in which protein production is growth associated and induction does not produce a fast severe metabolic burden of host cell.

References.

[1] O.Durany, C.de Mas, J.Lopez-Santin, Fed-batch production of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 40 (2005) 707-716.

[2] Z.M.Nor, M.I.Tamer, J.M.Scharer, M.Moo-Young, E.J.Jervis, Automated fed-batch culture of *Kluyveromyces fragilis* based on a novel method for on-line estimation of cell specific growth rate, Biochemical Engineering Journal 9 (2001) 221-231.

[3] J.Lee, S.Y.Lee, S.Park, A.P.Middelberg, Control of fed-batch fermentations, Biotechnology Advances 17 (1999) 29-48.

[4] L.Cazzador, V.Lubenova, Nonlinear estimation of specific growth rate for aerobic fermentation processes, Biotechnology and Bioengineering 47 (1995) 626-632.

[5] D.Riesenberg, et al., High cell density fermentation of recombinant *Escherichia coli* expressing human interferon alpha 1, Applied Microbiology and Biotechnology 34 (1990) 77-82.

[6] M.A.Gordillo, et al., Enhancement of *Candida rugosa* lipase production by using different control fed-batch operational strategies, Biotechnology and Bioengineering 60 (1998) 156-168.

[7] S.B.Petkov, R.A.Davis, On-line biomass estimation using a modified oxygen utilization rate, Bioprocess and Biosystems Engineering 15 (1996) 43-45.

[8] R.S.Donovan, C.W.Robinson, B.R.Glick, Review: optimizing inducer and culture conditions for expression of foreign proteins under the control of the *lac* promoter, Journal of Industrial Microbiology and Biotechnology 16 (1996) 145-154.

[9] W.Kramer, G.Elmecker, R.Weik, D.Mattanovich, K.Bayer, Kinetic studies for the optimization of recombinant protein formation, Recombinant DNA Biotechnology III: the Integration of Biological and Engineering Sciences 782 (1996) 323-333.

[10] E.García-Junceda, G.J.Shen, T.Sugai, C.H.Wong, A New Strategy for the Cloning, Overexpression and One-Step Purification of 3 DHAP-Dependent Aldolases - Rhamnulose-1-Phosphate Aldolase, Fuculose-1-Phosphate Aldolase and Tagatose-1,6-Diphosphate Aldolase, Bioorganic & Medicinal Chemistry 3 (1995) 945-953.

[11] J.Baeza, D.Gabriel, J.Lafuente, An expert supervisory system for a pilot WWTP, Environmental Modelling and Software 14 (1999) 383-390.

[12] E.Heinzle, Present and potential applications of mass spectrometry for bioprocess research and control, Journal of Biotechnology 25 (1992) 81-114.

[13] O.Durany, G.Caminal, C.de Mas, J.Lopez-Santin, Studies on the expression of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 39 (2004) 1677-1684.

[14] P.N.Royce, Effect of changes in the pH and carbon dioxide evolution rate on the measured respiratory quotient of fermentations, Biotechnology and Bioengineering 40 (1992) 1129-1138.

[15] R.D.Kiss, G.Stephanopoulos, Metabolic activity control of the L-lysine fermentation by restrained growth fed-batch strategies, Biotechnology Progress 7 (1991) 501-509.

[16] C.Lipson, N.J.Sheth, Statistical design and analysis of engineering experiments, Mc. Graw-Hill Inc. (1971).

5. INDUCTION STRATEGIES IN *Escherichia coli* FED-BATCH CULTURES FOR RECOMBINANT RHAMNULOSE 1-PHOSPHATE ALDOLASE PRODUCTION.

Abstract.

Different induction strategies for recombinant protein production under the control of the strong T5 promoter in *Escherichia coli* have been investigated. When the production of the recombinant protein is growth-related, the productivity of the process can be strongly reduced due to the negative effects of protein expression on cell growth. IPTG pulse induction as well as inducer dosage have been implemented for recombinant rhamnulose 1–phosphate aldolase (RhuA) production, and their advantages and drawbacks have been highlighted. Both strategies led to high levels of the recombinant protein, approximately 1000 AU·gDCW^{-1}. Inducer concentration and inducer to biomass ratio were identified as the parameters influencing the rate of protein production and final enzymatic activity per gram of biomass, respectively. In pulse induction, the maximum enzymatic activity was found at inducer concentration of 70 μM, while in experiments where continuous induction was applied, inducer concentrations between 4 and 12 μM were identified as the working range within cell growth and recombinant protein accumulation occurred simultaneously. On the other hand, the amount of IPTG per gram of biomass required was 1.6 μmol IPTG·gDCW^{-1} in pulse induction and between 0.3-0.5 μmol IPTG·gDCW^{-1} in continuous induction.

5.1. Introduction and objectives.

Recombinant rhamnulose 1-phosphate aldolase (RhuA) had been previously cloned under the control of the weak *trc* promoter in *E. coli* XL1 Blue MRF' pTrcrham [1]. Preliminary expression studies had shown low RhuA yields when using this expression system, so the *rhaD* gene was excised from pTrcrham and cloned under the control of the strong T5 promoter in the *E. coli* M15 [pREP4] pQE40 system, obtaining significantly higher target protein yields (up to 50% of the total intracellular proteins) [2]. Owing to these promising results, the obvious next step was to develop the production process using *E. coli* M15 [pREP4] pQErham to maximise the productivity, *i.e.*, to obtain both high cell densities as well as high specific product concentrations.

During the production process development stage, *E. coli* M15 [pREP4] pQErham showed lower activities of target protein per gram of biomass in fed-batch cultures than in the case of batch cultures [3]. These results were related to the fact that, in fed-batch mode, part of the cells had lost the pQErham plasmid after IPTG pulse-induction. In order to avoid frequent antibiotic addition into the bioreactor to maintain the expression plasmid, the expression system was modified to the antibiotic-free plasmid maintenance construction *E. coli* M15 ∆glyA [pREP4] pQEαβrham [4]. The *glyA* gene (coding for serine hydroxymethyl transferase –SHMT-) deletion from the host chromosome resulted in *E. coli* M15 ∆glyA, which is a glycine auxotrophic strain. After that, the *glyA* gene was inserted into the expression vector under the control of the P3 constitutive promoter leading to *E. coli* M15 ∆glyA [pREP4] pQEαβrham, an improved expression system which prevents plasmid loss avoiding the need of antibiotic supplementation in defined media.

Even though the expression system based on the strong T5 promoter required a lower amount of IPTG than in the case of the weak *trc* promoter to obtain higher specific RhuA yields, it was found to be more sensitive to IPTG induction. It is well known that one of the major constraints of recombinant protein production in *Escherichia coli* is the metabolic burden imposed on host cells by the maintenance and expression of the foreign DNA [5]. The host metabolism can be unbalanced due to higher demand of resources, and a loss of productivity might be observed if cell growth is affected before maximum product levels are obtained. According to our results, expression systems based on strong promoters are prone to suffer from metabolic burden when strong induction of protein expression is performed, especially when grown under glucose limiting conditions (under substrate limitation, cells already experience a shortage of raw materials and energy). Hence, the period of recombinant gene expression can be significantly reduced, and therefore, maximal yield cannot be attained.

To extend the production phase and achieve optimal yields in systems using strong promoters, alternative strategies based on the modulation of the transcription rate by continuous inducer feeding in a constant ratio to biomass had been successfully applied [6].

In this chapter, the above concepts are presented for the production of recombinant rhamnulose 1-phosphate aldolase. The aim was to evaluate pulse induction and inducer dosage strategies in order to exploit host cell at its maximum capabilities, highlighting the advantages and drawbacks of both strategies when applied to substrate-limited high cell density cultures.

5.2. Materials and methods.

Bacterial strain and plasmids.

The K-12 derived strain *E. coli* M15 ΔglyA [pREP4] harbouring the vector pQEαβrham was used for rhamnulose 1-phosphate aldolase overexpression. Frozen stock aliquots containing glycerol prepared from exponential phase cultures grown in Luria-Bertani media (LB) were stored at -80ºC. For further details on the expression system, see section *3.1. Strains and vectors*.

Media composition.

LB medium was used for preinocula preparation. A defined mineral medium was used for inocula, shake flask and bioreactor experiments. Their composition is detailed in *3.2. Media composition*.

Cultivation conditions.

All preinocula, inocula and bioreactor experiments were performed as described in *3.4. Cultivation conditions*.

Induction was performed by adding IPTG in two different ways depending on the experiments (the strategy employed in each case is indicated in section *5.3. Results and discussion*). For the first group of experiments, IPTG was added in a single pulse. For the second group, it was slowly supplied as a continuous addition, regulating the induction action. Thus, the software used to monitor and control the experimental setup was accordingly modified for actuation over this new manipulated variable (Figure 5.1). A second microburette (Crison, Burette 1S-D) with a 1 mL syringe (Hamilton) was used to implement the IPTG feeding profile, which was either predefined to keep a constant ratio of IPTG to biomass (Eq. 5.1) or to keep a constant volumetric IPTG concentration (Eq. 5.2), depending on the case.

$$F_I(t_i) = \frac{f_{I/X} \cdot (XV_i - XV_{i-1})}{I_F \cdot \Delta t}$$
(Eq. 5.1)

where $F_I(t_i)$ is the inducer feed rate (L·h^{-1}) at a time t_i, $f_{I/X}$ is the ratio of IPTG to biomass set point (μmol·gDCW^{-1}), XV_i is the total biomass in the bioreactor (gDCW) at a time t_i, XV_{i-1} is the total biomass in the bioreactor (gDCW) at a time t_{i-1}, I_F is the inducer feed concentration (μmol·L^{-1}) and Δt is the actuation interval ($\Delta t = t_i - t_{i-1}$) (h).

$$F_I(t_i) = \frac{I_{fix} \cdot (V_i - V_{i-1})}{I_F \cdot \Delta t}$$

(Eq. 5.2)

where I_{fix} is the IPTG set point concentration ($\mu mol \cdot L^{-1}$), V_i is the culture volume (L) at a time t_i and V_{i-1} is the culture volume (L) at a time t_{i-1}.
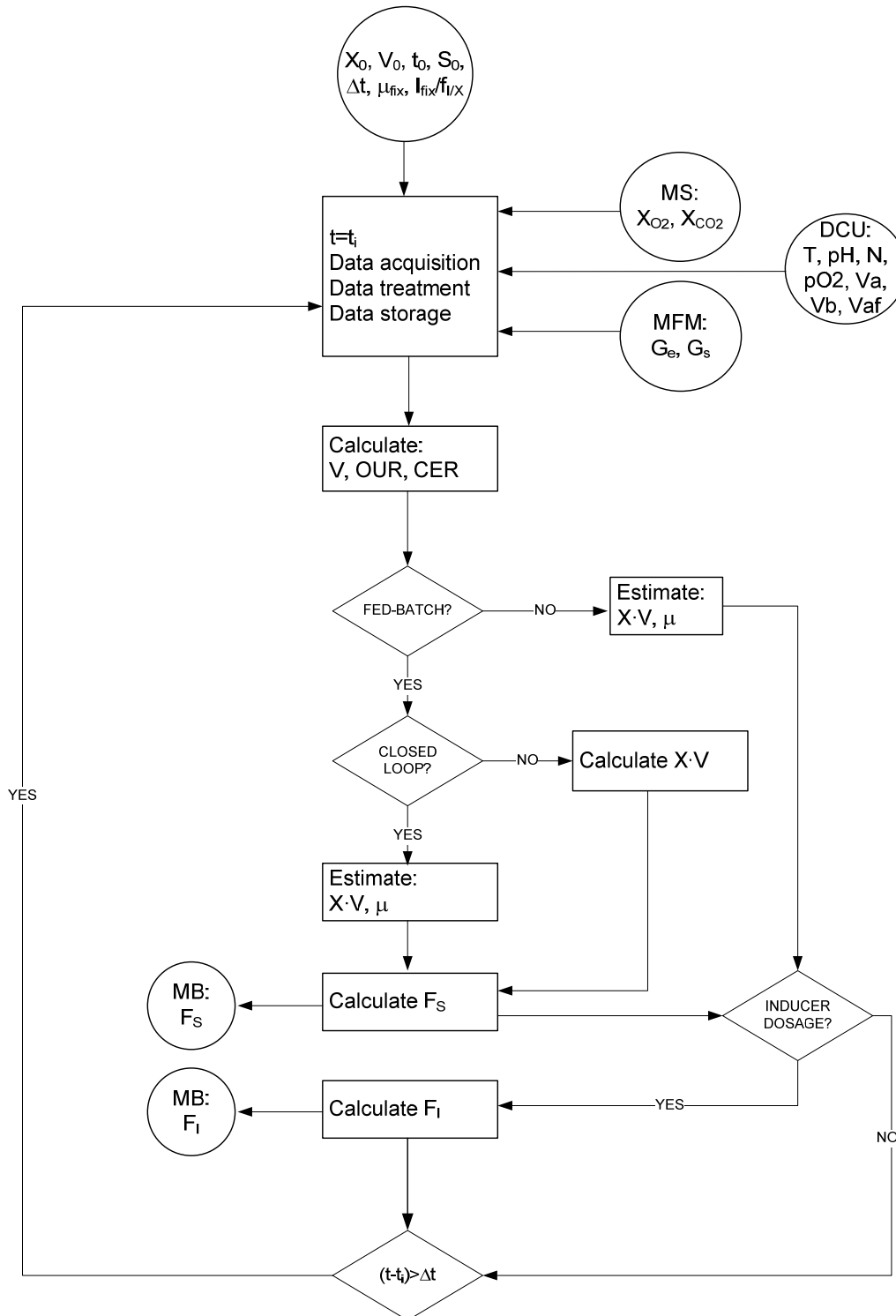

Figure 5.1. Flowchart of the developed software.

Analytical methods.

Growth was followed by optical density at 600 nm (OD600) and by dry cell weight (DCW) measurements as described in *3.5. Analytical methods*.

Glucose, organic acids, ammonium and phosphate measurements were performed as described in *3.5. Analytical methods*.

Determination of RhuA activity was performed as described in section *3.5. Analytical methods*. One unit of RhuA activity was defined as the amount of enzyme required to convert 1 μmol of rhamnulose 1-phosphate to DHAP per minute at 25ºC under the assay conditions.

5.3. Results and discussion.

5.3.1. Preliminary expression studies in shake flask cultures.

Preliminary studies were performed in shake flasks to characterize RhuA overexpression after IPTG induction. Six parallel cultures were grown at 5 g·L^{-1} of initial glucose concentration and, during the exponential growth phase, an early pulse of IPTG (at 0.13 gDCW·L^{-1} approximately) was added into five of them so that all cultures could potentially duplicate (while induced) at least twice before growth was limited by other factors than carbon source (cells grew unlimited for 4 h after induction). The volumetric IPTG concentration after the pulse addition ranged from 0 to 50 μM depending on the culture, which supposed specific inducer concentrations at induction time from 0 to 373 μmol·gDCW^{-1}. Biomass, substrate and product concentrations were measured along time in all cultures. From these data, maximum specific activities, maximum specific production rates and specific growth rates relative to the uninduced control culture were calculated. Figure 5.2 depicts the relationship of these variables to IPTG concentration and to IPTG to biomass ratio.

The best results were observed in the culture induced with a 20 μM or 153 μmol·gDCW^{-1} IPTG pulse. In this case, the maximum accumulated specific activity was 1560 AU·gDCW^{-1}. Below this optimal IPTG level, the maximum specific activities decreased depending on the pulsed amount of inducer, and they did not increase above it. When looking at the maximum specific production rates, results showed that they increased with the pulsed IPTG, but between 10 and 20 μM (corresponding to 79 and 153 μmol·gDCW^{-1}, respectively) they remained nearly constant and close to 1000 AU·gDCW^{-1}·h^{-1}. The specific growth rates measured in the different cultures also showed different behaviour below and above this inducer range. While below 20 μM (or 153 μmol·gDCW^{-1}) growth was unaffected with respect to the control culture, the specific growth rate was significantly reduced above this critical inducer levels as the pulsed amount of IPTG increased.

In order to try to explain this behaviour, one could take into account inducer uptake. IPTG can be transported into the cells by different mechanisms [7]. It has been shown that lactose permease (the *lacY* gene product) takes part into transporting the IPTG across the cell membrane [8]. Since this is an active transport mechanism, IPTG induces the *lac* operon and can be taken up even at low concentrations. However, when the inducer concentration is high enough, other unspecific mechanisms of transport (or diffusion) are also active [9]. Both mechanisms are usually active, but one can predominate over the other depending on the culture conditions.
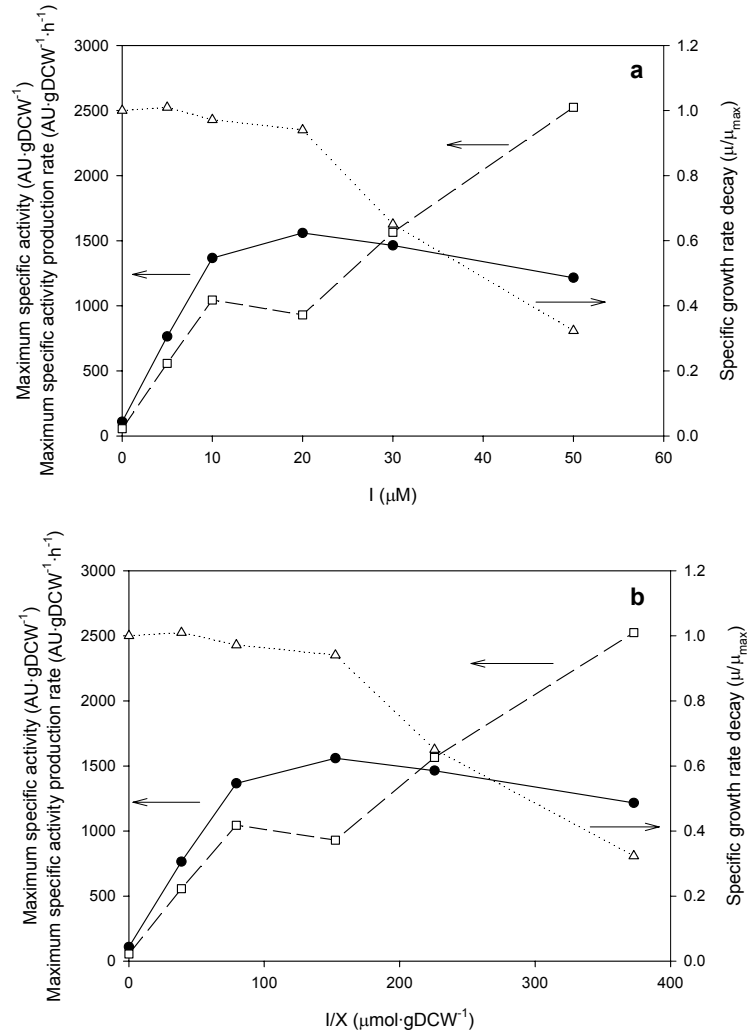
Figure 5.2. Maximum specific activities (—●—), specific activity production rates (- □ -) and specific growth rate reduction due to induction (····△····) in shake flask cultures versus IPTG concentration (a) and versus IPTG to biomass ratio (b).

In our case, it seemed that IPTG was mainly uptaken through lactose permeases when its concentration was below 20 μM. In those cultures, it took some time to reach the maximum specific production rates (Figure 5.3), behaviour which was attributed to the lactose operon dynamics [10]. In contrast, at IPTG levels above this critical value, the specific production rates showed a sudden peak after induction to progressively decrease as the specific growth rate was being affected. This behaviour suggested that, under those conditions, the inducer was mainly transported by unspecific mechanisms or diffusion and the *lac* operon was fully induced in a short time. Since RhuA production is growth-related, the specific growth rate decrease was translated into lower process productivities.
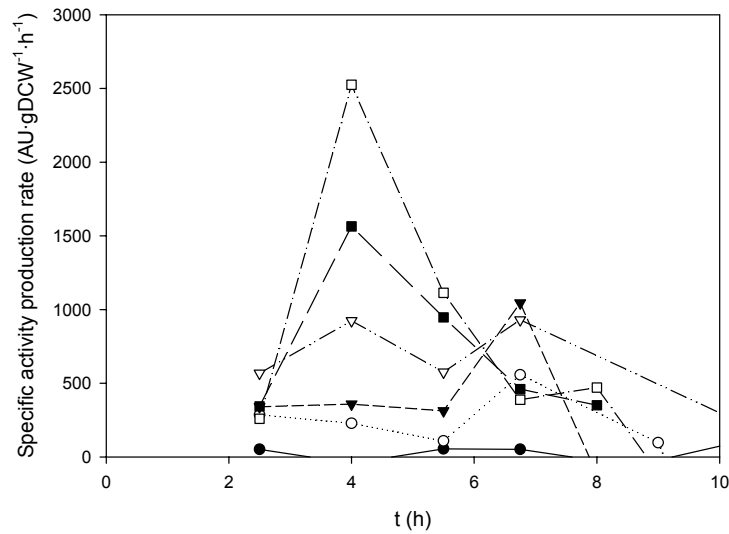
Figure 5.3. Specific activity production rate profiles in shake flask cultures induced at 0 (●), 5 (○), 10 (▼), 20 (▽), 30 (■) and 50 (□) μmol IPTG·L^{-1}.

5.3.2. Fed-batch cultures.

Once the conditions which maximised the activity per gram of dry cell weight had been identified, high cell density cultures were performed to increase the productivity of the process. However, non-induced cultivations were first run to ensure a proper glucose feeding strategy to reach high biomass concentrations. Factors like solubility of substrates in water and limitation or inhibition of cell growth by these can affect the fed-batch culture performance. Therefore, a well designed media and feeding strategy are needed to achieve high cell densities [11, 12]. Another motivation to develop a feeding strategy was to decrease acetate accumulation, since high acetate concentration can inhibit growth and recombinant protein production [13].

Several uninduced fed-batch cultures using an exponential feed were performed. In order to keep the specific growth rate to the desired set point, the substrate to biomass yield as well as the maintenance coefficient had to be set in the feeding model (Eq. 3.1). Determination of these parameters led to 0.35 gDCW·g^{-1} for biomass-substrate yield and 0.10 g·gDCW^{-1}·h^{-1} for maintenance coefficient. For phosphorus, a yield of 16.8 gDCW·g phosporus^{-1} was determined. Accumulation of acetate was detected in cultures performed at specific growth rates above 0.26 h^{-1}, and the specific accumulation rate increased with the specific growth rate (Figure 5.4).

Below this critical specific growth rate, values of OD600 above 200 were reached (Fig. 5.5), so undesired limitations or inhibitions from any other of the media or feed components than glucose could be excluded.
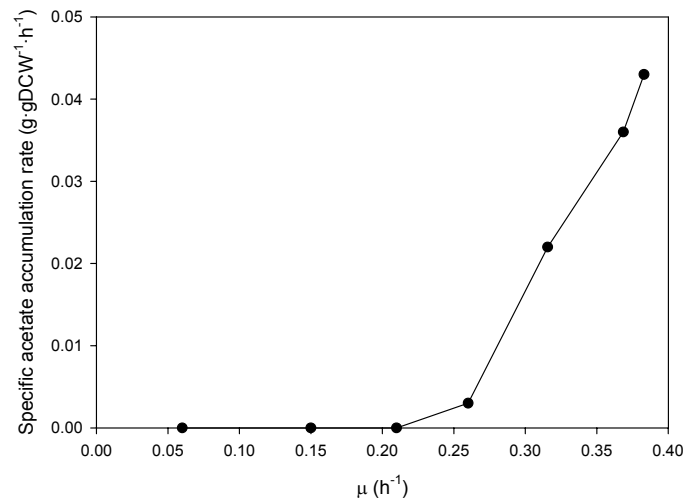
Figure 5.4. Specific acetate accumulation kinetics in substrate limited fed-batch cultures. Each point comes from a different culture performed at constant specific growth rate.
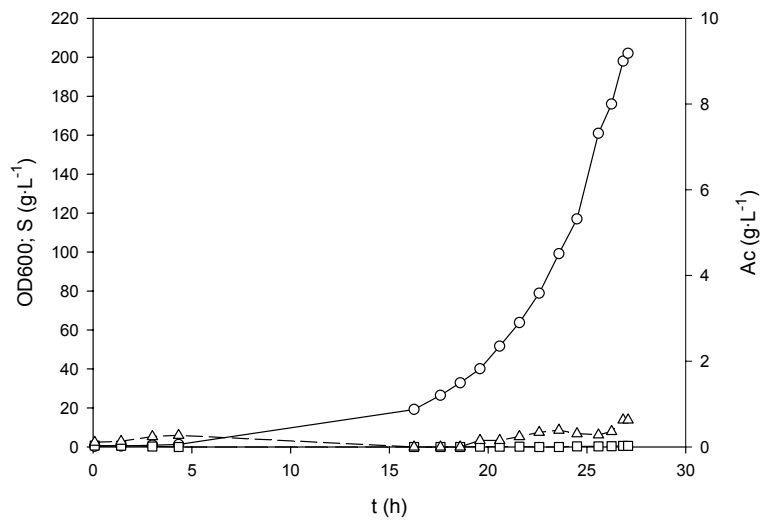
Figure 5.5. Biomass (OD600, —○—), substrate (S, −□−) and acetate (Ac, -△-) concentration profiles in a non-induced fed-batch culture performed at 0.26 h^{-1}.

5.3.2.1. Pulse induction.

The easiest approach to overexpress the recombinant protein was to feed limiting amounts of glucose exponentially (keeping the specific growth rate below the critical one to avoid acetate accumulation) until high cell densities were reached, and then add an IPTG pulse. Using this strategy, a series of fed-batch runs at μ= 0.25 h^{-1} were performed as explained in section *3.4. Cultivation conditions*. Induction was performed at concentrations ranging from 12 to 150 μM, which corresponded to specific IPTG concentrations ranging from 0.7 to 4.9 μmol·gDCW^{-1}.

Biomass, glucose, product and acetate were followed along time, so that specific activities and specific activity production rates could be calculated. Their maximum values were plotted versus inducer concentration and versus the IPTG ratio to biomass for each experiment (Figure 5.6). In contrast to the results obtained from preliminary expression studies, cell growth ceased after induction in all cases, evidencing that pulse induction clearly overloaded the host cell metabolism. The length of the production phase, however, depended on the amount of IPTG pulsed (data not shown). In the cases where more inducer was added, the specific growth rate decay was so fast that the maximum specific production rates were even lower than in those cases where lower amounts of inducer were added.
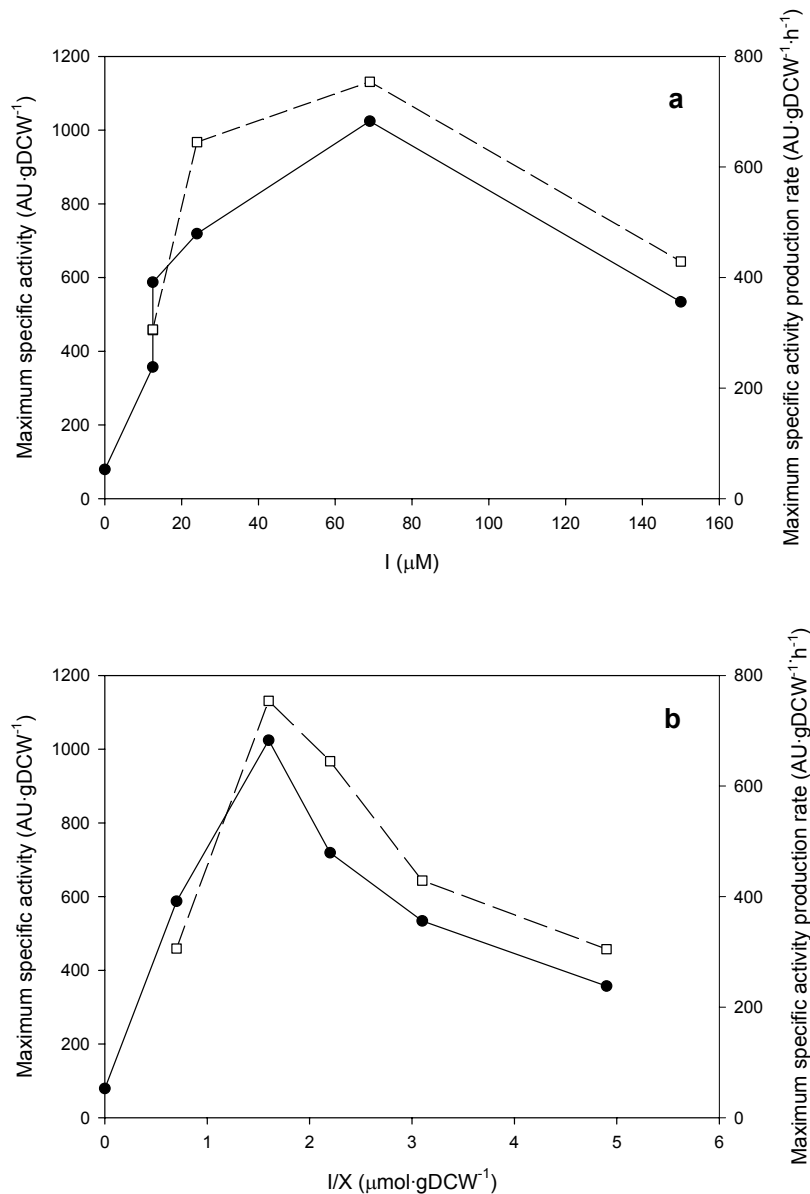

Figure 5.6. Maximum specific activities (—●—) and specific activity production rates (-□-) in glucose-limited fed-batch cultures versus IPTG concentration (a) and versus IPTG to biomass ratio (b). Each point comes from a different pulse-induced fed-batch experiment at $\mu=0.25$ h^{-1}.

Interestingly, similar IPTG concentrations (12-13 μM) did not lead to similar specific activities when induced at different biomass levels (Figure 5.6a), suggesting that volumetric IPTG concentration was a key parameter over inducer transport rate and over the specific production rate, but not over the protein content within the cells. Table 5.1 shows the observed differences between these two cultures. The fact that one of them had a ratio of IPTG to biomass of 0.7 μmol IPTG·gDCW^{-1} and the other one of 4.9 μmol IPTG·gDCW^{-1} made that the first one could grow for 3 h and the second one only for 2 h after induction, due to overload. This longer induced phase growth could explain why the first one reached a maximum specific activity of 587 AU·gDCW^{-1} while the second one only reached 357 AU·gDCW^{-1}.

I (μM)	I/X (μmol·gDCW^{-1})	Maximum specific activity production rate (AU·gDCW^{-1}·h^{-1})	Maximum specific activity (AU·gDCW^{-1})
12	0.7	306	587
13	4.9	305	357

Table 5.1. Comparison between cultures induced at different cell densities and similar IPTG concentrations.

Figure 5.7 depicts the evolution of the culture which led to the highest amount of the target enzyme. Induction with an IPTG pulse of 70 μM or 1.6 μmol·gDCW^{-1} showed the highest specific activity (1023 AU·gDCW^{-1}) as well as the highest specific production rate peak (754 AU·gDCW^{-1}·h^{-1}). Below and above this optimal IPTG levels, the maximum specific activities as well as the maximum specific production rates were lower depending on the pulsed amount of inducer.

When comparing the performance of preliminary shake flask experiments and fed-batch cultures after pulse induction, several differences were found. First, cell growth was affected in all fed-batch cultivations conducted under substrate limitation before maximum recombinant protein accumulation had taken place. As a consequence, the maximum specific activities were lower. Second, the maximum specific production rates were also lower and showed an optimum, while in the preliminary shake flask experiments they increased according to the pulsed amount of IPTG.

A possible explanation of these differences can be drawn from IPTG transport considerations. Under substrate limiting conditions, carbon catabolite repression and inducer exclusion, two well known phenomena which affect the *lac* operon when glucose is present [14-16], can be excluded. Thus, active IPTG transport should take place. In addition, higher IPTG concentrations increased the driving force for inducer diffusion with respect to the case of the

shake flasks. Both mechanisms probably led to high intracellular inducer levels very quickly, affecting cell growth. Our results suggested that probably only a small fraction of the IPTG was uptaken by cells in the preliminary shake flask studies, since the optimal ratio of IPTG to biomass required to induce the cells was two orders of magnitude higher than in glucose limited cultures (153 μmol IPTG·gDCW^{-1} and 1.6 μmol IPTG·gDCW^{-1}, respectively).
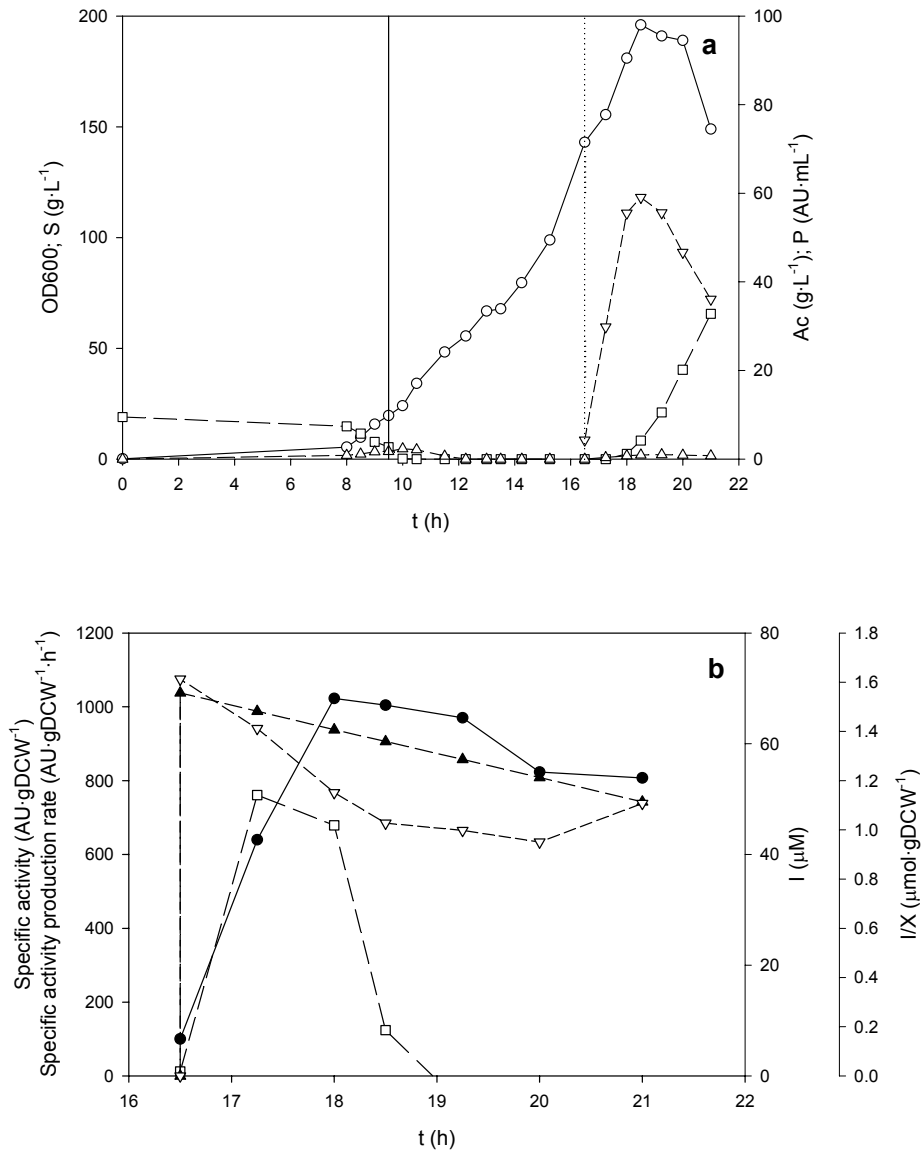
Figure 5.7. (a) Biomass (OD600,—○—), glucose (S, - -□ -), acetate (Ac, -△ - -) and activity profiles (P, - ▽ -) in a glucose limited fed-batch culture. The start of the feeding is represented by a continuous vertical line, while the vertical dotted line represents a 70 μM IPTG pulse addition. (b) Specific activity (—●—), specific activity production rate (- -□- -), inducer concentration (I, - ▲ -) and specific inducer concentration (I/X, - ▽ -) profiles.

In the case of shake flask experiments, the critical specific production rate above which cell growth was affected was close to 1000 AU·gDCW^{-1}·h^{-1}, while in the case of substrate-limited fed-batch cultures, even specific production rates of 300 AU·gDCW^{-1}·h^{-1} were enough to stop

biomass build-up. In shake flask cultures cells grew at high specific growth rates with an excess of substrate, while in fed-batch cultures the specific growth rate was maintained below a critical value by limiting substrate availability, a factor which is known to increase the stress of cells [17-19]. It has been reported that stringent regulation causes altered growth behaviour, and a specific growth rate drop can be observed due to segregation of cells to a non-dividing state [20]. In our case, it looked that starvation decreased the stress levels produced by recombinant protein overexpression that cells could stand. The maximum specific activities obtained in fed-batch cultures were also lower than in the case of shake flask cultures, probably due to shorter induction phase lengths. In fed-batch cultures, the maximum length of the induced phase was 1 doubling time (2.5 h), while in induced shake flask cultures growth was sustained for 2 generations in the worst case (4 h).

5.3.2.2.   Continuous induction.

To extend the production phase and to try to improve yields, adjustment of recombinant gene expression by modulation of the transcription rate can be applied [6]. Therefore, a continuous inducer dosage to tune the productivity of this strong expression system below the observed critical values was implemented in substrate-limited fed-batch cultures. From the results obtained in pulse-induced cultures run under substrate limitation, it was concluded that the maximum specific activity production rate had to be controlled below 300 AU·gDCW^{-1}·h^{-1}.

Preliminary experiments were performed using different inducer feed strategies aiming to control either the specific IPTG concentration or the volumetric IPTG concentration at constant values. Results from these cultures showed that an inducer concentration around 4 μM triggered the accumulation of RhuA in the cells (data not shown). On the other hand, substrate started to accumulate (evidencing that cell metabolism was affected) when the inducer concentration was higher than 12 μM. These results are consistent with those of others who determined the threshold inducer concentration to maintain induction at 3-5 μM and 20 μM to fully induce the *lac* operon [21, 22].

Fed-batch cultures run at constant I/X set points ranging from 0.2 to 0.6 μmol IPTG·gDCW^{-1} were able to keep the maximum specific activity production rate at very low values in all cases (data not shown). Continuous IPTG addition minimised metabolic overload of host cells and the induction phase was prolonged as pretended.

Based on these results, the inducer feeding was predefined to progressively increase IPTG concentration within previously observed working range (from 4 to 15 μM) together with

specific inducer levels from 0.3 to 0.5 μmol·gDCW^{-1}. The time profiles of the experiment, performed at μ=0.13 h^{-1}, are shown in Figure 5.8. Cell growth stopped at an IPTG concentration of 12 μM, 1.5 generations after inducer feed had been switched on (12 h), and most of the recombinant protein production took place at specific inducer values between 0.3 and 0.4 μmol·gDCW^{-1}. In this case, the maximum specific activity was close to 1000 AU·gDCW^{-1}, similar to the maximum specific activity achieved with a 70 μM IPTG pulse but reducing nearly 6-fold the amount of inducer added. The inducer feeding policy proposed allowed the control of the specific production rate below the previously observed critical values, extending the induction phase and obtaining good RhuA specific yields.
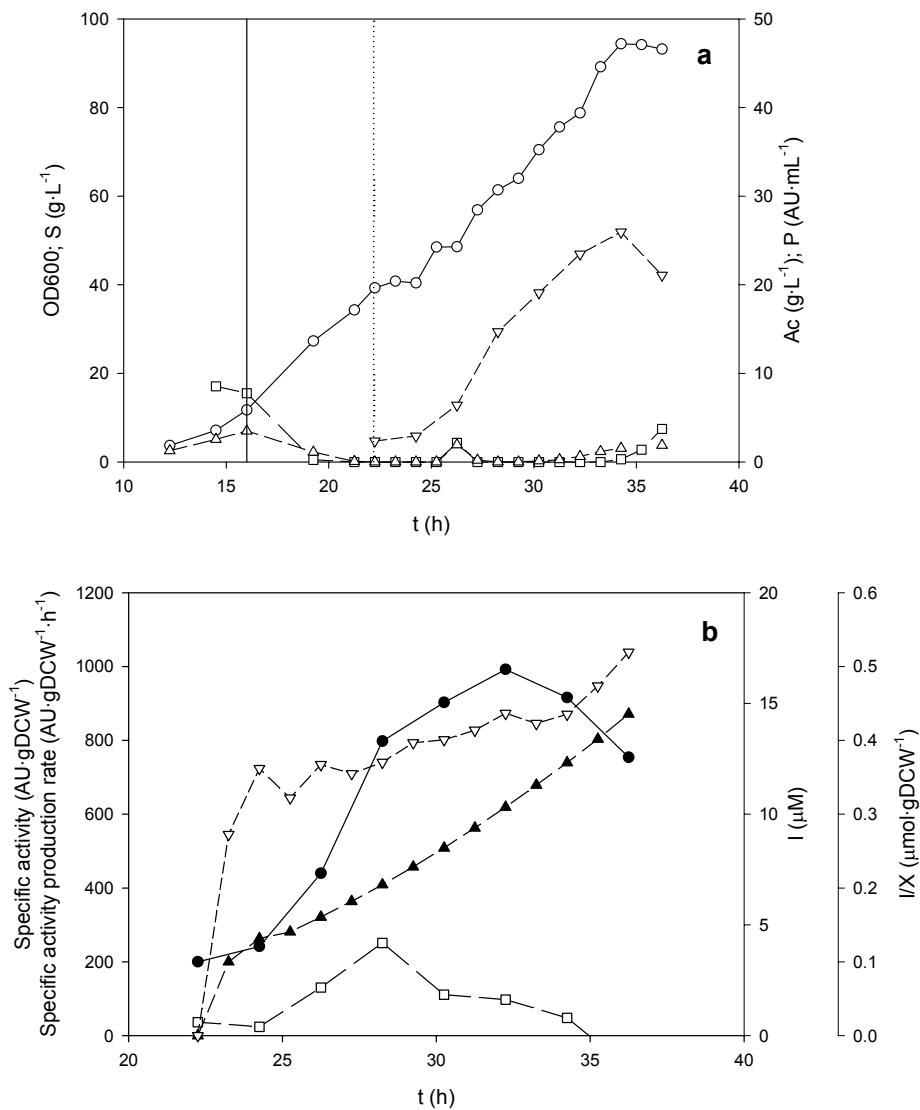
Figure 5.8. (a) Biomass (OD600,—○—), glucose (S, -□- -), acetate (Ac, - -△- -) and activity profiles (P, - -▽- -) in a glucose limited fed-batch culture. The vertical continuous line represents when glucose feed started, while the vertical dotted line represents the start of inducer feeding. (b) Specific activity (—●—), specific activity production rate (- -□- -), inducer concentration (I, - -▲- -) and specific inducer concentration profiles (I/X, - -▽- -).

5.4. Conclusions.

The induction strategy had a strong influence on recombinant rhamnulose 1-phosphate aldolase production process in *E. coli* M15 ΔglyA [pREP4] pQE$\alpha\beta$rham. Since protein production was growth-related, exploitation of host cell metabolism was strongly dependent on the induction policies.

The classical pulse addition of inducer clearly overloaded the host cell capacity affecting cell growth when transcription was under the control of the strong T5 promoter. Optimisation of the IPTG pulse in substrate-limited cultures led to maximum specific activities 1/3 lower than the maximum ones previously measured in shake flask cultures.

An alternative continuous inducer dosage strategy was developed in order to minimise the metabolic burden imposed in host cells by full induction of RhuA overexpression. This strategy allowed to control the maximum activity production rates below critical values, so that the induction phase could be successfully extended. It led to maximum specific activities similar to those obtained with the optimal IPTG pulse, but the required amount of inducer was significantly reduced.

The inducer concentration could be identified as a key parameter affecting inducer transport, which in turn affected the specific activity production rate. This variable was directly related to the metabolic burden and, thus, the extension of the induction phase depended on it. A critical maximum specific activity production rate was identified in glucose in excess cultivations (1000 AU·gDCW^{-1}·h^{-1}) and this critical value was lower when cells were induced under glucose limitation (300 AU·gDCW^{-1}·h^{-1}), probably due to starvation-induced stress.

A range of inducer concentrations approximately between 4 and 12 μM was identified as the working range within cell growth and recombinant protein accumulation occurred simultaneously. Higher IPTG concentrations triggered cell segregation into a viable but not culturable state (see results from chapter 6) although the specific activity production rate had been controlled at low values and maximum specific RhuA activity was below the maximum which cells could stand. On the other hand, the inducer to biomass ratio was also identified as a key parameter affecting the final specific activities. Therefore, the optimal induction strategy should take both variables into account.

Inducer transport behaved differently in substrate-limited cultures than in batch cultures. The amount of IPTG per gram of biomass required to trigger expression was two orders of magnitude lower in substrate-limited fed-batch cultivations (around 1 μmol IPTG·gDCW^{-1}) when

compared to non-limited experiments (100-200 μmol IPTG·gDCW^{-1}). This represents an important cost saving not only because of IPTG cost but also because of IPTG removal costs as well since, depending on the product application, it has to be completely eliminated due to its potential toxicity.

Even though the productivity of the process was increased in substrate-limited fed-batch cultures because of the higher final biomass concentrations, the maximum specific activities achieved were approximately 66% of the maximum ones obtained in batch cultivations. Therefore, further investigations are required to find out the reasons for this undesired behaviour.

References.

[1] E.García-Junceda, G.J.Shen, T.Sugai, C.H.Wong, A New Strategy for the Cloning, Overexpression and One-Step Purification of 3 DHAP-Dependent Aldolases - Rhamnulose-1-Phosphate Aldolase, Fuculose-1-Phosphate Aldolase and Tagatose-1,6-Diphosphate Aldolase, Bioorganic & Medicinal Chemistry 3 (1995) 945-953.

[2] L.Vidal, et al., High-level production of recombinant His-tagged rhamnulose 1-phosphate aldolase in *Escherichia coli*, Journal of Chemical Technology and Biotechnology 78 (2003) 1171-1179.

[3] L.Vidal, P.Ferrer, G.Alvaro, M.D.Benaiges, G.Caminal, Influence of induction and operation mode on recombinant rhamnulose 1-phosphate aldolase production by *Escherichia coli* using the T5 promoter, Journal of Biotechnology 118 (2005) 75-87.

[4] L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

[5] W.E.Bentley, N.Mirjalili, D.C.Andersen, R.H.Davis, D.S.Kompala, Plasmid-Encoded Protein - the Principal Factor in the Metabolic Burden Associated with Recombinant Bacteria, Biotechnology and Bioengineering 35 (1990) 668-681.

[6] G.Striedner, M.Cserjan-Puschmann, F.Potschacher, K.Bayer, Tuning the transcription rate of recombinant protein in strong *Escherichia coli* expression systems through repressor titration, Biotechnology Progress 19 (2003) 1427-1432.

[7] L.H.Hansen, S.Knudsen, S.J.Sorensen, The effect of the *lacY* gene on the induction of IPTG inducible promoters, studied in *Escherichia coli* and *Pseudomonas fluorescens*, Current Microbiology 36 (1998) 341-347.

[8] P.R.Jensen, H.V.Westerhoff, O.Michelsen, The Use of *Lac*-Type Promoters in Control Analysis, European Journal of Biochemistry 211 (1993) 181-191.

[9] J.Beckwith, The lactose operon, in: F.C.Neidhart (Ed.), *Escherichia coli* and *Salmonella typhimurium*, American Society of Microbiology, Washington, D.C. (1987) 1444-1452.

[10] N.Yildirim, M.C.Mackey, Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data, Biophysical Journal 84 (2003) 2841-2851.

[11] J.Shiloach, R.Fass, Growing *E. coli* to high cell density--a historical perspective on method development, Biotechnology Advances 23 (2005) 345-357.

[12] D.Riesenberg, High-Cell-Density Cultivation of *Escherichia coli*, Current Opinion in Biotechnology 2 (1991) 380-384.

[13] R.A.Majewski, M.M.Domach, Simple Constrained-Optimization View of Acetate Overflow in *Escherichia coli*, Biotechnology and Bioengineering 35 (1990) 732-738.

[14]   M.H.Saier, Protein-Phosphorylation and Allosteric Control of Inducer Exclusion and Catabolite Repression by the Bacterial Phosphoenolpyruvate - Sugar Phosphotransferase System, Microbiological Reviews 53 (1989) 109-120.

[15]   M.Santillan, M.C.Mackey, Influence of catabolite repression and inducer exclusion on the bistable behavior of the *lac* operon, Biophysical Journal 86 (2004) 1282-1292.

[16]   P.Wong, S.Gladney, J.D.Keasling, Mathematical model of the *lac* operon: Inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose, Biotechnology Progress 13 (1997) 132-143.

[17]   D.Chatterji, A.K.Ojha, Revisiting the stringent response, ppGpp and starvation signaling, Current Opinion in Microbiology 4 (2001) 160-165.

[18]   H.J.Chung, W.Bang, M.A.Drake, Stress response of *Escherichia coli*, Comprehensive Reviews in Food Science and Food Safety 5 (2006) 52-64.

[19]   J.Ihssen, T.Egli, Specific growth rate and not cell density controls the general stress response in *Escherichia coli*, Microbiology-Sgm 150 (2004) 1637-1648.

[20]   L.Andersson, L.Strandberg, S.O.Enfors, Cell segregation and lysis have profound effects on the growth of *Escherichia coli* in high cell density fed batch cultures, Biotechnology Progress 12 (1996) 190-195.

[21]   E.M.Ozbudak, M.Thattai, H.N.Lim, B.I.Shraiman, A.van Oudenaarden, Multistability in the lactose utilization network of *Escherichia coli*, Nature 427 (2004) 737-740.

[22]   J.M.G.Vilar, C.C.Guet, S.Leibler, Modeling network dynamics: the *lac* operon, a case study, Journal of Cell Biology 161 (2003) 471-476.

6. INFLUENCE OF PROCESS TEMPERATURE ON RECOMBINANT RHAMNULOSE 1-PHOSPHATE ALDOLASE QUALITY IN *Escherichia coli* FED-BATCH CULTURES.

Abstract.

The influence of process temperature on recombinant rhamnulose 1-phosphate aldolase (RhuA) quality has been studied. Progressive induction by means of continuous isopropyl-β-D-thiogalactopyranoside (IPTG) dosage in *E. coli* fed-batch cultures led to high specific levels of recombinant protein. However, the specific activity profile did not correlate to the specific protein contents when the process was run at 37ºC, evidencing a decrease of the enzyme activity along the induction phase. Specific activity loss depending on the presence of an energy source was observed at short term, but protein degradation due to the action of energy independent metalloproteases occurred after a longer time period. The effects of lowering the temperature were analysed on both mechanisms, and a reduction of the specific activity loss was observed when the process temperature was decreased to 28ºC. In high cell density fed-batch cultures performed under these conditions, lower plasmid copy number and specific production rates probably alleviated the metabolic load on host cell during recombinant protein overexpression, leading to a significant increase of the enzyme activity.

Universität für Bodenkultur Wien
University of Natural Resources and Applied Life Sciences, Vienna

6.1. Introduction and objectives.

As described in the previous chapter, recombinant RhuA overexpression yields in substrate-limited fed-batch cultures were strongly dependant on the induction strategy. Even though induction was optimised for pulse induction as well as for inducer dosage, the specific activity profiles (activity units per gram of dry cell weight) presented a maximum followed by a decrease of this variable in both cases (see Figure 5.7 and 5.8). Plasmid losses were discarded as the reason for this behaviour because of the use of *E. coli* M15 ΔglyA [pREP4] pQEαβrham, an expression system which ensures the maintenance of the plasmid in defined media [1] (as it will be shown further on in Figure 6.5).

In spite of the extensive knowledge on recombinant protein production in *E. coli*, not every gene can be expressed efficiently due to the unique structural features of the protein, its ease of folding and its degradation by host cell proteases [2]. Since recombinant proteins have to be active for their use, the objectives of the production process should not only include the maximisation of the amount of target protein, but also the maximisation of its biological activity. Some authors have attributed losses in enzyme activity to the partial intracellular denaturation of the protein driven by high temperature or by the absence of a critical factor (prosthetic group, subunit or chaperone) [3]. Others have attributed this fact to increased degradation of the protein in cells under starvation conditions [4], as is the case of substrate-limited fed-batch cultures. Therefore, it was decided that both the amount of protein and its activity should be measured in our case to find out whether the decrease of the activity units per gram of dry cell weight had been caused by denaturation or by degradation of the target protein.

Proteolysis is one of the mechanisms used by cells to get rid of misfolded, incorrectly synthesised or unwanted proteins. Its role, mechanisms and targets have been studied [5, 6] and some strategies to control it have been suggested, but little is known about the influence of cultivation conditions, especially in fed-batch cultures. Several techniques like designing a more stable protein or modification of host cell have been successfully applied to control proteolysis [7], but their effects on the specific activity of the recombinant protein are not known beforehand.

The aim of this chapter was to avoid the specific RhuA activity losses previously observed in substrate-limited fed-batch cultures. Therefore, it was first focused on the understanding of this behaviour. Since temperature reduction has been reported to improve protein folding and activity as well as to and reduce protein degradation, the work was then focused on the study of the process temperature and its effects on protein quality in substrate-limited fed-batch cultures.

6.2. Materials and methods.

Bacterial strain and plasmids.

The K-12 derived strain *E. coli* M15ΔglyA [pREP4] harbouring the vector pQEαβrham was used for rhamnulose 1-phosphate aldolase overexpression (see section *3.1. Strains and vectors* for further details on the expression system). Frozen stock aliquots containing glycerol prepared from exponential phase cultures grown in Luria-Bertani media (LB) were stored at -80ºC.

Media composition.

LB medium was used for preinoculum preparation and prepared as detailed in *3.2. Media composition*.

A defined mineral medium, utilising glucose as the sole carbon source, was used for all experiments. It contained 3 g·L^{-1} KH$_2$PO$_4$ and 6 g·L^{-1} K$_2$HPO$_4$·3H$_2$O. The other components were added in relation of gram dry cell weight (gDCW): sodium citrate (trisodium salt ·2H$_2$O) 0.25 g, MgSO$_4$·7H$_2$O 0.10 g, CaCl$_2$·2H$_2$O 0.02 g, trace element solution 50 μL and glucose·H$_2$O 3 g. The batch medium was calculated to produce 22.5 gDCW and was supplemented with 0.15 g yeast extract per gDCW to accelerate initial growth of the cell population. For the feeding phase 8 L of minimal medium were prepared according to the amount of biological dry matter (365 gDCW) to be produced in the feeding phase, whereby phosphate salts were again added per litre. Trace element solution prepared in 5 N HCl (g·L^{-1}): FeSO$_4$·7H$_2$O 40.0, MnSO$_4$·H$_2$O 10.0, AlCl$_3$·6H$_2$O 10.0, CoCl$_2$ 4.0, ZnSO$_4$·7H$_2$O 2.0, Na$_2$MoO$_2$·2H$_2$O 2.0, CuCl$_2$·2H$_2$O 1.0, H$_3$BO$_3$ 0.50. Moreover, thiamine was added to 0.003 g·gDCW^{-1} and a surplus of 0.339 gZnSO$_4$·7H$_2$O·gDCW^{-1} was added to ensure the 0.1 mgZn·gDCW^{-1} necessary to have high levels of RhuA fully active [8]. IPTG was supplied in a separate feed (630 μM).

Cultivation conditions.

The cells were grown in a 20 L (12 L net volume, 4 L batch volume) computer controlled bioreactor (MBR; Wetzikon, CH) equipped with standard control units. The pH was maintained at a set-point of 7.0 ± 0.05 by addition of 25 % (w/v) ammonia solution, the temperature was set to 37 ± 0.5 °C or 28 ± 0.5 °C depending on the case. In order to avoid oxygen limitation the dissolved oxygen level was stabilised above 30 % saturation by stirrer speed and aeration rate control. Foaming was suppressed by addition of antifoam suspension (Glanapon 2000, Bussetti, Vienna) with a concentration of 1 mL·kg feed medium^{-1}. For inoculation, a deep frozen (-80 °C)

working cell bank vial was thawed and 1 mL (OD600= 1) was transferred aseptically to the bioreactor. Feeding was started when the culture, grown to a bacterial dry matter of 30 g in 4 L batch medium, entered stationary phase. Fed-batch regime with an exponential substrate feed was used to provide a constant growth rate of 0.1 h^{-1} during 4 doubling times (28 h). The substrate feed was controlled by increasing the pump speed according to the exponential growth algorithm $X=X_o.e^{\mu t}$, with superimposed feedback control of weight loss in the substrate tank [9]. The feed medium provided sufficient components to yield another 365 g of bacterial dry matter.

Analytical methods.

Growth was followed by optical density measurements at 600 nm (OD600). The dry cell weight (DCW) was measured by centrifugation of 10 mL of cell suspension, re-suspension in distilled water followed by centrifugation, and re-suspension for transfer to a pre-weighted beaker, which was then dried at 105 °C for 24 h and re-weighted.

Total Cell Number (TCN) and percentage of dead cells (DC) were determined using flow cytometric methods. All measurements were performed with a FACSCalibur flow cytometer (four-color system; Becton Dickinson), equipped with an air-cooled laser providing 15 mW at 488 nm and the standard filter setup. For acquisition and analysis of data, the CellQuest software package (version 4.1, Becton Dickinson) was used. Forward light scatter (FSC), right-angle light scatter (side scatter; SSC), and fluorescence were collected as pulse height signals (four decades of a logarithmic scale). Side light scatter (SSC) served as the primary detection parameter (threshold value 125). In addition, software discriminators were set on side scatter light signal to further reduce electronic and small particle noise. Discrimination of dead cells was done via staining with propidium iodide [10], while total cell number was determined via ratiometric counting. Samples were spiked with a known amount of fluorescent counting beads (Becton Dickinson, USA) so that the absolute cell number could be back-calculated.

Plasmid copy number (PCN) was calculated from plasmid and chromosomal DNA. Plasmid DNA was isolated from the bacterial cells using a commercial miniprep kit (plasMix) and quantified with the Agilent Bioanalyser DNA 7500 LabChip® Kit. Chromosomal DNA was determined by a fluorescence assay with HOECHST dye H33258 after cell disintegration with lysozyme and SDS [11].

Plasmid-containing cells were determined by cultivation of bacterial cells on LB-agar plates and on plates containing 100 mg·mL^{-1} ampicillin and by counting colony forming units (CFU) after incubation at 37ºC for 24 h. To confirm overexpression of the recombinant protein in plasmid-

carrying cells, CFUs were also determined after recombinant gene induction by IPTG (200 mg·mL^{-1}) on ampicillin-containing plates. Since overexpression of recombinant protein is lethal to the cell, the failure to form colonies on agar plates containing high amounts of IPTG was the criterion for the determination of cells producing high amounts of recombinant protein.

The two different methods used to quantify the product in each sample are described in section *3.5. Analytical methods*. The first method was used to quantify the amount of protein, while the second one was used to quantify its activity. One unit of RhuA activity was defined as the amount of enzyme required to convert 1 μmol of rhamnulose 1-phosphate in DHAP per minute at 25ºC under the assay conditions.

Proteolysis tests.

In order to find out whether RhuA degradation existed, *in vitro* tests were performed. Three aliquots from samples withdrawn from the bioreactor were centrifuged and frozen for further processing. After thawing, the pellet was resuspended in 30 mM Tris·HCl (pH=8.2) and incubated under gentle mixing at 37, 33 or 28ºC depending on the case. The first aliquot was used to monitor OD600. 2 mM ATP and 2 mM MgCl$_2$ were added to the second one to activate proteases, since many proteases require ATP and/or Mg^{2+} for activity [12]. 2 mM ATP, 2 mM MgCl$_2$ and 50 mM EDTA were added to the third aliquot to check whether metalloproteases activity was significant. OD600 was measured periodically from the first aliquot, and samples from the rest were centrifuged to freeze the pellet for further analysis. SDS-PAGE gels were run from these samples as described in section *3.5. Analytical methods* to check for RhuA degradation.

In order to determine RhuA activity loss along the induction phase, *in vivo* tests were carried out. Samples taken every generation after inducer feed had started were withdrawn from the bioreactor and used to inoculate two 500 mL shake flasks containing 100 mL of defined media to OD600=3, but only one of them contained glucose (10 g·L^{-1}). This was done to check whether the process affecting RhuA activity was energy dependant. Chloramphenicol (100 mg·L^{-1}) was added to both flasks to stop protein synthesis and therefore, quantify only RhuA degradation. The shake flasks were incubated at 28ºC or 37 ºC and 200 rpm. Samples were withdrawn periodically during incubation to measure the optical density at 600 nm, the specific activity and total intracellular RhuA (as described in section *3.5. Analytical methods*).

6.3. Results and discussion.

The expression of RhuA was performed at 37°C in a glucose limited fed-batch process. Once the glucose from the batch phase had been depleted, an exponential substrate feed was used to provide a constant growth rate of 0.1 h^{-1} during 4 doubling times (28 h). According to the results obtained in chapter 5, an inducer feed was switched on one generation after the substrate feed had been started to linearly increase the specific inducer concentration from 0.3 to 0.6 μmol IPTG·gDCW^{-1} during the next 3 doublings [13]. Although the continuous inducer feed kept the specific production rate at low values to avoid the metabolic burden effect (data not shown), cell growth stopped during the last doubling time (Figure 6.1).
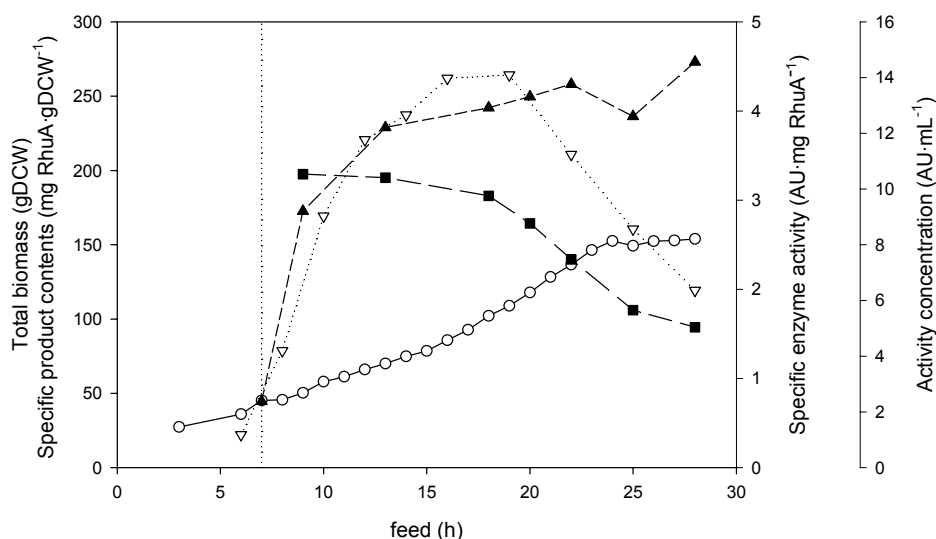

Figure 6.1. Total biomass (−○−), specific product contents (−▲−), specific enzyme activity (−■−) and activity concentration (·▽··) profiles along time in a substrate-limited fed-batch culture performed at 37ºC. The vertical dotted line indicates start of induction.

Flow cytometry analysis showed that the percentage of non-viable cells was below 1% during the whole cultivation, suggesting that the whole population entered a viable but not culturable state [14]. Size changes were detected when cells stopped growing, as the forward scatter signal showed a sharp increase (26%) at that point (results presented in chapter 7, see Figure 7.15). The reason for this behaviour could have been the high contents of recombinant RhuA accumulated in the cytoplasm (36% of the total soluble protein, corresponding to 258 mg RhuA·gDCW^{-1}). Expression of serine hydroxymethyl transferase (results presented in chapter 7, see Figures 7.11b and 7.12) triggered additional load independent from the induction strategy and could therefore contribute to the break-down of the system. Most of the recombinant protein, 235 mg RhuA·gDCW^{-1} accumulated during the first doubling after induction and the

specific activity of the recombinant enzyme was close to 3 AU·mg RhuA^{-1}. From that moment on, the protein contents within the cells remained nearly constant, but significant activity loss of the recombinant protein was observed although cells continued to grow. This behaviour was translated into a sharp decrease of the volumetric activity concentration. It became evident that the specific activity of the recombinant protein decreased along the induction phase, and since 6.2 AU·mg RhuA^{-1} had been reported as the maximum specific activity of pure RhuA [15], the reasons for the decrease of the specific activity had to be investigated.

As possible reduction of aldolase activity due to the lack of RhuA prosthetic group (Zn) had been previously avoided by keeping 0.1 mgZn·gDCW^{-1} in the medium [8], partial proteolysis was considered as a possible explanation to the observed activity losses. In consequence, *in vitro* tests were performed to check RhuA degradation. Different aliquots from samples withdrawn at feed=19 h (cells in a still growing state) and feed=28 h (cells in a viable but not culturable state) were used to compare recombinant protein degradation at different stages of the culture. Cells were incubated at 37ºC, and samples were taken periodically to measure OD600 as well as intracellular protein content. Densitometries of the SDS-PAGE were used to calculate the percentage of intracellular RhuA along incubation time. In the case where cells were incubated in the presence of EDTA, steady specific contents of RhuA were observed. However, the same samples incubated under the same conditions but in the absence of EDTA showed significant RhuA degradation, although it took around 20 h to trigger proteolytic activity. Figure 6.2 shows the percentage of normalised specific RhuA levels after 24 h of incubation with respect to the initial ones.
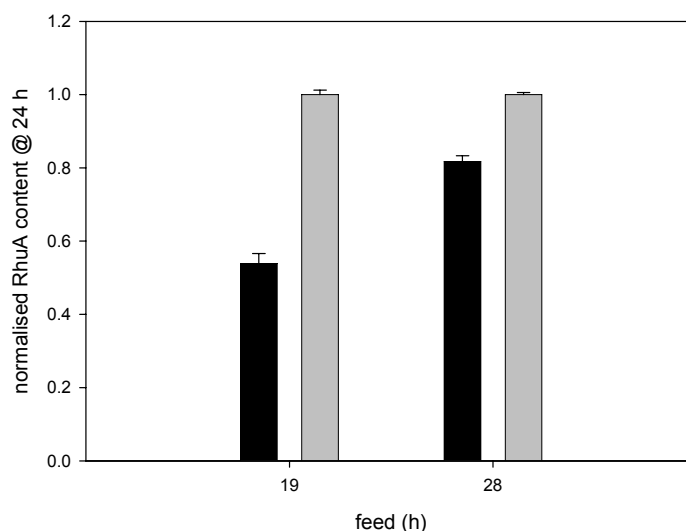

Figure 6.2. Normalised remaining intracellular RhuA after 24 h of incubation in samples withdrawn at different stages of the culture. Black bars depict results obtained in incubation buffer (▬▬▬), while grey bars symbolise results obtained in incubation buffer supplemented with EDTA (▭▭▭).

EDTA efficiently prevented degradation, suggesting that metalloproteases were responsible for RhuA disappearance. Moreover, depending on the stage of the culture significant differences were found in the amount of remaining RhuA. Noteworthy, higher proteolysis was detected in samples withdrawn from the growing culture (feed=19 h, 46% degradation) compared to non dividing cells from the end of the process (feed=28 h, 18% degradation).

The mechanism of proteolysis in *E. coli* usually consists of a first step comprising partial cleavage of the protein by energy-dependant proteases followed by a second step where partially cleaved proteins are degraded into amino acids by energy independent proteases [6]. Our results suggested that partial proteolysis could be responsible for the reduced specific activity yields during cultivation.

Several approaches to decrease proteolysis have been proposed in the literature. They include protein sequence modification [16], inclusion body formation [17], the use of host strains deficient in proteases [18] or in the stringent response [19], coexpression of protease inhibitors [20] or secretion to the periplasm [21]. However, the effect of these solutions on enzymatic activity is not known beforehand and many of the protease deficient strains lack the required stability for commercial recombinant protein production [7, 12]. Therefore, we focused on the control of proteolysis on cultivation level. It has been reported that target recognition by proteases depends on the secondary and tertiary structure of proteins and that protein folding is improved at lower temperatures [22]. Moreover, protein degradation rate will decrease with decreased temperature in accordance to the Arrhenius' Law, and protein misfolding often occurs at high temperatures due to either heat damage or due to an increase of protein synthesis rate [7]. For these reasons, the most straightforward as well as promising strategy to improve RhuA quality was to optimise the production process temperature.

To verify the influence of temperature on RhuA degradation an *in vitro* test was carried out using sample aliquots taken from the 37°C experiment at feed=19 h (Figure 6.1) whereby incubation temperature was set to 28, 33 or 37ºC. Again, EDTA prevented recombinant protein degradation. Results also showed that, at least for this metalloprotease which degraded the protein after long term incubation, activity was higher at 33 ºC (83% @24h) than at 37 ºC (46% @24h) and at 28 ºC (31% @24h) RhuA degradation was reduced compared to higher temperatures (Figure 6.3).

With the aim to improve the recombinant protein specific activity, another culture was performed and induced at 28ºC using the same operational strategy as in the previous experiment. In this case (Figure 6.4), cell growth was not affected and specific RhuA levels increased along the whole induction phase up to 203 mg RhuA·gDCW^{-1} at the end of the

culture, attaining a specific activity of the recombinant enzyme of 4.2 AU·mg RhuA^{-1}. Although the specific protein content (mg RhuA·gDCW^{-1}) was lower than in the process run at 37ºC, a higher yield of active RhuA was obtained due to the higher specific protein activity and the higher biomass yield.
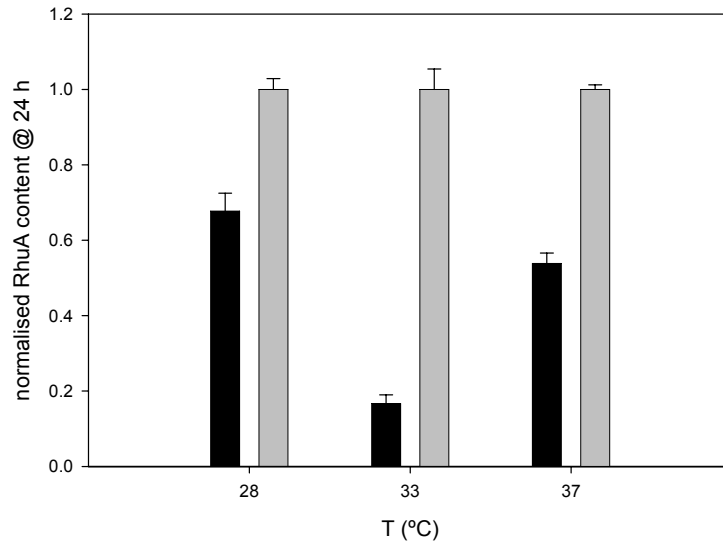
Figure 6.3. Normalised remaining intracellular RhuA after 24 h of incubation at 28, 33 or 37ºC. Black bars depict results obtained in incubation buffer (▬▬▬), while grey bars symbolise results obtained in incubation buffer supplemented with EDTA (▭▭▭).

Figure 6.4. Total biomass (−○−), specific product contents (−▲−), specific enzyme activity (−■−) and activity concentration (··∇··) profiles along time in a substrate-limited fed-batch culture performed at 28ºC. The vertical dotted line indicates start of induction.

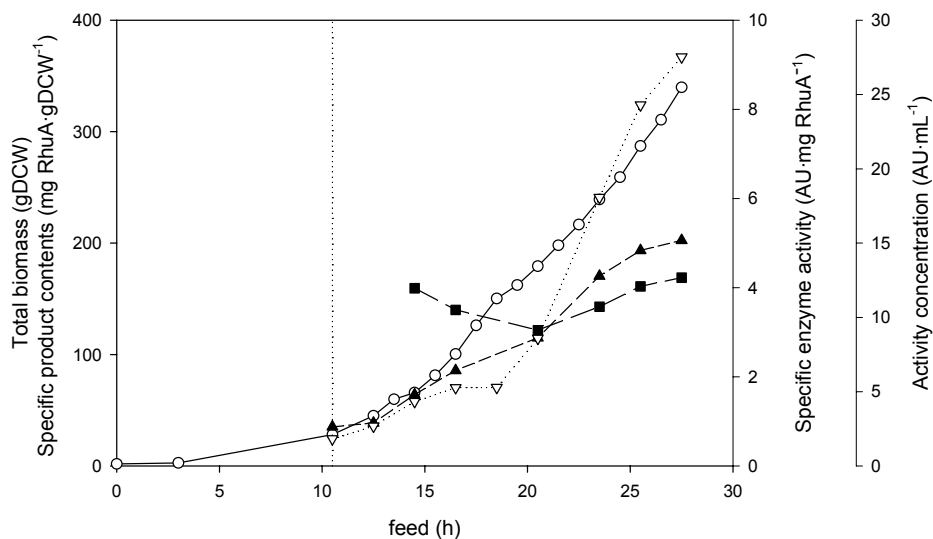The experiments performed at 37 and 28°C are compared in Table 6.1. At a cultivation temperature of 28 °C, a 60% decrease of the maximum specific production rate and a 26% decrease of the specific protein contents were observed. The effects of cultivation temperature on PCN (Figure 6.5) could be responsible for lower production rates at 28° C. The attenuation of the production rate presumably helped to balance the synthesizing machinery resulting in higher specific activities and higher activity concentration. These results agree with the findings that the cellular protein folding capacity represents the bottleneck at high synthesis rates [7, 23] and proteolytic activity is strongly related to the level of misfolded protein [24, 25].
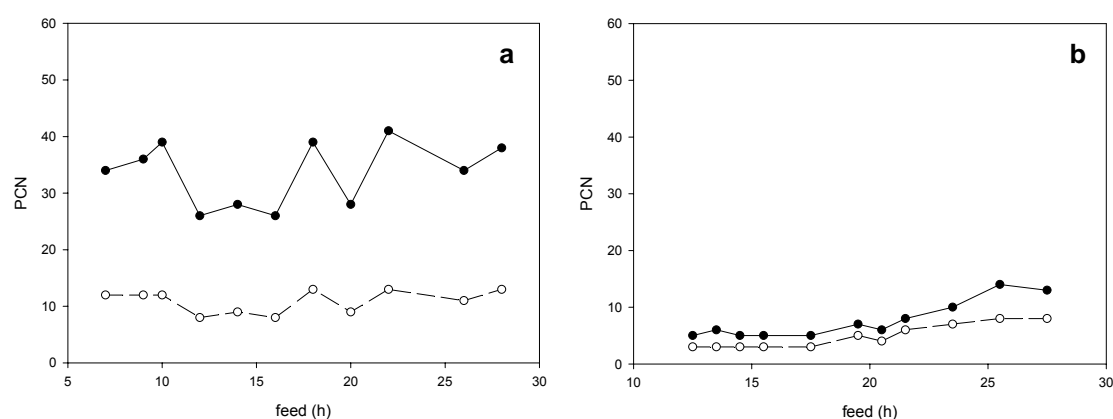

Figure 6.5. Plasmid copy number profiles for the fed-batch culture performed at 37ºC (a) and for the one performed at 28ºC (b). pQEαβrham (−●−) represents the number of production plasmid copies, while pREP4 (−○−) represents the helper plasmid copies (see section *3.1 Strains and vectors* for further details).

	37 ºC	28 ºC
Maximum specific production rate (mg RhuA·gDCW^{-1}·h^{-1})	82	33
Maximum specific product contents (mg RhuA·gDCW^{-1})	275	203
Maximum specific enzyme activity (AU·mg RhuA^{-1})	3.3	4.2
Maximum specific activity (AU·gDCW^{-1})	908	853
Maximum volumetric activity (AU·mL^{-1})	14	27.5
pQEαβrham production plasmid (PCN·cell^{-1})	~ 30	5 - 12
pREP4 helper plasmid (PCN·cell^{-1})	~ 12	4 - 8

Table 6.1. Comparison of the maximum averaged values for specific production rates, specific RhuA contents, enzyme activities, activities, volumetric activity concentrations and plasmid copy number (PCN) in cultures performed at 37 and 28 ºC.

Process operation at a temperature of 28°C was very successful and the final product yield could be improved, but there were still open questions concerning the problem of decreasing RhuA activities along the induction phase. As the *in vitro* test for proteolytic activity was not

suited to answer this question the two fed-batch experiments were conducted a second time to check the specific activity loss *in vivo*. Both processes were successfully reproduced (data not shown) and samples for *in vivo* proteolysis tests were withdrawn 1 and 2 generations after the inducer feed had been switched on, respectively. Since cells maintained growth in both cases, comparable protein degradation along the induction phase was expected. Each sample was transferred to a shake flask with defined medium with and without glucose and cultivated at the same temperature as in the process. The medium was supplemented with 100 mg·L^{-1} chloramphenicol to stop protein synthesis in order to quantify solely protein degradation. In the absence of glucose neither cell growth nor RhuA degradation was observed, confirming the energy dependency of the activity loss of the recombinant protein (data not shown). In the case where glucose was present, specific RhuA contents remained nearly constant and activity loss was observed, suggesting that either the protein had been partially cleaved by energy-dependant proteases or that it was subject to conformational changes of its structure. Cells grown at 37° and induced for two generations showed a higher rate of activity loss compared to those induced for 1 generation (Figure 6.6).
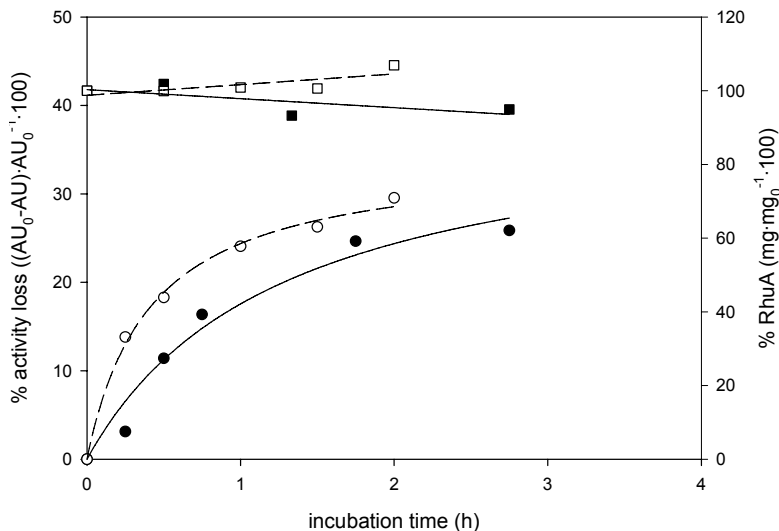

Figure 6.6. Percentages of specific activity loss and RhuA contents versus incubation time in cells harvested at different stages of a culture induced at 37ºC. Full circles represent the percentage of activity loss in cells from first generation (●), open circles the percentage of activity loss in cells from second generation (○), full squares the percentage of intracellular remaining RhuA in cells from first generation (■) and open squares the percentage of intracellular remaining RhuA in cells from second generation (□). Incubation was performed at 37ºC in defined medium with 10 g·L^{-1} glucose and in the presence of 100 mg·L^{-1} of chloramphenicol.

In contrast to these observations, no increase of proteolytic activity during the induction phase of the 28°C experiment was detected (Figure 6.7). Even though the incubation temperature was 9°C lower the obtained activity loss profiles almost comply with that of the 37° sample one

doubling past induction. Therefore, either the proteolytic activity measured in the *in vivo* assay is independent from the temperature (which would be inconsistent with the results obtained the *in vitro* experiments) or the differences observed in the *in vitro* experiments (Figure 6.3) were not significant.
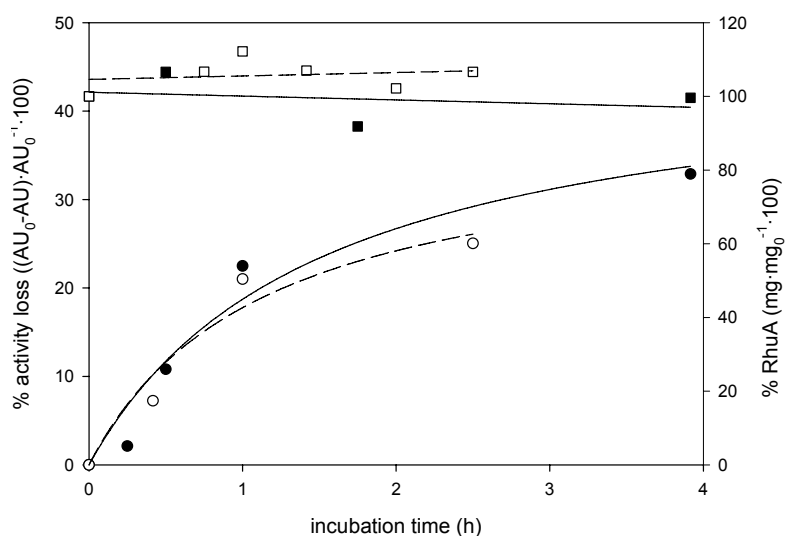
Figure 6.7. Percentages of specific activity loss and RhuA contents versus incubation time in cells coming from different stages of a culture induced at 28ºC. Full circles represent the percentage of activity loss in cells from first generation (●), open circles the percentage of activity loss in cells from second generation (○), full squares the percentage of intracellular remaining RhuA in cells from first generation (■) and open squares the percentage of intracellular remaining RhuA in cells from second generation (□). Incubation was performed at 28ºC in defined medium with 10 g·L^{-1} glucose and in the presence of 100 mg·L^{-1} of chloramphenicol.

6.4. Conclusions.

A decrease of specific aldolase activity was observed during fed-batch cultivation employing continuous inducer supply at 37ºC. By reducing the process temperature to 28°C, higher overall yields were obtained as cell growth and aldolase production rates could be maintained during the whole cultivation. At reduced temperature, lower copies of plasmid per cell probably helped to alleviate the host metabolism from the stress imposed by induction, and that could be the reason for increased enzyme activities.

Possible proteolytic degradation of the recombinant protein was considered to explain activity losses. *In vitro* and *in vivo* tests clearly proved energy dependent proteolytic activity but the temperature dependency of protease activity could not be clearly correlated. To better understand these results, further studies on protein folding could be considered. Moreover, currently available standard tests for proteolysis are not sufficiently elaborated to allow distinct statements concerning causes and effects responsible for RhuA activity loss. Therefore, development of reliable and reproducible assays for determination of proteolytic activity must be forced in future to investigate recombinant protein degradation during the bioprocess.

References.

[1] L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

[2] S.C.Makrides, Strategies for achieving high-level expression of genes in *Escherichia coli*, Microbiological Reviews 60 (1996) 512-&.

[3] A.Mitraki, J.King, Protein Folding Intermediates and Inclusion Body Formation, Nature Biotechnology 7 (1989) 690-697.

[4] A.L.Goldberg, A.C.John, Intracellular Protein Degradation in Mammalian and Bacterial Cells: Part 2, Annual Review of Biochemistry 45 (1976) 747-804.

[5] S.Gottesman, M.R.Maurizi, Regulation by Proteolysis - Energy-Dependent Proteases and Their Targets, Microbiological Reviews 56 (1992) 592-621.

[6] S.Gottesman, Proteases and their targets in *Escherichia coli*, Annual Review of Genetics 30 (1996) 465-506.

[7] A.Rozkov, S.O.Enfors, Analysis and Control of Proteolysis of Recombinant Proteins in *Escherichia coli*, Physiological Stress Responses in Bioprocesses, Springer Berlin/Heidelberg (2004) 163-195.

[8] L.Vidal, P.Ferrer, G.Alvaro, M.D.Benaiges, G.Caminal, Influence of induction and operation mode on recombinant rhamnulose 1-phosphate aldolase production by *Escherichia coli* using the T5 promoter, Journal of Biotechnology 118 (2005) 75-87.

[9] M.Cserjan-Puschmann, W.Kramer, E.Duerrschmid, G.Striedner, K.Bayer, Metabolic approaches for the optimisation of recombinant fermentation processes, Applied Microbiology and Biotechnology 53 (1999) 43-50.

[10] G.N.Caron, P.Stephens, R.A.Badley, Assessment of bacterial viability status by flow cytometry and single cell sorting, Journal of Applied Microbiology 84 (1998) 988-998.

[11] Z.Rymaszewski, W.A.Abplanalp, R.M.Cohen, P.Chomczynski, Estimation of Cellular DNA Content in Cell Lysates Suitable for RNA Isolation, Analytical Biochemistry 188 (1990) 91-96.

[12] G.L.Jordan, S.W.Harcum, Characterization of up-regulated proteases in an industrial recombinant *Escherichia coli* fermentation, Journal of Industrial Microbiology & Biotechnology 28 (2002) 74-80.

[13] J.Pinsach, C.de Mas, J.Lopez-Santin, Induction strategies in fed-batch cultures for recombinant protein production in *Escherichia coli*: Application to rhamnulose 1-phosphate aldolase, Biochemical Engineering Journal 41 (2008) 181-187.

[14] L.Andersson, L.Strandberg, S.O.Enfors, Cell segregation and lysis have profound effects on the growth of *Escherichia coli* in high cell density fed batch cultures, Biotechnology Progress 12 (1996) 190-195.

[15] L.Vidal, et al., High-level production of recombinant His-tagged rhamnulose 1-phosphate aldolase in *Escherichia coli*, Journal of Chemical Technology and Biotechnology 78 (2003) 1171-1179.

[16] H.Hellebust, M.Murby, L.Abrahmsen, M.Uhlen, S.O.Enfors, Different Approaches to Stabilize A Recombinant Fusion Protein, Bio-Technology 7 (1989) 165-168.

[17] M.Murby, M.Uhlen, S.Stahl, Upstream strategies to minimize proteolytic degradation upon recombinant production in *Escherichia coli*, Protein Expression and Purification 7 (1996) 129-136.

[18] M.Kanemori, K.Nishihara, H.Yanagi, T.Yura, Synergistic roles of Hs1VU and other ATP-dependent proteases in controlling in vivo turnover of sigma(32) and abnormal proteins in *Escherichia coli*, Journal of Bacteriology 179 (1997) 7219-7225.

[19] N.Dedhia, R.Richins, A.Mesina, W.Chen, Improvement in recombinant protein production in ppGpp-deficient *Escherichia coli*, Biotechnology and Bioengineering 53 (1997) 379-386.

[20] L.D.Simon, B.Randolph, N.Irwin, G.Binkowski, Stabilization of Proteins by A Bacteriophage-T4 Gene Cloned in *Escherichia coli*, Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 80 (1983) 2059-2062.

[21] K.Talmadge, W.Gilbert, Cellular Location Affects Protein Stability in *Escherichia coli*, Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 79 (1982) 1830-1833.

[22] S.Gottesman, M.R.Maurizi, CELL BIOLOGY: Enhanced: Surviving Starvation, Science 293 (2001) 614-615.

[23] J.Klein, P.Dhurjati, Protein aggregation kinetics in an *Escherichia coli* strain overexpressing a *Salmonella typhimurium CheY* mutant gene, Applied Environmental Microbiology 61 (1995) 1220-1225.

[24] R.Rosen, et al., Protein aggregation in *Escherichia coli*: role of proteases, FEMS Microbiology Letters 207 (2002) 9-12.

[25] C.Squires, C.L.Squires, The Clp proteins: proteolysis regulators or molecular chaperones?, Journal of Bacteriology 174 (1992) 1081-1085.

7. PRELIMINARY STUDIES ON MODELLING RECOMBINANT RHAMNULOSE 1-PHOSPHATE ALDOLASE PRODUCTION IN *E. coli*: IDENTIFICATION OF KEY PROCESS FEATURES.

Abstract.

The original aim of this work was to optimise recombinant rhamnulose 1-phosphate aldolase (RhuA) production process. The challenge was to obtain a recombinant protein production model for *Escherichia coli*, so that the optimal substrate and inducer feeding policies which maximised the productivity of the process could be determined. The description of the process was based on mass balances, transport phenomena, growth and protein production kinetics as well as stoichiometry. However, as experiments were performed, it became clear that it was hard to come up with a model which was able to describe all different situations which could potentially occur. Under these circumstances, the lack of robustness of the model seriously compromised the optimisation results. Thus, the understanding of the different behaviours which cells showed under those situations became of major importance. The main objective of this work was then focused on the identification and description of the most important phenomena involved in recombinant protein production process.

Growth was modelled considering glucose inhibition for *E. coli* M15 ΔglyA [pREP4] pQE$\alpha\beta$rham. Other inhibitions or limitations by acetate, oxygen and/or carbon dioxide were avoided because they negatively affected the process performance and, therefore, were excluded from the model. The description of the induction phase was based on experimental observations: segregation of cell population into a viable-but-not culturable state, IPTG transport, product inhibition, plasmid maintenance and copy number, viability, cell size changes, recombinant protein inactivation and degradation, etc. were studied and included into the proposed model when necessary. However, there were some features of the process which could not be properly described, evidencing that more experimental data was required. Amongst them, IPTG transport, protein degradation and inactivation as well as segregation of cells into a viable but not culturable state (VBNC) were the most significant ones and should be studied with more detail. A better understanding of these processes would not only help to obtain a robust model of the process, but also to improve the process itself by avoiding (or minimising) undesirable situations.

This work was partially done under the supervision of Prof. Jayant Modak in the Biotechnology Department of the Indian Institute of Science of Bangalore, India.

7.1. Introduction and objectives.

One of the goals in the study of recombinant bacteria is to maximise the production of foreign proteins. Since experiments are usually expensive and time-consuming, their number must be reduced as far as possible. As an alternative, mathematical techniques based on sufficiently accurate models can be used to determine an optimal process [1-4].

An individual cell is a complicated self-regulated biochemical reactor that can alter its biosynthetic machinery to meet the demands of a changing environment. Although significant efforts have been made to describe the dynamics of recombinant protein production, the establishment of fully mechanistic models to describe the cellular behaviour is, at least to date, not possible and therefore all models are based on simplifications [5].

Structured models [6-10] are developed to elucidate the interrelation between cell compartments, and metabolic flux analysis is used to understand the regulation of the central metabolism of cells [11]. Recent advances in "omics" data acquisition (genomics, transcriptomics, proteomics, metabolomics and localizomics) provide measurements of virtually all classes of biomolecules, but still a lot of work has to be carried out to model, simulate and analyse the whole cell at that level of description [12].

On the other hand, unstructured models are still useful for process design when knowledge is limited [13], but previously unstructured models [14-17] did not consider all factors influencing our process and, therefore, they were not reliable when using our experimental data to determine the optimal operation policies.

Other approaches used black- or grey-box models [16, 18-20] where no prior knowledge is required, but the performance of those models is strongly dependant on the training experiments, which are not necessarily close to the optimum. Moreover, identification of process bottlenecks becomes even more difficult when these kinds of models are used. For industrial process control, hybrid models which combine mathematical models and process data (real-time or historical) are very useful, but still not reliable enough to optimise the process.

As a matter of fact, a model that is sufficient for optimising the process might be much simpler than a model which comprehensively describes all the available experimental data, because there is no need to describe accurately the features which are not met by the process and do not influence

its performance significantly. Modelling of biological systems is a complex task due to non-linear dynamics and the ability of the microorganisms to adapt to changing environmental conditions, but efforts can be saved by avoiding the investigation of situations which are far from the optimum operation [21].

Thus, the proposed model pretends to be as simple as possible, but at the same time it pretends to describe the most important features influencing the process. It is an unstructured model based on mass balances, transport phenomena, growth and production kinetics as well as stoichiometry. However, not all the experiments were described by a single set of parameters, suggesting that still not all the necessary information had been included into the model. For this reason, the following work only pretends to set up the basics for further model development, identifying some important features of RhuA production process which should be taken into account. It also indicates some points which are not clear enough and, therefore, further work should be focused on them to improve their understanding.

Obviously, part of the data obtained in other chapters has been used in this chapter. The reader might find some experimental data which has been already presented, but it has been included to help the reader to follow the present discussion (these cases are explicited in the text and referenced to the chapter where the data had been previously presented).

7.2. Materials and methods.

Software.

The modelling tasks were performed using gPROMS ModelBuilder (Process Systems Enterprise Ltd., UK). The mathematical methods used to solve the combination of differential and algebraic equations defined in the model were the preset standard ones in gPROMS for dynamic model-based activities. The differential-algebraic equation solver uses both a non-linear equation solver and a linear equation solver for parameter estimation ("PESolver") or dynamic simulation model-based activities ("DASolver", based on variable time step/variable order backward differential formulae). At the same time, the non-linear equation solver requires two different sub-solvers: one for the initialisation of dynamic problems ("NLSOL") and another one for reinitialisation of the problem after discontinuities ("SPARSE", which provides a sophisticated implementation of a Newton-type method). The linear equation solver also uses two different sub-solvers: one for the linear algebraic equations ("MA48", based on LU-factorisation algorithms) and another one for the non-linear algebraic equations ("SPARSE"). The parameters used in the solvers were also the preset ones, except for the tolerance of the parameter estimation method, which was set to $1 \cdot 10^{-5}$. The precision of the parameter estimates was calculated considering 95% confidence intervals.

Experimental data.

All the experiments used were performed as described in previous chapters. From those experiments, the following data were required to run parameter estimation problems:

- Initial conditions: biomass concentration (X_0), substrate concentration (S_0), system volume (V_0), inducer concentration (I_0), and product concentration (P_0).
- Time-invariant controls: substrate feed concentration (S_F) and inducer feed concentration (I_F).
- Piecewise-constant controls: substrate feed rate (F_S) and inducer feed rate (F_I).
- Measured data along time: biomass concentration ($X(t)$), substrate concentration ($S(t)$), and specific product contents ($P/X(t)$).

In order to simulate the process' variables along time, only the initial conditions and controlled variables were required.

All these data were introduced to the software without prior calculations (there was no need for interpolation).

7.3. Results and discussion.

Since the simplest RhuA production model which was considered included a significant number of parameters to be estimated, we first focused on modelling cell growth of *E. coli* M15 ∆glyA [pREP4] pQEαβrham in non-induced cultures.

Once growth model had been calibrated, protein production and its effects on cell growth were studied.

7.3.1. Modelling the growth phase.

For the sake of simplicity, biomass and substrate were the only variables chosen to describe cell growth at 37ºC. Acetate, CO_2 and O_2 are other variables which could influence the process performance [16, 21-25], but different strategies were used to minimise their influence.

A defined medium utilising glucose as carbon source was used in all experiments, and pure oxygen when required in fed-batch experiments to ensure aerobic conditions. However, *E. coli* is known to produce many by-products when grown aerobically on glucose. Amongst them, acetate has been identified as a cause for reduced biomass as well as protein yields when accumulated in the medium [26-28]. To reduce acetate formation, strategies ranging from strain improvement to the control of cultivations conditions have been successfully applied [29-32]. In particular, the control of the specific growth rate by means of exponential substrate feeding had been previously applied to avoid acetate accumulation in *E. coli* M15 ∆glyA [pREP4] pQEαβrham fed-batch cultures [33, 34].

Preliminary shake flask cultures were performed to characterize growth and acetate production of *E. coli* M15 ∆glyA [pREP4] pQEαβrham at 37ºC in defined media. Ten non induced shake flask cultures were grown at initial glucose concentrations ranging from 0 to 90 $g·L^{-1}$ in order to define the growth kinetics due to substrate consumption (μ_S). Biomass, substrate and acetate concentrations were measured along time in all cultures.

Similar initial specific growth rates were measured in the case of low initial glucose concentrations. However, large differences were observed in the case of cultures grown at high initial glucose concentrations. When plotting the initial specific growth rates (μ_0) versus initial glucose

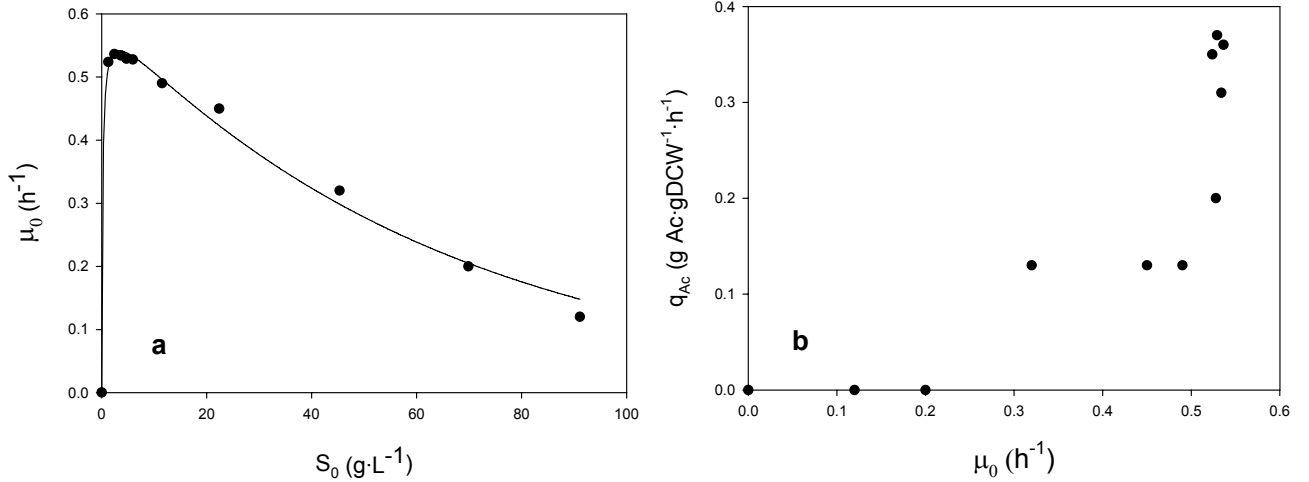concentrations (S_0) for each culture, substrate inhibition growth kinetics was obtained (Figure 7.1a).

Figure 7.1. Initial specific growth rate vs. initial glucose concentration (a) and specific acetate production rate vs. initial specific growth rate (b) in shake flask cultures. Experimental data is represented by full circles (●) and model predictions by continuous line (−).

Values of μ_{max}=0.60±0.02 h^{-1}, K_S=0.18±0.07 g·L^{-1} and K_{SI}=65.06±3.81 g·L^{-1} were determined when the experimental results were fitted to the Aiba-Edwards model [35, 36], which was the growth equation (Eq. 7.1) which best described the experimental data.

$$\mu_S = \frac{\mu_{max} \cdot S}{K_s + S} \cdot e^{-\frac{S}{K_{SI}}}$$
(Eq. 7.1)

The standard error for K_S was considered as meaningless due to the limited precision (0.1 g·L^{-1}) of the HPLC method used to measure glucose concentration.

These preliminary shake flask experiments also allowed the identification of the conditions under which acetate was accumulated. During the exponential growth phase of the previously presented cultures, specific acetate production rates (q_{Ac}) were correlated to the specific growth rates (Figure 7.1b). High specific acetate production rates were measured at specific growth rates above 0.50 h^{-1}. At specific growth rates between 0.49 to 0.32 h^{-1}, the specific accumulation rates remained constant and close to 0.13 g·gDCW^{-1}·h^{-1}. Acetate accumulation was switched off at a specific growth rate between 0.32 h^{-1} and 0.20 h^{-1}, since it was not detected below this range. This critical

specific growth range was in accordance to the critical specific growth rate determined in substrate-limited fed-batch cultures [33], although in that case the specific acetate accumulation rates were lower due to the ability of the cells to assimilate acetate under glucose limitation. For this particular strain, inhibiting acetate concentrations could be avoided growing at specific growth rates up to 0.25 h^{-1}.

Therefore, the model was based on the following mass balances in fed-batch cultures:

$$\frac{dS}{dt} = \frac{F_S}{V} \cdot (S_F - S) - q_S \cdot X_1 \qquad \text{(Eq. 7.2)}$$

$$\frac{dX_1}{dt} = \mu_1 \cdot X_1 - \frac{F_S}{V} \cdot X_1 \qquad \text{(Eq. 7.3)}$$

$$\frac{dV}{dt} = F_S \qquad \text{(Eq. 7.4)}$$

where S stands for substrate concentration ($g \cdot L^{-1}$), X_1 for biomass concentration ($gDCW \cdot L^{-1}$), V for the system volume (L), F_S for the substrate feed ($L \cdot h^{-1}$) and S_F for the substrate feed concentration ($g \cdot L^{-1}$).

The specific growth rate (μ_1), substrate consumption kinetics (q_S) and the biomass concentration were defined as:

$$\mu_1 = \mu_S \qquad \text{(Eq. 7.5)}$$

$$q_S = \frac{1}{Y_{XS}} \cdot \mu_S + m_1 \qquad \text{(Eq. 7.6)}$$

$$X_1 = OD600 \cdot f_{gDCW/OD600} \qquad \text{(Eq. 7.7)}$$

where Y_{XS} represents the substrate yield to biomass ($gDCW \cdot g^{-1}$), m_1 is the maintenance coefficient ($g \cdot gDCW^{-1} \cdot h^{-1}$), OD600 is the optical density at 600 nm and $f_{gDCW/OD600}$ is the factor relating the OD600 to the dry cell weight concentration.

The optical density at 600 nm and the dry cell weight concentration were experimentally measured in different cultures, and both variables could be correlated within a broad range of conditions. The conversion factor ($f_{gDCW/OD600}$) was calculated as 0.31 $gDCW \cdot L^{-1} \cdot OD600^{-1}$ (Figure 7.2).
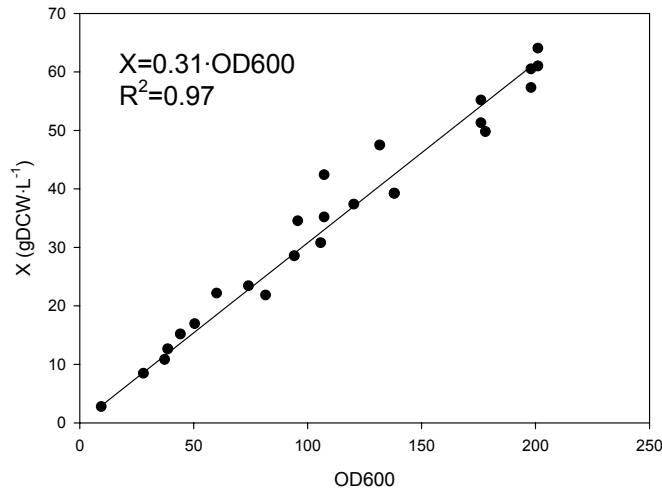
Figure 7.2. Dry cell weight concentration versus optical density at 600 nm in non-induced fed-batch cultures. Dots represent the experimental data (●) and continuous line the linear regression (−).

The kinetic parameters describing cell growth from preliminary shake flask experiments were used as initial estimates for the dynamic model calibration in fed-batch mode. 6 different bioreactor experiments performed under different conditions were employed to calibrate the growth model proposed from Eq. 7.1 to Eq. 7.7 (Figure 7.3). The experiments were performed in such a manner that the substrate concentration, the final dry cell weight concentration and the specific growth rate were different in each case in order to cover a broader range of conditions. The aim was to obtain a single set of parameters for the proposed model which could describe all those different situations at the same time. Thus, the parameter estimation problem was run using 6 experiments at the same time. Results were acceptable for most of the parameters except for the K_S, due to the precision of the substrate analysis method:

$$\mu_{S,\max} = 0.59 \pm 0.01 \qquad h^{-1}$$
$$K_S = 0.002 \pm 0.009 \qquad g \cdot L^{-1}$$
$$K_{SI} = 64.36 \pm 0.009 \qquad g \cdot L^{-1}$$
$$Y_{XS} = 0.34 \pm 0.02 \qquad g \cdot g^{-1}$$
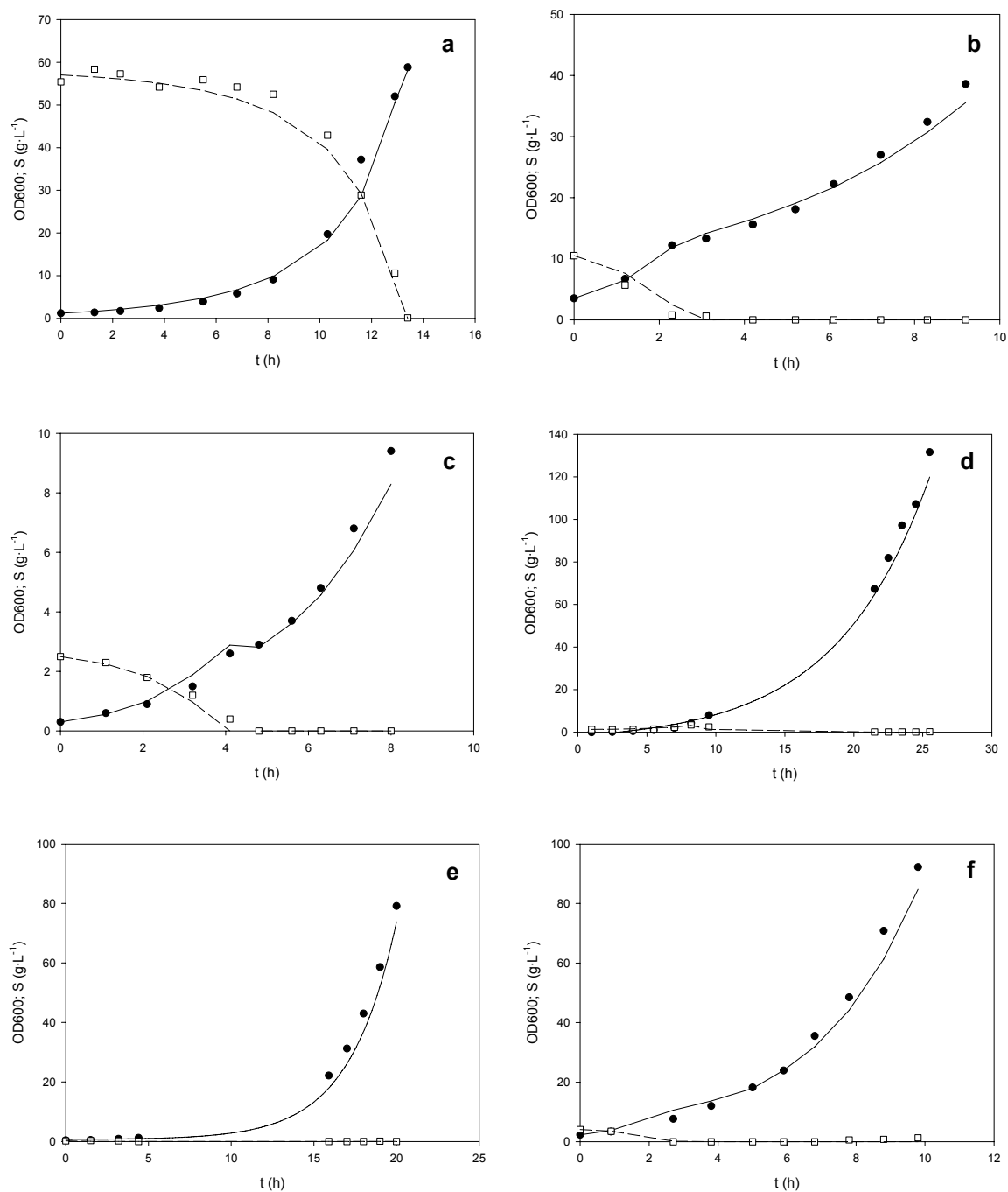$$m_1 = 0.104 \pm 0.0002 \quad g \cdot g^{-1} \cdot h^{-1}$$

Figure 7.3. Experimental values for optical density at 600 nm (●) and substrate concentration (□) in different fed-batch cultures performed at a range of initial substrate concentrations from 0 to 60 g·L⁻¹ and at specific growth rates up to 0.3 h⁻¹. Continuous lines depict the model estimation for OD600 (−) while dashed lines represent the model estimation for glucose concentration (− −).

The model was then validated by comparing the experimental data of two high cell density fed-batch cultures with model simulations (Figure 7.4).
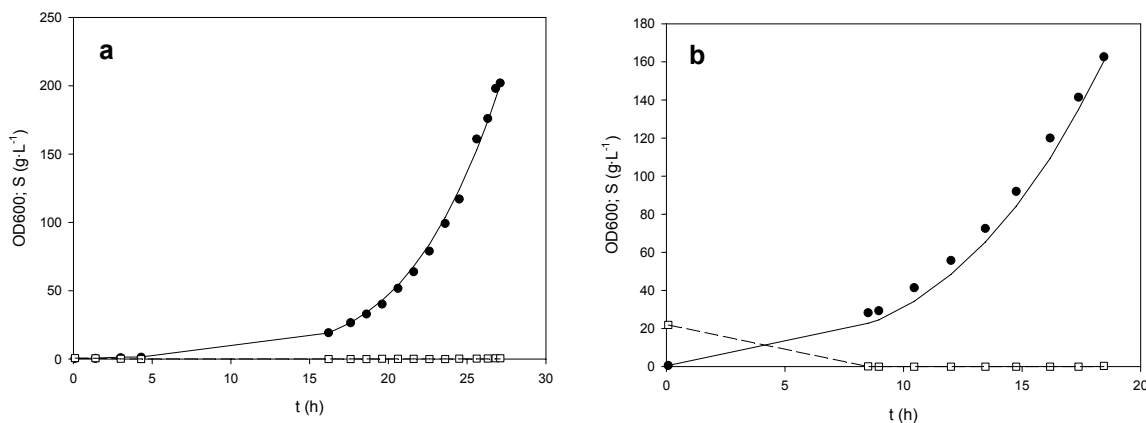

Figure 7.4. Experimental values for optical density at 600 nm (●) and substrate concentration (□) in high cell density fed-batch cultures performed at specific growth rates up to 0.3 h^{-1}. Continuous lines depict the model simulation for OD600 (−) while dashed lines represent the model simulation for glucose concentration (− −).

Considering that this set of parameters described fairly well the growth behaviour under different conditions even at high cell densities, the structure of the growth model was fixed to investigate RhuA production and its effects over cell growth.

7.3.2.   Modelling the production phase.

Lac-derived inducible systems are commonly used for recombinant protein production when using *E. coli* as a host. In these cases, protein overexpression is usually switched on by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to the culture once the cells have grown to a certain biomass concentration. As explained in chapter 5, RhuA overexpression was under the control of a strong *lac*-derived promoter [37], and substrate-limited fed-batch cultures were induced either with a pulse or with a continuous feed of IPTG depending on the case [33].

Initially, a relatively simple model (unstructured and non-segregated) was considered to describe RhuA production at 37ºC. It was based on previous work published by Lee and Ramirez [15], which took into account the effects of inducer on cell growth and foreign protein production. Although cell growth was also affected after IPTG addition in our system, this model was not able to describe our

experimental data. Further observations increased the complexity of the model by considering segregation of cells into a viable but not culturable (VBNC) state, IPTG transport into the cell, protein deactivation as well as protein degradation. When considering all these phenomena, the ability of the model to capture reality significantly improved, but still some questions remained to be answered.

To describe RhuA production, in addition to the variables considered when growth was modelled, the inducer as well as the product had to be included into the model. Thus, the variables chosen to describe recombinant protein production were biomass, substrate, product and inducer concentrations as well as the system volume.

7.3.2.1.   The biomass.

Cell segregation into viable-but-not-culturable (VBNC) state.

Results from chapter 6 (example already presented in Figure 6.1) indicated that progressive induction led to high levels of recombinant protein in fed-batch cultures where growth was controlled by limiting the glucose addition but, finally, cells stopped growing, indicating their inability to divide (Figure 7.5).
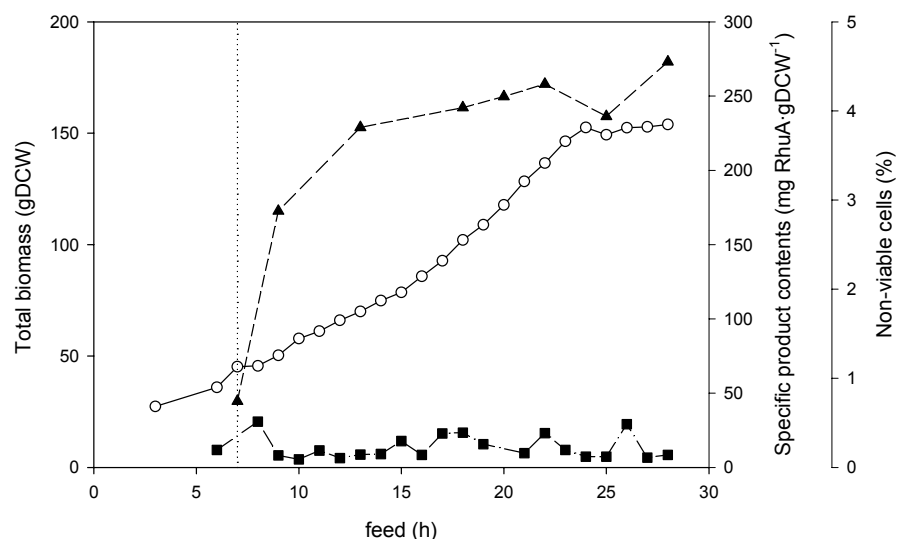

Figure 7.5. Total biomass (−○−), specific RhuA contents (−▲−) and non-viability (−■−) of *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells in a substrate-limited fed-batch culture with progressive induction. The vertical dotted line shows when the continuous induction was switched on.

Previous segregated models proposed in the literature which described growth arrest [38, 39] considered up to four different cell subpopulations: dividing, non-dividing, lysed and divinding cells which had lost the plasmid. In our case, only the first two subpopulations were used to describe the whole of the biomass. Lysed cells were not considered because there was not a significant decay of the total biomass after cells stopped dividing. Plasmid-free cells were also neglected because the host and vector used (*E. coli* M15 ΔglyA [pREP4] pQEαβrham) ensured the plasmid maintenance in defined media [34]. Thus, only dividing and non-dividing cells were considered (all of them bearing plasmids). In addition, flow cytometric analyses indicated that non-viable cells were below 1% all the time, even after biomass build-up had stopped (Figure 7.5). Therefore, based on experimental biomass and viability measurements, non-dividing cells were considered as viable-but-not culturable (VBNC) ones.

When taking into account these considerations, the total measured biomass became the sum of the viable-and-culturable population (X_1) plus the viable-but-not-culturable one (X_2), as indicated in Eq. 7.8.

$$X = X_1 + X_2 \qquad\qquad \text{(Eq.7.8)}$$

It was considered that X_1 were able to grow and overexpress RhuA, while X_2 were neither able to grow nor able to overexpress RhuA. The conversion rate from X_1 to X_2 was named α (Figure 7.6).
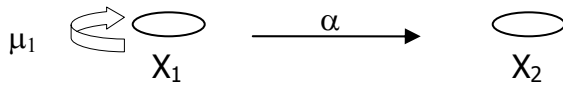

Figure 7.6. Schematic representation of the segregated model.

The mass balances for both cell populations in fed-batch cultures became:

$$\frac{dX_1}{dt} = (\mu_1 - \alpha) \cdot X_1 - \left(\frac{F_S + F_I}{V} \right) \cdot X_1 \qquad\qquad \text{(Eq. 7.9)}$$

$$\frac{dX_2}{dt} = \alpha \cdot X_1 - \left(\frac{F_S + F_I}{V} \right) \cdot X_2 \qquad\qquad \text{(Eq. 7.10)}$$

where F_S and F_I are the substrate and inducer feed rates, respectively (L·h^{-1}).

Growth and segregation rates.

The specific growth rate measurements were considered as "apparent" ones, since they depended on the specific growth rate of active cells and the fraction they represented over the total biomass:

$$\mu = \mu_1 \cdot \frac{X_1}{X}$$
(Eq. 7.11)

where μ_1 is defined by Eq. 7.1.

Considering that growth arrest happened because of segregation of cells into a VBNC state, the cause for this segregation had to be included into the model in order to describe the biomass behaviour. Segregation had been previously attributed to different causes:

   i.   Maintenance of foreign DNA.
   ii. Expression of foreign DNA.
   iii. Intracellular accumulation of recombinant protein.
   iv. IPTG effects on host metabolism.

Although high values of the specific activity production rate coincided with segregation of cells to a VBNC state, it often coincided with high inducer concentrations and, sometimes, with high intracellular RhuA activities. Therefore, each one of the potential factors affecting segregation was investigated:

i.   Raw material and energy are withdrawn from the host metabolism for maintenance of the foreign DNA [40]. Gene dosage has been reported as a cause for increased metabolic load in *E. coli* when using strong promoters [41]. Excessive plasmid replication has been observed after inducing recombinant gene expression in ColE1-derived plasmids [42], and the pQE-derived plasmid used in this work is one of them (www.qiagen.com).

As part of the work presented in chapter 6, the plasmid copy number (PCN) was monitored along susbtrate-limited fed-batch cultures with progressive induction to study the potential relationship between this variable and growth arrest (Figure 6.5a). Results showed steady values of PCN in cultures conducted at 37ºC, even for the helper plasmid. These steady values of the PCN could be also explained by a simultaneous increase of the PCN in a certain

cell subpopulation while other cells could have lost the plasmid, but colony counts indicated that 100% of the cells were bearing the pQE-plasmid (they were resistant to ampicillin all over the run; data not shown). Since the metabolic load due to plasmid maintenance was not increasing along the induction phase, gene maintenance was discarded as a cause for segregation and not considered for modelling purposes.

ii. Overexpression of recombinant proteins has been related to a metabolic burden on the host cell. It can be explained by a rapid increase in the demand of metabolites [43]: the host might be burdened by protein synthesis due to shortage of certain aminoacid pools [44]. Thus, growth arrest might be related to high specific production rates (q_P).

Experimental observations showed that the process strategy strongly influenced the way in which RhuA overexpression rate affected the specific growth rate. Results from chapter 5 indicated that, in the case of pulse induced shake flask cultures, specific growth rate decay was observed at specific activity production rates above 1000 AU·gDCW^{-1}·h^{-1} (Figure 5.2).

However, when looking at substrate-limited fed-batch cultures induced with an IPTG pulse, growth was arrested in all cases, even though the maximum specific activity production rates were between 300 and 800 AU·gDCW^{-1}·h^{-1} in all experiments (Figure 5.6). Under those circumstances, cell growth never lasted for more than 4 h during the induction phase (data not shown), suggesting that limiting substrate levels during the induction phase of the cultivation might favour segregation. To avoid this sudden metabolic burden, other substrate-limited fed-batch experiments were induced with a continuous IPTG feed in order to tune q_P at low values (results also from chapter 5). The tuning of the specific production rate below 100 AU·gDCW^{-1}·h^{-1} significantly extended the induction phase (Figure 7.7), indicating that in pulse induced experiments where growth was limited by substrate addition the expression rates were above the critical values which cells could stand. In the case of continuos induction, cell growth did not stop when the maximum q_P was reached, but later on, suggesting that biomass build-up stopped probably due to other reasons (data not shown).
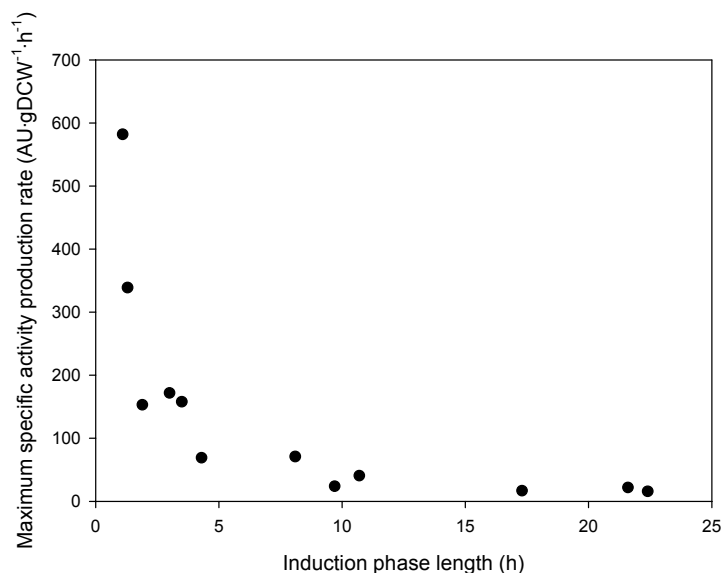
Figure 7.7. Maximum specific activity production rate versus induction phase length in substrate-limited fed-batch cultures conducted at 37ºC and induced with a continuous IPTG feed.

iii.   It is also accepted that high intracellular contents of recombinant protein might be toxic to the host cell. Some authors include a product inhibition term in growth kinetic equations [14] to take this fact into account. Then, α could also depend on the accumulated amount of intracellular RhuA (specific product contents, P/X). But RhuA was not the only recombinant protein overexpressed by the host cells: expression of the auxotrophy selection marker (SHMT) under the control of the constitutive P3 promoter inserted in the pQE plasmid could be another reason for an increased metabolic load on the host.

Results from chapter 5 showed that, in shake flask cultures, growth stopped when specific activities above 1000 AU·gDCW^{-1} were measured (Figure 5.2). In substrate-limited fed-batch cultures induced with an IPTG pulse, biomass build-up stopped in all cases, although in some cultures the maximum specific activity was below 400 AU·gDCW^{-1} (Figure 5.6). When the specific production rate was tuned by continuous IPTG dosage in substrate-limited fed-batch cultures, cells stopped dividing even at specific activities below that value (Figure 7.8). Apparently, the accumulated intracellular activity did not look to be responsible for growth arrest, at least in those cultures in which low specific activities were measured.

However, interesting results were obtained when comparing the intracellular RhuA levels to specific activities (results from chapter 6). The specific activity did not correlate to the specific protein contents and, although growth stopped at relatively low values of intracellular activity, RhuA contents within cells was considerably high (Figure 7.9 clearly shows it for the experiment already presented in Figure 6.1). Experimental data revealed a decay of the specific RhuA activity along the induction phase. Those results indicated that elevated RhuA levels within the cells (but not specific activity contents) could not be discarded as a reason for segregation, and that specific activity losses were taking place and should be considered.
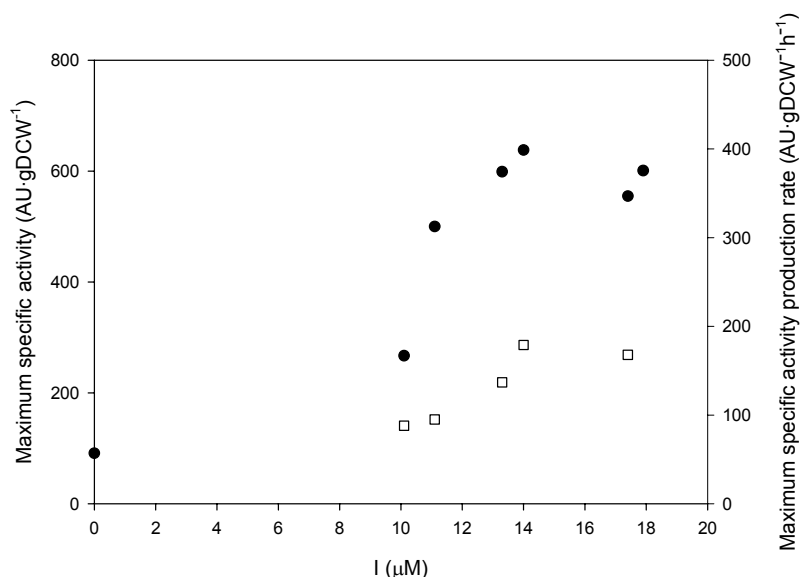

Figure 7.8. Maximum specific activity (●) and specific activity production rate (□) versus final inducer concentration in substrate-limited fed-batch cultures performed at 37ºC with continuous IPTG addition.

But, as indicated before, RhuA was not the only protein overexpressed by host cells. The auxotrophy-based system for plasmid maintenance had been developed by deleting the *glyA* gene from the host chromosome and cloning it into the pQE- expression vector under the control of the P3 constitutive promoter [34]. When comparing the non auxotrophic expression system (*E. coli* M15 [pREP4] pQErham) to the auxotrophic one (*E. coli* M15 ΔglyA [pREP4] pQEαβrham), remarkable differences were observed in fed-batch cultures (even though similar behaviour was observed in shake flask experiments, data not shown). First, when cultivated and induced under the same conditions, the old system did not show cell segregation into a viable but not culturable state, while the new one did (Figure 7.10).
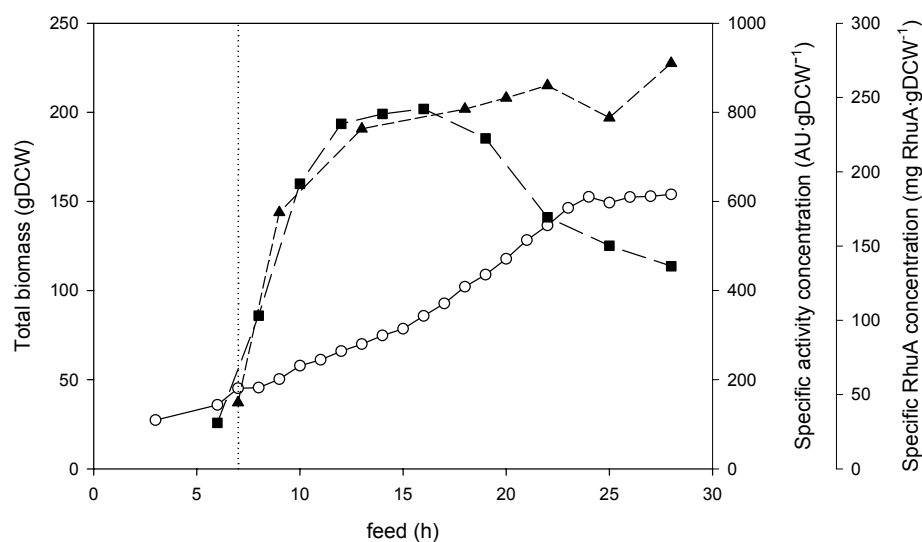
Figure 7.9. Total biomass (−○−), specific activity concentration (−■−) and specific RhuA concentration (−▲−) in a substrate-limited fed-batch culture with progressive induction conducted at 37ºC. The vertical dotted line shows when the continuous IPTG feed was switched on.

The non-auxotrophic construction showed a plasmid loss around 25% at the end of the culture, but colony counts indicated that the percentage of plasmid-bearing cells was nearly 100% for the auxotrophic strain (data not shown). These results explained why the auxotrophic strain was not able to keep on growing, and strongly suggested that the plasmid was somehow related to cell segregation. SDS-PAGE gels showed that the auxotrophic strain accumulated RhuA to 36% and SHMT to 27% of the total intracellular soluble protein (63% altogether), while the old system only accumulated RhuA to 40% (Figure 7.11).
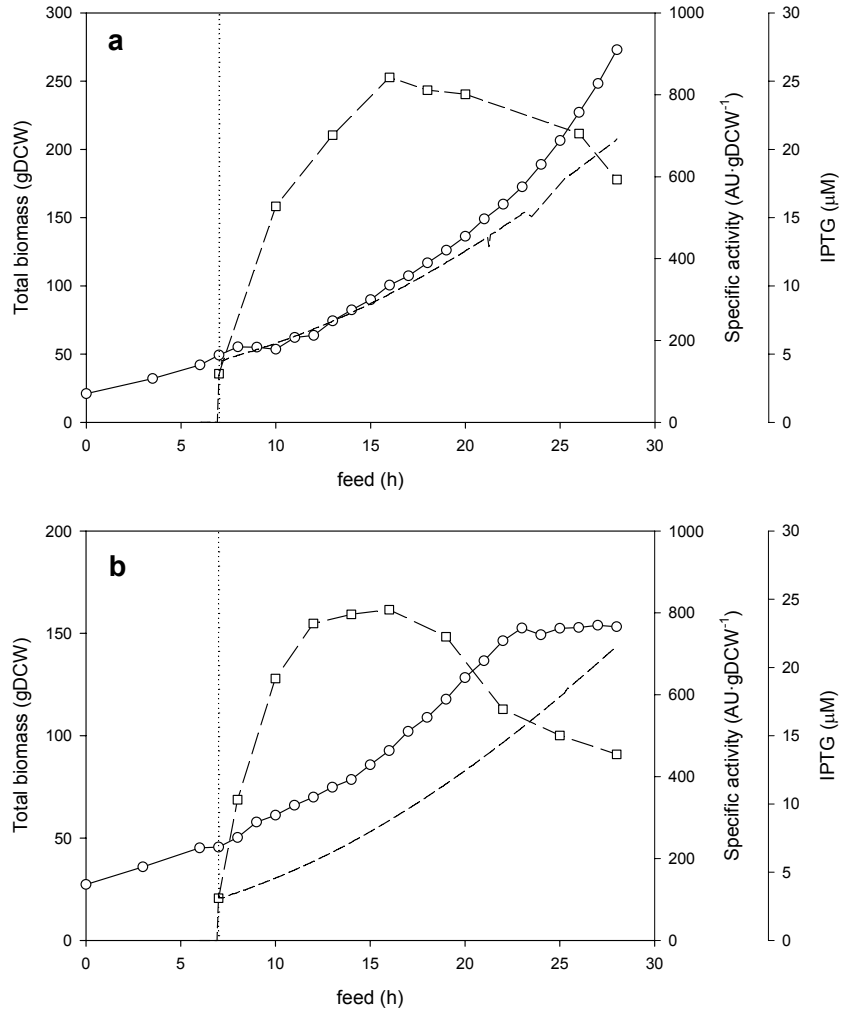
Figure 7.10. Substrate-limited fed-batch cultivations for *E. coli* M15 [pREP4] pQErham (a) and *E. coli* M15 ΔglyA [pREP4] pQEαβrham (b) strains at a controlled specific growth rate of 0.1 h⁻¹ and at 37ºC. The vertical dotted line indicates when the IPTG feed pump was switched on, open circles represent the total biomass (−○−), open squares the specific activity (−□−) and the dashed line represents the volumetric IPTG concentration (- - -).

Thus, *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells experienced an increased metabolic load due to excessive SHMT overexpression, which was identified as another reason contributing to cell segregation. In accordance to pQEαβrham plasmid copy number determinations (Figure 6.5a), intracellular SHMT levels inside the cells were steady, but extremely high (Figure 7.12). Due to these steady values, increased metabolic load caused by SHMT was not considered for modelling purposes, but one should keep in mind that the maximum attainable RhuA levels are strongly influenced by SHMT.
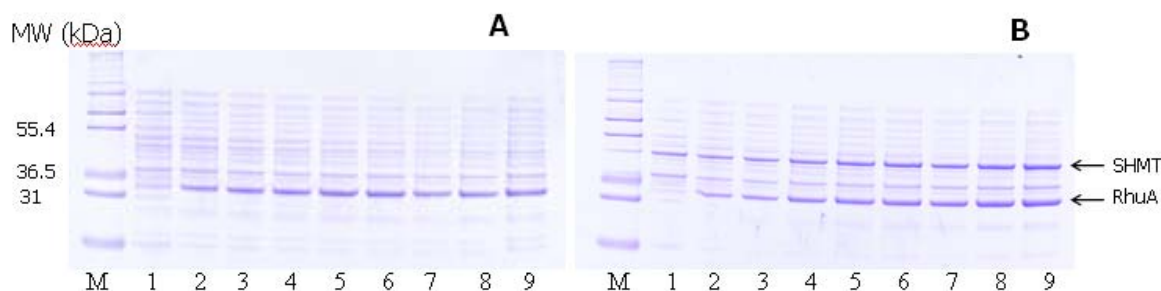
Figure 7.11. SDS-PAGE of substrate-limited fed-batch cultivations with *E. coli* M15 [pREP4] pQErham (A) and *E. coli* M15 ΔglyA [pREP4] pQEαβrham (B) strains at a controlled specific growth rate of 0.1 h^{-1}. Lane M: molecular weight markers. Lanes 1 to 9 correspond to samples withdrawn at feed (h)= 7 (before induction), 10, 13, 16, 19, 22, 25 and 28, respectively. RhuA can be identified at 32 kDa and SHMT at 46 kDa.
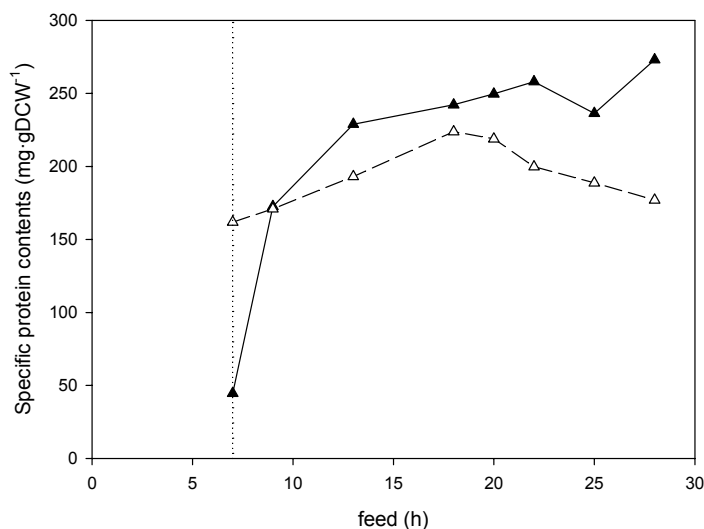
Figure 7.12. Specific RhuA (−▲−) and SHMT (−Δ−) levels in an *E. coli* M15 ΔglyA [pREP4] pQEαβrham fed-batch culture at a specific growth rate of 0.1 h^{-1} and at 37ºC with continuous induction.

iv.   Finally, the loss of the ability to divide had been also related by some authors to the amount of inducer added to the culture [15, 39]. It has been reported that IPTG can influence the host cell metabolism even when no plasmid is present [45]. Therefore, the inducer itself was investigated as a possible reason for segregation.

As described in chapter 5, cell growth stopped at IPTG concentrations above 20 μM in the case of shake flask cultures. In substrate limited fed-batch cultures induced with an IPTG

pulse, cell growth stopped even at lower IPTG concentrations, as in substrate-limited fed-batch experiments with continuous induction.

In order to check whether the IPTG itself influenced the host metabolism without the influence of foreign DNA maintenance and expression, IPTG was supplemented to *E. coli* M15 ΔglyA cultures. Those were plasmid-free cells which required glycine supplementation to the synthetic media in order to grow. 5 parallel shake flask cultures plus a control (without glycine supplementation) were induced with an IPTG pulse which ranged from 0 to 100 μM depending on the case. At the range of concentrations studied, the IPTG did not significantly affect growth in any case (Figure 7.13). Thus, the inducer itself was discarded as a reason for the segregation of the cells into a VBNC state, at least within the studied IPTG concentration ranges.
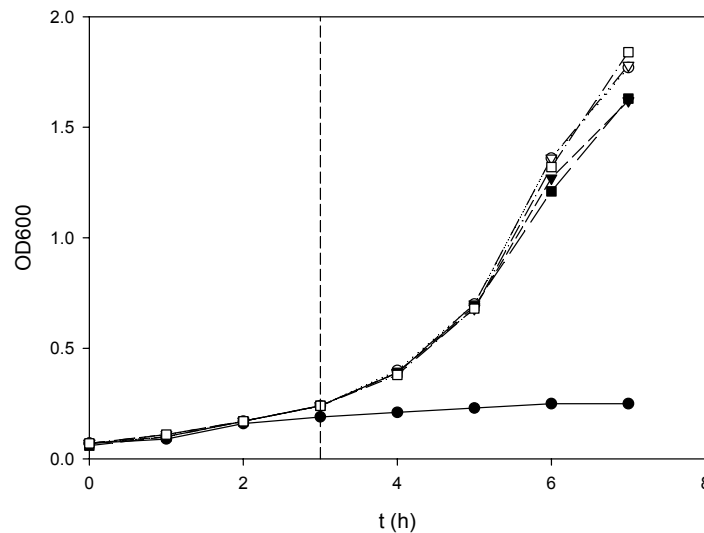
Figure 7.13. Growth profiles in induced shake flask cultures of plasmid-free cells. The dashed line represents the corresponding IPTG pulse addition in the following cases: control culture without glycine supplementation (−●−), non-induced culture (−○−), 10 μM IPTG (−▼−), 40 μM IPTG (−▽−), 70 μM IPTG (−■−) and 100 μM IPTG (−□−).

To sum up, the experimental data suggested that α could be influenced by:

-   Intracellular recombinant protein concentration (P/X, not necessarily related to intracellular activity). In this case, both RhuA and SHMT should be considered, but since SHMT levels

were measured as steady along the induction phase, only the intracellular RhuA levels were included into the model.

- RhuA specific production rate (q_P, whose influence depends on substrate levels).

Thus, α could be described by:

$$\alpha = a \cdot \left(1 - \frac{1}{1 + e^{\left(\frac{P}{X} - \frac{P}{X}\big|_{crit} \right)}} \right) + b \cdot q_P \cdot \frac{c}{c + S} \qquad \text{(Eq. 7.12)}$$

where a, $(P/X)_{crit}$, b and c were parameters to be estimated. The first term of Eq. 7.12 represents a sigmoidal function where segregation into a VBNC state is switched on at P/X values above $(P/X)_{crit}$, while the second term represents the influence of high specific production rates, specially at low substrate levels.

7.3.2.2.   The substrate.

The substrate mass balance when considering the segregated model depicted in Figure 7.6 became:

$$\frac{dS}{dt} = \frac{F_S}{V} \cdot (S_F - S) - \frac{F_I}{V} \cdot S - q_S \cdot X_1 - m_2 \cdot X_2 \qquad \text{(Eq. 7.13)}$$

where m_2 represented the maintenance coefficient of the VBNC population (g·gDCW^{-1}·h^{-1}) and q_S is described by Eq. 7.6. However, the substrate to biomass yield as well as the maintenance coefficient for X_1 should be different in the case of induced cultures because part of the substrate is used for RhuA production and the whole metabolism is altered.

7.3.2.3.   The product.

Although high intracellular levels of RhuA were attained, specific enzyme activity losses were observed during induction (results presented in chapter 6 and in Figure 7.9). Since the target of the

process was to obtain active RhuA for its further use as biocatalyst, both product concentration (P, in mg·L^{-1}) and activity concentration (AC, in AU·L^{-1}) were measured. The enzyme activity (EA, in AU·mg RhuA^{-1}) was defined as:

$$EA = \frac{AC}{P}$$
(Eq. 7.14)

As described in chapter 6, experiments were performed to study these undesired activity losses. Cells withdrawn from the bioreactor at different stages of the induction phase (before and after growth had been arrested) were incubated in shake flasks in order to study protein inactivation and degradation. Chloramphenicol was supplemented to the incubation media to stop protein synthesis, so that the specific RhuA variations accounted only for degradation. Although RhuA degradation was not significant during these short-time tests, activity losses were observed. This inactivation was significant in both X_1 and X_2, but it happened faster in VBNC cells (Figure 6.6). Results also showed that, after 24 h of incubation, approximately 45% of the recombinant protein was degraded within cells which were still growing, while RhuA degradation within VBNC cells was about 20% (Figure 7.14).
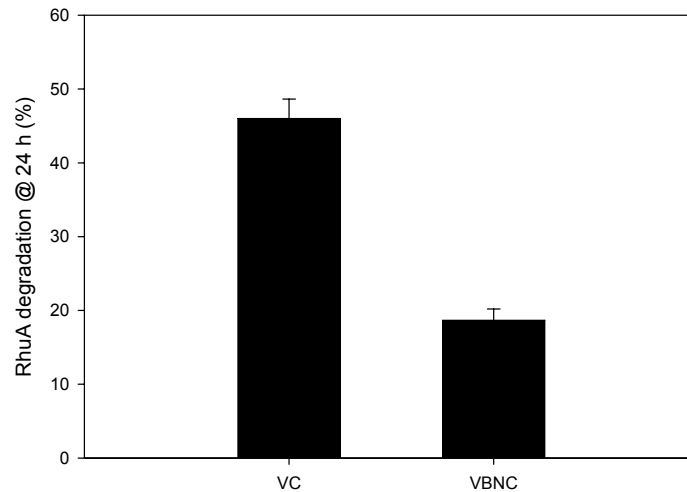
Figure 7.14. Percentage of RhuA (measured as mg RhuA·gDCW^{-1}) degradation in cells withdrawn from the bioreactor and incubated for 24 h in shake flasks in the presence of chloramphenicol. VC represents the viable and culturable cells (X_1), while VBNC represents viable-but-not-culturable cells (X_2).

It became evident that activity losses as well as protein degradation should be included into the model to describe recombinant protein production, but further experiments should be carried out in

order to better understand (and avoid) both phenomena. To simplify the modelling task, the product was modelled as protein contents. It was considered that the protein production should be properly described before trying to understand activity losses.

Taking into account these considerations, the mass balance for RhuA in fed-batch mode led to:

$$\frac{dP}{dt} = q_{P1} \cdot X_1 - \left(\frac{F_S + F_I}{V} \right) \cdot P \qquad \text{(Eq. 7.15)}$$

where P represented the product concentration in mg RhuA·L^{-1}, q_{P1} the specific production rate of X_1 in mg RhuA·gDCW^{-1}·h^{-1}, the second term accounted for dilution and a third term for RhuA degradation, which mainly happened in the fraction of producing cells, could be missing.

The specific production rate measurements were considered as "apparent", since q_P was the result of RhuA production by the fraction of active biomass:

$$q_P = q_{P1} \cdot \frac{X_1}{X} \qquad \text{(Eq. 7.16)}$$

The intrinsic specific production rate depended on intracellular IPTG concentration and on the specific growth rate of the producing biomass, since previous work indicated production of RhuA was growth-related [33]:

$$q_{P1} = \left(q_{P0} + \frac{q_{P\max} \cdot I_i}{K_I + I_i} \right) \cdot \frac{\mu_1}{\mu_{\max}} \qquad \text{(Eq. 7.17)}$$

where q_{P0} accounts for RhuA expression rate before induction (mg·gDCW^{-1}·h^{-1}), q_{Pmax} is the maximum specific RhuA production rate (mg·gDCW^{-1}·h^{-1}) and I_i is the intracellular IPTG concentration (g·L^{-1}).

7.3.2.4.   The inducer.

In accordance to some of the studies found in the literature [17, 39, 46], preliminary modelling of our process showed that production could not be only described in terms of the amount of the IPTG added into the culture, so intracellular inducer concentration had to be considered. Results

from chapter 5 also suggested that inducer transport behaved differently in substrate-limited cultivations than in the case of glucose in excess [33].

It has been described that IPTG can enter the cell by active transport and/or by diffusion, depending on the case [47]. Active transport takes place through specific permeases (the *lacY* gene product), which are subject to carbon catabolite repression when glucose is in excess, reducing the AMP and CRP levels [48, 49]. When glucose is in excess, IPTG transport into the cell is expected to be mainly by diffusion because permease expression should be repressed under those conditions [50, 51], but active transport had to be considered when glucose levels were low.

Taking into account both mechanisms, the mass balances for IPTG became:

$$\frac{dI}{dt} = \frac{F_I}{V} \cdot \left(I_F - I\right) - \frac{F_S}{V} \cdot I - AT \cdot X_1 - DIF \cdot X \qquad \text{(Eq. 7.18)}$$

$$\frac{dI_i}{dt} = AT \cdot \rho \cdot \frac{X_1}{X} + DIF \cdot \rho - \mu_1 \cdot I_i \qquad \text{(Eq. 7.19)}$$

where I_F represented the IPTG concentration in the inducer feed (g·L^{-1}), I accounted for the inducer concentration in the medium (g·L^{-1}), AT described IPTG uptake by active biomass using active transport (g·gDCW^{-1}·h^{-1}), DIF accounts for IPTG diffusion into all cells (g·gDCW^{-1}·h^{-1}) and ρ is the dry cell density (g·L^{-1}). The last term of Eq.7.19 explained the dilution of intracellular IPTG due to cell growth.

ρ was fixed to 450 gDCW·L^{-1} in other works [52], but experimental flow cytometry measurements indicated that, in our case, cell size (measured as FSC) progressively increased during the induction phase, shifting significantly when the cell population segregated into a viable but not culturable state (Figure 7.15). Thus, the reasons for this increase of cell size should be understood and included into the model as well.

The active transport was considered to depend on permease uptake kinetics (AT), and the number of permeases per gram of biomass (g) subject to induction and to catabolite repression:

$$AT = \left(\frac{g \cdot I}{h + I}\right) \qquad \text{(Eq. 7.20)}$$

$$g = i + \frac{j \cdot I_i}{k + I_i} \cdot \frac{n}{n + S} \qquad \text{(Eq. 7.21)}$$

where g accounts for permease production. Basal permease activity is described by i, and the last term describes carbon catabolite repression, being n the parameter to determine.
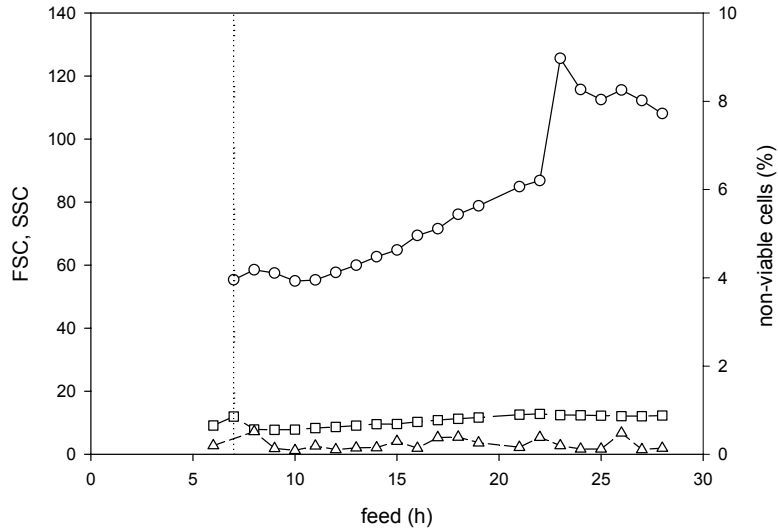

Figure 7.15. Forward light scatter (FSC, −○−), side scatter (SSC, −□−) and percentage of non-viable cells (−△−) along the induction phase in the substrate limited fed-batch culture with continuous induction presented in Figure 7.5. The vertical dotted line depicts when the inducer feed was switched on.

Diffusion was modelled using Fick's Law:

$$DIF = m \cdot (I - I_i) \qquad \text{(Eq. 7.22)}$$

The model to describe recombinant protein production would be described by Eq. 7.8 - Eq.7.22 plus Eq. 7.23:

$$\frac{dV}{dt} = F_S + F_I \qquad \text{(Eq. 7.23)}$$

The proposed model for substrate-limited cultures induced at 37ºC (Box 7.1) was not successfully calibrated because of the limited knowledge of cell behaviour under certain circumstances. The proposed model was able to describe each one of the experiments independently (example

presented in Figure 7.16), but there was not a single set of parameters which could describe them all. However, this preliminary work on modelling allowed identifying some important features of RhuA production which should be taken into account.
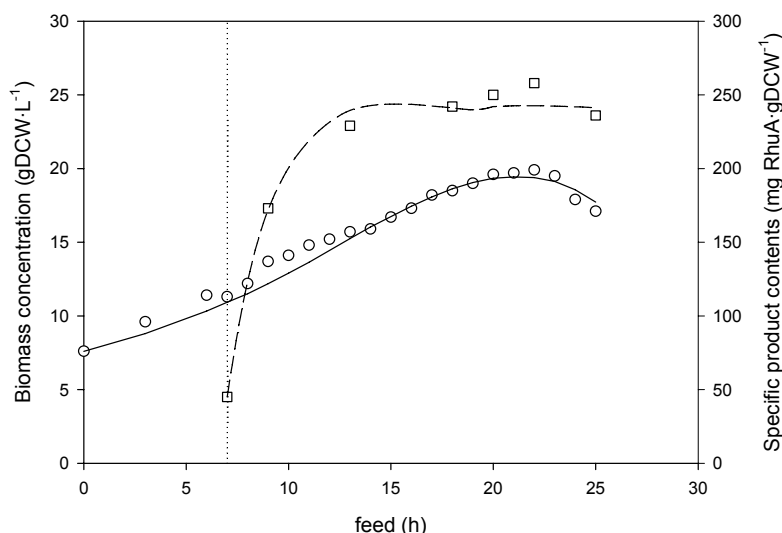

Figure 7.16. Experimental values for biomass concentration (○) and specific product contents (□) in a fed-batch culture performed at specific growth rate of 0.1 h^{-1}. The continuous line depict the model estimation for the biomass concentration (−) while dashed lines represent the model estimation for glucose concentration (− −). The vertical dotted line depicts when the inducer feed was switched on.

Since RhuA is used as biocatalyst, the aim of the model should be to optimise activity production/productivity. Thus, the product should be modelled in terms of activity instead of amount of protein, but protein degradation and deactivation should be first fully understood.

On the other hand, some other strategies like inducing at higher substrate levels (see next chapter 8) or inducing at lower temperatures (results from chapter 6) showed good results when trying to avoid protein degradation and deactivation. Thus, since the process performance is improved by those alternative strategies, future modelling work should include the corresponding additional variable(s).

$$X = X_1 + X_2$$

$$\frac{dX_1}{dt} = (\mu_1 - \alpha) \cdot X_1 - \left(\frac{F_S + F_I}{V} \right) \cdot X_1$$

$$\frac{dX_2}{dt} = \alpha \cdot X_1 - \left(\frac{F_S + F_I}{V} \right) \cdot X_2$$

$$\mu_1 = \frac{0.59 \cdot S}{0.002 + S} \cdot e^{\frac{-S}{64.4}}$$

$$\frac{dS}{dt} = \frac{F_S}{V} \cdot (S_F - S) - \frac{F_I}{V} \cdot S - q_S \cdot X_1 - m_2 \cdot X_2$$

$$q_S = \frac{1}{0.34} \cdot \mu_S + 0.10$$

$$\alpha = a \cdot \left(1 - \frac{1}{1 + e^{\left(\frac{P}{X} - \frac{P}{X} \big|_{crit} \right)}} \right) + b \cdot q_P \cdot \frac{c}{c + S}$$

$$\frac{dP}{dt} = q_{P1} \cdot X_1 - \left(\frac{F_S + F_I}{V} \right) \cdot P$$

$$q_{P1} = \left(q_{P0} + \frac{q_{P\max} \cdot I_i}{K_I + I_i} \right) \cdot \frac{\mu_1}{\mu_{\max}}$$

$$\frac{dI}{dt} = \frac{F_I}{V} \cdot (I_F - I) - \frac{F_S}{V} \cdot I - AT \cdot X_1 - DIF \cdot X$$

$$\frac{dI_i}{dt} = AT \cdot \rho \cdot \frac{X_1}{X} + DIF \cdot \rho - \mu_1 \cdot I_i$$

$$AT = \left(\frac{g \cdot I}{h + I} \right) \qquad g = i + \frac{j \cdot I_i}{k + I_i} \cdot \frac{n}{n + S}$$

$$DIF = m \cdot (I - I_i)$$

$$\frac{dV}{dt} = F_S + F_I$$

Box 7.1. Proposed model equations for RhuA production in *E. coli* M15 ΔglyA [pREP4] pQEαβrham substrate-limited fed-batch cultures induced with IPTG at 37ºC.

7.4. Conclusions.

An unstructured model which describes the growth at 37ºC of *E. coli* M15 ΔglyA [pREP4] pQE$\alpha\beta$rham was defined, calibrated and validated using different high cell density cultures simultaneously. The ability of the model to describe cell growth under different situations (including substrate inhibition conditions) demonstrated its robustness.

This model was further developed to describe recombinant protein production in substrate-limited fed-batch cultures induced by IPTG addition. The proposed model was not able to describe all possible cultivation conditions and, thus, it couldn't be used for process optimisation. However, important information was gathered during the modelling process which could be used to increase the process performance:

- Cell segregation into a viable-but-not culturable state was observed after high levels of intracellular RhuA were achieved. This undesired effect was translated into growth arrest which didn't allow continuing the process. It has been shown that the auxotrophy marker (SHMT) expression contributes to this fact, although no term was included into the model.

- Activity losses as well as protein degradation were observed, negatively affecting the final bioactive RhuA yields.

- The continuous inducer feed used did not allow to control induction properly, since IPTG transport into the cytoplasm was found to be a complex process with a strong impact on process performance.

Based on these learnings, it woudn't be recommendable to continue the modelling task before studying these aspects in depth, and specific experiments should be planned and performed to obtain the necessary data. This could help not only to improve the process description, but also to avoid undesired situations.

References.

[1] B.H.L.Betlem, P.Mulder, B.Roffel, Optimal mode of operation for biomass production, Chemical Engineering Science 57 (2002) 2799-2809.

[2] V.Galvanauskas, R.Simutis, A.Lubbert, Model-based design of biochemical processes: simulation studies and experimental tests, Biotechnology Letters 19 (1997) 1043-1047.

[3] A.Lubbert, S.B.Jorgensen, Bioreactor performance: a more scientific approach for practice, Journal of Biotechnology 85 (2001) 187-212.

[4] J.M.Modak, H.C.Lim, Optimal-Mode of Operation of Bioreactor for Fermentation Processes, Chemical Engineering Science 47 (1992) 3869-3884.

[5] A.K.Gombert, J.Nielsen, Mathematical modelling of metabolism, Current Opinion in Biotechnology 11 (2000) 180-186.

[6] M.Jenzsch, R.Simutis, A.Lubbert, Optimization and control of industrial microbial cultivation processes, Engineering in Life Sciences 6 (2006) 117-124.

[7] S.T.Browning, M.L.Shuler, Towards the development of a minimal cell model by generalization of a model of *Escherichia coli*: Use of dimensionless rate parameters, Biotechnology and Bioengineering 76 (2001) 187-192.

[8] A.Kremling, et al., The organization of metabolic reaction networks. III. Application for diauxic growth on glucose and lactose, Metabolic Engineering 3 (2001) 362-379.

[9] W.E.Bentley, D.S.Kompala, A Novel Structured Kinetic Modeling Approach for the Analysis of Plasmid Instability in Recombinant Bacterial Cultures, Biotechnology and Bioengineering 33 (1989) 49-61.

[10] J.Nielsen, A.G.Pedersen, K.Strudsholm, J.Villadsen, Modeling Fermentations with Recombinant Microorganisms - Formulation of A Structured Model, Biotechnology and Bioengineering 37 (1991) 802-808.

[11] K.Schugerl, Progress in monitoring, modeling and control of bioprocesses during the last 20 years, Journal of Biotechnology 85 (2001) 149-173.

[12] A.R.Joyce, B.O.Palsson, The model organism as a system: integrating 'omics' data sets, Nature Reviews Molecular Cell Biology 7 (2006) 198-210.

[13] R.Portner, T.Schafer, Modelling hybridoma cell growth and metabolism--a comparison of selected models and data, Journal of Biotechnology 49 (1996) 119-135.

[14] H.J.Chae, et al., Framework for online optimization of recombinant protein expression in high-cell-density *Escherichia coli* cultures using GFP-fusion monitoring, Biotechnology and Bioengineering 69 (2000) 275-285.

[15] J.Lee, W.F.Ramirez, Mathematical-Modeling of Induced Foreign Protein-Production by Recombinant Bacteria, Biotechnology and Bioengineering 39 (1992) 635-646.

[16] D.Levisauskas, V.Galvanauskas, R.Simutis, A.Lubbert, Model based calculation of substrate/inducer feed-rate profiles in fed-batch processes for recombinant protein production, Biotechnology Techniques 13 (1999) 37-42.

[17] D.M.Ramirez, W.E.Bentley, Characterization of stress and protein turnover from protein overexpression in fed-batch *E. coli* cultures, Journal of Biotechnology 71 (1999) 39-58.

[18] D.C.Psichogios, L.H.Ungar, A Hybrid Neural Network-1st Principles Approach to Process Modeling, Aiche Journal 38 (1992) 1499-1511.

[19] R.Simutis, A.Lubbert, Exploratory analysis of bioprocesses using artificial neural network-based methods, Biotechnology Progress 13 (1997) 479-487.

[20] B.Zelic, N.Bolf, D.Vasic-Racki, Modeling of the pyruvate production with *Escherichia coli*: comparison of mechanistic and neural networks-based models, Bioprocess and Biosystems Engineering 29 (2006) 39-47.

[21] V.Galvanauskas, R.Simutis, N.Volk, A.Lubbert, Model based design of a biochemical cultivation process, Bioprocess Engineering 18 (1998) 227-234.

[22] M.J.Guardia, E.G.Calvo, Modeling of *Escherichia coli* growth and acetate formation under different operational conditions, Enzyme and Microbial Technology 29 (2001) 449-455.

[23] U.Lendenmann, T.Egli, Kinetic models for the growth of *Escherichia coli* with mixtures of sugars under carbon-limited conditions, Biotechnology and Bioengineering 59 (1998) 99-107.

[24] K.V.Venkatesh, P.Doshi, R.Rengaswamy, An optimal strategy to model microbial growth in a multiple substrate environment, Biotechnology and Bioengineering 56 (1997) 635-644.

[25] B.Xu, M.Jahic, S.O.Enfors, Modeling of overflow metabolism in batch and fed-batch cultures of *Escherichia coli*, Biotechnology Progress 15 (1999) 81-90.

[26] E.B.Jensen, S.Carlsen, Production of Recombinant Human Growth-Hormone in *Escherichia coli* - Expression of Different Precursors and Physiological-Effects of Glucose, Acetate, and Salts, Biotechnology and Bioengineering 36 (1990) 1-11.

[27] R.A.Majewski, M.M.Domach, Simple Constrained-Optimization View of Acetate Overflow in *Escherichia coli*, Biotechnology and Bioengineering 35 (1990) 732-738.

[28]   A.J.Wolfe, The acetate switch, Microbiology and Molecular Biology Reviews 69 (2005) 12-+.

[29]   W.R.Farmer, J.C.Liao, Reduction of aerobic acetate production by *Escherichia coli*, Applied and Environmental Microbiology 63 (1997) 3205-3210.

[30]   G.L.Kleman, J.J.Chalmers, G.W.Luli, W.R.Strohl, A Predictive and Feedback-Control Algorithm Maintains A Constant Glucose-Concentration in Fed-Batch Fermentations, Applied and Environmental Microbiology 57 (1991) 910-917.

[31]   K.Konstantinov, M.Kishimoto, T.Seki, T.Yoshida, A Balanced DO-Stat and Its Application to the Control of Acetic-Acid Excretion by Recombinant *Escherichia coli*, Biotechnology and Bioengineering 36 (1990) 750-758.

[32]   T.Paalme, K.Tiisma, A.Kahru, K.Vanatalu, R.Vilu, Glucose-Limited Fed-Batch Cultivation of *Escherichia coli* with Computer-Controlled Fixed Growth-Rate, Biotechnology and Bioengineering 35 (1990) 312-319.

[33]   J.Pinsach, C.de Mas, J.Lopez-Santin, Induction strategies in fed-batch cultures for recombinant protein production in *Escherichia coli*: Application to rhamnulose 1-phosphate aldolase, Biochemical Engineering Journal 41 (2008) 181-187.

[34]   L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

[35]   S.Aiba, M.Nagatani, Kinetics of Product Inhibition in Alcohol Fermentation, Biotechnology and Bioengineering 10 (1968) 845-&.

[36]   V.H.Edwards, Influence of High Substrate Concentrations on Microbial Kinetics, Biotechnology and Bioengineering 12 (1970) 679-&.

[37]   L.Vidal, et al., High-level production of recombinant His-tagged rhamnulose 1-phosphate aldolase in *Escherichia coli*, Journal of Chemical Technology and Biotechnology 78 (2003) 1171-1179.

[38]   L.Andersson, L.Strandberg, S.O.Enfors, Cell segregation and lysis have profound effects on the growth of *Escherichia coli* in high cell density fed batch cultures, Biotechnology Progress 12 (1996) 190-195.

[39]   Z.Y.Zheng, S.J.Yao, D.Q.Lin, Using a kinetic model that considers cell segregation to optimize hEGF expression in fed-batch cultures of recombinant *Escherichia coli*, Bioprocess and Biosystems Engineering 27 (2005) 143-152.

[40]   W.E.Bentley, N.Mirjalili, D.C.Andersen, R.H.Davis, D.S.Kompala, Plasmid-Encoded Protein - the Principal Factor in the Metabolic Burden Associated with Recombinant Bacteria, Biotechnology and Bioengineering 35 (1990) 668-681.

[41] R.Grabherr, E.Nilsson, G.Striedner, K.Bayer, Stabilizing plasmid copy number to improve recombinant protein production, Biotechnology and Bioengineering 77 (2002) 142-147.

[42] A.Teich, H.Y.Lin, L.Andersson, S.Meyer, P.Neubauer, Amplification of ColE1 related plasmids in recombinant cultures of *Escherichia coli* after IPTG induction, Journal of Biotechnology 64 (1998) 197-210.

[43] B.R.Glick, Metabolic Load and Heterologous Gene-Expression, Biotechnology Advances 13 (1995) 247-261.

[44] M.Cserjan-Puschmann, W.Kramer, E.Duerrschmid, G.Striedner, K.Bayer, Metabolic approaches for the optimisation of recombinant fermentation processes, Applied Microbiology and Biotechnology 53 (1999) 43-50.

[45] M.J.Kosinski, U.Rinas, J.E.Bailey, Isopropyl-β-D-thiogalactopyranoside Influences the Metabolism of *Escherichia coli*, Applied Microbiology and Biotechnology 36 (1992) 782-784.

[46] J.M.G.Vilar, C.C.Guet, S.Leibler, Modeling network dynamics: the *lac* operon, a case study, Journal of Cell Biology 161 (2003) 471-476.

[47] J.Beckwith, The lactose operon, in: F.C.Neidhart (Ed.), *Escherichia coli* and *Salmonella typhimurium*, American Society of Microbiology, Washington D.C. (1987) 1444-1452.

[48] B.M.Hogema, et al., Inducer exclusion by glucose 6-phosphate in *Escherichia coli*, Molecular Microbiology 28 (1998) 755-765.

[49] H.Ishizuka, A.Hanamura, T.Kunimura, H.Aiba, A lowered concentration of cAMP receptor protein caused by glucose is an important determinant for catabolite repression in *Escherichia coli*, Molecular Microbiology 10 (1993) 341-350.

[50] P.R.Jensen, K.Hammer, Artificial promoters for metabolic optimization, Biotechnology and Bioengineering 58 (1998) 191-195.

[51] A.M.Sanden, M.Bostrom, K.Markland, G.Larsson, Solubility and proteolysis of the Zb-MaIE and Zb-MaIE31 proteins during overproduction in *Escherichia coli*, Biotechnology and Bioengineering 90 (2005) 239-247.

[52] J.Hiller, E.Franco-Lara, D.Weuster-Botz, Metabolic profiling of *Escherichia coli* cultivations: evaluation of extraction and metabolite analysis procedures, Biotechnology Letters 29 (2007) 1169-1178.

8. PROCESS STRATEGIES IN *E. coli* IMPROVING PROTEIN QUALITY AND DOWNSTREAM YIELDS.

Abstract.

Overexpression of recombinant proteins under the control of strong promoters in *Escherichia coli* may burden the host cell before maximum product accumulation occurs. This can be translated into reduced product yields and/or reduced product quality. When expressing recombinant rhamnulose 1-phosphate aldolase (RhuA) under the control of a strong T5 promoter in *Escherichia coli* with a hexa histidine-tag (6xHis-tag) at its N-terminus, reduced product yields and product quality were observed when induction was applied to standard substrate-limited cultures. Alternative operational strategies to minimise these deleterious effects were implemented in fed-batch cultures, having an important impact on both RhuA yields and downstream processing when preparing the enzyme for its use as biocatalyst.

First, a continuous inducer feed was implemented in substrate limited fed-batch cultures to tune RhuA overexpression. By doing so, reduced product synthesis rates avoided sudden metabolic burden and the production phase was successfully extended. The final specific RhuA levels were quite high, but the final specific enzyme activity (1.7 AU·mg RhuA^{-1}) was considerably lower than expected. Moreover, only 55% of immobilization yield was achieved when immobilised metal affinity chromatography (IMAC) was used to purify and immobilise RhuA from cellular lysate in a single step. Western blot analyses showed a decrease of the 6xHis-tag in recombinant RhuA along induction phase and, at the end of the culture, only 20% of the initial 6xHis-tag·mg RhuA^{-1} was detected.

Two alternative strategies were implemented to minimise these process bottlenecks:

i)   Temperature reduction to 28ºC in substrate limited fed-batch cultures alleviated the metabolic load on host cells probably due to lower production rates and improved protein folding. It also led to high RhuA contents within the cells, the enzyme activity increased to 4 AU·mg RhuA^{-1} and purification-immobilization yield also increased to 93%.

ii)   A novel fed-batch operational procedure working at high glucose concentration was implemented. The aim was to alleviate starvation-induced stress by using an excess of glucose. Acetate accumulation to toxic levels was avoided by inhibiting the specific growth rate at high substrate concentrations. RhuA overexpression was improved under these conditions. Comparable aldolase levels to those obtained using the previous strategies were reached, but the final enzyme activity increased to 4.8 AU·mg RhuA^{-1} and 95% of immobilization yield was achieved.

In both cases, western blot analyses showed that 80-100% of initial 6xHis-tag·mg RhuA^{-1} was present at the end of the culture, confirming that both alternative process strategies improved the recombinant protein quality.

8.1. Introduction and objectives.

As explained in previous chapters, the *rhaD* gene coding for rhamnulose 1-phosphate aldolase (RhuA) had been previously cloned in *E. coli* M15 ΔglyA [pREP4] pQErham and overexpressed as a fusion protein to a hexa-hisitidine tag [1]. Thus, one-step immobilization and purification of clear lysate on IMAC (immobilised metal affinity chromatography) support [2] could be used after fermentation broth processing. This technology has not only the advantage of reduction in enzyme loss and time processing during the purification stages, but also allows to obtain an immobilised enzymatic derivative suitable for application in synthetic processes in a single step.

The production strategy of recombinant rhamnulose 1-phosphate aldolase in *E. coli* was already developed in previous chapters. In order to get a cost-effective process, feeding strategies based on substrate limitation were implemented to obtain high biomass concentrations. However, a common constraint was the metabolic burden caused by strong overexpression of the recombinant protein which quickly led to growth arrest, particularly when pulse induction was performed. The development of continuous inducer feeding strategies (see chapter 5) successfully extended the production phase and reduced the required amount of IPTG, but overexpression of the target protein was found to trigger other stress responses not detected as cell death: even though final specific RhuA levels were high, specific enzyme activities were lower than expected [3]. Moreover, low immobilisation yields were achieved when IMAC was used to purify and immobilise RhuA from cellular lysate in a single step, with the concomitant loss in immobilised biocatalyst yields.

Limitations still exist to produce recombinant proteins at high levels in a biological active form, and proteolysis has been reported to impact these processes [4-6]. Little work has been done on the control of proteolysis in high cell density fed-batch cultures [7] and although affinity tags have been reported to have a positive impact on yield, solubility and even folding and reduction of proteolysis [8], partial proteolysis has been identified as the cause for reduced IMAC recovery yields [9]. Partial intracellular denaturation of proteins driven by high temperature may lead to a loss in enzyme activity, but proteolysis [10] and degradation of proteins is also increased in cells under starvation conditions [11]. Therefore, proteolysis may be induced by starvation and/or high temperatures.

Several techniques like designing a more stable protein sequence or modification of host cell have been successfully applied to control proteolysis [7], but their effects on the specific activity of the recombinant protein are not known beforehand. As a consequence, we have focused on the study of different cultivation strategies to overcome this problem.

As described in chapter 6, decreasing the process temperature from 37ºC to 28ºC improved protein quality in substrate-limited fed-batch cultures with continuous inducer addition [12]. Therefore, the analysis of temperature reduction on protein quality and downstream yields was one of the aims of this chapter.

Results from chapter 5 suggested that starvation-induced stress due to limited carbon source [13] compromised cell growth yielding lower specific activities when compared to unlimited cultures. Thus, it would be desirable to induce high cell density cultures in which cells had enough resources available, *i.e.*, an excess of glucose. Under those circumstances, carbon catabolite repression takes place [14, 15], and IPTG transport into the cell is expected to be mainly by diffusion because permease (*lacY* gene products) expression should be repressed [16, 17]. For the same reason, the expression of the recombinant aldolase should take place at a lower level than in absence of glucose [18]. If IPTG transport and protein expression rates were compatible with growth in the presence of glucose in excess, a novel strategy for recombinant protein production could be developed. Although it would imply an increase of fermentation costs due to the presence of residual glucose at the end of the process, any possible increase in productivity and downstream yields (even moderate) would justify its use because the target protein is a high-added value product.

When having a look at production and downstream processing as a whole, cost breakdown showed that downstream processing can hamper the economical viability of some recombinant protein production process [19]. Thus, the development of integrated production and downstream process strategies becomes of major importance.

In present chapter, the linkage between recombinant 6xHis-tagged RhuA production strategy and downstream process by one-step purification and immobilization was investigated in order to improve the yield and quality of the final product.

This work was done in collaboration with the PhD student Jordi Ruiz, who was working on downstream processes at that time.

8.2. Materials and methods.

Bacterial strain and plasmids.

The K-12 derived strain *E. coli* M15ΔglyA [pREP4] harbouring the vector pQEαβrham was used for rhamnulose 1-phosphate aldolase overexpression (see section *3.1. Strains and vectors* for further details). Frozen stock aliquots containing glycerol prepared from exponential phase cultures grown in Luria-Bertani media (LB) were stored at -80ºC.

Media composition.

LB medium was used for preinoculum preparation. For inoculum and bioreactor experiments, a defined mineral medium utilising glucose as the sole carbon source was used. Their composition is described in *3.2. Media composition*.

Cultivation conditions.

Preinocula, inocula and bioreactor experiments were carried out as described in section *3.4. Cultivation conditions*, except for the substrate inhibited cultures. In this latter case, feed was pulsed to reach 60 g·L^{-1} of glucose and the exponential feeding strategy was continued in order to maintain an almost constant concentration inside the bioreactor and the same specific growth rate. Induction was started 1 hour later.

In all cases, IPTG was continuously fed into the bioreactor to tune the transcription rate below values which burdened the host cell metabolism (see section *5.2. Materials and methods*). The continuous inducer feed was designed to progressively increase the IPTG concentration maintaining a relation between 0.4 - 0.6 μmol IPTG. gDCW^{-1} [3].

Broth processing.

The fermentation broth was centrifuged at 10000 rpm for 20 minutes at 4ºC using a Beckman J2-21M/E centrifuge. Harvested cells were resuspended in lysis buffer: 43 mM Na_2HPO_4, 7 mM NaH_2PO_4, 20 mM Imidazol, 300 mM NaCl (pH= 8) at a ratio of 1mL buffer: 0.3 g harvested cells.

Resuspended cells were lysed by one-shot high pressure disruption (Constant Systems LTD One Shot) at 2.57 kbar and at constant temperature of 4ºC.

The crude cell lysate was centrifuged at 14000 rpm for 35 minutes at 4ºC and cell debris was rejected. Enzymatic activity of the supernatant was measured and sodium azide was added to keep a concentration of 0.02 % (w/w) to avoid biological degradation of the clear lysate.

RhuA purification-immobilization.

10 mL aliquots of clear lysate at the appropriate activity concentration were employed for one-step purification-immobilization on Co-IDA support (Chelating Sepharose FF Amersham Biosciencies-GE Healthcare with Co^{2+} chelated on it). 1 mL sample was used as reference, keeping it under mild horizontal agitation at 4ºC. Its activity was measured both at the beginning and end of the immobilization process. The rest of the sample (9 mL) was added to 1 mL of Co-IDA support, and the residual activity of the suspension and supernatant was measured until the adsorption-desorption equilibrium was reached.

Analytical methods.

Growth was followed by optical density at 600 nm (OD600) and by dry cell weight (DCW) measurements as described in section *3.5. Analytical methods*.

Glucose, organic acids, ammonium and phosphate were also measured as described in section *3.5. Analytical methods*.

Product concentration during the cultures was determined as total protein content, percentage of RhuA amongst the rest of intracellular soluble proteins and RhuA activity. All three methods are described in section *3.5. Analytical methods*.

Western blots were performed after SDS-PAGE proteins were transferred to a nitrocellulose membrane. 6xHis-tag proteins were detected with mouse anti-6xHis antibodies (Roche); anti-mouse IgG alkaline phosphatase conjugate (Sigma) was used as a secondary antibody. Detection was performed with the Alkaline Phosphatase Conjugate Substrate Kit (BioRad).

8.3. Results and discussion.

8.3.1.   Substrate-limited fed-batch cultures.

Several glucose-limited fed-batch processes for RhuA production were performed at 37ºC. The cultures were operated in a similar way to those already presented in chapters 5 and 6, using an exponential substrate feed and a continuous inducer feed to progressively increase the inducer concentration up to 15 μM [3].

The behaviour of the cultures was in accordance to previous results: even though the continuous inducer feed allowed an extension of the induction phase, cell growth was finally arrested. The contents of recombinant RhuA at that point was quite high, but the specific enzyme activity was below 2 AU·mg RhuA^{-1} (see Figure 6.1 as an example).

After processing the fermentation broth, the intracellular lysate was submitted to stirred tank adsorption on Co-IDA resin in order to purify and immobilize 6xHis-tag RhuA in a single step [2]. Total immobilisation yields between 55-65% were achieved (Figure 8.1). These results suggested that part of RhuA had lost the histidine tag and was not able to adsorb on the IMAC support.
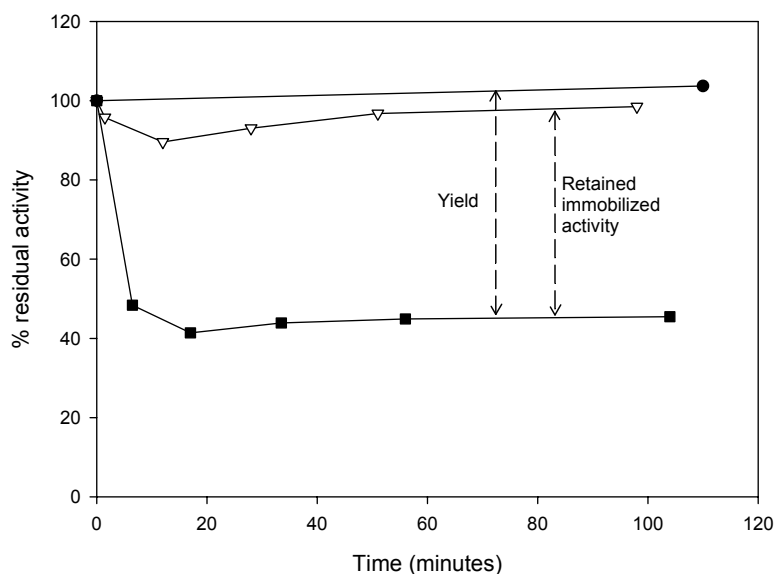

Figure 8.1. Time evolution of one step purification-immobilization of RhuA (from substrate-limited fermentation conducted at 37ºC) on Co-IDA offering 5 AU·mL^{-1} of support. The immobilisation process was performed at 4ºc and pH=8. The percentage of residual activity was measured for the reference (●), for the suspension (▽) and for the supernatant (■).

Thus, western blot analyses were performed to check for the 6xHis-tag in RhuA along the induction phase. In Figure 8.2, western blot data are presented as fraction of initial specific intensity (calculated for the first sample taken after induction had started) along time.
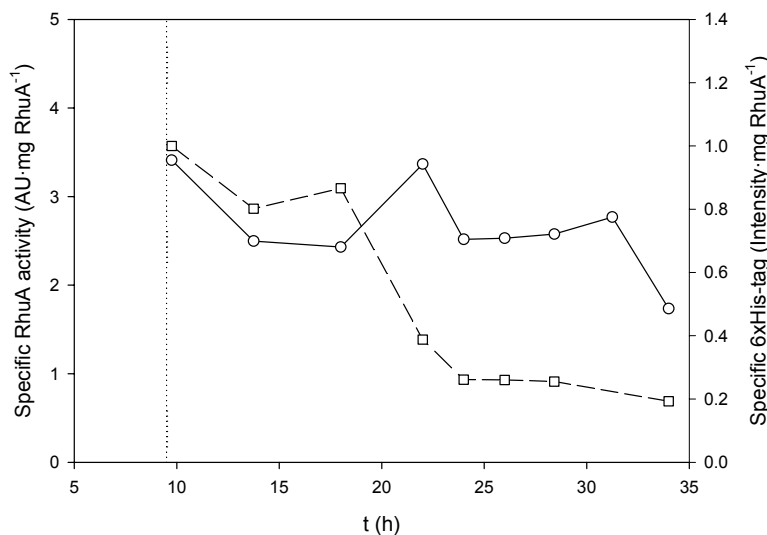

Figure 8.2. Specific RhuA activity (○) and specific 6xHis-tag contents in recombinant RhuA (□) along the induction phase in a substrate-limited fed-batch culture at 37ºC. The vertical dotted line shows when continuous IPTG supply was started.

As can be seen, the specific content of 6xHis-tag in RhuA decreased along the production phase and, at the end of the culture, only 20% of the initial 6xHis-tag·mg RhuA^{-1} was detected. This fact could be attributed to a partial proteolysis affecting the recombinant protein tag. The results are compatible with the obtained immobilization yields of 55% as the western blot is specific for 6 histidine residues per tag. A fraction of the partially proteolysed protein could keep enough histidine residues attach to the column and account for the remaining 35% yield.

In fact, this hypothetical partial proteolysis did not only affect the affinity tag, but also the biological activity of the protein. The specific activity of the recombinant enzyme (also shown in Figure 8.2) was never above 3.4 AU·mg RhuA^{-1}, and as low as 1.7 AU·mg RhuA^{-1} at the end of the culture, evidencing the decrease of the specific enzyme activity already observed in previous cultures (see Figure 6.1). Proteolysis could probably be switched on because of the stress caused by the recombinant RhuA synthesis combined to the starvation experienced by cells because of substrate limitation [4, 5, 20].

To minimise the deleterious effects on protein quality, i.e., activity units per mg of RhuA and integrity of the hexahistidine tag, we focused on the following alternative strategies:

a) Lowering process temperature, in order to improve protein folding and reduce the activity of proteases (production strategy already developed in chapter 6).

b) Substrate-in-excess fed-batch cultures to minimise the starvation-induced stress.

8.3.2.   Substrate-limited fed-batch cultures at 28ºC.

Reduction of the production process temperature was considered to improve RhuA quality because it has been reported that the conformational quality and functionality of protein increase in parallel to reduced culture growth temperature [21]. Moreover, protein degradation rate should decrease with decreased temperature in accordance to the Arrhenius' Law, and protein misfolding often occurs at high temperatures due to either heat damage or due to an over increase of protein synthesis rate [7]. From results presented in chapter 6, we knew that by decreasing the process temperature protein synthesis rates were reduced (probably due to lower gene dosage), and this helped to improve the specific activity of RhuA.

Therefore, an additional culture was performed reducing the fed-bath phase temperature from 37ºC to 28ºC, applying the same inducer dosage strategy as in the previous experiment (data not shown; cultures performed at 28ºC have been already presented in chapter 6, see Figure 6.4). In this case, the specific RhuA levels increased along the whole induction phase to similar levels to those obtained at 37ºC, but the enzyme activity at the end of the culture was increased 2-fold (above 3.5 AU·mg RhuA^{-1}, Figure 8.3).

The fermentation broth was submitted to the same downstream processing than the one performed at 37ºC and one-step purification-immobilisation using intracellular lysate was carried out. The obtained activity immobilisation yield was around 93%, allowing an almost complete recovery of the produced aldolase in active immobilized form. Western blot results, presented in Figure 8.3, agreed with the immobilization data. As can be seen, the specific content of 6xHis-tag in RhuA is nearly constant along the induction phase, indicating that all the obtained active aldolase contained enough histidine molecules to be linked to the Co-IDA resin.
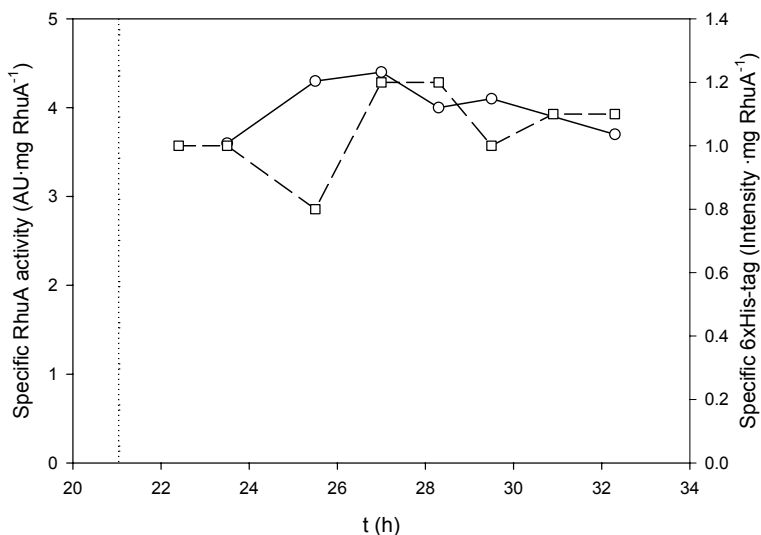
Figure 8.3. Specific RhuA activity (○) and specific 6xHis-tag contents in recombinant RhuA (□) along the induction phase in a substrate-limited fed-batch culture at 28ºC. The vertical dotted line shows when continuous IPTG supply was started.

The above data are confirming the suitability of temperature reduction to improve recombinant protein quality, significantly increasing the downstream yields.

8.3.3.   Substrate in excess fed-batch cultures.

As described in chapter 5, IPTG addition triggered induction in cells which were already short of resources and stressed due to limited carbon source [22], compromising cell growth and yielding lower specific protein levels in glucose-limited cultures than in the case of unlimited cultures [3]. Thus, it would be desirable to overexpress RhuA in unlimited high cell density cultures, but acetate accumulation might inhibit growth and protein production under these circumstances.

Shake flask cultures performed to characterize cell growth in defined medium at 37ºC were presented in chapter 7. Those results indicated that the critical specific growth rate for acetate accumulation was close to 0.25 h^{-1}, and that the specific growth rate was below that value when the substrate concentration in the media was above 60 $g \cdot L^{-1}$ (Figure 7.1). Therefore, the control of the specific growth rate below 0.25 h^{-1} to avoid acetate accumulation could be achieved when keeping the substrate concentration at inhibiting levels and, as a consequence, this strategy could be potentially used to obtain high cell density cultures.

Owing to these results, an alternative operational strategy to minimise the deleterious effects of substrate limitation on recombinant RhuA was implemented in fed-batch cultures by using an excess of glucose. The challenge was to control cell growth and acetate accumulation under substrate in excess conditions while induction was performed.

Since this was a new operational strategy, operation mode during both growth and induction phases were studied at 37ºC. Preliminary experiments were performed to develop the induction strategy before continuous inducer addition could be implemented to high cell density cultures (data not shown).

From these preliminary results, a process strategy could be defined for both growth and induction phases at 37ºC. Figure 8.4 shows that growth started as usual, using a predefined exponential feed addition profile. As described in section *8.2. Materials and methods*, growth was inhibited by pulsing feed to reach 60 g glucose·L^{-1} one hour before IPTG feed was switched on. After the glucose pulse, the specific growth rate was kept at low enough values to avoid acetate accumulation and the IPTG concentration was increased from 10 to 25 μM during the 5.5 hours induction phase. The IPTG concentration range was determined from previous experiments at glucose-inhibiting conditions (data not shown), which showed that 10 μM IPTG was necessary to trigger RhuA expression while bacterial growth was affected above 25 μM. At the end of the culture, the biomass concentration was above 40 gDCW·L^{-1}. The biomass build-up stopped when the specific RhuA levels reached 233 mg·$gDCW^{-1}$ (corresponding to 39% of the total soluble protein). Figure 8.4 also shows that the specific activity increased similarly to the specific content of RhuA reaching 1130 AU·$gDCW^{-1}$ at the end of the culture. In contrast to the case of substrate limitation, no specific activity decrease was observed when cell growth stopped.

Using this novel strategy, the specific activity production rate could be maintained around 200 AU·$gDCW^{-1}$·h^{-1} for nearly 5 hours, confirming that both growth and production are compatible under glucose in excess conditions. Continuous addition of inducer did not improve specific activity when compared to the preliminary pulse-induced experiments performed under substrate inhibiting conditions, but the amount of IPTG added to the culture could be considerably reduced. Substrate limited cultures induced in a similar way (I/X between 0.3-0.6 μmol·$gDCW^{-1}$) led to slightly lower specific activities of 1000 AU·$gDCW^{-1}$ (results from chapter 5). Nevertheless, in that case, lower inducer concentrations (4 to 15 μM) were required. It has to be taken into account that inducer uptake is supposed to be different depending on the substrate concentration. In substrate in excess cultures, the inducer was not uptaken as effectively as in the case of glucose limitation, and

probably a lower fraction of the IPTG added was transported into the cell. Carbon catabolite repression might have influenced inducer uptake when glucose levels were high, and IPTG diffusion was probably the limiting step for RhuA overexpression. In contrast, at limiting glucose levels, IPTG was also transported by permeases, which enhanced its transport into the cytoplasm [17, 23].
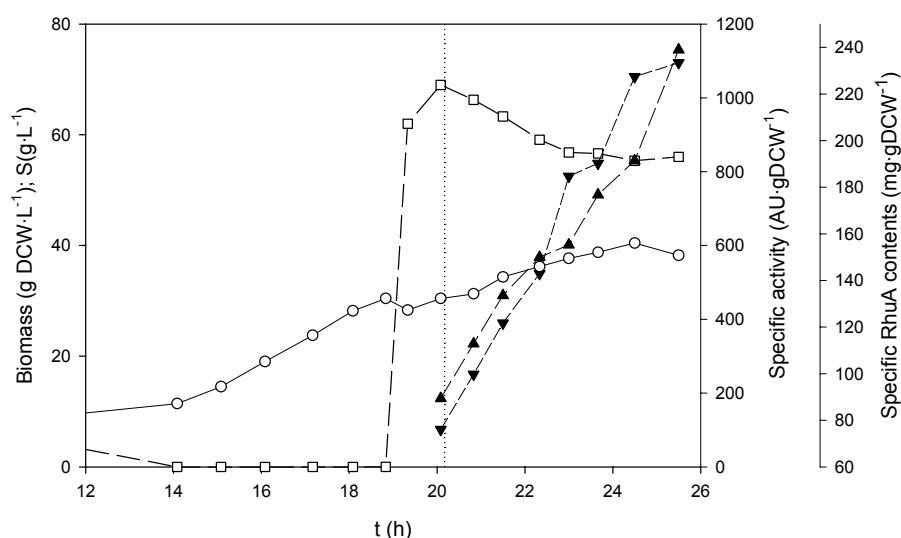

Figure 8.4. Biomass (○), substrate (□), specific activity (▲) and specific RhuA content (▼) profiles along time in a substrate-inhibited fed-batch culture. Glucose concentration was increased to 65 g.L^{-1} at 19 hours, by adding 100 mL of a concentrated glucose solution. The vertical dotted line shows when continuous IPTG supply was started.

After processing the fermentation broth, a single step purification-immobilization using clear lysate was carried out, and the immobilization yield achieved in this case was 95%. As in the substrate-limited experiments performed at 28ºC, almost all the activity was able to adsorb on the IMAC support.

Proteolysis effects over the recombinant protein decreased, and western blot analyses (Figure 8.5) confirmed that the specific content of 6xHis-tag in RhuA was nearly constant along the production phase. At the end of the culture, about 80% of the initial 6xHis-tag·mg RhuA^{-1} was detected. These results are compatible with the 95% activity immobilization yield, which included the attachment of 6xHis tagged protein and also partially hydrolyzed protein still containing enough histidine residues to adsorb onto the IMAC support.

The negative effects of proteolysis over RhuA biological activity were also diminished, as presented in Figure 8.5. The average specific activity of the recombinant enzyme was close to 4 AU·mg RhuA^{-1} during most of the time, and as high as 4.8 AU·mg RhuA^{-1} at the end of the culture. Thus, RhuA quality was significantly improved reaching similar levels than the operation at low temperature.
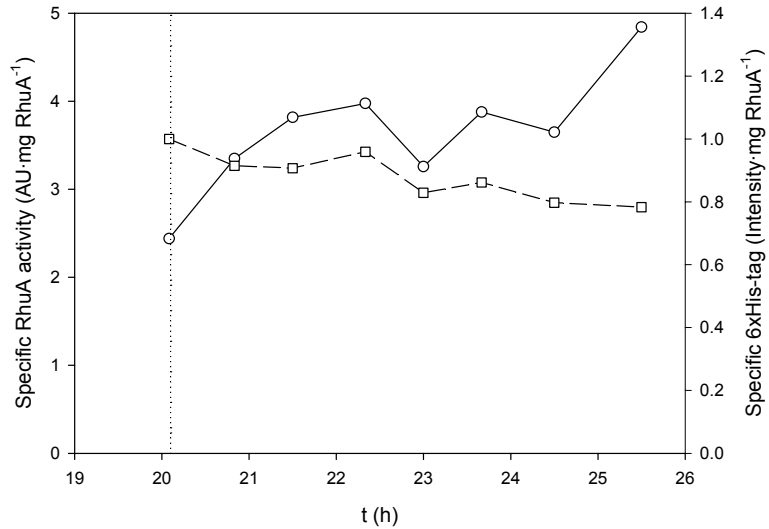

Figure 8.5. Specific RhuA activity (○) and specific 6xHis-tag contents in RhuA (□) along the induction phase in a substrate-inhibited fed-batch culture. The vertical dotted line shows when continuous IPTG supply was started.

8.4. Conclusions.

The standard substrate-limited fed-batch strategy at 37ºC was found to have a negative impact on recombinant 6xHis-tagged RhuA production and downstream. Starvation-induced proteolysis affected both the enzyme biological activity as well as its recovery yield. When global specific yield was calculated, it was found that only around 150-200 AU·gDCW^{-1} were finally immobilised as active biocatalyst. To overcome these problems, two alternative operational procedures have been proposed.

The operation of substrate limited fed-batch cultures at low temperature (28ºC) allowed increasing the specific enzyme activity when compared to the same strategy at 37ºC. In addition, the recombinant enzyme could be almost quantitatively recovered in a single step. At the end of the whole process, around 750-800 AU·gDCW^{-1} were immobilised as active biocatalyst, significantly improving (nearly 5-fold) the global yield when compared to the process run at 37ºC.

On the other hand, a novel growth strategy at non-limiting substrate levels was implemented in high cell density cultures at 37ºC. Cultures were grown to high biomass concentrations and submitted to continuous inducer supply for RhuA overexpression. When compared to substrate-limited cultures, higher inducer concentrations were required, although still much lower than the standard IPTG concentrations commonly used -in the range 500 μM to 1 mM- [18]. The required minimum value for aldolase expression moved from 4 to 10 μM and the IPTG amount affecting cell growth from 12 to 25 μM. Successful operation at high glucose levels allowed obtaining slightly higher protein levels and enzyme quality than in the process run at 28ºC. Quantitative recovery of the enzyme in active immobilized form was achieved, and the global specific yield was increased to approximately 1000 AU·gDCW^{-1} of immobilised active biocalalyst (nearly 6-fold compared to the substrate-limited process at 37ºC). These findings could justify the higher glucose costs associated with this operation mode.

Therefore, two different alternative operational strategies to substrate-limited fed-batch cultures performed at 37ºC were successfully implemented for recombinant RhuA production. The novel strategies have been proven to improve product yield and quality, especially when analysing the production and downstream processes altogether.

References.

[1] L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

[2] I.Ardao, M.D.Benaiges, G.Caminal, G.Alvaro, One step purification-immobilization of fuculose-1-phosphate aldolase, a class II DHAP dependent aldolase, by using metal-chelate supports, Enzyme and Microbial Technology 39 (2006) 22-27.

[3] J.Pinsach, C.de Mas, J.Lopez-Santin, Induction strategies in fed-batch cultures for recombinant protein production in *Escherichia coli*: Application to rhamnulose 1-phosphate aldolase, Biochemical Engineering Journal 41 (2008) 181-187.

[4] G.L.Jordan, S.W.Harcum, Characterization of up-regulated proteases in an industrial recombinant *Escherichia coli* fermentation, Journal of Industrial Microbiology & Biotechnology 28 (2002) 74-80.

[5] D.M.Ramirez, W.E.Bentley, Characterization of stress and protein turnover from protein overexpression in fed-batch *E. coli* cultures, Journal of Biotechnology 71 (1999) 39-58.

[6] A.Rozkov, T.Schweder, A.Veide, S.O.Enfors, Dynamics of proteolysis and its influence on the accumulation of intracellular recombinant proteins, Enzyme and Microbial Technology 27 (2000) 743-748.

[7] A.Rozkov, S.O.Enfors, Analysis and Control of Proteolysis of Recombinant Proteins in *Escherichia coli*, Physiological Stress Responses in Bioprocesses, Springer Berlin/Heidelberg (2004) 163-195.

[8] D.S.Waugh, Making the most of affinity tags, Trends in Biotechnology 23 (2005) 316-320.

[9] W.E.Bentley, et al., Generation of a histidine-tagged antibotulinum toxin antibody fragment in *E. coli*: Effects of post-induction temperature on yield and IMAC binding-affinity, Journal of Industrial Microbiology & Biotechnology 21 (1998) 275-282.

[10] A.Mitraki, J.King, Protein Folding Intermediates and Inclusion Body Formation, Nature Biotechnology 7 (1989) 690-697.

[11] A.L.Goldberg, A.C.John, Intracellular Protein Degradation in Mammalian and Bacterial Cells: Part 2, Annual Review of Biochemistry 45 (1976) 747-804.

[12] J.Pinsach, C.de Mas, J.Lopez-Santin, G.Striedner, K.Bayer, Influence of process temperature on recombinant enzyme activity in *Escherichia coli* fed-batch cultures, Enzyme and Microbial Technology 43 (2008) 507-512.

[13] L.Andersson, S.J.Yang, P.Neubauer, S.O.Enfors, Impact of plasmid presence and induction on cellular responses in fed batch cultures of *Escherichia coli*, Journal of Biotechnology 46 (1996) 255-263.

[14] B.M.Hogema, et al., Inducer exclusion by glucose 6-phosphate in *Escherichia coli*, Molecular Microbiology 28 (1998) 755-765.

[15] H.Ishizuka, A.Hanamura, T.Kunimura, H.Aiba, A lowered concentration of cAMP receptor protein caused by glucose is an important determinant for catabolite repression in *Escherichia coli*, Molecular Microbiology 10 (1993) 341-350.

[16] P.R.Jensen, K.Hammer, Artificial promoters for metabolic optimization, Biotechnology and Bioengineering 58 (1998) 191-195.

[17] A.M.Sanden, M.Bostrom, K.Markland, G.Larsson, Solubility and proteolysis of the Zb-MaIE and Zb-MaIE31 proteins during overproduction in *Escherichia coli*, Biotechnology and Bioengineering 90 (2005) 239-247.

[18] R.S.Donovan, C.W.Robinson, B.R.Glick, Review: optimizing inducer and culture conditions for expression of foreign proteins under the control of the *lac* promoter, Journal of Industrial Microbiology and Biotechnology 16 (1996) 145-154.

[19] R.V.Datar, T.Cartwright, C.G.Rosen, Process Economics of Animal-Cell and Bacterial Fermentations - A Case-Study Analysis of Tissue Plasminogen-Activator, Bio-Technology 11 (1993) 349-357.

[20] S.Gottesman, M.R.Maurizi, CELL BIOLOGY: Enhanced: Surviving Starvation, Science 293 (2001) 614-615.

[21] S.Gottesman, M.R.Maurizi, Regulation by Proteolysis - Energy-Dependent Proteases and Their Targets, Microbiological Reviews 56 (1992) 592-621.

[22] L.Andersson, L.Strandberg, S.O.Enfors, Cell segregation and lysis have profound effects on the growth of *Escherichia coli* in high cell density fed batch cultures, Biotechnology Progress 12 (1996) 190-195.

[23] P.R.Jensen, H.V.Westerhoff, O.Michelsen, The Use of *Lac*-Type Promoters in Control Analysis, European Journal of Biochemistry 211 (1993) 181-191.

9. GENERAL CONCLUSIONS AND RECOMMENDATIONS.

In this work, the production process of two recombinant aldolases in *E. coli* has been improved. The main conclusion is that the expression system as well as the target protein strongly influence the process productivity and, therefore, suitable process strategies must be developed for every particular case.

On the one hand, production of recombinant fuculose 1-phosphate aldolase (FucA) under the control of a weak IPTG inducible promoter in *E. coli* XL1 Blue MRF' pTrcfuc was automated by a simple feedback specific growth rate control method. The control strategy, based on mass balances and exhaust gas analysis, might be applicable to other systems in which protein production is growth associated, particularly when induction does not burden the host cells.

On the other hand, process strategies for recombinant rhamnulose 1-phosphate aldolase (RhuA) production under the control of a strong IPTG inducible promoter in *E. coli* M15 ΔglyA [pREP4] pQEαβrham were developed. High productivities with low IPTG requirements were achieved when using this expression system based on a strong promoter, but it was found to be very sensitive to the process strategy.

First, optimisation of the induction strategy showed that pulse induction in substrate-limited fed-batch cultures caused severe metabolic burden in host cells. As a consequence, reduced specific activity yields were obtained when compared with preliminary expression studies performed in shake flask cultures. To overcome this problem, an alternative induction strategy based on continuous IPTG dosage was implemented. Although the tuning of the transcription rate below critical values decreased inducer requirements, minimised the metabolic burden in host cells and prolonged the induction phase, the specific RhuA activity yields did not reach the same levels obtained in preliminary expression studies.

However, the continuous inducer addition led to high levels of protein yields, evidencing a decrease of specific aldolase activity along the induction phase when the process was run at 37°C. By reducing the process temperature to 28°C, high protein yields were also achieved, but the target protein quality was significantly improved due to the reduction of temperature-driven stress and lower gene dosage. Proteolysis, protein folding and plasmid copy number were variables which depended on process temperature and significantly affected RhuA quality.

Besides, a model describing RhuA production to optimise the process strategy was developed. A robust model describing *E. coli* M15 ΔglyA [pREP4] pQEαβrham growth at 37°C was obtained, but it was not possible to obtain a production model robust enough to optimise the process. However, the modelling task helped to identify key process features which were not properly understood or should be improved. Cell segregation into a viable-but-not-culturable state, activity losses and protein degradation, IPTG transport into the cell, plasmid copy number and overexpression of the auxotrophy selection marker (SHMT) were factors which strongly affected the process performance and, therefore, should be studied with more detail. A better understanding of these aspects would not only improve the description of the process, but it would also help to avoid some of the process bottlenecks.

Finally, the whole process to obtain immobilised RhuA ready to be used as biocatalyst was analysed. The production process strategy was found to have a significant impact on downstream yields, evidencing the need to take into account both stages when developing the process. When the production process was run at 37°C under substrate limiting conditions, reduced biological activity of the enzyme and low enzyme recovery yields were achieved. To overcome these problems, two alternative production strategies were implemented. First, reduction of the process temperature to 28°C decreased temperature-driven stress in host cells, leading to an improved biological activity of the enzyme and nearly complete immobilisation yield. Then, a novel production strategy at non-limiting substrate levels was developed to decrease the starvation-induced stress in host cells. The strategy was based on the control of the specific growth rate at inhibiting substrate concentrations to avoid acetate accumulation in fed-batch cultures performed at 37°C. Results showed an increase on protein activity and quantitative immobilisation of active RhuA. Therefore, both strategies led to improved protein qualities, significantly increasing immobilised RhuA yields.

As in the case of the automation of FucA production, the process strategies developed for RhuA production could be applied in other cases, particularly when expression systems based on strong promoters are used to overexpress proteins which need to be active for their further use.

Recommendations for future process development are based on aldolase production bottlenecks identified during this work regarding growth, induction and protein quality.

When developing the growth strategy, acetate accumulation was identified as a limiting step. Therefore, the use of a host strain which possesses an acetate production self-control mechanism could be considered (like *E. coli* BL21). That should allow rapid growth on excess of glucose to reach high cell densities. This properties could be translated into shorter process times (higher

productivities) and decreased starvation-induced stress experienced by cells (higher protein quality).

When developing the induction strategy, the use of strong promoters significantly decreased the amount of inducer and time required to obtain high yields of target protein, but significant efforts were required to tune protein expression in order to avoid a sudden metabolic burden of the host cells. As a consequence, it is also recommended to look for alternative inducers to IPTG (they should simplify the induction strategy, be non-toxic and cheaper). Interesting induction strategies are based on the *araB* promoter (which is induced by arabinose) as well as on the *proU* promoter (which is inducible by NaCl) controlling T7 RNA polymerase expression, for instance.

On the other hand, gene maintenance and dosage was identified as another factor influencing the induction strategy. The glycine auxotrophy plasmid maintenance system has been proven to be effective, but the expression of SHMT gene should be tuned properly. As an alternative, cloning the gene of the target protein into the host chromosome could be considered (this could avoid plasmid maintenance as well as gene dosage considerations).

Protein misfolding and degradation were identified as factors which can negatively influence the target protein quality. A decrease of the process temperature as well as induction in the presence of an excess of glucose presented a significant improvement of the protein quality, but other strategies like the use of protease deficient strains could be also considered (for instance, *E. coli* BL21 strains lack the Lon protease as well as the OmpT outer membrane protease).

Finally, other process bottlenecks like the oxygen transfer limitations could be studied to sustain the culture to higher cell densities and, as a consequence, increase the process yields.

10. APPENDICES.

10.1. List of publications.

J.Pinsach, C.de Mas, J.Lopez-Santin, A simple feedback control of *Escherichia coli* growth for recombinant aldolase production in fed-batch mode, Biochemical Engineering Journal 29 (2006) 235-242.

J.Pinsach, C.de Mas, J.Lopez-Santin, Induction strategies in fed-batch cultures for recombinant protein production in *Escherichia coli*: Application to rhamnulose 1-phosphate aldolase, Biochemical Engineering Journal 41 (2008) 181-187.

J.Pinsach, C.de Mas, J.Lopez-Santin, G.Striedner, K.Bayer, Influence of process temperature on recombinant enzyme activity in *Escherichia coli* fed-batch cultures, Enzyme and Microbial Technology 43 (2008) 507-512.

J.Pinsach, J.Ruiz, G.Alvaro, G.Gonzalez, C.de Mas, D.Resina, J.Lopez-Santin, Alternative production process strategies in *E. coli* improving protein quality and downstream yields, Process Biochemistry (submitted).

L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.

10.2. Galactose as an alternative inducer.

As described in section *1.1.3. Induction of recombinant protein expression*, both expression systems used in this work were based on *lac*-derived promoters and IPTG was always used to induce recombinant protein expression.

When using *E. coli* XL1 Blue MRF' pTrcfuc, the amount of IPTG required to optimise FucA production was considerably high [1] because transcription was under the control of the weak *trc* promoter (see section *3.1. Strains and vectors* for further details).

When using *E. coli* M15 ΔglyA [pREP4] pQEαβrham, the amount of IPTG required to optimise RhuA production was considerably low [2] because transcription was under the control of the strong phage T5 promoter (see section *3.1. Strains and vectors* for further details). However, modulation of the transcription rate by implementing an IPTG dosage strategy to avoid the metabolic burden imposed on host cells increased process and operation complexity (results from chapter 5).

When considering all these facts plus the high cost and potential toxicity of IPTG (depending on the application, complete IPTG removal might increase downstream costs significantly), it would be interesting to replace it by other inducers before process scale-up. Therefore, research of alternative inducers to IPTG was focused on RhuA production using the expression system based on the strong T5 promoter, *E. coli* M15 ΔglyA [pREP4] pQEαβrham.

Unfortunately, *E. coli* M15 ΔglyA [pREP4] pQEαβrham has a partial deletion in β-galactosidase gene (it is a *lacZ* mutant). Since β-galactosidase is the enzyme responsible for the conversion of lactose into allolactose -the true inducer of the *lac* operon [3]- (Figure 10.1), neither cell growth nor induction of protein expression was expected when culturing this host on lactose. Anyway, growth and induction on lactose was tested because it has been reported that a second gene showing β-galactosidase activity exists in some *E. coli* K12 *lacZ* mutants. The enzyme coded by this gene is named Ebg, and significantly differs from β-galactosidase characteristics, but previous works showed high β-galactosidase activity [4]. In our case, however, preliminary results obtained in shake flask cultures using lactose as the only carbon source and inducer did not show any ability of *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells to degrade lactose (data not shown).
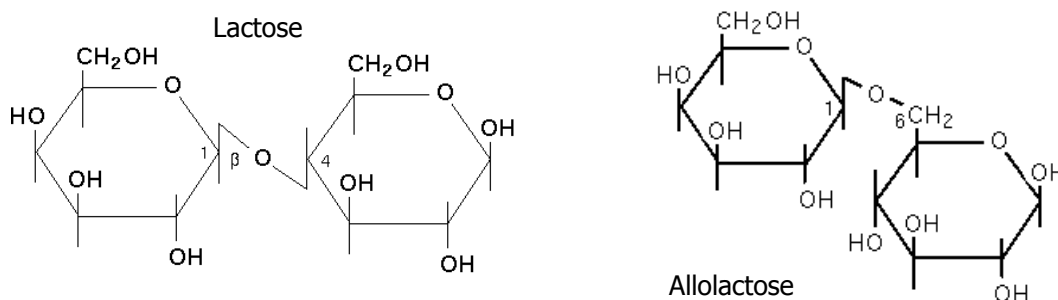
Figure 10.1. Lactose vs. allolactose.

Looking for other potential inducers of the *lac* operon, it was noticed that both IPTG and allolactose contained a galactose moiety (Fig. 10.1 and 10.2). Even though galactose had never been reported as an inducer of the *lac* operon in *E. coli*, it had been reported that it effectively induced the *lac* operon in *Klebsiella aerogenes* [5]. Therefore, galactose was also tested as a potential alternative inducer to IPTG in *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells.
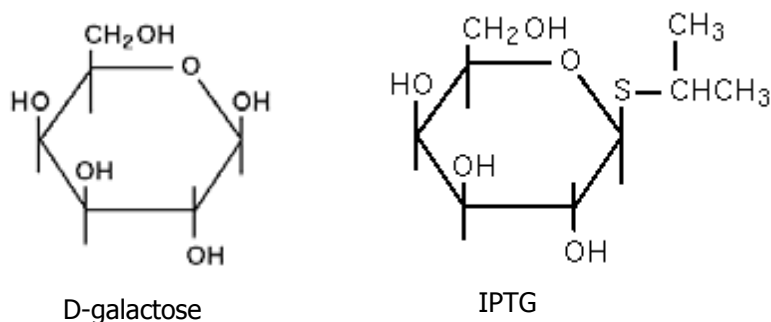
Figure 10.2. Galactose vs. IPTG.

Preliminary experiments were conducted in 5 parallel shake flask cultures in order to characterise cell growth and RhuA production. Galactose was added as a sole carbon source and inducer at initial concentrations ranging from 1 to 20 g·L^{-1} into 4 of them, while the fifth culture was run using a mixture of glucose and galactose, each one supplemented at 2.6 g·L^{-1} of initial concentration (the rest of the medium components were added to the same concentrations as in the defined media described in section *3.2. Media composition* in all cases, since glucose and galactose were expected to present comparable substrate to biomass yields).

Cells grew exponentially in all cases for the first 8 h, until other factors than carbon source limited cell growth (Fig. 10.3).
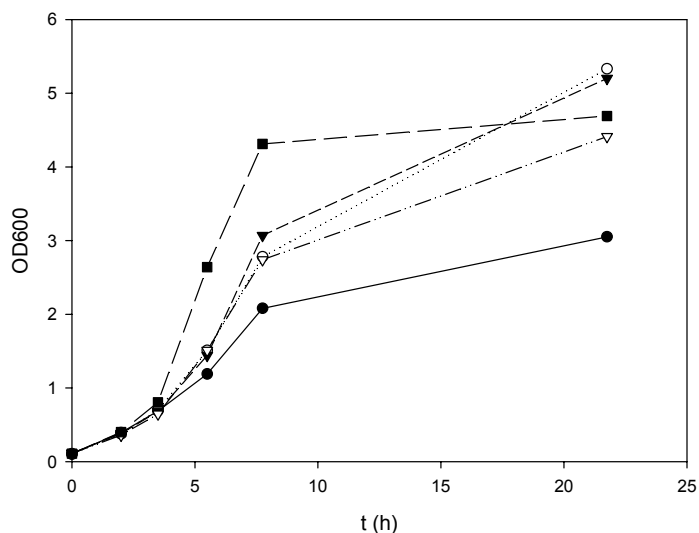

Figure 10.3. Cell growth of *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells in shake flask cultures supplemented with initial galactose concentrations of 1.6 (●), 3.9 (○), 8.4 (▼) and 17.3 (▽) $g \cdot L^{-1}$ and with initial galactose and glucose of 2.6 $g \cdot L^{-1}$ (■) each.

From these results, a Monod growth equation could be fitted (μ_{max}= 0.41 h^{-1}, K_S= 0.68 $g \cdot L^{-1}$) to experimental data when plotting the initial specific growth rates versus the initial galactose concentrations in the cases where the latter was the only carbon source (Fig. 10.4).
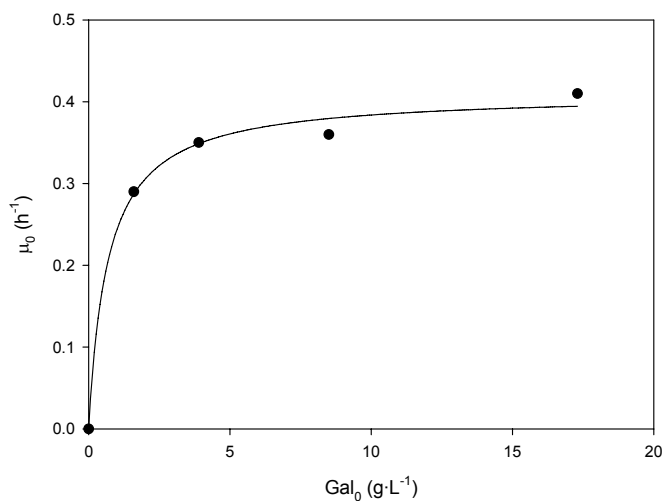
Figure 10.4. Experimental (●) and modelled (−) growth kinetics of *E. coli* M15 ΔglyA [pREP4] pQEαβrham growing on galactose at 37ºC.

Cell growth was faster (μ_{max}=0.57 h^{-1}) in the culture growing on both glucose and galactose than in the rest of the cultures, both substrates were consumed 8 h after inoculation and acetate was accumulated to 0.8 g·L^{-1} (data not shown).

When analysing the activity contents within the cells along time, several observations were made. First, significant specific RhuA activity was measured in some cases, although it took some time to reach it. A final specific activity close to 1400 AU·gDCW^{-1} was determined in the culture performed using 3.9 g·L^{-1} of initial galactose concentration (comparable to the maximum specific RhuA activity ever obtained using IPTG as inducer -see results from chapter 5-), and 900 AU·gDCW^{-1} were obtained when the initial galactose concentration was 1.6 g·L^{-1} (Fig. 10.5a). Cultures conducted at higher initial galactose concentrations did not show significant RhuA activity. That could be explained either because process variables were not under control –like pH- or because of acetate accumulation in the media, which was close to 2.5 g·L^{-1} at the end of those cultures (data not shown).

When calculating the specific activity production rates (Fig. 10.5b), results suggested that RhuA production depended on both galactose concentration as well as on the specific growth rate. The specific activity production rate increased at higher galactose concentrations and specific growth rates, but it was not maintained in those cases where acetate was accumulated.
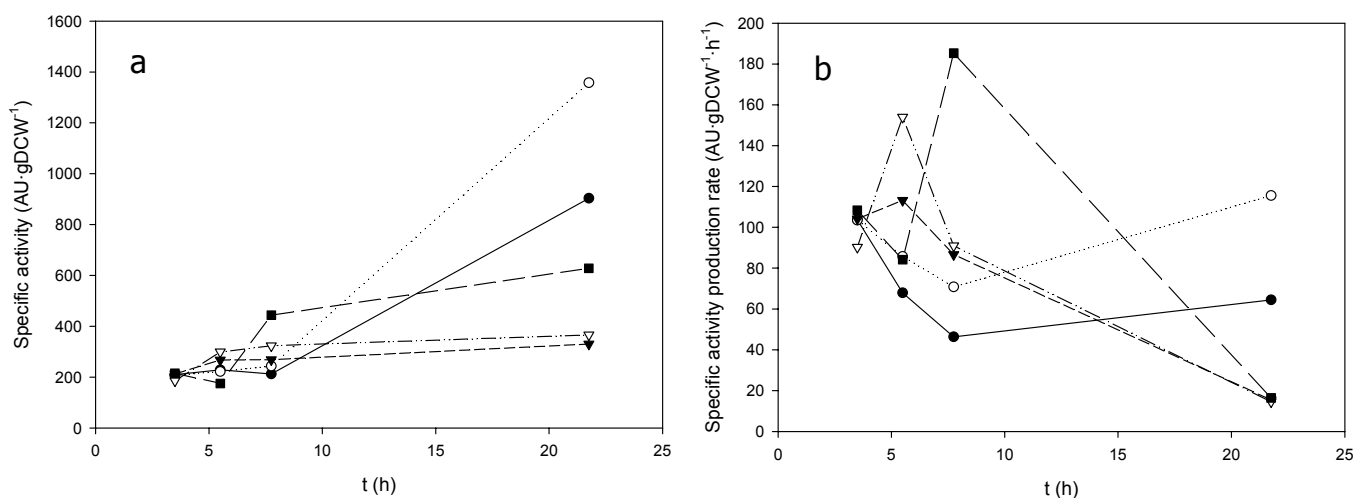

Figure 10.5. Specific activity (a) and specific activity production rate (b) profiles of *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells growing in shake flask cultures supplemented with initial galactose concentrations of 1.6 (●), 3.9 (○), 8.4 (▼) and 17.3 (▽) g·L^{-1} and with initial galactose and glucose of 2.6 g·L^{-1} (■) each.

These results showed that galactose was a potential alternative inducer of RhuA expression as well as an alternative carbon source to glucose. Although production was slower than in the case of IPTG, significant RhuA activity was achieved. Similarly to the cases where glucose was used as carbon source and IPTG as inducer, results suggested that protein yields were sensitive to acetate accumulation.

The next step was to conduct a galactose fed-batch to extend the production phase while controlling acetate accumulation. A fed-batch performed using galactose as the only carbon source and inducer showed an exponential cell growth at $\mu=0.21$ h^{-1} (Fig. 10.6). However, RhuA was not overexpressed even though growth was sustained for 30 h, galactose concentration was kept above 2 $g \cdot L^{-1}$ and acetate accumulation was not detected during the whole culture (data not shown).
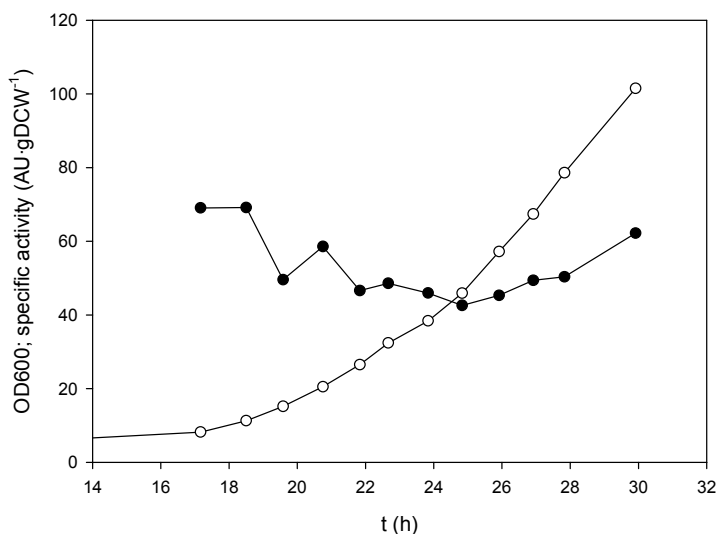

Figure 10.6. Cell growth (○) and specific RhuA activity (●) profiles during a fed-batch culture of *E. coli* M15 ΔglyA [pREP4] pQEαβrham conducted at $\mu=0.21$ h^{-1} using galactose as the sole carbon source and inducer.

Therefore, it was concluded that there was some other variable (not only induction time and acetate accumulation) influencing the RhuA overexpression when galactose was used as inducer. Baldauf and collaborators [5] described a better induction in *Klebsiella aerogenes gal* mutant strains than in those showing a functional galactose catabolism due to rapid degradation of the inducer.

To find out the conditions under which galactose acted as an effective inducer, serial shake flask experiments using galactose as the sole carbon source and inducer at 5 $g \cdot L^{-1}$ of initial concentration

were conducted (Fig. 10.7). Cells grown in a first culture (A) were used to inoculate a second one (B) 8 h later as well as to inoculate a third (C) and fourth (D) cultures 22 after inoculation of A. The only difference between C and D was the inoculum volume used in each case (in C it was twice of that used in D).
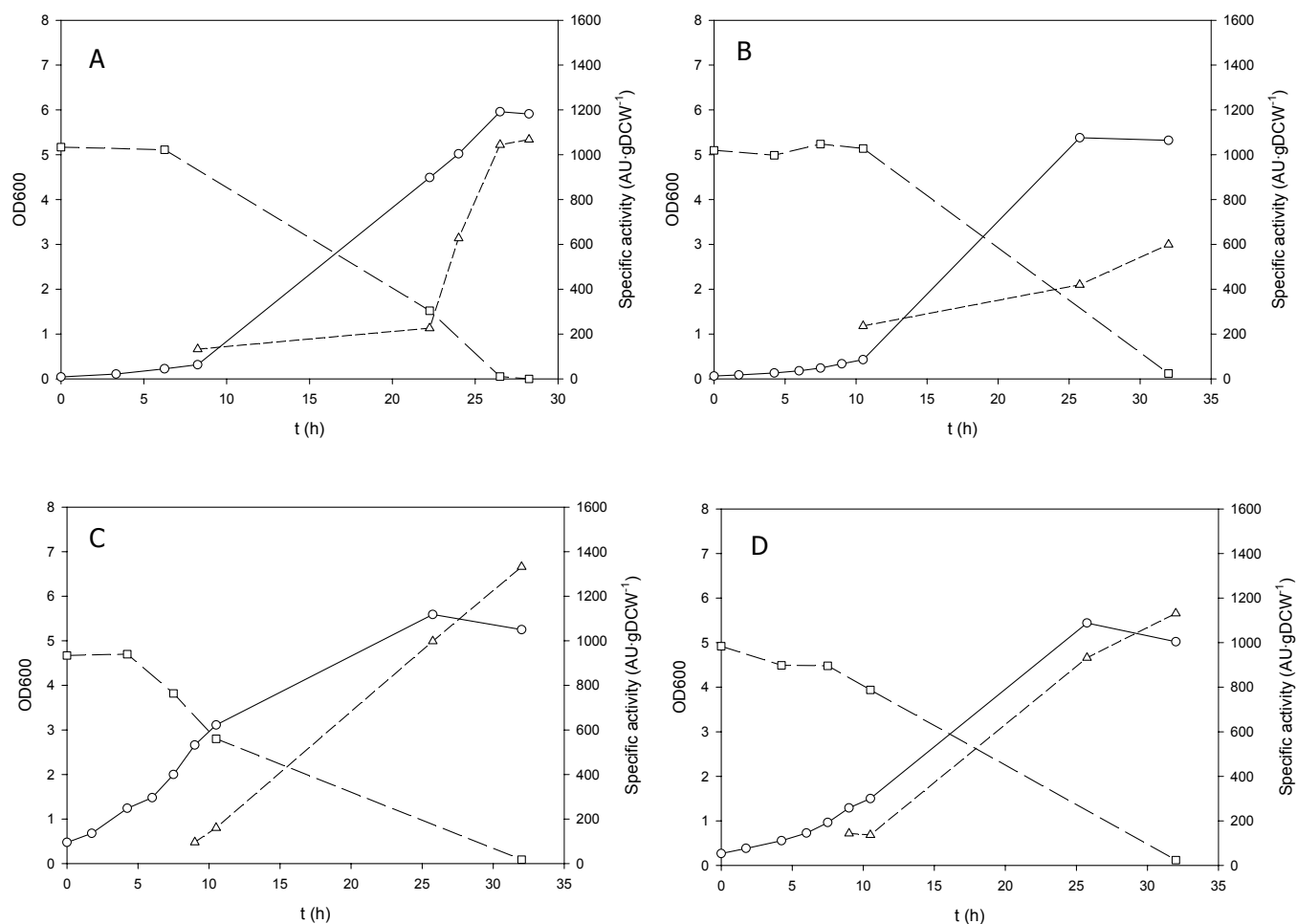

Figure 10.7. Optical density at 600 nm (○), galactose concentration (□) and specific RhuA activity (Δ) profiles in serial shake flask cultures of *E. coli* M15 ΔglyA [pREP4] pQEαβrham performed at 37°C using galactose as the only carbon source and inducer. Cultures B, C and D were inoculated from cells withdrawn from culture A at t= 8, 22 and 22 h, respectively.

Again, high specific RhuA activities were measured (as high as 1332 AU·gDCW^{-1} in culture C), confirming the effectiveness of galactose as an inducer of RhuA overexpression. These experiments suggested that RhuA overexpression was not only a matter of induction time: activity production was only detected when there was galactose available but there was a slow growth, probably due to oxygen transfer limitations. This latter assumption is based on the fact that it had been

repeatedly observed before that there was a specific growth rate decay of *E. coli* M15 ∆glyA [pREP4] pQEαβrham cells in shake flask cultures at biomass concentrations above 3 units of optical density at 600 nm (data not shown).

Thus, further research could be carried out to use galactose as an effective RhuA overexpression inducer, either:

- by introducing a *gal* mutation in the host and simply using glucose as the carbon source and galactose as the inducer, or
- by performing fed-batch cultures in which, when high cell densities were reached, only galactose was added as carbon source and inducer while limiting the oxygen supply to minimise cell growth and galactose consumption (hoping to avoid acetate accumulation).

10.3. Monitoring of *E. coli* M15 ΔglyA [pREP4] pQEαβrham cell growth.

Since RhuA production is growth-related when using glucose as substrate and IPTG as inducer (results from chapter 5), in-line monitoring of cell growth would be interesting to achieve a good process control.

Results from chapter 4 showed that exhaust gas analysis was a good method to monitor cell growth, but only when cell growth was not influenced by recombinant protein expression. Since IPTG pulse induction strongly burdened the metabolism of *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells, a sterilizable turbidity sensor (InPro 8200 from Mettler-Toledo) was used to try to monitor cell growth in-line (Fig. 10.8). The turbidity measurements of this optical probe were done at 550 nm, reason why a correlation between these values and the experimental measurements of optical density at 600 nm was required.
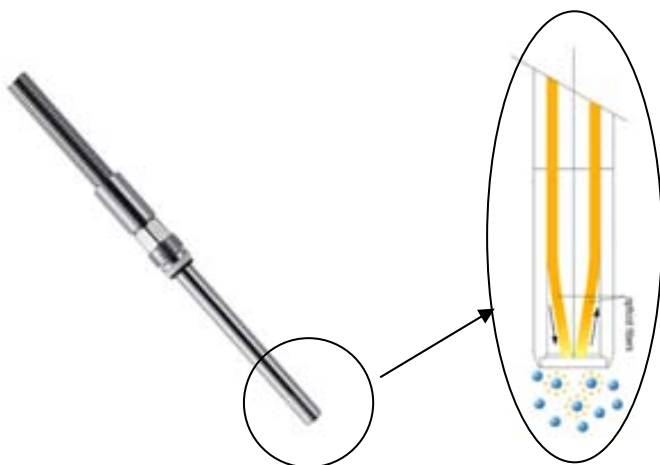

Figure 10.8. InPro 8200 turbidity sensor (Mettler-Toledo), based on backscattered light technology measurements by its dual optical fiber system.

As an example, a fed-batch culture performed at 37ºC limiting the glucose feeding to control the specific growth rate at $\mu=0.20$ h^{-1} is presented (Fig. 10.9). Cells grew exponentially to optical densities at 600 nm above 100. Even though an IPTG feed was switched on at t=18 h, induction did not burden the cells until t>25 h (at that point substrate started to accumulate –data not shown- and the specific growth rate could not be maintained). During the whole culture a constant gas flow rate of 1.5 LN·min^{-1} was fed into the bioreactor, being air until t=21 h and pure oxygen after this point.
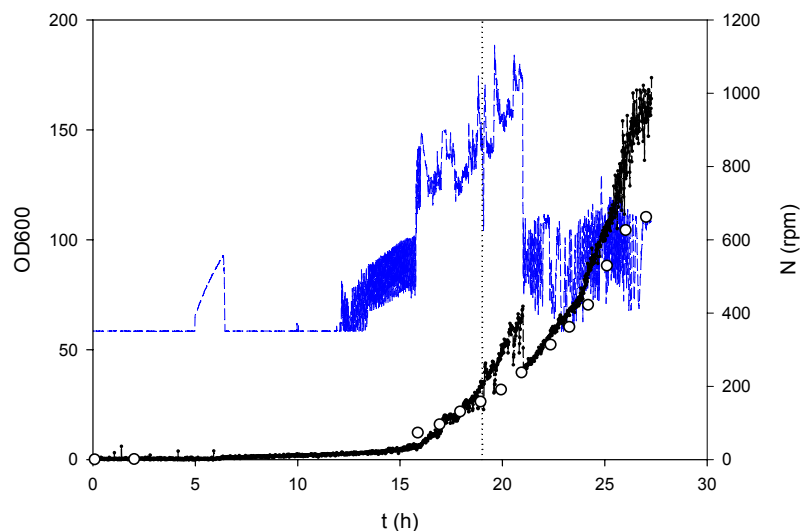
Figure 10.9. Off-line –experimental- (○) and in-line -optical probe- (▪) biomass concentration measurements, and stirring rate (---) in a glucose-limited fed-batch culture. The vertical dotted line depicts when the IPTG feed was switched on.

The optical probe did not lead to a robust prediction of the experimental biomass measurements. The cell concentration was underestimated at low cell densities, while it was overestimated at higher biomass concentrations. Moreover, a significant influence of the stirring speed was observed, especially when there were sudden changes of the latter (like when pure oxygen was fed into the bioreactor at t=21 h).

To overcome this problem, another fed-batch was carried out at constant stirring speed (N=1000 rpm). The specific growth rate was again controlled at μ=0.20 h^{-1} until cell density reached OD600=150 (Fig. 10.10). The IPTG feed was switched on at t=20 h, and the inlet gas flow rate was kept constant throughout the culture at 1.5 LN·min^{-1}, being air at the beginning which was progressively enriched with pure oxygen to maintain the dissolved oxygen signal within the reactor at 50% of saturation.

In this case, the biomass estimation from the optical probe improved up to 90 units of optical density at 600 nm, but significant deviations were observed from this point to the end of the culture.
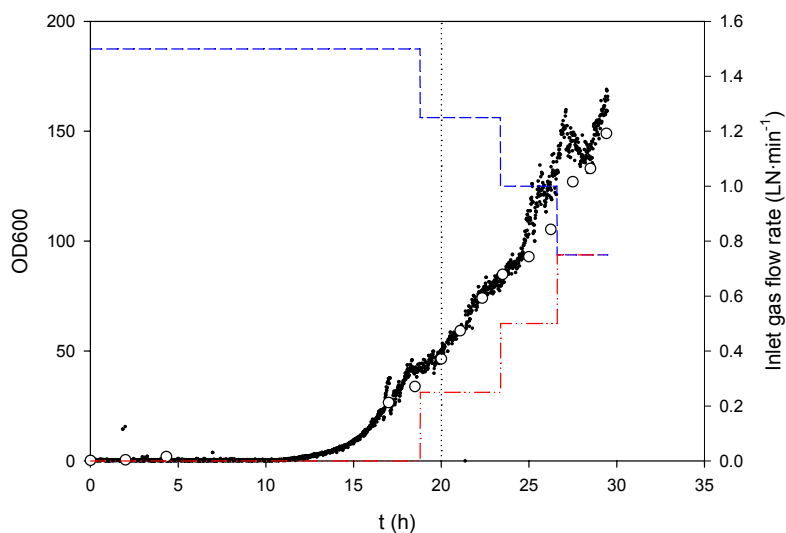
Figure 10.10. Off-line –experimental- (○) and in-line -optical probe- (▪) biomass concentration measurements, inlet air flow rate (---) and inlet oxygen flow rate (– ▪ – ▪ –) in a glucose-limited fed-batch culture. The vertical dotted line depicts when the IPTG feed was switched on.

Therefore, the stirring speed was not the only variable affecting the optical probe measurements, and further research should be done to find out and control other factors influencing them. In this sense, experimental observations could suggest:

- either PLP (pyridoxal-5'-phosphate), a yellow (yellow colour is detected at a wavelength of about 580 nm) SHMT cofactor secreted into the medium by *E. coli* M15 ΔglyA [pREP4] pQEαβrham cells [6, 7] interfered with the measurements at 550 nm from the optical probe or,

- cell size changes caused by RhuA overexpression affected the backscattered light in-line measurements.

References.

[1] O.Durany, C.de Mas, J.Lopez-Santin, Fed-batch production of recombinant fuculose-1-phosphate aldolase in *E. coli*, Process Biochemistry 40 (2005) 707-716.

[2] J.Pinsach, C.de Mas, J.Lopez-Santin, Induction strategies in fed-batch cultures for recombinant protein production in *Escherichia coli*: Application to rhamnulose 1-phosphate aldolase, Biochemical Engineering Journal 41 (2008) 181-187.

[3] C.Burstein, M.Cohn, A.Kepes, J.Monod, Role of lactose and its metabolic products in the induction of the lactose operon in *Escherichia coli*, Biochimica et Biophysica Acta 95 (1965) 634-639.

[4] J.H.Campbell, J.A.Lengyel, J.Langridge, Evolution of A Second Gene for β-Galactosidase in *Escherichia coli*, Proceedings of the National Academy of Sciences of the United States of America 70 (1973) 1841-1845.

[5] S.L.Baldauf, M.A.Cardani, R.A.Bender, Regulation of the Galactose-Inducible *Lac* Operon and the Histidine Utilization Operons in *Pts* Mutants of *Klebsiella aerogenes*, Journal of Bacteriology 170 (1988) 5588-5593.

[6] L.Vidal, J.Calveras, P.Clapes, P.Ferrer, G.Caminal, Recombinant production of serine hydroxymethyl transferase from *Streptococcus thermophilus* and its preliminary evaluation as a biocatalyst, Applied Microbiology and Biotechnology 68 (2005) 489-497.

[7] L.Vidal, J.Pinsach, G.Striedner, G.Caminal, P.Ferrer, Development of an antibiotic-free plasmid selection system based on glycine auxotrophy for recombinant protein overproduction in *Escherichia coli*, Journal of Biotechnology 134 (2008) 127-136.