Escola de Engenharia
Universidade do Minho

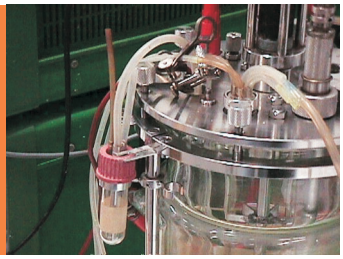
Isabel Cristina de Almeida Pereira da Rocha

MODEL-BASED STRATEGIES FOR COMPUTER-AIDED OPERATION OF A RECOMBINANT *E. COLI* FERMENTATION

Braga, 2003

Escola de Engenharia
Universidade do Minho

Isabel Cristina de Almeida Pereira da Rocha
MODEL-BASED STRATEGIES FOR COMPUTER-AIDED OPERATION OF
A RECOMBINANT *E. COLI* FERMENTATION



MODEL-BASED STRATEGIES FOR COMPUTER-AIDED OPERATION OF RECOMBINANT *E. COLI* FERMENTATION

Isabel Cristina de Almeida Pereira da Rocha

Dissertação para Doutoramento
em Engenharia Química e Biológica

Departamento de Engenharia Biológica

Escola de Engenharia
Universidade do Minho

2003

Provas de Doutoramento realizadas em 17 de Novembro de 2003.

Tese realizada sob a orientação do doutor Eugénio Manuel de Faria Campos Ferreira, Professor Associado
Universidade do Minho, Escola de Engenharia, Departamento de Engenharia Biológica.

Constituição do Júri

Presidente	Doutor António Sérgio Pouzada Professor Catedrático, Presidente da Escola de Engenharia da Universidade do Minho
Vogais	Doutor Georges Bastin Professor Catedrático do Department d'Ingénierie Matématique da Université Catolique de Louvain, Bélgica
	Doutor Joaquim Manuel Sampaio Cabral Professor Catedrático do Departamento de Engenharia Química da Universidade Técnica de Lisboa
	Doutor Sebastião José Cabral Feyo de Azevedo Professor Catedrático do Departamento de Engenharia Química da Universidade do Porto
	Doutor José António Couto Teixeira Professor Catedrático do Departamento de Engenharia Biológica da Universidade do Minho
	Doutor Eugénio Manuel de Faria Campos Ferreira Professor Associado do Departamento de Engenharia Biológica da Universidade do Minho
	Doutora Isabel Maria Pires Belo Professora Auxiliar do Departamento de Engenharia Biológica da Universidade do Minho

Universidade do Minho. Braga, Novembro 2003.

AGRADECIMENTOS

Ao finalizar esta tese, não posso deixar de agradecer a todos quantos me acompanharam e apoiaram ao longo destes últimos anos. Em primeiro lugar, gostaria de agradecer ao meu orientador científico, Professor Eugénio Ferreira, pelo apoio concedido ao longo destes anos. Sem dúvida que as suas excepcionais qualidades científicas e humanas tornaram a realização deste trabalho numa tarefa bem mais aliciante. Os seus constantes incentivos nas alturas menos boas foram determinantes para o bom termo deste trabalho. Ao director do Lab. de Fermentações e, simultaneamente do Departamento de Engenharia Biológica, Professor José Teixeira, pelas facilidades na cedência de alguns equipamentos e pela boa disposição sempre presente. Ao Professor Manuel Mota, na qualidade de director do Centro de Engenharia Biológica, pelo acolhimento como investigadora deste centro.

Gostaria ainda de agradecer às várias pessoas que, ao longo das tarefas relativas ao trabalho experimental, ajudaram de uma ou outra forma. Em relação à condução de fermentações, numa primeira fase ao Paulo Margarido e à Cláudia Santos, os meus agradecimentos pela disponibilidade de algumas das vossas noites de sono no acompanhamento destes processos. Os mesmos agradecimentos são devidos à Sandra Carvalho e à Ana Veloso. A ambas, gostaria ainda de expressar o meu reconhecimento pela amizade incondicional e pela disponibilidade sempre demonstradas.

Ao Doutor José Catita, da empresa Paralab, pela importante ajuda na implementação dos métodos de FIA.
Ao Bruno Sommer Ferreira, do Instituto Superior Técnico, pelos conselhos valiosos prestados relativamente ao funcionamento do MS.
Ao meu irmão Miguel Rocha, por algumas sugestões interessantes relativas à aplicação de algoritmos genéticos.
To Velislava Lyubenova I would like to thank for the cooperation in the application of her algorithms to the developed model.

Ao sr. Manuel Santos, técnico do DEB, gostaria de expressar o meu agradecimento pela ajuda na montagem de parte da instalação experimental.

Gostaria de agradecer ainda às pessoas que durante a fase de escrita desta tese, me prestaram o seu apoio: ao Pablo, pela ajuda imprescindível nas questões gráficas, e ao António Peres pelas sugestões ao nível da escrita. À Lígia Rodrigues, pela amizade e por me ter substituído em alguns compromissos profissionais.

São devidos também os agradecimentos às instituições que contribuíram, pelos apoios concedidos, para a realização deste trabalho: à FCT pela bolsa PRAXIS XXI/16961/98 e ao FSE no Âmbito do III Quadro Comunitário de Apoio. À Fundação Calouste Gulbenkian pela atribuição de uma bolsa relativa a deslocação a uma conferência no estrangeiro.

AGRADECIMENTOS (CONT.)

À Agência de Inovação pelo apoio concedido no âmbito do projecto PROTEXPRESS.
Aos parceiros desse projecto, nomeadamente a empresa Biotecnol e o CEQB (IST) pela colaboração em algumas tarefas.

Gostaria ainda de agradecer à minha família e aos meus amigos todo o apoio e incentivo prestados ao longo destes anos.

Finalmente, gostaria de dedicar este trabalho ao meu marido Pablo e à memória dos meus pais, Teresa e Francisco.

Isabel

ABSTRACT

MODEL-BASED STRATEGIES FOR COMPUTER-AIDED OPERATION OF RECOMBINANT *E. COLI* FERMENTATION

The main objectives of this thesis were the development of model-based strategies for improving the performance of a high-cell density recombinant *Escherichia coli* fed-batch fermentation. The construction of a mathematical model framework as well as the derivation of optimal and adaptive control laws were used to accomplish these tasks. An on-line data acquisition system was also developed for an accurate characterization of the process and for the implementation of the control algorithms. The mathematical model of the process is composed of mass balance equations to the most relevant state variables of the process. Kinetic equations are based on the three possible metabolic pathways of the microorganism: glucose oxidation, fermentation of glucose and acetate oxidation. A genetic algorithm was used to derive the kinetic structure and to estimate both yield and kinetic coefficients of the model, minimizing the normalized quadratic differences between simulated and real values of the state variables.

After parameter estimation, a sensitivity function analysis was applied to evaluate the influence of the various parameters on model behavior. Sensitivity functions revealed the sensitivity of the state variables to variations in each model parameter. Thus, essential parameters were selected and the model could be re-written in a simplified version that could also describe accurately experimental data.

A system for the on-line monitoring of the major state variables was also developed. Glucose and acetate concentrations were measured with a developed Flow Injection Analysis system, while the carbon dioxide and oxygen transfer rates were calculated from data obtained with exhaust gas analysis. The fermentation culture weight was also continuously assessed with a balance, allowing the use of more precise mass-based concentrations, while environmental variables like pH, dissolved oxygen and temperatures were controlled and assessed via a Digital Control Unit. The graphical programming environment LabVIEW was used to acquire and integrate these variables in a supervisory computer, allowing the performance of integrated monitoring and control of the process.

A model-based adaptive linearizing control law was derived for the regulation of acetate concentration during fermentations. The non-linear model was subjected to transformations in order to obtain a linear behavior for the control loop when a non-linear control is applied. The implementation of the control law was performed through a C script embedded in the supervisory LabVIEW program.

Finally, two optimization techniques for the maximization of biomass concentration were compared: a first order gradient method and a stochastic method based on the biological principle of natural evolution, using a genetic algorithm. The former method revealed less efficient concerning to the computed maximum, and dependence on good initial values.

Key-words: bioprocess modeling; model identification; sensitivity functions; adaptive control; optimal control; stochastic optimization methods; Genetic Algorithms; Flow Injection Analysis; *Escherichia coli*; on-line monitoring and control system.

SUMÁRIO

ESTRATÉGIAS BASEADAS EM MODELOS PARA A OPERAÇÃO ASSISTIDA POR COMPUTADOR DA PRODUÇÃO DE PROTEÍNAS HUMANAS RECOMBINADAS

A presente tese teve como principais objectivos o desenvolvimento de estratégias baseadas em modelos para melhorar o desempenho da fermentação em modo semi-contínuo em altas densidades celulares de *Escherichia coli* recombinada. Para o efeito, foi construído um modelo matemático representativo do processo e a partir deste foram desenvolvidos algoritmos de controlo óptimo e adaptativo. De forma a possibilitar a implementação de leis de controlo em linha e a caracterização do processo fermentativo, foi desenvovido um sistema informático de aquisição e envio de dados.

O modelo matemático representativo do processo em estudo foi elaborado tendo por base as equações dinâmicas de balanço mássico para as variáveis de estado mais relevantes, contemplando as três possíveis vias metabólicas do microrganismo. A estrutura cinética, bem como os parâmetros do modelo foram determinados por recurso a uma abordagem sistemática tendo por base a minimização das diferenças quadráticas entre dados reais e dados simulados, com recurso a uma ferramenta de optimização estocástica denominada de Algoritmos Genéticos. Após a etapa de identificação do modelo matemático, foram calculadas as sensibilidades relativas ao longo do tempo das variáveis de estado do modelo relativamente aos vários parâmetros determinados. Os resultados desta análise de sensibilidade possibilitaram avaliar a relevância de cada um dos parâmetros em causa, permitindo propor uma estrutura de modelo menos complexa, por exclusão dos parâmetros menos importantes.

O sistema elaborado para a aquisição e envio em linha de dados da fermentação inclui um sistema de FIA (*Flow Injection Analysis*) desenvolvido para a medição das concentrações de acetato e glucose, uma unidade de controlo digital que controla as variáveis físicas mais relevantes para o processo, e um equipamento de Espectrometria de Massas para analisar as correntes gasosas de entrada e saída do fermentador. O sistema dispõe ainda de duas balanças, uma das quais para a aferição em linha do peso do caldo de fermentação, permitindo o uso de concentrações mássicas que proporcionam resultados mais exactos. A aquisição e integração destas variáveis medidas são efectuadas através de um *software* de supervisão elaborado no ambiente de programação gráfico LabVIEW.

Adicionalmente, foi elaborada uma lei de controlo adaptativo linearizante para a regulação da concentração de acetato no meio de fermentação. A síntese da lei de controlo não linear foi efectuada por técnicas de geometria diferencial com linearização do sistema por retroacção de estado. A adaptação foi feita tendo por base a estimação de parâmetros variáveis no tempo, nos quais se concentram as incertezas do modelo. A implementação ao processo real da referida lei de controlo foi efectuada por recurso a um programa elaborado em C incluído no programa supervisor elaborado em LabVIEW. Finalmente, para a optimização da quantidade de biomassa formada no final da fermentação por manipulação do caudal de alimentação, foram estudadas duas ferramentas de optimização: um método de gradiente e uma ferramenta baseada em Algoritmos Genéticos. Esta última revelou-se mais eficaz tanto na convergência para o valor óptimo, como na estimativa inicial fornecida.

Palavras-chave: modelação de bioprocessos; identificação de modelos; funções de sensibilidade; controlo adaptativo; controlo óptimo; ferramentas estocásticas de optimização; Algoritmos Genéticos; *Flow Injection Analysis*; *Escherichia coli*; sistema de monitorização e envio de dados.

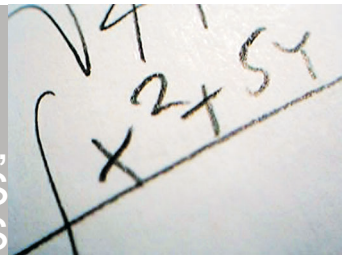


TABLE OF CONTENTS, LIST OF FIGURES, LIST OF TABLES AND LIST OF SYMBOLS

TABLE OF CONTENTS

TABLE OF CONTENTS
(CONT.)

TABLE OF CONTENTS
(CONT.)

TABLE OF CONTENTS
(CONT.)

LIST OF FIGURES

LIST OF FIGURES
(CONT.)

LIST OF FIGURES
(CONT.)

LIST OF FIGURES (CONT.)

LIST OF FIGURES (CONT.)

LIST OF FIGURES (CONT.)

LIST OF FIGURES
(CONT.)

LIST OF FIGURES
(CONT.)

LIST OF TABLES

LIST OF TABLES
(CONT.)

LIST OF SYMBOLS

Uppercase Latin letters

A	Acetate concentration	(ML^{-3} or MM^{-1})
A^*	Acetate concentration set-point	(ML^{-3} or MM^{-1})
A_{crit}	Critical acetate concentration	(ML^{-3} or MM^{-1})
A_{exp}	Experimental acetate concentration	(ML^{-3} or MM^{-1})
$\overline{A}_{exp}$	Average experimental acetate concentration	(ML^{-3} or MM^{-1})
A_{sim}	Simulated acetate concentration	(ML^{-3} or MM^{-1})
B	Bicarbonate concentration in the liquid phase	(ML^{-3} or MM^{-1})
C	Dissolved carbon dioxide concentration	(ML^{-3} or MM^{-1})
C	Concentrations of sample materials after dispersion in a FIA apparatus	(ML^{-3} or MM^{-1})
C_0	Concentrations of sample materials before dispersion in a FIA apparatus has taken place	(ML^{-3} or MM^{-1})
C_{eq}	Liquid phase concentration of carbon dioxide in equilibrium with the gas phase	(ML^{-3} or MM^{-1})
C_{in}	Concentration of carbon dioxide in the liquid inflow	(ML^{-3} or MM^{-1})
C_T	Total dissolved carbon dioxide concentration, including both forms of carbon dioxide and bicarbonate	(ML^{-3} or MM^{-1})
D	Dilution rate	(T^{-1})
D	FIA dispersion coefficient	
$\mathbf{F}$	Liquid mass flow feed vector (per unit of weight inside the reactor)	
F_a	Liquid mass flow of acid solution added to the bioreactor	(L^3T^{-1} or MT^{-1})
F_b	Liquid mass flow of base solution added to the bioreactor	(L^3T^{-1} or MT^{-1})
F_{evp}	Liquid mass flow evaporated from the bioreactor	(L^3T^{-1} or MT^{-1})
$\mathbf{F_f}$	Liquid mass flow feed vector for the components that exhibit slow dynamics	
F_{gas}	Liquid mass flow taken from the reactor due to gas transferences	(L^3T^{-1} or MT^{-1})
F_i	Mass inflow of component *i* in the bioreactor per unit of weight	($MM^{-1}T^{-1}$)
$F_{in,i}$	Liquid mass flow of the stream containing component *i* fed into the bioreactor	(L^3T^{-1} or MT^{-1})
$F_{in,S}$	Liquid mass flow of the stream containing glucose fed into the bioreactor	(L^3T^{-1} or MT^{-1})

$F_{in,tot}$	Total liquid mass flow fed into the bioreactor	(L^3T^{-1} or MT^{-1})
$\mathbf{F_s}$	Liquid mass flow feed vector for the components that exhibit fast dynamics	
F_{smp}	Liquid mass flow taken from the reactor due to sampling	(L^3T^{-1} or MT^{-1})
G_{in}	Molar gas inflow rate	(MT^{-1})
G_{out}	Molar gas outflow rate	(MT^{-1})
H^+	Protons concentration in the liquid phase	(ML^{-3} or MM^{-1})
H^{CO2}	Henry constant for carbon dioxide	(L^2T^{-2})
J	Objective function in a given optimization method	
$\mathbf{K}$	Yield or stoichiometric coefficients matrix	
K_A	Affinity constant for the substrate (acetate)	(ML^{-3} or MM^{-1})
K_{acid}	Carbonic acid dissociation constant	(ML^{-3} or MM^{-1})
$\mathbf{K_f}$	Yield coefficients matrix for the components that exhibit fast dynamics	
$\mathbf{K_s}$	Yield coefficients matrix for the components that exhibit slow dynamics	
$K_{i,O}$	Inhibition constant of acetate on oxygen uptake	(ML^{-3} or MM^{-1})
$K_{i,S}$	Inhibition constant of acetate on glucose uptake	(ML^{-3} or MM^{-1})
K_La^{CO2}	Carbon dioxide mass transfer coefficient from the liquid to the gas phase	(T^{-1})
K_La^{O2}	Oxygen mass transfer coefficient from the gas phase to the liquid phase	(T^{-1})
$K_{m,S}$	Inhibition constant of the specific growth rate on glucose on the maintenance coefficient	(T^{-1})
K_S	Affinity constant for the substrate (glucose)	(ML^{-3} or MM^{-1})
L_j	Number of reactants in reaction j	
M_{CO2}	Molar mass of carbon dioxide	(MM^{-1})
M_{O2}	Molar mass of oxygen	(MM^{-1})
O	Dissolved oxygen concentration	(ML^{-3} or MM^{-1})
O_{eq}	Liquid phase concentration of oxygen in equilibrium with O^G	(MM^{-1})
O_{in}	Concentration of oxygen in the liquid inflow	(ML^{-3} or MM^{-1})
O_{in}^G	Oxygen concentration in the inlet gas	(ML^{-3})
O_{out}^G	Oxygen concentration in the outlet gas	(ML^{-3})
P	Total pressure of the gas phase	($MT^{-2}L^{-1}$)
$\mathbf{Q}$	Vector of rate of mass flow in gaseous form (per unit of weight inside the reactor)	
$\mathbf{Q_f}$	Vector of rate of mass flow in gaseous form for the components that exhibit fast dynamics	

Q_i	Mass net gaseous flow of component *i*.	$(MM^{-1}T^{-1})$
$\mathbf{Q_{in}}$	Vector of rate of mass inflow in gaseous form (per unit of weight inside the reactor)	
Q_{in}	Volumetric inflow to the bioreactor in gaseous form	(L^3T^{-1})
Q_{out}	Volumetric outflow from the bioreactor in gaseous form	(L^3T^{-1})
$\mathbf{Q_{out}}$	Vector of rate of mass outflow in gaseous form (per unit of weight inside the reactor)	
Q_p	Total bioreactor productivity	(MT^{-1})
$\mathbf{Q_s}$	Vector of rate of mass flow in gaseous form for the components that exhibit slow dynamics	
R	Ideal gas constant	$(T^{-2}L^2\theta^{-1})$
R	Linear regression coefficient	
R_j	Number of products on reaction *j*	
RSD_i	Relative standard deviation of the component *i*	$(ML^{-3}$ or $MM^{-1})$
S	Substrate (glucose) concentration	$(ML^{-3}$ or $MM^{-1})$
$S_{1/2}$	Volume of sample solution that is necessary to reach a dispersion of 2	$(ML^{-3}$ or $MM^{-1})$
S_{crit}	Critical specific growth rate	$(ML^{-3}$ or $MM^{-1})$
S_{exp}	Experimental value of glucose concentration	$(ML^{-3}$ or $MM^{-1})$
$\overline{S}_{exp}$	Average value of the experimental glucose concentration	$(ML^{-3}$ or $MM^{-1})$
S_{in}	Substrate concentration in the inflow	$(ML^{-3}$ or $MM^{-1})$
S_{sim}	Simulated value of glucose concentration	$(ML^{-3}$ or $MM^{-1})$
T	Temperature	(θ)
T	Time span between the sample injection and the peak maximum in a FIA system	(T)
T	Sampling period	
V	Liquid volume inside the bioreactor	(L^3)
V_0	Liquid volume inside the bioreactor at the beginning of the fed-batch phase	(L^3)
V_G	Gas volume inside the reactor	(L^3)
W	Weight of liquid inside the bioreactor	(M)
X	Biomass concentration	$(ML^{-3}$ or $MM^{-1})$
X_0	Biomass concentration at the beginning of the fed-batch phase	$(ML^{-3}$ or $MM^{-1})$
X_{exp}	Experimental value of biomass concentration	$(ML^{-3}$ or $MM^{-1})$
$\overline{X}_{exp}$	Average value of the experimental biomass concentration	$(ML^{-3}$ or $MM^{-1})$
X_{sim}	Simulated value of biomass concentration	$(ML^{-3}$ or $MM^{-1})$

Lowercase Latin letters

c_a	Amount of acid necessary to equilibrate the pH due to acetate production, per amount of biomass formed	(MM^{-1})
c_b	Amount of base necessary to equilibrate the pH due to acetate production, per amount of biomass formed	(MM^{-1})
dif	Normalized difference between simulated and real data	
difA	Normalized difference between simulated and real data obtained for acetate	
difS	Normalized difference between simulated and real data obtained for glucose	
difX	Normalized difference between simulated and real data obtained for biomass	
i	Index of the components inside the bioreactor	
j	Index of the reactions taking place inside the reactor	
k	Time index	
k_{AS}	Acetate yield on glucose	(MM^{-1})
k_{AX}	Acetate yield on biomass	(MM^{-1})
k_{ij}	Yield coefficients or stoichiometric coefficients of component i in the reaction j	(MM^{-1})
k_{XA}	Biomass yield on acetate, obtained when acetate is being oxidized	(MM^{-1})
$k_{XS,F}$	Biomass yield on glucose, considering a fermentative pathway	(MM^{-1})
$k_{XS,OX}$	Biomass yield on glucose, considering a fully oxidative pathway	(MM^{-1})
m	Number of reactions taking place inside the bioreactor	
m_S	Maintenance coefficient (seen as a growth rate)	(T^{-1})
$m_{S,max}$	Maximum maintenance coefficient	(T^{-1})
n	Number of components inside the bioreactor	
$p_{CO2,out}$	Partial pressure of carbon dioxide in the gas phase	$(MT^{-2}L^{-1})$
p_j	Model parameter	
q	Specific consumption rate	$(MM^{-1}T^{-1})$
q_{AC}	Specific acetate uptake rate	$(MM^{-1}T^{-1})$
$q_{AC,max}$	Maximum specific acetate uptake rate	$(MM^{-1}T^{-1})$
q_{AP}	Specific acetate production rate	$(MM^{-1}T^{-1})$
$q_{AP,max}$	Maximum specific acetate production rate	$(MM^{-1}T^{-1})$
$q_{m,S}$	Part of the specific glucose uptake rate that is not used by the cells for growth	$(MM^{-1}T^{-1})$

$q_{O,max}$	Maximum specific oxygen uptake rate	($MM^{-1}T^{-1}$)
q_{OS}	Specific oxygen uptake rate used for glucose oxidation	($MM^{-1}T^{-1}$)
q_p	Specific product formation rate	($MM^{-1}T^{-1}$)
q_S	Substrate specific consumption rate	($MM^{-1}T^{-1}$)
q_{SF}	Substrate specific consumption rate that follows the fermentative pathway	($MM^{-1}T^{-1}$)
$q_{SF,en}$	Substrate specific consumption rate that follows the fermentative pathway used for energy formation	($MM^{-1}T^{-1}$)
q_{SOX}	Substrate specific consumption rate that follows the fully oxidative pathway	($MM^{-1}T^{-1}$)
r	reaction rates vector	
r_j	Reaction rate of equation j	($ML^{-3}T^{-1}$ or $MM^{-1}T^{-1}$)
s	Standard deviation	
t	Time	(T)
t_0	Initial time for fed-batch fermentation phase	(T)
$y_{C,in}$	Gas phase molar fraction of carbon dioxide in the outflow	(MM^{-1})
$y_{C,out}$	Gas phase molar fraction of carbon dioxide in the outflow	(MM^{-1})
$y_{N,in}$	Gas phase molar fraction of nitrogen in the inflow	(MM^{-1})
$y_{N,out}$	Gas phase molar fraction of nitrogen in the outflow	(MM^{-1})
$y_{O,in}$	Gas phase molar fraction of oxygen in the inflow	(MM^{-1})
$y_{O,out}$	Gas phase molar fraction of oxygen in the outflow	(MM^{-1})

Greek Letters

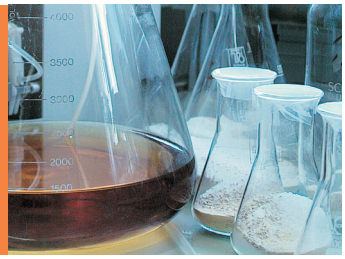
α	Constant	
β	Constant	
ϕ_i	Regressor associated to θ_i, depending on F and Q	
γ	Positive definite estimator gain	
λ	Adaptive controller gain	
μ	Specific growth rate	(T^{-1})
μ_{AC}	Specific growth rate on acetate	(T^{-1})
$\mu_{AC,max}$	Maximum specific growth rate on acetate	(T^{-1})
μ_{crit}	Critical specific growth rate	(T^{-1})

μ_{max}	Maximum specific growth rate	(T^{-1})
μ_S	Specific growth rate on glucose	(T^{-1})
$\mu_{S,max}$	Maximum specific growth rate on glucose	(T^{-1})
μ_{set}	Desired specific growth rate	(T^{-1})
θ_i	Adaptive control parameters (variable functions of the yield coefficients)	
$\hat{\theta}_i$	Estimated adaptive control parameters	
τ	Natural period of oscillation	
$\boldsymbol{\xi}$	State vector composed by the concentrations of the components	
$\xi_{exp,i}$	Experimental concentration of component *i*	(ML^{-3} or MM^{-1})
$\overline{\xi}_{exp,i}$	Average experimental value of the state variable ξ	(ML^{-3} or MM^{-1})
ξ_f	Vector of the state space components that exhibit fast dynamics	(ML^{-3} or MM^{-1})
ξ_i	State variable (concentration of the component *i*)	(ML^{-3} or MM^{-1})
$\xi_{in,i}$	Concentration of the component *i* in the corresponding inflow stream	(ML^{-3} or MM^{-1})
$\xi_{in,i}^{G}$	Concentration of the component *i* in the gaseous inflow	(ML^{-3})
$\xi_{out,i}^{G}$	Concentration of the component *i* in the gaseous outflow	(ML^{-3})
ξ_s	Vector of the state space components that exhibit slow dynamics	
$\xi_{sim,i}$	Simulated concentration of component *i*	(ML^{-3} or MM^{-1})
ζ	Damping coefficient	

Abbreviations

ADP	Adenosine di-phosphate
AK	Acetate kinase
ASCII	American Standard Code for Information Interchange
ATP	Adenosine tri-phosphate
CER	Carbon Dioxide Evolution Rate (per liquid volume or weight)
CTR	Carbon Dioxide Transfer Rate from the liquid to the gas phase (per liquid volume or weight)
DCU	Digital control unit

DNA	Deoxy-ribonucleic acid
DO	Dissolved oxygen
FIA	Flow injection analysis
GA	Genetic algorithms
GDH	Glucose dehydrogenase
GOD	Glucose oxidase
HPLC	High performance liquid chromatography
I/O	Input/output
IPTG	Isopropyl-β-D-thiogalactopyranoside
MIMO	Multiple input / multiple output
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry / mass spectrometer
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide reduced form
OD	Optical density
OTR	Oxygen transfer rate from the gas phase to the liquid phase (per liquid volume or weight)
OUR	Oxygen Uptake Rate (per liquid volume or weight)
PC	Personal Computer
PK/LDH	Pyruvate kinase/lactate dehydrogenase
POD	Glucose peroxidase
rDNA	Recombinant deoxy-ribonucleic acid
RSD	Relative Standard Deviation
SIA	Sequential Injection Analysis
SO	Sarcosine oxidase
TCA	Tricarboxylic acid cycle
tRNA	Transport ribonucleic acid
UV/VIS	Ultra-violet / visible
VI	Virtual Instrument
VISA	Virtual Instrument Software Architecture
WJE	Wall Jet Electrode



CHAPTER 1
INTRODUCTION

"If the study of all these sciences (...), should ever bring us to their mutual association and relationship, and teach us the nature of the ties which bind them together, I believe that the diligent treatment of them will forward the objects which we have in view, and that the labor, which otherwise would be fruitless, will be well bestowed."

Plato.

This thesis addresses the development of a mathematical framework able to describe the main features of the fed-batch fermentation of a recombinant *E. coli* strain, in order to derive optimal and model-based adaptive control approaches to improve heterologous protein production by computer-aided process operation. The applications of mathematical models in biotechnology is not yet a common issue, although it should be expected that a model-based approach would enhance the performance of the production of valuable biotech products

1.1 CONTEXT AND MOTIVATION

Biotechnology has been defined at the United Nations Conference on Environment and Development in 1992 as “a body of techniques that use biological systems, living organisms, or derivatives thereof to make or modify products or processes for specific use”. Agenda 21, the work program adopted by that conference, asserted that biotechnology “promises to make a significant contribution in enabling the development of, for example, better health care, enhanced food security through sustainable agricultural practices, improved supplies of potable water, more efficient industrial development processes for transforming raw materials, support for sustainable methods of aforestation and reforestation, and detoxification of hazardous wastes” (Juma and Konde, 2001).

Also, biotechnology has been considered by the European Union (Commission of the European Communities, 2002) to provide in the next future, a major contribution to achieve the European Community's Lisbon Summit’s objective of becoming a leading knowledge-based economy. In Stockholm in March 2001 the European Council confirmed this and invited the Commission, together with the Council, “to examine measures required to utilize the full potential of biotechnology and strengthen the European biotechnology sector's competitiveness in order to match leading competitors while ensuring that those developments occur in a manner which is healthy and safe for consumers and the environment, and consistent with common fundamental values and ethical principles”.

Indeed, life sciences and biotechnology are widely recognized to be, after information technology, the next wave of the knowledge-based economy, creating new opportunities for societies and economies, and opening up new applications in several fields.

Some of these fields covered by biotechnology include the agro-food area, where biotechnology has the potential to deliver improved food quality and environmental benefits through agronomically improved crops. Biotechnology also has the potential to improve non-food uses of crops as sources of industrial feedstocks or new materials such as biodegradable plastics. New ways to protect and improve the environment are offered by biotechnology including bioremediation of polluted air, soil, water and waste as well as development of cleaner industrial products and processes, e.g. based on the use of enzymes (biocatalysis).

However, one of the most promising applications of biotechnology is related with health care, where there is a huge need for novel and innovative approaches to meet the needs of ageing

populations and poor countries. There are still no known cures for half of the world's diseases, and even existing cures such as antibiotics are becoming less effective due to the increase of resistance to treatments. Biotechnology already enables cheaper, safer and more ethical production of a growing number of traditional as well as new drugs and medical services (e.g. human growth hormone without risk of Creutzfeldt-Jacobs Disease, treatment for haemophiliacs with unlimited sources of coagulation factors free from AIDS and hepatitis C virus, human insulin, and vaccines against hepatitis B and rabies). Stem cell research and xenotransplantation offer the prospect of replacement tissues and organs to treat degenerative diseases and injury resulting from stroke, Altzheimer's and Parkinson's diseases, burns and spinal-cord injuries.

Some estimates suggest that by the year 2005 the European biotechnology market could be worth over € 100 billion. By the end of the decade, global markets, including sectors where life sciences and biotechnology constitute a major portion of new technology applied, could amount to over € 2000 billion. Pharmaceutical biotechnology itself is expected to worth € 818 billion.

Among pharmaceutical applications of biotechnology, the ones that are expected to grow at a higher rate are the recombinant DNA applications. There are already several approved biopharmaceuticals produced with this technology, especially using microbial vectors. Among them, the bacterium *Escherichia coli* (*E. coli*) represents the universal cell-factory for the fermentative production of bio-pharmaceuticals (Sandkvist and Bagdasarian, 1996; Weickert et al., 1996; Hannig and Makrides, 1998, and Baneyx, 1999).

Important aspects in the development and optimisation of these processes are the production yields and product quality and purity. These aspects can be enhanced both at the biochemical and cultivation levels. While the former aspect is a field for physiologists, the latter one is usually a subject for process engineers. However, although the importance of the sector should be an incentive to the optimisation of cultivation processes in pharmaceutical biotechnology, several factors have hampered the development of this field:

- The lack of reliable sensors for the *in situ* monitoring of fermentation processes, and virtually the inexistence of sensors for the on-line monitoring of biological properties (Sonnleitner, 1999 and Scheper et al., 1999).
- The high degree of complexity of bioprocesses, and the non-linear behaviour exhibited by the majority of the fermentation processes.

Together, these factors limit the development of accurate mathematical models, which could serve as the basis for process optimisation and control (Lübert and Jorgensen, 2001 and Lee et al., 1999).

1.2 RESEARCH AIMS

The main purposes of this thesis were the development of a mathematical framework able to describe the main features of the fed-batch fermentation of a **recombinant *E. coli*** strain, in order to derive optimal and **model-based** adaptive control approaches to improve heterologous protein production by **computer-aided process operation**. The dynamical description of the fermentation process was based on the General State-Space Dynamical model proposed by Bastin and Dochain (1990), as well as the adaptive control algorithm developed. For the definition of the kinetic structure of the model, as well as for the determination of the model parameters, a systematic approach based on the application of a stochastic optimisation method using Genetic Algorithms was proposed to approximate real and simulated data. The performance of the mentioned optimisation method was compared with another one, based on gradients, for the determination of an optimal feeding profile for the carbon source.

In order to obtain reliable data for process modelling and to accurately implement the developed control actions, an integrated monitoring and control system was developed, that comprises, besides off-gas analysis and common physical and chemical determinations, a developed Flow Injection Analysis System (Ruzicka and Hansen, 1987) for on-line analysis of the carbon source – glucose – and the major metabolic by-product – acetate.

The proposed modelling technique, as well as optimal and adaptive control methods and the developed monitoring and control system are intended to be applicable not only to the studied microorganism, but also to a vast class of bioprocesses.

1.3 OUTLINE OF THE THESIS

This thesis is organized in seven chapters that cover the research aims stated above. The thesis subjects are introduced in this chapter, while chapter 7 respects to the main

conclusions and recommendations extracted from the current work. In the other chapters, the research fields are covered as follows:

- In chapter 2, a brief description of the fermentation process is given, in what metabolic, genetic and physiological characteristics concerns. The main advantages and pitfalls related with the utilization of *E. coli* to produce recombinant proteins are enumerated and the current methods used to grow it until high-cell densities are described.
- In chapter 3, the developed on-line Flow Injection Analysis method for the determination of acetate and glucose during the fermentation of *E. coli* is described. For acetate measurement, a physical / chemical method was chosen, while an enzymatic methodology was used to determine glucose concentration. The optimisation of both methods for the desired purpose is described, and the individual characteristics such as method detection level, linearity limit and interferences are discussed.
- The main equipment used to monitor and control the fed-batch fermentation is described in chapter 4. The Mass Spectrometry technique was used to analyze the bioreactor's off-gas, and the digital control unit of the fermenter controlled the main physical variables of the fermentation. The software developed to acquire and control the relevant data from the fermentation is also described in this chapter. The development tool used was the LabVIEW environment, which allowed a modular and flexible data acquisition and control software to be developed.
- In chapter 5, the mathematical model developed is presented. The structure of the model is based on mass balances to the main state variables of the process, namely biomass, glucose, acetate, oxygen and carbon dioxide concentrations. The structure of the process kinetics and the model parameters are identified by using a stochastic optimisation method: Genetic Algorithms. The assumptions made for model derivation are explained and the sensitivity of the model outputs to the individual parameters is assessed.
- The application of optimisation methods to the determination of the optimal feeding profile of glucose is evaluated in chapter 6. Two optimisation methods are compared. The optimal feeding profile was used to simulate a fermentation and the corresponding results are evaluated. Also, an adaptive control law for regulating the concentration of acetate in the medium is presented, and both simulated and real results of its implementation are shown.

1.4 REFERENCES

Baneyx, F. Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology.* 10, 411-421. 1999.

Bastin, G. and Dochain, D. *On-line estimation and adaptive control of bioreactors*. Elsevier Science Publishers, Amsterdam. 1990.

Comission of the European Comunities. *Life sciences and biotechnology - a strategy for Europe*. European Comission, Brussels. 2002.

Hannig, G. and Makrides, S. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends in Biotechnology.* 16, 54-60. 1998.

Juma, C. and Konde, V. *The new bioeconomy - Industrial and environmental biotechnology in developing countries*. United Nations Conference on Trade and Development, 2001.

Lee, J., Lee, S. Y., Park, S., and Middelberg, A. Control of fed-batch fermentations. *Biotechnology Advances.* 17, 29-48. 1999.

Lübert, A. and Jorgensen, S. B. Bioreactor performance: a more scientific approach for practice. *Journal of Biotechnology.* 85, 187-212. 2001.

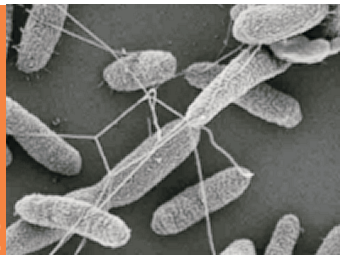
Ruzicka, J. and Hansen, E. H. *Flow Injection Analysis*. John Wiley & Sons, New York. 1987.

Sandkvist, M. and Bagdasarian, M. Secretion of recombinant proteins by Gram negative bacteria. *Current Opinion in Biotechnology.* 7, 505-511. 1996.

Scheper, T., Hitzmann, B., Stärk, E., Ulber, R., Faurie, R., Sosnitza, P., and Reardon, K. F. Bioanalytics: detailed insight into bioprocesses. *Analytica Chimica Acta.* 400, 121-134. 1999.

Sonnleitner, B. Instrumentation of biotechnological processes. *Advances in Biochemical Engineering / Biotechnology.* 66, 1-64. 1999.

Weickert, M., Doherty, D., Best, E., and Olins, P. Optimization of heterologous protein production in *Escherichia coli*. *Current Opinion in Biotechnology.* 7, 494-499. 1996.



CHAPTER 2
PRODUCTION OF RECOMBINANT PROTEINS WITH *ESCHERICHIA COLI* - SOME IMPORTANT FEATURES

"Once we understand the biology of Escherichia coli, we will understand the biology of an elephant".

Jacque Monod.

In this chapter, the bacterium *Escherichia coli* is briefly characterized with respect to basic metabolic and genetic characteristics. A special emphasis is given to the fermentative metabolism, as its main product, acetate, constitutes one of the major problems when trying to achieve a high productivity in the production of recombinant products with this microorganism. The most common cultivation methods used to grow *E. coli* and produce recombinant proteins are revised and their advantages and disadvantages are discussed.

2.1 INTRODUCTION

Following the discovery in the late 1960s of the tools for manipulating DNA sequences *in vitro*, the expression of genes in microbial cells has been established as a key technique to produce, via fermentation, bioactive proteins in an amount and quality that, previously, was difficult or impossible to achieve by their isolation from natural sources.

In 1982 human insulin was the first marked human health care product derived from the recombinant DNA (rDNA) technology. The commercialization of this technology increased greatly until the present, and human insulin now accounts for approximately 70% of all insulin used by diabetic patients in the USA (Balgi et al., 1999).

Other successful examples of therapeutic products produced with this technology include blood proteins like erythropoietin; human hormones like epidermal growth factor; immune modulators like interferons and several vaccines.

However, the successful commercialization of these proteins from an industrial perspective requires a genetically stable recombinant culture, a high-productivity fermentation process, and cost-effective recovery and purification procedures.

Since the beginning, the bacterium *E. coli* has become the host organism of choice for the expression of many heterologous genes, primarily due to the ease with which it may be manipulated genetically, its rapid growth rate, high-cell densities attainable, simple nutritional requirements, completed genomes sequence, and well established metabolic pathways (Blight and Holland, 1994 and Yang et al., 1998). Although *E. coli* cannot be used to produce some large or complex proteins, there are many other highly valuable proteins like interferons, interleukins, or hormones that can be produced using this microorganism.

However, it has also some drawbacks (Akesson et al., 1999a) that make it necessary to accurately control the fermentation conditions of this microorganism in order to improve their efficiency and performance by obtaining high cell concentrations and high specific product levels. Such goals are crucial for the economical and efficient use of recombinant microorganisms.

In the next sections a brief characterization of this microorganism concerning physiological, genetic and metabolic aspects is presented as well as cultivation methodologies employed in order to allow its growth to high-cell densities levels.

2.2 PHYSIOLOGICAL AND METABOLIC ASPECTS

E. coli was originally described in 1885 by Theodore Escherich, a German bacteriologist, and named *bacterium coli commune.* It is a typical Gram negative bacterium, possessing an inner membrane (cytoplasmic membrane) surrounded by a murein wall and by an outer membrane. The space between both membranes is called periplasm and significantly differs from the cytosol in terms of composition. It is a facultative anaerobic, short, and straight bacillus. As a mesophile it can grow at temperatures ranging from 8 to 48°C with a maximum growth rate at 39ºC at a minimum water activity of 0.95. It is fairly acid tolerant and can grow at pH values ranging from 4.4 to 10. A more detailed information about *E. coli* physiology can be found in Neidhardt et al. (1990).

A photograph and a scheme of an *E. coli* cell are shown in Figure 2.1.

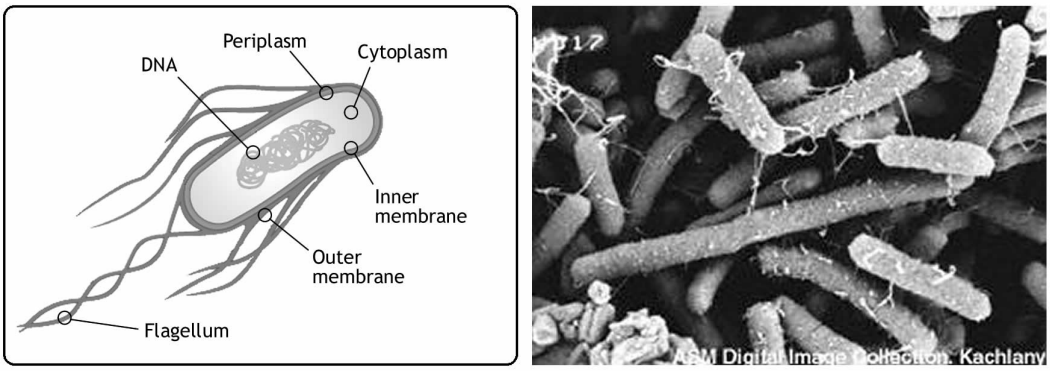


Figure 2.1 Scheme and electronic microscopic photography of an *E. coli* cell.

It follows an overview about the main features of *E. coli* metabolism, with the purpose of generating the tools for a better understanding of this microorganism's behaviour during fermentations. Extensive descriptions about *E. coli* metabolism can be found for example in Gottschalk (1986). Several databases, containing most of the compiled knowledge about this microorganism, can be found in the internet. For example the EcoCyc database[1] is a bioinformatics database that describes the genome and the biochemical machinery of *E. coli* and is funded by a grant from the American National Institute of Health. A less extensive

[1] http://biocyc.org/ecocyc/

approach can also be found for example in Paalme et al. (1997a) or as an Annex of Varma and Palsson (1993).

E. coli is able to grow with different substrates in the presence of oxygen or in its absence. Glucose is the most commonly used substrate since it can be easily metabolized to provide carbon and energy for biosynthesis and also due to its low price.

2.2.1 Oxidation of Glucose

Under certain aerobic conditions, part of the glucose is oxidized to carbon dioxide with oxygen as the terminal electron acceptor. This process is exergonic and allows the formation of ATP, which is required for the biosynthesis of cellular constituents. If glucose is the substrate, about 50% is oxidized to carbon dioxide; this results in enough ATP to convert the other 50% into cell material.

The main reactions for oxidation of glucose are illustrated in Figure **2.2** and this process can be divided into a number of functional blocks of reactions:

A. Transport of glucose into the cell by the phosphoenolpyruvate phosphotransferase system. The cytoplasmic membrane of *E. coli* is not permeable to glucose. Instead, it possesses a transport system that picks up glucose at the medium. This transport process is coupled to the phosphorilation of glucose to glucose-6-phosphate.

B. Degradation of glucose-6-phosphate to pyruvate via glycolysis or Embden-Meyerhof-Parnas pathway. In the process precursor metabolites are also produced, 3-phosphoglycerate, phosphoenolpyruvate and pyruvate that are used for the biosynthetic reactions. Net metabolic energy is also generated by these reactions.

C. Oxidative decarboxylation of pyruvate to acetyl-coenzyme A (acetyl-CoA) by pyruvate dehydrogenase;

D. Oxidation of the acetyl moiety of acetyl-CoA to carbon dioxide via the tricarboxylic acid (TCA) cycle. It generates energy and reducing power. Intermediates to the TCA cycle such as oxaloacetate and α-ketoglutarate are consumed in the production of macromolecules. Replenishment of these intermediates is the function of other pathways. During growth on glucose, there are several pathways that accomplish the generation of TCA cycle intermediates.

E. Oxidation of the reduced coenzymes formed in steps B to D in the respiratory chain. Respiration is an ATP-generating process in which compounds act as electron

donors through a chain of electron transfer to electron acceptors. Aerobic respiration uses oxygen as the final acceptor. During the chain of electron transfer, protons (H^+) are transported outside the cytoplasmic membrane, generating a proton motive force. Upon passage of protons back into the cytoplasm, the proton motive force energy is captured as ATP, catalyzed by a multi-subunit ATPase.

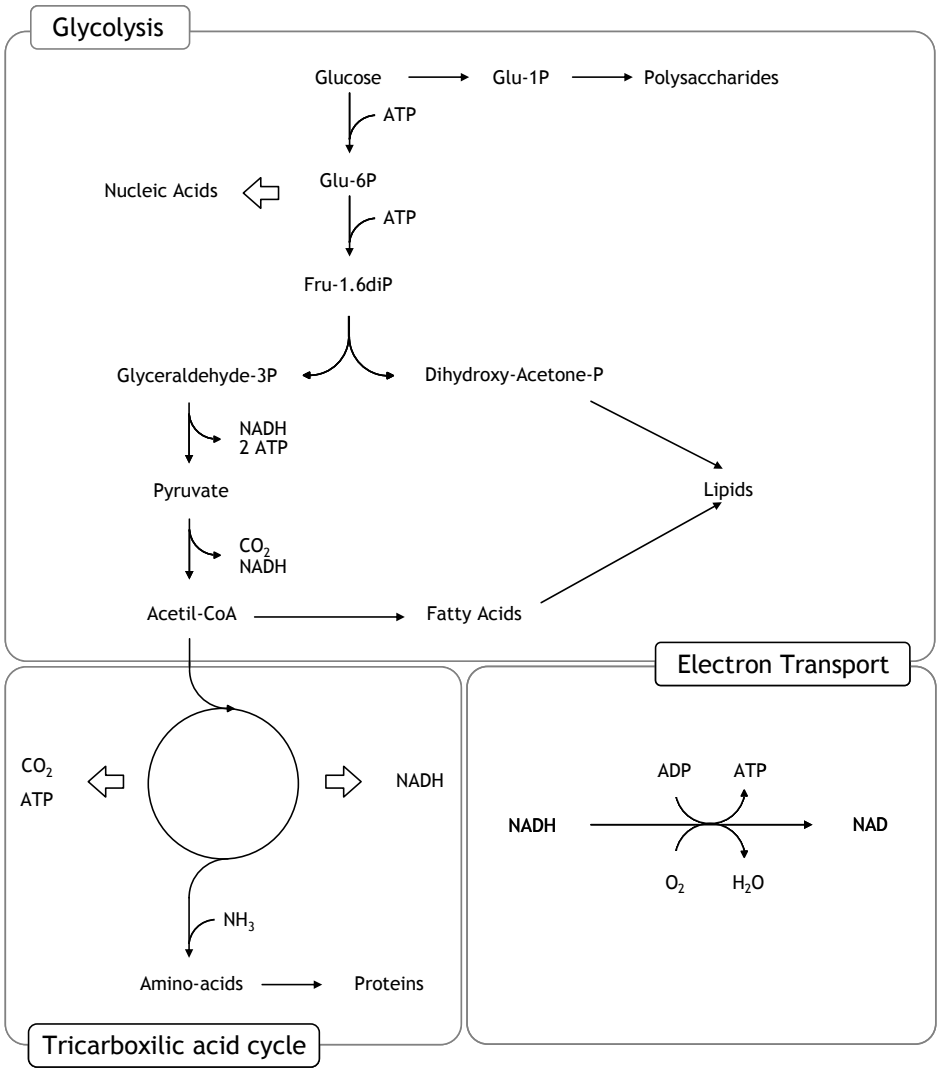


Figure **2.2** Main steps in the oxidation of glucose. A brief description of the process can be found in the text (steps A to D). For simplifying the scheme, both A and C steps were included in the glycolysis step.

The net reaction of glucose oxidation in a molar basis is presented on Equation 2.1. It should be noticed that the stoichiometry of this reaction ignores the utilization of glucose for biosynthetic pathways.

$$\text{Glucose} + \text{Oxygen} \longrightarrow \text{Carbon Dioxide} + \text{Water} + \text{Energy}$$
$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O + \text{Energy} \qquad \text{Equation 2.1}$$

2.2.2 Fermentation of Glucose

When glucose is in excess or under anaerobic conditions, glucose can follow a fermentative pathway after glycolysis. It is converted into acidic metabolites like succinate, formate, acetate, lactate, ethanol and hydrogen gas (Blight and Holland, 1994). Amongst these fermentation products, acetic acid is by far the most commonly found during *E. coli* fermentations.

During acetic acid production, flux of acetyl-CoA is directed to acetic acid, via acetylphosphate, instead of entering the TCA cycle. By two subsequent enzymatic reactions, catalyzed by acetyl-CoA:orthophosphatase acetyltransferase and acetate kinase, one acetic acid is converted from an acetyl-CoA. Concomitantly with this conversion, one ATP is formed. So, from the maximal energy conservation point of view, aerobic acetic acid production yields 4 ATP molecules per each glucose molecule consumed, when compared with 2 ATP molecules per each glucose molecule for ethanol production, if the anabolic use of glucose is not considered. Also, the major difference between acetic acid synthesis and the synthesis of other partially oxidized glucose metabolites like ethanol or lactic acid is the absence of the NADH oxidation step in the acetic acid synthesis pathway. The NADH is oxidized during the conversion of pyruvate to ethanol or lactic acid, while NADH is generated during the conversion of pyruvate to acetic acid. Therefore, acetic acid formation requires the NADH oxidation step and so the consumption of oxygen.

The main reactions involved in *E. coli* fermentation are illustrated in Figure **2.3**.

The net reaction of the fermentation of glucose into acetate is as follows, once again ignoring the anabolic utilization of the sugar:

$$\text{Glucose} + \text{Oxygen} \longrightarrow \text{Acetate} + \text{Carbon Dioxide} + \text{Water} + \text{Energy}$$
$$C_6H_{12}O_6 + 2O_2 \longrightarrow 2CH_3COOH + 2CO_2 + 2H_2O + \text{Energy} \qquad \text{Equation 2.2}$$

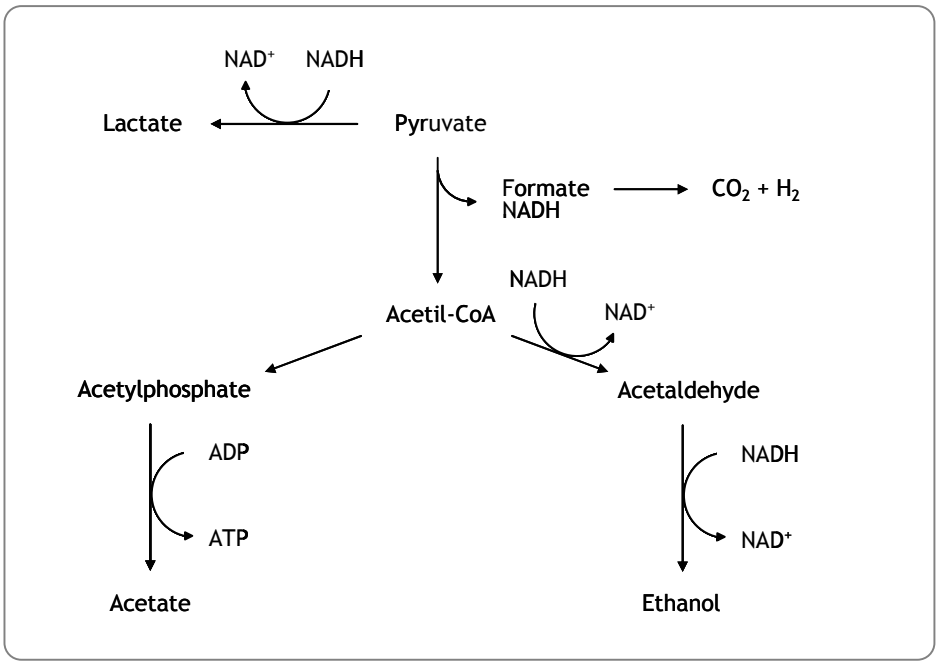


Figure **2.3** Main possible reactions during the fermentation of glucose by *E. coli*.

Nowadays, a huge amount of information is available concerning aerobic acetic acid formation in *E. coli* fermentation. Together with its counterpart phenomenon in yeast, *i.e.*, the aerobic ethanol production (Sonnleitner and Käppeli, 1986), this process is now known as glucose overflow metabolism. High specific growth rate (Alba and Calvo, 2000), high specific glucose uptake rate, high dissolved carbon dioxide concentration (Wipf et al., 1994), and glucose concentration above 1-2 $g \cdot L^{-1}$ (Aristidou et al., 1999), have been suggested to trigger the acetic acid overflow metabolism. Although the last condition can be easily found in the literature, it is more likely that it is a high glucose flow into the cells or the correspondent high specific growth rate that induces acetic acid production, rather than an absolute value of the glucose concentration. The threshold of specific growth or glucose uptake rates, known as critical specific growth or glucose uptake rates, depends on the strain, culture conditions, and medium compositions (Han et al., 1992).

In the origin of this phenomenon can be a lack of regulation in the maximum glucose uptake velocity by the phosphotransferase system (described in step A of section 2.2.1) in *E. coli* at high glucose concentrations. Consequently, the flux of glucose into the central metabolic pathways produces more acetyl-CoA than the amount that can be oxidatively used for

biosynthesis and energy generation. However, it is still not unanimous where this limitation specifically occurs.

One possible explanation suggests that the respiratory system, where NADH is reoxidized, has a limited capacity. As the flux to the TCA cycle results in a significant NADH production, and the flux to acetic acid does not, redirection of acetyl-CoA flux to acetic acid would be necessary to avoid accumulation of NADH when the respiration saturates (Akesson et al., 1998). Another explanation that has been suggested is that the TCA cycle has a limited capacity and that this limitation is reached before that of the respiration. When the TCA cycle saturates, increasing glucose uptake will again result in flux from acetyl-CoA to acetic acid. In this case, NADH production and respiration can increase further until the maximum respiration capacity or the maximum glucose uptake is reached (Xu et al., 1999).

If the production of acetate is due to a limited respiratory capacity, one of the visible effects of this overflow mechanism should be, besides production of acetic acid, a stabilization of oxygen uptake rate in its maximum value when glucose uptake rate reaches its critical or threshold value. In the work of Akesson (Akesson et al., 1998; Akesson et al., 1999a and Akesson et al., 1999b) not only that phenomenon was detected but also used for the construction of a control algorithm with the aim of maintaining acetate at low levels concentrations. However, as it was already mentioned, production of acetate results in the production of NADH, which has to be oxidized with the concomitant consumption of oxygen in the respiratory chain. Therefore NADH production is much lower than during total glucose oxidation (two molecules of oxygen per glucose molecule during acetate formation when compared to six molecules produced during complete oxidation) and consequently the increase in oxygen consumption for specific growth rates above the critical value should be smaller. Han et al. (1992) verified experimentally this phenomenon and supported the theory of limitations in the TCA cycle. According to these authors, the TCA cycle is sufficient at low specific growth rates to meet both anabolic and catabolic requirements, but at high dilution rates, this does not happen. Under these conditions, it is proposed that *E. coli* reorganizes the oxidative metabolism to ensure first that the anabolic requirement is satisfied with the oxidative metabolism, while the catabolic demand is met using both the remaining oxidative metabolism capacity and the acetic acid production metabolism. According to this proposal, the acetic acid formation could be reduced by either lowering the anabolic demand or by enhancing the maximum capacity of the oxidative metabolism.

In conclusion, although *E. coli* is nowadays one of the most studied microorganisms, still there is not a unanimous justification for one of the most important phenomena observed

during its growth. However, evidence suggests that a limitation in the TCA cycle is the most likely justification for the aerobic formation of acetic acid.

Whatever the justification behind this fact, it is clear that the production of acetic acid in aerobic fermentations of *E. coli* constitutes one of the major problems avoiding the achievement of high yield and high volumetric productivity.

Accumulation of acetic acid in the culture medium has been shown to inhibit several physiological properties of the culture itself, and the reported effects include reduction of maximum specific growth rate (Luli and Strohl, 1990 and Turner et al., 1994), glucose uptake rate (Rhee et al., 1997), recombinant protein production (Jensen and Carlsen, 1990; Tomson et al., 1995; Akesson et al., 1999a; Aristidou et al., 1999 and Gschaedler et al., 1999) and growth yield (Rinas et al., 1989) or a combination of the above mentioned effects. Depending on the strain and growth conditions, the threshold of acetic acid concentration for those effects can vary from 5 (Luli and Strohl, 1990) to 15 $g \cdot L^{-1}$ (Konstantinov et al., 1990).

There is an evidence suggesting that acetic acid acts as an uncoupler of the proton motive force of *E. coli* (Luli and Strohl, 1990 and Xu et al., 1999). After it is formed, acetic acid exists in the fermentation medium at the neutral pH in most fermentation processes involving *E. coli* in both the ionized (CH_3COO^-) and protonated (CH_3COOH) states. The protonated form can pass through the lipid membrane into the interior of the cell, where it dissociates at the higher internal pH (*ca*. 7.5) to CH_3COO^- and H^+, thereby decreasing the intracellular pH. When the protonated acid moves across the membrane into the cells, additional protonated acid is formed in the medium by the equilibrium, causing a net electroneutral hydrogen ion influx. The overall external pH would not change drastically because of the large volume of buffered medium, and the decrease in intracellular pH would cause an uncoupling effect.

2.2.3 Oxidation of Acetic Acid

If the acetic acid concentration is not at inhibitory levels and if the TCA cycle capacity is not fully in use by glucose, it is used as a secondary carbon source prolonging the growth of the cell (Alba and Calvo, 2000). This theory is in good agreement with the TCA overflow hypothesis and implies that acetate cannot be produced and consumed at the same time. However, this fact is not unanimously accepted. Shiloach et al. (1996) and Van de Walle and Shiloach (1998) explain their experimental data on a particular strain of *E. coli* (BL21) with a simultaneous consumption and production of acetate. According to these authors, acetate concentration is controlled by *E. coli* cells around a certain level by an adequate balance of

its consumption / production. However, this theory is not supported by the majority of the published results and the observations made by those authors may also be justified by the overflow theory.

In fed-batch cultures, acetic acid uptake system is present in glucose-grown cells; thus, it is not catabolite repressed and cells can grow in a mixture of glucose and acetic acid, as long as the oxidative capacity is not saturated. According to Xu et al. (1999), in batch cultures glucose represses cellular use of exogenous acetic acid, which is consumed only after glucose has been exhausted. It has also been reported (Luli and Strohl, 1990) that not every *E. coli* strain is capable of assimilating acetic acid.

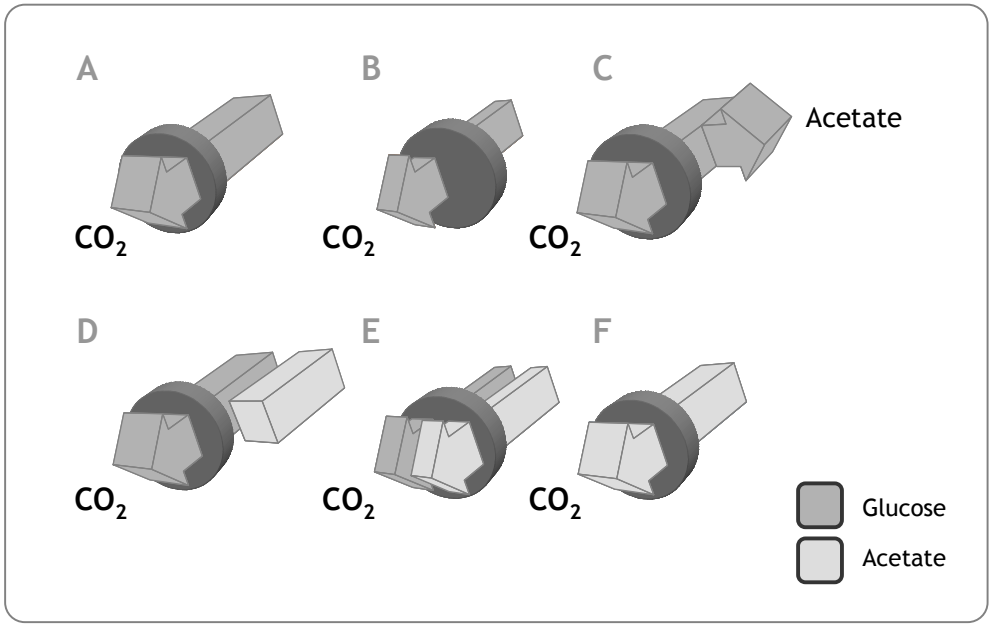


Figure **2.4** Schematic representation of the oxidative bottleneck of *E. coli*. In A, B and C there is only glucose in the medium. In case A, bottleneck is full with glucose, while in B there is still some free capacity and in C glucose exceeds the oxidative capacity, leading to the production of acetate. In case D, oxidative capacity is full with glucose and consequently acetate in the medium is not oxidized, while the opposite happens in E. In case F only acetate is present in the medium and is fully oxidized.

Acetic acid is converted back to acetyl-CoA, which then enters the aerobic energy metabolism at phase D of glucose oxidation mentioned in section 2.2.1 and represented on Figure **2.2**. During growth only on acetic acid, the TCA intermediates are generated by the glyoxalate shunt, which has to be activated under these circumstances.

The net reaction for the oxidation of acetic acid is:

$$\text{Acetate} + \text{Oxygen} \longrightarrow \text{Carbon Dioxide} + \text{Water}$$
$$CH_3COOH + 2O_2 \longrightarrow 2CO_2 + 2H_2O \qquad \text{Equation 2.3}$$

An illustration of glucose overflow under aerobic conditions is illustrated in Figure **2.4**, as well as acetic acid consumption when the oxidative capacity is not saturated with glucose.

2.3 GENETIC ASPECTS

According to the American National Institute of Health, the rDNA technology consists generally in the construction of DNA molecules outside of living cells by joining natural or synthetic DNA segments to DNA molecules that can replicate in a living cell, or molecules that result from their replication. The process of *in vitro* recombination makes it possible to cut different strands of DNA with a restriction enzyme and join the DNA molecules together *via* complementary base pairing. Figure **2.5** represents the main steps involved in rDNA process. A detailed description of these phases, as well as of other aspects related with rDNA technology can be found, for example, in Walker and Gingold (1993) and in Glick and Pasternak (1998).

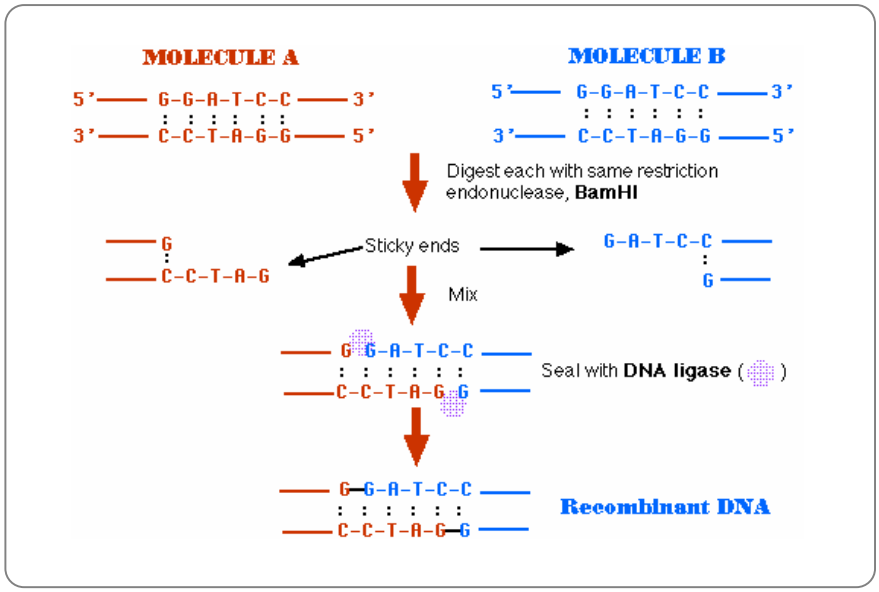


Figure **2.5** Pictorial representation of the recombinant DNA process. (Adapted from http://www.unh.edu/ehs/BS/).

One of the basic strategies of molecular cloning is to move the desired genes from a large, complex genome to a small simple one. A cloning vector is a DNA molecule that carries foreign DNA into a host cell (usually bacterial or yeast), where it replicates, producing many copies of itself along with the foreign DNA.

There are several types of cloning vectors of which the most commonly used are plasmids. They are small, circular, extra-chromosomal DNA molecules found naturally in bacteria, which can replicate on their own, outside of a host cell. They have a cloning limit of 100 to 10,000 base pairs. A plasmid vector is made from natural plasmids by removing unnecessary segments and adding essential sequences. It contains a set of genetic elements that affect protein production, like the promoter that allows the regulation of the gene expression, transcriptor terminals and ribosome binding sites.

Factors regulating the expression of foreign genes in a specific host-vector system include the strength of transcriptional promoters, the strength of translational binding to the ribosome, plasmid stability, plasmid copy number (Ryan and Parulekar, 1991 and Patnaik, 2002), plasmid size, the stability of expressed foreign protein in the host, among others (Ryan and Parulekar, 1991 and Chen et al., 1992). There have been numerous reviews about the effects of these aspects in the robustness of the expression system (Zabriskie and Arcuri, 1986; Georgiou, 1988; Weickert et al., 1996; Hannig and Makrides, 1998 and Baneyx, 1999).

Amongst those factors, the stability of the plasmids is one of the most important features in recombinant fermentation processes. A microorganism containing a recombinant plasmid is compelled to assign a fraction of its limited resources for maintenance and replication of the plasmid. It has been reported (Ryan et al., 1989; Jensen and Carlsen, 1990; Park et al., 1990 and Patnaik, 1998) that the presence of certain plasmids constitutes a metabolic burden that results in slower growth of plasmid-bearing cells than that of plasmid-free cells under comparable conditions, especially in the case of high expression plasmids and in continuous cultures (Weber and San, 1987 and Ryan and Parulekar 1990) particularly at high dilution rates (Nancib and Boudrant, 1992). Owing to the growth advantage they have over plasmid-bearing cells, plasmid-free cells, once formed due to segregational instability during cell division (Kim et al., 1992a), outgrow plasmid bearing cells (Rhee et al., 1997) and eventually take over the entire population, causing a decline in productivity of the plasmid-encoded product (Ryan and Parulekar, 1991). Contradictory results were found about the increase (Hopkins et al., 1987) or decrease (Lin and Neubauer, 2000) of this effect in large-scale bioreactors.

The introduction of an antibiotic resistance gene in the plasmid facilitates selection of the plasmid-bearing cells by growing recombinant cells in selective medium by inhibiting the growth of plasmid-free cells. The most common antibiotics used for the cultivation of recombinant proteins are kanamicin and ampicilin. Hasenwinkle et al. (1997) found that the insertion of a kanamicin resistance gene in the place of ampicilin resistance gene imparted a greater stability to the plasmid.

Usually, recombinant microorganisms are more unstable under expressed condition than under the repressed condition for the cloned gene. One strategy to deal with the bioprocesses involving these recombinants is to separate the growth stage from the production stage and control the level of the cloned gene expression through a proper genetic switch. In order to separate growth and production phases, thermal and chemically inducible promoters can be used. Isopropyl-β-D-thiogalactopyranoside (IPTG) is a non-hydrolyzable lactose analog and thus constitutes an effective inducer of the powerful *lac*-derived promoters like *tac*, and *trc* promoters (Fu et al., 1993). In large scale, however, this inducer is not ideal because of its high cost, although this is not generally a problem when producing high-added-value products. A more serious issue concerns IPTG toxicity. This can be circumvented by utilizing lactose as an inducer or by using thermo-sensitive promoters. These temperature-inducible expression systems have been intensively used (Lee et al., 1988; Park et al., 1990 and Strandberg and Enfors, 1991), despite the fact that the increase of the temperature necessary to induce the expression system also amplifies the activity of proteases. Other promoters exist and, in the already mentioned reviews (Weickert et al., 1996 and Hannig and Makrides, 1998), it can be found a examination of the most common. Recently, Schroeckh et al. (1999) described a new promoter based on the nitrogen concentration that is easily applicable to controlled high-cell density cultures. In recent years, the pET vectors (commercialized by Novagen, USA), have gained increasing popularity. They allow a very high productivity and are often incorporated in specific strains containing a prophage (λDE3), like the strain BL21(DE3). A description of this production system can be found in Baneyx (1999).

When the expression system is optimised by the appropriate choice of antibiotic resistance and promoter, *E. coli* has the ability to accumulate many proteins up to 80% of the total soluble protein (Nancib and Boudrant, 1992). However, as the techniques for the over-expression of recombinant proteins became more sophisticated, a further difficulty became apparent: the formation of intracellular inclusion bodies. The recombinant *E. coli* cells produce the protein in the form of biologically inactive, dense, insoluble inclusion bodies (Blight and Holland, 1994 and Balgi et al., 1999), that need to be converted to the active

proteins by costly denaturation and refolding processes (Jeong and Lee, 1999). The inclusion bodies are separated either by breaking open the cellular membranes or by dissolution into appropriate extracellular medium. Carrió and Villaverde (2002) made a recent review on the conditions of development, and strategies to minimize formation of these aggregates. For this reason, certain tags, such as a polyhistidine tail, β-galactosidase (Feliu et al., 1998), maltose binding proteins, or a cellulose binding domain (Hasenwinkle et al., 1997) are commonly joined to recombinant proteins. Those purification tags can provide not only for localization of the target protein to preferred compartments in the host cell (Blight et al., 1994), for example in the periplasm (Tokugawa et al., 1994), cell wall, or culture medium, making them easily separable from cell extracts, but also a means for protection against proteolytic degradation (Park et al., 1999), a means for increasing the solubility of the product, a means for rapidly recovering the active gene product, and for the formation of disulfide bounds (Sako, 1985). Sandkvist and Bagdasarian (1996) have reviewed the current advancement in the molecular mechanisms by which proteins are secreted to the extracellular medium by Gram-negative bacteria, while Hockney (1994) has reviewed methods for protein folding and secretion.

Other authors optimised the secretion of recombinant proteins by manipulating the culture and medium conditions (Jensen and Carlsen, 1990; Yang et al., 1998 and Yazdani and Mukherjee, 1998) and or by using purification methods (Hirata et al., 1997).

Georgiou and Valax (1996) have reviewed relevant techniques for the expression of correctly folded complex proteins in *E. coli*. Kurland and Gallant (1996) have summarised the available information concerning the occurrence of errors in the translation of heterologous protein expression.

2.4 CULTIVATION MODES

In the production of recombinant proteins, not only a high expression of the target protein is important, but also a final high-cell density is vital for the economic viability of the production process. When compared with a low density culture, a high-cell density culture has many advantages, such as higher product concentration, increased productivity, decreased production costs, enhanced downstream processing, reduced wastewater, and reduced investment in equipment (Lee, 1996). Also, it is important when the recombinant proteins are produced after an induction step, because the bacteria cannot be reused after induction. Therefore, it is the preferred mode of operation in industrial application.

Due to the negative effects mentioned in section 2.2.2, the main problem in achieving high-cell density cultures during the cultivation of recombinant *E. coli* is acetate accumulation. It was already discussed that the conditions that trigger acetate formation are mainly oxygen limitation and glucose overflow. Thus, it is necessary to avoid both situations if a high-cell density is required.

There are three alternatives to grow *E. coli* cells: continuous, batch and semi-continuous or fed-batch cultivation. Mainly, these methods differ in the way nutrients are supplied to the culture: in batch all nutrients except for oxygen are in the fermentation vessel at the time of inoculation; in continuous cultures the medium is constantly added to the culture and the same amount of fermentation broth is removed, while in fed-batch there is an addition of certain nutrients along the fermentation but there is no removal of broth. In Figure **2.6** there is a schematic description of these modes of operation.

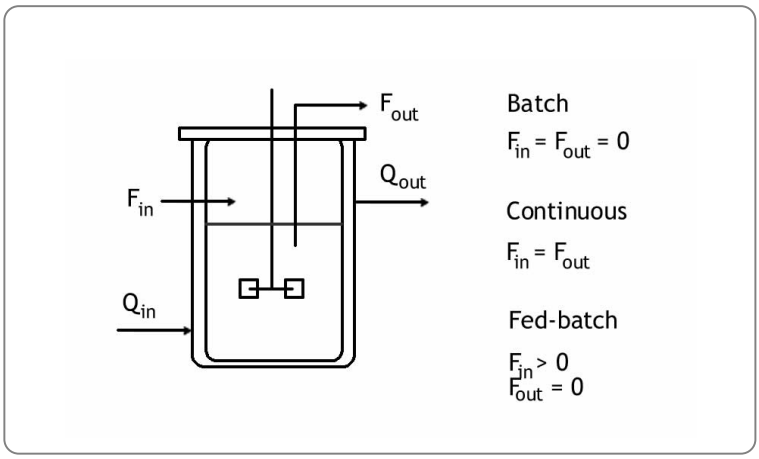


Figure **2.6** Differences between fed-batch, continuous and batch modes of fermentation.

Although it is very versatile in terms of the imposition of a desired specific growth rate, the continuous cultivation method is very time-consuming, as the recommended stabilization time to achieve a steady-state (chemostat) is at least three generation times (Paalme et al., 1995). Another disadvantage is that plasmid instability is very high, and the probabilities of contamination are significant. Yet, continuous culture is still an important tool for the characterization of the strains as well as of its production characteristics. For example, in Flickinger and Rouse (1993), a continuous culture was used to study protein synthesis during very slow growth.

Additionally, for inducible systems, the continuous culture has to be done in two stages, one for bacterial growth and another for protein production. This approach increases the possibility of contamination besides being very complex in terms of setup, since it requires two fermentation systems. Moreover, in this case, not all cells leaving the vessel have experienced an equal residence time. This distribution of residence times could result in a distribution of expression levels within the population since cells would spend different times in the induction stage.

Nevertheless, there are several works using two-stage continuous cultures. For example, in Fu et al. (1993), a stable continuous overproduction of β-lactamase for 50 days using the strain *E. coli* K-12 has been performed in two-stage chemostats. The second stage culture was continuously induced and the final production achieved was 300 $mg \cdot L^{-1}$. In Kim et al. (1992b), production of human epidermal growth factor is also described in a two-stage chemostat.

Park et al. (1990) have chosen a temperature-inducible production system to evaluate the effect of the growth rate on the performance of a two-stage continuous culture of recombinant *E. coli*. In this system, the first stage was operated at temperatures lower than 37ºC for cell growth without cloned-gene expression, while the second stage was maintained at temperatures higher than 38ºC to inactivate the heat-sensitive repressor molecules and, hence, to induce the biosynthesis and accumulation of the cloned gene product. Also in Ryu and Seigel (1990), a temperature-inducible system was used in two-stage continuous cultures of *E. coli*.

Regarding batch cultivation, it is still the most widespread method in industrial processes of microbial synthesis due to its simplicity of operation and to the virtual absence of contamination. However, the dry cell weights of traditional batch fermentations of *E. coli* are only 1-2 $g \cdot L^{-1}$ with a maximum achievable of 12$g \cdot L^{-1}$ (Tomson et al., 1995) because of nutrient limitation and/or growth inhibition.

Over the past few years, the fed-batch mode has been introduced in an increasing number of fermentation processes in order to prevent the accumulation of by-products allowing the achievement of high density of biomass within a short cultivation time. Additionally, it allows easier maintenance of sterility than continuous cultures.

Nutrient inhibition and by-product formation can be avoided in fed-batch mode due to the possibility of restricting the growth rate, thus also avoiding limitations in oxygen transfer as well as heat transfer (Strandberg and Enfors, 1991).

Finally, theoretical analysis has revealed that the use of fed-batch fermentation when compared with batch fermentation implies an increase in productivity and a concomitant decrease in product manufacturing cost (Longobardi, 1994).

Fed-batch fermentations have also gained popularity because of the improved control possibilities using computer-coupled fermentation systems. The use of computers to monitor and control fermentation processes has, in recent years, become a commonplace. The power of modern computers allows the more development of sophisticated control schemes that use multiple growth parameters and mathematical models to control nutrient addition to fed-batch cultures.

In fed-batch cultures, the nutrient feeding is either predetermined (feed-forward control) or is feed-back controlled. Two types of feed-back-controlled fed-batch fermentations have been introduced (Agawal et al., 1989). These are classified as indirect feed-back controlled and direct feed-back controlled fed-batch fermentations. In the fed-batch operation with indirect feed-back control, parameters as dissolved oxygen, pH, respiratory quotient, or partial pressure of carbon dioxide in the exhaust gas are used to automatically regulate the feed rate of the nutrients. In the second type of feed-back-controlled fermentation, the concentration of the limiting substrate in the culture medium is directly used as a feed-back parameter. An extended revision of fed-batch modes and theory can be found in Yamanè and Shimizu (1984) and Lee et al. (1999).

Different nutrient feeding strategies, either with or without feed-back control, have been employed to obtain high-cell densities. At present and depending on the strain, high-cell densities greater than 100 $g \cdot L^{-1}$ can be achieved for recombinant *E. coli* producing different proteins. It follows a brief description of the most common types of fed-batch fermentation used to grow recombinant *E. coli* to high-cell densities.

2.4.1 Exponential Feeding

Fed-batch culture with exponential feeding allows cells to grow at constant specific growth rates, and acetate formation can be minimized by controlling the specific growth rate below the critical value of acetate formation. The theoretical deduction of the feeding equation is as follows.

During fed-batch operating, the substrate concentration S follows from the feeding rate F_{in} and the specific substrate consumption rate q_S:

$$\frac{dS}{dt} = \frac{F_{in}}{V}(S_{in} - S) - q_S X$$ Equation 2.4

where S_{in} is the substrate concentration in the inflow, V is the reactor volume, and X is the biomass concentration. Specific substrate consumption rate depends on microbial specific growth rate (μ):

$$q_S = \frac{\mu}{k_{XS}}$$ Equation 2.5

where k_{XS} the global yield coefficient defined as the ratio of grams of biomass produced per grams of substrate consumed.

In the fed-batch phase the cultivation is operated under glucose limitation. Therefore, *S=0* and *dS/dt=0* and the last equations' rearrangement gives:

$$F_{in}S_{in} = \left(\frac{\mu_{set}}{k_{XS}}\right)VX$$ Equation 2.6

where μ_{set} is the desired specific growth rate.

Starting the fed-batch phase at t_0 with a cell density X_0 and cultivation volume V_0, exponential growth with the desired specific growth rate (μ_{set}) in the variable volume system gives:

$$VX = V_0 X_0 e^{\mu_{set}(t-t_0)}$$ Equation 2.7

Hence, the variable substrate feeding rate for constant growth rate in the fed-batch phase is given by:

$$F_{in}S_{in} = \left(\frac{\mu_{set}}{k_{XS}}\right)V_0 X_0 e^{\mu_{set}(t-t_0)}$$ Equation 2.8

This exponential regimen is usually used in an open loop mode, keeping the specific growth rate at a pre-determined value using a feed-forward approach. This has been used, for example, by Yang (1992) and Martinez et al. (1998). Using this same methodology, Hellmuth et al. (1994) obtained a product concentration 10-fold and a productivity 20-fold higher than in continuous runs. Paalme et al. (1989) used a similar feeding equation, although biomass concentration in a certain instant is not calculated from the initial value but from the total amount of glucose fed to the culture in a certain moment.

This open-loop exponential approach is probably one of the most widely used methods to grow cells to high densities because of the ease of implementation. However, the inherent features of a feed-forward method limit the application of this feeding scheme.

For example, Gregory and Turner (1993) found a discrepancy between the predicted specific growth rate and the real profile. The main errors they found using Equation 2.8 were related with the fact that maintenance was not contemplated in the equation and erroneous estimates of initial biomass concentration were used. Thus, some authors include the maintenance coefficient in the feeding equation (Hellmuth et al., 1994).

In Korz et al. (1995), the authors found that volume estimation is the main source of errors when applying this approach to a laboratory scale reactor. In fact, Equation 2.8 implies that the cultivation volume only depends on the feeding rate, which is not always true with lab-scale reactors, where sample volume, sampling frequency and ammonia addition are also important for volume variations. In that work, V was calculated from the weights of base and samples, and the corresponding densities.

Also, the application of this methodology may not adequately control bioprocesses because model parameters, such as the yield coefficients and even the critical specific growth rate, must be given to compute the feeding rate and these values may change as the cultivation proceeds.

Therefore, the exponential feeding strategy is sometimes compensated for by incorporating some appropriate feed-back control. In Lee et al. (1997), the authors developed a simple method to adjust feeding rate based on the calculation of the actual specific growth rate, and correcting the feed rate using a proportional action. A similar approach was proposed by Agawal et al. (1989). Fieschko and Ritch (1986) applied an exponential rate by feeding 1 $mL \cdot h^{-1}$ per gram of biomass and recalculating the biomass concentration at defined intervals.

Another feature of the majority of the exponential feeding methods, either with feed-back control or not, is the utilization of step changes in the feed rate to simulate the desired exponential profile. A drawback of the step changed feeding is that the cells might experience repeated cycles of large-extent excessive-to-insufficient nutritive environments. As a result, inhibitory effects, such as formation of acetic acid, might be induced during over-dosed period, and cell concentration as well as cell productivity would accelerate to fall behind the expected profiles as the fermentation proceeds. Cheng et al. (2002) used a two-mixing chambers system to vary smoothly the feed rate, keeping F_{in} constant while varying S_{in}.

2.4.2 pH-stat

The pH change in the culture broth may reflect the level of glucose concentration in the aerobic cultivation of *E. coli*. The overflow of carbon flux mentioned in section 2.2.2 channels the metabolic pathways of cells to the fermentative pathway to excrete acetate and other organic acids which lower the pH in the culture broth. In contrast, glucose deprivation makes the cells catabolize organic acids or amino acids as carbon or energy sources, which raises the pH in the culture broth. Thus, the decrease in the pH in the culture broth corresponds to abundant glucose, while the increase in pH corresponds to a lack of glucose in the culture broth. On the other hand, the rate of pH change can be employed to estimate the residual glucose concentration in the culture broth so that the feed rate of glucose can be adjusted for optimal cell growth. The mode of cultivation that uses this data for controlling acetate concentration is called pH-stat.

This approach has been successfully used by Choi and Lee (1999) for the production of P(3HB-co-3HV) – a biodegradable polymer – in a chemically defined medium. In Wang and Lee (1997) and Lee and Wang (1998), the authors use this strategy to produce another biopolymer: PHB.

Production of human leptin, a protein that influences body weight, was done using the pH-stat in Jeong and Lee (1999). A nutrient feeding solution was added when the pH rose to a value greater than its set point (pH 6.8) by 0.08 units due to the depletion of glucose. The appropriate volume of the feeding solution was added to increase the glucose concentration in the broth to 0.7 $g \cdot L^{-1}$.

The main disadvantage with the pH-stat approach is that it doesn't follow the exponential needs of the culture, and eventually the specific growth rate decreases as the fermentation proceeds. However, there are methods to overcome this problem. For example, Shin et al. (1997) increased stepwise the volumetric media feed rate during the feeding-on period as the cell concentration increased during the pH-stat production of mini-proinsulin with *E. coli*.

Another problem with this methodology is that intermittently the cell produces acetate that is consumed when glucose is exhausted. Although the levels of production are not inhibitory to the cell, the global biomass yield is reduced when compared with the situation where glucose is only oxidized.

2.4.3 DO-stat

The DO-stat technique aims the control of the dissolved oxygen (DO) by manipulating the feeding addition. In the literature, a great number of papers can be found where DO is used as a feed-back parameter. This fact can be explained probably due to the existence of reliable DO electrode in any fermentation facility. Also, that method is very simple and inexpensive when compared with other fed-batch techniques.

A general operation of this methodology is: if the DO increases slightly owing to a decrease in substrate in the medium, the controller starts the addition of the feeding solution, and the substrate increases with a resulting decrease in the DO value. When that value reaches a pre-imposed minimum, glucose addition is stopped.

This methodology was used, for example, by Gleiser and Bauer (1981) in *E. coli* fermentation and by Yano et al. (1991), applied to other microorganisms.

However, Konstantinov et al. (1990) found that using this approach, acetate production was still significant (15 $g \cdot L^{-1}$). This was due to glucose overfeed. These authors proposed a new fed-batch method: the balanced DO-stat, with two interconnected loops, one for the DO, controlled by the agitation speed, and the other for the control of the agitation speed, that was kept at a fixed set point by manipulating the glucose feed rate. These authors were able to obtain high cell and product concentrations without production of acetate.

Still, the main disadvantage of these methods is that, by keeping constant the agitation speed, it does not follow the needs of an exponentially-growing culture, and so the specific growth rate is expected to decrease during the fermentation.

Other more sophisticated methodologies based on the DO-stat approach are reported in the literature. For example, in Shi and Shimizu (1992), a neuro-fuzzy control was used to recognize the patterns of the change in DO concentration to calculate the feeding rate in *Saccharomyces cerevisiae* and recombinant *E. coli*.

2.4.4 Other Feed-back Techniques

Application of feed-back techniques other than pH-stat, DO-stat and feed-back controlled exponential rate exist in the literature.

As an example of direct feed-back fermentation mode, in Kleman and Strohl (1994), a glucose feed-back fed-batch fermentation system was used to control glucose at 0.5 $g \cdot L^{-1}$ in the medium. On-line analysis of glucose was performed with a biochemical analyser.

In a large scale (1000 L) fed-batch system for the production of interferon-α-2a, glucose concentration was controlled by estimating its consumption from the integration of the off-gas carbon dioxide concentration (Wipf et al., 1994). A particularity of this work is that the authors claim that producing therapeutics where glucose is fed in limiting quantities can lead to glucose starvation and subsequent induction of proteases. Thus, the limitation of growth rate by lowering the temperature to 25ºC is used by these authors.

Feed-back control of feeding rate by on-line measurements of biomass using a laser turbidimeter was used by Yamanè et al. (1992).

2.4.5 A-stat

The A-stat technique is a combination of continuous and fed-batch cultures that was described in Paalme and Vilu (1992) and Paalme et al. (1995). Its main application in what recombinant *E. coli* concerns, like in the case of continuous cultures, is not the achievement of high cellular and product concentrations but the characterization of the kinetic and production capabilities of a particular strain.

It is basically a continuous culture with smooth change of desired growth rate. At first, like in a chemostat, a steady-state culture is obtained. After that, the computer controlled smooth change of dilution rate, keeping its time derivative constant, is started and continued up to almost complete washout. This technique showed to be a powerful technique for the quantitative study of cell physiology, being at the same time considerably less time consuming and more informative than the conventional chemostat. Also, cultures seem to react better to a smooth rather than an abrupt change in the dilution rate.

In Paalme et al. (1995), this approach is used to study the growth characteristics of *E. coli* K-12, while in Paalme et al. (1997b), the effects of ethanol/glucose ratio on the growth characteristics of *Saccharomyces cerevisiae* were studied using the A-stat technique.

2.4.6 Pre- and Post-induction Strategies

After induction, growth and metabolic activities of the host cells are often influenced by the expression of recombinant proteins in the cell. In certain cases, IPTG, often used as an inducer, could further complicate the metabolism of *E. coli*. Therefore, growth rates and consequently nutrient feeding methods during the pre-induction and the post-induction phases are very important and subjected to studies. According to Hellmuth et al. (1994), the growth rate during production is also important because of favourable intracellular conditions for folding of proteins that can occur at some specific growth rates.

For example, in Andersson et al. (1996), induction with 0.1 mM IPTG led to a marked reduction in growth. Induction caused a decline in the number of dividing cells to less than 10% of the population within 10 h.

In Wong et al. (1998), it was found that post-induction cell growth rate was independent of nutrient feeding rate but feeding profile influenced protein production. Linearly changing post-induction feeding rate with a suitable slope allowed the highest production obtained.

A 3-stage feeding was used by Wangsa-Wirawan et al. (1997) to produce insulin-like growth factor 1 from the strain JM101. The initial stage consisted of an exponential feeding to give a constant specific growth rate of 0.25 h^{-1} prior to induction. Immediately following induction, a linearly-decreasing feed rate was used for 0.5 h, followed by exponential feeding at a specific growth rate of 0.1 h^{-1}.

Lim and Jung (1998) found that with a temperature-inducible promoter, sustaining the specific growth rate above a critical level after induction was a key factor for achieving a high productivity. Using an exponential feeding for cell growth and a constant feeding after temperature induction, these authors achieved a high concentration of interferon-γ.

Regarding pre-induction growth rate, Curless et al. (1990) showed that when a low growth exists in the moment of induction, the production of recombinant consensus interferon in *E. coli* K-12 is low.

2.4.7 Cyclic Fed-batch

When one of the previous techniques is used to increase the *E. coli* growth until high-cell densities, induction is usually conducted at the end of the growth phase in the same bioreactor.

However, that is not the only possibility. The two-stage, cyclic fed-batch process is a fed-batch process where some volume is periodically removed from the culture vessel and transferred to an induction tank while leaving some cell broth in the growth tank. The volume of liquid in the growth tank can then be replenished to its pre-transfer volume while induction is taking place in the production tank. The specific growth rate in the growth tank can be kept constant using an exponential feeding. In addition, the feed substrate concentration of the growth tank can be selected in order that the desired cell density is maintained. Conditions in the production tank, such as pH, temperature, and feed media composition, can be operated independently from conditions in the growth tank. This can be advantageous since optimal growth conditions are not necessarily optimal production conditions. In Curless et al. (1991), the authors proved that a two-stage fed-batch is twice as productive as a single fed-batch fermentation. They applied this approach to the production of lymphokine with K-12 *E. coli.*

Choi et al. (2001) used a two-stage fed-batch process for the production of human granulocyte-colony stimulating factor. They optimised the pre-induction growth rate and the feeding strategy during the production stage.

2.5 EXAMPLES OF HIGH-CELL DENSITY CULTURES

The development of the fed-batch techniques described in the previous section had led to efficient, high-level production of various proteins and non-protein products. Table **2.1** represents a non-exhaustive compilation of some of the most relevant production systems described in the literature. A much more extensive revision can be found for example in Lee (1996).

2.6 INTERFERON ALFA-2B

The interferons are a family of naturally occurring small proteins of molecular weights of approximately 15,000 to 27,600 daltons that belong to the cytokines family. They have strong antiviral, anti-proliferative and immuno-modulatory activities and are produced by cells in small quantities in response to viral infections and to synthetic or biological inducers (Müller et al., 1994). An explanation of the mechanism of actuation of interferons can be found in Darnell Jr. et al. (1994).

Table **2.1** Examples of high-cell density and high productivity cultures of recombinant *E. coli*.

Strain	Cellular Concentration	Product	Product Concentration	Cultivation method	Reference
XL1-Blue	157 $g \cdot L^{-1}$	P(3HB)	157 $g \cdot L^{-1}$	pH-stat	Wang and Lee (1997)
BL21(DE3)	52 $g \cdot L^{-1}$	Human leptin	9.7 $g \cdot L^{-1}$	pH-stat	Jeong and Lee (1999)
BL21(DE3)	70 $g \cdot L^{-1}$	Mini-proinsulin	7 $g \cdot L^{-1}$	pH-stat with exponential correction	Shin et al. (1997)
JM103	72 $g \cdot L^{-1}$	β-galactosidase	4150 $U \cdot L^{-1}$	pH-stat and fuzzy control	Jin et al. (1994)
TOPP5	72 $g \cdot L^{-1}$	Cellulose Binding Domain	8.2 $g \cdot L^{-1}$	DO-stat	Hasenwinkle et al. (1997)
AT2741	36 $g \cdot L^{-1}$	Phenylalanine	24 $g \cdot L^{-1}$	Balanced DO-stat	Konstantinov et al. (1990)
M5248	29 $g \cdot L^{-1}$	Interleukin-2	1.20 $g \cdot L^{-1}$	DO-stat	Seo et al. (1992)
TG1	100 $g \cdot L^{-1}$	Fused aprotinin::β-galactosidase	2000 $U \cdot mL^{-1}$	Exponential	Hellmuth et al. (1994)
X90	92 $g \cdot L^{-1}$	Trypsin	56 $mg \cdot L^{-1}$	Exponential	Yee and Blanch (1993)
AM-7	70 $g \cdot L^{-1}$	Alpha consensus interferon	5.5 $g \cdot L^{-1}$	Exponential	Fieschko and Ritch (1986)

Thus, the potential medical application of these compounds is obvious and, as soon as the rDNA technology made it possible to produce large amounts of these proteins, several therapeutic applications were tested and approved, while many others are still in the trials phase.

Although there are nowadays many attempts to produce interferons using a variety of expression systems, including mammal cells, yeasts or fungi, these proteins are still mainly produced by *E. coli*, because they do not require any post-translational modification. Indeed, only one interferon molecule requires post-translational modifications, but even though, its non-glycolized form is equally biologically active.

There are two types of interferon molecules that can be distinguished by their biological activity. Type I includes the subtypes α and β, while type II is constituted by the subtype γ. The human interferons α (IFN-α) include more than 15 different molecules, being IFN-α-2a and IFN-α-2b the most studied ones. This last one in particular is produced by Schering Corporation (USA) from *E. coli* and is commercialized since 1986 under the name

INTRON®A for patients with hairy cell leukaemia, malignant melanoma, follicular lymphoma, condylomata acuminate, AIDS-related Kaposi's sarcoma and chronic hepatitis B and C. It has 165 amino acids residues and a molecular weight of 19,271 daltons.

A schematic representation of this protein can be found in Figure **2**.**7**.

Figure **2**.**7** Schematic representation of the interferon-α-2b (from RasWin molecular graphics).

Production of interferons in *E. coli* has been widely studied, and productivity was enhanced either by lowering the temperature (Mizukami et al., 1986) and recycling the biomass (Ohtaguchi et al., 1992) during the production of IFN-β, or by optimising the feeding profile (Fieschko and Ritch, 1986) and cultivating *E. coli* under the DO-stat method (Riesenberg et al., 1990) for the production of IFN-α. However, many of some new and promising strains or fermentation techniques developed in the past years were still not applied for the production of interferons in *E. coli*, while a robust mathematical model that would allow optimisation and characterization of this recombinant process can not be found in the literature.

2.7 REFERENCES

Agawal, P., Koshy, G., and Ramseier, M. An algorithm for operating a fed-batch fermentor at optimum specific-growth rate. *Biotechnology and Bioengineering.* 33, 115-125. 1989.

Akesson, M., Hagander, P., and Axelsson, J. P. A probing feeding strategy for *Escherichia coli* cultures. *Biotechnology Techniques.* 13, 523-528. 1999a.

Akesson, M., Karlsson, E. N., Hagander, P., Axelsson, J. P., and Tocaj, A. On-line detection of acetate formation in *Escherichia coli* cultures using dissolved oxygen responses to feed transients. *Biotechnology and Bioengineering.* 64, 590-598. 1999b.

Akesson, M., Tocaj, A., Hagander, P., and Axelsson, J. P. Acetate formation and dissolved oxygen responses to feed transients in *Escherichia coli* fermentations: modeling and experiments. *Horizon of BioProcess Systems Engineering in 21st Century (7th Int. Conf. Computer Applications on Biotechnology, Osaka, Japan).* (Yoshida, T. and Shioya, S., Eds.). Pergamon, Oxford. 107-111. 1998.

Alba, M. J. G. and Calvo, E. G. Characterization of bioreaction processes: aerobic *Escherichia coli* cultures. *Journal of Biotechnology.* 84, 107-118. 2000.

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

Aristidou, A., San, K., and Bennett, G. Improvement of biomass yield and recombinant gene expression in *Escherichia coli* by using fructose as the primary carbon source. *Biotechnology Progress.* 15, 140-145. 1999.

Balgi, G., Reynolds, J., Mayer, R. H., Cooley, R. E., and Sevickmuraca, E. M. Measurements of multiply scattered-light for online monitoring of changes in size distribution of cell debris suspension. *Biotechnology Progress.* 15, 1106-1114. 1999.

Baneyx, F. Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology.* 10, 411-421. 1999.

Blight, M. A., Chervaux, C., and Holland, B. Protein secretion pathways in *Escherichia coli*. *Current Opinion in Biotechnology.* 5, 468-474. 1994.

Blight, M. A. and Holland, I. B. Heterologous protein secretion and the versatile *Escherichia coli* haemolysin translocator. *Trends in Biotechnology.* 12, 450-455. 1994.

Carrió, M. M. and Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. *Journal of Biotechnology.* 96, 3-12. 2002.

Chen, H.-C., Hwang, C.-F., and Mou, D.-G. High-density *Escherichia coli* cultivation process for hyperexpression of recombinant porcine growth hormone. *Enzyme and Microbial Technology.* 14, 321-326. 1992.

Cheng, L.-C., Wu, J.-Y., and Chen, T.-L. A pseudo-exponential feeding method for control of specific growth rate in fed-batch cultures. *Biochemical Engineering Journal.* 10, 232. 2002.

Choi, J. I. and Lee, S. Y. High-level production of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) by fed-batch culture of recombinant *Escherichia coli*. *Applied and Environmental Microbiology.* 65, 4363-4368. 1999.

Choi, S. J., Park, D. H., and Jung, K.-H. Development and optimization of two-stage cyclic fed-batch culture for hG-CSF production using L-arabinose promoter in *Escherichia coli*. *Bioprocess and Biosystems Engineering.* 24, 51-58. 2001.

Curless, C., Fu, K., Swank, R., Menjares, A., Fieschko, J., and Tsai, L. Design and evaluation of a two-stage cyclic, recombinant fermentation process. *Biotechnology and Bioengineering.* 38, 1082-1090. 1991.

Curless, C., Pope, J., and Tsai, L. Effect of preinduction specific growth rate on recombinant alpha consensus interferon synthesis in *Escherichia coli*. *Biotechnology Progress.* 6, 149-152. 1990.

Darnell Jr., J. E., Kerr, I. M., and Stark, G. R. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science.* 264, 1415-1421. 1994.

Feliu, J. X., Cubarsi, R., and Villaverde, A. Optimized release of recombinant proteins by ultrasonication of *E. coli* cells. *Biotechnology and Bioengineering.* 58, 536-540. 1998.

Fieschko, J. and Ritch, T. Production of human alpha consensus interferon in recombinant *Escherichia coli*. *Chemical Engineering Communications.* 45, 229-240. 1986.

Flickinger, M. and Rouse, M. P. Sustaining protein synthesis in the absence of rapid cell division: an investigation of plasmid-encoded protein expression in *Escherichia coli* during very slow growth. *Biotechnology Progress.* 9, 555-572. 1993.

Fu, J., Wilson, D. B., and Shuler, M. L. Continuous, high level production and excretion of a plasmid-encoded protein by *Escherichia coli* in a two-stage chemostat. *Biotechnology and Bioengineering.* 41, 937-946. 1993.

Georgiou, G. Optimizing the production of recombinant proteins in microorganisms. *AIChE Journal.* 34, 1233-1248. 1988.

Georgiou, G. and Valax, P. Expression of correctly folded proteins in *Escherichia coli*. *Current Opinion in Biotechnology.* 7, 190-197. 1996.

Gleiser, I. E. and Bauer, S. Growth of *E. coli* W to high cell concentration by oxygen level linked control of carbon source concentration. *Biotechnology and Bioengineering.* 23, 1015-1021. 1981.

Glick, R. and Pasternak, J. J. *Molecular Biotechnology*. American Society of Microbiology, Washinton DC. 1998.

Gottschalk, G. *Bacterial Metabolism*. Springer Verlag, New York. 1986.

Gregory, M. E. and Turner, C. Open-loop control of specific growth rate in fed-batch cultures of recombinant *E. coli*. *Biotechnology Techniques.* 7, 889-894. 1993.

Gschaedler, A., Robas, N, Boudrant, J, and Branlant, C. Effects of pulse addition of carbon sources on continuous cultivation of *Escherichia coli* containing a recombinant *E. coli gapA* gene. *Biotechnology and Bioengineering.* 63, 712-720. 1999.

Han, K., Lim, H. C., and Hong, J. Acetic acid formation in *Escherichia coli* fermentation. *Biotechnology and Bioengineering.* 39, 663-671. 1992.

Hannig, G. and Makrides, S. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends in Biotechnology.* 16, 54-60. 1998.

Hasenwinkle, D., Jervis, E., Kops, O., Liu, C., Lesnicki, G., Haynes, C. A., and Kilburn, D. G. Very high-level production and export in *Escherichia coli* of a cellulose binding domain for use in a generic secretion-affinity fusion system. *Biotechnology and Bioengineering.* 55, 854-863. 1997.

Hellmuth, K., Korz, D. J., Sanders, E. A., and Deckwer, W.-D. Effect of growth rate on stability and gene expression of recombinant plasmids during continuous and high cell density cultivation of *Escherichia coli* TG1. *Journal of Biotechnology.* 32, 289-298. 1994.

Hirata, R. D. C., Hirata, M. H., Levy, D. E., and Nguyen, N. Y. Optimized production of the soluble human interferon α receptor (IFNAR) expressed in *E. coli*. *Biotechnology Techniques.* 11, 301-305. 1997.

Hockney, R. C. Recent developments in heterologous protein production in *Escherichia coli*. *Trends in Biotechnology.* 12, 456-463. 1994.

Hopkins, D. J., Betenbaugh, M. J., and Dhurjati, P. Effects of dissolved oxygen shock on the stability of recombinant *Escherichia coli* containing plasmid pKN401. *Biotechnology and Bioengineering.* 29, 85-91. 1987.

Jensen, E. B. and Carlsen, S. 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, 1-11. 1990.

Jeong, K. J. and Lee, S. Y. High-level production of human leptin by fed-batch cultivation of recombinant *Escherichia coli* and its purification. *Applied and Environmental Microbiology.* 65, 3027-3032. 1999.

Jin, S., Ye, K., and Shimizu, K. Efficient Fuzzy control strategies for the application of pH-stat to fed-batch cultivation of genetically engineered *Escherichia coli*. *Journal of Chemical Technology and Biotechnology.* 61, 273-281. 1994.

Kim, B.-S., Choi, C. Y., and Kim, B. G. Plasmid stability in a recombinant mammalian cell bioprocess. *Biotechnology Letters.* 14, 351-356. 1992a.

Kim, E., Fu, J, Wilson, D. B., and Shuler, M. L. Expression of human epidermal growth factor by *Escherichia coli* in continuous culture. *Biotechnology Letters.* 14, 339-344. 1992b.

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

Konstantinov, K., Kishimoto, M., Seki, T., and Yoshida, T. A balanced DO-stat and its application to the control of acetic acid excretion by recombinant *Escherichia coli*. *Biotechnology and Bioengineering.* 36, 750-758. 1990.

Korz, D. J., Rinas, U., Hellmuth, K., Sanders, E. A., and Deckwer, W.-D. Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. *Journal of Biotechnology.* 39, 59-65. 1995.

Kurland, C. and Gallant, J. Errors of heterologous protein expression. *Current Opinion in Biotechnology.* 7, 489-493. 1996.

Lee, J., Lee, S. Y., and Park, S. Fed-batch culture of *Escherichia coli* W by exponential feeding of sucrose as a carbon source. *Biotechnology Techniques.* 11, 59-62. 1997.

Lee, J., Lee, S. Y., Park, S., and Middelberg, A. Control of fed-batch fermentations. *Biotechnology Advances.* 17, 29-48. 1999.

Lee, S. B., Ryu, D. D. Y., Seigel, R., and Park, S. H. Performance of recombinant fermentation and evaluation of gene expression efficiency for gene product in two-stage continuous culture. *Biotechnology and Bioengineering.* 31, 805-820. 1988.

Lee, S. Y. High cell-density culture of *Escherichia coli*. *Trends in Biotechnology.* 14, 98-105. 1996.

Lee, S. Y. and Wang, F. High cell density culture of metabolically engineered *Escherichia coli* for the production of poly(3-hydroxybutyrate) in a defined medium. *Biotechnology and Bioengineering.* 58, 325-328. 1998.

Lim, H.-K. and Jung, K.-H. Improvement of heterologous protein productivity by controlling postinduction specific growth rate in recombinant *Escherichia coli* under control of the P_L promoter. *Biotechnology Progress.* 14, 548-553. 1998.

Lin, H. Y. and Neubauer, P. Influence of controlled glucose oscillations on a fed-batch process of recombinant *Escherichia coli*. *Journal of Biotechnology.* 79, 27-37. 2000.

Longobardi, G. P. Fed-batch versus batch fermentation. *Bioprocess Engineering.* 10, 185-194. 1994.

Luli, G. W. and Strohl, W. R. Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Applied and Environmental Microbiology.* 56, 1004-1011. 1990.

Martinez, A., Ramirez, O., and Valle, F. Effect of growth rate on the production of B-galactosidase from *Escherichia coli* in *Bacillus subtilis* using glucose-limited exponentially fedbatch cultures. *Enzyme and Microbial Technology.* 22, 520-526. 1998.

Mizukami, T., Komatsu, Y., Hosoi, N., Itoh, S., and Oka, T. Production of active human interferon-β in *E. coli.* I. Preferential production by lower culture temperature. *Biotechnology Letters.* 8, 605-610. 1986.

Müller, U., Steinhoff, U., Reis, L. F. L., Hemmi, S., Pavlovic, J., Zinkernagel, R. M., and Aguet, M. Functional role of type I and type II interferons in antiviral defense. *Science.* 264, 1918-1924. 1994.

Nancib, N. and Boudrant, J. Effect of growth rate on stability and gene expression of a recombinant plasmid during continuous culture of *Escherichia coli* in a non-selective medium. *Biotechnology Letters.* 14, 643-648. 1992.

Neidhardt, F. C., Ingraham, J. L., and Schaechter, M. *Physiology of the bacterial cell - a molecular approach*. Sinauer Associates, Sunderland. 1990.

Ohtaguchi, K., Sato, H., Hirooka, M., and Koide, K. A high density cell cultivation of *Escherichia coli* for the production of recombinant human interferon-β. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 375-378. 1992.

Paalme, T., Elken, R., Kahru, A., Vanatalu, K., and Vilu, R. The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach. *Antoine van Leeuwenhoek.* 71, 217-230. 1997a.

Paalme, T., Elken, R., Vilu, R., and Korhola, M. Growth efficiency of *Saccharomyces cerevisiae* on glucose/ethanol media with a smooth change in the dilution rate (A-stat). *Enzyme and Microbial Technology.* 20, 174-181. 1997b.

Paalme, T., Kahru, A., Elken, R., Vanatalu, K., Tiisma, K., and Vilu, R. The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat). *Journal of Microbiological Methods.* 24, 145-153. 1995.

Paalme, T., Tiisma, K., Kahru, A., Vanatalu, K., and Vilu, R. Glucose-limited fed-batch cultivation of *Escherichia coli* with computer-controlled fixed growth rate. *Biotechnology and Bioengineering.* 35, 312-319. 1989.

Paalme, T. and Vilu, R. A new method of continuous cultivation with computer-controlled change of dilution rate. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 299-301. 1992.

Park, S. H., Ryu, D. D. Y., and Kim, J. Y. Effect of cell growth rate on the performance of a two-stage continuous culture system in a recombinant *Escherichia coli* fermentation. *Biotechnology and Bioengineering.* 36, 493-505. 1990.

Park, S. J., Georgiou, G., and Lee, S. Y. Secretory production of recombinant protein by a high cell density culture of a protease negative mutant *Escherichia coli* strain. *Biotechnology Progress.* 15, 164-167. 1999.

Patnaik, P. R. Dispersion-induced behavior in sub-critical operation of a recombinant fed-batch fermentation with run-away plasmids. *Bioprocess Engineering.* 18, 219-226. 1998.

Patnaik, P. R. Comparative evaluation of batch and fed-batch bioreactors for GAPDH production by recombinant *Escherichia coli* with distributed plasmid copy number. *Chemical Engineering Journal.* 87, 357-366. 2002.

Rhee, J. I., Bode, J., Diaz-Ricci, J. C., Poock, D., Weigel, B., Kretzmer, G., and Schügerl, K. Influence of the medium composition and plasmid combination on the growth of recombinant *Escherichia coli* JM109 and on the production of the fusion protein *Eco*RI::SPA. *Journal of Biotechnology.* 55, 69-83. 1997.

Riesenberg, D., Menzel, K., Schulz, V., Schumann, K., Veith, G., Zuber, G., and Knorre, W. A. High cell density fermentation of recombinant *Escherichia coli* expressing human interferon alpha 1. *Applied Microbiology and Biotechnology.* 34, 77-82. 1990.

Rinas, U., Kracke-Helm, H.-A., and Schügerl, K. Glucose as a substrate in recombinant strain fermentation technology. *Applied Microbiology and Biotechnology.* 31, 163-167. 1989.

Ryan, W. and Parulekar, S. J. Effects of culture conditions on plasmid stability and production of a plasmid-encoded protein in batch and continuous cultures of *Escherichia coli* JM103[pUC8]. *Annals of New York Academy of Sciences.* 589, 91-110. 1990.

Ryan, W. and Parulekar, S. J. Recombinant protein synthesis and plasmid instability in continuous cultures of *Escherichia coli* JM103 harboring a high copy number plasmid. *Biotechnology and Bioengineering.* 37, 415-429. 1991.

Ryan, W., Parulekar, S. J., and Stark, B. C. Expression of B-lactamase by recombinant *Escherichia coli* strains containing plasmids of different sizes - effects of PH, phosphate, and dissolved oxygen. *Biotechnology and Bioengineering.* 34, 309-319. 1989.

Ryu, D. D. Y. and Seigel, R. Scale-up of fermentation processes using recombinant microorganisms. *Annals of New York Academy of Sciences.* 73-82. 1990.

Sako, T. Overproduction of staphylokinase in *Escherichia coli* and its characterization. *European Journal of Biochemistry.* 149, 557-563. 1985.

Sandkvist, M. and Bagdasarian, M. Secretion of recombinant proteins by Gram negative bacteria. *Current Opinion in Biotechnology.* 7, 505-511. 1996.

Schroeckh, V., Wenderoth, R., Kujau, M., Knüpfer, U., and Riesenberg, D. The use of elements of the *E. coli* Ntr-system for the design of an optimized recombinant expression system for high cell density cultivations. *Journal of Biotechnology.* 75, 241-250. 1999.

Seo, Jin, D., Chung, B. H., Hwang, Y. B., and Park, Y. H. Glucose-limited fed-batch culture of *Escherichia coli* for production of recombinant human interleukin-2 with the DO-stat method. *Journal of Fermentation and Bioengineering.* 74, 196-198. 1992.

Shi, Z. P. and Shimizu, K. Fuzzy control of fed-batch fermentation with the aid of neural networks. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 167-172. 1992.

Shiloach, J., Kaufman, J., Guillard, A. S., and Fass, R. Effect of glucose supply strategy on acetate accumulation, growth, and recombinant protein production by *Escherichia coli* BL21 (λDE3) and *Escherichia coli* JM109. *Biotechnology and Bioengineering.* 49, 421-428. 1996.

Shin, C. S., Honh, M. S., Bae, C. S., and Lee, J. Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of *Escherichia coli* BL21 (DE3) [pET-3aT2M2]. *Biotechnology Progress.* 13, 249-257. 1997.

Sonnleitner, B. and Käppeli, O. Growth of *Saccharomyces cerevisiae* is controlled by its limited respiratory capacity: formulation and verification of a hypothesis. *Biotechnology and Bioengineering.* 28, 927-937. 1986.

Strandberg, L. and Enfors, S.-O. Batch and fed batch cultivations for the temperature induced production of a recombinant protein in *Escherichia coli*. *Biotechnology Letters.* 13, 609-614. 1991.

Tokugawa, K., Ishii, T., Nakamura, K., Masaki, H., and Uozumi, T. A model system for the continuous production of a heterologous protein using a novel secretion promoting factor which operates in *Escherichia coli*. *Journal of Biotechnology.* 37, 33-37. 1994.

Tomson, K., Paalme, T., Laakso, P. S., and Vilu, R. Automatic laboratory-scale fed-batch procedure for production of recombinant proteins using inducible expression systems of *Escherichia coli*. *Biotechnology Techniques.* 9, 793-798. 1995.

Turner, C., Gregory, M. E., and Turner, M. K. A study of the effect of specific growth rate and acetate on recombinant protein production of *Escherichia coli* JM107. *Biotechnology Letters.* 16, 891-896. 1994.

Van de Walle, M. and Shiloach, J. Proposed mechanism of acetate accumulation in two recombinant *Escherichia coli* strains during high density fermentation. *Biotechnology and Bioengineering.* 57, 71-78. 1998.

Varma, A. and Palsson, B. O. Metabolic capabilities of *Escherichia coli*: I. Synthesis of biosynthetic precursors and cofactors. *Journal of Theoretical Biology.* 165, 477-502. 1993.

Walker, J. M. and Gingold, E. B. *Molecular biology and biotechnology*. Royal Society of Chemistry, Cambridge. 1993.

Wang, F. and Lee, S. Y. Production of poly(3-hydroxybutyrate) by fed-batch culture of filamentation-supressed recombinant *Escherichia coli*. *Applied and Environmental Microbiology.* 63, 4765-4769. 1997.

Wangsa-Wirawan, N., Lee, Y., Falconer, R., Mansell, C., O'Neill, B., and Middelberg, A. Novel fed-batch strategy for the production of insulin-like growth factor 1 (IGF-1). *Biotechnology Letters.* 19, 1079-1082. 1997.

Weber, A. and San, K. Persistence and expression of the plasmid pBR322 in *Escherichia coli* K12 cultured in complex medium. *Biotechnology Letters.* 9, 757-760. 1987.

Weickert, M., Doherty, D., Best, E., and Olins, P. Optimization of heterologous protein production in *Escherichia coli*. *Current Opinion in Biotechnology.* 7, 494-499. 1996.

Wipf, B., Weibel, E. K., and Vogel, S. Computer controlled large scale production of α-interferon by *E. coli*. *Bioprocess Engineering.* 10, 145-153. 1994.

Wong, H. H., Kim, Y. C., Lee, S. Y., and Chang, H. N. Effect of post-induction nutrient feeding strategies on the production of bioadhesive protein in *Escherichia coli*. *Biotechnology and Bioengineering.* 60, 271-276. 1998.

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

Yamanè, T., Hibino, W., Ishihara, K., Kadotani, Y., and Kominami, M. Fed-batch culture automated by uses of continuously measured cell concentration and culture volume. *Biotechnology and Bioengineering.* 39, 550-555. 1992.

Yamanè, T. and Shimizu, S. Fed-Batch Techniques in Microbial Processes. *Advances in Biochemical Engineering / Biotechnology.* 30, 148-194. 1984.

Yang, J., Moyana, T., Mackenzie, S., Xia, Q., and Xiang, J. One hundred seventy-fold increase in excretion of an FV fragment-tumor necrosis factor alpha fusion protein (sFV/TNF-a) from *Escherichia coli* caused by the synergistic effects of glycine and Triton X-100. *Applied and Environmental Microbiology.* 64, 2869-2874. 1998.

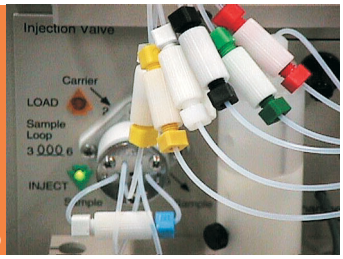
Yang, X.-M. Optimization of a cultivation process for recombinant protein production by *Escherichia coli*. *Journal of Biotechnology.* 23, 271-289. 1992.

Yano, T., Kurokawa, M., and Nishizawa, Y. Optimum substrate feed rate in fed-batch culture with DO-stat method. *Journal of Fermentation and Bioengineering.* 71, 345-349. 1991.

Yazdani, S. S. and Mukherjee, K. J. Overexpression of streptokinase using a fed-batch strategy. *Biotechnology Letters.* 20, 923-927. 1998.

Yee, L. and Blanch, H. W. Recombinant trypsin production in high cell density fed-batch cultures in *Escherichia coli*. *Biotechnology and Bioengineering.* 41, 781-790. 1993.

Zabriskie, D. W. and Arcuri, E. J. Factors influencing productivity of fermentation employing recombinant microorganisms. *Enzyme and Microbial Technology.* 8, 706-717. 1986.



CHAPTER 3
FLOW INJECTION ANALYSIS SYSTEM

"Art and science have their meeting point in method".

Edward Bulwer-Lytton

IIn this section, the applicability of Flow Injection Analysis (FIA) systems to the high-cell density fed-batch fermentations of recombinant *E. coli* is discussed. Methods for on-line measurement of both glucose and acetate were developed. Acetate measurement was performed with a modified and optimised version of an existing method, based on acetate diffusion through a gas-diffusion chamber into a stream containing an acid-base indicator. The subsequent decrease in the absorbance was detected with an incorporated photometer. Commercially packed glucose oxidase was used for the amperometric measurement of glucose. After optimisation, both methods revealed acceptable linearity limits and detection levels. Also, no significant interferences were detected when the results were compared with other reference methods. Parallel and serial combination of both methods was tested in order to optimise on-line simultaneous measurement. The application of the parallel configuration revealed a better option in terms of analysis time and thus was chosen for the application to the fed-batch *E. coli* fermentation. In several experiments, reliable results were obtained.

3.1 THEORY

3.1.1 Flow Injection Analysis Principles

Flow Injection Analysis (FIA) is based on the injection of a liquid sample into a moving, non-segmented continuous carrier stream of a suitable liquid. The injected sample forms a zone, which is then transported toward a detector that continuously records the absorbance, electrode potential, or other physical parameter as it continuously changes due to the passage of the sample material through the flow cell (Ruzicka and Hansen, 1987).

Additional processes, such as chemical reaction or extraction, may also occur between the sample and the carrier in order to generate the detected species. The measurement is performed under non-equilibrium conditions, as neither physical homogeneity nor chemical equilibrium have been attained by the time of the detection (Hall, 1991). In this way, reliability of FIA is based on the principle that all samples are sequentially processed in exactly the same way during the passage through the analytical system.

The advent of this technique took place in 1974 and had to do with the need to automate chemical analysis and to make it compatible with personal computers and on-line process control. Comparing with traditional batch methods, FIA is also much easier to miniaturize and to control, and provides an environment for highly reproducible operations.

In the basis of FIA there was a continuous-flow technique that used air segmentation in the tubes in order to minimize carryover between adjacent samples. Once it was found that this segmentation, as well as the attainment of a steady state is not a prerequisite for continuous flow analysis, FIA became a universally accepted methodology.

FIA belongs to a family that embraces, besides FIA, also chromatography and electrophoresis (Ruzicka, 1992). The individual members of this family differ in one fundamental respect, the function of the chemical modulator, which alters the original square-wave input, provided by sample injection, into a chromatogram or fiagram (Figure 3.1). Thus, the function of the chemical modulator is associated with the name of individual techniques such as chromatography (column), electrophoresis (electric field), and flow injection (reactor).

Although FIA is the youngest of the flow-based techniques, it has already been applied to a full range of reagent assays from inorganic to enzymatic, from ions to proteins and from traces to highly concentrated analytes. Industrial, clinical, scientific research, environmental and biotechnological applications of FIA have been described in many papers and

monographs and discussed at numerous meetings, although for some authors it is still regarded as an academic method (Van der Pol et al., 1996).

The referred success is probably due, not only to the already mentioned computer compatibility and reproducibility, but also to the speed of analysis (sampling frequency varies from 2 to 0.5 min^{-1}), the low sample volumes required (from 1 to 200 μL), the low reagent consumption (no more than 0.5 mL per analysis), and the low relative standard deviation (RSD), typically better than 1% (Ruzicka and Hansen, 1987 and Ruzicka, 1992).

The simplest FIA system consists on a pump, which is used to propel the carrier stream through a narrow tube; an injection valve by means of which a well-defined volume of a sample solution is injected into the carrier stream in a reproducible manner; and a micro-reactor in which the sample zone disperses and reacts with the components of the carrier stream, forming a species that is detected by a flow-through detector and recorded (Figure 3.1). If several reagents must be added in succession, additional streams are confluenced and coils are added. Amongst the first procedures adapted to FIA were classical colorimetric methods (e. g. for ammonia, phosphate, glucose and ethanol), where chart recorders were used to obtain a fiagram. During the last years the range of detectors has increased and also has the variety of reactor designs in order to accommodate solvent extraction, gas diffusion, photodegradation, colorimetric reagent generation and titrations. In addition, miniaturized packed reactors, containing solid reagents, reductants, immobilized enzymes, ion exchangers or silica materials, were introduced to convert, catalyse or pre-concentrate analyte molecules.

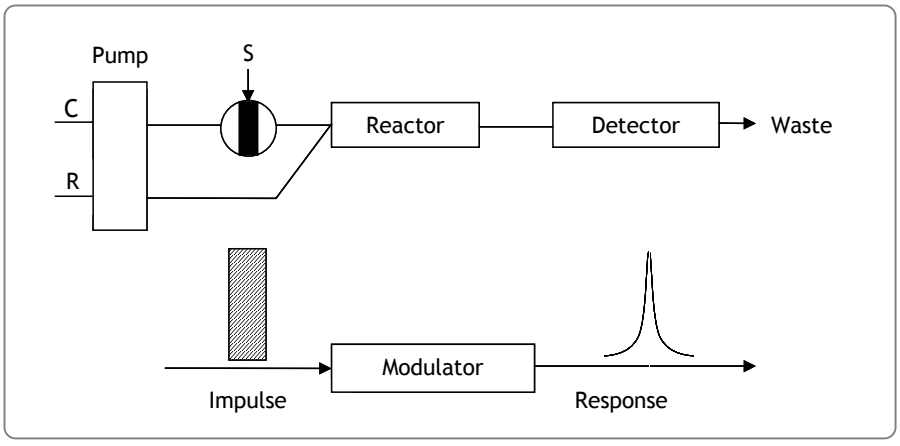


Figure 3.1 Scheme of Flow Injection Analysis. The sample (S) is injected into a carrier stream (C), reacting with a reagent stream (R) inside a reactor. A detectable species is then formed that migrates to the detector and causes a peak-shaped impulse (adapted from Ruzicka, 1992).

In FIA, the peak shaped measurement signal is evaluated by using the signals maximum height, the integral or the width at a specific weight (Figure 3.2). In the absence of chemical reactions, when the detector responds linearly to the injected species, it does not make any difference whether peak height, area or peak width is being measured, since they all yield useful information, although the concentration of the injected material is related to each of these parameters in a different manner. The same applies to assays based on production of measurable species (such as in photometry or electrochemical detection) provided that via the dispersion process an excess of reagent is available throughout the entire length of the dispersion zone. Peak height is the most frequently measured peak dimension, since it is easily identified and directly related to the detector response (Ruzicka and Hansen, 1987).

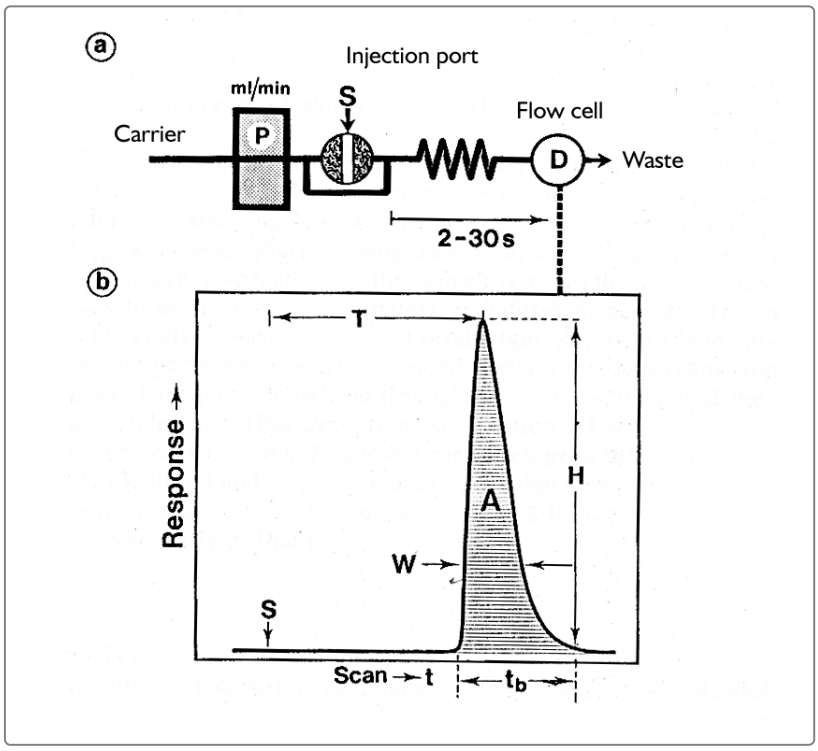


Figure 3.2 The analog output of a FIA measurement system has the form of a peak, the recording starting at S (corresponding to the time of injection). H is the peak height, W is the peak width at a selected level, and A is the peak area. T is the residence time corresponding to the peak height measurement, and t_b is the peak width at the baseline. (Adapted from Ruzicka and Hansen, 1987).

In the flow-injection apparatus, two processes take place simultaneously: physical dispersion of the sample zone within the carrier stream of the reagent, and a chemical reaction between the analyte and the reagent molecules. The parabolic profile established during the injection process expands during sample zone transport trough an open-tubular channel, ultimately

leading to an undesired zone broadening and consequent loss of sensitivity and sampling frequency.

In order to design an FIA system rationally, it is important to know how much the original sample is diluted or dispersed on its way to the detector, and how much time has elapsed between the sample injection and the readout. For the first purpose, the dispersion coefficient *D* has been defined as the ratio of concentrations of sample materials before (C_0), and after (*C*) the dispersion process has taken place:

$$D = \frac{C_0}{C} \qquad \text{Equation 3.1}$$

For the second purpose, the residence time *T* is evaluated, being defined as the time span between sample injection and peak maximum, during which the reaction takes place (Figure 3.2).

A limited dispersion is obtained injecting a sufficient volume of sample. Thus, the variable $S_{1/2}$ can be defined as the volume of sample solution that is necessary to reach 50% of the steady state, that is to say a *D* equals 2. For a certain FIA system, the maximum sample requirement will not exceed $S_{1/2}$. Other variables influence *D*, namely, the length and the diameter of the FIA tubes, the flow rate and the channel geometry. In order to decrease the sample dispersion, a narrow and short tube should be chosen, and the pumping rate should be decreased. The introduction of mixing chambers in the system increases largely the dispersion.

3.1.2 Application of FIA in Bioprocesses

In order to make microbial processes more efficient, several parameters, which give information about the physical and chemical environment as well as about growth and production, have to be determined continuously. So far, mostly physical (like agitation speed, air flow, temperature) and a few chemical parameters (such as pH value, and oxygen pressure) are accessible to on-line measurements and control, and general possibilities for the continuous monitoring of substrate concentration decrease, product concentration increase and other metabolites concentration evolution are scarcely available.

The methods for specific determination of these concentrations are mostly based on enzymatic reactions. Since enzymes are very sensitive towards changes of physiological conditions such as temperature or pH value, the direct and continuous substrate and product monitoring with enzyme electrodes or *in situ* biosensors is very limited and has only been

applied in a few cases so far (Van Brunt, 1987; Riedel et al., 1989 and Ogbomo et al., 1990). Additional complications encountered with enzyme electrodes are the problems related with their sterilization, their possible dependence on oxygen pressure within the bioreactor and the clogging of the protecting membrane. Often, the small linearity range of the enzyme response does not meet the concentration range of the substrate/product in the medium. Finally, the biological part of enzyme electrodes may be inhibited by the components of the process solution.

Thus, those small molecules are often measured off-line after the fermentation, using different methods. Manual off-line analysis, or on-line determinations with slow analysis times, will not typically provide the timely results necessary for real time control of concentration. As fermentations can run for many hours, it is impractical to use even rapid manual methods for continuous control. Off-line analyses may also be inadequate for observing transient culture events because of insufficient sample frequency or because cells and the medium are not separated rapidly enough to prevent sample degradation (Forman et al., 1991).

These problems directed the research towards external continuously working analytical systems. Flow Injection Analysis is a very promising method for on-line process control. In the case of enzyme-based FIA systems, longer stability of the biological component of the sensor is expected, since inhibiting substances can be eliminated, and modifications of signal response due to variations in pH value or oxygen pressure of the cultivation medium can be eliminated. High signal frequencies can be obtained, and a simple and rapid recalibration of the sensor is easily possible.

3.1.3 Sample Pre-treatment

All methods developed for performing on-line analysis with FIA have one common requirement: the sample to be analysed must describe accurately what happens inside the bioreactor. For automated sampling there are mainly two different methods: direct methods, where the withdrawn sample contains all three phases present in the broth (dissolved gas, liquid and solids/cells) and membrane based methods, where a particle-free liquid or gaseous sample is obtained.

Gründig et al. (1993) measured glucose in gluconic acid production fermentation by only automatically diluting the sample before injection and performing a cleaning procedure between every two samples, with no apparent interferences. A similar approach was used by Ding et al. (1993) when determining microbial activity in an *E. coli* fermentation. However, a

major obstacle of these direct sampling methods is the need to stop all cellular turnover of substrates and product formation during retrieval of sample. This can be achieved either by cooling the sample before injecting it (Nielsen et al., 1990) or adding a chemical inhibitor (Christensen et al., 1999).

The described direct methods are suitable for bacteria and yeast fermentation on a synthetic medium growing until low cell densities, but for viscous broth or high-cell density systems, as in the penicillin fermentation or *E. coli* fermentations, a membrane-based sampling system is preferred (Christensen et al., 1991).

In the membrane-based methods, no metabolic turnover of the sample occurs after its retrieval, because the membrane acts as an aseptic barrier. However, attention has to be paid when interpreting measurements carried out on cell-free samples, because in some cases biomass constitutes a substantial part of the reactor volume. Also, the retrieval of cell-free samples from small bioreactors growing cultures at a low specific growth rate can cause an artificial increase in the biomass concentration, that has nothing to do with growth (Christensen et al., 1999).

There are other troubles related to this sample pre-treatment. Common problems are the clogging of the membrane system, and the possible introduction of errors in the measurements, the occurrence of dead volumes, the need of rather large sample volumes, and the difficulties of maintaining aseptic conditions in the culture medium (Ogbomo et al., 1990 and Christensen et al., 1999).

The membrane module may be placed either inside the bioreactor (*in situ* probes) or in a recycle loop connected to the bioreactor. With an *in situ* placed module, a true picture of the conditions in the bioreactor is obtained, but it is not possible to change the membrane during fermentation. Another major drawback of *in situ* membrane probes is their relatively high response time (Christensen et al., 1991).

Several membrane probes have been designed and some are now commercially available. Advanced Biotechnology, Co. developed the most used one that is now commercialized by Eppendorf-Netheler-Hinz (Germany). It consists on a membrane mounted on a stainless steel support unit. Liquid passing the membrane is collected, enters a central channel of the probe and a continuous sample stream to the analyser leaves the unit from the top of the probe. The membrane has the form of a tubular stocking made of polypropylene, polycarbonate or a ceramic material. It accommodates filtrate flow rates from 0.5 to 2.0 $mL \cdot min^{-1}$, but a constant flow of 1.0 $mL \cdot min^{-1}$ may be obtained throughout the fermentation. Total delay time is about 5 minutes. The same membrane can be used for many fermentation experiments if the membrane is washed after each fermentation. The hydraulic resistance is kept almost

constant after more than 1200 h of use. Huang et al. (1995) and Christensen et al. (1999) used this device to obtain particle free samples for the measurement of glucose in an *E. coli* fermentation. Tothill et al. (1997) applied another similar filtration module designed by Persep (Techsep, France) to the same purpose.

With the membrane module placed in a loop outside the bioreactor it is possible to change the membrane during fermentation, but it may have an effect on the performance of the microbial culture owing to the recirculation in the loop. At high-cell densities, the dissolved oxygen is consumed rapidly by the cells in the loop, and the oxygen limitation may cause the change from the oxidative to the oxidative-fermentative metabolism. This phenomenon can occur in *S. cerevisiae* or *E. coli* (Schügerl, 1993 and Schügerl et al., 1996). Also, microorganisms can grow in the tubes of the loop, causing consumption of the medium components and clogging of the system. For example, a growing cell suspension of *E. coli* in the sample can be expected to consume several grams per litre per minute at the end of a batch phase (Bradley et al., 1991). The shear might also influence the morphology of the microorganism (Christensen et al., 1991). Additionally, some of these systems have long response times, need high filtrate rates and the volume of the sample is too high to allow the use of the system with small fermenters (Garn et al., 1989). It is therefore important to minimize both the residence time in the loop and the ratio of loop volume to working volume of the bioreactor.

Several filtration devices that can be used in an external loop are commercially available and were used by different authors in many microbial processes.

The Biopem device was designed by B. Braun, Germany but is no longer commercialized. It consists on a small mixing chamber in which liquid flow at the membrane surface is created by a small height adjustable magnetic rod placed above the membrane and driven by an external magnetic stirrer. It is claimed that it performs well for bacteria and yeasts, with the drawbacks of having high response time (1 – 2 minutes) as a consequence of the large dead volume (200 mL).

Freitag et al. (1991) used this device together with a tangential filtration unit connected in series during a fermentation of hybridoma cells where protein concentration was measured on-line with FIA. The main problem with this application was the lost of permeability of the device for the proteins.

Christensen et al. (1991) also used this device during penicillin fermentations but it revealed very poor for sampling during fermentation with filamentous fungi, because the filtrate decreased substantially during the fermentation.

Other devices exploit the concept of cross-flow or tangential filtration. In those devices, a sufficiently high flow rate prevents the formation of membrane deposits as well as the existence of anaerobic conditions or substrate consumption.

Among the available devices of this kind, the mostly used is the A-SEP device, commercialized by Applikon (Germany). This device is represented in Figure 3.3. The overall dimensions of A-SEP are 5.5 × 5.5 × 2.5 cm. A 47-mm diameter hydrophilic membrane filter disc is sandwiched between stainless steel plates, each with grooves machined into them. The larger channel on one plate carries the process flow; the smaller channel on the other plate carries the filtrate to the outside where tubing is attached.

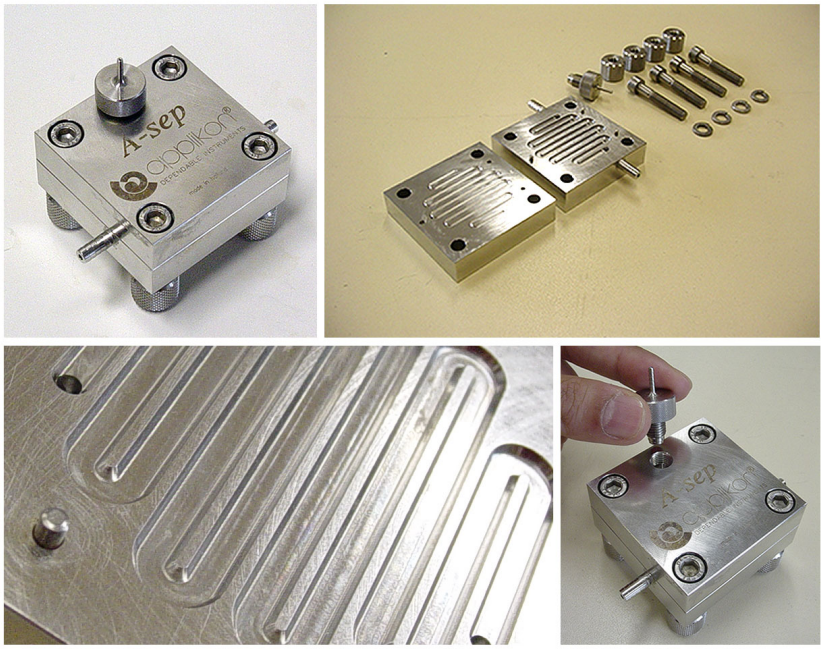


Figure 3.3 A-SEP sample pre-treatment tangential flow filtration device.

According to Forman et al. (1991), this device has the following advantages:

- the use of cross flow design ensures efficient operation by reducing membrane fouling;
- the use of standard-sized filter discs, available in a wide range of materials;
- allows the use of a linear recycle rate to minimize the period of time that the culture is out of the fermenter;
- the use of materials that can be sterilized in an autoclave.

These authors used A-SEP for the on-line determination of acetate and phosphate with FIA in an *E. coli* fermentation and observed no fouling of the membrane. The filtrate rates were sufficient for on-line analysis and the response time was only 25 seconds. Zigova et al. (1999) also used A-SEP with success during the analysis of glucose during *S. cerevisiae* fermentation.

Other cross or tangential flow devices are described in the literature with successful applications to several bioprocesses with even shorter response times than the ones referred. Christensen et al. (1991) compared the performance of a developed cross-flow device with Biopem and an *in situ* probe during filamentous *fungi* fermentation and obtained better results in terms of response time with the first one (only 10 seconds).

Garn et al. (1989) used a micro-filtration module similar to A-SEP that had been conceived by Millipore, USA, but was never commercialized for the analysis of several components in *E. coli*, *S. cerevisiae* and *S. Pilosus* fermentations. The sampling system had a bypass, in order to sterilize it *in situ* at any time.

Chung et al. (1991) had a similar approach for the measurement of glucose, ammonia and proteins in a *S. cerevisiae* fermentation.

A more sophisticated but also expensive alternative is the Minitan device, commercialized by Millipore. This device can accommodate several membranes and allows filter regeneration. Unlike any other device mentioned in this revision, it allows the returning of the filtrate not used in analysis to the bioreactor.

This equipment was used by Valero et al. (1990) for the measurement of glucose in a *Candida rugosa* fermentation and in Gabriel et al. (1998) during on-line analysis of nitrite and nitrate in wastewater samples. During the bioprocesses, no increase in pressure or plugging was observed in the microfiltration equipment.

Additionally to the devices described, other non-conventional sample pre-treatment equipments have also been referred in the literature for very specific bioprocesses, namely vapour generation, dialysis, and microdialysis (Fang, 1999).

3.1.4 Glucose Analysis

Glucose is an analyte of central importance in medicine, food and bioprocess analysis as it is one of the most important carbon sources of energy for cell growth. However, an excessive amount of glucose causes production of acetate during bacteria fermentation, which reduces biomass as well as productivity. Therefore, it is desirable to be able to monitor glucose

concentration, and thereby control its feeding during fermentation. Also, on-line glucose monitoring can be used to generate a signal for the beginning of the feeding phase in fed-batch fermentations. In order to achieve this, it is necessary to use an automated on-line monitoring and control system.

Glucose level monitoring in fermentation can be based on direct or indirect monitoring techniques. Indirect monitoring can be done by analysis of off-gas components or pH measurements, from which the glucose level of the fermentation broth can be deduced. Another indirect approach is based on software sensors that can predict glucose concentrations from other state variables measurements through mathematical model manipulations. However, recent advances have started a trend towards more direct glucose measurements, usually based on enzymes. In such cases, the glucose measurement can be done outside the fermenter or *in situ*.

As mentioned before, commercially available *in situ* biosensors used for glucose measurements, usually do not fulfil some of the essential requirements for an *in situ* probe, as they do not assure sterility, they do not allow readings over a wide range of concentration, and they do not give rapid results (Tothill et al., 1997). Also, complications result from the high complexity of reaction mixtures in the bioreactor, which are composed of many different metabolites whose influence on the enzymatic analysis is still unknown.

Thus, an approach based on the combination of the available enzymatic systems with on-line FIA measurements is preferred since it gives more accurate results with a lower response time (Bilitewski et al., 1993), besides the fact that it does not alter the sterility of the culture. A number of this kind of applications to biotechnological processes has been studied with some success and a non-exhaustive approach of what exists in the literature will be given throughout this section.

There are mainly two enzyme systems available for the measurement of glucose that allow an easy detection method due to their reaction products: glucose oxidase (GOD) and NAD-dependent glucose dehydrogenase (GDH) (Schmid and Künnecke, 1990 and Schügerl et al., 1996). Reactions involved in these processes are as follows:

$$\beta\text{-D-Glucose} + O_2 \xrightarrow{GOD} \text{Glucono-}\delta\text{-lactone} + H_2O_2 \qquad \text{Equation 3.2}$$

$$\beta\text{-D-Glucose} + NAD^+ \xrightarrow{GDH} \text{Gluconic Acid} + NADH + H^+ \qquad \text{Equation 3.3}$$

A non-negligible characteristic of these reactions is that both enzymes are specific for β-D-Glucose. The equilibrium reaction between both anomers is the following:

$$\alpha\text{-D-Glucose} \xrightarrow{\text{mutarotase}} \beta\text{-D-Glucose} \qquad \text{Equation 3.4}$$

Due to this, in order to obtain reliable results during glucose analysis, one has to be sure that all standards and samples are in mutarotational equilibrium, especially since some microorganisms have anomeric specificity with regard to the uptake of glucose. For some applications it is therefore necessary to co-immobilize mutarotase together with GOD or GDH (Zigova et al., 1999).

In terms of set-up, FIA systems usually consist of the immobilized enzyme column positioned after the mixing coil for the analyte solution and the reagents. The enzymatically formed indicator reagent is then transported to and assayed in the detection unit. If the immobilized enzyme and the detector are in sufficiently close contact, the detector becomes a “biosensor” (Schmid and Künnecke, 1990). Although there are several descriptions of biosensors in the literature, in most enzyme-supported FIA designs, immobilized enzyme and detector unit are kept separate. The reason for this is a 20-fold higher operational stability of enzyme reactors compared with enzyme electrodes. Another obvious advantage is that both modules can be used separately.

Between both available enzymes, GOD and GDH, the most commonly used is the first one because, like the majority of oxidases but unlike GDH, it can be used without a coenzyme, simplifying greatly the flow system. However, application of GDH can still be found in the literature coupled with fluorescence detector to measure the formed NADH. Zigova et al. (1999) used this approach for on-line analysis of glucose during *S cerevisiae* fermentation.

Glucose oxidase is usually obtained from *Aspergillus niger* and has a molecular weight of 180 kDa. It is a homo-dimeric enzyme and is highly specific for β-D-Glucose. Other hexoses, mannose, galactose, and 2-deoxy-glucose, are poorly oxidized. It is usually immobilized on controlled-pore glass beads, using the gluteraldehyde procedure, inside of a packed-bed reactor. Commercial reactors of this kind are available and have a good performance in terms of stability and durability.

As can be inferred from the reaction described on Equation 3.2, detection with GOD implicates the presence of molecular oxygen in the medium, raising another problem for this kind of measurements, because the FIA system has to be operated in a way which assures a surplus of oxygen in all samples regardless of the oxygen tension prevailing in the bioreactor (Haouz and Stieg, 2002). In cases where there is no oxygen control during the bioprocess, the possible insufficiency of oxygen can be circumvented by diluting the carrier flow with oxygen-saturated buffer or by considering other electron acceptors (Christensen et al., 1991).

Another pitfall in this kind of measurement system has to do with the fact that glucono-δ-lactone acts as a weak competitive inhibitor of GOD. In order to solve this, most of the commercial packed enzymes have immobilized lactonase besides glucose oxidase, assuring fast hydrolysis of that compound to gluconic acid, according to the following reaction:

$$\text{Glucono-}\delta\text{-Lactone} + H_2O \xrightarrow{\text{lactonase}} \text{Gluconic Acid} \qquad \text{Equation 3.5}$$

In the GOD-based FIA system, there are several ways to determine the formed or consumed species, each one with particular advantages and drawbacks. In the literature, there are descriptions of the applications of these approaches to different kinds of bioprocesses.

Hitzmann et al. (1997) measured glucose by detecting a pH decrease caused by the formation of gluconic acid. Due to the change in the pH of the samples, in this kind of measurements, the peak heights, widths or areas were not adequate to obtain good results. Thus, in this case, artificial neural networks, that correlated several characteristic contours of the peaks with the analyte concentration, evaluated the measurement signals.

The drop on the dissolved oxygen caused by the reaction catalysed by glucose oxidase can also be the basis for an amperometric measurement of glucose. Traditionally, Pt electrodes poised at -600 mV and + 700 mV, have been used to monitor the following reaction:

$$O_2 + 4H^+ + 4e^- \longrightarrow 2H_2O \qquad \text{Equation 3.6}$$

Haouz and Stieg (2002) measured both glucose and lactate based in this principle.

However, the detection of the formed hydrogen peroxide has become the most popular method for the measurement of glucose. Also in this case, the approaches vary significantly, but for all of them some difficulties have to be taken into account related to the fact that the detection is often hampered by interfering reactions that consume some amount of the hydrogen peroxide. It is the case of penicillin, manganese ions and ammonium.

In the simplest version for the detection of hydrogen peroxide, this compound can be detected amperometrically in the same way as oxygen by direct electrochemical oxidation at Pt electrodes at large over-potentials (500-700 mV *versus* Ag/AgCl/sat. KCl) following the reaction:

$$H_2O_2 \longrightarrow O_2 + 2e^- + 2H^+ \qquad \text{Equation 3.7}$$

Hall (1991) referred several applications where hydrogen peroxide was detected amperometrically by a Pt electrode (sulphite, lactate, and glucose), mentioning some of the interferences that this kind of measurement has. Gründig et al. (1993) developed a biosensor

for the simultaneous monitoring of glucose and sucrose with automatic detection based in the amperometric principle.

Matsumoto et al. (1990) also used a similar method, optimising some important variables like the pH and obtained reliable results with a linearity from 0.02 up to 10 mM and a RSD of 1.1%. Huang et al. (1995) presented a FIA/WJE (Wall Jet Electrode) amperometric detector coupled with a packed-bed reactor for the on-line measurement of glucose during *E. coli* fermentations. The authors achieved linearity until 4 $g \cdot L^{-1}$ using an internal standard for minimizing the errors caused by deterioration of the enzyme reactor and changes in the surface of the Pt electrode. The immobilized enzyme reactor was stable for more than one year. Using this set-up, glucose concentration was on-line controlled and it was possible to achieve high-cell densities, although there were glucose pulses and some acetate accumulation. Lüdi et al. (1990) also used this approach for the determination of glucose in *S. cerevisiae* fermentation.

For this kind of detection, however, the high positive potentials used increase the risk of interference from several oxidizable compounds such as ascorbate in samples. Furthermore, in flow analysis, a long time is required to stabilize the background current of Pt electrodes (Schügerl, 1993).

Some authors tried to solve these problems by electrically reducing the hydrogen peroxide, thus facilitating the electron transference, either by using electron mediators or a direct reduction with glucose peroxidase (POD) without mediators.

Tothill et al. (1997) used a new catalytic powder based on promoted rhodium, dispersed on carbon, which selectively reduce the potential required to oxidize hydrogen peroxide for the construction of an amperometric biosensor. This biosensor was applied to the measurement of glucose in the fermentation of *E. coli*, *S. cerevisiae*, and lactic acid fermentation in different growth media. The results were compared with the ones obtained with other off-line enzymatic methods and the correlations were very close to 1 except when the samples were not diluted, increasing interferences from the matrix.

Céspedes et al. (1995) used a biocomposite material for the electrocatalytic oxidation of hydrogen peroxide. This amperometric glucose biosensor was integrated in a FIA system and used for discrete glucose monitoring of *Candida rugosa* fermentation. The analytical system showed a linear response for glucose in the range 10^{-5} – 10^{-2} $mol \cdot L^{-1}$. The reproducibility of the signal was below 5% and the sampling frequency was about 20 $samples \cdot h^{-1}$. The results were compared with glucose sensor from Yellow Spring Instruments (Model 2000) and to a HPLC system with a good agreement.

Yamamoto et al. (2000) proposed a peroxidase/ferrocene-embedded carbon paste electrode covered with a membrane to reduce bioelectrocatalytically the hydrogen peroxide at 100 mV. The linearity of the calibration was retained up to 100 μM. The detection limit was 75 nM. The RSD was 2.7% and the lifetime of the GOD immobilized column was over 1 month. The FIA system was applied to the determination of glucose in a human serum sample and the glucose concentration was similar to the one obtained with other methods.

Min et al. (1998) built an amperometric biosensor based on the co-immobilization of GOD with POD in a carbon paste matrix. The method was linear between 0.01 and 0.1 $g \cdot L^{-1}$. The potential applied to the working electrode was -50 mV *versus* Ag/AgCl (much lower than the usual +600 mV, due to the presence of the immobilized POD).

Besides the amperometric detection, hydrogen peroxide can also be determined by indirect methods, either photometrically or by chemiluminescence.

Both Garn et al. (1989) and Chung et al. (1991) based their glucose determination on the so-called Trinder reaction, where a glucose oxidase-peroxidase (GOD-POD) system allowed the peroxidase-catalysed oxidative reaction of hydrogen peroxide to yield a coloured product measured photometrically at 505 nm. In the first work, a dual-wavelength, double-injection and stopped-flow system was used for a general bioprocess. In the second case, it was used during fed-batch fermentations of *E. coli* and *S. cerevisae*, and the results obtained were successfully used as a parameter for on-line estimation of biomass concentration in the fermenter and for prediction of the end of the batch phase.

A very similar reaction was used with reliable results by Valero et al. (1990) with non-immobilized enzymes in the measurement of glucose in a fermentation of *Candida rugosa*.

Vodopivec et al. (2000) applied a new immobilization technique (Convective Interaction Media) to the GOD and POD enzymes in order to apply the above-mentioned reaction to the successful on-line monitoring of glucose during cultivation of *S. cerevisiae* and *Aspergillus niger*.

An optical biosensor for glucose was developed by Lenarczuk et al. (2001) where a Prussian Blue (PB) film with chemical linked molecules of GOD constituted the biological part of the biosensor. The PB changed its colour in the presence of hydrogen peroxide, allowing the optical detection of glucose at 720 nm. The biosensor was linear from 0.1 to 2 mM and revealed very specific for glucose and applicable to pharmaceutical products, drinks and body fluids.

A different approach is to detect the hydrogen peroxide formed by the chemiluminescence reaction:

$$\text{Luminol} + 2H_2O_2 \longrightarrow \text{3-Aminophthalate} + 3H_2O + \text{Light}$$ Equation 3.8

One of the disadvantages of this method is that, for most process applications, it is too sensitive, and the sample has to be diluted (Schügerl et al., 1996). However, some authors (Nielsen et al., 1990 and Christensen et al., 1999) were able to analyse glucose over a wide linear range (1 $mg \cdot L^{-1}$ – 2 $g \cdot L^{-1}$ and 0.25 – 500 $mg \cdot L^{-1}$, respectively) with no observable interferences in a number of different bioprocesses. Spohn et al. (1995) also applied this method successfully with the integration of an optical fibre in animal cell cultures.

Finally, a different approach to eliminate the interferences during the measurement of glucose with FIA was proposed by Hitzmann et al. (1998) and Schöngarth and Hitzmann (1998). These authors used a multiple injection FIA system to compensate the interferences of the matrix in complex samples. Three different solutions were injected in a fast sequence, being the first and the third one standards. Calculation of the sample concentration was done by Artificial Neural Networks or employing Partial Least-Squares regression. These techniques enabled a fast adaptation to changes in the reaction system (it is similar to the use of internal standards). It was applied to the determination of glucose while the temperature of the reaction coil was changed. Glucose oxidase with concomitant detection of the exhausted oxygen with a dissolved oxygen electrode was used.

3.1.5 Acetate Analysis

Acetic acid can be the major metabolic by-product of *E. coli* cultures grown on glucose either in the presence of high substrate concentrations or when dissolved oxygen is very low. High acetate concentration has been associated with low cell yield or reduced productivity. An on-line analyzer can help to identify causes of observed culture problems and allow the test of various metabolic control strategies. Acetate is also important in other bioprocesses, as a non-desirable metabolite or as the major biotechnological product, like in yeast fermentation or in biological vinegar production, respectively. Despite this importance, only few methods allowing the on-line monitoring of this metabolite during fermentation are available, especially because the adaptation of the enzymatic determination method to a flow system is not straightforward.

The combination of immobilized acetate kinase (AK), and pyruvate kinase/lactate dehydrogenase (PK/LDH) reactors, with amperometric detection of NADH was described by Tang et al. (1997) and applied to *E. coli* cultivations. The corresponding reactions are as follows:

$$\text{Acetate} + \text{ATP} \overset{\text{AK}}{\rightleftharpoons} \text{Acetylphosphatase} + \text{ADP} \qquad \text{Equation 3.9}$$

$$\text{ADP} + \text{Phosphoenolpyruvate} \overset{\text{PK}}{\rightleftharpoons} \text{Pyruvate} + \text{ATP} \qquad \text{Equation 3.10}$$

$$\text{Pyruvate} + \text{NADH} + \text{H}^{+} \overset{\text{LDH}}{\rightleftharpoons} \text{Lactate} + \text{NAD}^{+} \qquad \text{Equation 3.11}$$

The amperometric measurements were carried out by means of a potentiostat connected to an electrochemical three-electrode cell of the wall-jet type. The working electrode was a graphite rod covered with an adsorbed mediator. A Pt wire served as the auxiliary electrode and a saturated calomel electrode was used as reference.

Although a very selective analysis should be expected due to the high specificity of AK, which converts, aside from acetate, only propionic acid, however in an extremely lower reaction rate, these authors found great interferences from pyruvate, also found in the broth during the batch phase of the *E. coli* fermentation. Pyruvate interfered with the measurement of acetic acid because it is an intermediate in the last step of the enzymatic reactions. The pyruvate concentrations were approximately 1% of that of acetate on a molar basis, but the analytical errors were found to be much larger because of the low yield in the AK reactor. The solution was to determine pyruvate separately, the sample bypassing the AK reactor. The major drawback of this solution was that it was necessary to make two calibration curves for pyruvate because of the increased dispersion that occurred as the fermentation samples passed through two columns.

Also, repeated injections of fermentation samples caused a drift downwards of the electrode baseline, probably caused by the interferences of this kind of detectors, already mentioned in the previous section.

Due to these drawbacks, the analytical results were only accurate for the relatively few measurements on each series reported by those authors: only twelve hours at a sampling frequency of 1 sample·hour^{-1}. Also, the system was only linear up to 480 mg·L^{-1}.

Moreover, although the stability of the PK/LDH reactor allowed it to run under fermentation conditions for 3 months, AK, which is the most expensive enzyme, is labile and lost 90% of its activity after one fermentation, under the mentioned conditions. Thus, a freshly immobilized AK had to be used for each cultivation.

Finally, the repeatability of the system was of 3.4% RSD and a comparison of the results with a commercial enzymatic kit showed that FIA results were 77 – 87% of the test kit data.

Better results were obtained with this system by Becker et al. (1993) in a vinegar fermentation process. Unlike the previous work, enzymes were immobilized in separate

reactors and the system was based on reversed FIA, in order to minimize enzyme and co-enzyme costs. The problem with enzyme AK was not as accentuated as previously, although it still lost 40% of its activity in the first 6 days (with continuous measurements at a frequency of 10 analyses per hour), being thus necessary to add a parallel enzyme system to the analyser that could be calibrated while the other was in use, and then automatically the system switched to this enzyme when the one being used lost its activity.

Also, it was necessary to store the NADH solution separately from the ATP/Phosphoenolpyruvate solution, because 30% of the original NADH solution was lost within 17 h when all three components were mixed.

Comparison of the results obtained with this FIA system with the ones obtained with the commercial enzymatic kit showed good agreement with a maximum deviation of 15%. After optimisation, the system was linear from 10 to 80 mM.

In a more recent work, Becker and Schmidt (2000) applied this approach to a system that also measured ethanol in a joined enzyme cascade system.

A completely different but much simpler approach was proposed by Tservistas et al. (1995): an inverse method was developed, where acetate acted as competitive inhibitor of the enzyme sarcosine oxidase (SO). Sarcosine, the natural substrate of SO, is an N-methylated derivative of glycine. The enzyme oxidizes it as follows:

$$\text{Sarcosine} + H_2O + O_2 \xrightarrow{SO} \text{Glycine} + \text{Formaldehyde} + H_2O_2 \qquad \text{Equation 3.12}$$

This reaction offers the same number of possibilities of determination that were discussed in the previous section, as it consumes oxygen and produces hydrogen peroxide. The authors chose to measure the decrease of the concentration of oxygen in the solution with an oxygen-electrode. In the presence of acetate, the enzyme is inhibited and the rate of reaction decreases. Based on that decrease of signal, the concentration of acetate can be determined.

Sarcosine could be added to the sample solution before passing the injection valve or added to the carrier solution. Depending on the method, acetate was measured as a positive peak whose height decreased with the increase of the concentration of acetate, or as a negative peak whose height increased with increasing acetate concentration. The method was linear up to 1.5 $g \cdot L^{-1}$.

This technique was applied to the measurement of acetate during *E. coli* batch fermentation with a great success, when the results were compared with gas chromatography.

However, although the last enzymatic method reveals promising in terms of a future commercial integration, none of these methods is yet commercially available and its application to a routine basis would implicate the constant immobilization of several enzymes (three in the first case). Moreover, the complexity of the first mentioned method makes it less flexible and causes difficulties with their integration in simultaneous determination of several analytes within the same FIA system.

For these reasons, a chemical method for the analysis of acetate is to be preferred. Forman et al. (1991) described a simple physical-chemical method for the analysis of acetate during *E. coli* fermentations that is illustrated in Figure 3.4.

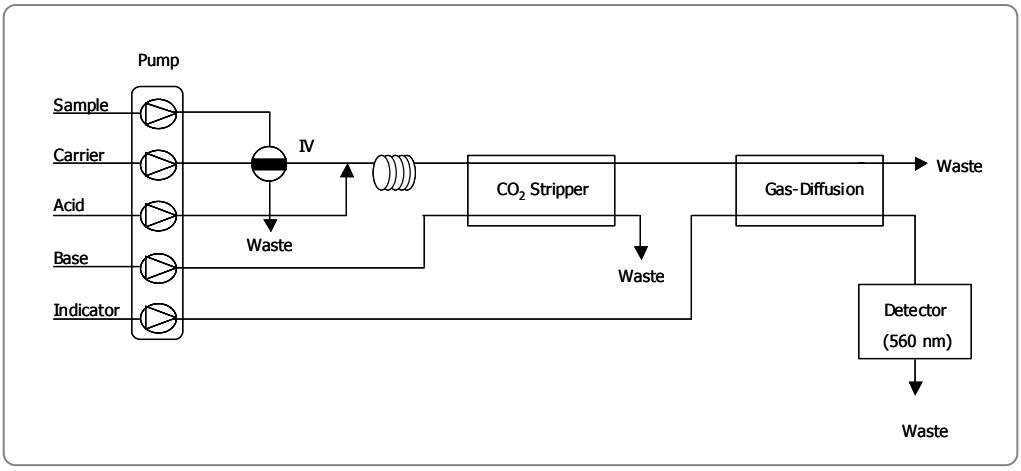


Figure 3.4 Scheme of the FIA method used by Forman et al. (1991) for the measurement of acetate. The sample was injected into a carrier stream (water) that was acidified by a sulphuric acid stream. The resulting stream passed trough a carbon dioxide stripper to eliminate the interference caused by this compound. Finally, acetate passed through a gas-diffusion chamber into an indicator stream, decreasing the red intensity, measured at 560 nm.

The sample containing the analyte was injected by means of a 10-μL sample loop on a carrier stream, consisting of distilled water, suffered an acidification with 1 M sulphuric acid and passed through a 15-cm knotted mixing coil. Consequently, the predominant form of acetic acid was the non-anionic form, which was able to cross a porous 0.2-μm hydrophobic membrane into an indicator stream containing a phenol red solution (0.1 mM phenol red in 1 mM sodium phosphate pH 7.2). At the neutral pH of this stream, acetate could no longer cross the membrane back and dissolved on the indicator, causing a pH drop, and consequently a colour change, detectable photometrically at 560 nm. Acetate concentration

was then proportional to the height of a negative peak. The gas diffusion process is schematically represented on Figure 3.5.

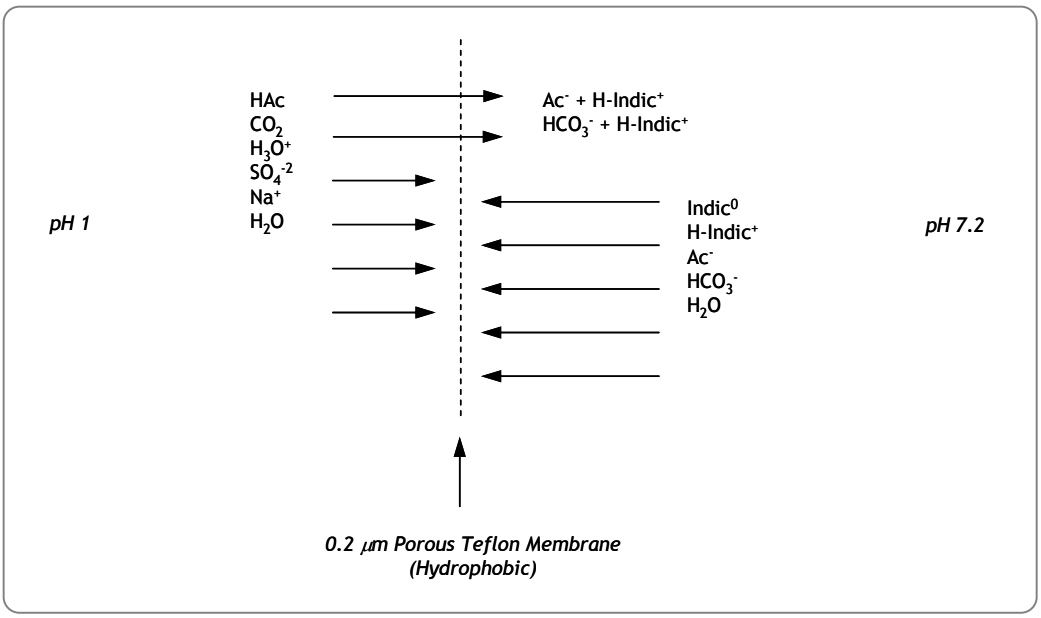


Figure 3.5 Gas-diffusion process. At a low pH, the non-ionic forms of volatile acids crossed the hydrophobic membrane, thus acquiring a positive charge at a higher pH value. In this new form, the molecules could no longer pass the membrane back (Forman et al., 1991).

As it can be seen from Figure 3.5, an important benefit of this method is that it excludes non-volatile coloured and buffering species, which might otherwise interfere with the colorimetric analysis.

However, an unfortunate problem with the assay is that any volatile acid can be detected. As carbon dioxide is the major end-product of aerobic glucose metabolism, it is present in the fermentation media, as hydrogen carbonate, often at concentrations higher than that of acetate. Further, according to the author, the molar response of carbon dioxide in the assay is 40-fold higher than that of acetic acid. To overcome this problem, an in-line carbon dioxide stripper was developed.

The design of the stripper is similar to that of the gas-diffusion device. A silicone-rubber tubing submerged in flowing 2 M sodium hydroxide permitted the rapid diffusion of carbon dioxide, with no observed loss of acetic acid. On the receiving side of the tubing the alkaline solution converted carbon dioxide into carbonate, which cannot re-cross the membrane back. This process is illustrated in Figure 3.6.

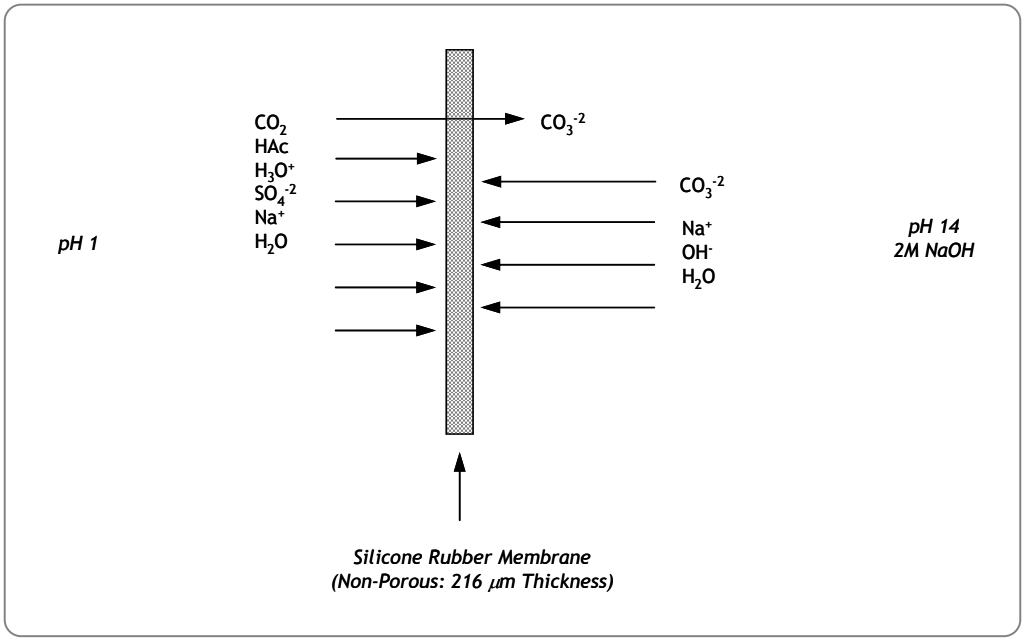


Figure 3.6 Carbon dioxide stripper. With the used thickness, only very small molecules could pass into the alkaline solution. The carbon dioxide is rapidly transformed in an ionic form that could no longer pass the membrane (Forman et al., 1991).

Applying this configuration to a real *E. coli* fermentation and comparing the results obtained with Ion Chromatography, the author found interferences in the order of 5-10 mM of acetate higher for FIA. The major responsible for this was formate that can also be determined by this method.

3.1.6 Simultaneous Determination

Although there is a vast amount of application of FIA to bioprocesses described in the literature, only few of them cover on-line measurements and an even less has focused on methods for simultaneous determination of multiple chemical compounds. Theoretically, when trying to develop an FIA system for several analytes, there are two possibilities: to do it in a serial or in a parallel design. The great part of the described applications is about parallel configurations, but it seems that the serial configuration, which would allow a significant saving in sample volumes, was not even tested in the majority of them. Another clear advantage of serial configurations is that they require much less system components, like

valves or pump channels. Regarding parallel configurations, the main benefits are connected with the speed of analysis, which is approximately twice than in the case of serial configurations, and the lower dispersion of the sample.

When designing such a simultaneous system, depending on the FIA equipment available, there are several optimisation variables that may differ significantly when methods are considered separately. Variables like pump speed and sample distribution in the case of parallel configuration, and sample volume and reagent interactions in the case of serial configuration have to be taken into account.

However, even the few authors that mention a simultaneous configuration do not give details about this optimisation process. Usually, the methods are described separately and only afterwards the configuration used is referred. A brief description of some of those applications can be found on Table 3.1 .

Table 3.1 Application of simultaneous FIA systems to several bioprocesses

Analytes	Process	Configuration	Comments	Reference
Nitrite and nitrate	Wastewater treatment	Serial	Same detector in both measurements	Gabriel et al. (1998)
Glucose and lactate	*E. coli*	Parallel	Oxygen electrode detector	Haouz and Stieg (2002)
Glucose, ethanol and lactate	Alcoholic beverages and serum	Parallel	Multichannel amperometric detector	Matsumoto et al. (1990)
Glucose and lactate	*Lactococcus lactis*	Parallel	Multichannel amperometric detector	Min et al. (1998)
Glucose, galactose, lactose, lactic acid, and proteins	Lactic acid fermentation (*Streptococcus cremosis*)	Parallel	Chemiluminescence detection	Nielsen et al. (1990)
Glucose and fructose	Fruit juices	Parallel	Redox reaction chain	Becker and Schmidt (2000)
Ethanol and acetate	Vinegar production	Parallel	Redox reaction chain	Becker and Schmidt (2000)
Glucose, lactate, ammonium, glutamine, and glutamate	Animal cell culture	Parallel	Chemiluminescence detection	Spohn et al. (1995)

3.1.7 New Trends in FIA Systems

Although FIA is still a recent technique, the “classical” approach that was described in the last sections has been suffering constant improvements and new developments in order to increase analyses performance and to cover new fields.

Optimisation of a FIA system is a balancing act, where dispersion of the sample zone and its mixing with a reagent must be weighted against the time required to achieve a desired chemical conversion of an analyte into detectable species. According to some authors (Ruzicka, 1992), “traditional” continuous FIA systems are difficult to optimise, thus leading to the notion of stopped-flow. As the name indicates, in a stopped-flow system, the carrier flow rate is stopped for a while in order to allow reaction to take place without increasing sample dispersion.

Several authors investigated this technique (Van Akker et al., 1999 and Castañon-Fernández et al., 2000) and concluded that many advantages are gained, namely an increase on the sensitivity of determination, a decrease on the consumption of solutions (and waste generation), and the possibility to perform reaction rate measurements, thus being the slope of the reaction rate curve recorded during the stopped-flow interval, instead of the peak height to quantify the analyte. This particular characteristic allows the elimination of the effect of complex matrices in fermentation processes.

Another methodology presented by Ruzicka (1992) is based on the development of flow-injection methodology based on dispersion and mutual penetration of sample and reagent zones – the Sequential Injection Analysis (SIA). This technique uses a selector (rather than injection) valve, through which precisely measured volumes of carrier solution, sample solution and reagent solution are aspirated into a holding coil by means of a pump which is capable of a precise controlled stop-go-forward-reverse movement.

SIA is mechanically simpler than conventional FIA, as it uses only a simple pump, a single valve and a single channel. The flow path does not have to be reconfigured if the injection volumes or reaction times are to be modified, as these parameters can be altered by changing the flow-rates, stopped-flow period and flow reversals via computer control. As additional reagents, reactors, and detectors can be clustered around the selector valve, multi-reagent chemistries and multi-detector assays can in principle be carried out in a single SIA system.

Other new developments are related with the application of chemometrics to the interpretation of results obtained with FIA (Schöngarth and Hitzmann, 1998) and the

development of fault diagnosis systems that can detect measurement problems like air bubbles and other interferences (Wu and Bellgardt, 1995).

In the next sections, the developed FIA system for on-line monitoring of glucose and acetate is described and the main results obtained with it are shown and discussed.

3.2 MATERIALS AND METHODS

3.2.1 Sampling System

The A-SEP device from Applikon (The Netherlands) was the system chosen at the pre-treatment step for the on-line measurement of glucose and acetate with FIA. Therefore, it was possible to work with low dead volumes and low response times as well as with a sterizable device. Another advantage was the possibility of changing the membrane during fermentation and the low cost of this device when compared with all the others mentioned in section 3.1.3.

The sampling system is illustrated in Figure 3.7. The bio-suspension was extracted from the fermenter using a peristaltic pump model 503U from Watson Marlow, UK, followed by the microfiltration unit A-SEP; inside this device, a small part of the bio-suspension passed through a standard 0.45 μm membrane with a 47 mm diameter from Pall Corp., USA, and the non-filtrated bio-suspension returned to the fermenter. In order to reduce dead volumes, which contribute to measurement bias, the sampling loop was kept as short as possible, and the peristaltic pump speed was maximized. Therefore, the retention time of a portion of bio-suspension outside the reactor never exceeded 1 min, and the corresponding volume never exceeded 40 mL. Also, to assure both a constant volume of filtrate for the FIA system and the correspondence of the samples collected to the most recent state inside the fermenter, a collecting vessel was designed to reject the filtrate volumes in excess. This vessel was permanently agitated with a magnetic stirrer and was always kept closed to avoid the evaporation of acetate.

3.2.2 FIA System

The FIA system used was the ASIA model from Ismatec, Switzerland. This is a modular commercial FIA apparatus, allowing the implementation of several set-ups within the same system. It was equipped with the following components (all from Ismatec):

- 1 vario pump (Model IS 7610)
- 1 fix pump (Model IS 7600)
- 3 injection valves (Model IS 7630)
- 1 six-way valve (Model IS 7620)
- 1 diffusion chamber (IS 2210)
- 1 photometric detector (IS 7595)
- 1 amperometric detector (IS 2996)
- 1 “T” junction
- 1 agitated dilution chamber

For both acetate and glucose measurements, the fix pump was used to propel the samples and standards into the sample loop of the injection valves through the 6-way valve, that selected the standard/sample to be analysed in a particular moment. Samples and standards tubing were made of Teflon with 0.8 mm of internal diameter and the fix pump tubing was a “yellow-yellow 3-stop colour-coded” tubing made of Tygon® R3603 (ref. SC 0064) with an internal diameter of 2.3 mm, both from Ismatec. The flow rate for the samples and standards was 2.3 $mL \cdot min^{-1}$.

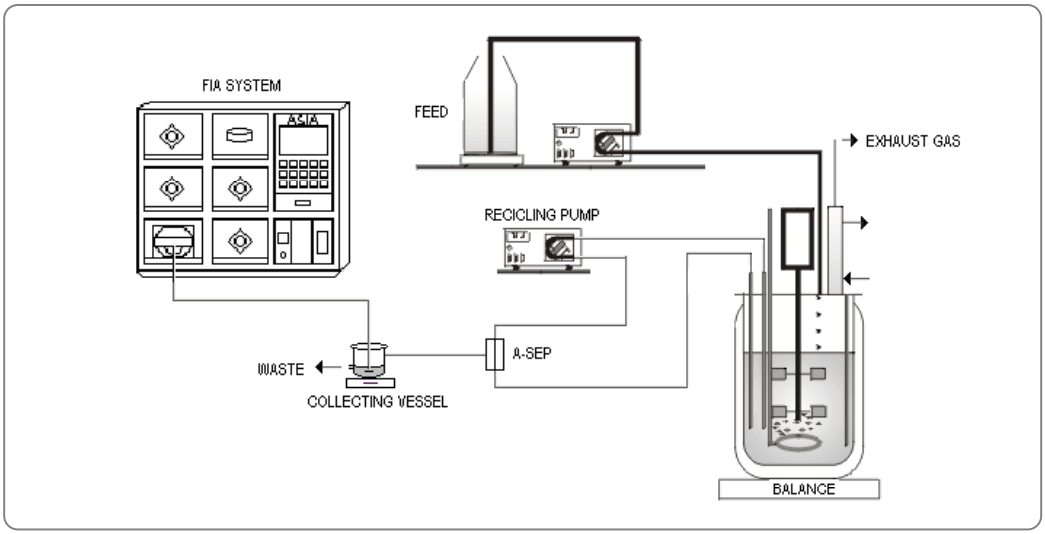


Figure 3.7 Bioreactor sampling system. The bio-suspension was extracted from the fermenter via a peristaltic pump and passed through the A-SEP device. The filtrate portion was conducted to a collecting vessel that excluded the exceeding volume. The remaining part of the bio-suspension returned to the fermenter.

The vario pump was used to drive the carrier and other reagent solutions into the detector. The material for carrier and reagent tubing was also Teflon with variable internal diameters, and vario pump tubing was a “2-stop colour-coded” type also made of Tygon® R3603.

The system is connected via a serial port with a personal computer equipped with the software package ASIA, from Ismatec. This software allows the control of all the relevant variables of the system, namely, pre- and loop load times, and analysis frequency. Calibration is also done by means of the software that injects all the standards in a pre-defined sequence, rejecting erroneous values and calculating the regression equation. During fermentations, the system can run in a fully automated way, based on routines written in PASCAL language. These routines determine the type of calibration, the number of samples to be analysed and the tests to be performed.

3.2.3 Acetate Measurement

The set-up for the measurement of acetate was adapted from the work by Forman et al. (1991), mentioned in section 3.1.5. Reagents were prepared according to Table 3.2. Standard preparation as well as a typical calibration can be found in Annex 1. The reagents used for this measurement were all of analytical grade. The set-up for the measurement is shown in Figure 3.8.

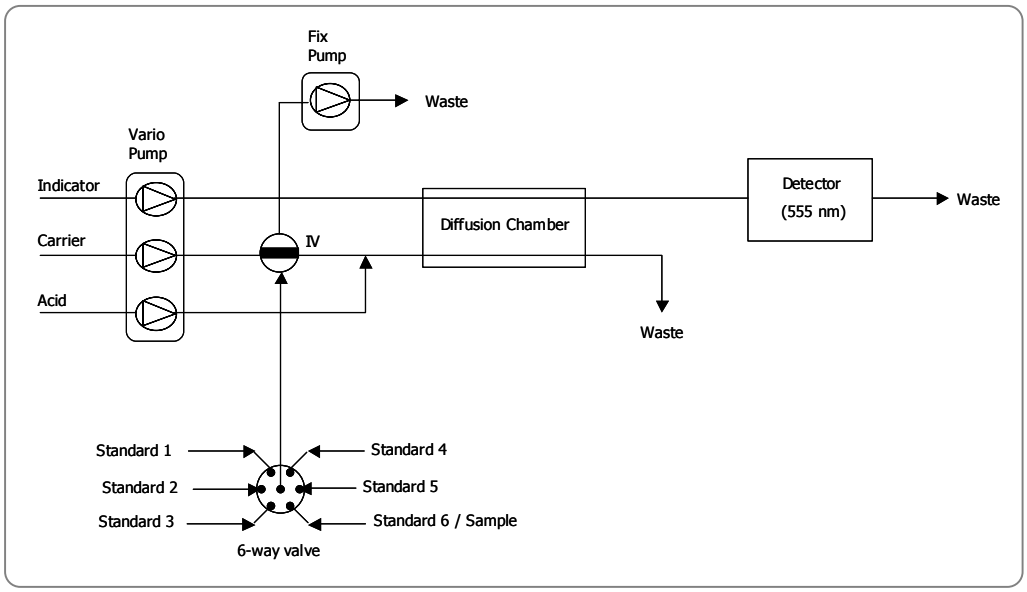


Figure 3.8 Set-up for the measurement of acetate. IV means Injection Valve.

Table 3.2 Reagents used for the analysis of acetate with FIA

Reagent	Stock solution	Final Solution
Sodium Phosphate (Na_2HPO_4)	Concentration: 0.25 M Preparation: - Dissolve 8.870 g with distilled water in a 250-mL volumetric flask - Fill with water - Store in the refrigerator (for 2 months maximum)	—
Phenol red	Concentration: 10 mM Preparation: - Dissolve 354 mg in 25 mL of a 70% ethanol solution in a 100-mL volumetric flask - Fill with distilled water - Filter in vacuum with a 0.45-μm membrane - Store in the refrigerator (for 2 months maximum)	Concentration: 80 μM Preparation: - Add 8 mL of stock solution in a 1-L volumetric flask - Add 1 to 5 mL of Sodium Phosphate stock solution - Fill with distilled water
Sulphuric acid 100%	—	Concentration: 1 M Preparation: - Add slowly 83 mL to 500 mL of distilled water in a 1-L volumetric flask - Fill with distilled water
Sodium Acetate ($3.H_2O$)	Concentration: 50 g acetate/kg solution Preparation: -Dissolve 57.583 g Sodium acetate 3-hydrate with water in a 500-mL volumetric flask - Add distilled water until 500 g - Store in the refrigerator (for 1 month maximum)	Several concentrations

Flow rates were optimised in order to achieve a compromise between sampling frequency and sample dispersion. The final values for the various streams are shown in Table 3.3. These rates allowed performing approximately 30 samples·h^{-1}.

In conclusion, several changes were made compared with the set-up proposed by Forman et al. (1991) shown in Figure 3.4:

- The wavelength of detection is 555 nm, instead of 560 nm.
- The injection volume was 10 μL, while in this work it was fixed in 50 μL.
- The phenol red concentration on the indicator solution proposed (0.1 M) was found to be too high, leading to a small sensitivity. The concentration used was of 80 μM.

- The acid concentration was kept at 1 M, while the phosphate concentration in the indicator stream was chosen to be an optimisation variable to increase the linear range or, on the other hand, to decrease the detection limit.

Table 3.3 Fix pump tubing diameters and flow rates for a pump speed of 20 rpm. Tube identification is based on the 2-stop colour-coded tubes from Ismatec

Stream	Tube identification	Internal Diameter (mm)	Flow rate ($mL \cdot min^{-1}$)
Carrier (water)	Orange-black	0.64	0.49
Acid	Orange-black	0.64	0.49
Indicator	White-white	1.02	1.14

3.2.4 Glucose Measurement

For glucose measurement, the set-up is shown in Figure 3.9. The sample containing the analyte was injected into a buffer carrier stream, which then reacted in a Glucose Oxidase and Lactonase enzymatic cartridge (Biocart Glucose LS - ref. 100012) from Anasycon (Germany) that converted glucose into gluconic acid and hydrogen peroxide. The concentration of this compound was then measured electrochemically with the amperometric detector. This detector is a self-filling wall-jet flow though cell with a volume of 5 μL and 3 electrodes (working, reference and counter electrodes). The working electrode is made of Platinum, while the reference electrode is an Ag/AgCl type with 3M of KCl as filling solution. The counter electrode is made of stainless steel. A constant potential of +600 mV is applied between the working and the reference electrode. The electric current resulting from the oxidation of hydrogen peroxide is then electronically converted and measured in mV, being proportional to the glucose concentration present in the sample. The detector allowed three different sensitivities for the measurement: 1, 10 and 1000 mA/V. The stability of the detector is only reached after one week of operation, being thus necessary to maintain the equipment on while it is not in use.

The composition of the buffer was initially of 17.9 $g \cdot L^{-1}$ of Na_2HPO_4 and 15.6 $g \cdot L^{-1}$ of $NaH_2PO_4.2H_2O$. The vario pump tubing was a “white-white” type, with an internal diameter of 1.02 mm and a flow rate of 1.14 $mL \cdot min^{-1}$ at 20 rpm. Analysis frequency was approximately 60 $samples \cdot h^{-1}$ Calibration standards preparation, as well as a typical calibration is also shown in Annex 1. The reagents used for this measurement were all of analytical grade.

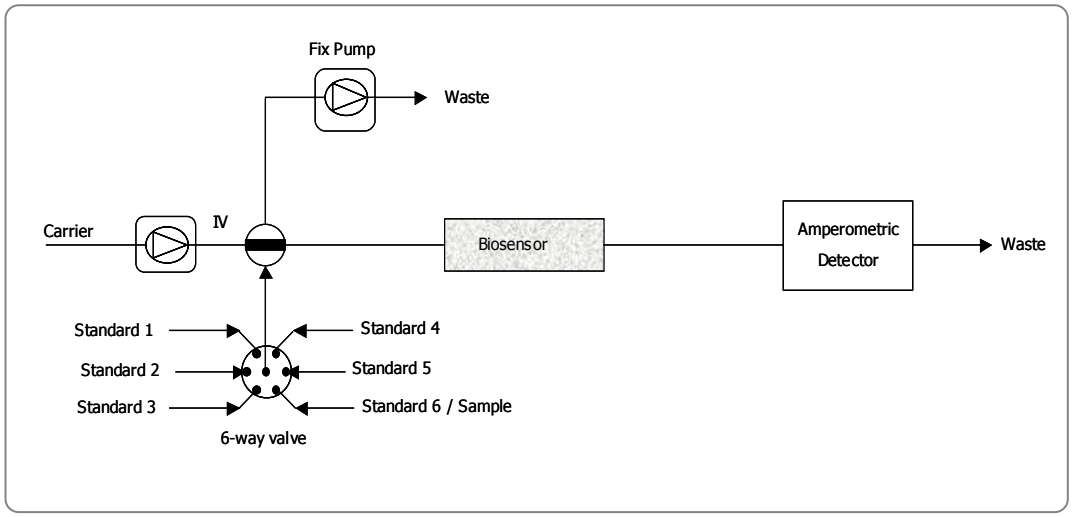


Figure 3.9 Set-up for the measurement of glucose with FIA. IV means Injection Valve.

3.2.5 Other Analytical Methods

For comparative purposes, acetic acid was also measured off-line with HPLC and with an enzymatic kit. The comparative methods selected for glucose analysis were the di-nitrosalicylic acid (DNS) method and an enzymatic kit.

A detailed description of these methods will be found in the next Chapter.

3.3 RESULTS AND DISCUSSION

3.3.1 Acetate Measurement Optimisation

The method proposed by Forman et al. (1991) included an alkali stripper to eliminate the carbon dioxide interference. However, because of the sample pre-treatment described in section 3.2.1, which included, besides the micro-filtration unit, a mixing vessel, carbon dioxide was not found to interfere greatly with acetate measurements.

During the optimisation of the acetate measurement with FIA, the main manipulated variable was the buffer capacity of the indicator solution, represented by its phosphate concentration. Several experiments were conducted at different buffer capacities of the indicator, covering a wide range of concentrations. Table 3.4 shows how the linearity and detection levels of the measurement can be influenced by that value. Method Detection Level (MDL) is defined as the constituent concentration that, when processed through the complete method, produces

a signal that has a 99% probability to be different from the blank (Clesceri et al., 1999). For the calculation of MDL in seven replicates of the sample, the condition was that mean had to be *3.14×s* above the blank, being *s* the standard deviation of the seven replicates. RSD was calculated, for each phosphate concentration, from 20 replicates of a standard in the middle of the linear range. The reference method for these experiments was the HPLC method.

Table 3.4 Influence of the buffer capacity of the indicator solution on the linearity and detection level for acetate measurements with FIA

Phosphate concentration (mM)	**Linearity limit ($g \cdot kg^{-1}$)**	**Equation of lines [Signal]=mV [Conc]=$g \cdot kg^{-1}$**	**R^2**	**MDL ($g \cdot kg^{-1}$)**	**RSD (%)**
0.250	2.5	Signal=–89.79×Conc – 7.97	0.9994	0.04	4.18
0.375	2.5	Signal=–125.38×Conc – 18.69	0.9971	0.02	4.56
0.500	5.0	Signal=–93.50×Conc – 26.04	0.9976	0.03	0.60
0.750	7.5	Signal=–76.41×Conc – 26.96	0.9973	0.04	2.56
1.000	10	Signal=–53.20×Conc – 7.39	0.9993	0.05	2.98
1.250	10	Signal=–46.23×Conc – 10.54	0.9995	0.10	1.26

The results obtained show the versatility of this method, and its applicability to processes where low acetate accumulations should be expected, such as batch or low cell densities fed-batch processes, and to other processes where acetate is present in a wide range of concentrations, like high-cell density fed-batch fermentation.

For batch fermentations, the recommended phosphate concentration would then be 0.375 mM, because at those buffer concentrations, the method presents a high sensitivity and a linearity limit sufficiently high to measure acetate concentrations in this kind of fermentation without a dilutions step. For evaluating the interferences that affect the proposed method when analysing acetate on fermentation samples, the correlation between HPLC and FIA measurements is presented in Figure 3.10. The results show a correlation coefficient very close to the unit (0.9879), indicating that the FIA method gives almost the same results as HPLC method.

For fed-batch processes, the most suitable phosphate concentration was considered to be 1 mM. With this phosphate concentration, it is possible to work in the linear zone of the method up to 10 $g \cdot kg^{-1}$, which corresponds to an acetate concentration commonly found in fed-batch fermentations. For this case, interferences affecting the method performance were also evaluated. The first approach was to compare the results obtained by analysing off-line

the same samples with FIA and with other conventional methods (HPLC and enzymatic kit). The results obtained for samples from a fed-batch fermentation are represented in Figure 3.11. It is clear that, for every sample, acetate concentrations obtained with FIA are slightly higher than those obtained with HPLC. This is mainly due to the non-selectivity of the method for acetate in complex samples mentioned in section 3.1.5. However, the trend along the fermentation is similar. The corresponding relative differences between the three methods were also calculated. The average deviation between FIA and HPLC is about 16%, oscillating between 4 and 36%, much less then the difference observed between HPLC and the enzymatic kit (with a minimum of 16, a maximum of 41 and an average of 27%). The magnitude of the deviation between HPLC and FIA was verified for other fed-batch fermentations and is around 20%.

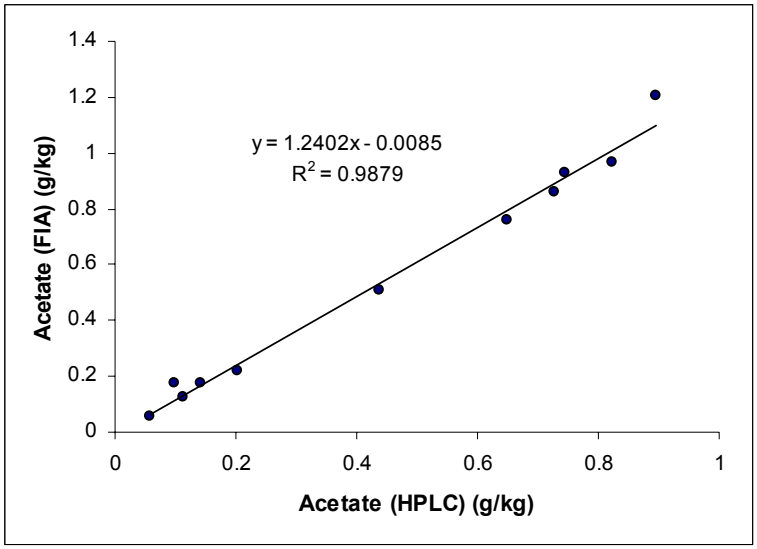


Figure 3.10 Correlation between FIA and HPLC measurements for a batch fermentation.

An example of another fermentation where the same samples were analysed off-line with both FIA and HPLC is shown in Figure 3.12. The particularity of this fermentation is that after reaching the FIA linearity limit (10 $g \cdot kg^{-1}$), samples were diluted before injection. It is clear that sample matrix interference is almost inexistent for high acetate concentrations, due to the dilution and also probably to the predominance of acetate when compared with other organic acids at that stage of the fermentation that are also detected by this technique.

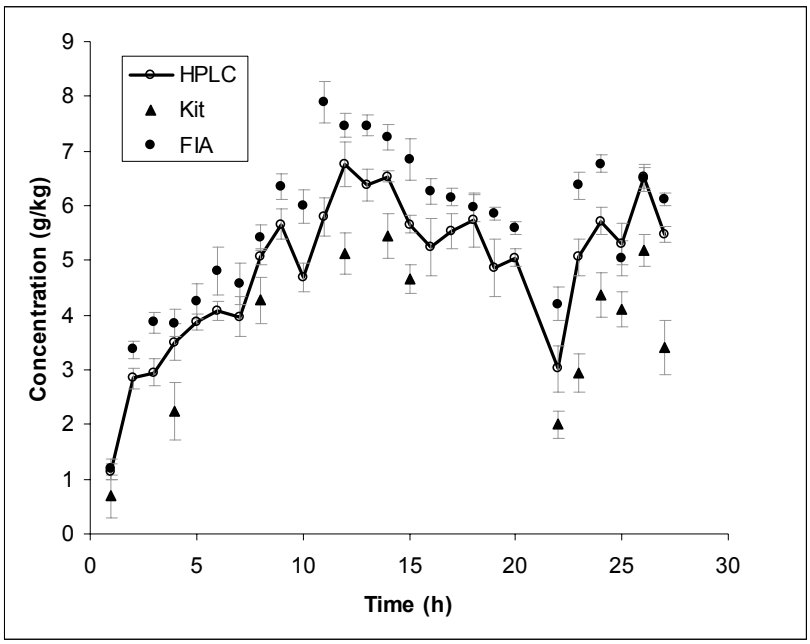


Figure 3.11 Comparison between several methods for the measurement of acetate in the course of a fed-batch fermentation. Error bars account for standard deviation between 2 (HPLC and enzymatic kit) or 3 samples (FIA).

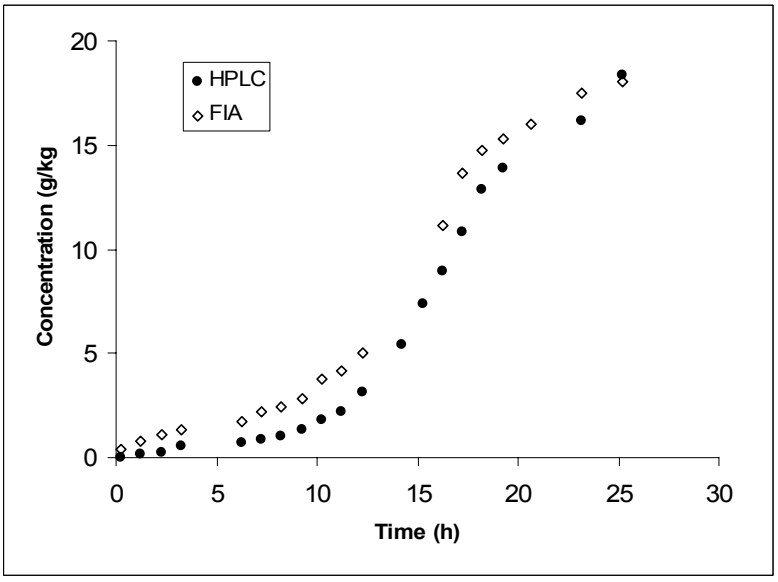


Figure 3.12 Comparison of FIA and HPLC results for a fermentation where acetate accumulated significantly and samples had to be diluted.

The linear correlation between both HPLC and FIA measurements for this case is shown in

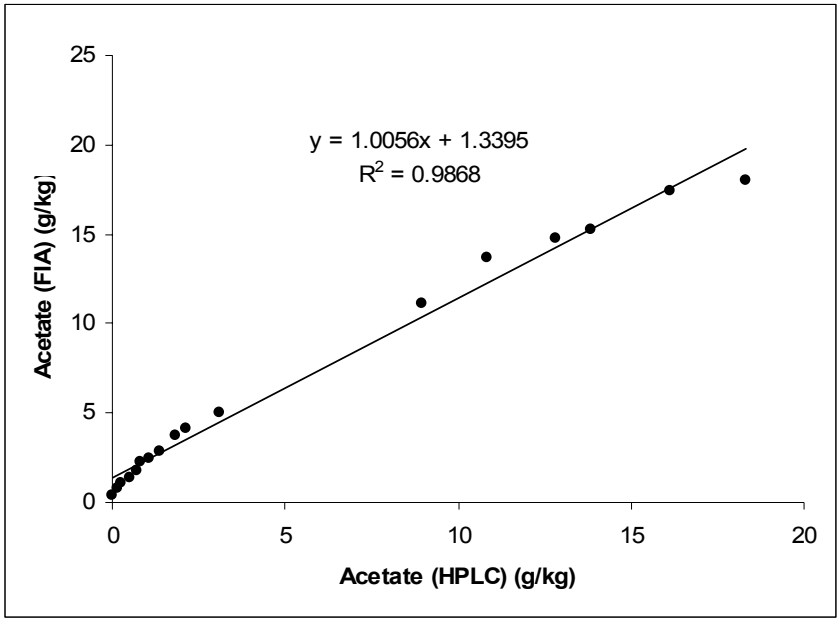


Figure 3.13: It can be concluded that, also for fed-batch samples, there is a good agreement between HPLC and FIA for acetate concentrations.

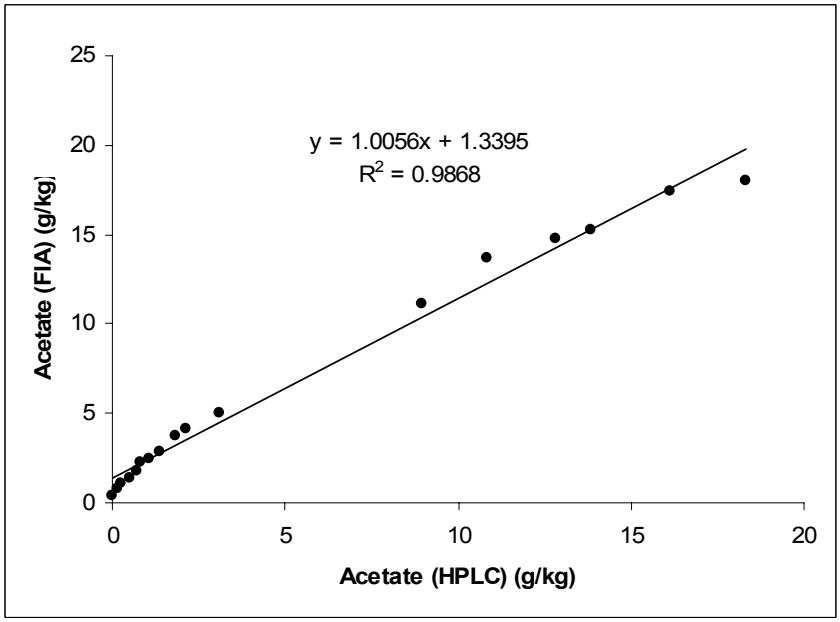


Figure 3.13 Linear correlation found between FIA and HPLC measurements for fed-batch samples.

Finally, the interference of the sampling system in the method's behaviour was evaluated, by comparing results obtained from FIA on- and off-line. These results are illustrated in Figure 3.14. Before the introduction of the collecting vessel in the sampling system, the filtrate was accumulated inside an ordinary recipient, without agitation or filtrate rejection. The differences between on- and off-line analyses are notorious, due to the accumulation phenomenon. After the introduction of the collecting vessel, on- and off-line results became very close, with no significant differences.

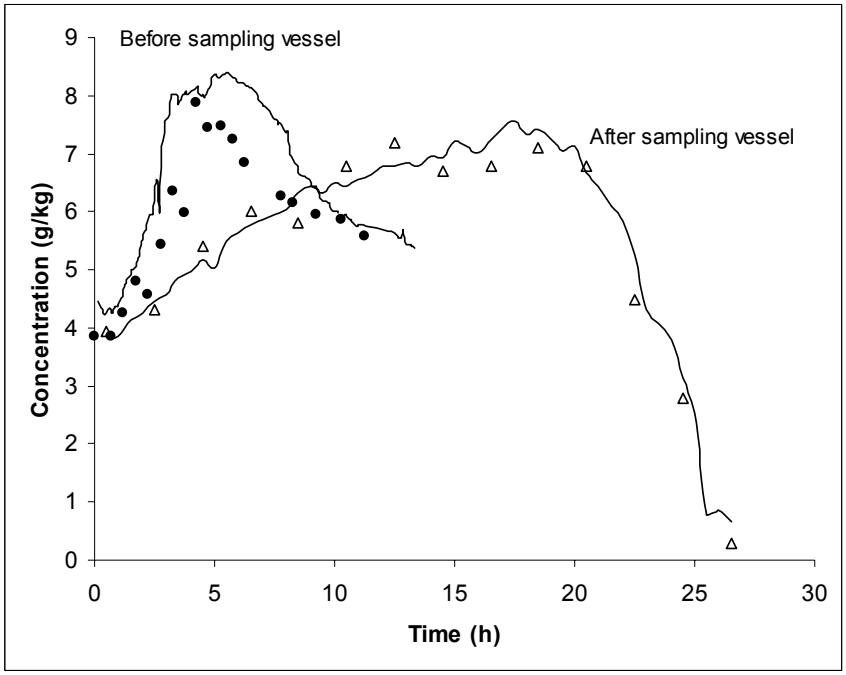


Figure 3.14 Comparison between both on-line (lines) and off-line (• and △) FIA acetate analyses before and after the introduction of the sampling vessel in the system.

3.3.2 Glucose Measurement Optimisation

The main purpose of this optimisation procedure was to increase the linearity limit of the method, in order to be able to determine both glucose and acetate in the same sample during a conventional *E. coli* fermentation. With a carrier composition of 17.9 $g \cdot L^{-1}$ of Na_2HPO_2, 15.6 $g \cdot L^{-1}$ of $NaH_2PO_4.2H_2O$, and 2 $mL \cdot L^{-1}$ of a 20% sodium azide solution, as recommended by the method described by Ismatec (1997), both carrier pH and pump speed were varied and the method detection level and linearity limit were determined. pH was varied by the addition

of sodium hydroxide or phosphoric acid to the carrier. The results obtained are represented on Figure 3.15 and Figure 3.16.

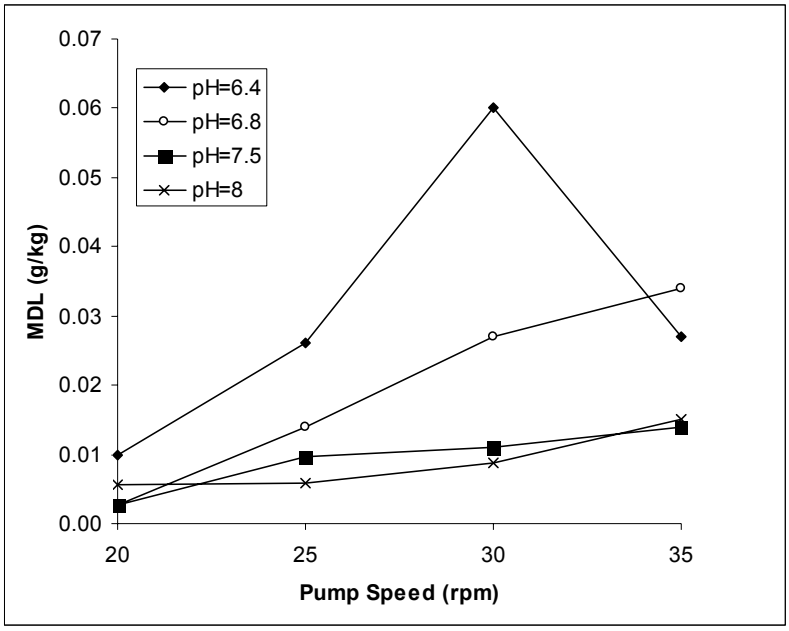


Figure 3.15 Variation of MDL with pH and pump speed for the glucose method.

Results show that at pH 6.4 it is possible to achieve much higher linearity limits, avoiding the need of introducing a dilution step in the measurement system. Also, the corresponding MDL was considered to be sufficient for this process, because glucose concentrations are not expected to drop below 0.02 $g \cdot kg^{-1}$. Pump speed was not chosen above 25 rpm in order to save carrier solution and also to make this method compatible with acetate method for simultaneous measurements.

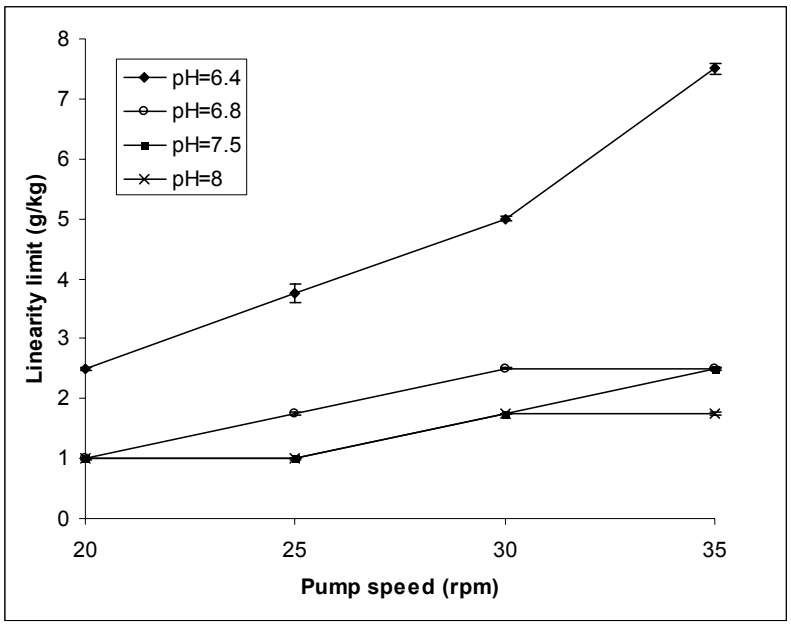


Figure 3.16 Linearity limits of the method as a function of pH and pump speed for the glucose method. Error bars represent the standard deviation between 5 injections.

However, when these conditions were applied to real samples obtained from *E. coli* fed-batch fermentation, a very strong interference was observed (Figure 3.17), that can not be explained by the complexity of the samples. After some experiments, it was found that the presence of sodium azide in the carrier was the main cause of these discrepancies. In fact, this compound somehow inhibited the enzyme glucose oxidase and this effect was attenuated in the presence of sample components. This inhibitor was primarily added to the carrier in order to avoid the clogging of the system, and to inhibit growth in the tubing walls. However, at this point, sodium azide had to be excluded from the carrier, while some cleaning procedures had to be implemented in the system on a daily basis to avoid the above-mentioned problems. The sensitivity of the detector had also to be changed from 1 to 10 mA/V.

Without sodium azide, further carrier optimisation had to be conducted in order to eliminate other interferences. It was found that the matrix effects of the samples were significant, and the contribution of each medium component to that interference was evaluated. In Figure 3.18, the differences observed in terms of detector response for a 1 $g \cdot kg^{-1}$ standard are shown. The first column of the chart corresponds to the signal obtained with a carrier composed of phosphates only. After that, the other medium growth components were added sequentially to check for individual interferences. It is clear that ammonium cloride is the

major contributor for the observed signal reduction. As already mentioned in section 3.1.4, ammonium is known to consume hydrogen peroxide and, consequently, to cause a decrease in the sensitivity of the method. This observation has to be taken into account in the interpretation of glucose FIA measurements, as ammonium concentration along the fermentation is not constant, due to its consumption by cells and also to the addition of that compound for pH control.

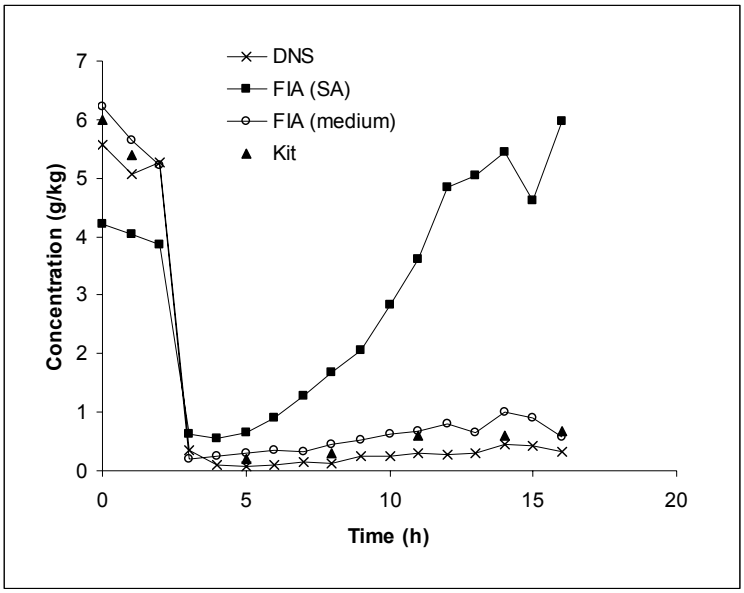


Figure 3.17 Comparison of DNS, FIA and enzymatic kit glucose analyses in fermentation samples. FIA (SA) means that the carrier contained sodium azide. FIA (medium) represents the analyses taken with the carrier composition similar to the growth medium.

To minimize those matrix effects, another carrier composition is proposed, combining the phosphates at the mentioned concentrations and the medium components referred in Figure 3.18. Thus, the above-mentioned samples were analysed with this new carrier composition and the results are also shown in Figure 3.17 (FIA medium). Although the trend is similar to the ones observed for the other methods, there is still a significant difference, especially when comparing with DNS data. However, since these concentrations did not cover the whole range of glucose method and were not obtained on-line, the reliability of glucose method was checked during the simultaneous determination, that is described in the next section.

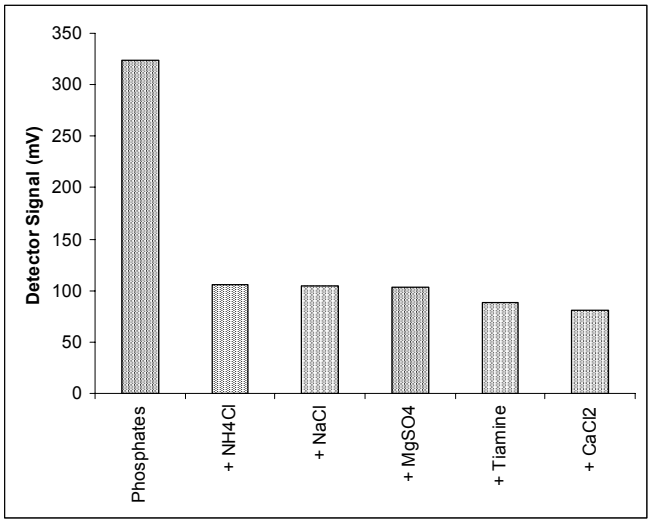


Figure 3.18 Contribution of the individual medium components to the observed matrix effect during glucose analysis. Each component was added sequentially to the carrier at the concentrations presented in the growth medium. The results presented correspond to the analysis of a 1 $g \cdot kg^{-1}$ glucose standard.

3.3.3 Simultaneous Determination

After the optimisation procedure described in the previous sections, both parallel and serial configurations were evaluated. The advantages of each case were discussed in section 3.1.6.

For the serial configuration, it was established that the samples would pass by glucose biosensor in the first place, in order to preserve the enzymatic cartridge from deterioration with sulphuric acid or phenol red. Thus, glucose method remained practically in the same arrangement as shown in Figure 3.9. However, for acetate measurement under this configuration, the carrier had to be switched from distilled water to phosphate buffer. Also, it should be expected that a higher dispersion occurs, caused by the increase in the length of the path from injection to detection. In order to evaluate the loss of sensitivity caused by this phenomenon, an experience was carried out that consisted on the comparison of the method response in the isolated case and in the serial configuration. The results obtained are shown in Table 3.5. It is clear that the linearity of the method was maintained, but the sensitivity decreased into almost half of its original value.

Table 3.5 Comparison of the performance of FIA acetate method before and after the serial configuration implementation. Signal is given in mV and concentration in g·kg^{-1}

	Isolated Method	**Serial Configuration**
Equation of lines	Signal=–74.79×Conc +4.62	Signal=–31.41×Conc +4.21
R^2	0.9998	0.9989
Linearity limit (g·kg^{-1})	10	10

In order to check for the influence of glucose analysis in acetate measurements, another experience was performed, that consisted in the injection of solutions containing variable glucose concentrations, while acetate concentration was kept unchanged. Results obtained from that experiment are shown in Table 3.6. It can be concluded that glucose measurement does not influence the response of the photometric detector under the serial configuration.

Table 3.6 Influence of glucose measurement in acetate method under the serial configuration. RSD refers to 5 injections of the same sample

Acetate (g·kg^{-1})	**Glucose (g·kg^{-1})**	**Average (mV)**	**RSD (%)**
	0.00	-240.26	1.41
	0.50	-236.32	2.25
5.00	1.25	.241.54	3.51
	2.50	-241.96	2.24
	5.00	-235.26	2.25

For parallel configuration, the only requirement was that the pump speed had to be the same for both analyses. This had been contemplated during optimisation, and the adopted value was 25 rpm. Sample injection was done simultaneously in both valves, with the help of a “T” configuration after the 6-way valve. The major disadvantage of this configuration was consumption of twice the sample volumes, when compared with the individual methods or with serial implementation. However, the sampling system was always able to provide enough quantities for the analyses, and thus this configuration was adopted for the process studied. Nevertheless, serial configuration should be considered as a valid alternative for situations where sample volume is scarce or when the FIA system does not accommodate a sufficient number of devices like valves.

The adopted configuration is shown in Figure 3.19. It allowed both glucose and acetate determinations in one sample within 2 min. The concentration values could then be obtained with a total delay of 3 minutes (including one minute for the sampling retrieval).

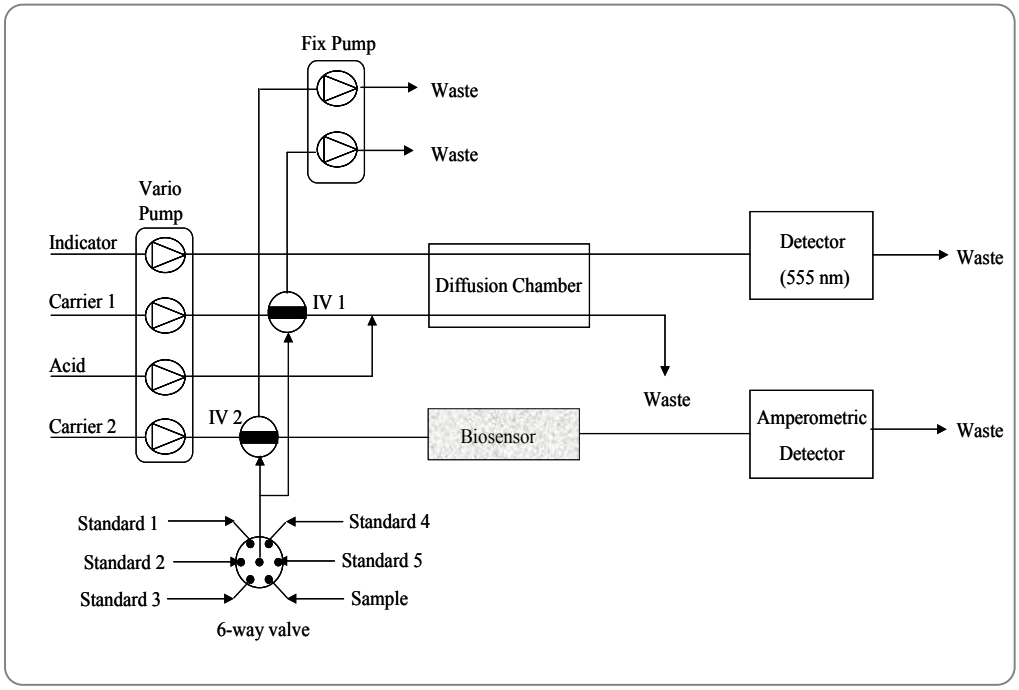


Figure 3.19 Parallel configuration for the simultaneous determination of acetate and glucose.

In order to validate the on-line simultaneous measurement of glucose and acetate, a fermentation was conducted and on-line results were compared with those obtained with conventional methods (HPLC for acetate and DNS for glucose). For this fermentation, the feeding was programmed with the purpose of obtaining glucose and acetate accumulation, in order to test both methods in a wide range of concentrations. When the metabolites accumulated in the medium at concentrations considered significantly high, the feeding was stopped to check for the measurements' robustness at high cellular concentrations, when certain fermentation by-products were present. Consumption of both components (and concomitant concentrations decline) started right after the stop of the feeding. The final cell dry weight obtained was of 40 $g \cdot kg^{-1}$, representing approximately 55 $g \cdot L^{-1}$. These results are illustrated in Figure 3.20. For both methods, a good approximation between on-line FIA measurements and off-line reference methods was obtained.

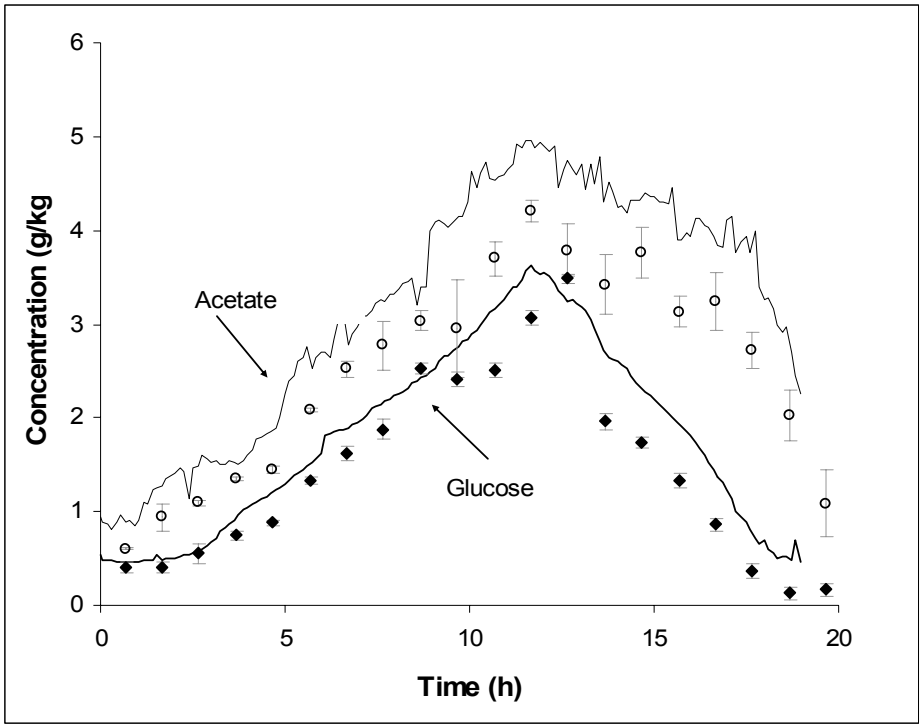


Figure 3.20 Validation of on-line simultaneous measurement of glucose and acetate during a fed-batch *E. coli* fermentation. The symbols represent off-line analyses performed with HPLC (for acetate) and DNS (for glucose), while the lines correspond to on-line measurements with FIA.

3.4 CONCLUSIONS

A Flow Injection Analysis method for the simultaneous on-line measurement of glucose and acetate was developed. Optimisation procedures allowed the elimination of major interferences by varying carrier and other reagents composition. However, both methods still exhibit some differences with respect to other reference methods, but their performances were considered satisfactory for on-line monitoring and control purposes.

A sampling system was also implemented that allowed the supply of cell-free samples to the analytical system with minimised delays.

Due to the flexibility of both methods, it is possible, without a dilution step and, by only varying carrier compositions, to apply this system to a diversity of fermentation processes, where both components can be present in a wide range of concentrations.

Both parallel and serial implementation for the simultaneous determination of glucose and acetate were investigated, and it was concluded that the former presents important advantages related with analysis time and ease of implementation. However, serial implementation can also be used when the availability of cell-free samples is scarce.

The low cost of reagents, together with the simplicity of the system, makes this technique an attractive alternative for the on-line monitoring of biotechnological processes, like the production of recombinant products with *E. coli*.

3.5 REFERENCES

Becker, T., Kittsteiner-Eberle, R., Luck, T., and Schmidt, H.-L. On-line determination of acetic acid in a continuous production of *Acetobacter aceticus*. *Journal of Biotechnology.* 31, 267-275. 1993.

Becker, T. and Schmidt, H.-L. New ways of enzymatic two-substrate determinations in flow injection systems. *Analytica Chimica Acta.* 421, 7-18. 2000.

Bilitewski, U., Drewes, W., Neermann, J., Schrader, J., Surkow, R., Schmid, R. D., and Bradley, J. Comparison of different biosensor systems suitable for bioprocess monitoring. *Journal of Biotechnology.* 31, 257-266. 1993.

Bradley, J., Stöcklein, W., and Schmid, R. D. Biochemistry based analysis systems for bioprocess monitoring and control. *Process Control and Quality.* 1, 157-183. 1991.

Castañon-Fernández, J., Fernández-Abedul, M. T., and Costa-Garcia, A. Determination of acid phosphatase activity in a double injection flow system with electrochemical detection. *Analytica Chimica Acta.* 406, 225-232. 2000.

Céspedes, F., Valero, F., Martínez-Fàbregas, E., Bartolí, J., and Alegret, S. Fermentation monitoring using a glucose biosensor based on an electrocatalytically bulk-modified epoxy-graphite biocomposite integrated in a flow system. *Analyst.* 120, 2255-2258. 1995.

Christensen, L. H., Marcher, J., Schulze, U., Carlsen, M., Min, R. W., Nielsen, J., and Viladsen, J. Semi-on-line analysis for fast and precise monitoring of bioreaction processes. *Biotechnology and Bioengineering.* 52, 237-247. 1999.

Christensen, L. H., Nielsen, J., and Viladsen, J. Monitoring of substrates and products during fed-batch penicillin fermentations on complex media. *Analytica Chimica Acta.* 249, 123-136. 1991.

Chung, S., Wen, X., Vilholm, K., Bang, M., Christian, G., and Ruzicka, J. Novel flow-injection analysis method for bioprocess monitoring. *Analytica Chimica Acta.* 249, 77-85. 1991.

Clesceri, L., Greenberg, A., and Eaton, A. *Standard methods for the examination of water and wastewater*. American Public Health Association, Washinton DC. 1999.

Ding, T., Bilitewski, U., Schmid, R. D., Korz, D. J., and Sanders, E. A. Control of microbial activity by flow injection analysis during high cell density cultivation of *Escherichia coli*. *Journal of Biotechnology.* 27, 143-157. 1993.

Fang, Z.-L. Trends of flow injection sample pretreatment approaching the new millennium. *Analytica Chimica Acta.* 400, 233-247. 1999.

Forman, L. W., Thomas, B. D., and Jacobson, F. S. On-line monitoring and control of fermentation processes by flow-injection analysis. *Analytica Chimica Acta.* 249, 101-111. 1991.

Freitag, R., Fenge, C., Scheper, T., Schügerl, K., Spreinat, A., Antranikian, G., and Fraune, E. Immunological on-line detection of specific proteins during fermentation processes. *Analytica Chimica Acta.* 249, 113-122. 1991.

Gabriel, D., Baeza, J., Valero, F., and Lafuente, J. A novel FIA configuration for the simultaneous determination of nitrate and nitrite and its use for monitoring an urban waste water treatment plant based on N/D criteria. *Analytica Chimica Acta.* 359, 173-183. 1998.

Garn, M., Gisin, M., Thommen, C., and Cevey, P. A Flow Injection Analysis System for fermentation monitoring and control. *Biotechnology and Bioengineering.* 34, 423-428. 1989.

Gründig, B., Strehlitz, B., Kotte, H., and Ethner, K. Development of a process-FIA system using mediator-modified enzyme electrodes. *Journal of Biotechnology.* 31, 277-287. 1993.

Hall, E. A. H. Flow injection analysis with immobilized reagents. *Current Opinion in Biotechnology.* 2, 9-16. 1991.

Haouz, A. and Stieg, S. Continuous monitoring of D-glucose and lactate by flow injection analysis. *Enzyme and Microbial Technology.* 30, 129-133. 2002.

Hitzmann, B., Ritzka, A., Ulber, R., Scheper, T., and Schügerl, K. Computational neural networks for the evaluation of biosensor FIA measurements. *Analytica Chimica Acta.* 348, 135-141. 1997.

Hitzmann, B., Ritzka, A., Ulber, R., Schöngarth, K., and Broxtermann, O. Neural networks as a modeling tool for the evaluation and analysis of FIA signals. *Journal of Biotechnology.* 65, 15-22. 1998.

Huang, Y. L., Foellmer, T. J., Ang, K. C., Khoo, S. B., and Yap, M. S. G. Characterization and application of an on-line flow injection analysis/wall-jet electrode system for glucose monitoring during fermentation. *Analytica Chimica Acta.* 317, 223-232. 1995.

Ismatec, SA. ASIA Methods - Glucose. 1997.

Lenarczuk, T., Wencel, D., Glab, S., and Koncki, R. Prussian blue-based optical glucose biosensor in flow-injection analysis. *Analytica Chimica Acta.* 447, 23-32. 2001.

Lüdi, H., Garn, M. B., Batailard, P., and Widmer, H. M. Flow injection analysis and biosensors: applications for biotechnology and environmental control. *Journal of Biotechnology.* 14, 71-79. 1990.

Matsumoto, K., Matsubara, H., Hamada, M., Ukeda, H., and Osajima, Y. Simultaneous determination of glucose, ethanol and lactate in alcoholic beverages and serum by amperometric flow injection analysis with immobilized enzyme reactors. *Journal of Biotechnology.* 14, 115-126. 1990.

Min, R. W., Rajendran, V., Larsson, N., Gorton, L., Planas, J., and Hahn-Hägerdal, B. Simultaneous monitoring of glucose and L-lactic acid during a fermentation process in an aqueous two-phase system by on-line FIA with microdialysis sampling and dual biosensor detection. *Analytica Chimica Acta.* 366, 127-135. 1998.

Nielsen, J., Nikolajsen, K., Benthin, S., and Viladsen, J. Application of flow-injection analysis in the on-line monitoring of sugars, lactic acid, protein and biomass during lactic acid fermentations. *Analytica Chimica Acta.* 237, 365-175. 1990.

Ogbomo, I., Prinzing, U., and Schmidt, H.-L. Prerequisites for the on-line control of microbial processes by flow injection analysis. *Journal of Biotechnology.* 14, 63-70. 1990.

Riedel, K., Renneberg, R., Wollenberger, U., Kaiser, G., and Scheller, F. W. Microbial sensors - fundaments and application for process control. *Journal of Chemical Technology and Biotechnology.* 44, 85-106. 1989.

Ruzicka, J. The second coming of flow-injection analysis. *Analytica Chimica Acta.* 261, 3-10. 1992.

Ruzicka, J. and Hansen, E. H. *Flow Injection Analysis*. John Wiley & Sons, New York. 1987.

Schmid, R. D. and Künnecke, W. Flow injection analysis (FIA) based on enzymes or antibodies - applications in the life sciences. *Journal of Biotechnology.* 14, 3-31. 1990.

Schöngarth, K. and Hitzmann, B. Simultaneous calibration in flow-injection signals evaluated by partial least squares. *Analytica Chimica Acta.* 363, 183-189. 1998.

Schügerl, K. Which requirements do flow injection analyzer/biosensor systems have to meet for controlling the bioprocess. *Journal of Biotechnology.* 31, 241-256. 1993.

Schügerl, K., Hitzmann, B., Jurgens, H., Kullick, T., Ulber, R., and Weigal, B. Challenges in integrating biosensors and FIA for on-line monitoring and control. *Trends in Biotechnology.* 14, 21-31. 1996.

Spohn, U., Preuschoff, F., Blankenstein, G., Janazek, D., Kula, M. R., and Hacker, A. Chemiluminometric enzyme sensors for flow-injection analysis. *Analytica Chimica Acta.* 303, 109-120. 1995.

Tang, X.-J., Tocaj, A., Holst, O., and Johansson, G. Process monitoring of acetic acid in *Escherichia coli* cultivation using electrochemical detection in a flow injection system. *Biotechnology Techniques.* 11, 683-687. 1997.

Tothill, I. E., Newman, J. D., White, S. F., and Turner, P. F. Monitoring of the glucose concentration during microbial fermentation using a novel mass-producible biosensor suitable for on-line use. *Enzyme and Microbial Technology.* 20, 290-596. 1997.

Tservistas, M., Weigel, B., and Schügerl, K. An on-line flow-injection analysis system for the determination of acetate. *Analytica Chimica Acta.* 316, 117-120. 1995.

Valero, F., Lafuente, J., Poch, M., Solá, C., Araújo, A. N., and Lima, J. L. On-line fermentation monitoring using flow injection analysis. *Biotechnology and Bioengineering.* 36, 647-651. 1990.

Van Akker, E. B., Bos, M., and Van der Linden, W. E. Continuous, pulsed and stopped flow in a μ-flow injection system (numerical vs experimental). *Analytica Chimica Acta.* 378, 111-117. 1999.

Van Brunt, J. Biosensors for bioprocesses. *BIO/TECHNOLOGY.* 5, 437-440. 1987.

Van der Pol, J., Gooijer, C. D., Biselli, M., Wandrey, C., and Tramper, J. Automation of selective assays for on-line bioprocess monitoring by flow-injection analysis. *Trends in Biotechnology.* 14, 472-477. 1996.

Vodopivec, M., Berovic, M., Jancar, J., Podgornik, A., and Strancar, A. Application of convective interaction media disks with immobilised glucose oxidase for on-line glucose measurements. *Analytica Chimica Acta.* 407, 105-110. 2000.

Wu, X. and Bellgardt, K.-H. Real-time recursive parameter estimation for fault detection in flow-injection analysis systems. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 29-34. 1995.

Yamamoto, K., Ohgaru, T., Torimura, M., Kinoshita, H., Kano, K., and Ikeda, T. Highly sensitive flow injection determination of hydrogen peroxide with a peroxidase-immobilized electrode and its application to clinical chemistry. *Analytica Chimica Acta.* 406, 201-207. 2000.

Zigova, J., Mahle, M., Paschold, H., Malissard, M., Berger, E. G., and Weusterbotz, D. Fed-batch production of a soluble Beta-1,4- galactosyltransferase with *Saccharomyces cerevisiae*. *Enzyme and Microbial Technology.* 25, 201-207. 1999.

CHAPTER 4
MONITORING SYSTEMS AND METHODS

"There does not exist a category of science to which one can give the name applied science. There are science and the applications of science, bound together as the fruit of the tree which bears it."

Luis Pasteur

The system for the on-line monitoring and control of the major state variables during *E. coli* fed-batch fermentation is described in this chapter. The fermentation vessel and the main features of its Digital Control Unit are depicted as well as the balances used for weighting the fermentation culture and the feeding solution. Mass Spectrometry was used for analysis of the in- and off-gas of the bioreactor, and the main aspects of this kind of measurements are discussed. These on-line measurements, together with data from the FIA system were integrated in a supervisory computer using the graphical programming environment LabVIEW.
Finally, the methods used for performing off-line analysis and the main steps undertaken to run a fed-batch fermentation are also described.

4.1 INTRODUCTION

The rapid development in biotechnology during the last few years was not always followed by a corresponding adaptation and improvement of the bioprocess engineering to the new demands. It quickly became clear that more efficient process monitoring was needed, to allow better process modelling and closer process control. The difference between on- and off-line analysis and the different types of signals generated with both can be found in Table 4.1.

Table 4.1 Concepts associated with on- and off-line measurements (Sonnleitner, 1999)

On-line = fully automatic	**Off-line = manual**
In-situ / bypass	Off-site
	After sampling
Continuous	Discrete = discontinuous
Real time	Delay, dead time
Non-invasive	Destructive
Closed-loop control	Open-loop control

In fact, when compared with other disciplines of engineering, sensors useful for *in situ* and/or on-line monitoring of biotechnological processes are comparatively scarce and they usually measure physical and chemical variables rather than biological ones. Typical on-line measurements include temperature, pH, various gas compositions, and flow rates of both liquid and gaseous streams. On-line measurement of substrates or products is quite unusual due to the scarcity of commercially available sensors. Extensive revisions about the main instrumentation used in bioprocesses can be found in Sonnleitner (1996); Sonnleitner (1997); Scheper et al. (1999); Sonnleitner (1999) and Schügerl (2001).

Besides the lack of reliable on-line sensors, another problem arises from the difficulty of coupling those few on-line measurement devices to computer systems and integrating those measurements in order to carry out supervisory process control. This difficulty is due to the existing diversity among the different sensors available.

Physical fermentation variables like pH, temperature, dissolved oxygen concentration, airflow rate and agitation speed, are usually directly controlled by bioreactor control units, and the first difficulty related to the computer data acquisition is precisely due to the specificities of

the communication protocols used by the different fermenter manufacturers. One way to acquire data from those control units is to use one of the several commercial software programs available from those manufacturers.

For example, the MFCS system, commercialised by B.Braun Biotech International (Germany), first developed to run in the DOS environment and now available for Windows™ is exclusive for the BIOSTAT bioreactors. An example of the application of this software for the monitoring and control of *E. coli* fermentation can be found in Pfaff et al. (1995). However, although this software is not limited to data acquisition from the Digital Control Unit (DCU) of the fermenter, it is very difficult if not impossible to implement sophisticated control algorithms or to integrate data from other devices in the same program. For example, those authors also analysed the composition of the exhaust gas and glucose concentration with FIA but those measurements were not integrated within the MFCS software with the other variables, and the algorithm for correcting glucose measurements had to be constructed in a separate program. A similar situation can be found in Huang et al. (1995), where the measured values of glucose concentration could not be integrated within the software, and thus the control action was implemented manually depending on the values of the measurements.

The MFCS software was also tested for the process described in this thesis, but it revealed inadequate for the purposes of data integration and adaptive control of the fermentation.

An alternative to control this kind of fermenters was used by Ding et al. (1993) and Korz et al. (1995): the UBICON (Universal Bio-process Control System) software developed by Electronic Systems Design, Germany. This software is able to integrate DCU measurements with exhaust gas analysis results from the paramagnetic and infrared gas analysis systems, mass flow meters, balances, and feeding pumps. It is also capable of performing data monitoring and storage, graphical representation, host communication, calculation of process variables and control loops for process variables (specific growth rate). Nevertheless, in the last case, FIA measurements that were conducted were not connected to that system. Moreover, this software is no longer commercialized.

New Brunswick Scientific Inc. (USA), another important manufacturer of fermentation units, commercialized the Advanced Fermentation Software, AFS to control some of its bioreactors. This software application was used by different authors, namely Yang (1992), Shin et al. (1997) and Claes and Van Impe (1998). However, all the three articles describe only very limited applications of this software, such as data recording and visualization and open-loop control of the feeding rate.

This software has been replaced by BioCommand, which is claimed to allow the development of Visual Basic based programs for advanced control applications. Still, the main disadvantage of this software is that it is specific for New Brunswick bioreactors.

Other types of commercial software were used by Dubach and Märkl (1992) and Gregory and Turner (1993), only to monitor and control standard variables like pH, temperature, and dissolved oxygen.

The FermExpert software developed by BioExpert Ltd. (Estonia) was developed for the application in different kinds of bioreactors and was used by Paalme et al. (1995) and Tomson et al. (1995). It allowed the integration of basic variables like temperature, pH, dissolved oxygen, among others, with the results from exhaust gas analysis. However, no sophisticated control algorithm was adopted in any of the cases.

A different solution is the application of expert systems to the monitoring and control of fermentation processes. For example, the G2 real-time expert system (Gensym Corporation, USA) was used by Fowler et al. (1992) in a pilot plant fermentation process. This product allows the management of the process by means of a model-based expert system (fuzzy rules) approach. Another knowledge-based diagnosis system was developed using a fuzzy expert system by Nakajima et al. (1992) for lactic acid fermentation.

An obvious solution for the stiffness of the commercial software packages is the development of dedicated software in several programming languages. However, this solution is often very time consuming and requires, besides strong programming skills, knowledge of the communication protocols used by the several devices that exist in a fermentation system, which is not always the case for the control units of the fermenters.

Nevertheless, several examples can be found in the literature concerning the construction of data acquisition and control solutions in several programming languages. Paalme et al. (1989) and Hsiao et al. (1990) developed different data acquisition software programs that allowed the acquisition of data from the bioreactors control units as well as from the exhaust gas analysis. In the first case, it was used for the implementation of an adaptive controller for the dissolved oxygen concentration in fed-batch *E. coli* fermentation.

Wipf et al. (1994) developed a FORTRAN based software to acquire data from several analog and digital controllers regarding temperature, pressure, volume, aeration rates, and automated sterilization cycles during a pilot-scale recombinant *E. coli* fermentation. Also, data from the off-gas carbon dioxide infrared and the oxygen paramagnetic measurements were acquired. This system allowed for data recording and additionally a supervisory control with

complex algorithms. It also allowed the integration of the off-gas data into set-point calculations with on line process data.

A similar approach can be found in Fagervik et al. (1998) in the so-called Fermentation Monitoring and Modeling System – FMMS. It is a model-based on-line fermenter supervisory system constructed in Visual Basic that integrates signal from the Digital Control Unit of the fermenter and from off-gas analyses.

In Konstantinov et al. (1990) and Konstantinov et al. (1991), the authors developed an application system for the fed-batch *E. coli* fermentation, designed especially for the development of new strategies for control of fermentation processes, that performed the following on-line functions: measurement, filtration, indirect variables calculation, control, data representation, printing, and archiving. These functions were performed as separated tasks written in C programming language. The software integrated data from the fermenter with off-gas analysis and feed balances.

Still, the integration of other on-line measurements related to substrate or product concentrations is not easily found in the literature. Nevertheless, Saucedo et al. (1995) constructed a software solution that integrated HPLC measurements with the important variables during fed-batch *E. coli* fermentation. The information obtained was used in a closed loop fed-batch fermentation operation to optimise the ethanol productivity by keeping the limiting substrate concentration in the fermenter equal to a predetermined optimum value. A similar approach was used by Kleman and Strohl (1994) to control the glucose in the fermentation medium during an *E. coli* fermentation using an on-line enzymatic glucose analyser.

Other examples include the control of a toxic substrate in *Pseudomonas cepacia* fermentation through on-line measurements of the substrate using a UV-spectrophotometer by Hagander and Holst (1992) and on-line determination of biomass by optical density and the integration of these measurements with off-gas data and other basic variables by Hopkins et al. (1987).

The integration of FIA measurements with other on-line physical and chemical data from the bioreactor was accomplished by the software FIACRE described in Busch et al. (1993).

Besides the solutions already mentioned, there are systems that combine the capacity of handling changing requirements of the user-developed software solutions with the ease of manipulation of the commercial software. Among them, the most commonly used are Labtech Notebook (Laboratory Technologies Corporation, USA) and LabVIEW (National Instruments, USA) development environments. These are graphical programming languages that already

have many data acquisition functions programmed that can be easily integrated in simple or advanced acquisition and control systems.

For example, in Lee et al. (1997), the authors used the Labtech Notebook for constructing a system for data acquisition and on-line display of cell concentration (from an on-line cell density device), dissolved oxygen, and pH (from a New Brunswick fermenter control unit). The two feed rate signals of glucose and ammonium hydroxide were also generated by this software.

The LabVIEW software has been used as an easy-to-use, powerful, and fast object-oriented language to program process data acquisition and monitoring systems. It is a virtual instrumentation package that uses graphical programming, allowing the construction of an instrumentation system with standard computers and cost-effective hardware.

LabVIEW programs consist of three parts: the front panel; the related program, which consists of a block diagram; and the icon/connector that is responsible for data flow between subroutines. Together, these three elements form the virtual instrument (VI), the basic element of a LabVIEW program.

The front panel acts as the graphical user interface (GUI) of the VI, and from it, the user can interact with the application using controls to change values on switches, knobs, sliders, among others. Data can also be displayed on the front panel in indicators, as graphs, strip charts, etc. The block diagram contains the source code for the VI. It is possible to select functions such as file I/O, instrument I/O, or data acquisition, and place them on the block diagram of the VI. Functions are coupled with “virtual” wires, in a flowchart-like environment, to define the execution of a VI. Every VI that is created can be re-used as a “subVI” in a higher level program. Therefore, it allows building an application with modular and reusable subVIs.

The LabVIEW environment can very easily take advantage of the serial port communication present in many instruments nowadays. Also, many data acquisition cards can be incorporated in LabVIEW programs. Additionally, the Virtual Instrument Software Architecture (VISA), provided with this software, offers a programming interface between the hardware and development environments. The several LabVIEW VISA functions constitute a standard for configuring, programming, and troubleshooting instrumentation systems comprising several communication protocols: GPIB, VXI, PXI, serial (RS232/485), Ethernet, and/or USB interfaces.

The graphical programming environment LabVIEW has been widely used in fermentation process data acquisition in the last years. For example, in Johnston et al. (2002), a starvation

dissolved-oxygen transient control was applied to fed-batch *E. coli* fermentation by using a data acquisition and control system developed in LabVIEW environment. In Turner et al. (1994a) and Turner et al. (1994b), the authors developed a program to control the feeding rate of a fed-batch *E. coli* fermentation based on on-line HPLC results that were integrated in the control algorithm developed in the LabVIEW environment. In Kellerhals et al. (1999), an automated on-line closed loop substrate control system was developed in LabVIEW and was applied to high-cell density fermentation systems of the microorganism *Pseudomonas putida*. Data from a Gas Chromatograph (GC) were acquired in a dedicated computer and sent via a serial port into a control computer, where the LabVIEW program integrated them with other relevant fermenter variables and calculated the feed rate. The program also performed the tasks of data storage and visualisation, among others.

In Claes and Van Impe (1998), the authors used LabVIEW to perform the tasks of logging and plotting the signal of an on-line biomass sensor, estimating the specific growth rate from an observer based estimator and performing dynamic control based on an optimal adaptive control algorithm of the influent flow rate in baker's yeast fermentation. However, this LabVIEW program runs in parallel with the New Brunswick Advanced Fermentation Software to acquire the main fermenter physical variables.

In the present work, and for the high-cell density fermentation of *E. coli*, the integration of the several on-line measurements in a supervisory program was thought essential for performing an accurate characterization of the bioprocess as well as for implementing advanced control actions. Thus, computer communication with the main devices used for on-line monitoring was implemented using the LabVIEW development system and the corresponding measurements were integrated in a supervisory program. This system is described in the following sections.

4.2 ON-LINE MONITORING AND CONTROL SYSTEM

All the data acquisition and control programs described in this section were constructed in the environment LabVIEW Professional Development System version 6.1. In runs in a Personal Computer (PC) with an AMD Duron processor running at 1.29 GHz and with 384 MB RAM in Windows XP™ environment. The data acquisition interfaces include standard (9-pins) and extra (25-pins) RS-232 serial ports derived from a C104P ISA multi-port serial board (Moxa

Technologies, Taiwan). A PCI data acquisition card is also included for multiport analog and digital input/output (I/O) (a PCI 6024E board from National Instruments, USA).

All functions used for the construction of the VIs for the data acquisition system belong to the VISA functions class.

4.2.1 Bioreactor and Digital Control Unit

All batch and fed-batch fermentations cultivations were conducted in a 5-L stirred tank bioreactor (Biostat MD, B.Braun, Germany). The reactor is equipped with a jacked for temperature control, an agitation rotor and mass flow controller for aeration. It also includes sensors for dissolved oxygen concentration, pH, temperature, redox potential, volume level and foam detection. The fermenter is connected to a DCU where some of those variables from the fermentation can be monitored and controlled. Each variable has its own control cycle, with proper parameters that can be altered directly by the user or through a remote communication. The peristaltic feeding pump (model 101 U/R from Watson Marlow, UK) is also connected to the DCU by a 25-pin analog I/O port and its rotation speed is one of the controlled variables (details about this connection can be found in Annex 2). A description of the components of the bioreactor can be found in Table 4.2.

The DCU also performs other functions such as amplification of sensor signals, and calibration of the sensors. The actuation on the final control elements is as well performed by this unit, namely on the pumps used for the addition of acid or base for pH control, on the heating of the jacket for temperature control, and on the agitation speed and/or the aeration rate for dissolved oxygen concentration control.

Furthermore, the BIOSTAT fermentation system incorporates another unit, responsible for the agitation power, water heating and pumping, and peristaltic pumps for base, acid and antifoam addition.

Dehumidified air supply to the bioreactor is obtained from the laboratory compressed air system and is controlled by a mass flow meter and controller (model Hi-Tec F201C-FB from Bronkhorst, The Netherlands). The inlet air line comprises two filtration devices: a 1-μm filter from Swagelok (USA) is used before the mass flow controller to remove dust particles and a 0.2-μm autoclavable hydrophobic membrane filter (ACRO 50 from Pall, USA) is placed before the bioreactor for air sterilization.

The bioreactor off-gas is dehumidified by the bioreactor's condenser and by two Dimroth type condensers. The exhaust gas line also includes silica and porous beads for final

dehumidication. The off-gas tubing is in polyethylene and stainless steel and has a total length of approximately 8 m.

Table 4.2 Description of the principal components of the bioreactor

Component	Description
Fermentation vessel (BIOSTAT MD, B. Braun, Germany)	Culture vessel of boro-silicate glass with an external jacket for temperature control and several ports for the probes, harvesting and sparging.
Stirrer	Driven by a 55 W DC motor directly mounted on the top plate of the vessel; 6-blade disk impellers.
Temperature probe	Platinum thermometer (Pt-100).
pH probe (Mettler Toledo, Switzerland)	Glass sterizable electrode; Potenciometric measurement.
Dissolved oxygen probe (Mettler Toledo, Switzerland)	Polarographic measurement.
Inlet gas system	Compressed air; 1-μm filter (Swagelok, USA); Mass flow controller Hi-Tec F201C-FB (Bronkhorst, The Netherlands); Membrane filter ACRO 50 (Pall, USA).
Outlet gas system	Reactor condenser; 2 Dymroth-type condensers; 1 silica and 1 porous bead.
Peristaltic pump (Watson Marlow, UK)	Substrate feeding.
pH control	2 peristaltic pumps for acid and base addition.

The communication of the DCU with the PC is accomplished by several dedicated LabVIEW VIs that constitute the instrument driver. The cable used for the communication is connected to the 9-pins “PERIPHERAL” port of the DCU and to one of the 25-pins serial port of the PC and the communication is done via the RS-232 protocol. The pin connection can be found in Annex 2. In order to enable the communication, several configurations have to be conducted on the DCU side. The communication parameters have to be defined according to the LabVIEW driver in both “HOST” and “PERIPHERAL” functions. The selected parameters are also presented in Annex 2.

The communication with the DCU is done through the communication protocol ASCII and the telegram format is shown in Figure 4.1.

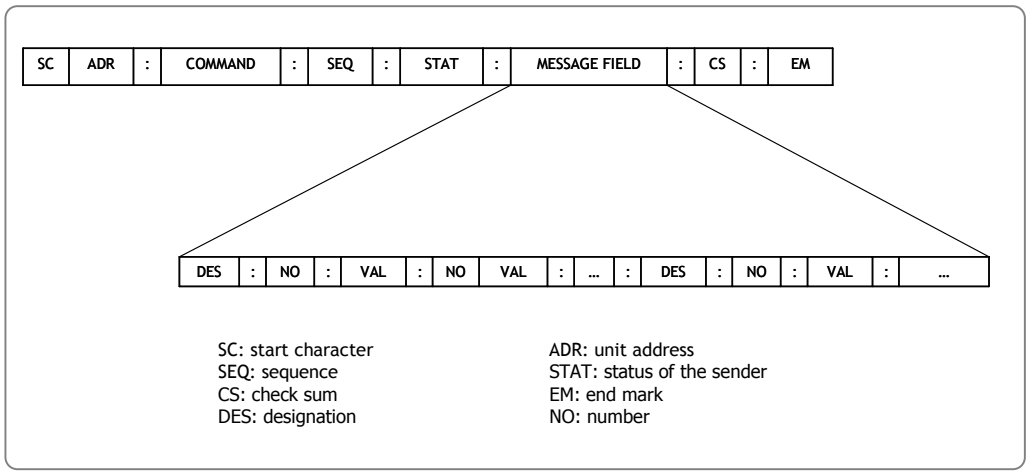


Figure 4.1 Telegram format for the communication with the fermenter DCU.

That telegram format is used for the communications in both directions. In the PC telegram, the start character is "*", while in the DCU telegram the start character is "#". The most important commands used in the communication from the PC to the DCU are the command "SU" for initialization, the command "DR" for data request and the command "DS" for data transmission. The initialization command initializes the communication by specifying in the "message field" the groups of data to be transmitted to the PC during the communication. The data request command asks the DCU for the process data that were included in the initialization. The data transmission command allows the fermenter to be fully controlled by the PC by specifying in the "message field" the control modes, set-points and parameters to the several control variables like pH, dissolved oxygen or stirrer speed. When a telegram with one of these commands is sent to the DCU, this equipment sends a response command within 500 ms that can be interpreted as an error telegram or a telegram containing data from the process.

The only extra requirement for the communication in the DCU side besides the communication parameters definition is turning on the "REM" button when the "DS" command is sent in order to allow the remote control of the control unit.

The "sequence" and "stat" fields of the telegram are predefined, while the "check sum" field is formed by the addition of all of the ASCII telegram values since the start character until the ":" before the check sum. The termination character is always "@".

Inside the "message field", the "designation" sub-field contains information about the kind of information to be transmitted. For example, it can be "PA" for process alarms, "CS" for controller set-points or "CM" for controller mode. The number corresponds to the sequence

number of the variables while the value corresponds to the value of the information corresponding to the variable.

As an example, the telegram "*1:DS:0:0:CS:3:625.0:CM:3:1:XXX@" is sent to the DCU to change the set-point of variable 3 to 625.0 and its control mode to 1 (automatic control mode). The characters "XXX" correspond to the check sum field. The answer in this case contains no relevant information.

In another example, the command "*1:DR:0:0:XXX@" can be sent to the DCU in order to request for data. The answer from the DCU contains information about all the variables that were initialized with the "SU" command.

For the construction of the "DS" command telegram or for the interpretation of the "DR" answering command, it is necessary to know the order of the variables attributed by the DCU, as well as the necessary conversion factors, since the DCU range for every variable is from 1 to 1000. This conversion was calculated and the corresponding information can be found in Table 4.3, as well as the order of the process variables used for the interpretation of the "DR" command answer.

Table 4.3 DCU Process variables (PVAL) table sequence and conversion

Variable Number	Variable	Real Range	Conversion
1	TEMP (reactor temperature)	0-150ºC	T (ºC) = 0.15 × Signal
2	JTEMP (jacked temperature)	0-150ºC	T (ºC) = 0.15 × Signal
3	STIRR (stirrer speed)	0-2000 rpm	STIRR (rpm) = 2 × Signal
4	MOTCR (motor)	0-100%	MOTCR (%) = 0.1 × Signal
5	pH	2-12 pH	pH = 2 + 0.01 × Signal
6	pO2 (dissolved oxygen tension)	0-100%	pO2 (%) = 0.1 × Signal
7	ACID (added acid)	0-1000 mL	ACID (mL) = Signal
8	BASE (added base)	0-1000 mL	BASE (mL) = Signal
9	AFOAM (added antifoam)	0-1000 mL	AFOAM (mL) = Signal
10	SUBS (feeding pump speed)	0-100%	SUBS (%) =0.1 × Signal
11	AIRFL (airflow rate)	0-10 slpm	AIRFL (slpm) = 0.01 × Signal
12	REDOX (redox potential)	-1000 mV-1000 mV	REDOX (mV) = -1000 + 2 × Signal
13	HARV (harvested volume)	0-1000 mL	HARV (mL) = Signal
14	FEEDP (added volume)	0-1000 mL	FEEDP (mL) = Signal

The controlled variables are much less than the process variables, and thus the "DS" command uses another sequence that can be found in Table 4.4, together with the type of control action that can be used for each case.

Table 4.4 Control variables table used by the bioreactor DCU and the corresponding control actions that can be taken in each case

Variable Number	Variable	Control action
1	TEMP	PID cascade with outputs for heating/cooling
2	STIRR	Set-point controller for external motor control
3	pH	PID with outputs for acid/base
4	pO2	PID cascade for agitation speed or airflow controller
5	AIRFL	Set-point controller for external mass flow controller
6	AFOAM	On/off controller for antifoam
7	SUBS	Set-point controller for external feeding pump
8	JTEMP	PID
9	LEVEL	PID

In order to accomplish an effective communication with the DCU, three LabVIEW programs were constructed to send and receive information related with the commands "SU" (Initialize VISA.vi), "DR" (DR VISA.vi) and "DS" (DS VISA.vi). The three programs were constructed to work as sub-VIs that can be integrated hierarchically to a data acquisition and process control program.

Basically, the VIs begin with the configuration of the communication, according to the parameters defined in the DCU. The next step is related with the break signal that has to be sent to the DCU before the communication. Then, the VISA Write function is used to send the appropriate string containing the command to the DCU. In the construction of the string, an additional VI was constructed to calculate the check sum value in each case. Finally, the answer from the DCU is received and the extraction of the valuable information is accomplished in the case of the DR VISA.vi. To acquire the strings as soon as they are available, the serial port is continuously monitored and the incoming characters counted. The character count is compared with the expected minimum value, which depends on the particular VI; if the character number is larger or equal to the preset value, meaning that the required amount of information is present, the string is processed further. If not, the function waits 5 seconds and then returns a timeout error. This and other communication errors are

treated by the error handling functions of VISA. Additionally, if a string containing an error is received (the second character of the answering command is identified as "E"), the error source is identified as "Instrument Error". As an example of the control panel and diagram of those VIs, the DS VISA.vi is presented in Figure 4.2 and Figure 4.3. The other programs can be found in the CD-ROM embedded to this thesis.

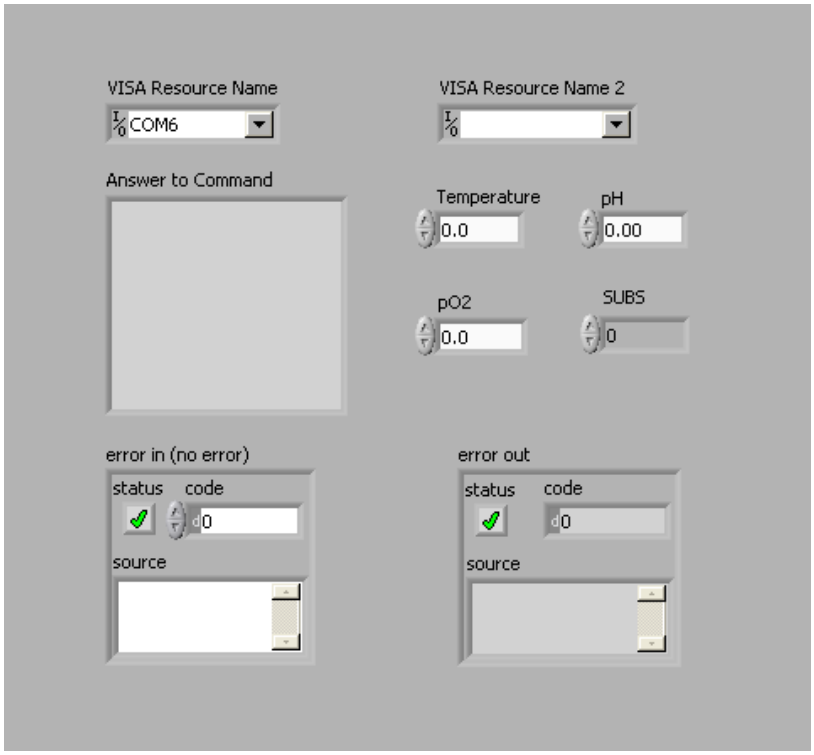


Figure 4.2 Control panel of the DS VISA.vi.

4.2.2 Balances

Two balances are used, one for measurement of the total weight of the bioreactor and the other for weighting the feeding solution reservoir.

The bioreactor's balance is a SB 32001 model from Mettler Toledo GmbH (Switzerland) with a 32-kg capacity and with a readability of 0.1 g. It has a platform with sufficient dimensions and stability to support the 5-L bioreactor and the corresponding attached devices. With this balance, it is possible to know in each moment the exact weight of the culture medium and biomass inside the bioreactor, allowing the use of more precise mass-based concentrations.

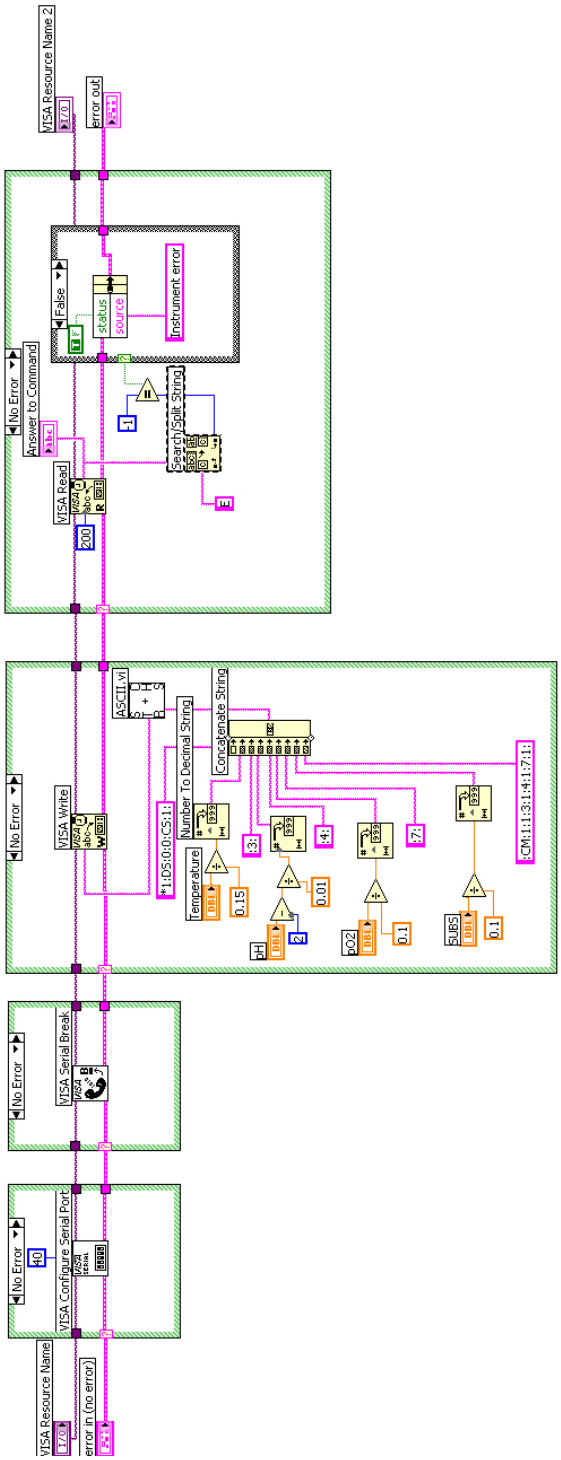


Figure 4.3 Diagram of the program `DS VISA.vi` developed in this work to send to the DCU information related with controllers' set-points, modes and parameters.

Communication with the bioreactor balance is also done via RS232 protocol and the communication cable is connected to one of the 25-pins serial port of the computer. Pin connection is shown in Annex 2.

This device continuously sends data as strings in the format: "`Ss  X.X g`". A VI called `Weight.vi` was constructed for the interpretation of these strings and the consequent acquisition of the bioreactor's weight. Initially, the communication was configured according to the balance parameters: 9600 bpm baud rate, words with 8 bits, 1 stop bit and no parity. In this case, a flow control was implemented based on the RTS/CTS mechanism. The termination character was the line feed. After the initialization, the `VISA Read function` was used for the acquisition of the balance string.

However, although, according to the balance's manufacturer, all the strings should have the same length (18 bits) and begin with the characters "`Ss`", it was verified that in many cases the strings were acquired in an incomplete mode. Thus, a function that is able to search for the numerical part of the string had to be included. After that function, another function that transformed the ASCII characters in decimal values was used to obtain the bioreactor weight in grams. It should be noticed that, due to the already mentioned problem of truncated strings, sometimes one or more digit of the weight value are not received, leading to erroneous weight values.

The diagram of the fermenter's weight acquisition VI is shown in Figure 4.4.

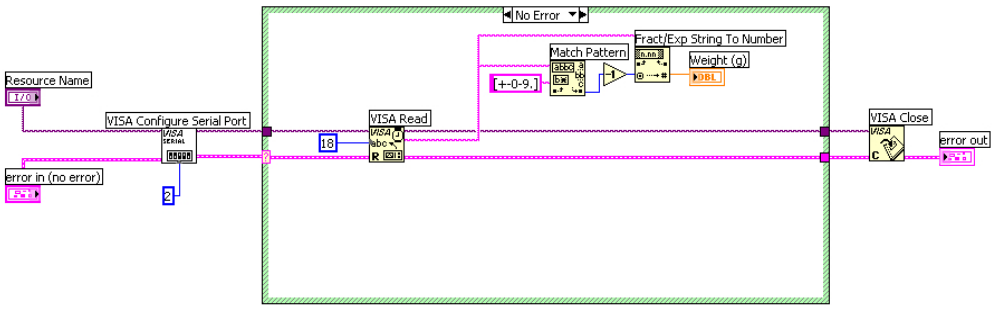


Figure 4.4 Diagram of the program `Weight.vi` that acquires data from the fermenter's balance.

The feeding balance is a PM 4800 Deltarange model, also from Mettler Toledo GmbH, with a precision of 0.001 (0-1 kg) or 0.1 (1-4 kg). It is used so that the exact feeding profile during fed-batch cultivations could be calculated.

The communication cable for the feeding balance is described in Annex 2. The program elaborated for the communication with this device is shown in Figure 4.5. In this case, the communication parameters are: 9600 bpm baud rate; words with 7 bits, 1 stop bit and even parity. The termination character is the line feed. After the initialization, the balance requires an input string constituted by a certain command followed by the characters carriage return and line feed. The commands can be "`S`" for requiring the weight value, "`U`" for changing the weighting units and "`T`" for taring the balance. In the VI developed for this communication, only the command "`S`" is sent to the balance, because is the only one required for the data acquisition system. The remaining part of the program is very similar to the previous case.

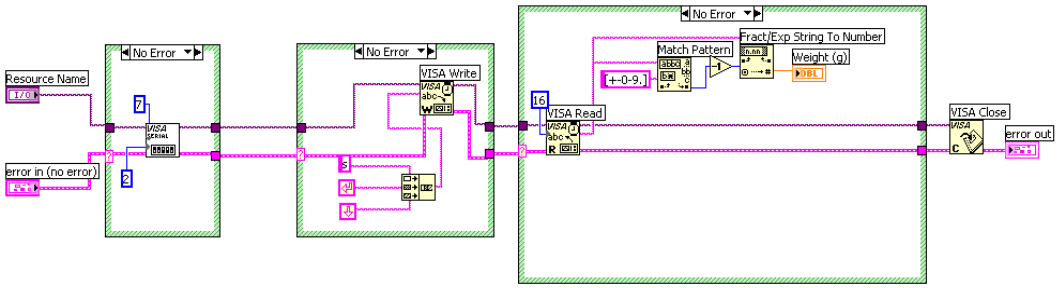


Figure 4.5 Diagram of the program `Weight_feed.vi` elaborated for the communication with the feeding balance.

4.2.3 Mass Spectrometer

After the off gas stream was filtered and dried, oxygen, nitrogen, and carbon dioxide in the exhaust gas stream were measured with a Mass Spectrometer (MS) (Bioquad from Ledamass, UK). This equipment is a quadrupole type, being able to analyse masses until 200 Da. A vacuum system, an analyser and an inlet capillary system constitute the equipment. A radio-frequency emissary, an amplifier, and a control unit (satellite) compose the analyser. In order to understand the basis of the MS measurements, a brief theoretical explanation of the equipment used in this work is presented.

4.2.3.1 Theory

Besides other applications in biotechnology, like the screening of microbial strains by Pyrolysis MS, protein structure/sequence evaluation and characterization of post-translational modification of proteins by Electrospray MS (ESMS) and Matrix Assisted Laser Desorption Ionisation MS (MALDI MS) (Schügerl, 2001) or on-line monitoring of the dissolved gases and volatile components (Heinzle et al., 1985; Oeggerli and Heinzle, 1992 and Oeggerli and Heinzle, 1994), Mass Spectrometry has been broadly applied for the analysis of oxygen, carbon dioxide and nitrogen in off-gas (Lee et al., 1997).

Mass Spectrometry was established by J.J. Thomson and F.W. Aston from 1897 onwards at the Cavendish Laboratories, Cambridge University and is a technique that can separate charged atoms or molecules according to their mass-to-charge ratio (m/z).

Basically, a mass spectrometric analysis can be envisioned to be made up of the following steps: Sample Introduction; Ionization; Mass Analysis and Ion Detection/Data Analysis. The description made in the next sections will only cover the equipment used in this work. For a more detailed information about MS, there are several tutorials, namely in the internet, like in the British Mass Spectrometry Society site[1]; the Cambridge University Mass spectrometry server[2]; the University of Arizona - Department of Chemistry site[3]; and the Vanderbilt University Mass Spectrometry Research Center site[4], among others. Also, the principles, sampling systems, control of the measuring device and application of MS for bioprocesses have been summarized in Heinzle (1987) and Heinzle (1992).

A general vision of the main constituents of a Mass Spectrometer like the one used in this work can be seen in Figure 4.6.

- Sample Introduction

The sample is introduced into the mass spectrometer, kept under high vacuum ($<10^{-5}$ mBar) via a capillary (stainless steel, heated, with 0.3 mm of internal diameter and 1 m large). The reduced pressure is kept in order to prevent collisions of ions with residual gas molecules in the analyzer during the flight from the ion source to the detector. It is achieved by two components: two mechanical pumps and a turbomolecular pump (from Elnor, Belgium).

---

[1] http://www.bmss.org.uk/
[2] http://www-methods.ch.cam.ac.uk/meth/ms/theory/index.html
[3] http://www.chem.arizona.edu/massspec/
[4] http://ms.mc.vanderbilt.edu/tutorials/ms/ms.htm

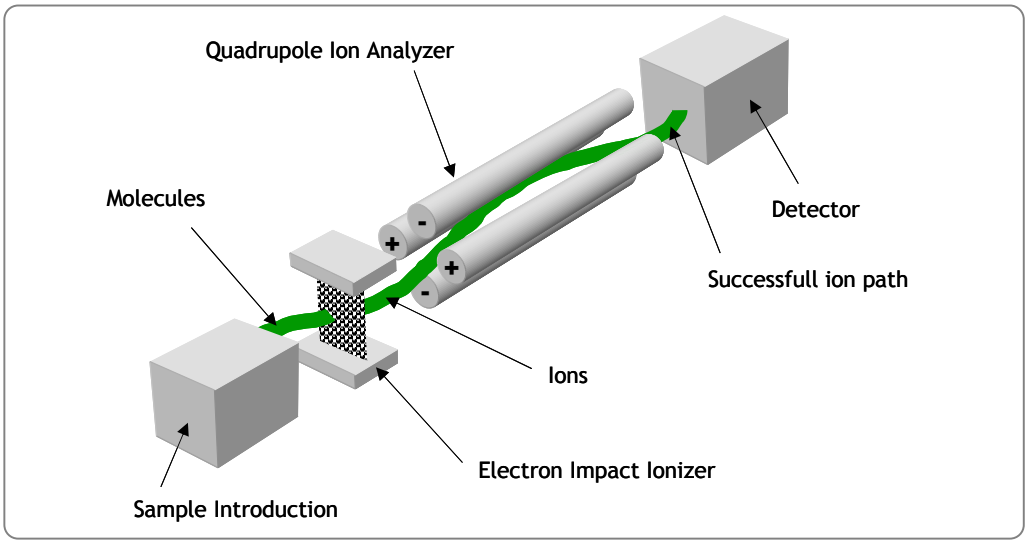


Figure 4.6 General scheme of a mass spectrometer. The line illustrates ions of a particular mass/charge ratio which reach the detector at a certain voltage combination (adapted from http://www.chem.arizona.edu/massspec/).

- Ionization

Many ionization techniques can be used to produce charged molecules in the gas phase. In the equipment used, the ionization technique is the Electron Ionisation (EI).

With this ionisation technique, a beam of energetic electrons is created inside a heated chamber by passing current through a tungsten filament. The Bioquad equipment has two twin filaments in order to decrease the frequency of the filament exchanging. These electrons leave the filament surface and are accelerated towards the ion source chamber that is held at a positive potential (equal to the accelerating voltage). The electrons acquire energy equal to the voltage between the filament and the source chamber – 70 electron volts (70 eV). The gaseous analyte molecules are introduced into the path of the electron beam. The resulting direct interaction causes the loss of one electron from the molecule, thus generating a positively charged molecular ion. Depending on the compound and the ionisation energy, the molecule ion may then fragment. Figure 4.7 shows a schematic diagram of an EI source.

The positive ion repeller voltage and the negative excitation voltage work together to produce an electric field in the source chamber such that ions will leave the source through the ion exit slit. Although both positive and negative ions may be generated at the same time, one polarity is chosen and either positive or negative ions are analysed and recorded. Molecules that do not ionise, i.e. remain neutral, are pumped away and will not be detected.

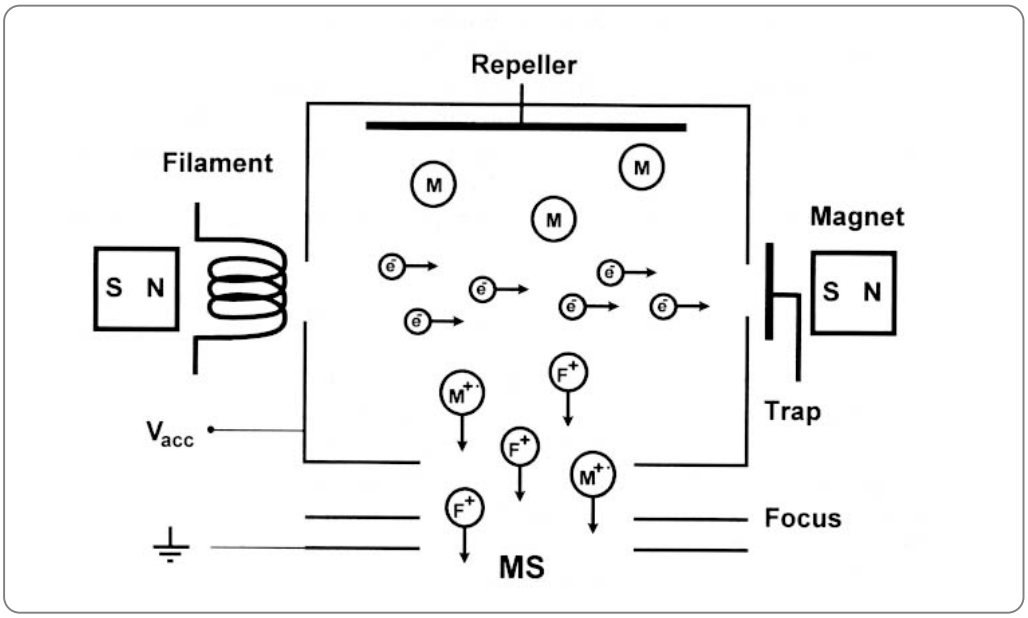


Figure 4.7 Schematic representation of an electron ionization ion source. M represents neutral molecules; e^-, electrons; M^+ , the molecular ion; F^+, fragment ions; Vacc, accelerating voltage (adapted from http://ms.mc.vanderbilt.edu/tutorials/ms/ms.htm).

- Mass Analysis

The analyser used in this work was a quadrupole-type mass analyser. In this analyser, only electric fields are used to separate ions according to their m/z values. A quadrupole consists of four parallel rods or poles through which the ions being separated are passed. The poles have a fixed DC and alternating RF (radio frequency) voltages applied to them. Depending on the produced electric field, only ions of a particular m/z will be focused on the detector, and all the other ions will be deflected into the rods. This phenomenon is illustrated in Figure 4.6. By varying the strengths and frequencies of electric fields, different ions will be detected thus making the mass spectrum.

The length and diameter of the rods will determine the mass range and ultimate resolution that can be achieved by the quadrupole assembly. However, the maximum mass range that is usually achieved is around 4000 Da with a resolution of around 2000.

- Detection

The detector of the Bioquad equipment is a Faraday cup-type that produces an electronic signal when it is struck by an ion. The ions are directed to a cup that measures directly the ion current.

Timing mechanisms which integrate those signals with the scanning voltages allow the instrument to report which m/z strikes the detector. Thus, the mass analyzer sorts the ions according to m/z and the detector records the abundance of each m/z.

4.2.3.2 Improved Instrument Set-up

Due to the inherent characteristics of the MS equipment (Ferreira et al., 1998) used throughout this work, the signal obtained from gas analysis was very noisy and had a significant drift. Thus, it was necessary to implement some changes in the set-up in order to make the signal acceptable for fermentation monitoring and control. In order to decrease the levels of noise and to increase the accuracy, the vacuum pumps were placed on a platform for stabilization, and an Uninterruptible Power Supply (UPS) (APC – American Power Conversion Corp., USA) was installed to avoid current peaks and oscillations.

Another change is related with the inlet system. The MS is used for the off-gas analysis and the subsequent calculation of the Oxygen Transfer Rate (OTR) and Carbon Dioxide Transfer Rate (CTR). However, as it will be described in the next chapter, on-line calculation of these variables requires the knowledge of both off- and in-gas composition. Due to the significant daily variations in the composition of atmospheric air, especially in what carbon dioxide respects, and also to the considerable oscillations in the MS performance, it was found undesirable to use average values for the composition of the inflow air. Thus, during the fermentations, the MS analysed sequentially both streams.

The MS equipment was initially designed to acquire gas for analysis from different inlets, being capable of performing the analysis of several points sequentially. It had a valve system that allowed the automatic selection of the appropriate point of analysis. However, the disadvantage of this set-up is that, depending on the pressure of the sampling points, the mass flow rate entering the MS varies significantly, leading to erroneous measurements. Thus, a new set-up for the stabilization of the inlet flow rate was designed, which is represented in Figure 4.8. It was decided to introduce a valve manifold (from Advantech Co., Ltd., Taiwan) before a mass flow controller (model HFC 202 from Teledyne Hastings Inc.,

USA) and thus, with only one of these devices, it was possible to control the mass flow rate entering the MS.

An ADAM relay output model from Advantech Co., Ltd. (Taiwan), which communicates through the RS485 protocol, controls the valve manifold. Thus, for controlling which valve is opened and closed, a RS485/RS232 converter (model ADAM 4520, also from Advantech) is connected with the valve controller and the communication is done via a serial port through the RS232 protocol. The communication pin connections are shown in Annex 2.

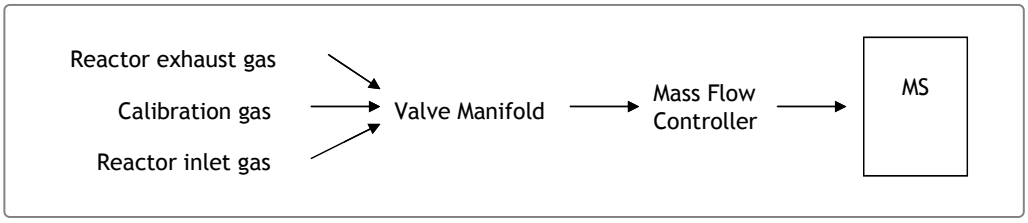


Figure 4.8 Improved set-up for gas analysis used in this work.

The communication with the valve manifold is also done by ASCII string decodification. The telegrams used for the communication with this device follow the format described in Figure 4.9. There are several commands that can be used, although for this application, only two types of telegrams are applied. The first is the "`Digital Data In`" telegram, expressed as "`$AA6(CR)`", where `AA` is the hexadecimal address, set to 01 and `CR` is the carriage return character. This telegram asks the module for the state of the valves and the answering telegram, after decodification, can be used for recording and visualizing the point of the sampling of the MS. The other telegram is the "`Digital Data Out`" telegram, expressed as "`#AABB(data)(CR)`", where `BB` is a variable set to `00` and data includes the value of the digital output in an hexadecimal base. This value determines the state of the valves to opened (1) or closed (0).

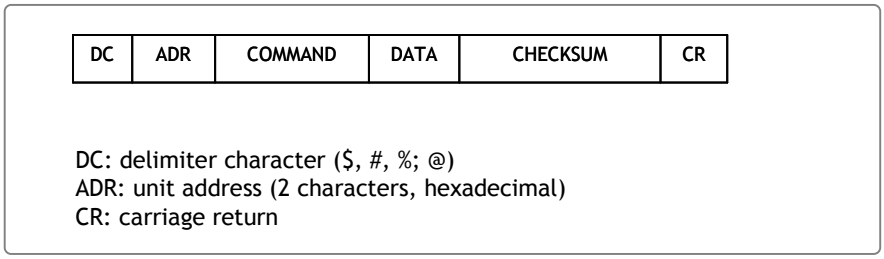


Figure 4.9 Format of the telegram used for the communication between the PC and the valve manifold.

The corresponding telegrams, as well as the constructed LabVIEW VIs can be found in Table 4.5. The structure of the VIs is very similar to the one of the DCU VIs. An example can be found in Figure 4.10, while the other VIs can be found in the CD-ROM embedded to this thesis.

Table 4.5 LabVIEW VIs used for the communication with the valve controller

VIs	Function	Telegram
ADAM 4060 Initialize.vi	Configures the communication	–
ADAM 4060 Open Channel 1.vi	Opens valve 1	#010001(CR)
ADAM 4060 Open Channel 2.vi	Opens valve 2	#010002(CR)
ADAM 4060 Open Channel 3.vi	Opens valve 3	#010004(CR)
ADAM4060 Close all Channels.vi	Closes all valves	#010000(CR)
ADAM4060 Read Channels.vi	Asks the module for the state of the channels. The answering telegram describes which valves are opened or closed	$016(CR)
ADAM4060 Close.vi	Ends the communication	–

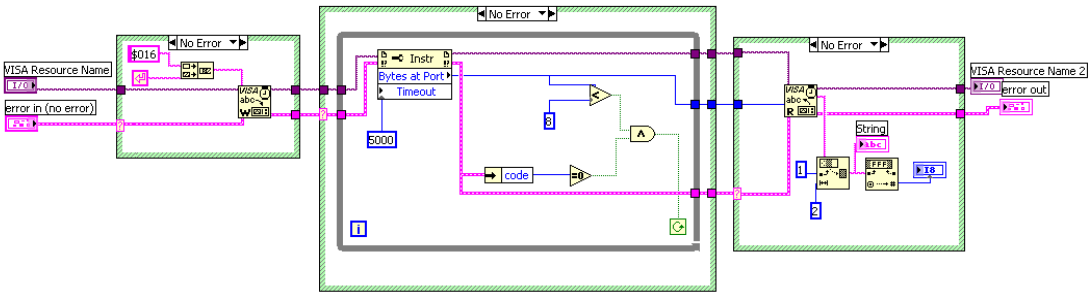


Figure 4.10 Diagram of the program ADAM4060 Read Channels.vi.

Communication with the mass flow controller is done via the NI card PCI 6024E via analog I/O. Connections with this device are shown in Annex 2. The analog I/O signals correspond to the controller measurements and set-point, respectively. Both signals are in the range 0-5 V. The calibration of the mass flow controller indicated that 5 V corresponds to 1.16 litres per minute of air at standard conditions (Annex 1). The set-point for the device was 250 mL·min^{-1}.

Regarding the communication, in this case, and for the sake of simplicity, two channels were defined in the Measurement & Automation Explorer version 1.2 software (from National Instruments, USA). This solution allowed one single definition of the I/O ranges, as well as the physical amounts (litres per minute) correspondence for any communication requiring this device. The LabVIEW VI elaborated for this communication is also in the CR-ROM attached to this thesis and uses the defined channels for data acquisition and set-point definition.

4.2.3.3 Quantitative Analysis and Calibration

The Mass Spectrometer generates spectra that are indicative of the components present in the streams under analysis, as well as of their quantity.

Each component, in the pure state, generates a single spectrum characteristic of that specie, constituted by several peaks that result from the fragmentation pattern. The characteristic spectrum for oxygen, carbon dioxide and nitrogen are shown in Figure 4.11. The x-axis of these bar graphs is the increasing m/z ratio. The y-axis is the relative abundance of each ion, which is related to the number of times an ion of that m/z ratio strikes the detector. Assignment of relative abundance begins by assigning the most abundant ion a relative abundance of 100%. All other ions are shown as a percentage of that most abundant ion. The basis for the quantification of a given compound is usually the most abundant ion (called major peak). Therefore, as in the case of oxygen, carbon dioxide and nitrogen, the most abundant ions are the single positive ones, and then the m/z ratio of those peaks corresponds to their molecular mass (32, 44 and 28, respectively).

In a mass spectrum of a gas mixture, the main assumption is that each peak is composed by the single contributions of the molecular ions and/or ionic fragments to that m/z value. The contributions are then added linearly and the total peak height for a given m/z in a mixture is equal to the sum of the contributions of the individual peaks that would be produced if the compounds were pure. This can be expressed by the following equation:

$$S_i = \sum_j S_{ij} \qquad \text{Equation 4.1}$$

Where S_i is the total intensity of a certain m/z peak and S_{ij} is the contribution of the component *j* to that peak.

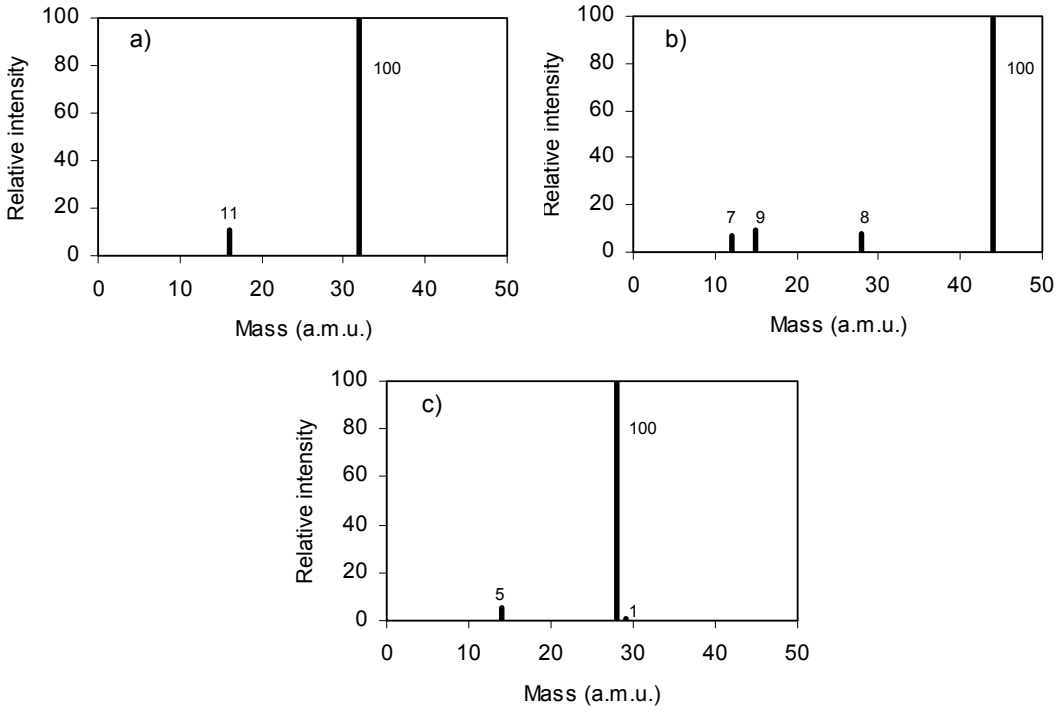


Figure 4.11 Mass spectra for the pure components: a) oxygen, b) carbon dioxide and c) nitrogen.

Thus, it is necessary to correct the interferences of the several components of a mixture in order to obtain the correct measurements. This can be done using the decoupling matrix proposed by Bartman, (1987). The decoupled signals can, accordingly, be represented as follows:

$$P = R^{-1}S \qquad \text{Equation 4.2}$$

Where **P** represents the decoupling signals vector, **S** is the original signals vector and $\mathbf{R^{-1}}$ is the decoupling matrix, an inverse of the so-called fragmentation matrix. For the present case, the fragmentation matrix is (Ferreira, 1995):

$$R = \begin{bmatrix} 1 & 0 & 0.08 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Equation 4.3}$$

Where the lines represent the basic peaks and the columns represent the fragmentation in those peaks.

And the decoupling matrix comes:

$$R^{-1} = \begin{bmatrix} 1 & 0 & -0.08 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Equation 4.4}$$

Thus, the decoupled signals for the 3 species will be:

$$p_{O_2} = S_{32} \qquad \text{Equation 4.5}$$
$$p_{CO_2} = S_{44}$$
$$p_{N_2} = S_{28} - 0.08 S_{44} \approx S_{28}$$

However, since the signals of the m/z 44 are usually two or three orders of magnitude smaller than the ones of m/z 28, the decoupled signal for nitrogen can be considered equal to the source signal, simplifying the calculations.

Still, the decoupled pressure signals have to be converted to partial pressures or molar fractions by proper calibration factors. The subject of calibrating this kind of MS equipments has been discussed in Ferreira et al. (1998). Accordingly, among the several alternatives for performing a calibration of the MS, the most recommended is the utilization of several standard mixtures. However, the calculation of the sensitivity coefficients is simplified when compared to what is described in that paper due to the assumptions made above concerning the absence of cross influence from m/z fragment peaks in the major peaks.

The composition of the calibration mixtures is shown in Table 4.6 and was chosen so that the molar fractions cover all the compositions found in the in- and off-gas of the reactor. The mixtures were obtained from Praxair, Inc. (Spain).

Table 4.6 Composition of the mixtures used for calibrating the MS

Mixture	Oxygen (molar fraction)	Carbon dioxide (molar fraction)	Nitrogen (molar fraction)
1	20.03	10	69.97
2	21.49	1.51	77
3	18.04	3	78.96
4	16.06	5.04	78.9
5	9.97	10.02	80.01
6	22.02	0.99	76.99

For the calibration, a mass flow-controlled acquisition from 3 of the mixtures was conducted for at least 40 hours and the pressure values obtained for each m/z value (28, 32 and 44) were averaged. Then a linear regression was conducted with these values against the known composition of the mixtures and a calibration factor was found that correlated the pressure values obtained in the MS for a given m/z with the composition. With this calibration factor (*F*), the composition of a given stream can be obtained as follows:

$$y_{O_2} = p_{O_2} / F_{O_2} = S_{32} / F_{O_2}$$
$$y_{CO_2} = p_{CO_2} / F_{CO_2} = S_{44} / F_{CO_2}$$
$$y_{N_2} = p_{N_2} / F_{N_2} \approx S_{28} / F_{N_2}$$

Equation 4.6

An example of an MS calibration is shown in Annex 1.

4.2.3.4 MS Data Acquisition

Information from the MS is acquired via serial port through a RS232 connection to the satellite of the equipment by the software RGA for Windows version 2.43 (Spectra, UK), running in multitask mode in the supervisory computer. This software is able to select the equipment valves (in the present work the MS valve used was always valve number one, due to the alteration on the MS set-up described in the last section), begin electron ionization by turning on the emission of one of the twin filaments, selecting the ranges of m/z to be analysed, and the accuracy of the acquisition. Each selected m/z is evaluated in terms of its partial pressure in torr.

It also allows the selection of one from various modes of acquisition, depending on the application required: the "`Vacuum Scan`" and the "`Bar Chart`" modes can be used to scan a selectable mass range and are useful for evaluating all the species in the gas streams under analysis. In contrast, the "`Peak Jump`" and the "`Multi Trend`" are channel based scanning modes, allowing the scan to be made only in selectable masses, and each of the channels has a set of parameters which may be set independently. These modes are more adequate for monitoring fermenter's off-gas, by selecting only the masses that correspond to oxygen, carbon dioxide and nitrogen and thus increasing the speed of a single scan. Other features of the "`Multi Trend`" mode, like the possibility of accompanying the evolution of the individual masses made it the mode chosen for this work.

RGA software also allows the interchange of data with other programs running on the same computer by the standard Windows Dynamic Data Exchange (DDE) protocol.

Thus, a VI was constructed to communicate with RGA using that protocol, with the objective of integrating MS data with the remaining fermentation variables. With this type of protocol, the DDE server (RGA) puts available its information by defining the variables service, topics and items that the LabVIEW VI identifies. The definition of those variables for the RGA `Multi Trend` mode, as well as the LabVIEW functions used to capture the corresponding values are shown in Table 4.7.

Table 4.7 Definition of the DDE variables in the RGA software and corresponding LabVIEW functions. # represents the sequence number of the m/z in "`Multi Trend`" mode

DDE variable	RGA definition	LabVIEW Function
Service	RGA4WIN	`DDE Open Conversation.vi`
Topic	Multi Trend	
Item	PeakData(#)	`DDE Advise Start.vi`

In the communication VI (represented in Figure 4.12), besides the initialization functions mentioned in Table 4.7, the `DDE Advise Check` function used allows the acquisition of the values of the variables defined by "PeakData(#)". This function has the particularity of being able to wait for a change to happen in the variable before reading it, instead of a timed acquisition that could be desynchronized with MS measurements. Thus, this function allows the acquisition in the LabVIEW VI of the values of the defined variables as soon MS acquires them. Also in this case, the values are acquired in the string format, which has to be converted to a numeric format. The acquisition of the values from the MS is inserted in a while loop that is interrupted by pressing a stop button in the control panel. When the acquisition is interrupted, the LabVIEW functions `DDE Advise Stop.vi` and the `DDE Close Conversation.vi` close the communication. Error handling is once again accomplished by error flux.

4.2.4 FIA

Data from the FIA system (glucose and acetate concentrations) are acquired in one dedicated computer with the ASIA software (ISMATEC, Switzerland) that also controls the entire system (testing and recalibrating when necessary), as already described in the previous chapter. In order to obtain the results in the supervisory computer and to integrate them with the remaining state variables, two VIs were developed for the communication between both computers.

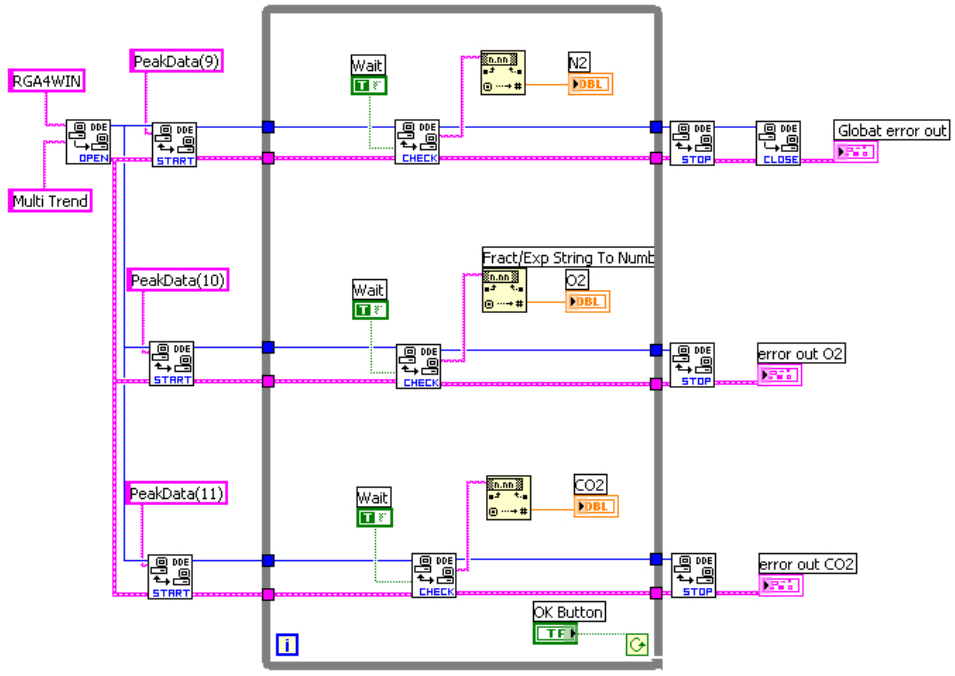


Figure 4.12 Diagram of the `MS.vi`, used to communicate with the Mass Spectrometer using the DDE protocol.

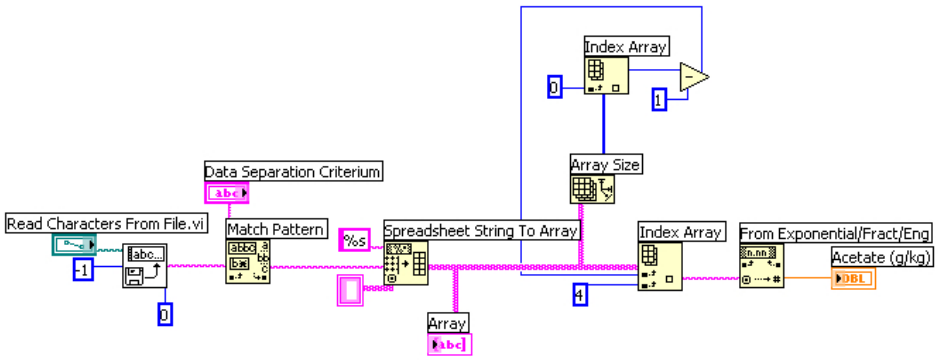


Figure 4.13 Diagram of the VI used for picking the acetate FIA measurements from the file created by the ASIA software.

The ASIA software saves the results of acetate and glucose measurements in a file with the extension ".con" that can be read with a text editor. Thus, a LabVIEW function called `Read Characters From File.vi` is able to open the corresponding file via Ethernet by specifying the path of that file (like "`\\Fia\C\ASIA21\CONC\fb1607.con`"). It is then interpreted as a string and other VIs are used to look for the value of the last acetate or glucose measurement in that string. The VI elaborated for acquiring acetate measurements is represented in Figure 4.13.

4.2.5 Integrated Data Acquisition System

The individual data acquisition and process control tasks just described were integrated hierarchically in a data acquisition and process control program called `Coli Monitor.vi`, and each sub-VI was responsible for the execution of a specific task. The program structure was hierarchical, meaning that VIs located at a higher level called sub-VIs to accomplish a specific subset of tasks. The program was modular, permitting sub-VIs to be called by more than one higher-level VI.

This supervisory VI also includes the algorithms for on-line calculation of the feeding profiles, which will depend on the type of experiment. Thus, variables like acetate measurements, bioreactor weight or MS measurements can be integrated in the supervisory VI as input information for control algorithms, being the controller output the set-point for the feeding pump. Also, this supervisory VI also allows the visualization and recording of the values of the process variables according to user-specified sampling rate and file format.

The experimental set-up showing all the equipments described and the correspondent connections is shown schematically in Figure 4.14, while a photograph with some equipment is shown in Figure 4.15.

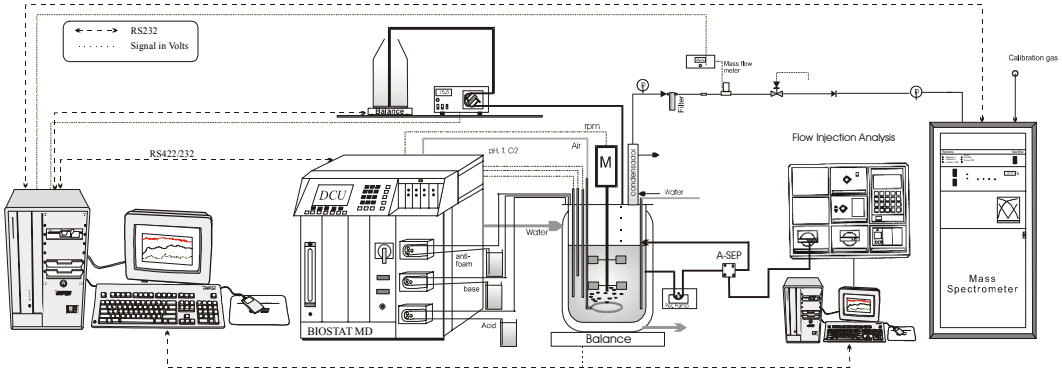


Figure 4.14 Experimental set-up.

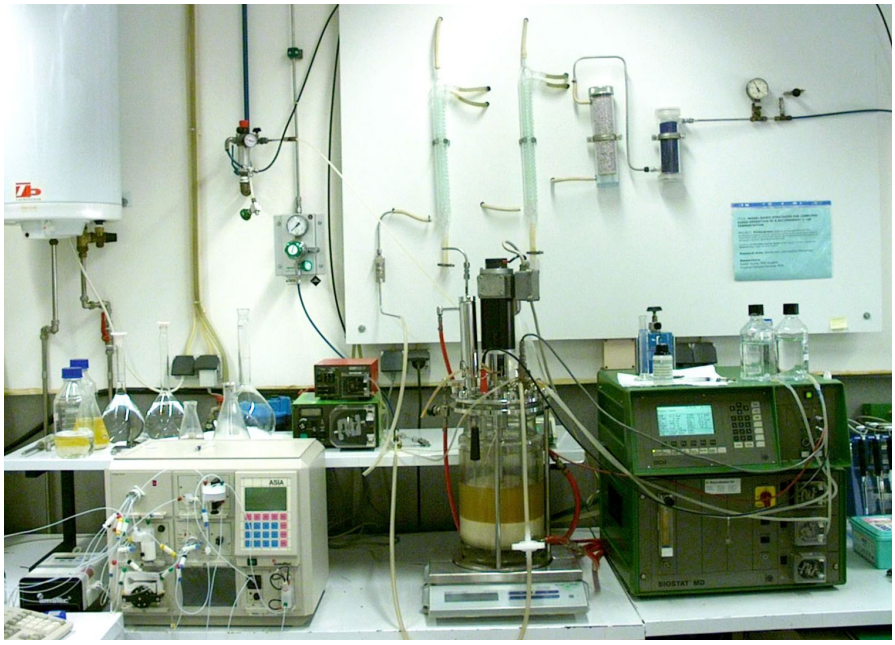


Figure 4.15 Photograph of the bioreactor, the Digital Control Unit and the FIA system.

The frontal panel and the diagram of the supervisory VI are shown in Figure 4.16 and Figure 4.17.

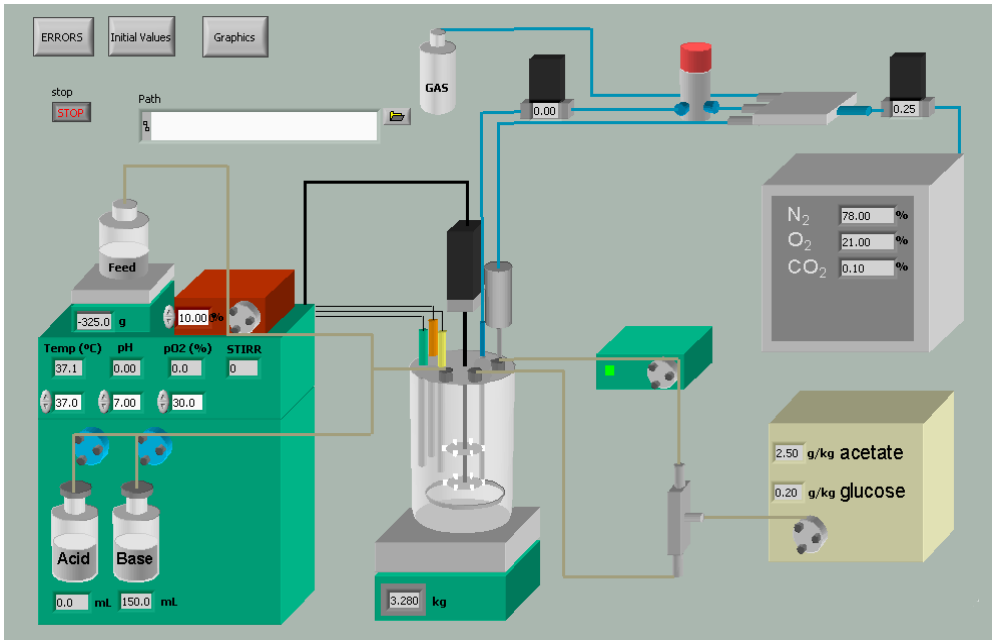


Figure 4.16 Frontal panel of the supervisory VI (`Coli Monitor.vi`).

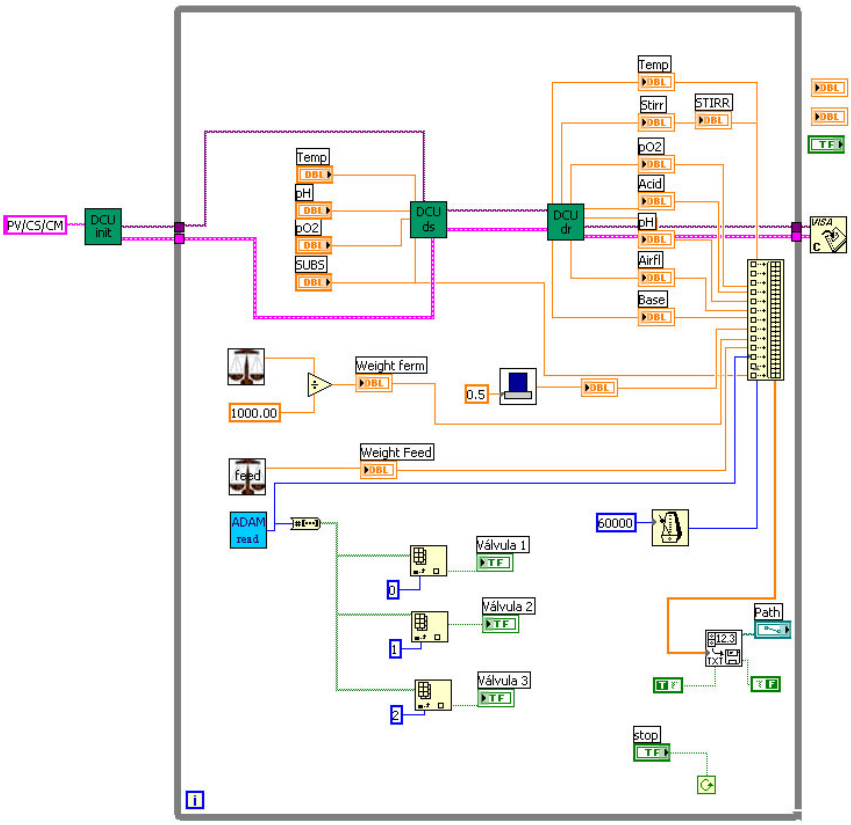


Figure 4.17 Diagram of the supervisory VI (`Coli Monitor.vi`).

4.3 OFF-LINE ANALYSIS

Besides the on-line measurements described in the previous sections, it was necessary to conduct several off-line measurements, especially to validate some of the on-line methods used. In fact, Biomass concentration was the only state variable that was measured exclusively off-line. In the following sections there is a brief description of the several off-line methods used throughout this thesis.

4.3.1 Biomass

Biomass determination was conducted by measuring the optical density (OD) of the samples at 600 nm in a UV/VIS spectrophotometer (model 7850 from Jasco, Inc., USA) and by correlating those measured values with dry cell weights. The blank used for these measurements was distilled water. Usually, the samples had to be diluted also with distilled water for fitting in the linear range of the method (0-0.6 absorbance units). Calibrations for finding the correct relation between OD and dry cell weight were performed for every fermentation and the corresponding method as well as an example can be found in Annex 1.

4.3.2 HPLC

HPLC measurements of acetate were conducted in a refractive index detector (Jasco, Canada) with a Chrompack organic acids column (Varian, USA) at 60ºC. The mobile phase consisted in a 0.01 N solution of H_2SO_4 at a flow rate of 0.75 $mL \cdot min^{-1}$. The typical retention time for acetate was 8.9 minutes.

Calibration was performed with standard solutions of sodium acetate before the analysis of the fermentation samples. Concentrations of the standards were in the range 0.1 to 5 $g \cdot kg^{-1}$. Standard preparation, as well as a calibration example can be found in Annex 1.

Due to the high level of interferences, especially from phosphates, analysis of glucose with HPLC was impossible.

4.3.3 Enzymatic Methods

Enzymatic measurements of acetate and glucose were performed with kits from R-Biopharm, Germany (ref. 148261 for acetate measurement and ref. 716251 for glucose measurement) according to the manufacturer's instructions. However, as the absorbance measurements were carried out in a Sunrise modular absorbance microplate reader (Tecan, Austria) it was possible to reduce to 10% the volumes of the reagents and samples indicated by the manufacturer.

4.3.4 Glucose Measurement by DNS

The di-nitrosalicylic acid (DNS) method was also used for glucose measurement. For this method, it is necessary to prepare a reagent according to Miller (1959), in the following way:

1. dissolve 5 g of dinitrosalysilic acid in 250 mL of distilled water at 80º C in a 500-mL volumetric flask;
2. wait until the solution reaches environmental temperature
3. add 100 mL of a 2 N NaOH solution
4. add 150 g of potassium sodium tartarate-4-hydrate
5. fill with distilled water until 500 mL

For each sample, 500 μL of the DNS reagent were added to 500 μL of sample and boiled at 100ºC during 5 minutes. Afterwards, 5 mL of deionised water were added and the absorbance of 300 μL of the resulting solution was measured at 540 nm in the Sunrise modular absorbance microplate reader. Calibration was performed with standards within the range 0.1-1 $g \cdot kg^{-1}$. An example of a DNS calibration can be found in Annex 1.

4.4 FERMENTATION OPERATION

The *E. coli* fermentations described throughout this thesis were operated in batch or fed-batch modes, with an initial weight of 3 kg. For the fed-batch mode, a maximum of 2 kg of feeding solution were added during the fermentations.

All the fed-batch fermentations began with a 5-7 hours long batch phase. The flag for the beginning of the feeding was the increase of the dissolved oxygen resulting from the exhaust of the glucose and acetate in the medium.

4.4.1 Microorganism

The microorganism was an *E. coli* JM109(DE3) strain harbouring the plasmid pET-IFNα, constructed by Castro (2003), with resistance to kanamicin.

4.4.2 Operating Set-points

The operating set-points common to all fermentations conducted are shown in Table 4.8. pH control was achieved by adding solutions of 12.5% ammonium hydroxide or phosphoric acid.

Dissolved oxygen was ensured to be above 30% air saturation by manipulating, in a first phase, the airflow rate and, when this variable achieved its maximum, the manipulated variable was shifted to the agitation speed. This control was automatically performed by the DCU. However, for ensuring a minimum of airflow and agitation inside the bioreactor, minimum set-points were imposed for these variables (3 $L \cdot min^{-1}$ and 150 rpm, respectively).

Foam was disrupted by addition of silicon antifoaming agent on demand, and foaming was not a general problem of this strain under the present conditions.

Table 4.8 Set-points for the controlled state variables

Variable	Set-point	Mode of control
Temperature	37ºC	PID
pH	7	PID
Dissolved Oxygen	30%	Cascade

4.4.3 Medium Composition

A modified M9 medium with addition of trace metals and vitamins was used for the batch cultivations of the microorganism and addition of glucose and ammonia was performed during the fed-batch phase (Riesenberg et al., 1990). In the following tables, the medium used in the fermentations is described. 300 g of inocula were prepared in the same way as batch medium, except that trace metal and vitamin solutions were not included in this medium formulation. For the preparation of the batch media, components marked with 1 in Table 4.9 were sterilized dissolved in distilled water inside the bioreactor in an autoclave for 40 minutes. The total weight of this solution was 2.4 kg. Another solution with a total weight of 0.5 kg was prepared with components 2 and was autoclave sterilized in a shake flask. Finally, solution with components 3 had a total weight of 0.1 kg and was filter sterilized with acrodisc syringe filters (from Pall Corp., USA). For the preparation of inocula and feeding media, the procedure was very similar.

Table 4.9 Composition of the media used in the fermentation. Components 1 and 2 are sterilized separately in the autoclave, while components 3 are filter sterilized

Component	Batch ($g \cdot kg^{-1}$)	Fed-Batch ($g \cdot kg^{-1}$)	Sterilization
Glucose	5.0	250	1
Na_2HPO_4	6.0	-	2
KH_2PO_4	3.0	-	2
NH_4Cl	1.0	10	2
NaCl	0.50	-	2
$MgSO_4.7(H_2O)$	0.12	4	3
Tiamine	0.34	-	3
$CaCl_2.2(H_2O)$	0.015	-	3
Kanamicin	0.050	-	3
Trace metals	2.0	-	3
Vitamins	2.0	-	3

Table 4.10 Composition of the trace metals solution

Component	Concentration ($g \cdot L^{-1}$)
$FeCl_3$	27
$ZnCl_2$	2.0
$CoCl_2$	2.0
$NaMoO_4$	2.0
$CaCl_2$	1.0
$CuCl_2$	1.0
H_3BO_3	0.50
HCl	100

Table 4.11 Composition of the vitamin solution

Component	Concentration ($g \cdot L^{-1}$)
Riboflavin	0.42
Pantothenic Acid	5.4
Nicotinic Acid	6.1
Piridoxin	1.4
Biotin	0.060
Folic Acid	0.042

4.4.4 Operating Steps

For operating a given fermentation, and after media sterilization, a certain sequence of steps was conducted in order to ensure the reproducibility of the experiments. These steps can be summarized as follows:

1. Sterile addition of media 2 and 3 (from Table 4.9) to medium 1
2. Turn stirrer controller on for a 150 rpm set-point
3. Initialize temperature control and wait for stable readings
4. Calibrate the dissolved oxygen probe for 0% by sparging the medium with nitrogen according to the DCU instructions.
5. Calibrate the dissolved oxygen probe for 100% by sparging the bioreactor with 10 $L \cdot min^{-1}$ of air and increasing the stirrer set-point to 500 rpm
6. Calibrate the pH probe outside the bioreactor with pH 4 and 7 standards
7. Introduce pH probe inside the bioreactor after sterilizing it with ethanol (70%, commercial)
8. Initialize pH control
9. Initialize the cascade control of dissolved oxygen
10. Set the bioreactor balance to zero
11. Inoculate for beginning the batch phase
12. When both substrates (glucose and acetate) are exhausted (dissolved oxygen increases sharply and the pH stabilizes after a smooth increase, corresponding to acetate consumption), prepare the feeding system (vessel, pump, and balance) for beginning the fed-batch phase
13. Initialize remote control and acquisition by starting the LabVIEW program `Coli Monitor.vi` that will command feeding addition.

4.5 CONCLUSIONS

A modular system for on-line data acquisition and control system was developed in the LabVIEW graphical programming environment to allow the integration of several on-line measurements in one supervisory computer. When compared with commercial software programs for the same task, this one presents several advantages, like the possibility of implementing complex control algorithms. Also, it allows the easy integration of virtually any new measurement that would result from an upgrade to the monitoring system, due to its modular structure. Finally, the graphical environment facilitates the elaboration of new programs, since only a minimum knowledge of programming techniques is required.

4.6 REFERENCES

Bartman, C. D. The application of a quadrupole mass spectrometer to biotechnology process control. *Mass Spectrometry in Biotechnological Process Analysis and Control*. (Heinzle, E. and Reuss, M., Eds.). Plenum Press, New York. 49-61. 1987.

Busch, M., Höbel, W., and Polster, J. Software FIACRE: bioprocess monitoring on the basis of flow injection analysis using simultaneously a urea optode and a glucose luminescence sensor. *Journal of Biotechnology.* 31, 327-343. 1993.

Castro, C. *Construction of a genetic expression system for the production of human interferon alpha 2b in Escherichia coli*. Universidade Técnica de Lisboa, Portugal. PhD thesis. 2003.

Claes, J. E. and Van Impe, J. F. On-line monitoring and optimal adaptive control of the fed-batch baker's yeast fermentation. *Horizon of BioProcess Systems Engineering in 21st Century (7th Int. Conf. Computer Applications on Biotechnology, Osaka, Japan).* (Yoshida, T. and Shioya, S., Eds.). Pergamon, Oxford. 405-410. 1998.

Ding, T., Bilitewski, U., Schmid, R. D., Korz, D. J., and Sanders, E. A. Control of microbial activity by flow injection analysis during high cell density cultivation of *Escherichia coli*. *Journal of Biotechnology.* 27, 143-157. 1993.

Dubach, A. and Märkl, H. Application of an extended kalman filter method for monitoring high density cultivation of *Escherichia coli*. *Journal of Fermentation and Bioengineering.* 73, 396-402. 1992.

Fagervik, K., Rydström, M., Rothberg, A., and Weegar, J. Computerized modeling and supervision of fermentation processes. *Horizon of BioProcess Systems Engineering in 21st Century (7th Int. Conf. Computer Applications on Biotechnology, Osaka, Japan).* (Yoshida, T. and Shioya, S., Eds.). Pergamon, Oxford. 469-174. 1998.

Ferreira, B. S., van Keulen, F., and Fonseca, M. M. R. Novel calibration method for mass spectrometers for on-line gas analysis. Set-up for the monitoring of a bacterial fermentation. *Bioprocess Engineering.* 19, 289-296. 1998.

Ferreira, E. C. *Identificação e controlo adaptativo de processos biotecnológicos*. Universidade do Porto, Portugal. PhD thesis. 1995.

Fowler, G. L., Higgs Jr, R. E., Clapp, D. L., Alford Jr, J. S., and Huber, F. M. Development of real time expert system applications for the on-line analysis of fermentation respiration data. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 173-178. 1992.

Gregory, M. E. and Turner, C. Open-loop control of specific growth rate in fed-batch cultures of recombinant *E. coli*. *Biotechnology Techniques.* 7, 889-894. 1993.

Hagander, P. and Holst, O. Substrate control in fed-batch cultivations using a model-based modification of a PI-regulator. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 77-82. 1992.

Heinzle, E. Mass spectrometry for on-line monitoring of biotechnological processes. *Advances in Biochemical Engineering / Biotechnology.* 35, 1-45. 1987.

Heinzle, E. Present and potential applications of mass spectrometry for bioprocess research and control. *Journal of Biotechnology.* 25, 81-114. 1992.

Heinzle, E., Kramer, H., and Dunn, I. J. State analysis of fermentation using a mass spectrometer with membrane probe. *Biotechnology and Bioengineering*. 27, 238-246. 1985.

Hopkins, D. J., Betenbaugh, M. J., and Dhurjati, P. Effects of dissolved oxygen shock on the stability of recombinant *Escherichia coli* containing plasmid pKN401. *Biotechnology and Bioengineering.* 29, 85-91. 1987.

Hsiao, J., Ahluwalia, M., Kaufman, J., Clem, T. R., and Shiloach, J. Adaptive control strategy for maintaining dissolved oxygen concentration in high density growth of recombinant *E. coli*. *Annals of New York Academy of Sciences.* 321-333. 1990.

Huang, Y. L., Foellmer, T. J., Ang, K. C., Khoo, S. B., and Yap, M. S. G. Characterization and application of an on-line flow injection analysis/wall-jet electrode system for glucose monitoring during fermentation. *Analytica Chimica Acta.* 317, 223-232. 1995.

Johnston, W., Cord-Ruwisch, R., and Cooney, M. J. Industrial control of recombinant *E. coli* fed-batch culture: new perspectives on traditional controlled variables. *Bioprocess and Biosystems Engineering.* 25, 111-120. 2002.

Kellerhals, M. B., Kessler, B., and Witholt, B. Closed-loop control of bacterial high-cell-density fed-batch cultures: production of mcl-PHAs by *Pseudomonas putida* KT2442 under single-substrate and cofeeding conditions. *Biotechnology and Bioengineering.* 65, 306-315. 1999.

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

Konstantinov, K., Kishimoto, M., Seki, T., and Yoshida, T. A balanced DO-stat and its application to the control of acetic acid excretion by recombinant *Escherichia coli*. *Biotechnology and Bioengineering.* 36, 750-758. 1990.

Konstantinov, K., Nishio, N., Seki, T., and Yoshida, T. Physiologically motivated strategies for control of the fed-batch cultivation of recombinant *Escherichia coli* for phenylalanine production. *Journal of Fermentation and Bioengineering.* 71, 350-355. 1991.

Korz, D. J., Rinas, U., Hellmuth, K., Sanders, E. A., and Deckwer, W.-D. Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. *Journal of Biotechnology.* 39, 59-65. 1995.

Lee, J. H., Hong, J., and Lim, H. C. Experimental optimization of fed-batch culture for poly-β-hidroxybutiric acid production. *Biotechnology and Bioengineering.* 56, 697-705. 1997.

Miller, G. L. Use of dinitrosalicylic acid reagent for the determination of reducing sugar. *Analytical Chemistry.* 31, 426-428. 1959.

Nakajima, M., Siimes, T., Yada, H., Asama, H., Nagamune, T., Linko, P., and Endo, I. Knowledge based diagnosis in lactic acid fermentation. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 179-183. 1992.

Oeggerli, A. and Heinzle, E. On-line analysis of volatiles in fermenter exhaust gas using mass spectrometry. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 295-298. 1992.

Oeggerli, A. and Heinzle, E. On-line exhaust gas analysis of volatiles in fermentation using mass spectrometry. *Biotechnology Progress.* 10, 284-290. 1994.

Paalme, T., Kahru, A., Elken, R., Vanatalu, K., Tiisma, K., and Vilu, R. The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat). *Journal of Microbiological Methods.* 24, 145-153. 1995.

Paalme, T., Tiisma, K., Kahru, A., Vanatalu, K., and Vilu, R. Glucose-limited fed-batch cultivation of *Escherichia coli* with computer-controlled fixed growth rate. *Biotechnology and Bioengineering.* 35, 312-319. 1989.

Pfaff, M., Wagner, E., Wenderoth, R., Knüpfer, U., Guthke, R., and Riesenberg, D. Model-aided on-line glucose monitoring for computer-controlled high cell density fermentations. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 6-11. 1995.

Riesenberg, D., Menzel, K., Schulz, V., Schumann, K., Veith, G., Zuber, G., and Knorre, W. A. High cell density fermentation of recombinant *Escherichia coli* expressing human interferon alpha 1. *Applied Microbiology and Biotechnology.* 34, 77-82. 1990.

Saucedo, V. M., Valentinotti, S. C., Karim, M. N., and Collins, H. W. Feedback control of a recombinant fed-batch fermentation using on-line HPLC measurements. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 24-28. 1995.

Scheper, T., Hitzmann, B., Stärk, E., Ulber, R., Faurie, R., Sosnitza, P., and Reardon, K. F. Bioanalytics: detailed insight into bioprocesses. *Analytica Chimica Acta.* 400, 121-134. 1999.

Schügerl, K. Progress in monitoring, modeling and control of bioprocesses during the last 20 years. *Journal of Biotechnology.* 85, 149-173. 2001.

Shin, C. S., Honh, M. S., Bae, C. S., and Lee, J. Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of *Escherichia coli* BL21 (DE3) [pET-3aT2M2]. *Biotechnology Progress.* 13, 249-257. 1997.

Sonnleitner, B. New concepts for quantitative bioprocess research and development. *Advances in Biochemical Engineering / Biotechnology.* 54, 155-188. 1996.

Sonnleitner, B. Bioprocess automation and bioprocess design. *Journal of Biotechnology.* 52, 175-179. 1997.

Sonnleitner, B. Instrumentation of biotechnological processes. *Advances in Biochemical Engineering / Biotechnology.* 66, 1-64. 1999.

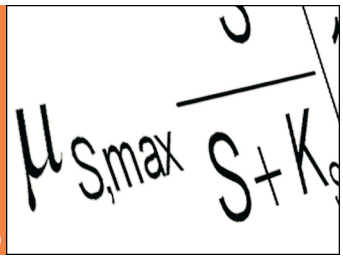
Tomson, K., Paalme, T., Laakso, P. S., and Vilu, R. Automatic laboratory-scale fed-batch procedure for production of recombinant proteins using inducible expression systems of *Escherichia coli*. *Biotechnology Techniques.* 9, 793-798. 1995.

Turner, C., Gregory, M. E., and Thornhill, N. Closed-loop control of fed-batch cultures of recombinant *Escherichia coli* using on-line HPLC. *Biotechnology and Bioengineering.* 44, 819-829. 1994a.

Turner, C., Gregory, M. E., and Turner, M. K. A study of the effect of specific growth rate and acetate on recombinant protein production of *Escherichia coli* JM107. *Biotechnology Letters.* 16, 891-896. 1994b.

Wipf, B., Weibel, E. K., and Vogel, S. Computer controlled large scale production of α-interferon by *E. coli*. *Bioprocess Engineering.* 10, 145-153. 1994.

Yang, X.-M. Optimization of a cultivation process for recombinant protein production by *Escherichia coli*. *Journal of Biotechnology.* 23, 271-289. 1992.



CHAPTER 5
DEVELOPMENT OF THE MATHEMATICAL MODEL

"Everything should be represented as simple as it is, but not simpler".

Albert Einstein.

In this chapter, the development of a mathematical model for the fed-batch fermentation of *E. coli* is described. It is composed of mass balance equations to the most relevant state variables of the process. Kinetic equations are based on the three possible metabolic pathways of the microorganism: glucose oxidation, fermentation of glucose and acetate oxidation. A genetic algorithm was used to estimate both yield and kinetic coefficients of the model, minimizing the normalized quadratic differences between simulated and real values of the state variables, by manipulating both yield and kinetic coefficients. After parameter estimation, a sensitivity function analysis was applied to evaluate the sensitivity of the state variables to variations in each model parameter. Thus, essential parameters were selected and the model was re-written in a more simplified form that could also describe accurately experimental data.

5.1 THEORY

5.1.1 Introduction

Several valuable products such as recombinant proteins, antibiotics and amino acids are produced commercially using fermentation techniques and there is an enormous economic incentive to optimise such processes.

However, fermentation processes are typically very complex, involving different transport phenomena, microbial components and biochemical reactions. It is thus, often very difficult to keep the process conditions at pre-determined values in an optimum state, without a proper prediction and description of the key process variables relevant to the system.

This could eventually be achieved by extensive experimentation, based on a trial-and error approach. However, considering the extremely high costs of industrial-scale and even lab-scale experimentation, this approach is not practicable.

For a proper design and operation of fermentation processes, a mathematical representation of the process can thus serve as a model to study the process. This study can enable the simulation of the behaviour of the actual process by computations. Such information can then be used to describe results precisely and to make accurate predictions of the results of changes in the values of the independent variables and parameters. In this way, the models are most useful for deriving an optimal process design as well as for elucidation of new physiological relationships. They can also serve as a basis for tuning control algorithms that can be used to improve the process performance. In the case of adaptive control laws, the model of the process is an essential framework.

In most cases, every model represents a certain approximation of the real system and the best model will combine enough complexity to capture the process characteristics with sufficient simplicity to allow the model to be easily initiated and understood.

According to this complexity, there are several classifications of the models used in biotechnology. The most commons divide the models in structured and unstructured and/or segregated and unsegregated. According to Bailey (1998), the term “segregated” is used to indicate explicit accounting for the presence of heterogeneous individuals in a population of cells, and the term “structured”, to designate a formulation in which cell material is composed of multiple chemical components. Nielsen et al. (1991) define unstructured models as models

where the biomass is described by one variable, and structured models as models where intracellular structure is considered.

The four different combinations of segregation and structure, or their absence, conveniently partition the most common types of mathematical descriptions and of experiments into four disjoint categories as stated in Figure 5.1. Most models and most measurements belong to the unsegregated class.

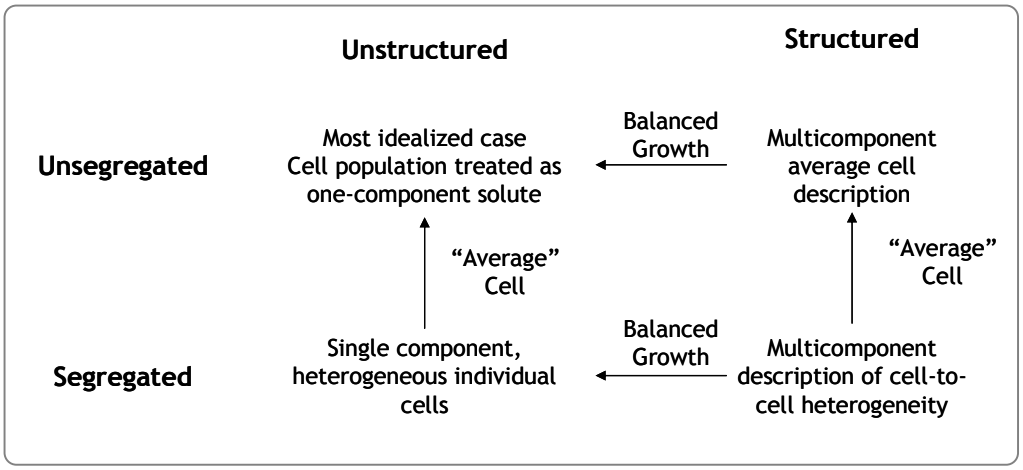


Figure 5.1 Classification of mathematical representation of biological processes according to Bailey (1998). The unstructured and unsegregated cases represent the simplest way to model a bioprocess.

Unstructured models include the most fundamental observations concerning microbial growth processes and account for key process variables (cell density, product and substrate concentrations), but do not recognize any internal structure of the cell. Therefore, the model fails whenever it is used inappropriately to describe situations where the cell composition is an important variable. In these cases, model predictions can be improved by using structured models. Many examples of these situations, like the ones associated with rapid changes in the environment of the cell and with the production of secondary metabolites or recombinant proteins, are discussed by Nielsen and Viladsen (1992). These authors also discuss many different attempts to structure the mathematical description of the metabolism of microorganisms. Those models are compared with a general framework to describe microbial kinetics. Bacterial, yeast and fungal structured models are also presented.

Structured models are as well most appropriate to describe the accumulated knowledge about metabolic pathways and to cope with the recent development in metabolic engineering (Schügerl, 2001).

However, the main disadvantage of structured models is their complexity. Often, the models incorporate too many equations and unknown parameters and provide a qualitative, rather than a quantitative description of the process (Tartakovsky et al., 1997).

Segregated models are usually applied when variant subpopulations coexist, for example, cells with or without plasmids or cells which differ in size as a consequence of asymmetric cell division processes like it happens in yeasts.

Several examples of structured and segregated models applied to recombinant microorganisms are revised in Zabriskie and Arcuri (1986).

Another classification divides models into white-, grey- and black-box models (Lübert and Jorgensen, 2001). Models that are based on heuristic knowledge are termed grey-box models to distinguish them from purely data driven black-box models and from mechanistically completely understood mathematical or white-box models. Grey-box models, usually based on rules-of-thumb or fuzzy rules, are used when that mechanistically-based knowledge about process components is unavailable. When properly formulated, these can easily be processed by means of computers. The most common applications of black-box models are based on artificial neural networks (ANN's) that use the tools of artificial intelligence to model a given bioprocess. Such ANN's are crisply defined non-linear relations. Their structure is characterized by the number of nodes and by the response functions through which these nodes respond to their averaged inputs. The parameters in these representations are termed weights, as they are used to compute weighted averages of inputs to the nodes. As in other black-box models, these parameters must be determined by fitting the model to experimental data. This process is usually referred to as the training procedure. A combination of grey- and black-models can also be applied, such as neural-fuzzy and fuzzy-expert systems.

Several researchers developed mathematical models for the description of the growth and host-plasmid interaction in recombinant *E. coli*. These models can be classified according to the distinctions made above.

A four compartment structured model was presented by Nielsen et al. (1991) to describe *E. coli* growth, based on the *Lactobacillus* model. In this model, the different biomass components are grouped in 4 intracellular variables. The structured model is described and its performance compared with experimental data from fermentations with recombinant *E. coli,* being able to express the majority of the observations made through fermentations with recombinant microorganisms. This model is especially suited for description of dynamic

changes in plasmid copy number, like in the work described in Patnaik (1998b), where that model was applied to the fermentative production of β-galactosidase by *E. coli*. In this case, in order to characterize heterogeneity, an earlier model visualizing the broth as a set of two reactors with internal recycle has been modified and included in the global model. In Rhee and Schügerl (1998), the growth of the recombinant *E. coli* JM109 and the production of the protein EcoRI::SPA were also simulated using the four-compartment model.

Other structured models were used by Bentley and Kompala (1989) and Betenbaugh and Dhurjati (1990) to describe some features of recombinant *E. coli* growth. Harcum (2002) presents a novel structured unsegregated kinetic model that is used to predict intracellular amino acid pool shortages during recombinant protein over-expression in *E. coli*, in order to avoid the protease activity.

In the work described by Tartakovsky et al. (1997), two novel approaches for modelling processes that can be described by a sequence of phases (metabolic states) are suggested and applied to *E. coli* fermentations. The first approach uses a multi-compartment model framework, coupled with knowledge-based logic. In the second approach, the multi-compartment model is reduced into the so-called Variable Structured Model consisting on a battery of alternative models, representing qualitatively each one of the process steps. Simulated intracellular process variables are then compared with the output of a multi-wavelength fluorosensor.

Togna et al. (1993) established a simple mathematical model to explain the complex population dynamics of an *E. coli* host-plasmid expression/excretion system for β-lactamase within single- and two-stage reactors. It integrates the regulatory, genetic, and population dynamics aspects of the system and is designed to predict operating conditions under which stable production and excretion of the product are likely to occur. Also in Lee et al. (1988), a genetically structured kinetic model was applied to a two-stage continuous culture system with recombinant *E. coli.*

Many of the tools for the development of structured models for *E. coli* can be found in the work described by Palsson and co-workers in Varma and Palsson (1993a); Varma and Palsson (1993b); Varma et al. (1993) and Varma and Palsson (1995). In Varma and Palsson (1994) a stoichiometric matrix for the catabolic reaction pathways of *E. coli* is presented that allows the use of linear programming to explore the boundaries of achievable metabolic performance.

The application of segregated or population models to *E. coli* fermentation is usually done in order to describe the changes in the plasmid content of the growing population. The simplest

of these models lumps all the cells containing plasmids and all the cells lacking plasmids into two separate populations. Other models consider different subpopulations.

In Mosrati et al. (1993), a segregated model was proposed to evaluate the specific growth rate of plasmid-carrying cells and plasmid-free cells as well as the probability of plasmid loss during continuous cultures. In Patnaik (1998a), the author describes a structured and segregated model that is a combination of the models described by Nielsen et al. (1991) and by Mosrati et al. (1993). This new model is used to analyze the effect of interruptions in substrate inflow in a fed-batch fermentation of *E. coli.* In Patnaik (2002), there is a description of another segregated model accounting for the Gaussian distribution of cells plasmid copy number.

In order to circumvent the already mentioned disadvantages of the structured and segregated models, both grey- and black-box models have been applied to the study of *E. coli* fermentation. For example, in Jin et al. (1994), several types of fuzzy control with pH-stat were considered for the production of β-galactosidase.

In Warnes et al. (1995), four different modelling techniques are considered for the *E. coli* fermentation process: Multiple Linear Regression (MLR), Principal Component Regression (PCR), Non-linear Auto-Regressive Moving Average with eXogeneous inputs (NARMAX) and ANN's. The models used industrial data from the process as inputs in order to forecast the concentration of biomass and recombinant protein.

The main objective of the present work is to develop a model that, at the same time, could be able to predict and simulate the process and to allow the application of optimal and adaptive control algorithms. In this particular case, the choice of the type of model to use is limited by factors such as the experimental data availability, simplicity of manipulation, parameter identifiability and computation time. Thus, both structured and segregated model types were considered too complex for these purposes and a simple unstructured and unsegregated approach was chosen. Black- and grey- box models could also perform well for process control, but the derivation of these models, although very useful in some situations, does not elucidate any details concerning the metabolic and physiological behaviour of the culture.

There can be found several unstructured and unsegregated models for describing *E. coli* fed-batch fermentations in the literature. For example, in Alba and Calvo (2000), it is presented a relatively complex unstructured and unsegregated white-box model that describes the main phenomena involved in a cultivation process, including chemical environment as substrate and product concentrations, and physical phenomena as fluid dynamics and mass transfer.

Stoichiometry, fluxes analysis, microbial kinetics, fluid dynamics and mass transfer are involved in the process description. The overall framework is used to predict the whole process in both stirred tank and airlift bioreactors.

Lee and Ramirez (1992) proposed a new generalized mathematical model for recombinant *E. coli* that has the particularity of including inducer effects on cell growth and foreign protein production. Another approach with a similar purpose can be found in Cockshott and Bogle (1999), applied to the production of bovine somatotropin with *E. coli.*

A model that resulted from an iterative bioprocess design is presented in Galvanauskas et al. (1998) and Galvanauskas et al. (1999), while another one, giving a particular relevance to the oxidative bottleneck of *E. coli* can be found in Akesson et al. (1999). An interesting approach to describe the same phenomenon was proposed by Xu et al. (1999), where the kinetics models are based on metabolic fluxes.

Some details of these models will be discussed later in a deeper detail in what the kinetic equations concerns. It follows a brief explanation of the main tool used for the construction of the *E. coli* model proposed in this work.

5.1.2 General State Space Dynamical Model

In a biological reactor, the variation of each component concentration can be explained as the result of the combination of two different phenomena (Lübert and Jorgensen, 2001):

- the transformation of some components into other components, due to chemical, biological or biochemical reactions;
- the transport and mass transfer processes, i.e. the spatial motions of the material in the reactor in order to bring the reagents into contact with each other.

In order to accurately describe these phenomena, the bioprocess model will always consist of three sets of equations (Andrews, 1993): the process mass conservation equations for each component that describe the type of bioreactor; the yield equations that describe the relationships between the rates of consumption and production of the various components; and the rate equations, describing the actual kinetics of the process. The yield and rate equations are independent of the type of bioreactor, at least theoretically, and together constitute a description of the metabolism and how it is affected by the cell's physicochemical environment.

A major difference between these three sets of equations is the reliability and certainty with which they can be derived and employed. The form of mass conservation equations depends mainly on the mixing conditions in the reactor, and they can be written with confidence for common types of bioreactors. The yield equations are based on the intracellular balances of energy and various elements, so they necessarily contain several assumptions about microbial metabolism. Thus, while the mass conservation and yield equations have a theoretical basis, the rate equations are purely empirical in nature.

In stirred tank reactors (Figure 5.2), the process is assumed to be in a completely mixed condition: this implies that the composition of the medium is homogeneous inside the reactor. Therefore, the present approach will not deal with heterogeneities in the medium outside the cells. However, the biochemical states exhibit complex changes with time, and this can be described by the use of dynamical models.

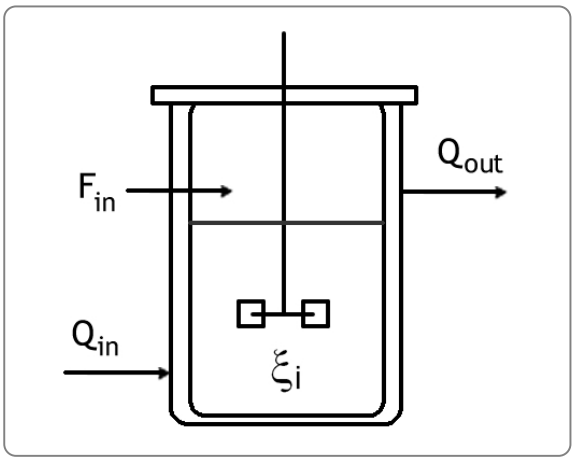


Figure 5.2 Fed-batch stirred tank bioreactor. The reactor is considered in a completely mixed condition, with the concentration of the components being represented by ξ_i. There is only one liquid stream entering the bioreactor (F_{in}), one gaseous inflow (Q_{in}) and one gaseous outflow (Q_{out}).

Those processes are commonly specified by the following reaction scheme for the j^{th} reaction (Bastin and Dochain, 1990):

$$\sum_{i \in L_j} k_{ij} \xi_i \xrightarrow{r_j} \sum_{i \in R_j} k_{ij} \xi_i \qquad j = 1, \ldots, m \qquad \text{Equation 5.1}$$

Where L_j denotes the set of indices of the components ξ_i which are the reactants of the j^{th} reaction; while R_j is the set of indices of the components ξ_i which are the reaction products of the j^{th} reaction. r_j is the reaction rate of the equation j.

The dynamical model for this kind of bioreactors can be obtained by mass balances to the components involved in a reactive scheme. The generic balance can be described as follows:

$$Accumulation = Conversion + (Input - Output) \qquad \text{Equation 5.2}$$

Given the reaction scheme of Equation 5.1 of a biotechnological process, the derivation of the dynamical model can be made systematically based on the General Dynamical Model of Bioreactors described by Bastin and Dochain (1990). For a reaction scheme involving n components and m reactions, the general model will be composed of n ordinary differential equations representing the balances to the components ξ_i.

The dynamics of the concentration of each component on a fed-batch reactor are then written as follows:

$$\frac{d\xi_i}{dt} = \sum_{j \in m} (\pm) k_{ij} r_j - \frac{F_{in,tot}}{W}\xi_i + \frac{F_{in,i}}{W}\xi_{in,i} + \frac{Q_{in}}{W}\xi_{in,i}^{G} - \frac{Q_{out}}{W}\xi_{out,i}^{G} \quad \text{i= 1,..., n;} \qquad \text{Equation 5.3}$$

Where:

ξ_i	state variable (concentration of the component i);
k_{ij}	yield coefficients or stoichiometric coefficients of component i in the reaction j;
$F_{in,tot}$	total liquid mass flow fed into the bioreactor;
$F_{in,i}$	liquid mass flow of the stream containing component i fed into the bioreactor;
$\xi_{in,i}$	concentration of the component i in the corresponding inflow stream;
Q_{in}	volumetric inflow to the bioreactor in gaseous form;
Q_{out}	volumetric outflow from the bioreactor in gaseous form;
$\xi_{in,i}^{G}$	concentration of the component i in the gaseous inflow;
$\xi_{out,i}^{G}$	concentration of the component i in the gaseous outflow;
W	weight of liquid inside the bioreactor.

The variable dilution rate D (D^{-1} is the residence time) can be defined as the quotient between the total liquid mass flow feed rate and the weight of liquid inside the reactor, simplifying the second term on the right side of Equation 5.3:

$$D \equiv \frac{F_{in,tot}}{W} \qquad \text{Equation 5.4}$$

The third term of Equation 5.3 can be simply referred to as F_i, corresponding to the mass inflow of component i in the bioreactor per weight unit inside the bioreactor, while the last two terms can be referred to as Q_i, corresponding to the mass net gaseous flow of component i per unit of weight inside the bioreactor.

k_{ij} are strictly positive constant yield coefficients. They have a "-" sign when ξ_i is a reactant and a "+" sign when ξ_i is a product of the reaction. These coefficients are without dimension (*i. e.* units of mass/mass) as a consequence of expressing the reaction rates in terms of the rate of one component, called nominal. For a given reaction, the coefficient associated with the nominal component is equal to the unit, while for components that do not participate in that reaction it is equal to zero.

The general model can also be written in the matrix form, as follows:

$$\frac{d\xi}{dt} = Kr(\xi,t) - D\xi + F + Q \qquad \text{Equation 5.5}$$

Where:

- **ξ** state vector composed by the concentrations of the components $\xi^T = [\xi_1, \xi_2, \ldots, \xi_n]$;
- **K** yield or stoichiometric coefficients matrix $K = [k_{ij}]$;
- **r** reaction rates vector $r^T = [r_1, r_2, \ldots, r_m]$;
- **F** liquid mass flow feed vector (per unit of weight inside the reactor): $F^T = [F_1, F_2, \ldots, F_n]$;
- **Q** vector of rate of mass flow in gaseous form (per unit of weight inside the reactor): $Q^T = [Q_1, Q_2, \ldots, Q_n]$.

In this equation there is an evidence of the two phenomena occurring inside a bioreactor: the first term **Kr** describes the kinetics of the biochemical and microbiological reactions, which are involved in the process, and the remaining terms **–Dξ + F – Q** describe the transport dynamics of the components through the reactor.

Additionally, it could be necessary, depending on the mode of operation, to use another equation for mass variation inside the reactor:

$$\frac{dW}{dt} = F_{in,tot} \qquad \text{Equation 5.6}$$

5.1.3 Kinetic Models for *E. coli*

At first glance, the solution of the simple ordinary equation system described by Equation 5.5 should not pose any difficulty. However, the rate expressions *r* usually lead to a strong non-linear coupling between the various equations. This non-linearity is due to the non-linear kinetic rate expressions for the key systems components.

A very simple example is the kinetic expression of Monod, which non-linearly relates the biomass growth rate to biomass and substrate concentrations, indicating that for large

concentrations of the limiting substrate the specific cell mass production rate becomes independent of that particular substrate concentration:

$$r_x = \mu X = \mu_{\max} \frac{S}{K_S + S} X \qquad \text{Equation 5.7}$$

Where:

r_X	growth rate;
μ	specific growth rate;
X	biomass concentration;
μ_{max}	maximum specific growth rate;
K_S	affinity constant for the substrate;
S	substrate concentration.

The parameters μ_{max} and K_m cannot be determined *a priori*, but must be inferred from experimental data, by a process known as model parameter identification.

Extensions of the simple Monod expression, available in the literature, address factors reducing growth rate, via metabolite repression and other inhibition and limitation effects. These factors significantly affect both microbial growth and product formation and, thus, impact on bioreactor performance. Kinetic models constructed with this type of equations are usually included in the unstructured type.

Another approach to model bioprocess kinetics is to derive analytical expressions to characterize the kinetic behaviour of the microorganisms only dependent on the biomass concentration. For instance, the logistic model, in which the growth rate decreases linearly as the biomass concentration increases, is frequently employed when the latter is the sole measured state variable. In Taylor et al. (1992) several biomass-dependent kinetic models were tested to describe *E. coli* growth in glycerol, and glucose/acetate. Another kinetic model based only on the biomass concentration for the estimation of biomass in a batch *E. coli* fermentation is also presented by Castan et al. (2002). These authors use an approach also based on the decomposition of cultures into regimes, each one described with a different model.

The previous expressions usually account for the influence of state variables in the reactions kinetics. Thus, the time dependence of the reaction rates expressed in Equation 5.5 is accounted by the variation of the state variables with time. However, it is commonly accepted that, especially at high-cell densities, the kinetic parameters like the ones described in Equation 5.7 are not constant over long periods of time and vary according to the culture characteristics, though in most cases this dependence is frequently considered negligible in order to simplify the kinetics modelling. Nevertheless, some approaches used to calculate

kinetic rates allow accounting for this dependence. For example, some algorithms used for on-line estimation of reaction rates using biomass concentration or other correlated variables measurements have been proposed that can account for this phenomenon. Pomerleau and Perrier (1990) and Pomerleau and Perrier (1992) proposed and validated experimentally an on-line estimation algorithm for multiple reaction rates. This procedure is applied to baker's yeast fermentation, and the algorithm required the on-line measurement of two or three state variables. Also in Lubenova (1999), the author describes a methodology for the design of a new parameter estimator of biomass growth rate and yield coefficient for oxygen consumption on the basis of the theory of adaptive estimation, for a class of aerobic bioprocesses in fed-batch or continuous mode.

In Ljubenova and Ignatova (1994), the authors proposed an approach for on-line growth rate estimation for a class of aerobic batch processes with dissolved oxygen control in the culture medium. The only required on-line measurement is the oxygen consumption rate.

In Oliveira et al. (1996), the authors describe an adaptive model-based algorithm for the on-line estimation of reaction rates, considering the yield coefficients invariable and known, based on the approach of Bastin and Dochain (1990) to stirred tank reactors. Another method based on adaptive control was used by Zeng et al. (1992) for estimating both biomass concentration and specific growth rate, assuming that the structure of the specific growth rate is known. This knowledge is then used to control the specific growth rate in a continuous fermentation process using a single measured variable: the substrate.

However, these last approaches are usually undertaken for control purposes, but are unable to generate characteristic parameters of the cultures and to constitute part of a model that can be used to entirely simulate the fermentation process. Thus, a complete set of equations derived from the Monod equation is the most suitable for application in the process modelling. Nevertheless, although there are many examples in the literature concerning the application of such equations to the growth of *E. coli*, there are very few complete sets of equations available that can explain the main phenomena observed during the fed-batch aerobic fermentation of this microorganism described in Chapter 2. Typically, authors develop models from batch or chemostat data that can hardly be applied to high-cell density fed-batch fermentation. In fact, it was verified by Liu (1999) that substrate-sufficient cultures like batch or continuous cultures have different metabolic behaviours from substrate-limited cultures like the majority of the fed-batch fermentations, with regard to the substrate consumption rate, growth yield and maintenance requirements, and models developed from that kind of data are not directly applied to fed-batch cultures.

In other cases, only one or two Monod type of equations are derived to describe glucose uptake or global growth rates without accounting for the glucose overflow or acetate consumption that usually occur in this kind of fermentation processes. Additionally, when applying these few kinetic models to the particular fermentation process studied in this thesis, the approximation between real and simulated results was significantly poor, thus motivating the development of a complete kinetic model that can be used to describe this particular fermentation process.

As already stated in Chapter 2, the aerobic growth of *E. coli* can follow three different metabolic pathways:

- Oxidative growth on glucose, with a net reaction as follows:

$$S + O \xrightarrow{\mu_1} X + C \qquad \text{Equation 5.8}$$

Where *S* means glucose concentration, *O* means dissolved oxygen concentration, *X* is the biomass concentration and *C* means dissolved carbon dioxide concentration.

- Fermentative growth on glucose, where the global reaction can be described as:

$$S + O \xrightarrow{\mu_2} X + C + A \qquad \text{Equation 5.9}$$

Where *A* means acetate concentration.

- Oxidative growth on acetate:

$$A + O \xrightarrow{\mu_3} X + C \qquad \text{Equation 5.10}$$

The first problem detected in some of the models described in the literature (for example, Cockshott and Bogle, 1992 and Cockshott and Bogle, 1999) is that the global specific growth rate on glucose is usually directly calculated from the state variables, thus not accounting for the differences between the different biomass yields obtained when the bacterium metabolizes this sugar under oxidative or fermentative pathways. Indeed, if Equation 5.7 is applied directly to the calculation of one single specific growth rate on glucose in *E. coli,* which would include μ_1 from Equation 5.8 and μ_2 from Equation 5.9, the oxidative bottleneck is not considered. In this case, models usually account for a global yield on glucose that is an average of the oxidative and fermentative yields. However, there are situations when only oxidative metabolism takes place and this global yield underestimates real values. Also, even when both pathways occur, the ratio between oxidative and fermentative growth is not constant, and thus a global yield is not adequate. Moreover, in these articles, for the calculation of a global glucose consumption rate to apply in the glucose balance equation, the authors include not only the uptake rate obtained from the global specific growth rate but also

a fermentative growth rate based on acetate production that has already been included in the global growth rate on glucose, duplicating the effect of fermentative growth on glucose in the global glucose uptake rate of the culture. Finally, the main drawback related with this kinetic modelling is that, when trying to optimise biomass production or biomass-related products, there is no “driving force” to direct glucose metabolism into the oxidative pathway, because the distinction between the yields obtained in both glucose pathways is not clearly stated.

In order to overcome these problems, in the model developed in this thesis, glucose uptake rate is first calculated as a function of glucose and acetate concentration and, depending on the value it assumes, this glucose flux can be totally directed to the oxidative metabolism, or it can be divided in oxidative and fermentative fluxes, in order to account for the oxidative bottleneck. Only after calculating these fluxes, specific growth rates are calculated for each case and a global (or observable) specific growth rate on glucose is obtained as the sum of the partial growth rates. A similar approach could only be found in Akesson et al. (1999) and Xu et al. (1999). However, the model described in the first case is specific for control applications and thus covers only some of the metabolic pathways described above. The second one is by far the most complete and is based on metabolic flux analysis. However, every flux through a certain metabolic pathway is further divided into anabolic and energy fluxes, adding many additional parameters to the model. Nevertheless, it can be shown that the model described by these authors can be in some features equivalent to some equations presented in this section, by considering the anabolic fluxes included in the calculated (or apparent) yield coefficients.

Using this approach, the relationship between specific uptake rates and growth rates is as follows:

$$q = \frac{\mu}{k} \qquad \text{Equation 5.11}$$

Where q is the specific uptake rate and k is the yield coefficient for that particular reaction.

In the model described in Galvanauskas et al. (1998); Galvanauskas et al. (1999) and Galvanauskas (1999), although the glucose uptake rate is calculated in the first place, the global specific growth rate on glucose is directly calculated from it using Equation 5.11, without accounting for any differences in the fermentative and oxidative yields, thus presenting the same problems as the one discussed above.

Another issue is related to the equations used to describe acetate consumption, as the ones described in the literature usually do not accurately illustrate both the situations that trigger acetate oxidation and the rate at which it takes place.

Therefore, the kinetic model derived for this particular bioprocess represents a compromise between the high complexity of structured or segregated models and the exaggerated simplicity that can be found in some unstructured models available in the literature. Also, some aspects that have been considered as inexactitudes like the ones described above, have been eliminated. Both in the literature revision and in the construction of the kinetic model, a basic nomenclature was adopted, in order to unify the several approaches found. Generally, *r* is used to designate growth rate, while μ means specific growth rate. The specific uptake or production rates of substrates and products are represented by *q*. Also, the kinetic parameters are represented by capital *K*, while yield coefficients are represented by lowercase *k*. In order to avoid confusions with the indexes of the yield coefficients of the final model deduced, and in a first stage, these yield coefficients are identified by the state variables to which they correspond, in the form k_{AB}, meaning amount of *A* per amount of *B*. The kinetic model has been divided in the following major steps:

5.1.3.1 Glucose Uptake Rate

For the majority of the authors, the total glucose uptake rate is assumed to follow the Monod model. Although not considered by all authors (Lee and Ramirez, 1992 and Akesson et al., 1999), the inhibition by acetate on glucose uptake is usually included and comes from experimental observations. This can be formulated into a non-competitive mode of inhibition (Galvanauskas et al., 1998; Galvanauskas, 1999; Galvanauskas et al., 1999 and Xu et al., 1999):

$$q_S = q_{S,\max} \frac{S}{S + K_S} \frac{K_{i,S}}{K_{i,S} + A} \qquad \text{Equation 5.12}$$

Where:

q_S specific uptake rate of glucose;
$q_{S,max}$ maximum specific uptake rate;
$K_{i,S}$ inhibition constant of acetate on glucose uptake.

In Dubach and Märkl (1992), the authors found a similar expression for glycerol uptake. As it was already mentioned, in the work of Cockshott and Bogle (1992) and Cockshott and Bogle (1999), glucose uptake rate is not calculated, and the specific growth rate on glucose is described by:

$$\mu_S = \mu_{S,\max} \frac{S}{S + K_S}\left(1 - \frac{A}{K_{i,S}}\right)$$ Equation 5.13

Thus, in this work, a different expression for acetate inhibition is used.

Another equation for the glucose uptake is the one described in Alba and Calvo (2000) as follows:

$$\mu_S = \mu_{S,\max} \frac{S}{S + K_S + \frac{K_S}{K_{i,S}} A^2}$$ Equation 5.14

In another work (Luli and Strohl, 1990), acetate inhibition of the growth rate at pH 7.0 was evaluated from a relationship where the growth rate is affected by a constant inhibition that is a logarithmic function of the acetate concentration.

There is also evidence (Kleman and Strohl, 1994) that, at very high concentration (50 $g \cdot L^{-1}$), glucose can also inhibit growth. However, as these concentrations are not usually used in fermentation processes, this phenomenon is not contemplated in the kinetic models.

5.1.3.2 Glucose Overflow

For better understanding and quantitative prediction of the acetate overflow phenomenon, a number of mathematical models (Majewski and Domach, 1990 and Varma and Palsson, 1994) were reported. However, as it was already stated, the majority of the models found in the literature does not account for the advantages of oxidative over fermentative pathways in terms of productivity. When a global specific growth rate on glucose is calculated by a Monod-type equation, the bottleneck is assumed to be reached when that rate reaches a critical value. This approach is used by Cockshott and Bogle (1999). From this point on, they consider that acetate is produced at a rate given by the following equation:

$$q_{AP} = \text{Min}\ q_{AP,max}, k_{AX}(\mu_S - \mu_{S,crit})$$ Equation 5.15

Where:

q_{AP} specific production rate of acetate;
$q_{AP,\max}$ maximum specific production rate of acetate;
k_{AX} acetate yield on biomass;
$\mu_{S,crit}$ critical specific growth rate.

A very similar approach is used by Galvanauskas et al. (1998); Galvanauskas et al. (1999) and Galvanauskas (1999), although only the second term of Equation 5.15 is considered.

For the majority of the authors, the critical specific growth rate is considered constant. However, for example in the work of Paalme et al. (1989), this value was observed to decrease with the presence of organic acids in the medium. Indeed, Axelsson et al. (1992), found that the oxidative bottleneck of the yeast *Saccharomyces cerevisiae* (a microorganism that has a similar behaviour to *E. coli* in what glucose consumption concerns) is inhibited by the presence of ethanol.

In Alba and Calvo (2000) acetate production rate and overflow phenomenon are accounted in the same equation:

$$q_{AP} = \beta S \frac{\mu_S}{\mu_{S,\max}} \left(\frac{S - S_{crit}}{S} \right) \qquad \text{Equation 5.16}$$

Where:

β constant;
S_{crit} minimum glucose concentration that gives the same specific growth rate as S.

When the specific uptake rates are calculated at a first step, usually the glucose flux is split into two metabolic ones, i.e., the fully oxidative metabolic flux, q_{SOX}, and the overflow, q_{SF}, the split being determined by a boundary condition. In Akesson et al. (1999), the authors propose the existence of a critical specific glucose uptake rate above which the glucose overflow takes place.

According to Xu et al. (1999), this boundary condition is determined by the oxidative capacity in the form $q_{OS}<q_{O,max}$, where the first is the specific oxygen uptake rate used for oxidative glucose degradation and the second is its maximum value. At low rates of sugar consumption, all sugar is channelled through the oxidative pathway and the corresponding glucose rate is q_{SOX}. When the bottleneck condition is achieved, the global glucose uptake rate is divided in q_{SOX} and q_{SF}. These authors also considered that $q_{O,max}$ is inhibited by acetate; thus, the non-competitive inhibition term with a constant $K_{i,O}$ is included. Therefore, the algorithm used for simulation includes a boundary condition:

$$q_{OS} \leq q_{O,\max} \frac{K_{i,O}}{A + K_{i,O}} \qquad \text{Equation 5.17}$$

The acetate formation rate is obtained by means of the stoichiometry of glucose conversion to acetate:

$$q_{AP} = q_{SF,en} k_{AS} \qquad \text{Equation 5.18}$$

Where $q_{SF,en}$ is the part of the fermentative glucose flux that follows to energy metabolism.

5.1.3.3 Acetate Consumption

According to some studies, there are strains that are not able to consume acetate. That is the case of Cockshott and Bogle (1999), where a strain derived from the K-12 is assumed not to be able to grow in acetate.

Also, according to Patnaik (2002), the bacterial strain gal K is only able to grow on acetate in the presence of peptone and after glucose has fallen bellow 0.01 $g \cdot L^{-1}$.

However, according to the glucose overflow theory, acetate is re-consumed by *E. coli* cells whenever the rate of glucose uptake is so low that the oxidative pathway is not saturated. Assuming that acetate uptake also follows the Monod model, the specific rate of acetate consumption would be (Xu et al., 1999):

$$q_{AC} = q_{AC,\max} \frac{A}{A + K_A}$$ Equation 5.19

Where:

q_{AC} specific acetate uptake rate;
$q_{AC,max}$ maximum specific acetate uptake rate;
K_A affinity constant for acetate.

After calculating this rate, these authors determine the corresponding oxygen flux necessary for the consumption of acetate flux. If the sum of oxygen requirements for both glucose and acetate oxidation is greater than the maximum oxygen uptake rate, $q_{O,max}$, then acetate consumption rate is recalculated in order to account for this limitation.

Other authors calculate directly the specific growth rate on acetate. In the case of Alba and Calvo (2000), there is an inhibition term for acetate concentration:

$$\mu_{AC} = \mu_{AC,\max} \frac{A}{A + K_A} \left(1 - \frac{A}{A_{crit}}\right)$$ Equation 5.20

In this case, above a critical acetate concentration A_{crit}, cells cease to grow and the uptake rate of acetate becomes zero. However, the total oxidative capacity is not accounted, and the calculated specific growth rate on acetate can be greater than the one imposed by the oxidative bottleneck.

In Patnaik (2002), a term for acetate inhibition is also included, but as in Equation 5.12.

According to Galvanauskas et al. (1998), Galvanauskas et al. (1999) and Galvanauskas (1999), acetate is always consumed at the maximum rate that fulfils the oxidative capacity, given by:

$$q_{AC} = q_{AC,\max} - \alpha\mu_S \qquad \text{Equation 5.21}$$

Where α is constant. The second term of this equation is related with the limited oxidative capacity of the cells. However, the global specific growth rate on glucose is not proportional to the flux of glucose following the oxidative pathway, and thus it is not straightforward that this equation can accurately describe acetate consumption rate.

5.1.3.4 Maintenance

The maintenance energy is usually defined as the amount of energy (ATP), glucose or oxygen necessary to maintain cells at zero growth rate (Paalme et al., 1997). Although some authors do not consider this phenomenon (Akesson et al., 1999 and Alba and Calvo, 2000), especially in the case of high-cell densities or slow growth, a maintenance term is normally included in the model to account for the consumption of those substrates not related with growth.

In the majority of the models available in the literature, (Nielsen and Viladsen, 1992; Varma and Palsson, 1993a; Cockshott and Bogle, 1999 and Patnaik, 2002), maintenance coefficient, q_m is accounted as the part of the specific substrate uptake rate that is not used by the cells for growth, and its relation with q and μ can be described as:

$$q = \frac{\mu}{k} + q_m \qquad \text{Equation 5.22}$$

The rate of a given substrate consumption is thus the sum of the requirements for growth and maintenance. In the work of, for example, Galvanauskas et al. (1998), the approach is very similar, although it is not considered a specific uptake rate but as a specific growth rate (m_S).

The majority of the authors only account for maintenance in the uptake of glucose. However, some authors (Patnaik, 2002) also include a similar term for the growth on acetate.

Usually, this maintenance term is considered constant, but in some cases it is considered to decrease hyperbolically as the specific growth rate increases (Patnaik, 2002):

$$m_S = m_{S,max}\frac{K_{m,S}}{K_{m,S} + \mu_S} \qquad \text{Equation 5.23}$$

Where $K_{m,s}$ and $m_{S,Max}$ are constants.

5.1.3.5 Specific Growth Rates

When substrate uptake rates or fluxes are calculated in a first place, like in Xu et al. (1999), the specific growth rate is obtained as a sum of the three substrate fluxes by means of the related yield coefficients:

$$\mu = (q_{SOX} - q_m)k_{XS,OX} + q_{SF}k_{XS,F} + q_{AC}k_{XA} \qquad \text{Equation 5.24}$$

A very similar approach is used by Akesson et al. (1999). A different mode of calculation is used by Alba and Calvo (2000). In this work, the authors applied a stoichiometric matrix to the knowledge about stoichiometry and metabolic reactions, and found that in *E coli* fermentation, all the rates are linearly dependent, and that the growth and acetate rates are independent and needed to calculate the remaining rates.

5.1.4 Dissolved Gases Dynamics

The dynamic of gaseous concentrations assumes a great importance in the aerobic *E. coli* fermentation, primarily because oxygen and carbon dioxide represent important substrates and products, respectively, in the reactions described in Equation 5.8 to Equation 5.10.

Oxygen is thus a growth limiting factor and below a critical value affects the growth rate, but at the same time can have inhibitory effects when present in excess. On the other hand, low levels of dissolved carbon dioxide are reported to stimulate growth, meanwhile increasing levels have progressive inhibitory effects. At elevated carbon dioxide concentrations, carboxylation alters various metabolite pools and has an impact on amino acid production and in acetate production. Blight and Holland (1994) used carbon dioxide enriched air in order to check for the influence of this gas is *E. coli* fermentations and found that the maximum specific growth rate was significantly reduced, and that the production of acetate and the biomass yield coefficient on glucose increased with the presence of significant concentrations of dissolved carbon dioxide. A similar conclusion can be found in Pan et al. (1987). In Jones and Greenfield (1982) and Onken and Liefke (1989), the authors revise the influence of the partial pressures of these compounds in the growth of several microorganisms.

Therefore, the ability to accurately predict and control the level of dissolved oxygen and carbon dioxide is then of considerable interest since it can result in the production of undesirable end products and lower system efficiency.

The levels of dissolved oxygen and carbon dioxide are affected by the consumption or production respectively, and the transference between phases. Important fermentation variables can then be defined using the corresponding rates: the oxygen transfer rate, abbreviated to *OTR*, can be defined as the amount (usually in grams or moles) of oxygen transferred from the gas phase to the liquid phase per units of time and of liquid volume or weight. The carbon dioxide transfer rate (*CTR*) has the same meaning for carbon dioxide, except that the transfer is done from the liquid phase into the gas phase. The terms oxygen uptake rate (*OUR*) and carbon dioxide evolution rate (*CER*) are usually applied to designate for the amount of oxygen consumed and of carbon dioxide produced by the cells, respectively, per units of time and of liquid volume or weight. In the present work, the units adopted for these rates are $g \cdot h^{-1} \cdot kg^{-1}$.

The mass balance equations for the species in liquid and gaseous phases where well-mixed phases can be considered are as follows (Alba and Calvo, 2000, and Heinzle and Dunn, 1990):

5.1.4.1 Oxygen

The balance to the gas phase comes:

$$V_G \frac{dO^G_{out}}{dt} = Q_{in} O^G_{in} - Q_{out} O^G_{out} - OTR \times W \qquad \text{Equation 5.25}$$

Where:

V_G	gas volume inside the reactor;
$O_{in}{}^G$	oxygen concentration in the inlet gas;
$O_{out}{}^G$	oxygen concentration in the outlet gas;
OTR	oxygen transfer rate;
W	liquid weight.

The oxygen transfer rate can be expressed as:

$$OTR = K_L a^{O_2} (O_{eq} - O) \qquad \text{Equation 5.26}$$

Where:

$K_L a^{O2}$	oxygen mass transfer coefficient from the gas phase to the liquid phase;
O	liquid phase concentration of oxygen;
O_{eq}	liquid phase concentration of oxygen in equilibrium with $O^G{}_{out.}$

The *OTR* is a very important variable, as it is often a limiting factor, and the bioreactor has to provide good mixing characteristic to reach a good gas–liquid transfer. Usually, the maximum feeding rate for the energy source is limited by the oxygen transfer capacity of the bioreactor. By increasing the oxygen transfer capacity of the bioreactor it is possible to achieve higher cell productivity and final biomass concentration.

According to Equation 5.26, the *OTR* can be increased by increasing both the values of $K_L a$ or O_{eq}. The value of $K_L a$ is usually controlled by the rate of agitation and the gas flow rate but, according to some authors (Alba and Calvo, 2000), high values of those variables may limit growth through cell damage caused by the shear stress acting on the cell due to fluid dynamics forces. However, according to flow cytometry results, in Hewitt et al. (1998), the authors concluded that over a wide range of aeration rates and power inputs employed and with dissolved oxygen levels down to just above the critical value, the physiology of *E. coli* does not change.

It is also possible to increase the driving force of oxygen transfer by increasing the pressure in the bioreactor. However, this method has the drawback that the partial pressure of carbon dioxide will be increased as well, which might have the undesired effects described before.

Another possibility to control the oxygen transfer rate is to increase the partial pressure of oxygen in the air supply, e. g. with oxygen enriched air, increasing the value of O_{eq}.

However, when oxygen enriched air is used in order to achieve high feed rate, the impact of high oxygen concentration needs to be investigated. As it was already mentioned, oxygen itself is potentially toxic also for aerobic microorganisms such as *E. coli*. Furthermore, the impact of dissolved oxygen concentration on the production of recombinant proteins is ambiguous in the literature. In Blight and Holland (1994) the authors found that in *E. coli* high-cell density cultures, exponential growth could proceed for a longer time and higher growth rates could be maintained with oxygen enriched air supply. However, a higher specific oxygen consumption rate per glucose was measured after the start of the oxygen enrichment, indicating higher maintenance. The total amount of product was decreased in 50%. Another disadvantage of employing oxygen enriched air is its high price, which increases significantly the costs of a fermentation run.

Another approach to overcome the limitations of growth due to oxygen transfer capabilities is to use metabolic engineering tools to enhanced growth under oxygen-limiting conditions (Hewitt et al*.,* 1998 and Jacobsen and Khosla, 1998).

For the calculation of the *OTR*, usually Equation 5.26 is not very adequate, as the value of $K_L a$, besides its dependence on the aeration rates and on the stirrer speed, also depends on

the viscosity of the culture and on the characteristics of the agitator. Thus, a more expedite way to calculate this variable comes from the gas mass balances.

Using the ideal gas law and if the temperatures and pressures can be considered equal in inlet and outlet, Equation 5.25 can be written in terms of moles as follows:

$$\frac{PV_G}{RT}\frac{dy_{o,out}}{dt} = G_{in}y_{O,in} - G_{out}y_{O,out} - \frac{OTR \times W}{M_{O_2}} \qquad \text{Equation 5.27}$$

Where:

P	total pressure of the gas phase;
R	ideal gas constant;
T	temperature;
$y_{O,out}$	gas phase molar fraction of oxygen in the outflow;
$y_{O,in}$	gas phase molar fraction of oxygen in the inflow;
G_{in}	molar gas inflow rate;
G_{out}	molar gas outflow rate;
M_{O2}	molar mass of oxygen.

Using the same equation for the balance of the inert gas nitrogen, and considering that it is not consumed nor produced by the cells and that the accumulation term is negligible, the molar outflow rate can be calculated from the molar inflow rate as follows:

$$G_{out} = G_{in}\frac{y_{N,in}}{y_{N,out}} \qquad \text{Equation 5.28}$$

Where:

$y_{N,in}$	gas phase molar fraction of nitrogen in the inflow;
$y_{N,out}$	gas phase molar fraction of nitrogen in the outflow.

When the reaction rates do not change dramatically with time, a steady state may be assumed for the gas phase and thus the accumulation term in Equation 5.27 can be considered negligible. Then, combining Equation 5.27 with Equation 5.28, the oxygen transfer rate can be calculated as follows:

$$OTR = \frac{G_{in} \times M_{O_2}}{W}\left(y_{O,in} - y_{O,out}\frac{y_{N,in}}{y_{N,out}}\right) \qquad \text{Equation 5.29}$$

This equation, together with the knowledge of the gas inflow rate, inflow and outflow concentrations allows the calculation of the oxygen transfer rate.

The liquid phase mass balance of dissolved oxygen can be obtained using an approach similar to the one found in Equation 5.3:

$$\frac{dO}{dt} = -OUR + DO_{in} + OTR - DO$$ Equation 5.30

Where:

O_{in} concentration of oxygen in the liquid inflow.

In fed-batch fermentation, the term DO_{in} can be considered negligible because usually the medium is autoclaved before the fermentation, causing the dissolved gases concentration to drop to values close to zero.

Also, since the solubility of oxygen in fermentation media is low, the driving force for oxygen transfer is also low and thus, the transfer is a slow process, compared with oxygen consumption. Simple calculations show that, during the main phase of a bacterial or yeast cultivation process, all the oxygen that can be dissolved in the medium is consumed by the cells within a few seconds. Thus, the dissolved oxygen concentration in the liquid, *O*, is observed to be constant ($dO/dt \approx 0$) (Heinzle and Dunn, 1990 and Lübert and Jorgensen, 2001) and the oxygen uptake rate matches the rate of oxygen transfer.

5.1.4.2 Carbon Dioxide

In an analogous form as in oxygen, the carbon dioxide transfer rate can be calculated from the balance equation to the gaseous phase (Heinzle and Dunn, 1990 and Royce, 1992):

$$CTR = \frac{G_{in} \times M_{CO_2}}{W}\left(y_{C,out}\frac{y_{N,in}}{y_{N,out}} - y_{C,in} \right)$$ Equation 5.31

Where:

M_{CO2} molar mass of carbon dioxide;
$y_{C,out}$ gas phase molar fraction of carbon dioxide in the outflow;
$y_{C,in}$ gas phase molar fraction of carbon dioxide in the inflow.

However, regarding the liquid phase, the situation for carbon dioxide is more complex than for oxygen, because that gas is 5 times more soluble than oxygen. That solubility is enhanced by its reaction with water to form bicarbonate ions (Royce, 1992):

Equation 5.32

$$\begin{array}{c} \text{cells} \\ \downarrow \quad \text{CER} \\ CO_2(g) \underset{K_L a^{CO_2}}{\overset{K_L a^{CO_2}}{\rightleftharpoons}} CO_2(l) \rightleftharpoons H_2CO_3 \overset{K_{acid}}{\rightleftharpoons} H^+ + HCO_3^- \\ \xleftarrow[\text{CTR}]{} \end{array}$$

Where:

K_{acid} carbonic acid dissociation constant.

Further dissociation of bicarbonate to form carbonate ions is considered negligible for pH around 7. The hydration equilibrium is very slow, whereas the bicarbonate forming equilibrium is almost instantaneous. Hence, the conversion between carbon dioxide and bicarbonate ions is described by the overall reaction:

$$CO_2(l) \rightleftharpoons H^+ + HCO_3^- \qquad \text{Equation 5.33}$$

The global equilibrium constant is:

$$K = \frac{B \times H^+}{C} \qquad \text{Equation 5.34}$$

Where B, C and H^+ mean the liquid phase concentrations of bicarbonate, carbon dioxide and protons.

Only at low pH values, when the concentration of bicarbonate may be negligible, can the equilibrium liquid phase reactions be neglected. However, the concentration of bicarbonate ions increases with increasing pH, being equal to that of dissolved carbon dioxide at a pH of about 6.3 and an order of magnitude greater at pH 7.3. At pH 7, the bicarbonate concentration is around 5 times that of carbon dioxide.

Hence, the balance of carbon dioxide in the liquid phase can be calculated from the following equation (Heinzle, 1987; Royce et al., 1988; Heinzle and Dunn, 1990 and Royce, 1992):

$$\frac{d(C+B)}{dt} = CER + DC_{in} - CTR - DC \qquad \text{Equation 5.35}$$

Where:

C_{in} concentration of carbon dioxide in the liquid inflow.

The term DC_{in} can, as in the case of oxygen, be considered negligible. However, for pH>6.5, the total concentration of carbon dioxide in solution, in the form of both carbon dioxide and bicarbonate ions, can be one or two orders of magnitude greater than that of oxygen, and

hence the differential terms in Equation 5.35 are not necessarily negligible. This poses a difficulty when trying to calculate the *CER* from the *CTR* values, as it is no longer equivalent as in the case of oxygen. This is particularly important for the calculation of the respiratory quotient (Yegneswaran et al., 1990), an important control variable in certain microorganisms like *S. cerevisiae*. In Neeleman et al. (2000) and Neeleman (2001), the author developed a Kalman filter for the estimation of *CER* from the values of *CTR*.

The transfer of carbon dioxide across the gas-liquid interface is a purely physical process as for oxygen, which is liquid-film limited:

$$CTR = K_L a^{CO_2} (C - C_{eq})$$ Equation 5.36

Where:

$K_L a^{CO2}$ carbon dioxide mass transfer coefficient from the liquid to the gas phase;
C_{eq} liquid phase concentration of carbon dioxide in equilibrium with the gas phase.

Considering well-mixed gas and liquid phases, the equilibrium concentration can be calculated:

$$C_{eq} = \frac{p_{CO_2,out}}{H^{CO_2}} = \frac{y_{C,out} P}{H^{CO_2}}$$ Equation 5.37

Where:

$p_{CO2,out}$ partial pressure of carbon dioxide in the gas phase;
H^{CO2} Henry constant for carbon dioxide.

The values of the Henry constant can be found in any physical-chemical handbook. In Royce and Thornhill (1991), there is an equation for the dependence of the Henry constant of carbon dioxide with temperature. However, it should be noticed that the solubility of carbon dioxide in the culture medium can be influenced by the presence of non-polar (sugars, ethanol, etc.) and ionic species (Jones and Greenfield, 1982 and Frick and Junker, 1999).

Since carbon dioxide and bicarbonate concentrations are not usually measured on-line, estimates have to be derived from gas analysis and pH measurement. The assumption of equilibrium conditions between gas and liquid phases provides an estimate of the concentration of bicarbonate, by combining Equation 5.34 with Equation 5.37:

$$B = \frac{K \times C}{H^+} = \frac{K \times y_{C,out} \times P}{H^{CO_2} \times H^+}$$ Equation 5.38

Thus, the concentration of carbon dioxide and bicarbonate becomes:

$$C + B = \frac{y_{C,out} P}{H^{CO_2}} \left(1 + \frac{K}{10^{-pH}} \right)$$ Equation 5.39

However, the phase equilibrium assumption is sometimes far from reality, especially in the case of high-viscosity broths (Dahod, 1993). Under non-equilibrium conditions, the concentrations of carbon dioxide and bicarbonate can be calculated as follows from Equation 5.36:

$$C = \frac{CTR}{K_L a^{CO_2}} + C_{eq}$$ Equation 5.40

That gives:

$$C + B = \left(\frac{CTR}{K_L a^{CO_2}} + \frac{y_{C,out} P}{H^{CO_2}} \right) \left(1 + \frac{K}{10^{-pH}} \right)$$ Equation 5.41

The mass transfer coefficient for carbon dioxide transfer can be expressed in terms of that for oxygen transfer, which can be calculated on-line during fermentations from an oxygen balance (Royce, 1992). According to Royce and Thornhill (1991), the correlation between both mass transfer coefficients can be expressed as:

$$\frac{K_L a^{CO_2}}{K_L a^{O_2}} = 0.89$$ Equation 5.42

However, some authors found that this relation between both mass transfer coefficients is not constant (Dahod, 1993).

5.1.5 Balance Equations for Fed-Batch *E. coli* Fermentation

As already stated, during the aerobic growth of *E. coli* with glucose as the only added substrate, the microorganism can follow three different metabolic pathways, as described in Chapter 2 (equations 2.1 to 2.3). However, the stoichiometric relations found in those equations do not account for the anabolic needs of the cultures, and are only valid if the substrates would follow the catabolic reactions. Nevertheless, especially under defined medium conditions, the anabolic needs are supplied by the same metabolites as the catabolic reactions. Thus, the real yield coefficients may differ significantly from the stoichiometric ones. Therefore, the mathematical model for the description of the fed-batch growth of *E. coli* should be based on the following reactions:

$$k_1 S + k_5 O \xrightarrow{\mu_1} X + k_8 C$$ Equation 5.43

$$k_2 S + k_6 O \xrightarrow{\mu_2} X + k_9 C + k_3 A$$ Equation 5.44

$$k_4 A + k_7 O \xrightarrow{\mu_3} X + k_{10} C$$ Equation 5.45

The associated dynamical model can be described by the following equations:

$$\frac{dX}{dt} = (\mu_1 + \mu_2 + \mu_3) X - DX$$ Equation 5.46

$$\frac{dS}{dt} = (-k_1\mu_1 - k_2\mu_2) X + \frac{F_{in,S}}{W} S_{in} - DS$$ Equation 5.47

$$\frac{dA}{dt} = (k_3\mu_2 - k_4\mu_3) X - DA$$ Equation 5.48

$$\frac{dO}{dt} = (-k_5\mu_1 - k_6\mu_2 - k_7\mu_3) X + OTR - DO$$ Equation 5.49

$$\frac{dC_T}{dt} = (k_8\mu_1 + k_9\mu_2 + k_{10}\mu_3) X - CTR - DC_T$$ Equation 5.50

In the last equation, the variable C_T designates the total dissolved carbon dioxide, including both forms of carbon dioxide and bicarbonate, as explained in Equation 5.35. The state space dynamical model can then be expressed in the form:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} + \frac{F_{in,S}}{W} \begin{bmatrix} 0 \\ S_{in} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ OTR \\ -CTR \end{bmatrix}$$ Equation 5.51

This model corresponds to the matrix form of the general state space dynamical model of Equation 5.5. It should be noticed that, for oxygen and carbon dioxide balances, the model equations are equivalent to Equation 5.30 and Equation 5.35, taking $OUR=(k_5\mu_1+k_6\mu_2+k_7\mu_3)X$ and $CER=(k_8\mu_1+k_9\mu_2+k_{10}\mu_3)X$. The correspondence between the matrices and the variables of the *E. coli* model are as follows:

$$\xi = \begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix};\ K = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{bmatrix}$$

Equation 5.52

$$F = \begin{bmatrix} 0 \\ \frac{F_{in,S}}{W} S_{in} \\ 0 \\ 0 \\ 0 \end{bmatrix};\ Q = \begin{bmatrix} 0 \\ 0 \\ 0 \\ OTR \\ -CTR \end{bmatrix}$$

Equation 5.53

$$r(\xi, t) = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X$$

Equation 5.54

According to the bottleneck theory, acetate can not be consumed and produced at the same time, and thus growth can occur according to two separate regimens: i) a oxidative-fermentative regimen, corresponding to an acetate production state, and ii) a oxidative regimen, corresponding to an acetate consumption pathway. The model of Equation 5.51 can then be separated in two distinct models:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -k_1 & -k_2 \\ 0 & k_3 \\ -k_5 & -k_6 \\ k_8 & k_9 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} + \frac{F_{in,S}}{W} \begin{bmatrix} 0 \\ S_{in} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ OTR \\ -CTR \end{bmatrix}$$

Equation 5.55

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -k_1 & 0 \\ 0 & -k_4 \\ -k_5 & -k_7 \\ k_8 & k_{10} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} + \frac{F_{in,S}}{W} \begin{bmatrix} 0 \\ S_{in} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ OTR \\ -CTR \end{bmatrix}$$

Equation 5.56

An additional equation that accounts for volume variations is usually also added to the model in order to account for weight (or volume) variations. In the case of the feeding of a single stream containing only one substrate, one could feel tempted to consider that weight variation

is only dependent on the nutrient feeding rate (considering $F_{in,tot}$ = $F_{in,S}$ and $F_{in,S}/W=D$). In fact, some authors (Lee and Ramirez, 1992; Akesson et al., 1999 and Patnaik, 2002) use this assumption in simulating these processes. However, it was verified by other authors and during the course of the present work, that in small-scale and high-cell density reactors, the amount of culture removed or added during sampling, base and acid additions, evaporation and the mass taken from the reactor due to gas exchanges can not be considered negligible. Thus, the more complete equation for calculating weight variations (adapted from Galvanauskas et al., 1998) is:

$$\frac{dW}{dt} = F_{in,tot} = F_{in,S} + F_b + F_a - F_{evp} - F_{gas} - F_{smp} \qquad \text{Equation 5.57}$$

Where:

F_b	liquid mass flow of base solution added to the bioreactor;
F_a	liquid mass flow of acid solution added to the bioreactor;
F_{evp}	liquid mass flow evaporated from the bioreactor;
F_{gas}	liquid mass flow taken from the reactor due to gas transferences;
F_{smp}	liquid mass flow taken from the reactor due to sampling.

Some authors (Cockshott and Bogle, 1999 and Xu et al., 1999) only consider some of these flows. According to the structure of the model derived in the present work, the proposed equations for the simulation of these variables are as follows:

$$F_b = c_b \mu_2 XW; F_a = c_a \mu_3 XW; F_{evp} = c_e; F_{gas} = (CTR - OTR)W \qquad \text{Equation 5.58}$$

Where c_b represents a constant that correlates the amount of alkali necessary to equilibrate the pH due to acetate production, per amount of biomass formed in the corresponding pathway; c_a has the same meaning, but for acid and acetate production; c_e represents the constant evaporation rate.

The structure of the kinetic equations can be expressed by one of the kinetic models described in section 5.1.3. The criteria for the selection of the most adequate equation will be described in more detail in the following sections.

For deriving the full model described in this chapter, the following assumptions were made:

- Yield coefficients are constant, as well as kinetic parameters.
- Glucose is the only limiting substrate, considering that oxygen concentration never drops to critical values.

- There is an oxidative bottleneck at the TCA cycle, that limits the fully oxidation of glucose. Thus, when glucose uptake exceeds a given limit, part of it follows the fermentative pathway, where acetate and carbon dioxide represent the major products. Else, glucose follows a fully oxidative pathway, where carbon dioxide is the main product. The yields on biomass of these two pathways differ significantly.
- When glucose uptake does not make full use of the oxidative pathway, and if acetate exists on the medium, it can be aerobically metabolized to carbon dioxide.
- pH oscillations in the culture only depend on acetate production and consumption.
- The evaporation rate is constant and does not depend on the stirrer speed or any other variable.
- The bioreactor contains two phases (gas-liquid system), i.e. microorganisms are not considered as a separate phase because of their small size and water-like density.
- The reactor contents are considered homogeneous in axial and radial directions.
- The mass transfer between gas phase and liquid phase are explained by the film model.

5.1.6 Model Identification

In the state space dynamical model described by Equation 5.51, there are clearly two different types of parameters: the yield coefficients, k_{ij}, and the kinetic parameters that inherently correlate the specific growth rates μ_j with the state variables according to what was described in section 5.1.3. In order to use the model for the prediction of system behaviour, those parameter values have either to be known or estimated. The most common approach for the estimation of either of this type of parameters is based on a non-systematic way, i. e., some parameters are calculated for example from chemostat data (Fieschko and Ritch, 1986 and Varma and Palsson, 1994), others from batch fermentation (for example in Dubach and Märkl, 1992), while others are obtained from stoichiometric relations or from the literature (some of the model parameters in Rhee and Schügerl, 1998 and Cockshott and Bogle, 1999 were obtained from the literature). In Xu et al. (1999), the parameters were obtained from individual approaches to the state variables.

In Alba and Calvo (2000), the calculation of yield coefficients was based on stoichiometry, while kinetic coefficients were calculated from continuous and batch culture data.

However, the utilization of the parameters from the literature can only be used if the objective of the simulation is to obtain general tendencies or to derive theoretical relationships, because both yield and kinetic coefficients are extremely dependent on the particular strain under study, and also on the cultivation medium used for a particular fermentation. Also, parameters obtained from batch or chemostat data can hardly be applied in the simulation of fed-batch fermentation, as the strains have very different behaviours under different fermentation conditions. Also, when calculating parameters from batch data, the errors associated with this determination are usually very high and, consequently, the results obtained are not accurate.

Thus, a more scientific approach to parameter estimation is needed if a proper simulation is to be conducted.

Every parameter estimation problem can be formulated as an optimisation problem. Hence, a large number of numerical optimisation techniques, developed within the last decades can also be applied for parameter estimation.

The general goal of parameter estimation is normally achieved by comparing measurement data with simulation results according to a given performance criterion. As performance criterion, the common practice is to use the sum of squared errors between measured and simulated data, given the same model input (in fed-batch culture this input is the feeding profile) and the parameter estimation problem becomes one of minimizing a system's output error by a least square estimation. A graphical representation of parameter estimation methods minimizing the output error is shown in Figure 5.3.

Some authors used this performance criterion to compare several sets of parameters, without applying an optimisation routine (Lee and Ramirez, 1992 and Warnes et al., 1995). The performance criterion can also be used to test for a developed model performance on fresh fermentation data (Lübert and Jorgensen, 2001).

A separation between the determination of the yield coefficients from the kinetic coefficients can be very useful and was adopted by Chen (1992) based on a state transformation on the state-space dynamical model. In that way, yield coefficients can be determined by a linear regression, minimizing the sum of the squared errors between experimental and simulated data. An application of optimal experimental design for the calculation of yield coefficients by that method on a fed-batch *E. coli* fermentation can be found in Veloso et al. (2003).

However, when trying to calculate both yield and kinetic coefficients simultaneously, those simple optimisation methods like linear regressions are not applicable, due to the non-linearities found in the state-space dynamical model. Thus, other optimisation methods have

to be employed. The simplex method has been used for estimating some parameters on a structured model of *E. coli* (Tartakovsky et al., 1997) and on an unstructured model of *Alcaligenes eutrophus* (Lee et al., 1997). However, in this last case, the method was not very successful on approximating experimental data.

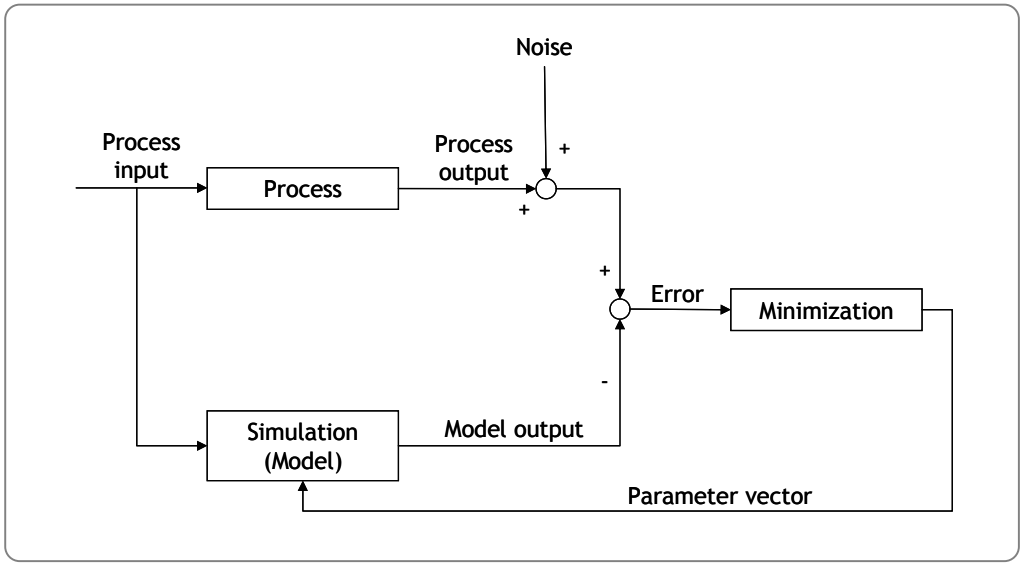


Figure 5.3 Graphical representation of a parameter estimation method (adapted from Schmidt and Isaacs, 1995). The process input is the same for both real and simulated process, and the corresponding outputs are compared by calculating an error value, representing the objective function in an optimisation problem that manipulates the values of the model parameters.

Some authors used other developed optimisation routines for this purpose (Nielsen et al., 1991 and Taylor et al., 1992).

Other optimisation methods that can be applied to this purpose include the ones based on the principles of biological evolution like Genetic Algorithms (GA's), which are direct (i.e. derivative free) optimisation techniques that apply basic elements of biological evolution like crossovers and random changes between elements (chromosomes) of a population to optimise technical systems. A more detailed explanation of the basic features of this optimisation technique is given in Chapter 6. An example for parameter optimisation by an evolutionary strategy in a high order biochemical model is presented by Schmidt and Isaacs (1995).

When parameter estimation is conducted, the sensitivity of the simulation results to model parameters has to be investigated, in order to check for the influence of those parameters' accuracy on the process prediction. However, the majority of the published results related

with parameter estimation does not mention this kind of results, although in some cases that sensitivity is obtained empirically by affecting the parameters separately by a given error and checking the corresponding averaged error on the simulated results (Dantigny et al., 1992; Galvanauskas et al., 1998; Cockshott and Bogle, 1999 and Xu et al., 1999).

Nevertheless, the sensitivity of the model to the parameters generates valuable information not only concerning which parameters and manipulated variables are most important, but also the time periods over which they matter most.

Therefore, sensitivity functions that give a time profile of the model sensitivity represent a more adequate and a more scientific approach. They can be defined as the model output sensitivities with respect to parameter variations, evaluated along the nominal output trajectories (Munack, 1991). One of the relevant applications of sensitivity functions is related with the optimal experimental design of fermentations for parameter estimation (Ejiofor et al., 1994), by using the Fisher information matrix (Schneider and Munack, 1995 and Ayesa et al. 1994).

Calculation of the sensitivity functions can be found in Pertev et al. (1997) and Smets et al. (2002) applied to the bacterium *Azopirillum brasilense* and to *Saccharomyces cerevisiae*, respectively. In these two cases, after sensitivity analysis, a model reduction was conducted by eliminating from the model the least significant parameters or by setting their values to literature data. However, it was not found in the literature an example of an application of this approach to fed-batch fermentation of *E. coli*.

Sensitivity functions can be computed, for the present case, as the sensitivity of the state variables ξ_i to small variations in the model parameters, expressed generically as p_j given the feeding profile input:

$$\frac{d}{dt}\left(\frac{\partial \xi_i}{\partial p_j}\right) = \frac{\partial}{\partial p_j}\left(\frac{d\xi_i}{dt}\right) \qquad \text{Equation 5.59}$$

Where $d\xi_i/dt$ are the differential equations given in Equation 5.51.

The purpose of this chapter is to use the derived model mass-balances structure described on Equation 5.51 in order to develop an integrated approach for kinetic structure derivation and model parameter identification. The method proposed is based on the application of optimisation routines to approximate experimental and simulated results by varying the kinetic structure and the yield and kinetic parameter values. For the better kinetic structure,

yield and kinetic coefficients are identified and then a sensitivity function analysis is conducted to understand which parameters influence the most model simulation results.

5.2 MATERIALS AND METHODS

5.2.1 Cultivation Conditions

Fermentations were operated as described in the previous chapter. For fermentations 1 and 2, the feeding profile was exponential according to the following equation:

$$F(t) = \frac{\mu_{crit}}{k_{XS}} \frac{X(t)W(t)}{S_{in}} \qquad \text{Equation 5.60}$$

The purpose of this feeding profile was to keep the cells growing for a long period at a dilution rate below the critical value and to characterize the oxidative, and eventually fermentative degradation of glucose. A constant value of 0.3 was used for k_{XS}, which represents an approximation for the biomass yield coefficient of the oxidative growth on glucose while the desired specific growth rate was $0.11h^{-1}$ in fermentation 1 and $0.13\ h^{-1}$ in fermentation 2.

The values of weight were acquired and corrected on-line using the bioreactor's balance, while the biomass values were furnished to the algorithm every time a sample was taken. The calculation of the feeding profile was done in a simple subVI that was included in the supervisory LabVIEW VI described on Chapter 4.

For fermentation 3, the feeding rate was varied in a way that allowed the cells to pass through all the metabolic states.

For all fermentations, biomass was determined according to the method described on Chapter 4, while glucose was measured by the DNS method and acetate was measured by HPLC (also described on Chapter 4). These methods were chosen for model identification and validation in detriment of FIA measurements, due to the high level of accuracy required for the purpose of this chapter.

Carbon dioxide and oxygen transfer rates were calculated from gas analysis data obtained using the Mass Spectrometer. During the fermentation, MS acquired samples from both the bioreactor inlet and off-gas, with a duration of 10 minutes for each line. For OTR and CTR computations, a MATLAB (version 6.0) routine (`off_gas.m`, on the CD-ROM embedded to this thesis) was developed that calculates the values of those variables for each sampling

point of the off-gas using an average of the previous in-gas sampling. OTR and CTR were calculated from Equation 5.29 and Equation 5.31.

The desired feeding was executed by the peristaltic pump described on Chapter 4, and the feeding rate was afterwards confirmed with the results acquired from the feeding balance.

5.2.2 Simulations and Optimisation

For simulating the process, a MATLAB routine for the numeric integration of model balance equations was developed using the Runge-Kutta 4th order integration method with fixed stepsize (the stepsize used for the simulations was 0.01 h). For each simulation, the following information is needed: initial biomass, acetate and glucose concentrations; glucose feeding profile; and weight profile obtained during the real experiment.

For parameters estimation, first the real fermentation data were analysed for any erroneous measurement and then saved as text files. A MATLAB routine was built to interpolate this data from different origins into sampling intervals of 0.01 h (program `int_data.m` on the CD-ROM). Afterwards, another MATLAB routine was developed that integrates process simulation with real data analysis (`coli_ident.m`). In a first stage, this program loads the file of fermentation data and identifies the feeding profile as an input for both real and simulated data. Afterwards, the initial values of the variables for the simulation are taken from real data at the initial time. Then, simulation begins by integrating the differential equations and by calling a kinetics function that calculates the specific growth rates as a function of the state variables (`kin.m`). Afterwards, a normalized difference between the real and the simulated data is calculated in order to check for the simulation accuracy. That difference takes the following generic formula:

$$dif = \sum_{i=1}^{n'} \frac{\sum_{j=1}^{p} \left(\frac{\xi_{sim,ij} - \xi_{exp,ij}}{\overline{\xi}_{exp,ij}} \right)^2}{RSD_{\xi_i}} \qquad \text{Equation 5.61}$$

This equation calculates the sum of the quadratic differences between simulated – $\xi_{sim,ij}$ –and real or experimental (interpolated) – $\xi_{exp,ij}$ – data for every point (p_j) of a given state variable within the simulation. For each point, the difference is normalized by dividing it by an average value ($\overline{\xi}_{exp,ij}$) of that state variable, in order to attribute the same importance to all state variable, regardless their magnitude. The global difference obtained for each state variable is

then divided by the corresponding relative standard deviation ($RSD_{\bar{g}}$) in order to confer more significance to the variables that have less errors associated. The value of relative standard deviation was used instead of the standard deviation commonly found in the literature because it was found more appropriate when using a normalized difference. Finally, the individual global differences obtained for each variable are added, being *n'* the number of state variables included in the optimisation goal.

As parameter estimation is considered an optimisation problem, the difference calculated from Equation 5.61 is subjected to a minimization using the Genetic Algorithms toolbox for MATLAB version 1.7 developed by Hartmut Pohlheim at the University of Sheffield, UK. Thus, the routine `coli_ident.m` works as a function that receives a parameter vector as an input and provides the objective function, i. e. the normalized difference between simulated and real data, as an output. On every iteration, GA's test a population of 150 chromosomes, each one containing a number of genes that corresponds to the individual parameters. Thus, every chromosome represents a set of parameters and a potential solution for the optimisation problem. For each chromosome, GA's attribute a single value, corresponding to the objective function calculated for that particular parameter set. The best solution of a given iteration is then stored and all chromosomes are subjected to genetic operations like mutations and crossovers for generating the population of the next iteration. After a given stop criterion is reached, the final solution is stored and represents the best solution for that optimisation. In this work, the optimisation stopped after reaching a pre-determined number of iterations, based on a preliminary approach. Usually 150 iterations were considered enough for optimising a given set of parameters.

In this toolbox, no initial value for the parameters set has to be given, as the algorithm generates a random population for the first iteration. However, upper and lower bounds for the solutions have to be defined and those are common for all the individuals of the chromosomes. Thus, as the individual parameters are expected to assume values of different magnitudes, a normalization of the parameters was conducted in the following way: the entire parameter set varied between 0.33 and 3, and for the simulation each parameter was multiplied by an expected value taken from the literature. This expected value is called initial value throughout this work, although this denomination does not fully correspond to the notion of initial values used in other optimisation methods.

5.3 RESULTS AND DISCUSSION

Model identification was conducted in several distinct steps, in order to separate the different phenomena. Due to the higher confidence on acetate, biomass and glucose measurements, when compared to MS measurements, balance equations to those state-variables were used to identify the corresponding parameters in the first place, thus allowing a more accurate identification of the kinetic structure and parameters. Also, real values of weight were used on the simulations in order to use the correct value for the dilution rate. For this identification, the model described on Equation 5.51 was divided on the sub-models of Equation 5.55 and Equation 5.56. Only then the yield coefficients related with oxygen and carbon dioxide were calculated and, at a final stage, parameters from Equation 5.58 were obtained.

5.3.1 Oxidative-Fermentative Regimen

The methodology used for the estimation of the kinetic structure and parameters related with oxidative-fermentative regimen is shown in Figure 5.4. The objective function in this case was:

$$dif = \frac{\sum_{j=1}^{p}\left(\frac{A_{sim,j}-A_{exp,j}}{\overline{A}_{exp,j}}\right)^2}{RSD_A} + \frac{\sum_{j=1}^{p}\left(\frac{X_{sim,j}-X_{exp,j}}{\overline{X}_{exp,j}}\right)^2}{RSD_X} + \frac{\sum_{j=1}^{p}\left(\frac{S_{sim,j}-S_{exp,j}}{\overline{S}_{exp,j}}\right)^2}{RSD_S} \qquad \text{Equation 5.62}$$

In this equation, the first term is related with the normalized difference obtained for acetate, and will be called *difA* throughout this work, and the same applies for the second (*difX*) and the third (*difS*) terms.

Characterization of the real fermentation is shown in Figure 5.5. It can be seen that biomass grows exponentially during almost all the fermentation. However, acetate production starts at t=20 h, and it continues until the end of the fermentation. An interesting phenomenon is the stabilization of the biomass concentration when acetate reaches approximately 10 $g \cdot kg^{-1}$. This could be due to an inhibition caused by the presence of acetate. However, this inhibition is not visible on glucose uptake, once glucose maintains a residual concentration along the fermentation. The hypothesis of an inhibition on the oxidation of glucose rather on the glucose uptake needs therefore to be tested.

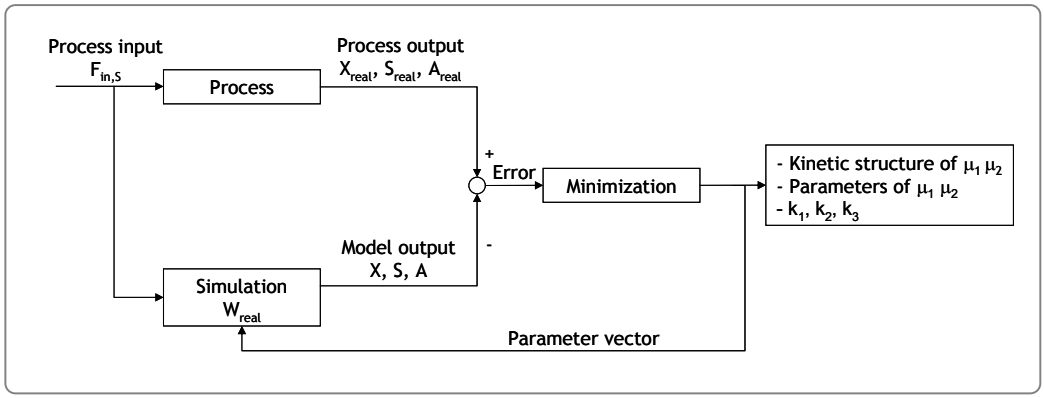


Figure 5.4 Methodology used for the derivation of the best kinetic structure and the yield and kinetic coefficients related to the oxidative-fermentative regimen. The input is common to both simulated and real fermentation, as well as the initial values of the state variables and the weight of the fermentation medium. Simulated and real values are compared according to the difference stated on Equation 5.62, and GA's operate on the parameter set in order to minimize that difference.

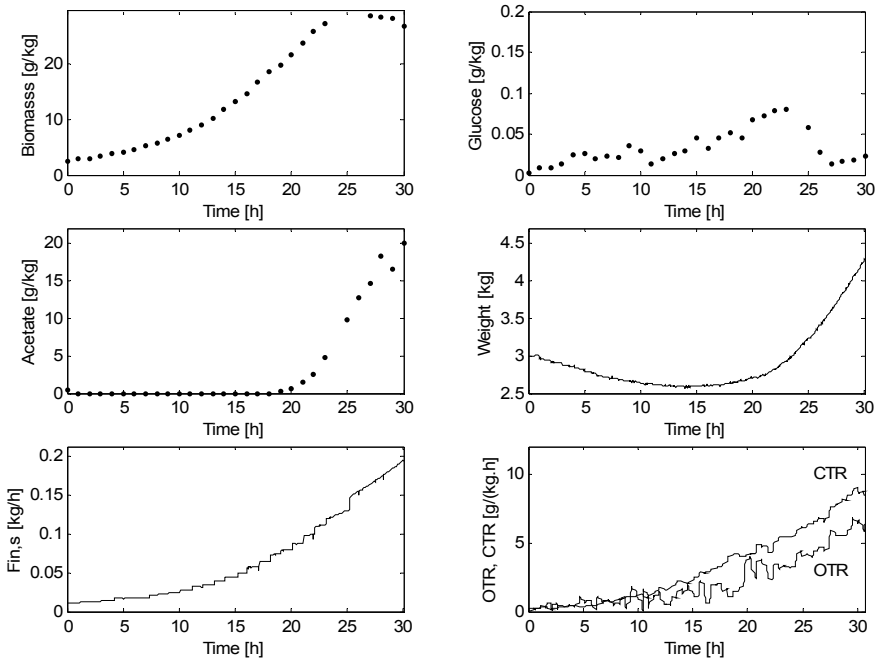


Figure 5.5 Evolution of the biomass, glucose, and acetate concentrations, weight, carbon dioxide and oxygen transfer rates, given the shown feeding profile for fermentation 1.

In order to compare the ability of the several kinetic equations presented on section 5.1.3 to describe the observed phenomenon, several combinations of the equations for glucose uptake, oxidative bottleneck, and maintenance were tested according to the procedure described in Figure 5.4. These combinations, called “cases”, are shown on Table 5.1.

All the combinations shown on Table 5.1 were subjected to an optimisation under the same circumstances, and the results of the simulations obtained with the optimised parameters are shown on Figure 5.6 and Figure 5.7. The normalized differences obtained from Equation 5.62 are shown on Table 5.2. It is clear that acetate inhibition has to be taken into account on the respiratory capacity of the bacteria, although only one reference was found in the literature to that possibility. No other explanation was found in the literature for the stabilization of biomass concentration happening when there is still a high glucose uptake, observed by the low levels of glucose found in the medium at the end of the fermentation. This phenomenon is followed by a continuous increase in acetate concentration, indicating a strong activity of the fermentative pathway.

Table 5.1 Combinations used for checking the best kinetic structure for the oxidative-fermentative metabolism of the strain

Case	Glucose Consumption Rate	Bottleneck	Maintenance
1	$q_S = q_{S,\max}\frac{S}{S+K_S}$	$q_{S,crit} = const$	$q_m = 0$
2		$q_{OS} \le q_{O,\max}\frac{K_{i,O}}{A+K_{i,O}}$	$q_m = 0$
3			$q_m = const$
4	$q_S = q_{S,\max}\left(\frac{S}{S+K_S}\right)\left(1-\frac{A}{K_{i,S}}\right)$	$q_{OS} \le q_{O,\max}\frac{K_{i,O}}{A+K_{i,O}}$	$q_m = 0$
5	$q_S = q_{S,\max}\frac{S}{S+K_S}\frac{K_{i,S}}{K_{i,S}+A}$	$q_{OS} \le q_{O,\max}\frac{K_{i,O}}{A+K_{i,O}}$	$q_m = 0$
6	$q_S = q_{S,\max}\frac{S}{S+K_S+\frac{K_S}{K_{i,S}}A^2}$	$q_{OS} \le q_{O,\max}\frac{K_{i,O}}{A+K_{i,O}}$	$q_m = 0$

The hypothesis that the critical glucose specific uptake rate decreases hyperbolically with the increase in acetate concentration was tested by using an approach similar to the one found in Xu et al. (1999). These authors consider that the bottleneck is related to a limitation on the respiratory chain, which is expressed by a limited oxygen uptake. However, the hypothesis stated on Chapter 2 is related to a limitation on the TCA cycle, and not in the oxygen uptake.

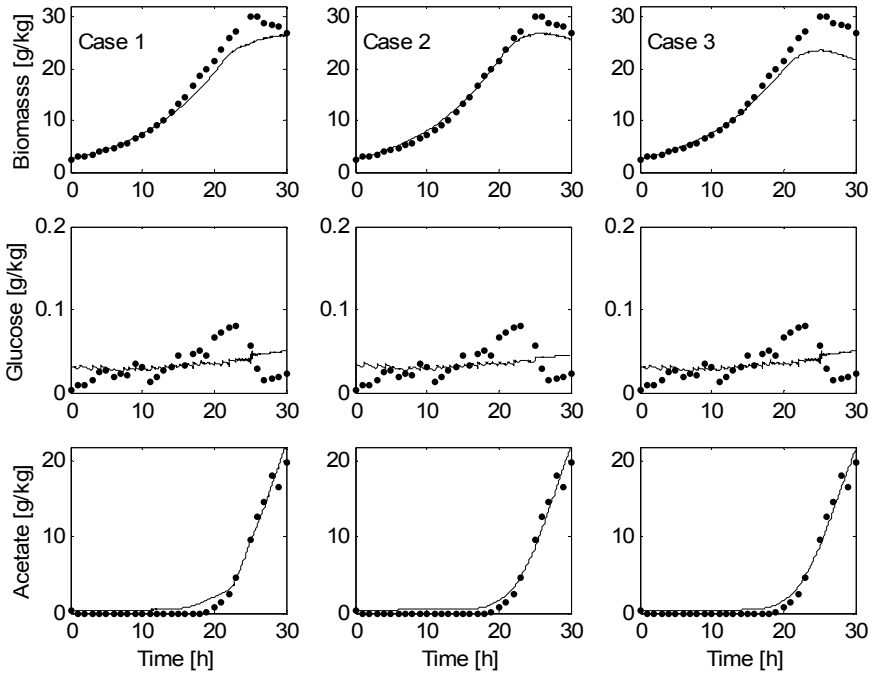


Figure 5.6 Comparison of real data and simulation results obtained after parameter estimation for biomass, glucose, and acetate concentrations for cases 1 to 3, as described on Table 5.1. Real values are represented by circles, while simulated values are the lines.

Nevertheless, equations developed for the former case can be applied to this hypothesis, only with a different meaning: instead of a limitation in the respiration inhibited by acetate concentrations, they can be viewed as a limitation in the TCA cycle also inhibited by that compound. Still, for the sake of simplicity, the nomenclature was maintained for those expressions.

In fact, it can be seen that the stabilization of biomass at the end of the fermentation is better described by using the above-mentioned approach (cases 2-6) than using a constant critical glucose specific uptake rate (case 1).

The introduction of a maintenance term could also describe in part this phenomenon, but the introduction of a constant maintenance term gave very poor results in what biomass approximation respects (case 3).

Addition of two different inhibition terms for acetate on glucose uptake (cases 4 and 5) gave a better approximation for acetate, in case 4, and for biomass in case 5. However, in both cases, there is a final tendency of an increase in glucose concentration, which was not

observed experimentally. Nevertheless, it should be stated that glucose concentration was not very well approximated in any of the tests, probably because experimental data had a significant relative error. Since glucose concentrations were always very low during the fermentation, the reliability of the measurements is poor due to the low reliability of the DNS method at these concentrations.

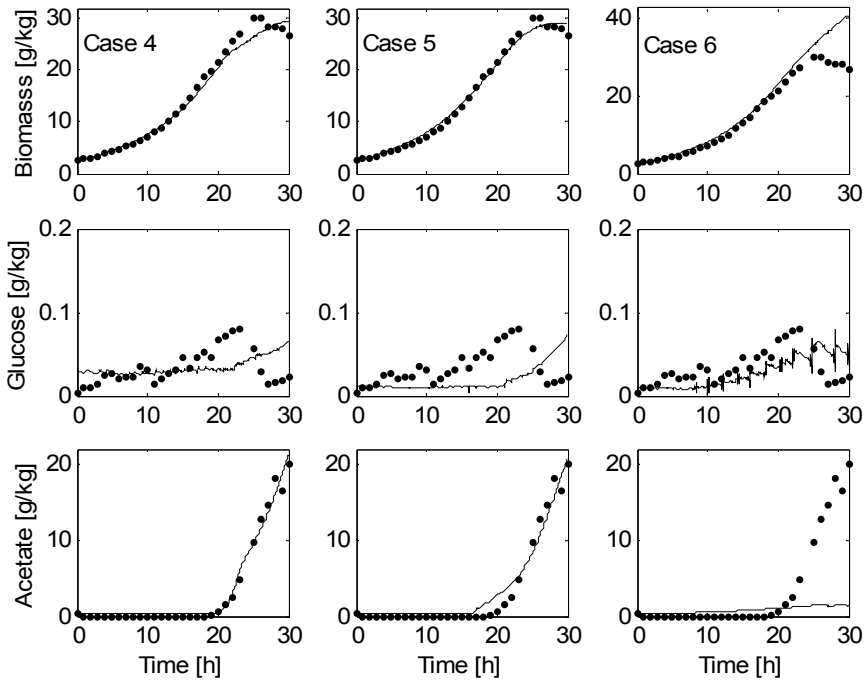


Figure 5.7 Comparison of real data and simulation results obtained after parameter estimation for biomass, glucose, and acetate concentrations for cases 4 to 6, as described on Table 5.1 Real values are represented by circles, while simulated values are the lines.

Finally, the utilization of a different expression for the description of glucose uptake (case 6) gave the worse results, thus indicating that a Monod-type equation, with or without an inhibition term should be used for glucose uptake.

Thus, while there are not many doubts relating the selection of the acetate-inhibited bottleneck mechanism and the absence of maintenance, the selection of the glucose uptake is not as evident, because cases 2, 4 and 5 gave very similar results, indicating that any of those expressions could be used to represent glucose uptake. However, the criterion for the

selection was the least value for the total objective function, *dif*, and clearly case 2 gave the smallest value.

Table 5.2 Values of the objective function (dif) of the optimisation routine and the corresponding partial differences related to the state variables biomass, glucose and acetate concentrations

	difX	**difS**	**difA**	**dif**
Case 1	62.88	238.16	208.16	509.20
Case 2	43.60	43.52	174.10	261.22
Case 3	122.15	108.68	162.35	393.18
Case 4	25.053	165.490	131.30	321.84
Case 5	20.72	93.38	207.59	321.69
Case 6	189.09	110.56	10077.00	10376.75

Figure 5.8 represents the time evolution of the specific uptake rate of glucose and the specific growth rates of the oxidation and fermentation of glucose for the selected kinetic structure of case 2. It can be seen that, due to the exponential feeding rate, glucose uptake presents a stable value during all the fermentation. Also, the initial values for the specific growth rate are very close to the set-point imposed by the feeding profile of Equation 5.60. However, a slight increase in the q_S values around 15-20 hours is sufficient to trigger acetate formation. After acetate accumulation starts, the critical specific glucose uptake rate decreases hyperbolically, as indicated by the decrease in the oxidative specific growth rate, beginning what can be seen as a cyclic phenomenon, where acetate accumulation originates more acetate accumulation, if an exponential feeding is maintained.

It should be mentioned that this fermentation was conducted, as already stated, with an exponential feeding profile, which is very commonly used in a feed-forward approach to grow *E. coli* to high-cell densities with limited acetate accumulation. However, even with on-line correction of the weight values and the introduction of the measured biomass measurements every hour, it was not possible to grow the cells to a high-cell density without acetate accumulation, leading to a reduced global yield on glucose and consequently to a decrease in the productivity of the fermentation. This very slight increase in the specific uptake rate can be due, for example, to an overestimated measurement of biomass, which is very likely to occur due to the dilutions necessary for the optical density measurements in high-cell density fermentations.

Thus, it can be concluded that, at least for this particular strain of *E. coli*, traditional feeding methods based on an exponential profile do not avoid the accumulation of acetate, except if

the imposed specific growth rate is significantly lower that the values that correspond to the maximum oxidative capacity. If this last solution is implemented, it would take much more time to achieve the desired cell density, compromising once again the productivity of the process.

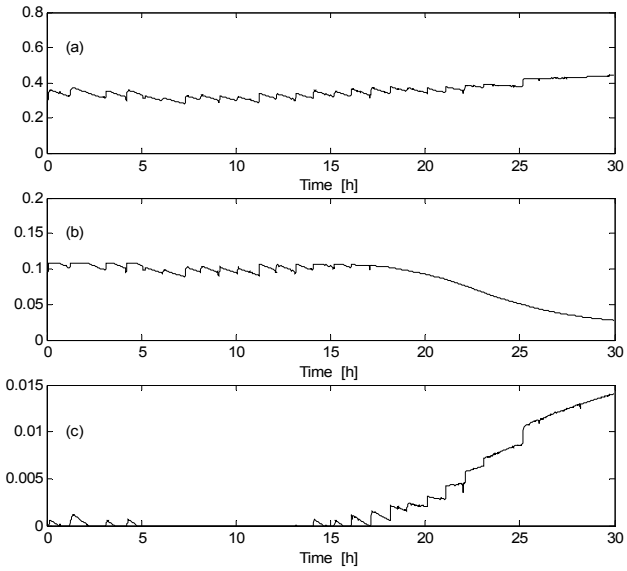


Figure 5.8 Time evolution of the specific rates during fermentation 1. (a) specific glucose uptake rate (q_S), (b) specific growth rate for oxidative glucose consumption (μ_1), and (c) specific growth rate corresponding to the fermentative glucose consumption (μ_2).

The oxidative-fermentative model that resulted from the kinetic structure and parameter estimation was then validated by comparing the real results of a different fermentation (fermentation 2) with the ones simulated using the adopted model, together with real initial values and the real feeding profile. In Figure 5.9 those results are shown, that indicate a good approximation between real and simulated values. However, the approximation is not sufficient, for example, for designing feed-forward control actions.

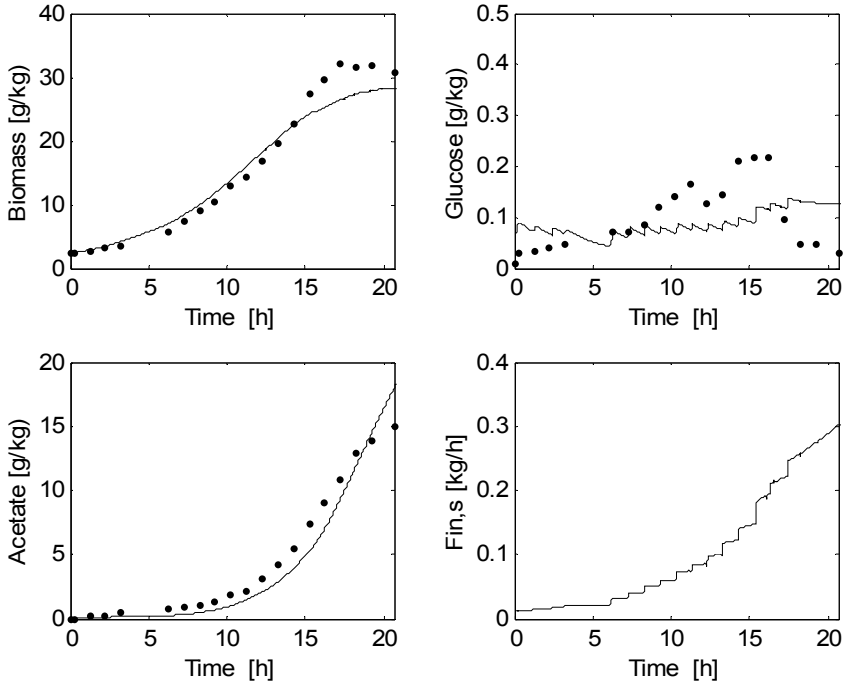


Figure 5.9 Validation of both model structure and identified parameters: comparison between real (circles) and simulated (lines) data for fermentation 2.

5.3.2 Oxidative Regimen

For selecting the best model structure and estimating the model parameters related to acetate consumption, a different fermentation was designed, that made the cells pass throughout all the three metabolic pathways. Then, an approach similar to the one described previously was applied to the experimental data. This approach is shown in Figure 5.10.

The experimental results obtained for this fermentation are shown in Figure 5.11. The increasing feeding profile in the beginning of the fermentation originated a significant acetate accumulation, which was re-consumed by the cells as soon as the feeding profile stopped. However, there is not an increase in the biomass concentration when acetate is being consumed, indicating a very low biomass yield for this metabolic pathway.

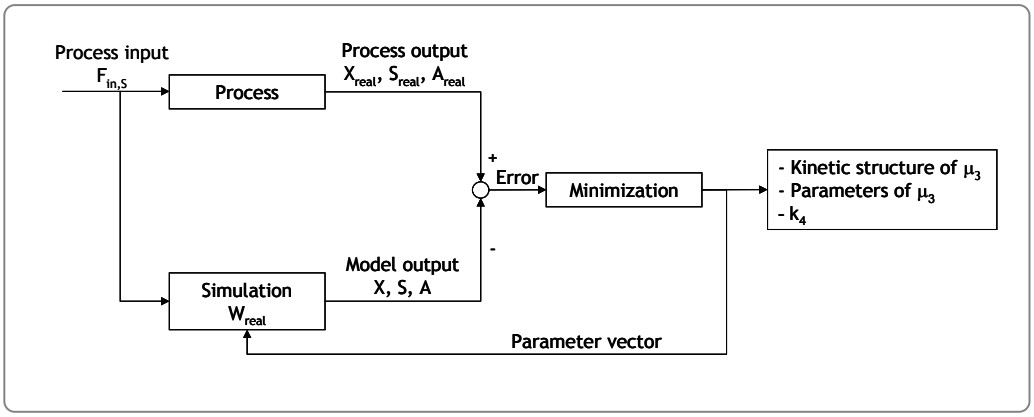


Figure 5.10 Methodology used for the determination of the kinetic structure and the estimation of model parameters for acetate consumption pathway.

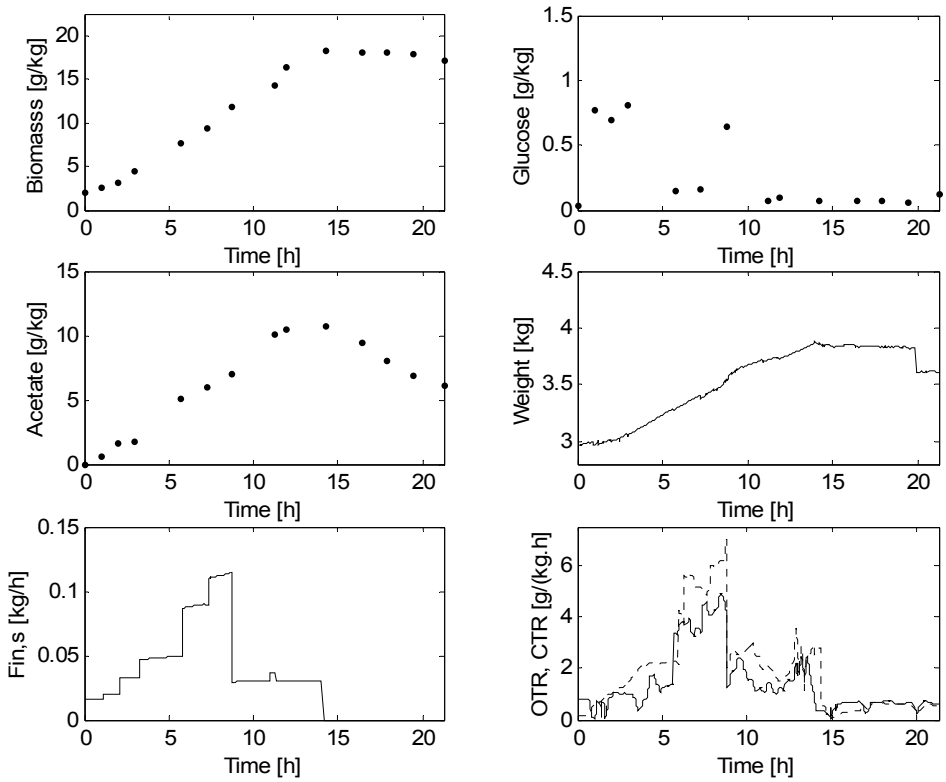


Figure 5.11 Real values of biomass, glucose and acetate concentrations obtained for fermentation 3. Also shown are the broth weight and the feeding profile. The last chart refers to oxygen (full line) and carbon dioxide (dotted line) transfer rates.

The kinetic structure and parameters obtained for the oxidative-fermentative model were applied to this fermentation, together with several alternatives for acetate consumption. The parameters related to this pathway were estimated according to the optimisation method already described. In this case, only few of the literature equations for describing acetate consumption were tested, because many others are not coherent with the bottleneck theory. Thus, acetate consumption is proposed to be consumed according to a specified rate every time it is the only substrate in the medium. Whenever glucose exists in the medium and its corresponding uptake rate does not fulfil the oxidative capacity, the same acetate consumption rate is calculated. Afterwards, the oxidative requirements corresponding to that consumption are calculated and added to the oxidative requirements of glucose. If the sum of those requirements exceeds the total oxidative capacity, acetate consumption is taken to be the maximum that the oxidative capacity can handle. If not, the values calculated for acetate consumption are considered valid. Thus, in terms of kinetic structure, the only variable to be evaluated in this case is the structure of acetate consumption rate. Three alternatives were considered that are shown in Table 5.3.

Table 5.3 Combinations used for checking the best kinetic structure for acetate consumption rate

Case	Acetate Consumption Rate
1	$q_{AC} = q_{AC,\max}\left(\frac{A}{A+K_A}\right)$
2	$q_{AC} = q_{AC,\max}$
3	$q_{AC} = q_{AC,\max}\frac{A}{A+K_A}\frac{K_{i,A}}{K_{i,A}+A}$

After the optimisation procedure, the best fit for each alternative is shown in Figure 5.12. It can be seen that the approximation for the oxidative-fermentative regimen is not very accurate in what acetate respects, indicating once again the variability of the behaviour of the strain. However, for the phase of acetate consumption (after hour 13, approximately), all the three alternatives give a good approximation. These results are in agreement with the values obtained for the objective function and the partial differences obtained for the state variables shown in Table 5.4. The differences subjected to optimisation respect only to the data points of acetate consumption, and glucose difference was taken from Equation 5.62 because equations for glucose uptake had already been derived. As the differences are so small, any of the presented equations could be chosen to represent acetate consumption. However, as

the criterion defined was the least value of *dif*, it was decided that the equation used in case 3 is more adequate for describing this phenomenon.

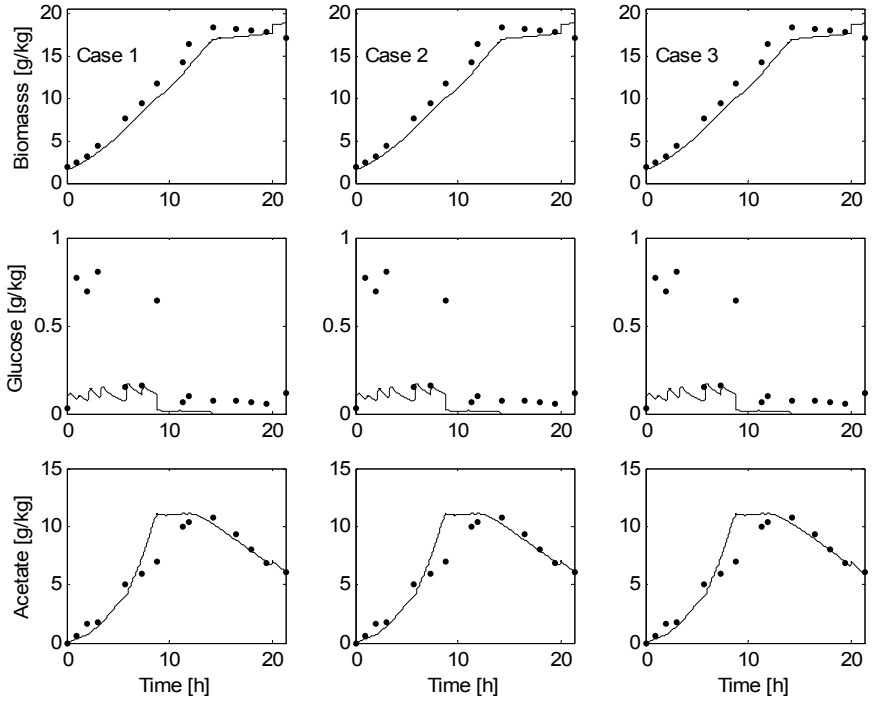


Figure 5.12 Comparison of real data and simulation results obtained after parameter estimation for biomass, glucose, and acetate concentrations for cases 1 to 3, as described on Table 5.3. Real values are represented by circles, while simulated values are the lines.

Table 5.4 Values of the objective function (dif) and partial values for X and A

	difX	**difA**	**dif**
Case 1	3.81	1.34	5.15
Case 2	3.81	1.31	5.12
Case 3	3.87	1.00	4.87

The time profiles obtained for specific glucose uptake rate (q_S) and for acetate consumption rate (q_{AC}) obtained for case 3 are shown in Figure 5.13, while the time evolution of the three specific growth rates is shown in Figure 5.14.

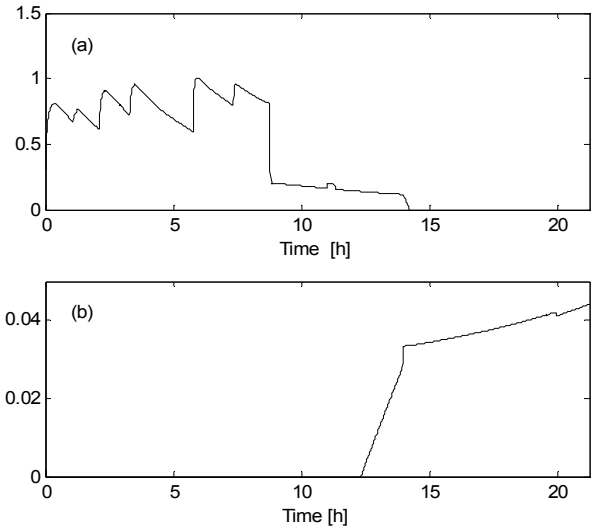


Figure 5.13 Time evolution of acetate and glucose consumption rates. (a) specific glucose uptake rate (q_S); (b) specific acetate uptake rate (q_{AC}).

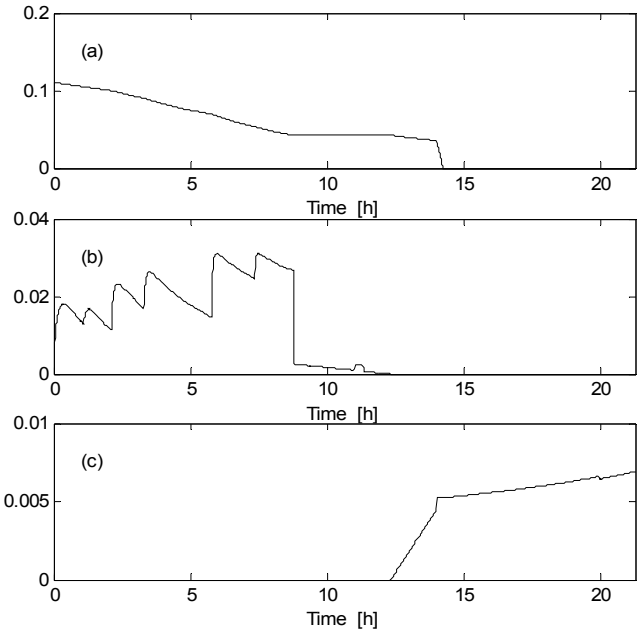


Figure 5.14 Representation of the time evolution of the specific growth rates: (a) μ_1,; (b) μ_2; (c) μ_3.

It can be seen that there is a strong activity of the fermentative pathway until 8 hours, due to the high values obtained for glucose uptake rate during that period. Oxidation of glucose decreases hyperbolically during that period, due to acetate accumulation and the corresponding inhibition of the TCA cycle. When the feeding profile is decreased to low values, provoking a decrease in the glucose uptake rate, the fermentative pathway decreases until reaching a zero value at 12.5 h. At that time, acetate consumption starts to increase, reaching a value close to its maximum when glucose consumption ends.

Thus the final expressions for the kinetic model are now derived, and the final kinetic model is represented in Figure 5.15.

5.3.3 Oxygen and Carbon Dioxide Related Parameters

In order to determine the yield coefficients related to oxygen uptake, it was considered that the dynamics of oxygen can be considered negligible, due to the reasons stated in section 5.1.4.1. Thus, the corresponding yield coefficients were obtained from linear regressions to experimental values of OTR, biomass concentration and specific growth rates, as follows:

$$OTR \approx OUR = \begin{bmatrix} k_5 & k_6 & k_7 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X \qquad \text{Equation 5.63}$$

Two linear regressions were conducted, one to the experimental values of fermentation 1 in order to determine k_5 and k_6, and the other to experimental values of equation 3, in order to determine k_7. Approximation between simulated and experimental values of OTR for fermentations 1 and 3 is shown in the upper part of Figure 5.17 and Figure 5.18. It can be seen that, although experimental data of OTR is corrupted with noise, simulated data approximate it in a satisfactory way.

For the yield coefficients related with carbon dioxide, the consideration made for oxygen is not straightforward, as it was discussed in section 5.1.4.2. Thus, an approach similar to the one used in the previous sections was designed in order to approximate real carbon dioxide and bicarbonate concentrations with simulated data, by manipulating k_8, k_9 and k_{10}. However, in this case, real data was not directly available, as there are no reliable commercial sensors for dissolved carbon dioxide measurements. Thus, several attempts to estimate this variable from other measurements were conducted, based on the expressions derived in section 5.1.4.2, and are shown in Figure 5.16.

$\mathbf{S>0, A>0}$

$$q_S = q_{S,\max}\frac{S}{K_S+S} \Rightarrow q_{OS} = q_S * k_{OS} \begin{cases} q_{OS} \le q_{O,\max}\dfrac{K_{i,O}}{K_{i,O}+A} \Rightarrow q_{AC} = q_{AC,\max}\dfrac{A}{K_A+A}\dfrac{K_{i,A}}{K_{i,A}+A} \Rightarrow q_{OA} = q_{AC} * k_{OA} \begin{cases} q_{OS}+q_{OA} > q_{O,\max}\dfrac{K_{i,O}}{K_{i,O}+A} \Rightarrow q_{AC} = \left(q_{O,\max}\dfrac{K_{i,O}}{K_{i,O}+A} - q_{OS}\right)/k_{OA} \Rightarrow \begin{cases} \mu_1 = q_S/k_1 \\ \mu_2 = 0 \\ \mu_3 = q_{AC}/k_4 \end{cases} \\ q_{OS}+q_{OA} \le q_{O,\max}\dfrac{K_{i,O}}{K_{i,O}+A} \Rightarrow \begin{cases} \mu_1 = q_S/k_1 \\ \mu_2 = 0 \\ \mu_3 = q_{AC}/k_4 \end{cases} \end{cases} \\ q_{OS} > q_{O,\max}\dfrac{K_{i,O}}{K_{i,O}+A} \Rightarrow q_{S,crit} = \dfrac{q_{O,\max}}{k_{OS}}\dfrac{K_{i,O}}{K_{i,O}+A} \begin{cases} \mu_1 = q_{S,crit}/k_1 \\ \mu_2 = (q_S - q_{S,crit})/k_2 \\ \mu_3 = 0 \end{cases} \end{cases}$$

$\mathbf{S>0, A=0}$

$$q_S = q_{S,\max}\frac{S}{K_S+S} \Rightarrow q_{OS} = q_S * k_{OS} \begin{cases} q_{OS} \le q_{O,\max}\dfrac{K_{i,O}}{K_{i,O}+A} \Rightarrow \begin{cases} \mu_1 = q_S/k_1 \\ \mu_2 = 0 \end{cases} \\ q_{OS} > q_{O,\max}\dfrac{K_{i,O}}{K_{i,O}+A} \Rightarrow q_{S,crit} = \dfrac{q_{O,\max}}{k_{OS}}\dfrac{K_{i,O}}{K_{i,O}+A} \begin{cases} \mu_1 = q_{S,crit}/k_1 \\ \mu_2 = (q_S - q_{S,crit})/k_2 \end{cases} \end{cases}$$

$$q_{AC} = 0 \Rightarrow \mu_3 = 0$$

$\mathbf{S=0, A>0}$

$$q_S = 0 \Rightarrow \begin{cases} \mu_1 = 0 \\ \mu_2 = 0 \end{cases}$$

$$q_{AC} = q_{AC,\max}\frac{A}{K_A+A}\frac{K_{i,A}}{K_{i,A}+A} \Rightarrow \mu_3 = q_{AC}/k_4$$

Figure 5.15 Representation of the kinetic model for *E. coli*. Three different cases are presented, related to the presence or absence of the two substrates, glucose and acetate. This model is coherent is the application of the bottleneck theory, with acetate inhibition reflected on the critical oxidative capacity and on acetate uptake rate.

First, only carbon dioxide concentration was calculated based on the assumption of equilibrium with the gas phase – "C" on the figure – (Equation 5.37), and then the sum of carbon dioxide and bicarbonate concentrations were calculated based several assumptions: for the data labelled with "C+B phase eq", the same equilibrium assumption was taken as stated in Equation 5.39, for "C+B KLa ct", carbon dioxide and bicarbonate were calculated from Equation 5.41 by considering a constant value of 89 h^{-1} for K_La^{CO2}, while for "C+B KLa var", the values of K_La^{CO2} were obtained from K_La^{O2} (taken from Equation 5.42 and Equation 5.26).

Considering that this last case is the best approximation obtained to real values, the first conclusion of these results is that bicarbonate concentration is, as predicted, significantly higher than dissolved carbon dioxide concentration. The second conclusion is that an assumption of equilibrium is not very far from reality, and gives a better approximation to real data than an assumption of a constant mass transfer coefficient. An explanation for this is that, during these fermentations, the stirrer speed is, through almost the entire fermentation, the manipulated variable for keeping dissolved oxygen above limiting values, and thus the mass transfer coefficients are not constant.

It was then decided to use the values of dissolved carbon dioxide and bicarbonate under the equilibrium assumption for the optimisation procedure, as the values obtained from the K_La^{CO2} were very noisy due to the propagation of the noise from OTR measurements. However, when trying to approximate these values by simulation using the calculated specific growth rates and experimental CTR, even if a very large range was given for the variation of the parameters k_8, k_9, and k_{10}, it was impossible to obtain an acceptable approximation. This fact can be due to many reasons, but it is more likely that some of the assumptions made for the computation of dissolved carbon dioxide and bicarbonate concentrations are not valid under the circumstances of this work. Another issue that should be investigated is the propagation of the experimental errors in this calculation, for example from pH measurements. However, without a reliable sensor for the on-line measurement of these variables, it will be very difficult to elucidate the best way to represent carbon dioxide and bicarbonate concentrations.

Thus, the only alternative left for the estimation of the yield coefficients related with carbon dioxide production is to make the same consideration as for oxygen, by using the following equation for obtaining the yield coefficients from a linear regression:

$$CTR \approx CER = \begin{bmatrix} k_8 & k_9 & k_{10} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X$$

Equation 5.64

The corresponding approximation between real values of CTR and simulated ones is shown in the lower part of Figure 5.17 and Figure 5.18. Although this approach represents a simplification that, *a priori*, should not be valid, the approximation between real and simulated values was good, and the identified yield coefficients were considered accurate enough for simulating the process.

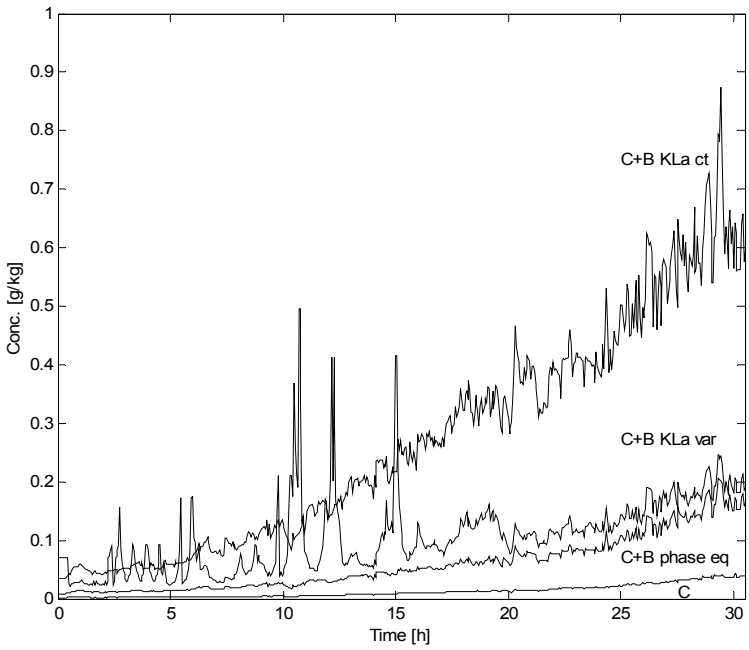


Figure 5.16 Comparison of several attempts to calculate dissolved carbon dioxide and bicarbonate concentrations. The first line from the bottom (C) respects to carbon dioxide calculated under the assumption of an equilibrium with the gas phase; the second line (C+B phase eq) represents the same assumption for the computation of both carbon dioxide and bicarbonate. Third (C+B KLa var) and fourth (C+B KLa ct) lines respect to calculations considering K_La^{CO2} variable and constant, respectively.

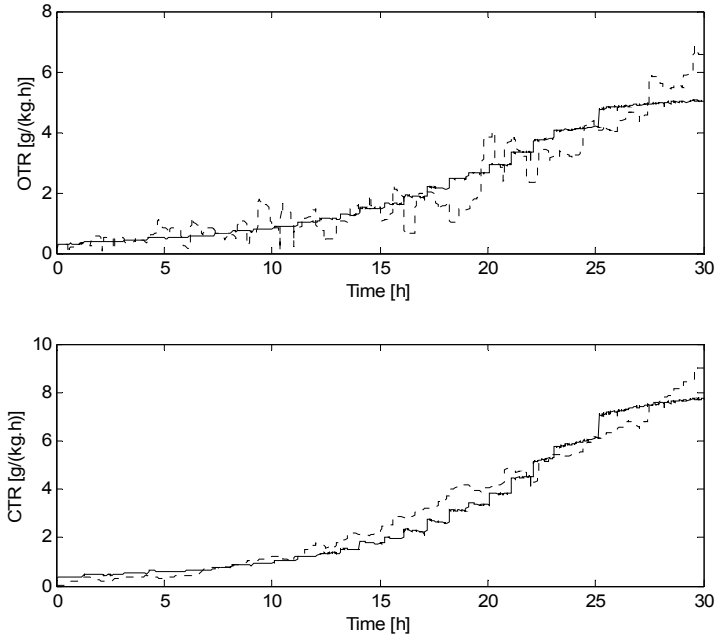


Figure 5.17 OTR and CTR for fermentation 1. Dotted lines represent original values, while full lines represent data simulated with the calculated yield parameters k_8 and k_9.

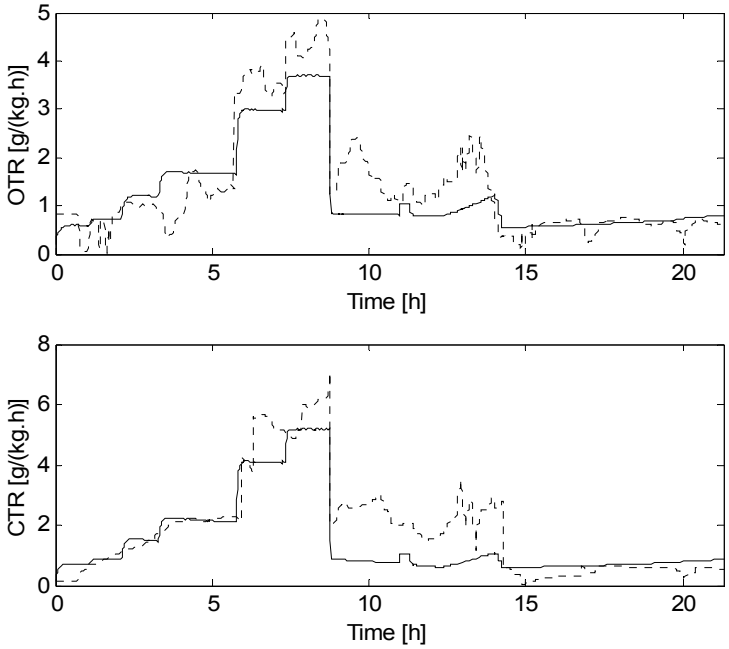


Figure 5.18 OTR and CTR for fermentation 3. Dotted lines represent original values, while full lines represent data simulated with the calculated parameters k_8, k_9 and k_{10}.

5.3.4 Feeding Rate Parameters

The importance of calculating weight variations considering a total feeding rate rather than only the substrate feeding rate was already addressed and an evidence of that for a fed-batch fermentation of *E. coli* can be found in the upper part of Figure 5.20, where real results of the time evolution of the fermenter's weight are compared with values simulated using the glucose feeding rate as the only source of weight variation. It is evident that these discrepancies can cause a significant error when integrating the model dynamic equations, because the dilution rate is present in the balances to all state variables. Thus, for simulation purposes, the expression stated in Equation 5.57 and Equation 5.58 must be used in order to obtain a good estimation of the reactor weight. However, the unknown parameters c_a, c_b, and c_e must be determined before those equations can be applied in a simulation. For this purpose, an optimisation approach was conducted, similar to the ones found in sections 5.3.1 and 5.3.2. In this method, represented in Figure 5.19, real values and simulated values of weight are compared by calculating the normalized difference. For the simulation, a matrix containing real values of the specific growth rates, biomass concentration, OTR, CTR and sampling volumes was used for the simulation, in order to isolate the influence of acid and base addition and evaporation. The results obtained for the parameters allowed approximation between real and simulated values shown in the lower part of Figure 5.20. Although there is a clear improvement in the approximation between real and simulated values, there is still a small difference between these two evolutions. One reason for that discrepancy could be the failure of the assumption that the evaporation rate is constant along the fermentation, which is probably not true due to differences in the stirrer speed and in other variables that can affect that rate, along the fermentation.

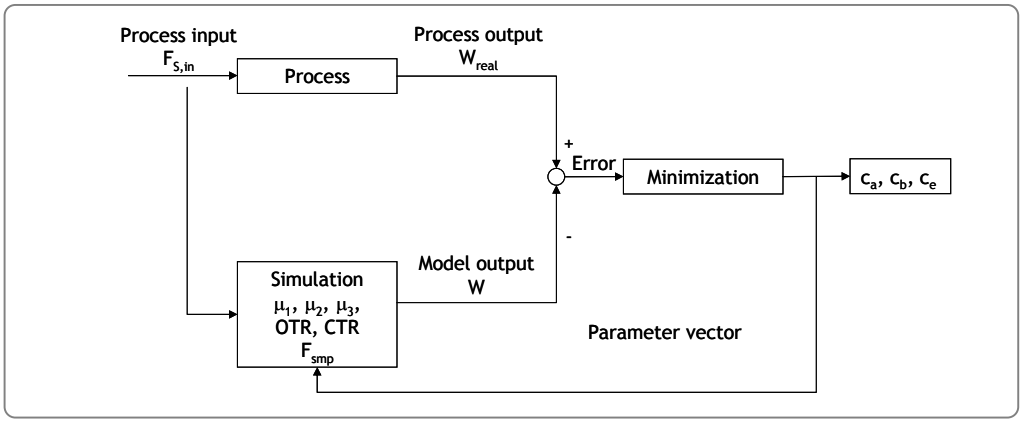


Figure 5.19 Methodology used for the estimation of the parameters related with weight variations. Simulation of the weight time evolution uses real values of OTR, CTR and sampling volumes, as well as calculated values for the specific growth rates.

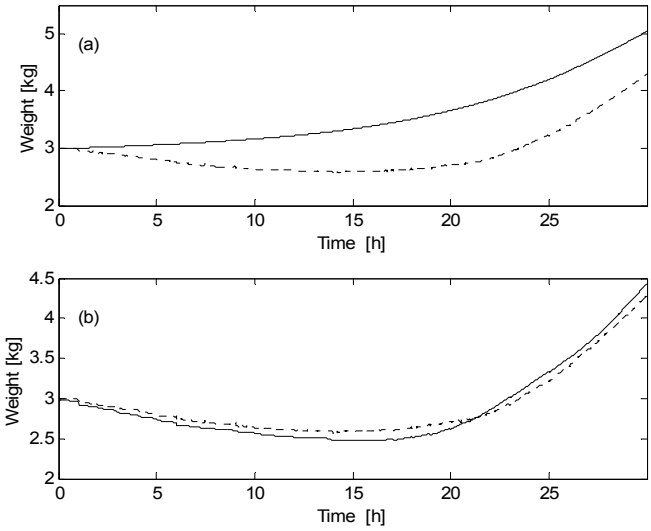


Figure 5.20 Comparison between simulated (full lines) and real (dotted lines) values of the weight profile. In (a), simulated values were calculated by considering that weight variations are only due to the glucose feeding rate, while (b) results from the application of both Equation 5.57 and Equation 5.58 with identified parameters.

5.3.5 Identified Parameters

The values of the 13 identified parameters in the previous sections are shown in Table 5.5. The meaning of initial values was explained in the Materials and Methods section, and only for the parameters obtained from an optimisation using GA's those values were used. Parameters that were calculated by linear regressions do not need an initial approximation.

Table 5.5 Identified values of the model parameters

Parameter	Initial value	Identified value
k_1	2.00	3.164
k_2	20.00	25.22
k_3	14.00	10.90
k_4	4.00	6.382
k_5	-	1.074
k_6	-	11.89
k_7	-	6.098
k_8	-	1.283
k_9	-	19.01
k_{10}	-	6.576
$q_{S,max}$	1.70	1.832
K_S	0.100	0.1428
k_{OS}	1.00	2.020
$q_{O,max}$	0.500	0.7218
$K_{i,O}$	4.00	6.952
$q_{ac,max}$	0.050	0.09670
K_A	1.00	0.5236
k_{OA}	1.00	1.996
$K_{i,a}$	4.00	5.850
c_b	-	0.1034
c_a	-	0.0950
c_e	0.0150	0.0444

5.3.6 Sensitivity Analysis

In order to evaluate the influence of the different parameters on the model outputs, a sensitivity analysis was conducted, based on the application of Equation 5.59. The model outputs, i. e., the state variables evaluated include biomass, glucose and acetate concentration, while 13 model parameters were tested, including all kinetic parameters, and the yield coefficients related with the evaluated state variables (k_1 – k_4). Thus, a total of 3×13 sensitivity functions were evaluated.

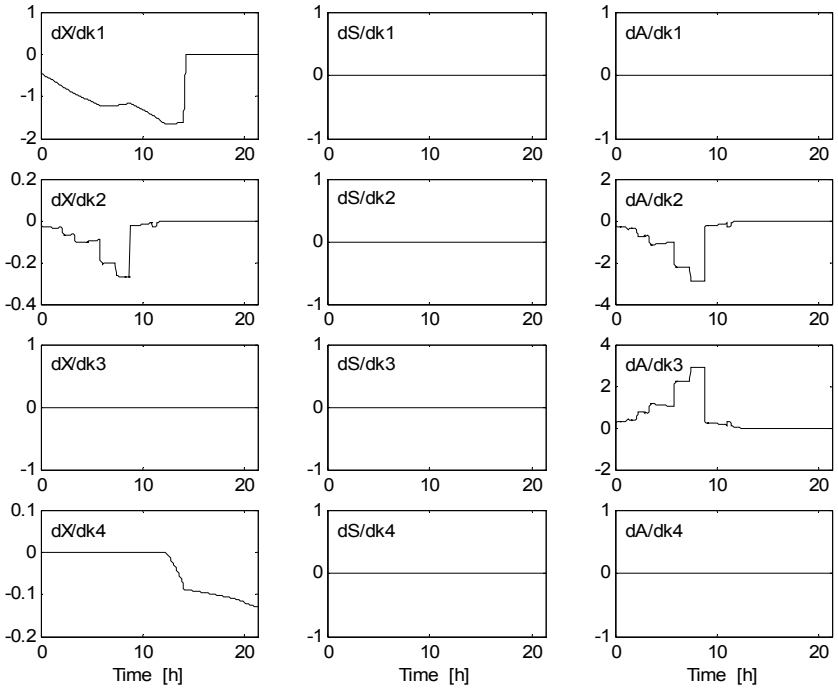


Figure 5.21 Time evolution of the sensitivity functions. In the left column, it is shown the sensitivity of biomass concentration to changes in the yield coefficients k_1 – k_4, while the sensitivity of glucose and acetate are in the middle and right columns, respectively.

For this computation, and in a first step, the model dynamic functions were written as functions of the model parameters, by replacing the adequate expressions for the specific growth rates. Thus, all the three cases represented in Figure 5.15 had to be contemplated for re-writing the model of Equation 5.51. Then, the Symbolic Math Toolbox release 4 for MATLAB was used to derive analytically those expressions regarding all the 13 parameters. The analytical expressions that resulted from this derivation can be found in Annex 3. Finally, these expressions were included in the function that calculates the kinetic expressions (a new function called `kin_sens.m` was obtained by introducing sensitivity functions in the kinetic function `kin.m`) for evaluating the sensitivity in every simulation point. For this purpose, the Euler method for the numerical integration of the model equations was used instead of the Runge-Kutta method, in order to simplify the methodology. The routine used for the simulation is called `coli_euler.m` and can also be found in the CD-ROM. For evaluating the sensitivity of the model outputs to all the 13 coefficients, a simulation of fermentation 3 was chosen, because all the metabolic pathways are represented in that experiment. For that

simulation, real values of the model input (glucose feeding profile) and of the initial values of state variables were used, while the parameter values were the ones stated on Table 5.5.

The results obtained for the sensitivity analysis are shown in Figure 5.21 to Figure 5.23. To allow for a better comparison between all sensitivities, each sensitivity function was rescaled by multiplying it with the parameter value under study.

By analyzing the sensitivity of the model outputs to variations in the yield coefficients, it is clear that these coefficients have no effects on glucose concentration, but have a significant effect on biomass or acetate, with the exception of parameter k_4, that only influences biomass, and in an order of magnitude smaller (10^{-1}) than the average order of magnitude (10^{0}). All other yield coefficients influence significantly at least one of the model outputs. Thus, it can be concluded that, if a literature value for k_4 is used during the simulations, the results obtained for the state variables will not differ much from the ones obtained with the value identified experimentally.

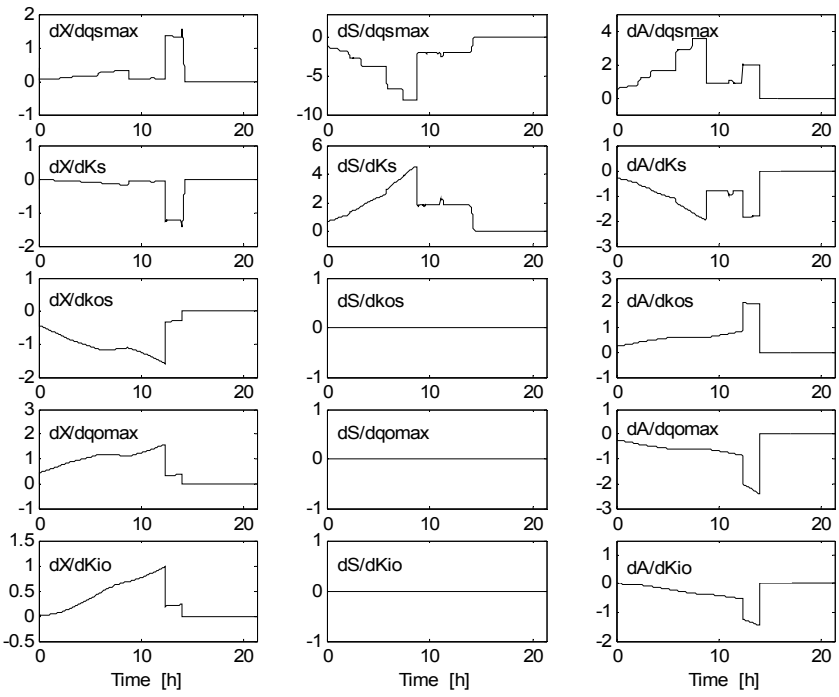


Figure 5.22 Time evolution of the sensitivity of biomass, glucose and acetate to glucose uptake and bottleneck parameters.

For glucose uptake and bottleneck parameters (Figure 5.22), it is clear that all of the parameters influence in the same order of magnitude at least two of the state variables, being the parameters K_S and $q_{S,max}$ the most critical ones. Thus, there must be a special attention in calculating these parameters, as a small variation in their values can be significant in terms of the simulation results for the state variables.

In what the acetate uptake parameters respects (Figure 5.23), it can be seen that the order of magnitude of the sensitivity of model outputs to parameters K_A and $K_{i,A}$ is smaller than the one related to the other parameters. This indicates that both the saturation and inhibition terms of acetate consumption can be ignored, and acetate consumption can be, for the range of acetate concentrations of this fermentation, considered to occur always at its maximum rate. This result agrees well with the conclusions reached on section 5.3.2, where the several options tested to describe acetate consumption rates gave almost the same results.

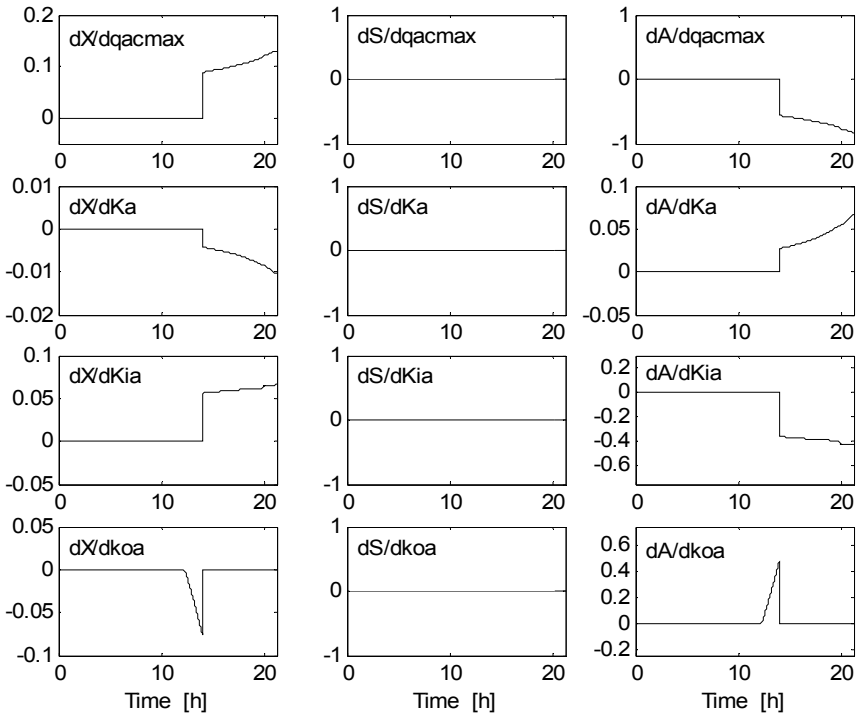


Figure 5.23 Sensitivity of biomass, glucose and acetate concentrations to acetate consumption parameters.

Nevertheless, these results are not sufficient to exclude *a priori* the saturation and inhibition terms from acetate consumption equation in every situations, because acetate concentration during that consumption only varied between 5 and 10 $g \cdot kg^{-1}$ for the fermentation under study. Thus, more experiments must be taken where acetate concentration varies in a wider range in order to conclude about that possibility.

5.4 CONCLUSIONS

From the work described in the previous chapters, the following conclusions can be derived:

- A complete mathematical model for representing growth of the bacterium *E. coli* in fed-batch cultures was developed. The dynamical equations have been derived based on the state-space dynamical model framework described by Bastin and Dochain (1990).
- The kinetic model that describes the oxidative growth on glucose and/or acetate and glucose overflow metabolism was selected from several literature options, being the selection based on an integrated approach, based on the choice of the best model that, after parameter estimation, fits the best the experimental data. The best kinetic model includes an inhibition of the critical glucose specific uptake rate by acetate, which causes a cyclic phenomenon, where acetate accumulation induces more acetate accumulation, if an exponential feeding profile is maintained. This conclusion raised the need of developing new feed-back feeding strategies that would avoid this situation.
- All model parameters were identified by using the Genetic Algorithms optimisation method for approximating real to simulated data for the state variables, by manipulating the parameter vector. This procedure was executed in several steps, in order to separate the different phenomena involved and to isolate the interference of some noisy data, like oxygen and carbon dioxide transfer rates.
- After parameter estimation, the simulations could well describe real biomass, glucose and acetate concentrations, as well as oxygen and carbon dioxide evolution rates. Also, the time profiles for the calculated specific uptake and growth rates are able to describe very well the main phenomena observed during *E. coli* fed-batch fermentations.

- However, when the final model was used to predict a fermentation different from the ones that had originated the model structure and parameters, the approximation between real and simulated values of the state variables was not excellent. This variability has not many possible explanations, because fermentations were conducted in the same exact conditions. These results once again motivate the development of a feed-back type of control with an adaptive capacity in order to cope with these process variabilities.
- Besides yield and kinetic parameters, also the parameters that correlate the total feeding rate with specific growth rates and with oxygen and carbon dioxide transfer rates were estimated. The motivation for this approach is to accurately simulate the variations of weight during the fermentation, which will have a profound effect on the simulation of the other state variables, and that are most often ignored for process simulations. This was only possible due to the existence of a balance that continuously acquires the weight from the fermenter and due to the data acquisition system described in the previous chapter.
- After model selection and parameter identification, a sensitivity analysis was conducted in order to check the influence of parameters changes on the model state variables. It was concluded that the parameters that have the strongest influence on the state variables are the kinetic parameters related with glucose consumption, while the least significant ones are one of the yield coefficients and the parameters related with acetate consumption. It can then extracted that acetate consumption at concentration between 5 and 10 $g \cdot kg^{-1}$ could be described in a simpler way, by considering that it occurs always at the maximum rate.
- The use of the methodology described in this chapter together with some of the tools of experimental design could be an interesting approach for improving the reliability of the parameter estimation procedure described.

5.5 REFERENCES

Akesson, M., Karlsson, E. N., Hagander, P., Axelsson, J. P., and Tocaj, A. On-line detection of acetate formation in *Escherichia coli* cultures using dissolved oxygen responses to feed transients. *Biotechnology and Bioengineering.* 64, 590-598. 1999.

Alba, M. J. G. and Calvo, E. G. Characterization of bioreaction processes: aerobic *Escherichia coli* cultures. *Journal of Biotechnology.* 84, 107-118. 2000.

Andrews, G. F. The yield equations in the modeling and control of bioprocesses. *Biotechnology and Bioengineering.* 42, 549-556. 1993.

Axelsson, J. P., Münch, T., and Sonnleitner, B. Multiple steady states in continuous cultivation of yeast. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 383-386. 1992.

Ayesa, E., Flórez, J., Larrea, L., and García-Heras, J. L. Evaluation of sensitivity and observability of the state vector for system identification and experimental design. *Water Science and Technology.* 28, 209-218. 1994.

Bailey, J. E. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnology Progress.* 14, 8-20. 1998.

Bastin, G. and Dochain, D. *On-line estimation and adaptive control of bioreactors*. Elsevier Science Publishers, Amsterdam. 1990.

Bentley, W. E. and Kompala, D. S. A novel structured kinetic modeling approach for the analysis of plasmid instability in recombinant bacterial cultures. *Biotechnology and Bioengineering.* 33, 49-61. 1989.

Betenbaugh, M. J. and Dhurjati, P. A comparison of mathematical model predictions to experimental measurements for growth and recombinant protein production in induced cultures of *Escherichia coli*. *Biotechnology and Bioengineering.* 36, 124-134. 1990.

Blight, M. A. and Holland, I. B. Heterologous protein secretion and the versatile *Escherichia coli* haemolysin translocator. *Trends in Biotechnology.* 12, 450-455. 1994.

Castan, A., Näsman, A., and Enfors, S.-O. Oxygen enriched air supply in *Escherichia coli* processes: production of biomass and recombinant human growth hormone. *Enzyme and Microbial Technology.* 30, 847-854. 2002.

Chen, L. *Modelling, identifiability and control of complex biotechnological systems*. Université Catholique de Louvain, Belgium. PhD thesis. 1992.

Cockshott, A. R. and Bogle, I. D. L. Modelling a recombinant *E. coli* fermentation producing bovine somatotropin. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 219-222. 1992.

Cockshott, A. R. and Bogle, I. D. L. Modelling the effects of glucose feeding on a recombinant *E. coli* fermentation. *Bioprocess Engineering.* 20, 83-90. 1999.

Dahod, S. K. Dissolved carbon dioxide measurement and its correlation with operating parameters in fermentation processes. *Biotechnology Progress.* 9, 655-660. 1993.

Dantigny, P., Ziouras, K., and Howell, J. A. A structured model of baker's yeast fed-batch growth. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 223-226. 1992.

Dubach, A. and Märkl, H. Application of an extended kalman filter method for monitoring high density cultivation of *Escherichia coli*. *Journal of Fermentation and Bioengineering.* 73, 396-402. 1992.

Ejiofor, A. O., Posten, C., Solomon, B. O., and Deckwer, W.-D. A robust fed-batch feeding strategy for optimal parameter estimation for baker's yeast production. *Bioprocess Engineering.* 11, 135-144. 1994.

Fieschko, J. and Ritch, T. Production of human alpha consensus interferon in recombinant *Escherichia coli*. *Chemical Engineering Communications.* 45, 229-240. 1986.

Frick, R. and Junker, B. Indirect methods for characterization of carbon-dioxide levels in fermentation broth. *Journal of Bioscience and Bioengineering.* 87, 344-351. 1999.

Galvanauskas, V. *Improvement of fed-batch biotechnological processes using advanced methods of experimental design, optimization and control*. Kaunas University of Technology, Lithuania. PhD thesis. 1999.

Galvanauskas, V., Simutis, R., and Lübert, A. Model-based design of biochemical processes: simulation studies and experimental tests. *Biotechnology Letters.* 19, 1043-1047. 1999.

Galvanauskas, V., Simutis, R., Volk, N., and Lübert, A. Model based design of a biochemical cultivation process. *Bioprocess Engineering.* 18, 227-234. 1998.

Harcum, S. W. Structured model to predict intracellular amino acid shortages during recombinant protein overexpression in *E. coli*. *Journal of Biotechnology.* 93, 189-202. 2002.

Heinzle, E. Mass spectrometry for on-line monitoring of biotechnological processes. *Advances in Biochemical Engineering / Biotechnology.* 35, 1-45. 1987.

Heinzle, E. and Dunn, I. J. Methods and instruments in fermentation gas analysis. *Biotechnology - Measuring, Modelling and Control*. 2nd ed. (Karl Schügerl, Eds.). In series: Biotechnology (Rehm, H. J., Reed, G., Pühler, A., and Stadler, P., Eds.) (vol. 4). VCH, Weinheim. 27-72. 1990.

Hewitt, C., Boon, L., McFarlane, C., and Nienow, A. The use of flow cytometry to study the impact of fluid mechanical stress on *Escherichia coli* W3110 during continuous cultivation in an agitated bioreactor. *Biotechnology and Bioengineering.* 59, 613-620. 1998.

Jacobsen, J. and Khosla, C. New directions in metabolic engineering. *Current Opinion in Chemical Biology.* 2, 133-137. 1998.

Jin, S., Ye, K., and Shimizu, K. Efficient Fuzzy control strategies for the application of pH-stat to fed-batch cultivation of genetically engineered *Escherichia coli*. *Journal of Chemical Technology and Biotechnology.* 61, 273-281. 1994.

Jones, R. P. and Greenfield, P. F. Effect of carbon dioxide on yeast growth and fermentation. *Enzyme and Microbial Technology.* 4, 210-223. 1982.

Kleman, G. L. and Strohl, W. R. Developments in high cell density and high productivity microbial fermentation. *Current Opinion in Biotechnology.* 5, 180-186. 1994.

Lee, J. and Ramirez, W. F. Mathematical modelling of induced foreign production by recombinant bacteria. *Biotechnology and Bioengineering.* 39, 635-646. 1992.

Lee, J. H., Hong, J., and Lim, H. C. Experimental optimization of fed-batch culture for poly-β-hidroxybutiric acid production. *Biotechnology and Bioengineering.* 56, 697-705. 1997.

Lee, S. B., Ryu, D. D. Y., Seigel, R., and Park, S. H. Performance of recombinant fermentation and evaluation of gene expression efficiency for gene product in two-stage continuous culture. *Biotechnology and Bioengineering.* 31, 805-820. 1988.

Liu, Y. Model of dissolved organic carbon distribution for substrate-sufficient continuous culture. *Biotechnology and Bioengineering.* 65, 474-479. 1999.

Ljubenova, V. and Ignatova, M. An approach for parameter estimation of biotechnological processes. *Bioprocess Engineering.* 11, 107-113. 1994.

Lubenova, V. N. Stable adaptive algorithm for simultaneous estimation of time-varying parameters and state variables in aerobic bioprocesses. *Bioprocess Engineering.* 21, 219-226. 1999.

Lübert, A. and Jorgensen, S. B. Bioreactor performance: a more scientific approach for practice. *Journal of Biotechnology.* 85, 187-212. 2001.

Luli, G. W. and Strohl, W. R. Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Applied and Environmental Microbiology.* 56, 1004-1011. 1990.

Majewski, R. A. and Domach, M. M. Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnology and Bioengineering.* 35, 732-738. 1990.

Mosrati, R., Nancib, N., and Boudrant, J. Variation and modeling of the probability of plasmid loss as a function of growth rate of plasmid-bearing cells of *Escherichia coli* during continuous cultures. *Biotechnology and Bioengineering.* 41, 395-404. 1993.

Munack, A. Optimization and sampling. *Measurements in Bioreactor Systems*. 2nd ed. (Rehm, H. J., Reed, G., Pühler, A., and Stadler, P., Eds.). In series: Biotechnology (vol. 2). VCH, Weinheim. 252-264. 1991.

Neeleman, R. Respiratory quotient: estimation during batch cultivation in bicarbonate buffered media. *Engineering and Manufacturing for Biotechnology.* 203-216. 2001.

Neeleman, R., van den End, E. J., and Van Boxtel, A. J. B. Estimation of the respiratory quotient in a bicarbonate buffered batch cell cultivation. *Journal of Biotechnology.* 80, 85-95. 2000.

Nielsen, J., Pedersen, A. G., Strudsholm, K., and Viladsen, J. Modelling fermentations with recombinant microorganisms: formulation of a structured model. *Biotechnology and Bioengineering.* 37, 802-808. 1991.

Nielsen, J. and Viladsen, J. Modelling of Microbial Kinetics. *Chemical Engineering Science.* 47, 4225-4270. 1992.

Oliveira, R., Ferreira, E. C., Oliveira, F., and Feyo de Azevedo, S. A study on the convergence of observer-based kinetic estimators in fed-batch fermentations. *Journal of Process Control.* 6, 367-371. 1996.

Onken, U. and Liefke, E. Effect of total and partial pressure (oxygen and carbon dioxide) on aerobic microbial processes. *Advances in Biochemical Engineering / Biotechnology.* 40, 135-169. 1989.

Paalme, T., Elken, R., Kahru, A., Vanatalu, K., and Vilu, R. The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach. *Antoine van Leeuwenhoek.* 71, 217-230. 1997.

Paalme, T., Tiisma, K., Kahru, A., Vanatalu, K., and Vilu, R. Glucose-limited fed-batch cultivation of *Escherichia coli* with computer-controlled fixed growth rate. *Biotechnology and Bioengineering*. 35, 312-319. 1989.

Pan, J. G., Rhee, J. S., and Lebeault, J. M. Physiological constraints in increasing biomass concentration of *Escherichia coli* B in fed-batch culture. *Biotechnology Letters.* 9, 89-94. 1987.

Patnaik, P. R. Analysis of the effect of interruptions in substrate inflow on a fed-batch fermentation with recombinant bacteria. *Biochemical Engineering Journal.* 1, 121-129. 1998a.

Patnaik, P. R. Dispersion-induced behavior in sub-critical operation of a recombinant fed-batch fermentation with run-away plasmids. *Bioprocess Engineering.* 18, 219-226. 1998b.

Patnaik, P. R. Comparative evaluation of batch and fed-batch bioreactors for GAPDH production by recombinant *Escherichia coli* with distributed plasmid copy number. *Chemical Engineering Journal.* 87, 357-366. 2002.

Pertev, C., Türker, M., and Berber, R. Dynamic modeling, sensitivity analysis and parameter estimation of industrial yeast fermenters. *Computers and Chemical Engineering.* 21, 739-744. 1997.

Pomerleau, Y. and Perrier, M. Estimation of multiple specific growth rates in bioprocesses. *AIChE Journal.* 36, 207-215. 1990.

Pomerleau, Y. and Perrier, M. Estimation of multiple specific growth rates: experimental validation. *AIChE Journal.* 38, 1751-1760. 1992.

Rhee, J. I. and Schügerl, K. Mathematical simulation of the growth of a three plasmid harboring *Escherichia coli* JM109 strain and the production of the fusion protein EcoRI::SPA with a four-compartment model. *Bioprocess Engineering.* 19, 261-267. 1998.

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

Royce, P. N., Anderson, M. M., Fish, N. M., and Thornhill, N. F. The use of carbon dioxide evolution rate in consistency checking other fermenter sensor outputs. *Modelling and Control of Biotechnological Processes 1988 (4th Int. Conf. Computer Applications in Fermentation Technology, London, UK).* (Fish, N. M., Fox, R. I., and Thornhill, N. F., Eds.). Elsevier, London. 443-447. 1988.

Royce, P. N. and Thornhill, N. Estimation of dissolved carbon dioxide concentrations in aerobic fermentations. *AIChE Journal.* 37, 1680-1686. 1991.

Schmidt, K. and Isaacs, S. H. An evolutionary algorithm for initial state and parameter estimation in complex biochemical models. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 239-242. 1995.

Schneider, B. and Munack, A. Improvements in the on-line parameter identification of bioprocesses. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 177-182. 1995.

Schügerl, K. Progress in monitoring, modeling and control of bioprocesses during the last 20 years. *Journal of Biotechnology.* 85, 149-173. 2001.

Smets, I., Bernaerts, K., Sun, J., Marchal, K., Vanderleyden, J., and Van Impe, J. F. Sensitivity function-based model reduction. A bacterial gene expression case study. *Biotechnology and Bioengineering.* 80, 195-200. 2002.

Tartakovsky, B., Sheintuch, M., Hilmer, J.-M., and Scheper, T. Modelling of *E. coli* fermentations: comparison of multicompartment and variable structure models. *Bioprocess Engineering.* 16, 323-329. 1997.

Taylor, K. B., Hazelrig, J., Brandt, T., and Li, X. Biomass, nutrient and product models for nutrient-limited batch growth of recombinant *E. coli*. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 335-337. 1992.

Togna, A. P., Fu, J, and Shuler, M. L. Use of a simple mathematical model to predict the behavior of *Escherichia coli* overproducing β-Lactamase within continuous single- and two-stage reactor systems. *Biotechnology and Bioengineering.* 42, 557-570. 1993.

Varma, A., Boesch, B., and Palsson, B. O. Biochemical production capabilities of *Escherichia coli*. *Biotechnology and Bioengineering.* 42, 59-73. 1993.

Varma, A. and Palsson, B. O. Metabolic capabilities of *Escherichia coli*: I. Synthesis of biosynthetic precursors and cofactors. *Journal of Theoretical Biology.* 165, 477-502. 1993a.

Varma, A. and Palsson, B. O. Metabolic capabilities of *Escherichia coli*: II. optimal growth patterns. *Journal of Theoretical Biology.* 165, 503-522. 1993b.

Varma, A. and Palsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Applied and Environmental Microbiology.* 60, 3724-3731. 1994.

Varma, A. and Palsson, B. O. Parametric sensitivity of stoichiometric flux balance models applied to wild-type *Escherichia coli* metabolism. *Biotechnology and Bioengineering.* 45, 69-79. 1995.

Veloso, A. C. A., Rocha, I., and Ferreira, E. C. Identification of yield coefficients in an *E. coli* model - an optimal experimental design using Genetic Algorithms. *9th Symposium on Computer Application in Biotechnology, Nancy, France (submitted).* 2003.

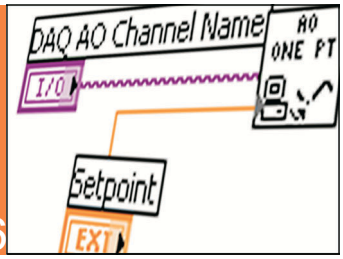
Warnes, M. R., Glassey, J., Montague, G. A., and Kara, B. Comparing different modelling techniques for the *Escherichia coli* fermentation process. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 142-147. 1995.

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

Yegneswaran, P. K., Gray, M. R., and Thompson, B. G. Kinetics of CO_2 hydration in fermentors: pH and pressure effects. *Biotechnology and Bioengineering.* 36, 92-96. 1990.

Zabriskie, D. W. and Arcuri, E. J. Factors influencing productivity of fermentation employing recombinant microorganisms. *Enzyme and Microbial Technology.* 8, 706-717. 1986.

Zeng, F. Y., Dahhou, B., Goma, G., and Nihtilä, M. T. Adaptive estimation and control of the specific growth rate of a nonlinear fermentation process via MRAC method. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 363-365. 1992.



CHAPTER 6
MODEL-BASED OPTIMAL AND ADAPTIVE CONTROL

"I didn't think; I experimented".

Wilhelm Roentgen.

In this chapter, both optimal and adaptive control approaches for operating the fed-batch fermentation of *E. coli* are discussed. For the first one, two optimisation techniques are compared: a first order gradient method and a genetic algorithm. The former method revealed less efficient concerning to the computed maximum, and dependence on good initial values. Thus, with the second one, an optimal feeding profile was derived that maximizes biomass concentration and, consequently, recombinant protein production. Also, initial fed-batch medium weight and concentrations of the state variables were optimized using the same methodology.
For the second case, a model-based adaptive linearizing control law was derived for the regulation of the acetate concentration. A model order reduction method was used to allow the development of the control algorithm without the knowledge of the kinetic structure being necessary. The control law requires on-line acetate and carbon dioxide and oxygen transfer rates measurements. Acetate measurements are achieved with the developed Flow Injection Analysis method. The gas transfer rates are calculated from gas analysis data obtained with the Mass Spectrometer. These calculations, as well as the implementation of the control law were performed through a C script embedded in the developed LabVIEW supervisory program.

6.1 THEORY

6.1.1 Introduction

As already mentioned in the previous chapters, acetate production during growth and recombinant protein production with *E. coli* is the principal pitfall in achieving high-cell densities and high productivities. Among the consequences of that production are the decrease of the biomass yield, the inhibition of the growth when acetate is present at high concentrations and the decrease of the production of recombinant proteins (Riesenberg et al., 1990 and Rothen et al., 1998).

For these reasons, the accurate control of acetate concentration is a very important issue when trying to achieve high productivities with this kind of processes.

Also, the non-linearity exhibited by these processes, coupled with the difficulties in model identification and the time varying properties of the model parameters is often a drawback in the application of model-based control strategies. To deal with these uncertainties, an adaptive approach of the model-based controller can be an attractive alternative.

On the other hand, the importance of operating near an optimal solution is evident due to the high value of the metabolites produced.

The main purposes of this work were to develop both optimal and adaptive control algorithms for improving process automation and performance.

6.1.2 Process Modelling

The model used for the development of the control algorithms is based on the general state-space dynamical model developed by Bastin and Dochain (1990).

$$\frac{d\xi}{dt} = Kr(\xi, t) - D\xi + F + Q \qquad \text{Equation 6.1}$$

That, applied to the fed-batch fermentation of *E. coli* originates the following model, as described in the last chapter:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & -k_6 & -k_7 \\ k_8 & k_9 & k_{10} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ A \\ O \\ C_T \end{bmatrix} + \frac{F_{in,S}}{W} \begin{bmatrix} 0 \\ S_{in} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ OTR \\ -CTR \end{bmatrix}$$

Equation 6.2

The kinetic equations and the values of both kinetic and yield parameters are also shown in chapter 5.

6.1.3 Optimal Control

A primary goal of recombinant fermentation research is to develop better methods for producing these products at high productivities. The total productivity (Q_p) can be defined as the units of product formed per unit of time (Chen et al., 1992; Kleman and Strohl, 1994 and Castan et al., 2002):

$$Q_p = q_p X V$$

Equation 6.3

Where q_p is the specific product formation rate. In order to maximize this factor, both final biomass concentration and specific product formation rate values have to be close to its maximum. If q_p is considered independent on the biomass concentration, then the problem of maximization of the fermentation productivity is transformed in maximization of final biomass concentration.

In fact, this is the basis for the utilization of high-cell density fermentations, where maximal cell densities are achieved and overall volumetric productivities are increased. Nevertheless, it should also be stated that the consideration made above for the independency of the specific product formation rate from biomass concentration is not always true and high productivities obtained in high-cell densities are achieved sometimes at the cost of somewhat lowered specific productivities. Nevertheless, the lack of an adequate model that includes product formation and the separation of the growth phase from the production phase in inducible systems motivate the previous approach.

In recent years, many efforts have been devoted to the optimisation of processes in biotechnology and bioengineering. A problem that has received major attention is the dynamic optimisation (open-loop optimal control) of fed-batch bioreactor. This optimisation has traditionally been done using the substrate feed rate as key manipulated variable in operation.

In the particular case of fed-batch fermentation of recombinant *E. coli*, the optimisation problem may consist on finding an expression or a sequence of values for the feeding rate that maximize the objective function:

$$J(t_f) = X(t_f)W(t_f) \qquad \text{Equation 6.4}$$

Subject to the constraints represented by the state-space dynamical model represented on Equation 6.2.

Several optimisation methods have been applied to solve this kind of problem. It has been shown that, for relatively simple bioreactor systems, which are expressed in differential equations models the optimisation problem can be solved analytically from the Hamiltonian function by applying the Minimum Principle of Pontryagin. However, in the majority of the cases reported, determination of the optimal feed rate profile has a problem of singular control, because the control variable (feed rate) appears linearly in the system of differential equations. Thus, this approach fails to provide a complete solution (Tartakovsky et al., 1995; Shukla and Pushpavanam, 1998 and Banga et al., 2002). For other applications, like the derivation of optimal profiles for pH and temperature, this problem does not exist (Breusegem and Bastin, 1990).

In several studies, singular arc properties were used to solve the analytical control problem. In Lee and Ramirez (1994), the authors used this approach to determine the optimal inducer and nutrient feeding profiles for the production of recombinant proteins in *E. coli*, while Park and Ramirez (1988) used a similar approach for optimisation of recombinant proteins production with yeasts. Modak et al. (1986) studied the general characteristics of the optimal feed rate profiles for various fermentation processes by analyzing singular controls and singular arcs. Another approach to circumvent the problem of singular control was adopted by Lee et al. (1997). These authors used a non-singular approach to determine the optimal feed rate profiles of glucose and ammonium hydroxide for the optimisation of PHB production in *Alcaligenes eutrophus*. With this approach, the problem of determining the optimal feed rates was first converted to that of determining the optimal substrate concentration profiles with the constraints of substrate concentrations. The feed rates were then determined from the nutrient concentrations and their derivatives.

However, those methodologies become too complicated when the number of state and control variables increases and the complexity of the system grow. Moreover, those methods can only be applied to differential equation models and do not satisfy for alternative bioreactor modelling methods based on neural networks and fuzzy models. Here, numerical methods as dynamic optimisation and dynamic programming algorithms are required. In

Banga et al. (2002), there is a small revision of the main numerical methods applied to the optimisation of bioprocesses and a comparison between some of them in terms of robustness and computational effort.

The numerical methods can be divided into deterministic and stochastic methods. Gradient algorithms based on the Hamiltonian function belong to the first category, and can be applied for bioprocess optimisation. In this approach, the control variables are discretized in time. For each moment, the local sensitivities of the objective function for changes in the values of the control variables are calculated. Subsequently, the local sensitivities are used to adjust the control trajectories in order to improve the objective function in an iterative procedure. However, these optimisation techniques may converge to local optima, especially if they are started far away from the global solution.

In contrast to the deterministic methods that find the next sampling point using the problem's features, stochastic search approaches do not use these features, but randomly sample the search space.

For example, dynamic programming methods discretize both time and control variables to a predefined number of values. A systematic backward search method in combination with the simulation of the system model equations is used to find the optimal path through the defined grid. However, in order to achieve a global minimum, the computational effort is very high. Still, there are several examples of the application of dynamic programming to the optimisation of bioprocesses. For example, Tremblay et al. (1991) developed an optimal feeding strategy for fed-batch cultivation of hybridoma cells using dynamic programming that was applied in Tremblay et al. (1993) and compared with a feed-back PI controller. The first one allowed to obtain a higher productivity. Also, Wang and Shyu (1997) applied these techniques to the optimisation of a fed-batch fermentation of *Zymomous mobilis*. Other applications include the development of optimal feeding policies to processes described on a hybrid model that combined balance relation and neural network parameter function model (Tholodur and Ramirez, 1996) and on a structured model with 15 state variables (Waldraff et al., 1997).

Genetic Algorithms (GA's), also in the class of stochastic algorithms, rely on the assumption that random changes in a system's properties will occasionally lead to an improvement of the system. They are based on biological evolution, inspired by Darwin's theory of 'survival of the fittest'. In fact, biological evolution itself can be regarded as an optimisation process where simple reproductive elements are optimised by mutation, crossover, and selection to highly complex beings. As the principles of biological evolution can be simulated on a computer, it is

possible to use this powerful optimisation tool for parameter estimation, as well as process optimisation.

In classical GA's formalism, first a set of candidate solutions (population) to the optimisation problem is generated randomly. The potential solutions are coded as vectors called chromosomes, the elements of which are called genes and are situated in well-defined positions. Goodness of each solution (individual) in the population is evaluated by utilizing a pre-specified fitness criterion, which is then used for selection. The selection is done by a so called roulette wheel selection, where a certain percentage of the individuals is chosen for genetic operations and the chance to be chosen is proportional to the goodness of the individual.

Then, a new generation of individuals is created from the actual population, by using two genetic operators, crossover and mutation. The crossover operator comprises the basic mechanism for redistributing genetic characteristics between the individuals. It is applied to two chromosomes called parents and creates two new individuals. They are generated by selecting one or more points at which the parent strings are 'cut'. The new individuals will have a genetic code in which substrings of genetic codes of both parents alternate. The mutation operator simply introduces the artificial concept of copying error. This operator is simulated by introducing a certain mutation probability according to which a gene is stochastically negated.

This new population generated is then evaluated in the next iteration, and eventually, the population will converge to the global optimum.

Generally, improvement of the chromosomes is mainly obtained by crossover. Mutation is not as important as crossover because it is more likely to produce harmful or destructive changes to individuals.

GA's, as optimisation technique, have been used in several control problems, usually related to fuzzy logic control problems, fuzzy modelling and controller design, and for parameter estimation. However, GA's can be applied virtually to any optimisation problem, as they do not require derivative information or a complete knowledge of the problem structure and parameters. Additionally, GA's consider many points in the search space simultaneously and, therefore, have a reduced chance of converging to local optima.

Nevertheless, premature convergence can be a problem of GA's and other optimisation methods. This happens if convergence occurs too fast and valuable information developed in part of the population is lost. Therefore, for each application, a balance between the

convergence to the global minimum and calculation effort by choosing the population size and selection mechanism has to be established.

Several applications of Genetic Algorithms to bioprocesses can be found in the literature. Integrated methods based on GA's for the maximization of cell mass production in recombinant *S. cerevisiae* and *Aureobasidium pullulans* fed-batch cultures are described in Na et al. (2002) and Ronen et al. (2002). First, the optimisation algorithm is designed and its performance is compared with others. Then, a model parameter estimation algorithm is developed and incorporated into the GA, to have an adaptive optimisation algorithm.

In Moriyama and Shimizu (1996), GA's were applied for on-line determination of optimal temperature profile for ethanol fermentation with *S. cerevisiae*.

Nevertheless, there are not many applications of GA's for the optimisation of the feeding profile in fed-batch fermentation described in the literature. For example, Roubos et al. (1999) compared the performance of GA's with first order gradient algorithm and with dynamic programming. The results show a good and often superior performance of GA's in comparison with other methods. These approaches were applied to hybridoma cells and to recombinant protein production with *E. coli*. A very similar work, only with poorer results, can be found in Nguang et al. (2001).

In Nougués et al. (2002), two optimisation problems were solved with GA's: the model parameters identification; and the optimal feeding profile applied to a chemical fed-batch process.

Angelov and Guthke (1997) present a new approach to optimisation of bioprocesses described by fuzzy rules. It is based on GA's and allows to determine optimal values or profiles of control variables and to optimise fuzzy rules. Also, a hybrid neural network model for the cultivation of *Bacillus thuringiensis* was used for the determination of the optimal feeding profile and for semi-real time optimisation in Zuo and Wu (2000).

A hybrid algorithm of evolutionary optimisation, called hybrid differential evolution, is developed in Chiou and Wang (1999) for the optimisation of the fed-batch fermentation of *Zymomous mobilis.*

The generic scheme for the GA's optimisation process for a fed-batch fermentation can be found in Figure 6.1. In a first step, the feeding trajectory has to be divided in intervals, each one represented by a single gene. Thus, the individual chromosome represents a sequence of feeding rates that constitutes a potential solution for optimising a given objective function. Usually, the feeding rate is considered constant along the interval defined initially. Each one of the individual chromosomes of the population is evaluated for its fitness by simulating the

process and calculating the corresponding objective function. Based on the values of the fitness criteria, the chromosomes are submitted to selection and a new population is formed.

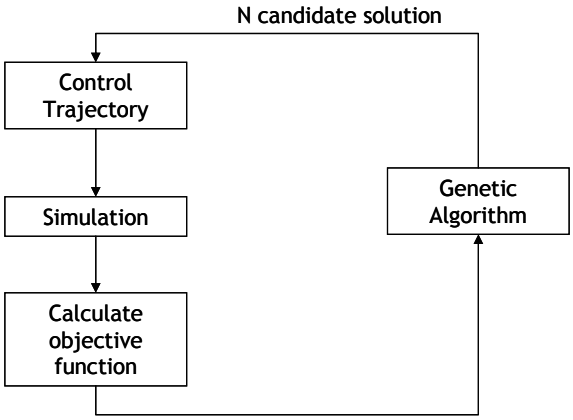


Figure 6.1 Generic scheme of the application of GA's to the optimisation of the fed-batch feeding profile. The given Genetic Algorithm generates a candidate solution composed by N chromosomes that are used to simulate the fermentation using an appropriate model. The fitness of each potential solution is evaluated by calculating the corresponding objective function. Results of this evaluation are then used by the GA's to generate the next population.

6.1.4 Adaptive Control

In order to avoid acetate accumulation, several strategies nave been implemented, either based on proper choices of strain (Luli and Strohl, 1990), carbon source (Tomson et al., 1995 and Aristidou et al., 1999), or other component of the medium (Han et al., 1992). Recently, the tools of metabolic engineering have been applied to reduce the formation of acetate in *E. coli*, either by decreasing the glycolitic fluxes, diverting it to acetoin or by increasing oxygen uptake capabilities (Khosla and Bailey, 1988 and Farmer and Liao, 1996).

Also, in fed-batch cultures, the feed rate can be manipulated to restrict formation of acetate. A number of feeding strategies have been developed and are often exponential strategies, either feed-forward or with some kind of measurement to correct the pre-imposed feeding profile (Chen et al., 1992). However, as seen in the previous chapter, these approaches are not adequate to avoid acetate formation for the particular *E. coli* strain studied in this thesis. Feed-back strategies based on indirect measurements (Konstantinov et al., 1990; Akesson et al., 1999a; Akesson et al., 1999b, and Johnston et al., 2002), have also been described in the literature, and a more detailed explanation of those methodologies can be found in chapter 2.

These indirect methods are mostly used due to the lack of reliable on-line sensors for acetate. Nevertheless, there are some analytical methods that have been used for semi-on-line measurement of this by-product, like HPLC (Shimizu et al., 1988 and Turner et al., 1994). Yet, those methodologies give very high sampling intervals and analysis delays, and are only applicable for giving some indication of the fermentation progress. Only relatively new analytical methods like Near Infrared or FIA, give sampling rates adequate for the implementation of a feed-back closed loop control. However, while the application of the former for the purpose of on-line monitoring of acetate in *E. coli* fermentations was described by Macaloney et al. (1997), application of acetate on-line FIA measurements could not be found in the literature, probably due to the inexistence of commercial methods to analyse this compound with that methodology.

Thus, with the developed FIA system for on-line measurement of acetate described on chapter 3, it is possible to develop a control law with the purpose of generating a feeding profile that regulates acetate concentration in the medium.

For implementing such a controller, several alternatives exist. For example, the use of traditional linear control approaches to bioprocesses can be found in the literature. PID control was used for regulating xylose concentration in *E. coli (*Saucedo et al., 1995), while in O'Connor et al. (1992), the classical PID controller was compared with other methodologies for baker's yeast production, and revealed to be the best method. An inferential PID controller for the specific growth rate was also proposed by Levisauskas (2001).

However, the non-linear behaviour and time-varying properties make bioreactors difficult to control with traditional techniques. These issues are of particular importance when dealing with recombinant *E. coli* high-cell density fermentations with glucose as the carbon source, as the process, besides the non-linearities exhibited, tends to change dramatically upon some events, like induction. Also, some of the inaccuracies in approximating real data with simulated data found in the last chapter may have their origin in the time-variance of both yield and kinetic parameters, which was not contemplated in the developed model.

An alternative cope with these model inaccuracies could be the utilization of fuzzy control schemes, that constitute model-free control systems (Konstantinov and Yoshida, 1992 and Lee et al., 1999) based in some process knowledge.

The model-based alternative to handle with these phenomena is adaptive control, which has been established as a performant substitute to linear control methodologies. Generally speaking, this technique is composed by two feed-back loops, one similar to conventional control techniques, and the other, a non-linear feed-back loop, that adapts the controller parameters according to a learning rule (Figure 6.2). There are several handbooks covering

the subject of adaptive control (Åström and Wittenmark, 1989; Narendra and Annaswamy, 1989 and Sastry and Bodson, 1989), where the several types of adaptive controllers are described in detail.

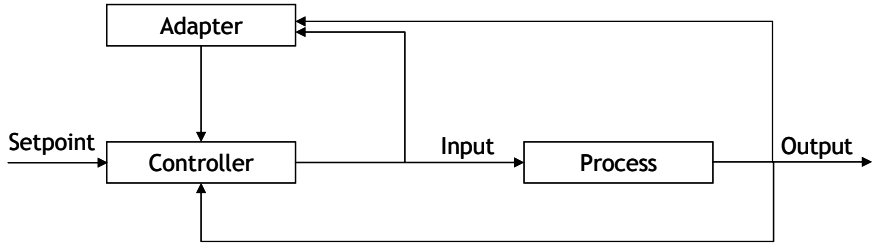


Figure 6.2 Adaptive control scheme.

For biotechnological applications, the main advantage of adaptive control techniques is that they are capable of adapting themselves to variations in process kinetics and parameters. Applications of adaptive control schemes in bioprocesses can be found for example in Montague et al. (1986) applied to penicillin fermentation, and to *S. cerevisiae* in Williams et al. (1986). Other examples include Zeng et al. (1992) for a given class of bioprocesses, Lee et al. (1992) for controlling the specific growth rate in *E. coli*, and Roux et al. (1992) for the control of the dilution rate and substrate concentration in a continuous flow fermentation process for the production of ethanol. The control of the Respiration Quotient (RQ) in *S. cerevisiae* has also been subject of adaptive control implementation (Jorgensen et al., 1992 and Zigova et al., 1999).

Moreover, some developments in non-linear system theory and its application to automatic control have given rise to linearizing control schemes based on differential geometry concepts, improving further the performance of adaptive controllers in bioprocess applications.

The methodology followed for the derivation of the following linearizing control law is based on the work described in Bastin and Dochain (1990). The difference between this technique and conventional control lies in the way that linearization is introduced in the problem. In a standard approach, first a linearized approximation of the model is calculated, and only then a linear controller is designed for this linear model. In contrary, the purpose of this approach is to derive a non-linear controller in order to obtain a dynamic linear behaviour for the controller closed loop. This approach warranties a stable behaviour whatever the operating point on the transient fed-batch trajectory. Figure 6.3 represents the main differences

between this methodology and the conventional approach. This methodology has been initially described for single input – single output stirred reactors. The extension for the multiple input – multiple output (MIMO) case is proposed in Dochain (1991).

Several applications of the linearizing control scheme to the control of fed-batch processes have been described. An approach similar to the one described in these sections can be found in Pomerleau and Viel (1992), where a linearizing adaptive non-linear control law based on ethanol and OTR measurements was applied to industrial baker's yeast production. An extension to the MIMO approach for baker's yeast can be found in Ferreira and Feyo de Azevedo (1996).

Another linearizing adaptive control law was derived and implemented experimentally by Pomerleau et al. (1995) for the control of substrate (methanol) and dissolved oxygen during the production of PHB with *Methylobacterium extorquens*.

In Siegwart et al. (1999), the design and development of a linearizing adaptive controller is presented for regulating glucose in fed-batch HEK-293 cells. Glucose was measured on-line with a FIA biosensor system.

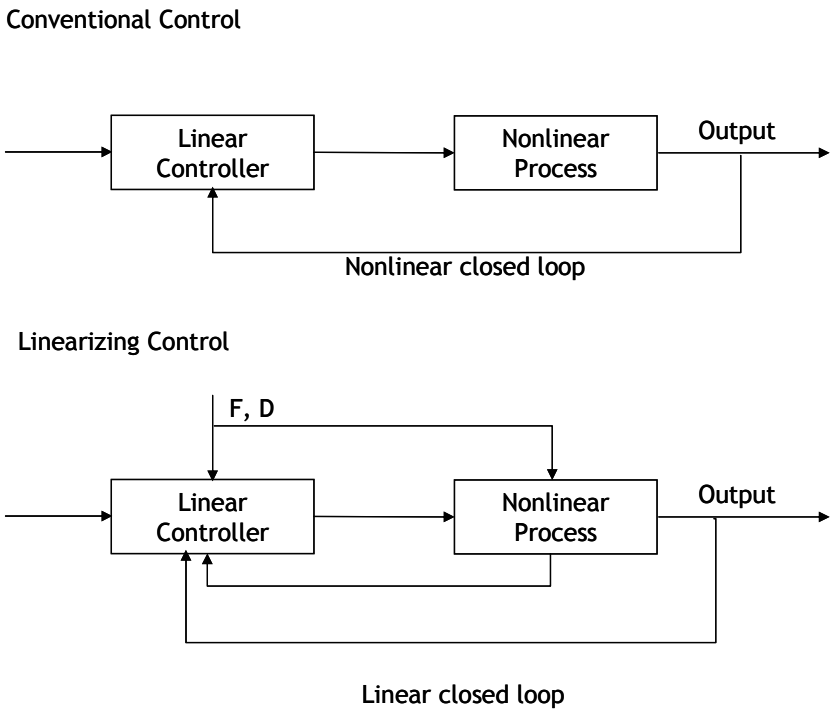


Figure 6.3 Conventional control *versus* linearizing control schemes (Bastin and Dochain, 1990).

A combination of adaptive linearizing schemes with optimal control was proposed by Van Impe (1993). An application of this methodology is described in Claes and Van Impe (1998), where it was adapted to baker's yeast fermentation. In Jadot et al. (1998), an optimal adaptive control law was derived for a type of bioprocesses where there exists a conflict between yield and productivity.

6.1.4.1 Model Order Reduction

The methodology is based on the state-space dynamical model of Equation 6.1 and, for this particular case, on the model described by Equation 6.2.

The objective is to control a scalar output variable, which, in the present case, is one of the state variables, *A(t),* to a reference value, *A*(t),* by manipulating the substrate feeding rate $F_{in,S}(t)$.

The first step of this approach is to derive an input/output model by appropriate manipulations of the state-space dynamical model in order to obtain a linear relationship between *A* and the manipulated variable. This is already the case for the state space dynamical model. However, it is desirable that the control law becomes independent of the specific growth rates, *r*, due to the empirical nature of the equations used to describe those variables. This can be achieved by applying the singular perturbation technique, by considering some components in the process as being "fast" (ξ_f) with respect to the dynamics of the other components ("slow" - ξ_s). Thus, those components are considered to be in pseudo steady-state and the corresponding dynamic term is set to zero.

In order to eliminate the three kinetic terms of Equation 6.2, the number of "fast" components has to be equal to or larger than three. Assuming that glucose, oxygen, and carbon dioxide exhibit fast dynamics relatively to the slow limiting dynamics of biomass and acetate gives:

$$\xi_f = \begin{bmatrix} S \\ O \\ C_T \end{bmatrix};\ \xi_s = \begin{bmatrix} X \\ A \end{bmatrix} \qquad \text{Equation 6.5}$$

It should be noticed that this consideration is coherent with the results presented in the previous chapter. In fact, glucose concentration maintains a value close to zero for all fermentations, while the assumption of steady state for oxygen and carbon dioxide gave a good approximation for OTR and CTR values.

The state-space dynamical model can then be written as:

$$\frac{d\xi_s}{dt} = K_s r - D\xi_s + F_s + Q_s; \qquad \text{Equation 6.6}$$

$$\frac{d}{dt}\begin{bmatrix} X \\ A \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & k_3 & -k_4 \end{bmatrix}\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} - D\begin{bmatrix} X \\ A \end{bmatrix}$$

And:

$$K_f r + F_f + Q_f = 0; \qquad \text{Equation 6.7}$$

$$\begin{bmatrix} -k_1 & -k_2 & 0 \\ -k_4 & -k_5 & -k_6 \\ k_7 & k_8 & k_9 \end{bmatrix}\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} + \frac{F_{in,S}}{W}\begin{bmatrix} S_{in} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ OTR \\ -CTR \end{bmatrix} = 0$$

In the above equations, the specific growth rates were replaced by growth rates to facilitate the calculations.

Since $\mathbf{K_f}$ is full rank, the vector **r** of the growth rates can be written as:

$$r = K_f^{-1}(-Q_f - F_f); \qquad \text{Equation 6.8}$$

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = K_f^{-1}\begin{bmatrix} \frac{-F_{in,S}}{W} S_{in} \\ -OTR \\ CTR \end{bmatrix}$$

Where $\mathbf{K_f^{-1}}$ is given by:

$$K_f^{-1} = \frac{1}{\tilde{k}}\begin{bmatrix} k_7k_9 - k_6k_{10} & k_2k_{10} & k_2k_7 \\ k_5k_{10} - k_7k_8 & -k_1k_{10} & -k_1k_7 \\ k_6k_8 + k_5k_9 & k_1k_9 - k_2k_8 & k_1k_6 - k_2k_5 \end{bmatrix} \qquad \text{Equation 6.9}$$

$$\tilde{k} = k_1k_6k_{10} - k_1k_7k_9 - k_5k_2k_{10} + k_8k_2k_7$$

Replacing Equation 6.8 in Equation 6.6 gives:

$$\frac{d\xi_s}{dt} = K_s K_f^{-1}\left(-Q - F\right) - D\xi_s \qquad \text{Equation 6.10}$$

Which gives, for acetate:

$$\frac{dA}{dt} = -DA + \theta_3 \frac{F_{S,in}}{W} S_{in} - \theta_2 OTR - \theta_1 CTR \qquad \text{Equation 6.11}$$

That represents the input/output reduced model where the parameters θ_1, θ_2, and θ_3 are variable functions of the yield coefficients. That theoretical dependency is represented on Table 6.1. A priori knowledge from those coefficients was used as initial values in the experiments.

Table 6.1 Theoretical dependency of the parameters θ

Parameter	Theoretical dependency
$\theta_1 \times \tilde{k}$	$k_1k_3k_7+k_4(k_1k_6-k_2k_5)$
$\theta_2 \times \tilde{k}$	$-k_3k_1k_{10}-k_4(k_1k_9-k_2k_8)$
$\theta_3 \times \tilde{k}$	$-k_3(k_5k_{10}-k_7k_8)-k_4(k_5k_9-k_6k_8)$

6.1.4.2 Adaptive Control Law

The linearizing control problem consists of deducing and implementing a non-linear law as a function of ξ, Q, and A^*, such that the controller convergence error, $\tilde{A} = (A^* - A)$, can be governed by a pre-specified stable linear differential equation, known as reference model. The control objective will consist of imposing a stable first order closed-loop dynamics of the form:

$$\frac{d}{dt}\left(A^* - A\right) + \lambda\left(A^* - A\right) = 0,\ \lambda > 0 \qquad \text{Equation 6.12}$$

Where λ is the controller gain.

For the design of the control law it is considered that the total feeding rate is equal to substrate feeding rate, and thus that $F_{in,S}/W = D$. Combining the input-output reduced model (Equation 6.11) with the reference model (Equation 6.12), will easily lead to the following linearizing regulation law:

$$D(t) = \frac{\theta_1 CTR + \theta_2 OTR + \lambda_1(A^* - A)}{\theta_3 S_{in} - A} \qquad \text{Equation 6.13}$$

The discretized form of the regulation law can be written as follows:

$$D_k = \frac{\theta_{1,k} CTR + \theta_{2,k} OTR + \lambda_1(A^* - A_k)}{\theta_{3,k} S_{in} - A_k} \qquad \text{Equation 6.14}$$

Employing the reduced model will lead to concentrate all the unknowns in the θ parameters that will be estimated on-line. Hence, in Equation 6.13 and Equation 6.14, the estimates $\hat{\theta}$ will be employed rather than the true values θ.

In the control law, the term $\lambda_1(A^*-A)/(\theta_3 S_{in}-A)$ can be seen as representing a variable gain proportional action, which performs the fine tuning for the feeding profile, while the term $(\theta_1 CTR+\theta_2 OTR)/(\theta_3 S_{in}-A)$ relates the substrate feed rate to the *CTR* and *OTR*, providing the required exponential feeding profile.

It should also be emphasized that for the application of this control law, there is no need of knowledge or estimators for the kinetics, and the only on-line measurements required are acetate concentration and gaseous transfer rates.

6.1.4.3 Estimation Laws

The basis for the method adopted is the so-called Lyapunov design as described by Bastin and Dochain (1990). It is assumed that each parameter θ_i is estimated by the following adaptive law:

$$\frac{d\hat{\theta}_i}{dt} = \gamma \phi_i^{-1}\left(A^* - A\right) \qquad \text{Equation 6.15}$$

Where ϕ_i is the regressor associated to θ_i, depending on *F* and *Q*, and γ a positive definite estimator gain. This equation is consequence of fixing the closed loop dynamics by choosing a real double pole, as proposed by Perrier and Dochain (1993), keeping coupled controller and estimator parameter adaptation. Based on this approach, Ferreira and Feyo de Azevedo (1996) proposed a tuning scheme that enforces a second order convergence dynamics with gains computed as:

$$\lambda = \frac{2\zeta}{\tau} - \frac{1}{\phi_i}\frac{d\phi_i}{dt};\ \gamma = \frac{1}{\tau^2} \qquad \text{Equation 6.16}$$

Which has the meaning that each $\hat{\theta}_i$ converges for the true θ_i through a second order dynamic trajectory, with a natural period of oscillation τ and a damping coefficient ζ.

In the discretized form, this adaptation law can be written as follows:

$$\hat{\theta}_{1,k+1} = \hat{\theta}_{1,k} + T\frac{A^* - A_k}{CTR_k\tau^2}$$ Equation 6.17

$$\hat{\theta}_{2,k+1} = \hat{\theta}_{2,k} + T\frac{A^* - A_k}{OTR_k\tau^2}$$ Equation 6.18

$$\hat{\theta}_{3,k+1} = \hat{\theta}_{3,k} - T\frac{A^* - A_k}{D_k S_{in}\tau^2}$$ Equation 6.19

being T the sampling period and k a time index.

6.2 MATERIALS AND METHODS

6.2.1 Simulations

Simulations for both optimal and adaptive control laws were performed in MATLAB (version 6.0). The dynamical equations of the model described in the previous chapter, together with the identified parameters, were integrated with a 4th order Runge-Kutta method with a fixed stepsize of 0.01 h. Typical initial values for the state variables were used and are shown in Table 6.2. The final time for all the simulations was set to 20 h. OTR and CTR were computed by considering a steady state for dissolved oxygen and carbon dioxide.

Table 6.2 Initial values for the state variables

Variable	Initial Value
X	5 g·kg^{-1}
S	0
A	0 or 0.5 g·kg^{-1}
W	3 kg

For the optimisation of the feeding profile, two methods were tested: a gradient method and a Genetic Algorithm. For the implementation of the first one, the MATLAB Optimization Toolbox (version 2.1) function '`fmincon`' was used. This function calculates the minimum of a function subject to several kinds of constraints starting at an initial estimate. It uses Sequential Programming methods, in which a Quadratic Programming (QP) sub-problem is

solved at each iteration. An estimate of the Hessian of the Lagrangian is updated at each iteration using the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) formula. It requires as inputs the objective function, the vector of initial estimates, constraints, upper and lower bounds for the variables and a previously defined vector of options.

The objective function was evaluated in the script '`coli_optim.m`' file (in the CD-ROM embedded to this thesis), a routine that calculated the final biomass production (the objective function of Equation 6.4) as a function of the input vector $F_{in,S}$. by integrating the dynamical model.

There were no constraint functions passed to the function '`fmincon`', as the final weight restriction (5 kg) was accounted by a penalty in the objective function. The upper and lower bounds of the variable $F_{in,S}$ were zero and the maximum feed rate (0.4 kg·h^{-1}), respectively.

For the implementation of the second method, the Genetic Algorithms toolbox for MATLAB (version 1.7) developed by Hartmut Pohlheim at the University of Sheffield, UK was used. It works with several genetic operators and supports binary, integer and real-valued representations. The representation chosen for this work was the real-valued, due to the known advantages (Roubos et al., 1999). The routine used as objective function was the same as in the previous method. The function '`tbxmpga`' was called from a high level script and conducted all the operations, namely, initialisation of the population, ranking and selection based on the objective function values, recombination/crossover, mutation, evaluation, reinsertion and migration. The number of individuals evaluated in any iteration was 200. Each individual was a sequence of values $F_{in,S}$, between upper and lower bounds mentioned above.

In both cases, in order to optimise the computation time, the integration step was different from the length of the input $F_{in,S}$. Thus, vector $F_{in,S}$ was divided in 20 elements (1 h^{-1}) that were interpolated 100 times for the calculation of the objective function value. The procedure of dividing the control input vector in relatively large intervals is usually found in the literature. However, the majority of the authors consider the feed rate constant along that interval, originating a step change in the feeding profile. The selection of the interpolation methods in detriment of that approach was physiologically motivated, because cells are much more likely to follow a continuous growth, rather than a discontinuous one.

For adaptive control purposes, a MATLAB routine ('`coli_control.m`') was developed to integrate model equations and to generate the feeding rate given by the control law. Adapting

parameters were computed according to the Euler's discretized form shown in Equation 6.17 to Equation 6.19, and the control law implementation followed Equation 6.14.

Initial values of the parameters θ were obtained based on the theoretical dependencies found in Table 6.1.

6.2.2 Cultivation

The adaptive control scheme has been implemented in a real fermentation environment. For this purpose, the supervisory VI described in chapter 4 was modified to implement the described control law. A subVI was developed to calculate on-line both OTR and CTR based on exhaust gas analysis and on the fermenter weight. In this case, and in order not to decrease the maximum sampling rate, the Mass Spectrometer acquired only from the exhaust gas. However, several measures had to be taken in order to have a good estimation of the inlet gas composition: first, Mass Spectrometer had to be stable for a long period, and second, acquisition from air had to be done directly prior to the fermentation.

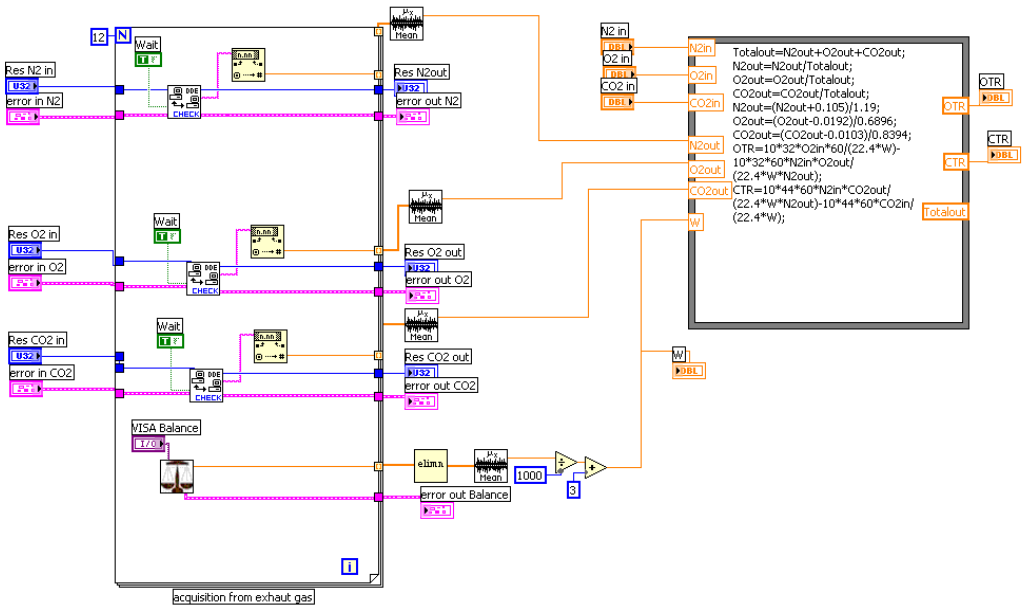


Figure 6.4 Diagram of the LabVIEW program `OTRCTR2.vi`, used to calculate on-line the values of CTR and OTR using MS and reactor balance raw data.

The subVI `OTRCTR2.vi` is shown in Figure 6.4 and works in the following way: 12 points are acquired from the mass spectrometer (approximately 3 minutes) and from the fermenter balance and are stored in arrays created by LabVIEW. For MS data, the corresponding pressure data are averaged and converted to molar fractions, according to the calculated calibration factors. For weight, an additional subVI had to be constructed to eliminate erroneous measurements, based on the standard deviation of the acquired values. The corresponding average values of exhaust gas composition and reactor weight, together with the estimative of the inlet gas composition are given as inputs to a C script embedded in the LabVIEW VI, which computes both CTR and OTR values.

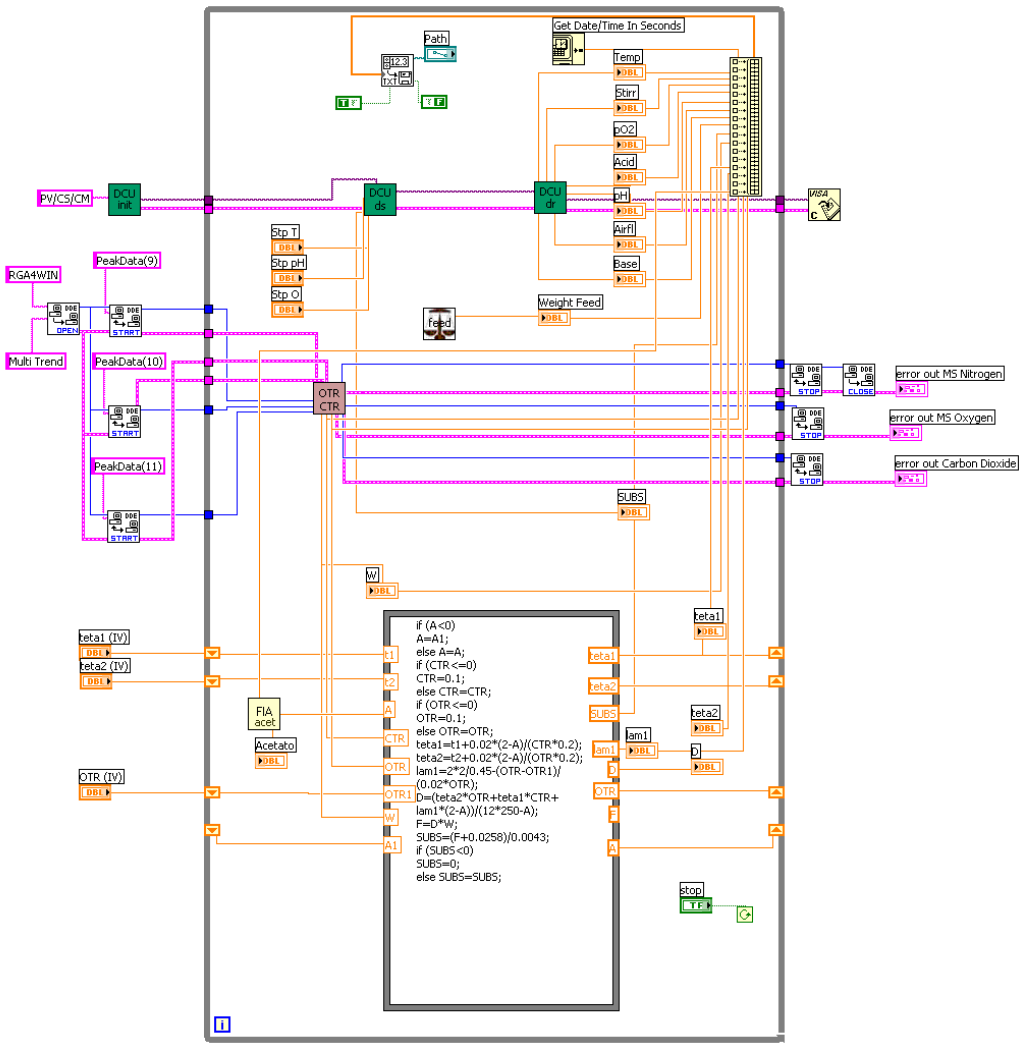


Figure 6.5 Modified supervisory VI (`control_coli.vi`), used for the implementation of the adaptive control law.

This subVI was then integrated in a modified supervisory VI, shown in Figure 6.5, together with the measurements of the other equipments. For the computation of the control law, another C script was embedded, this time directly in the supervisory VI. This script requires the values of CTR, OTR, and weight as inputs and the previous values of the estimated parameters and of OTR. This was achieved by using “shift registers” in the while loop, that are able to store values from one cycle to the following one. Acetate concentration is also an input to the control law computation, and is directly given by the FIA subVI. The output of the C script was the pump feeding rate, which was directly sent to the `DCU DS.VI`.

The limiting step of the cycle was OTR and CTR computations, and thus it was possible to calculate the feeding profile every 3 minutes, approximately.

6.3 RESULTS

6.3.1 Optimal Control

In order to select the best optimisation method, a preliminary study was conducted with the routines described. Afterwards, the results obtained with the selected method were used to simulate the optimised fermentation.

6.3.1.1 Gradient Method

For this optimisation method, results are presented in Table 6.3. The function converged to a solution, although it took 1000 iterations to achieve a high final biomass production.

Table 6.3 Results obtained with the gradient method after 100 and 1000 iterations

Number of iterations	Objective function (g)
Initial values	34.15
100	186.29
1000	241.89

It should also be mentioned that the values shown were obtained for an initial vector $F_{in,S}$ linearly increasing with time. With an initial constant vector $F_{in,S}$ (for example of 0.05 kg·h^{-1}), the function did not converge to a high objective function, remaining in the values of 30 g

even after 1000 iterations. This is due to the already mentioned sensitivity of the method to initial values with linear objective functions.

6.3.1.2 Genetic Algorithm

The results obtained with the GA method are illustrated in Table 6.4. It took only 350 iterations to achieve a final solution slightly higher than the one obtained after 1000 iterations with the gradient method. Moreover, it was not necessary to give a good initial estimation of the solution, as the first estimation is randomly generated by GA's.

It can then be concluded that, for this dynamical optimisation problem, Genetic Algorithms are more advantageous in terms of convergence and initial solution. The convergence to the optimum for this method is shown in Figure 6.6.

Table 6.4 Results obtained with the GA method after 100 and 350 iterations.

Number of iterations	Objective function (g)
Initial values	24.67
100	201.53
350	242.00

For the feeding profile obtained with the best solution found (350th iteration with GA's), a simulation of the fermentation was conducted and gave the results shown in Figure 6.7. The shape of the input vector as well as of the biomass concentration profile along the fermentation, are clearly exponential. Also, the only active pathway is the oxidative growth on glucose, with the corresponding specific growth rate always at 0.11 h^{-1}. This corresponded to a null production of acetate.

These results are coherent with the difference found on biomass yields for both oxidation and fermentation of glucose. In fact, when the fermentative pathway is activated, a much smaller amount of biomass is produced per unit of glucose. Also, the acetate formed in that case inhibits the oxidative pathway, originating a hyperbolic decrease in the global biomass yield. In concordance with these assumptions, it can be seen that, for this simulated fermentation, the inexistence of acetate allowed the cells to grow up to approximately 40 $g \cdot kg^{-1}$, while in the results presented in the previous chapter, it was impossible to obtain more than 30 $g \cdot kg^{-1}$.

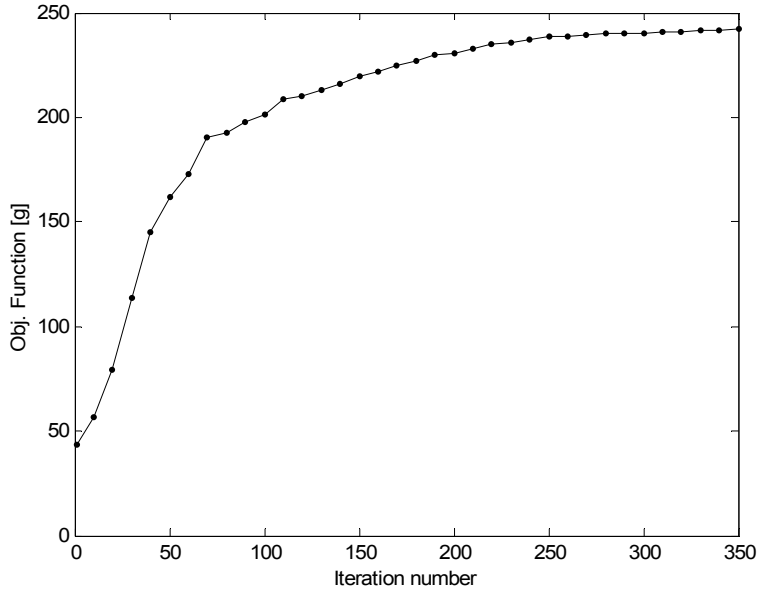


Figure 6.6 Convergence to the optimum value of biomass using the GA's as the optimisation method.

It should also be noticed that optimisations conducted with GA's for any of the literature models where the glucose overflow is not clearly defined, originated feeding profiles with dispersed values, instead of the clear exponential profile obtained in this case (results not shown). This fact reinforces the importance of clearly defining a model structure that can cover precisely the most relevant phenomena occurring during the fermentation of a given microorganism.

However, this feeding profile was not implemented in a real fermentation in a feed-forward approach, because the proximity of the glucose specific uptake rate to the critical value would, almost inevitably, generate acetate production caused by the smallest experimental error. Nevertheless, the purpose of this study was to derivate an optimal profile in order to evaluate how close to the optimum a given fermentation is operated, and, at the same time, to find a suitable optimisation framework for other biotechnological applications.

One of these applications is related with the optimisation of the initial conditions of the fed-batch phase in terms of the state variables and of the fermenter's weight. These variables are most often determined in an empirical way, and it is desirable to know their influence on the performance of the fed-batch fermentation. For this purpose, another optimisation was conducted for the determination of the best initial values for the fed-batch phase of this particular fermentation. The corresponding results are shown in Table 6.5.

Other applications of this optimisation technique could be the computation of the best induction time and the inducer feeding profile optimisation. Also for estimating control parameters, an optimisation technique can save many trial-and error experiments.

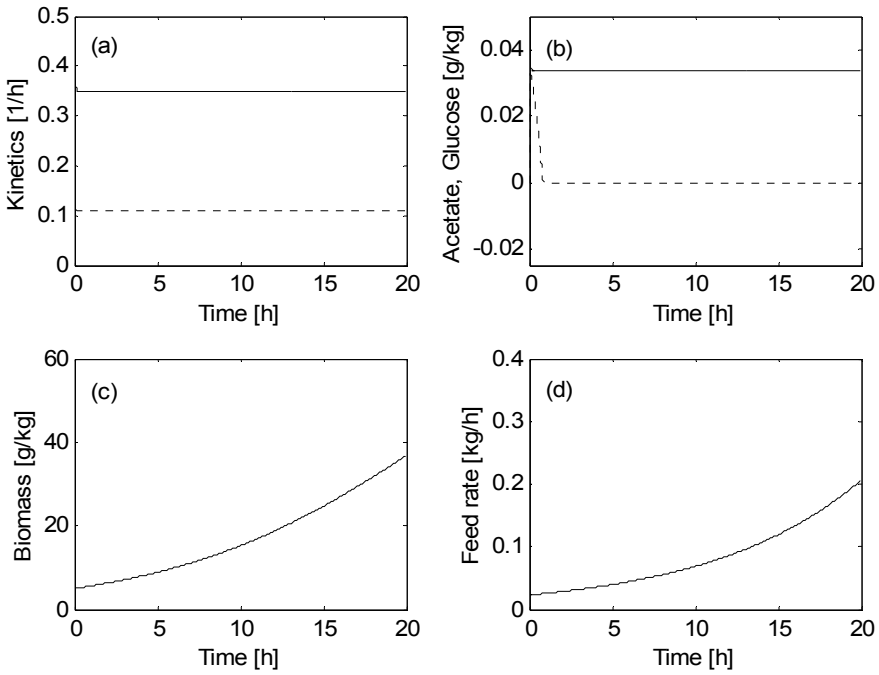


Figure 6.7 Time evolution of the optimised fermentation. (a) specific glucose uptake rate – q_S – (full line) and specific growth rate relative to glucose oxidation – μ_1 – (dotted line); (b) glucose (full line) and acetate (dotted line) concentrations; (c) biomass concentration and (d) glucose solution feeding rate.

Table 6.5 Optimised initial values for the fed-batch phase for biomass, glucose, and acetate concentrations and fermenter's weight

State Variable	Optimised Initial Value
X	5.00 $g \cdot kg^{-1}$
S	0.115 $g \cdot kg^{-1}$
A	0.00 $g \cdot kg^{-1}$
W	3.17 kg

6.3.2 Adaptive Control

The control algorithm derived was firstly used in simulations to control acetate concentration. The chosen set-point was 0.5 $g \cdot kg^{-1}$ (below inhibition levels).

An example of the fermentation behaviour is shown in Figure 6.8. Due to the control of acetate concentration, the microorganism is able to grow without acetate inhibition until the end of the fed-batch phase. The final biomass concentration achieved was around 20 $g \cdot kg^{-1}$. Glucose accumulation is also very limited. However, it should be noticed that the biomass concentration is very far from the theoretical maximum obtainable and presented in the previous section. Also, even when compared with the results shown in the last chapter, biomass production is lower. Nevertheless, acetate accumulation is limited, and thus glucose consumption is also limited, decreasing the costs of the fermentation, when compared with a non-controlled one.

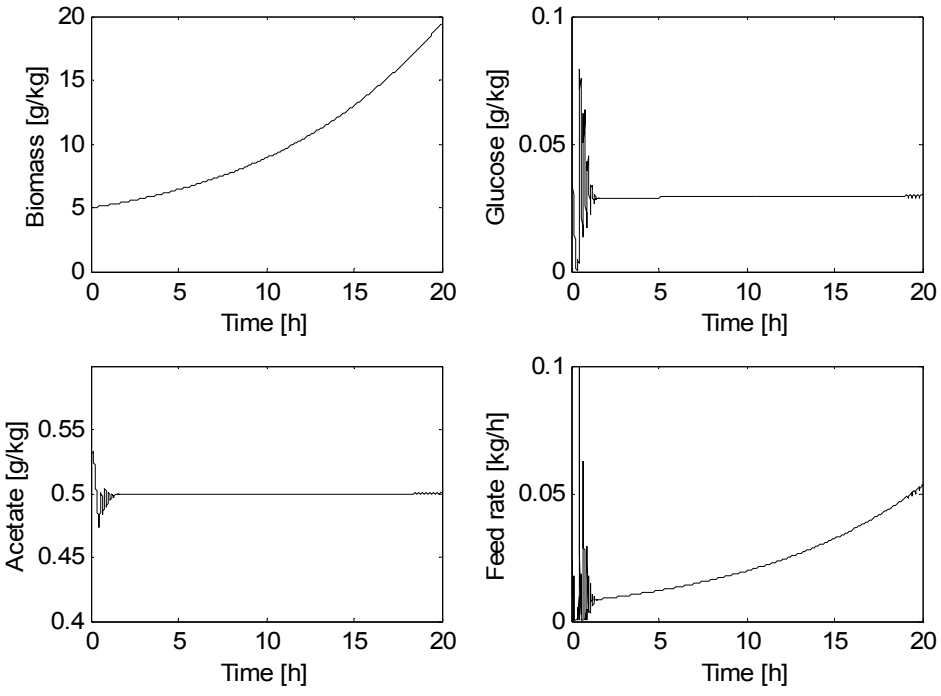


Figure 6.8 Example of a controlled fermentation.

In order to search for the best alternative for the controller design, different approaches were studied regarding the control parameters, namely the simultaneous estimation of the three θ parameters; or keeping constant one (θ_3), and two (θ_2 and θ_3) of them, based on their theoretical dependency on the yield coefficients described in Table 6.1. Comparative results are shown in Figure 6.9. In all cases, a good convergence of the controlled variable to the set-point is obtained. However, a slightly smaller divergence occurs at the beginning of the fermentation when the three parameters are adapted on-line.

The convergence of the control parameters to the initial values is also good, as seen in Figure 6.10.

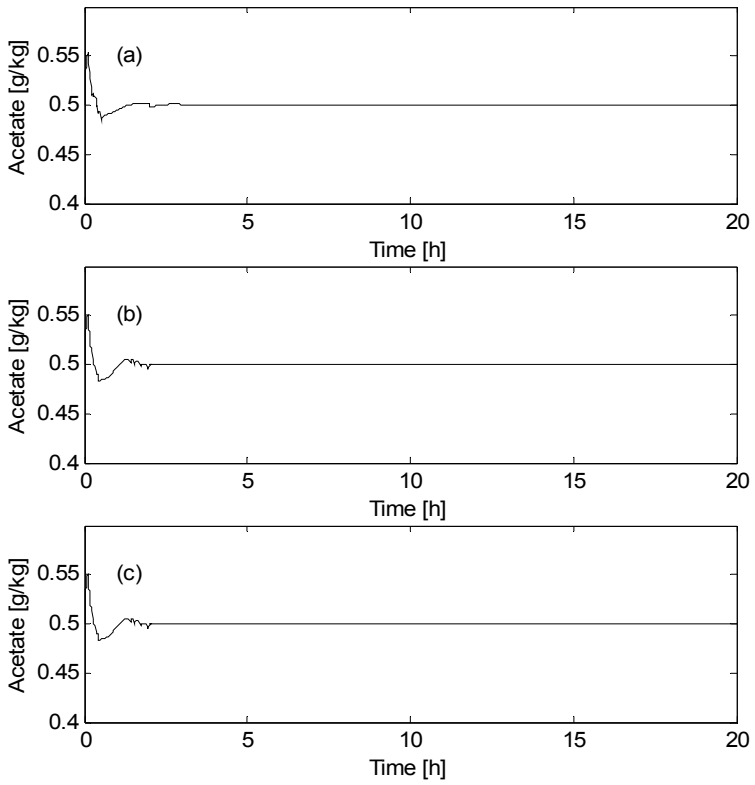


Figure 6.9 Controller performance under: (a) estimation of the three θ; (b) holding only θ_3, (c) holding of θ_2 and θ_3.

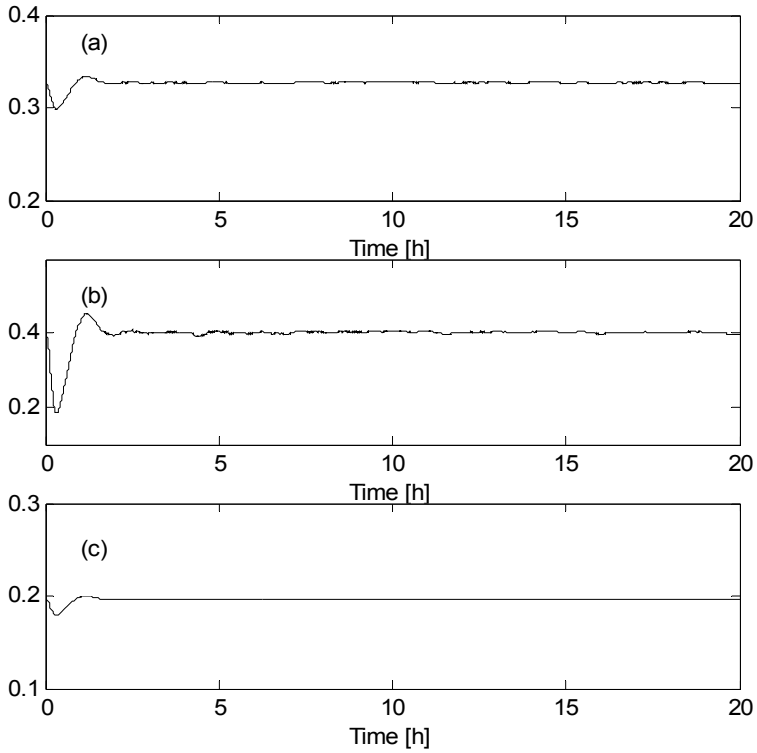


Figure 6.10 Convergence of the controller parameters to the initial values. (a) represents the time evolution of θ_1 (b) of θ_2 and (c) of θ_3.

With the purpose of testing the controller's robustness, other simulations were performed (Figure 6.11). First, a 50% change in the glucose concentration in the feeding at 10 h was studied. This alteration is likely to happen in a real case (although not with such an intensity), caused by a modification in the feeding pump or even by the degradation of the feeding solution. A small decrease in acetate concentration is verified, although the controller rapidly imposes a convergence to the set-point.

In another simulation, acetate set-point was switched from 0.5 to 1 $g \cdot kg^{-1}$ at 10 h of fermentation run, with a good set-point tracking behaviour (while keeping the same tuning parameters).

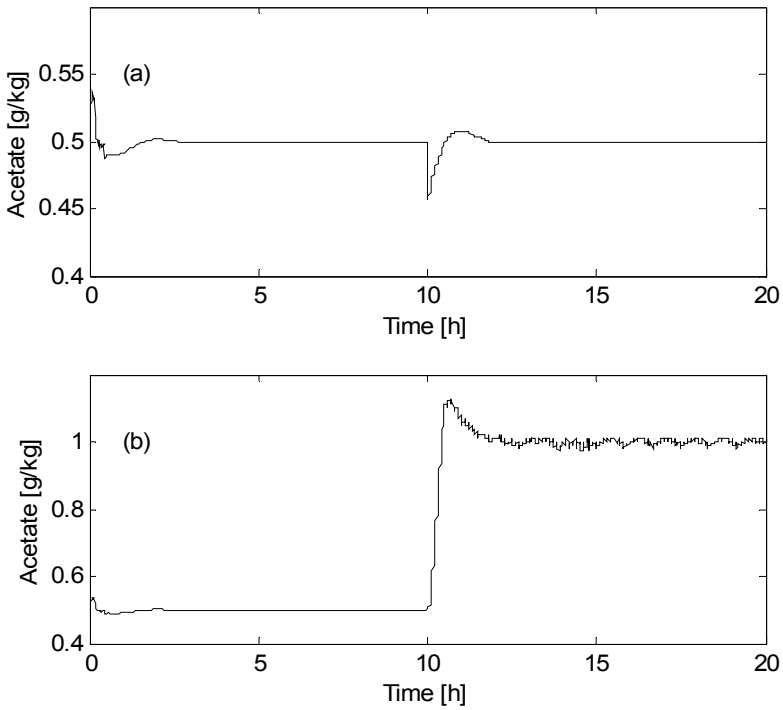


Figure 6.11 (a) Effects of a 50% change in S_{in} at 10h. (b) Set-point tracking from 0.5 to 1 g·kg^{-1}.

In order to check for the effect of signal to noise ratio on the controller performance, a 5% white noise was added to the measured variables (acetate and exhausted oxygen and carbon dioxide). It is clear from Figure 6.12 that the controller still converges to the set-point, although the variance increases as the fermentation approaches the end.

Also, and with the aim of testing the reliability of the acetate measurements conducted with FIA and their applicability to control purposes, other simulations were conducted. In this case, also represented in Figure 6.12, the “real” acetate values were affected by a 20% signal increase (the average of the difference obtained between HPLC and FIA acetate values) plus 5% of white noise (much higher than what is observed with acetate FIA measurements). In this case, although the simulation used the “real” values of acetate concentration, the controller had no knowledge of the 20% drift in the measurements. The results show a good convergence, although the set-point of “real” acetate suffers a negative deviation of 16-20%. This deviation was considered to be negligible for control purposes and acetate measurements with FIA were considered to be reliable for on-line control applications.

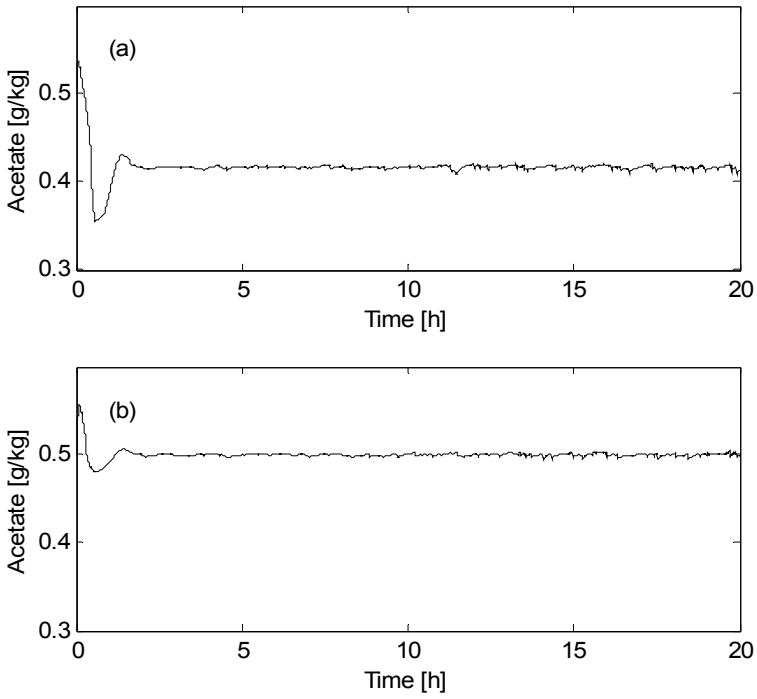


Figure 6.12 (a) 20% drift in acetate plus 5% white noise; (b) Effect of a 5% signal to noise in measured variables.

Finally, the developed methodology was applied to a real fermentation. The desired set-point of 0.5 $g \cdot kg^{-1}$ was impossible to implement due to the high background signal found in the FIA measurements caused by the presence of other organic acids and probably also some carbon dioxide in the form of hydrogen carbonate that was not eliminated in the filtration step. These results do not correspond to the ones obtained by simulations, where interferences in FIA were accounted. However, simulations considered that the maximum value of interference was 20%, which, at low concentrations of acetate, is certainly an underestimation.

Thus, the set-point had to be changed to 2 $g \cdot kg^{-1}$. Although this value is still not inhibiting, the FIA system should be enhanced to enable accurate measurements at low acetate concentrations, due to the low biomass yield on glucose when this compound is deviated to the fermentative pathway.

The results obtained using the new set-point are shown in Figure 6.13. It is clear that acetate concentration was kept around the set-point, after an initial positive deviation. An interesting

phenomenon is the improvement of the controller's performance close to the end of the fermentation. There is no clear justification for this occurrence, but a possible explanation is that the increase in biomass concentration causes an increase in the values of OTR and CTR, and the MS data are affected by a smaller relative error in those situations, improving the controller's accuracy. Thus, a possible way of improving the controller's performance is to accumulate higher concentrations of biomass during the batch phase.

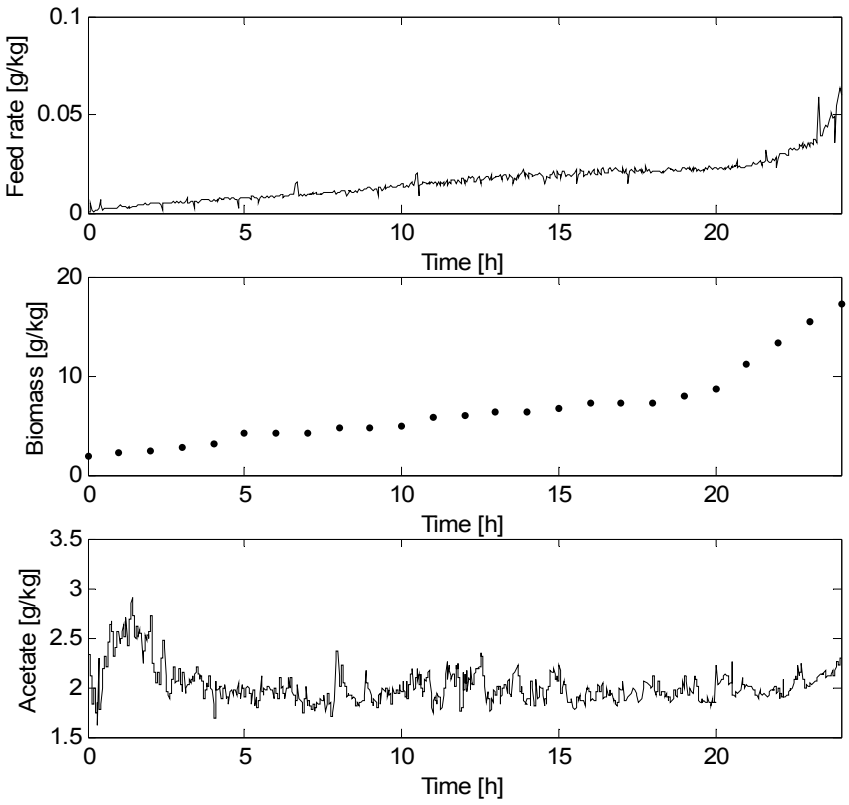


Figure 6.13 Fermentation performed with adaptive control of acetate concentration at a set-point of 2 $g \cdot kg^{-1}$.

Biomass growth is therefore much stronger at the end of the fermentation, although the amount of biomass achieved at the end is still much lower than the optimum value, and even smaller than the value obtained with exponential feeding (described on Chapter 5), in agreement with the simulation results.

Nevertheless, other experiments should be conducted in order to evaluate accurately the performance of the proposed control scheme.

Also, new simulation studies should be conducted for improving the trial and error tuning procedure since controller's performance is very sensitive to controller parameters. An optimisation procedure could be conducted in order to identify both the tuning parameters and the initial estimates of the θs.

6.4 CONCLUSIONS

Two optimisation methods were evaluated for deriving an optimal feeding profile of glucose. Both methods under comparison revealed useful in achieving an optimal solution for the fed-batch fermentation of *E. coli*. However, the GA's revealed more efficient concerning the computed maximum, and dependence on good initial values.

The results obtained with this method were used to simulate a fed-batch fermentation, and a cellular density of 40 $g \cdot kg^{-1}$ was obtained. The corresponding feeding profile was clearly exponential, and the cells followed only the glucose oxidation pathway. These results are in agreement with the knowledge available about *E. coli*, and consequently validate in a certain degree the selected mathematical model.

However, the obtained optimal profile was not implemented in a feed-forward approach in a real fermentation, because there was a high probability of acetate accumulation due to small experimental errors.

The selected optimisation technique can also be applied to other variables in fermentation processes, and an example is given for the calculation of the optimal values for the state variables at the beginning of the fed-batch phase.

An algorithm for the on-line adaptive control of the undesirable acetate production was also developed. It is based in the application of Lyapunov estimation design for parameter adaptation of a linearizing regulation algorithm using a reduced version of the state space dynamical model.

Simulation studies indicate that the developed algorithm is able to maintain acetate below toxic levels, converging rapidly to the imposed set-point.

The controller's robustness was assessed by changing the set-point and the influent glucose concentration, as well as introducing a 5% noise signal in the measured variables. The responses obtained were considered satisfactory.

The implementation of the control law was conducted and a limited accumulation of acetate was obtained. Nevertheless, it was impossible to impose a set-point lower that 2 $g \cdot kg^{-1}$, due to the significant background signal of FIA measurements. Therefore, it is recommended to further optimise the method developed for acetate measurements if a lower set-point is required.

Also, biomass accumulation was much lower than the maximum attainable, both in real and simulated results. Therefore, an improve in the design of the control law is required.

An interesting approach that would combine the main advantages of both optimal and adaptive control schemes would be an optimal adaptive control scheme, that could force the system into a near optimal solution and, at the same time, would be capable of adapting itself to process variations (Van Impe, 1993).

6.5 REFERENCES

Akesson, M., Hagander, P., and Axelsson, J. P. A probing feeding strategy for *Escherichia coli* cultures. *Biotechnology Techniques.* 13, 523-528. 1999a.

Akesson, M., Karlsson, E. N., Hagander, P., Axelsson, J. P., and Tocaj, A. On-line detection of acetate formation in *Escherichia coli* cultures using dissolved oxygen responses to feed transients. *Biotechnology and Bioengineering.* 64, 590-598. 1999b.

Angelov, P. and Guthke, R. A genetic-algorithm-base approach to optimization of bioprocesses described by fuzzy rules. *Bioprocess Engineering.* 16, 299-303. 1997.

Aristidou, A., San, K., and Bennett, G. Improvement of biomass yield and recombinant gene expression in *Escherichia coli* by using fructose as the primary carbon source. *Biotechnology Progress.* 15, 140-145. 1999.

Åström, K. J. and Wittenmark, B. *Adaptive Control*. Addison-Wesley, Reading. 1989.

Banga, J. R., Balsa-Canto, E, Moles, C. G., and Alonso, A. A. Efficient and robust numerical strategies for the optimal control of non-linear bio-processes. *Proceedings of the 10th Mediterranean Conference on Control and Optimization, Lisbon, Portugal.* 2002.

Bastin, G. and Dochain, D. *On-line estimation and adaptive control of bioreactors*. Elsevier Science Publishers, Amsterdam. 1990.

Breusegem, V. V. and Bastin, G. Optimal control of biomass growth in a mixed culture. *Biotechnology and Bioengineering.* 35, 349-355. 1990.

Castan, A., Näsman, A., and Enfors, S.-O. Oxygen enriched air supply in *Escherichia coli* processes: production of biomass and recombinant human growth hormone. *Enzyme and Microbial Technology.* 30, 847-854. 2002.

Chen, H.-C., Hwang, C.-F., and Mou, D.-G. High-density *Escherichia coli* cultivation process for hyperexpression of recombinant porcine growth hormone. *Enzyme and Microbial Technology.* 14, 321-326. 1992.

Chiou, J.-P. and Wang, F. S. Hybrid method of evolutionary algorithms for static and dynamic optimization problems with application to a fed-batch fermentation process. *Computers and Chemical Engineering.* 23, 1277-1291. 1999.

Claes, J. E. and Van Impe, J. F. On-line monitoring and optimal adaptive control of the fed-batch baker's yeast fermentation. *Horizon of BioProcess Systems Engineering in 21st Century (7th Int. Conf. Computer Applications on Biotechnology, Osaka, Japan).* (Yoshida, T. and Shioya, S., Eds.). Pergamon, Oxford. 405-410. 1998.

Dochain, D. Design of adaptive controllers for non-linear stirred tank bioreactors: extension to the MIMO situation. *Journal of Process Control.* 1, 41-48. 1991.

Farmer, W. and Liao, J. Progress in metabolic engineering. *Current Opinion in Biotechnology.* 7, 198-204. 1996.

Ferreira, E. C. and Feyo de Azevedo, S. Adaptive linearizing control of bioreactors. *IEE Conference Conference Publication No. 427 (UKACC International Conference on CONTROL'96, Exeter, UK).* Vol. II, 1184-1189. 1996.

Han, K., Hong, J., and Lim, H. C. Relieving effects of glycine and methionine from acetic acid inhibition in *Escherichia coli* fermentation. *Biotechnology and Bioengineering.* 41, 316-324. 1992.

Jadot, F., Bastin, G., and Van Impe, J. F. Optimal adaptive control of a bioprocess with yield-productivity conflict. *Journal of Biotechnology.* 65, 61-68. 1998.

Johnston, W., Cord-Ruwisch, R., and Cooney, M. J. Industrial control of recombinant *E. coli* fed-batch culture: new perspectives on traditional controlled variables. *Bioprocess and Biosystems Engineering.* 25, 111-120. 2002.

Jorgensen, S. B., Moller, H. E., and Andersen, M. Y. Adaptive control of continuous yeast fermentation, near critical dilution rate. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 107-112. 1992.

Khosla, C. and Bailey, J. E. Heterologous expression of a bacterial haemoglobin improves the growth properties of recombinant *Escherichia coli*. *Nature.* 331, 633-635. 1988.

Kleman, G. L. and Strohl, W. R. Developments in high cell density and high productivity microbial fermentation. *Current Opinion in Biotechnology.* 5, 180-186. 1994.

Konstantinov, K., Kishimoto, M., Seki, T., and Yoshida, T. A balanced DO-stat and its application to the control of acetic acid excretion by recombinant *Escherichia coli*. *Biotechnology and Bioengineering.* 36, 750-758. 1990.

Konstantinov, K. and Yoshida, T. Knowledge-based control of fermentation processes. *Biotechnology and Bioengineering.* 39, 479-486. 1992.

Lee, J., Lee, S. Y., Park, S., and Middelberg, A. Control of fed-batch fermentations. *Biotechnology Advances.* 17, 29-48. 1999.

Lee, J. and Ramirez, W. F. Optimal fed-batch control of induced foreign protein production by recombinant bacteria. *AIChE Journal.* 40, 899-907. 1994.

Lee, J. H., Hong, J., and Lim, H. C. Experimental optimization of fed-batch culture for poly-β-hidroxybutiric acid production. *Biotechnology and Bioengineering.* 56, 697-705. 1997.

Lee, S. C., Kim, C. G., Chang, Y. K., and Chang, H. N. Dissolved oxygen concentration and growth rate control in fed-batch fermentation process. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 267-270. 1992.

Levisauskas, D. Inferential control of the specific growth rate in fed-batch cultivation processes. *Biotechnology Letters.* 23, 1189-1195. 2001.

Luli, G. W. and Strohl, W. R. Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Applied and Environmental Microbiology.* 56, 1004-1011. 1990.

Macaloney, G., Hall, J. W., Rollins, M. J., Draper, I., Anderson, K. B., Preston, J., Thompson, B. G., and McNeil, B. The utility and performance of near-infra red spectroscopy in simultaneous monitoring of multiple components in a high cell density recombinant *Escherichia coli* production process. *Bioprocess Engineering.* 17, 157-167. 1997.

Modak, J. M., Lim, H. C., and Tayeb, Y. J. General characteristics of optimal feed rate profiles for various fed-batch fermentation processes. *Biotechnology and Bioengineering.* 28, 1396-1407. 1986.

Montague, G. A., Morris, A. J., Wright, A. R., Aynsley, M., and Ward, A. C. Modelling and adaptive control of fed-batch penicillin fermentation. *The Canadian Journal of Chemical Engineering.* 64, 567-580. 1986.

Moriyama, H. and Shimizu, K. On-line optimisation of culture temperature for ethanol fermentation using a genetic algorithm. *Journal of Chemical Technology and Biotechnology.* 66, 217-222. 1996.

Na, J.-G., Chang, Y. K., Chung, B. H., and Lim, H. C. Adaptive optimization of fed-batch culture of yeast by using genetic algorithms. *Bioprocess and Biosystems Engineering.* 24, 299-308. 2002.

Narendra, K. S. and Annaswamy, A. M. *Stable Adaptive Systems*. Prentice-Hall, Englewood Cliffs. 1989.

Nguang, S. K., Chen, L., and Chen, X. D. Optimisation of fed-batch culture of hybridoma cells using genetic algorithms. *ISA Transactions.* 40, 381-389. 2001.

Nougués, J. M., Grau, M. D., and Puigjaner, L. Parameter estimation with genetic algorithm in control of fed-batch reactors. *Chemical Engineering and Processing.* 41, 303-309. 2002.

O'Connor, G. M., Sanchez-Riera, F., and Cooney, C. L. Design and evaluation of control strategies for high cell density fermentations. *Biotechnology and Bioengineering.* 39, 293-304. 1992.

Park, S. and Ramirez, W. F. Optimal production of secreted protein in fed-batch reactors. *AIChE Journal.* 34, 1550-1558. 1988.

Perrier, M. and Dochain, D. Evaluation of control strategies for anaerobic digestion processes. *International Journal of Adaptive Control and Signal Processing.* 7, 309-321. 1993.

Pomerleau, Y., Perrier, M., and Bourque, D. Dynamics and control of the fed-batch production of poly-β-hydroxybutyrate by *Methylobacterium extorquens*. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 107-112. 1995.

Pomerleau, Y. and Viel, G. Industrial applications of adaptive nonlinear control for bakers' yeast production. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp .and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 315-318. 1992.

Riesenberg, D., Menzel, K., Schulz, V., Schumann, K., Veith, G., Zuber, G., and Knorre, W. A. High cell density fermentation of recombinant *Escherichia coli* expressing human interferon alpha 1. *Applied Microbiology and Biotechnology.* 34, 77-82. 1990.

Ronen, M., Shabtai, Y., and Guterman, H. Optimization of feeding profile for a fed-batch bioreactor by an evolutionary algorithm. *Journal of Biotechnology.* 97, 253-263. 2002.

Rothen, S. A., Sauer, M., Sonnleitner, B., and Witholt, B. Growth characteristics of *Escherichia coli* HB101 [pGEc47] on defined medium. *Biotechnology and Bioengineering.* 58, 92-100. 1998.

Roubos, J. A., Van Straten, G., and Van Boxtel, A. J. B. An evolutionary strategy for fed-batch bioreactor optimization: concepts and performance. *Journal of Biotechnology.* 67, 173-187. 1999.

Roux, G., Dahhou, B., Najim, K., and Queinnec, I. Adaptive linear quadratic gaussian (LQG) control of a bioreactor. *Journal of Chemical Technology and Biotechnology.* 53, 133-141. 1992.

Sastry, S. and Bodson, M. *Adaptive Control. Stability, Convergence, and Robustness*. Prentice-Hall, Inc, Englewood Cliffs. 1989.

Saucedo, V. M., Valentinotti, S. C., Karim, M. N., and Collins, H. W. Feedback control of a recombinant fed-batch fermentation using on-line HPLC measurements. *6th Int. Conf. Computer Applications on Biotechnology, Garmisch-Partenkirchen, Germany.* (Munack, A. and Schügerl, K., Eds.). Pergamon, Oxford. 24-28. 1995.

Shimizu, N., Fukuzono, S., Fujimori, K., Nishimura, N., and Odawara, Y. Fed-batch cultures of recombinant *Escherichia coli* with inhibitory substance concentration monitoring. *Journal of Fermentation Technology.* 66, 187-191. 1988.

Shukla, P. K. and Pushpavanam, S. Optimisation of biochemical reactors - an analysis of different approximations of fed-batch operation. *Chemical Engineering Science.* 53, 341-352. 1998.

Siegwart, P., Cote, J., Male, K., Luong, J. H. T., Perrier, M., and Kamen, A. Adaptive control at low glucose concentration of Hek-293 cell serum-free cultures. *Biotechnology Progress.* 15, 608-616. 1999.

Tartakovsky, B., Ulitzur, S., and Sheintuch, M. Optimal control of fed-batch fermentation with autoinduction of metabolite production. *Biotechnology Progress.* 11, 80-87. 1995.

Tholodur, A. and Ramirez, W. F. Optimization of fed-batch bioreactors using neural network parameter function model. *Biotechnology Progress.* 12, 302-309. 1996.

Tomson, K., Paalme, T., Laakso, P. S., and Vilu, R. Automatic laboratory-scale fed-batch procedure for production of recombinant proteins using inducible expression systems of *Escherichia coli*. *Biotechnology Techniques.* 9, 793-798. 1995.

Tremblay, M., Perrier, M., Chavarie, C., and Archambault, J. Fed-batch culture of hybridoma cells: comparison of optimal control approach and closed loop strategies. *Bioprocess Engineering.* 9, 13-21. 1993.

Tremblay, M., Perrier, M., and Dochain, D. Application of dynamic programming to the optimization of an animal cell culture fed-batch reactor. *Proc. 13th IMACS World Congress on Computation and Applied Mathematics (Dublin, Ireland)*. 1493-1494, 1991.

Turner, C., Gregory, M. E., and Thornhill, N. Closed-loop control of fed-batch cultures of recombinant *Escherichia coli* using on-line HPLC. *Biotechnology and Bioengineering.* 44, 819-829. 1994.

Van Impe, J. F. *Modeling and optimal adaptive control of biotechnological processes*. Katholieke Universiteit Leuven, Belgium. PhD thesis. 1993.

Waldraff, W., King, R., and Gilles, E. D. Optimal feeding startegies by adaptive mesh selection for fed-batch bioprocesses. *Bioprocess Engineering.* 17, 221-227. 1997.

Wang, F. S. and Shyu, C. H. Optimal feed policy for fed-batch fermentation of ethanol production by *Zymomous mobilis*. *Bioprocess Engineering.* 17, 32-68. 1997.

Williams, D., Yousefpour, P., and Wellington, M. H. On-line adaptive control of a fed-batch fermentation of *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering.* 28, 631-645. 1986.

Zeng, F. Y., Dahhou, B., Goma, G., and Nihtilä, M. T. Adaptive estimation and control of the specific growth rate of a nonlinear fermentation process via MRAC method. *Modelling and Control of Biotechnical Processes 1992 (2nd IFAC Symp. and 5th Int. Conf. Computer Applications in Fermentation Technology, Keystone, CO, USA).* (Karim, M. N. and Stephanopoulos, G., Eds.). Pergamon, Oxford. 363-365. 1992.

Zigova, J., Mahle, M., Paschold, H., Malissard, M., Berger, E. G., and Weusterbotz, D. Fed-batch production of a soluble Beta-1,4- galactosyltransferase with *Saccharomyces cerevisiae*. *Enzyme and Microbial Technology.* 25, 201-207. 1999.

Zuo, K. and Wu, W. T. Semi-realtime optimization and control of a fed-batch fermentation system. *Computers and Chemical Engineering.* 24, 1105-1109. 2000.

CHAPTER 7
GENERAL CONCLUSIONS

"There is one thing even more vital to science than intelligent methods; and that is, the sincere desire to find out the truth, whatever it may be."

Charles Sanders Pierce.

In this chapter, the major conclusions extracted from the present thesis are addressed. More detailed conclusions can be found at the end of each individual chapter. Also, some suggestions for further research in this field are given.

7.1 CONCLUSIONS

The objective of the present thesis was to develop model-based strategies for the computer-aided operation of the production of recombinant proteins with *Escherichia coli*. For covering these aims, several subjects were studied and different strategies were successfully implemented. The main conclusions that can be derived are the following:

- A Flow Injection Analysis method for the simultaneous on-line measurement of glucose and acetate was developed. Optimisation procedures allowed the elimination of major interferences by varying carrier and other reagents composition, while the developed sampling system permitted the analysis with minimised delays. Due the flexibility of both methods, it is possible, without a dilution step and, by only varying carrier compositions, to apply this system to a diversity of fermentation processes, where both components can be present in a wide range of concentrations.
- The low cost of reagents, together with the simplicity of the system, makes this technique an attractive alternative for the on-line monitoring of biotechnological processes, like the production of recombinant products with *Escherichia coli*.
- A modular system for on-line data acquisition and control system was developed in the LabVIEW graphical programming environment to allow the integration of several on-line measurements in one supervisory computer. When compared with commercial software programs for the same task, this one presents several advantages, like the possibility of implementing complex control algorithms. Also, it allows the easy integration of virtually any new measurement that would result from an upgrade to the monitoring system, due to its modular structure. Finally, the graphical environment facilitates the development of new programs, since a minimum knowledge of programming techniques is required.
- A complete unstructured mathematical model for representing growth of the bacterium *E. coli* in fed-batch cultures was developed. The dynamical equations have been derived based on the state-space dynamical model framework.
- The kinetic model that describes the oxidative growth on glucose and / or acetate and glucose overflow metabolism was selected from several literature options, being the selection based on an integrated approach, based on the choice of the best model that, after parameter estimation, fits the best the experimental data. The best kinetic model includes an inhibition of the critical glucose specific uptake rate by

acetate, which causes a cyclic phenomenon, where acetate accumulation induces more acetate accumulation, if an exponential feeding profile is maintained. This conclusion raised the need of developing new feed-back feeding strategies that would avoid this situation.

- All model parameters were identified by using the Genetic Algorithms optimisation method for minimizing the quadratic difference between real and simulated data for the state variables. This procedure was executed in several steps, in order to separate the different phenomena involved and to isolate the interference of some noisy data, like oxygen and carbon dioxide transfer rates.
- After parameter estimation, the simulations could well describe biomass, glucose and acetate concentrations, as well as oxygen and carbon dioxide evolution rates. Also, the time profiles for the calculated specific uptake and growth rates are able to describe very well the main phenomena observed during *E. coli* fed-batch fermentations.
- However, when the final model was used to predict a fermentation different from the ones that had originated the model structure and parameters, the approximation between real and simulated values of the state variables was not excellent, although both fermentations were conducted in the same exact conditions. These results once again motivated the development of a feed-back type of control with an adaptive capacity in order to cope with these process variabilities.
- Besides yield and kinetic parameters, also the parameters that correlate the total feeding rate with specific growth rates and with oxygen and carbon dioxide transfer rates were estimated. The motivation for this approach is to accurately simulate the variations of weight during the fermentation, which will have a profound effect on the simulation of the other state variables, and that are most often ignored for process simulations. This was made possible due to the existence of a balance that continuously acquires the weight from the fermenter and due to the data acquisition system developed.
- After model selection and parameter identification, a sensitivity analysis was conducted in order to check the influence of parameters changes on the model state variables. It was concluded that the parameters that have the strongest influence on the state variables are the kinetic parameters related with glucose consumption, while the least significant ones are one of the yield coefficients and the parameters related with acetate consumption. It can then extracted that acetate consumption at

concentration between 5 and 10 $g \cdot kg^{-1}$ could be described in a simpler way, by considering that it occurs always at a maximum rate.

- Two optimisation methods were evaluated for deriving an optimal feeding profile of glucose with the aim of maximizing the fermentation productivity. Both methods under comparison revealed useful in achieving an optimal solution for the fed-batch fermentation of *E. coli*. However, the Genetic Algorithms revealed more efficient when compared to the gradient method concerning to the computed maximum, and dependence on good initial values.
- Simulation results show, as expected, that an exponential feeding just below the critical specific growth rate yields maximum biomass production. However, the implementation of this feeding profile in a feed-forward approach is not expected to correspond to this maximum value, due to the vicinity of the operational conditions that trigger acetate accumulation. Nevertheless, these results give a very good indication about the theoretical maximum production achievable with the strain of *E. coli* used.
- An algorithm for the on-line adaptive control of the undesirable acetate production during the recombinant *E. coli* fed-batch fermentation was developed. It is based in the application of Lyapunov estimation design for parameter adaptation of a linearizing regulation algorithm using a reduced version of the state space dynamical model. Simulation studies indicate that the developed algorithm is able to maintain acetate below toxic levels, converging rapidly to the imposed set-point.
- The controller's robustness was assessed by changing the set-point and the influent glucose concentration, as well as introducing a 5% noise signal in the measured variables. The responses obtained were considered satisfactory.
- The implementation of the designed control law in a real fermentation was possible thanks to the developed FIA system for on-line monitoring of acetate. However, due to the non-negligible background signal of this method, it was only possible to control this metabolite at a set-point of 2 $g \cdot kg^{-1}$. Nevertheless, the control law revealed effective to accomplish the desired objective, and biomass growth was very close to the simulation results.
- However, biomass accumulation was much lower than the maximum attainable for this process, both in real and simulated results. Therefore, an improvement in the design of the control law is required.

7.2 RECOMMENDATIONS

- The approach selected for model identification gave very interesting results for deriving the best kinetic structure and estimating the parameters of a mathematical model. However, the proposed model estimation procedure could be improved by using optimal experimental design tools based on Fisher Information Matrix, in order to maximize the information obtained for each fermentation.
- The methodology developed for feeding profile optimisation can also be successfully applied to the optimisation of many other parameters, like induction feeding rate and starting time, initial fed-batch conditions and final fermentation time, among other relevant problems in the fermentation field, which are, most of the times, empirically determined.
- In what adaptive control concerns, if a set-point lower than 2 $g \cdot kg^{-1}$ is required for acetate concentration, it is recommended to further optimize the method developed for acetate FIA measurements.
- In order to improve the performance of the adaptive controller, related with an increase in biomass accumulation, an interesting approach that would combine the main advantages of both optimal and adaptive control schemes could be used. This is the case of an optimal adaptive control scheme, that could force the system into a near optimal solution and, at the same time, would be capable of adapting itself to process variations

ANNEX

Annex 1 — Calibrations
Annex 2 — Communication Parameters
Annex 3 — Sensitivity Functions
Annex 4 — CD Contents

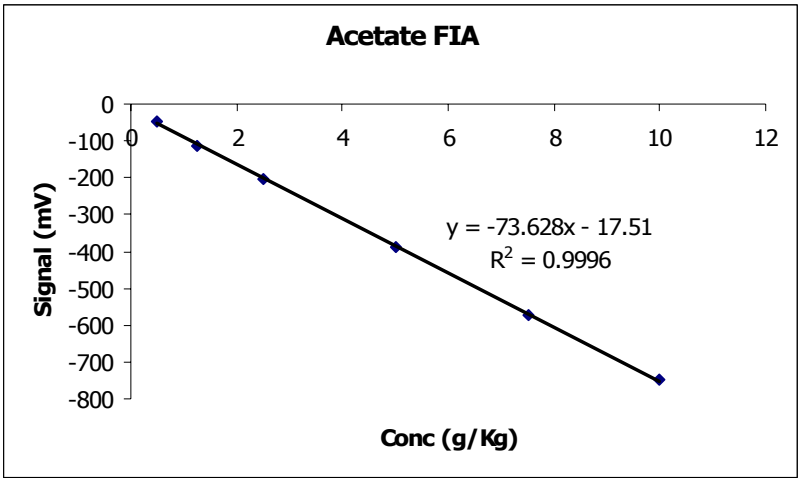
1. Calibrations

Examples of calibrations for several off-line methods are given in this section. Those examples are shown in the format of the Excel™ sheets used to obtain the calibration coefficients. Also, for some methods, a schematic description is given for a better understanding of the adopted methodologies.

1.1 ACETATE MEASUREMENTS WITH FIA

Solution	Added Mass	from solution	Dilution	Final Mass (g)	Conc. (g/Kg)
1	Stock solution			100	50.0
2	1.0	1	100	100	0.500
3	2.5	1	40	100	1.25
4	5.0	1	20	100	2.50
5	10	1	10	100	5.00
6	15	1	6.7	100	7.50
7	20	1	5.0	100	10.0

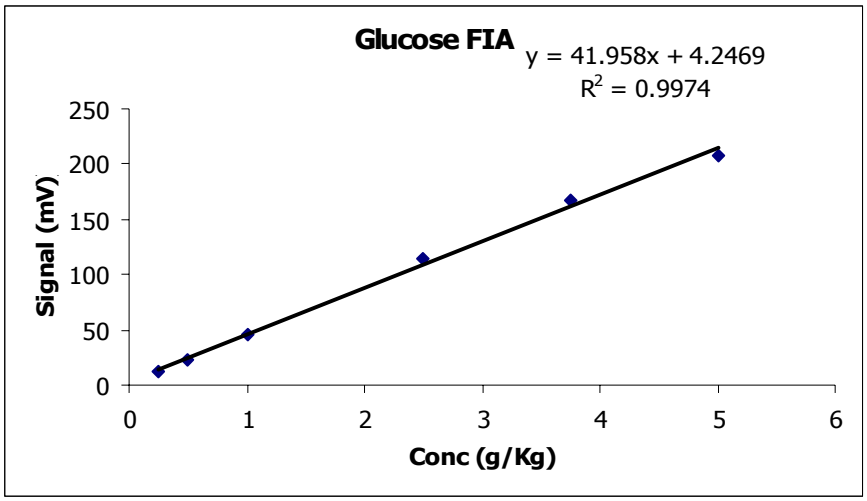
Solution	Conc. (g/Kg)	Height (mV)
2	0.500	-45.8
3	1.25	-114.5
4	2.50	-205.0
5	5.00	-386.2
6	7.50	-573.4
7	10.0	-749.7



1.2 GLUCOSE MEASUREMENTS WITH FIA

Solution	Added Mass	from solution	Dilution	Final Mass (g)	Conc. (g/Kg)
1	Stock solution			100	50.0
2	0.50	1	200	100	0.250
3	1.0	1	100	100	0.500
4	2.0	1	50.0	100	1.00
5	5.0	1	20.0	100	2.50
6	7.5	1	13.3	100	3.75
7	10	1	10.0	100	5.00

Solution	Conc. (g/Kg)	Height (mV)
2	0.250	12.73
3	0.500	23.60
4	1.00	46.00
5	2.50	113.7
6	3.75	166.6
7	5.00	208.3



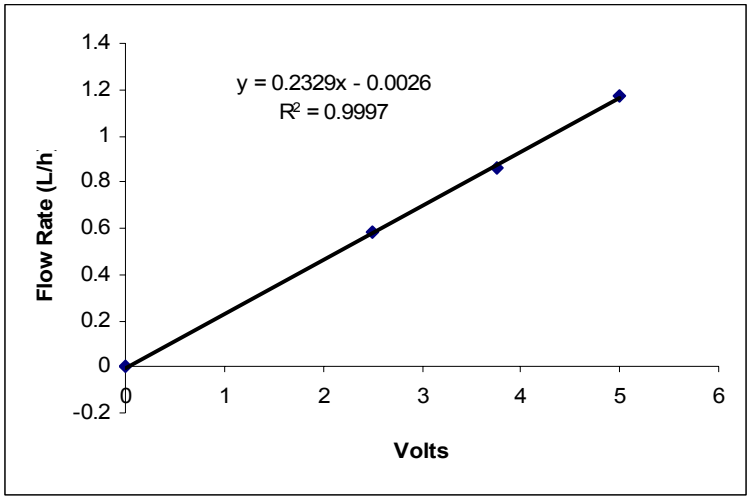
1.3 MASS FLOW CONTROLLER (HASTINGS)

Voltage	Time	Volume	Flow Rate
(V)	(s)	(L)	(L/min)
2.50	50	0.5	0.60
	105	1	0.57
	104	1	0.58
		AVERAGE	0.58

Voltage	Time	Volume	Flow Rate
(V)	(s)	(L)	(L/min)
3.75	69	1	0.87
	70	1	0.86
		AVERAGE	0.86

Voltage	Time	Volume	Flow Rate
(V)	(s)	(L)	(L/min)
5.00	25	0.5	1.20
5.00	26	0.5	1.15
5.00	52	1	1.15
5.00	52	1	1.15
		AVERAGE	1.17

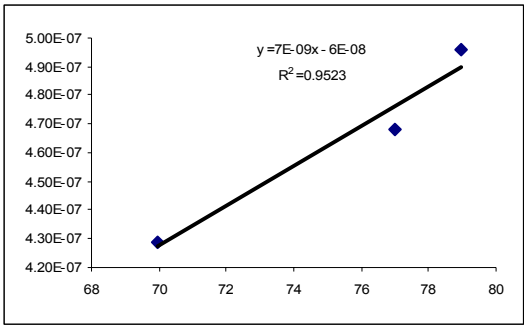
Voltage	Flow Rate
(V)	(L/min)
0	0
2.50	0.580
3.75	0.860
5.00	1.17



1.4 MASS SPECTROMETER

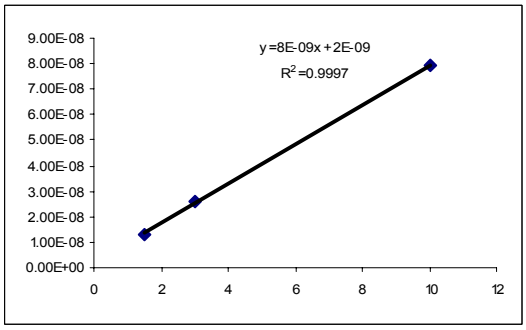
Nitrogen

Molar Fraction	Pressure m/z=28 (torr)
69.97	4.29E-07
77.00	4.68E-07
78.96	4.96E-07



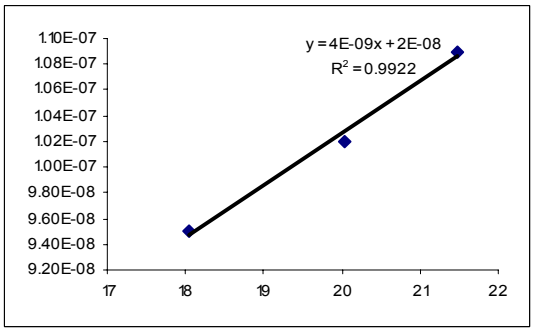
Carbon Dioxide

Molar Fraction	Pressure m/z=44 (torr)
1.51	1.32E-08
3.00	2.59E-08
10.0	7.95E-08



Oxygen

Molar Fraction	Pressure m/z=44 (torr)
18.04	9.50E-08
20.03	1.02E-07
21.49	1.09E-07



1.5 BIOMASS

1.5.1 Method

Biomass Calibration

1 Place 6 numbered filters (45 um) at 105ºC until constant weight

2 Weight the filters in the analytical balance

3 Take 50 mL of fermenter's broth and keep in ice

4 Filter a constant weight of broth into filter 1

5 Wash the filter with distilled water

6 Repeat 4 and 5 for filters 2-6

7 Prepare standard solutions at different Optical Densities from the removed broth:

$$C_i V_i = C_f V_f \Leftrightarrow DO_i V_i = DO_f V_f$$

$$V_i = \frac{DO_f . V_f}{DO_i}, \ \mathrm{DO_f}=0.1,\ 0.2,\ 0.4,\ 0.6,\ 0.8 \ and \ 1$$

$$V_f = 50 \ or \ 100 mL$$

8 Calculate

$$X_{Culture} = \frac{Average\ Cell\ Dry\ Weight}{Filtered\ Volume}$$

$$S\tan dard\ Sol\ Conc = \frac{X_{Culture} . V_i}{V_f}$$

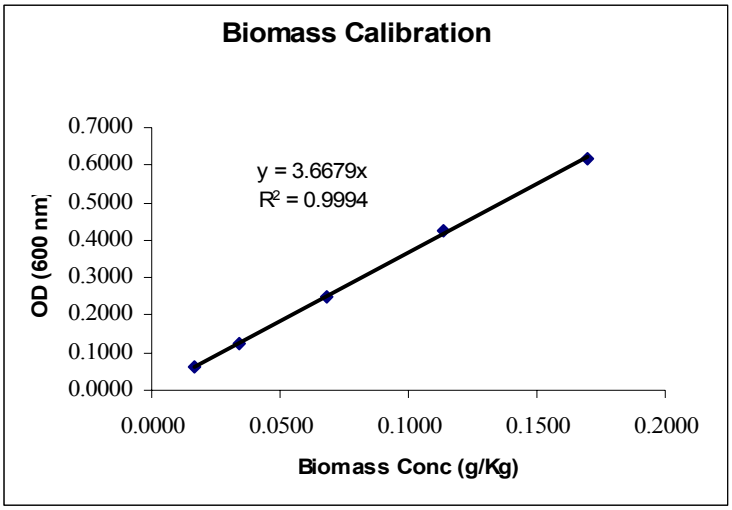
9 Build a chart X versus DO

10 Select linear region and make linear regression

1.5.2 Calibration

Filter	Weight (g)				Biomass Conc. (g/kg)
	Filter	Filter + Biomass	Biomass	Total Filtered	
1	0.0648	0.09550	0.0307	3.0000	10.220
2	0.0639	0.1210	0.0571	3.0000	19.033
3	0.0629	0.1207	0.0578	3.0000	19.267
4	0.0626	0.1211	0.0585	3.0000	19.500
5					
		Average	0.051015		17.005

Flask	Volume (mL)		Biomass Conc	Optical Density
	Added	Flask	g/kg	
1	1	1000	0.01700	0.06300
2	1	500	0.03400	0.1255
3	1	250	0.06800	0.2505
4	1	150	0.1134	0.4250
5	1	100	0.1701	0.6170
6	1	50	0.3401	1.131

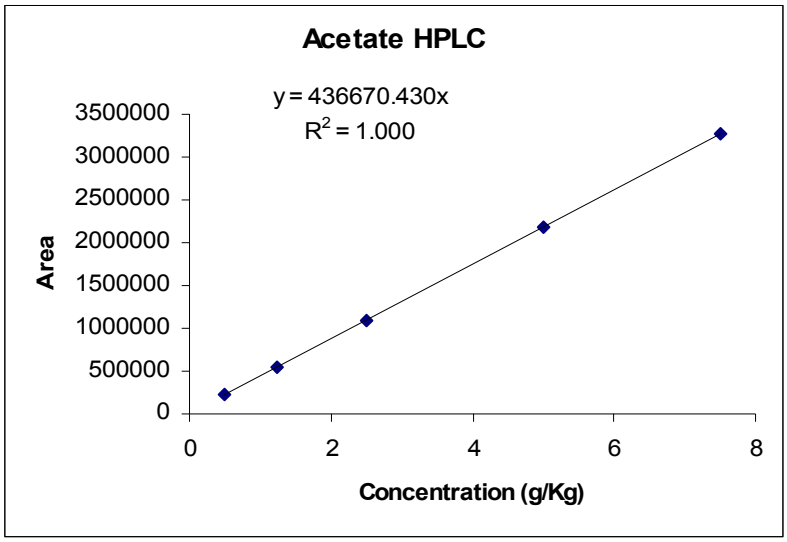


1.6 ACETATE MEASUREMENTS WITH HPLC

Solution	Added Mass	From solution	Dilution	Final Mass	Conc. (g/Kg)
1	Stock solution			100	50.0
2	1.0	1	100	100	0.500
3	2.5	1	40	100	1.25
4	5.0	1	20	100	2.50
5	10	1	10	100	5.00
6	15	1	6.7	100	7.50

Results (RI detector)

Solution	Conc. (g/Kg)	Area 1	Area 2	Average Area
2	0.500	222432	230519	226476
3	1.25	544540	543896	544218
4	2.50	1097574	1084593	1091084
5	5.00	2183527	2196598	2190062
6	7.50	3273137	3267822	3270480

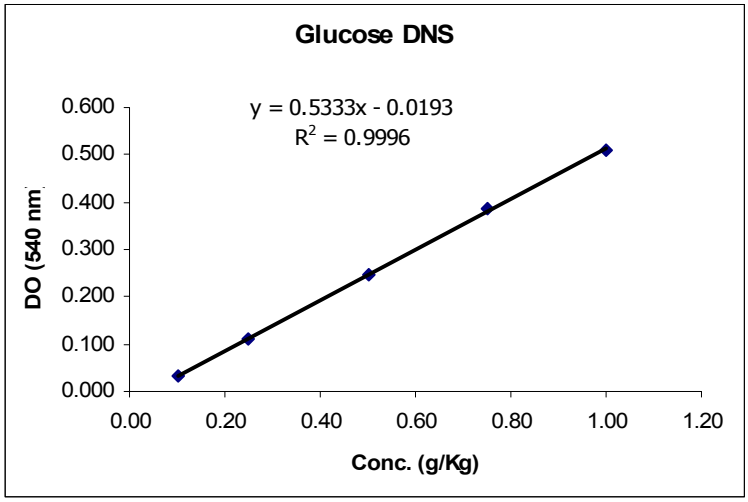


1.7 GLUCOSE MEASUREMENT WITH DNS

Solution	Added Mass	from solution	Dilution	Final Mass	Conc. (g/Kg)
1	Stock solution			100	5.0
2	2	1	50	100	0.10
3	5	1	20	100	0.25
4	10	1	10	100	0.50
5	15	1	6.7	100	0.75
7	20	1	5	100	1.0

Results

Solution	Conc. (g/kg)	OD1	OD2	OD
2	0.10	0.0330	0.0350	0.0340
3	0.25	0.110	0.116	0.113
4	0.50	0.245	0.246	0.246
5	0.75	0.396	0.379	0.388
7	1.0	0.500	0.520	0.510



2. Communication Parameters

In this section, the communication parameters between the components of the set-up are depicted. The pin connection is shown for the communication cables, and the nomenclature used is the following:

A/O Analog Output
A/I Analog Input
GND Signal Ground
TXD Transmit Data
RXD Receive Data
RTS Request to Send
CTS Clear to Send
DTR Data Terminal Ready
DTS Data Set Ready

Table A.1 Pin Connection between peristaltic pump and the DCU

Pump (15 pins)		DCU (25 pins)	
Pin number	**Function**	**Pin number**	**Function**
2	A/O	2	A/I
10	GND	16	GND

Table A.2 Pin connection between DCU and PC

DCU (9 pins)		PC (25 pins)	
Pin number	**Function**	**Pin number**	**Function**
2	GND	7	GND
3	TXD	3	RXD
4	RXD	2	TXD
8	CTS	20	DTR
9	DTR	5	CTS

Table A.3 Communication parameters defined in the DCU

Host Communication Parameters	Peripheral Communication Parameters
Address: 1	Baud rate: 9600
Baud rate: 9600	Data bits: 8 bits
Data bits: 8 bits	Stop bits: 1
Stop bits: 1	Parity: no parity
Parity: no parity	XOFFT: 2
Break: yes	

Table A.4 Pin connection between the bioreactor's balance and the PC

Balance (9 pins)		PC (25 pins)	
Pin number	**Function**	**Pin number**	**Function**
2	TXD	3	RXD
3	RXD	2	TXD
4	CTS	4	RTS
5	GND	7	GND
6	RTS	5	CTS
7	Xanted with	8	-

Table A.5 Pin connection between the feeding balance and the PC

Balance (9 pins)		PC (25 pins)	
Pin number	**Function**	**Pin number**	**Function**
2	RXD	2	TXD
12	TXD	3	RXD
13	GND	7	GND

Table A.6 Pin connection between the valve manifold and the PC

Valve Manifold (9 pins)		PC (25 pins)	
Pin number	**Function**	**Pin number**	**Function**
2	TXD	3	RXD
3	RXD	2	TXD
5	GND	7	RTS
7	RTS	4	GND

Table A.6 Pin connection between the mass flow controller and the data acquisition card

Mass Flow Controller (15 pins)		**PCI 6024E**	
Pin number	**Function**	**Pin number**	**Function**
6	AO	ch 1	AI
7	GND	ch 1	AI
9	-15 VDC	-	Source
11	+ 15 VDC	-	Source
14	AI	ch1	AO

3. Sensitivity functions

In this Annex, the sensitivity functions for biomass, glucose and acetate described on chapter 5 are presented. These functions were obtained using the Symbolic Math Toolbox of MATLAB. The nomenclature used is the one described on chapter 5.

A) BIOMASS

CASE 1 – S>0; A=0; oxidative regimen

```
dXdqsmax=1/k1*S/(Ks+S)*X
dXdk1=-qsmax/k1^2*S/(Ks+S)*X
dXdKs=-qsmax/k1*S/(Ks+S)^2*X
```

CASE 2 – S=0; A>0

```
dXdqacmax=1/k4*A/(KA+A)*Kia/(Kia+A)*X
dXdk4=-qacmax/k4^2*A/(KA+A)*Kia/(Kia+A)*X
dXdKA=-qacmax/k4*A/(KA+A)^2*Kia/(Kia+A)*X
dXdKia=qacmax/k4*A/(KA+A)/(Kia+A)*X-
qacmax/k4*A/(KA+A)*Kia/(Kia+A)^2*X
```

CASE 3 – S>0; A>0; oxidative regimen

CASE 3.1 – Acetate consumption does not exceed oxidative capacity

```
dXdqsmax=1/k1*S/(Ks+S)*X
dXdk1=-qsmax/k1^2*S/(Ks+S)*X
dXdKs=-qsmax/k1*S/(Ks+S)^2*X
dXdqacmax=1/k4*A/(KA+A)*Kia/(Kia+A)*X
dXdk4=-qacmax/k4^2*A/(KA+A)*Kia/(Kia+A)*X
dXdKA=-qacmax/k4*A/(KA+A)^2*Kia/(Kia+A)*X
dXdKia=qacmax/k4*A/(KA+A)/(Kia+A)*X-
qacmax/k4*A/(KA+A)*Kia/(Kia+A)^2*X
```

```
CASE 3.2 - Acetate consumption exceeds oxidative capacity
dXdqsmax=1/k1*S/(Ks+S)*X-1/KOA/k4*S/(Ks+S)*KOS*X
dXdk1=-qsmax/k1^2*S/(Ks+S)*X
dXdKs=-qsmax/k1*S/(Ks+S)^2*X+qsmax/KOA/k4*S/(Ks+S)^2*KOS*X
dXdqomax=1/KOA/k4*Kio/(Kio+A)*X
dXdKOA=-
qomax/KOA^2/k4*Kio/(Kio+A)*X+qsmax/KOA^2/k4*S/(Ks+S)*KOS*X
dXdk4=-
qomax/KOA/k4^2*Kio/(Kio+A)*X+qsmax/KOA/k4^2*S/(Ks+S)*KOS*X
dXdKio=qomax/KOA/k4/(Kio+A)*X-qomax/KOA/k4*Kio/(Kio+A)^2*X
dXdKOS=-qsmax/KOA/k4*S/(Ks+S)*X

CASE 4 – Fermentation of Glucose
dXdqomax=1/KOS/k1*Kio/(Kio+A)*X-1/KOS/k2*Kio/(Kio+A)*X
dXdKOS=-
qomax/KOS^2/k1*Kio/(Kio+A)*X+qomax/KOS^2/k2*Kio/(Kio+A)*X
dXdk1=-qomax/KOS/k1^2*Kio/(Kio+A)*X
dXdKio=qomax/KOS/k1/(Kio+A)*X-qomax/KOS/k1*Kio/(Kio+A)^2*X-
qomax/KOS/k2/(Kio+A)*X+qomax/KOS/k2*Kio/(Kio+A)^2*X
dXdqsmax=1/k2*S/(Ks+S)*X
dXdk2=-qsmax/k2^2*S/(Ks+S)*X+qomax/KOS/k2^2*Kio/(Kio+A)*X
dXdKs=-qsmax/k2*S/(Ks+S)^2*X

B) GLUCOSE
dSdqsmax=-S/(Ks+S)*X
dSdKs=qsmax*S/(Ks+S)^2*X

C) ACETATE
CASE 2 – S<0; A>0
dAdqacmax=-A/(KA+A)*Kia/(Kia+A)*X
dAdKA=qacmax*A/(KA+A)^2*Kia/(Kia+A)*X
dAdKia=-
qacmax*A/(KA+A)/(Kia+A)*X+qacmax*A/(KA+A)*Kia/(Kia+A)^2*X
```

```
CASE 3 – S>0; A>0; oxidative regimen
CASE 3.2 - Acetate consumption exceeds oxidative capacity
dAdqomax=-1/KOA*Kio/(Kio+A)*X
dAdYOA=qomax/KOA^2*Kio/(Kio+A)*X-qsmax/KOA^2*S/(Ks+S)*KOS*X
dAdKio=-qomax/KOA/(Kio+A)*X+qomax/KOA*Kio/(Kio+A)^2*X
dAdqsmax=1/KOA*S/(Ks+S)*KOS*X
dAdKs=-qsmax/KOA*S/(Ks+S)^2*KOS*X
dAdYOS=qsmax/KOA*S/(Ks+S)*X

CASE 4 – Fermentation of Glucose
dAdk3=qsmax/k2*S/(Ks+S)*X-qomax/YOS/k2*Kio/(Kio+A)*X
dAdqsmax=k3/k2*S/(Ks+S)*X
dAdk2=-
k3*qsmax/k2^2*S/(Ks+S)*X+k3*qomax/YOS/k2^2*Kio/(Kio+A)*X
dAdKs=-k3*qsmax/k2*S/(Ks+S)^2*X
dAdqomax=-k3/YOS/k2*Kio/(Kio+A)*X
dAdYOS=k3*qomax/YOS^2/k2*Kio/(Kio+A)*X
dAdKio=-
k3*qomax/YOS/k2/(Kio+A)*X+k3*qomax/YOS/k2*Kio/(Kio+A)^2*X
```

4. CD Contents

Due to the impracticability of referring all the desired information about the computer programs developed in the body of this thesis, a CD is also presented as part of the thesis. The contents of the aforementioned CD are divided into folders treating different subjects as follows:

Data Acquisition System (LabVIEW VIs)

DCU
- DR VISA.vi
- Initialize VISA.vi
- DS VISA.vi
- Checksum.vi

Valve Manifold
- ADAM 4060 initialize
- Opens channel 1
- Open channel 2
- Open channel 3
- Close all channels
- ADAM4060 Read ChannelsA
- ADAM4060 Close.vi
- Flow_contr.vi

Balances
- Weight.vi
- Weight_feed.vi

Mass Spectrometer
- MS.vi
- MS2.vi

FIA
- FIA Acetate.vi

Supervisory
- Coli Monitor.vi
- Control coli.vi

Simulations (MATLAB)
- Coli_control.m
- Coli_euler.m
- Coli_ident.m
- Coli_optim.m
- Kin.m
- Kin_sens_A.m
- Kin_sens_X.m
- Offgas.m

Departamento de Engenharia Biológica
Universidade do Minho
MMIII