

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

ANTONIO CARLOS LUPERNI HORTA

SISTEMA AUTOMÁTICO DE SUPERVISÃO E CONTROLE  
DE CULTIVOS DE ALTA DENSIDADE CELULAR DE *E. coli*  
RECOMBINANTE

SÃO CARLOS

2011

ANTONIO CARLOS LUPERNI HORTA

SISTEMA AUTOMÁTICO DE SUPERVISÃO E CONTROLE  
DE CULTIVOS DE ALTA DENSIDADE CELULAR DE *E. coli*  
RECOMBINANTE

Tese apresentada ao Programa de Pós-Graduação em Engenharia Química da Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Doutor em Engenharia Química.

Orientadora:

Profa. Dra. Teresa Cristina Zangirolami

Co-orientador:

Prof. Dr. Roberto de Campos Giordano

SÃO CARLOS

2011

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária/UFSCar**

H821sa

Horta, Antonio Carlos Luperni.

Sistema automático de supervisão e controle de cultivos de alta densidade celular de *E. coli* recombinante / Antonio Carlos Luperni Horta. -- São Carlos : UFSCar, 2012.  
170 f.

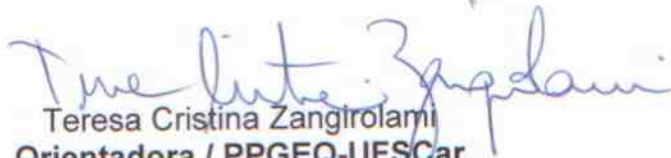
Tese (Doutorado) -- Universidade Federal de São Carlos, 2011.


1. Engenharia química. 2. Sensor de capacitância. 3. Controle de biorreatores. 4. Comitê de redes neurais. 5. Cultivo de alta densidade celular. 6. *Escherichia coli* recombinante. I. Título.

CDD: 660 (20<sup>a</sup>)

MEMBROS DA BANCA EXAMINADORA DA TESE DE DOUTORADO DE **ANTONIO CARLOS LUPERNI HORTA**, APRESENTADA AO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA DA UNIVERSIDADE FEDERAL DE SÃO CARLOS EM 22 DE DEZEMBRO DE 2011.

BANCA EXAMINADORA:


  
Teresa Cristina Zangrolami  
Orientadora / PPGEQ-UFSCar

  
Roberto de Campos Giordano  
Co-orientador / PPGEQ-UFSCar

  
Ana Maria Frattini Fileti  
FEQ/UNICAMP

  
Aldo Tonso  
EP/USP

  
Ruy de Sousa Júnior  
DEQ/UFSCar

  
Antonio José Gonçalves da Cruz  
PPGEQ/UFSCar

## AGRADECIMENTOS

À professora Doutora Teresa Cristina Zangirolami.

Ao professor Doutor Roberto de Campos Giordano.

À minha esposa Cibele e meus filhos, Ana Paula e Gabriel.

Aos meus pais, Paulo e Cassia.

À minha avó Ondina, meu exemplo de sabedoria,

Ao meu irmão mais velho, Paulo, minha referência como pessoa e profissional;

Aos meus irmãos, Ana Maria, Pedro, Luiz, Giordano;

À minha madrasta Dominique.

Aos professores Ruy, Antonio, Alberto, Raquel e à Dra Viviane.

A todos os meus colegas e amigos, Adilson, Cíntia, Ana Maria, Manu, Juliana, Inti, Carlos, Maurício, Alline, Esteban, Lio, Gilson, Gabriel, Gianni, Marina, Dasci, Sandra, Sandrinha, Ivana, Rodrigo, Nestor, Elvis, Sabrina, Edson, Mirela, Chanel, Amadeus, Eudoro, Adilson, Oscar, Evlin, Alcione, Tiago, Daniel, Isis, Osvaldo e todos os demais que também fizeram parte de minha vida durante estes quatro anos.

À CAPES pelo financiamento.

Mesmo que não seja lembrado, espero que meu trabalho seja útil à humanidade.

(Livre do Autor)

## RESUMO

Cultivos de alta densidade celular de *E. coli* recombinante constituem uma técnica rápida e econômica para produção de proteínas recombinantes. Por meio deste bioprocesso, são obtidos produtos de alto valor agregado e de grande importância na indústria farmacêutica, tais como insulina, hormônios de crescimento humano e bovino, antígenos proteicos para formulação de vacinas, enzimas, dentre outros. Entretanto, manter estes cultivos dentro das condições desejadas se torna um grande desafio, em função da dificuldade de controlar variáveis como a concentração de oxigênio dissolvido (COD) e a concentração de substrato nos níveis desejados. Por isso, o desenvolvimento e a implementação de sistemas automáticos de supervisão e controle são requisitos fundamentais para o bom desempenho de um cultivo de alta densidade. O presente trabalho teve como principais objetivos estudar estratégias de alimentação para cultivos de alta densidade celular de *Escherichia coli* recombinante e desenvolver uma ferramenta computacional para suporte na execução das estratégias escolhidas, realizando o monitoramento, controle e supervisão dos cultivos. Os cultivos em batelada alimentada realizados sob supervisão da ferramenta foram conduzidos em biorreator de 5 L, equipado com sensores de temperatura, oxigênio dissolvido, pH, pressão e biomassa (sensor que mede a concentração de células viáveis a partir dos dados de permissividade), bombas peristálticas e conectado a analisador de gases. A ferramenta foi desenvolvida com os programas LabView 8.0 e MatLab 6.5, sendo a aquisição e a comunicação com os diferentes acessórios do biorreator realizada via compact Field Point (National Instruments). Vinte e dois cultivos em batelada alimentada com 5 diferentes clones de *E. coli*, BL21(D3) expressando a enzima penicilina G acilase (PGA) assim como proteínas antigênicas de *Streptococcus pneumoniae* (PspA3, PspA245 e PspA4Pro) e de *Erysipelothrix rhusiopathiae* (SpaA) foram realizados durante o desenvolvimento da ferramenta e dos estudos de estratégia de alimentação, empregando tanto meio definido (HDF modificado) como meio complexo (ZYM-5052 modificado), tendo glicerol ou glicose como principal fonte de carbono e IPTG ou lactose como indutores. Em todos os cultivos, amostras foram coletadas para quantificar a concentração de células (método de massa seca em filtro

de 0,22 $\mu$ m e leitura da densidade ótica a 600 nm), de ácidos orgânicos, glicose, glicerol e lactose (HPLC) e a expressão da proteína (densitometria e método NIPAB para a PGA) e a estabilidade de plasmídeo (plaqueamento). A ferramenta SUPERSYS\_HCDC<sup>®</sup> (registrada como software livre) desenvolvida, implementada e validada nos cultivos realizados, desempenha as funções básicas de softwares de supervisão de biorreatores, tais como: monitoramento e aquisição de dados de pressão, temperatura, pH, COD, fração de CO<sub>2</sub> e de O<sub>2</sub> nos gases de saída; estimativa em tempo real do quociente respiratório, das velocidades de consumo de oxigênio e de produção de CO<sub>2</sub>. Esta ferramenta apresenta as seguintes funcionalidades especiais: i) controle automático das vazões de ar e de oxigênio de acordo com a demanda celular; ii) acionamento automático da bomba de alimentação ao final da batelada; iii) controle automático da vazão de alimentação em função da velocidade específica de crescimento inferida em tempo real; iv) controle automático da alimentação com restrições pela concentração de oxigênio dissolvido; v) alarmes sonoros indicando falhas no processo; vi) envio de mensagens de falhas por email; vii) controle automático da concentração de oxigênio dissolvido; viii) controle de segurança da pressão do biorreator, e ix) controle da temperatura do banho. Em relação aos estudos das estratégias de alimentação visando ao aumento da produtividade em biomassa em cultivos de alta densidade celular de *E. coli* recombinante, com o auxílio da ferramenta de supervisão desenvolvida aliada a modificações na composição do meio de cultivo sintético disponível na literatura, foram alcançadas concentrações celulares maiores que 150 g/L em menos de 24 h de tempo total de cultivo, levando a uma produtividade de 9,2 g/Lh, a qual é superior aos valores relatados na literatura, sem acúmulo de acetato e possibilitando elevada produção da proteína recombinante.

**Palavras-chave:** Sensor de Capacitância. Controle de Biorreator. Comitê de Redes Neurais. *Escherichia coli* Recombinante. Identificação das Fases de Crescimento. Cultivo de Alta Densidade Celular.



## ABSTRACT

High cell density cultivations of recombinant *E. coli* are a fast and economical way to produce recombinant proteins. Through this bioprocess, products with high added value and pharmaceuticals of great importance such as insulin, human and bovine growth hormone, protein antigens for formulation of vaccines, enzymes, among others, are obtained. However, keeping these cultivations within the desired conditions becomes a major challenge, since some variables such as dissolved oxygen concentration (DOC) and substrate concentration are difficult to control. Therefore, the development and implementation of an automatic monitoring and control tool are key requirements for the performance of high density cultivation. The present work has as main objectives to study feeding strategies for high cell density cultivation of recombinant *Escherichia coli* and develop a computational tool capable of ensuring the implementation of the chosen strategies, performing the monitoring, control and supervision of the cultivations. Fed batch cultivations were carried out under the supervision of the tool in a 5 L in-house bioreactor, equipped with sensors for temperature, dissolved oxygen, pH, pressure and biomass (sensor that measures the concentration of viable cells based on permittivity measurements), peristaltic pumps and connected to the gas analyzer. The tool was developed with LabView 8.0 and MatLab 6.5, being the acquisition and communication with the different bioreactor accessories via compact Field Point. Twenty two fed-batch cultivations with 5 different clones of *E. coli*, BL21(D3) expressing the enzyme penicillin G acylase (PGA) as well as antigenic proteins of *S. pneumoniae* (PspA3, PspA245 and PspA4Pro) and *E. rhusiopathiae* (SpaA) were performed during the development of the tool and the studies of feeding strategy. Both defined medium (HDF modified) as complex medium (ZYM-5052 modified), usually having glycerol as main carbon source and IPTG or lactose as inducers were used. In all cultivations, samples were collected to quantify the concentration of cells (dry weight method in filter of 0.22  $\mu\text{m}$  and optical density at 600 nm), organic acids, glucose, glycerol and lactose (HPLC) as well as protein expression (densitometry and NIPAB method for PGA) and plasmid stability (plating). The tool SUPERSYS\_HCDC<sup>®</sup> (registered as a free software) developed, implemented and

validated in the performed cultivations, carries out the basic functions of bioreactor supervision software, such as monitoring and data acquisition of pressure, temperature, pH, DOC, fraction of CO<sub>2</sub> and O<sub>2</sub> in the outlet gas as well as real-time estimate of the respiratory quotient, the rate of oxygen consumption and CO<sub>2</sub> production. However, it also has the following special features, including: i) automatic control of air and oxygen flow according to cellular demand, ii) automatic activation of the feed pump at the end of the batch; iii) automatic control of feeding flow rate as function of the specific growth rate inferred in real time; iv) automatic control of feeding flow rate constrained by the concentration of dissolved oxygen, v) audible alarms indicating failures in the process; vi) failure messages sent via email; vii) automatic control of dissolved oxygen concentration; viii) control of the bioreactor pressure; and ix) control of bath temperature. Regarding the studies of feeding strategies aimed at biomass productivity increase in high cell density cultivations of recombinant *E. coli*, using the supervision tool developed together with changes in the composition of the synthetic culture medium available in the literature, a cellular concentrations greater than 150 g/L was achieved in less than 24 hours of cultivation, corresponding to a productivity of 9.2 g/Lh. This value, which is higher than the reported in the literature, was obtained without acetate accumulation and allowing high production of recombinant protein.

**Key words:** Capacitance Sensor. Bioreactor Control. Neural Network Committee. Recombinant *Escherichia coli*. Growth Phase Identification. High Cell Density Cultivation.

## LISTA DE FIGURAS

Figura 2.1.1. Distribuição das técnicas óticas e espectroscópicas no espectro eletromagnético. ....	24
Figura 2.2.1. Diagrama do método de <i>feedback</i> (loop simples) indireto. ....	34
Figura 2.2.2. Diagrama da estratégia de controle da vazão de alimentação baseada em demanda ou consumo de substrato. ....	35
Figura 2.3.1. Telas do programa BioCommand® .....	40
Figure 4.1.1. Experimental set-up: "A" - the computer program SUPERSYS_HCDRC, "B"- pH meter, "C" - exhaust gas analyzer, "D" - cFP; "E" and "F" - acid and base pumps, "G" - dissolved oxygen sensor, "H" - mass flow controller of air supply, "I" - mass flow controller of oxygen supply, "J"- biomass sensor; "K" - stirring speed controller, "L" - feed pump, "M" - cooling/heating water bath, "N"- bioreactor, "O" - pressure transducer and "P" – temperature transducer. ....	50
Figure 4.1.2. Pseudocode for combined control of feed flow rate .....	53
Figure 4.1.3. Pseudo-code for the dissolved oxygen concentration control module. ....	54
Figure 4.1.4. Pseudo-code for the control of air and oxygen flow rates .....	54
Figure 4.1.5. Pseudo-code for the set of alarms. ....	55
Figure 4.1.6. Fed-batch cultivation of <i>E. coli</i> BL21(D3) expressing a fragment of SpaA antigenic protein. Culture carried out with auto-induction complex medium, containing glucose, glycerol and lactose as carbon sources, at 37 °C (Silva et al, 2011). a) permittivity (perm – pF/cm), carbon evolution rate (CER – mmol/min), and cellular concentration (Cx – g/L); b) CER (mol/min), oxygen uptake rate (OUR – mol/min) and respiratory quotient (RQ); c) specific growth rate ( $\mu - h^{-1}$ ). ....	57
Figure 4.1.7. Fed-batch cultivation 1 of <i>rE. coli</i> BL21(DE3) expressing a fragment of PspA3 antigenic protein. Culture carried out with modified HDF medium using glycerol as carbon source, at 30°C (Sargo et al, 2011); A) Stirrer speed (Agit -rpm), dissolved oxygen concentration (DOC - %saturation); B) Air flow rate (Qair – L/min), oxygen flow rate (QO2 – L/min) and stirrer speed. Culture carried out with auto-induction complex medium, containing glucose, glycerol and Lactose. ....	58
Figure 4.1.8. Typical time profiles of stirring speed, feed flow rate, air and oxygen flow rates and other variables for the combined DOC control. Fed-batch cultivation 2 of	

*rE. coli* BL21(DE3) expressing a fragment of PspA3 antigenic protein. Culture carried out with modified HDF medium using glycerol as carbon source, at 35°C (Sargo et al, 2011). A) air and oxygen flow rates in L/min; B) feed flow rate in L/min; C) stirrer speed in rpm; D) dissolved oxygen concentration as % of saturation; E) pH; F) biomass concentration on-line estimated from permittivity data and measured off-line by dry weight method; G) on-line specific growth rate inferred from both permittivity data and dry weight measurements; H) substrate concentration (glycerol); I) cumulative volume of NH<sub>4</sub>OH added; J) acetate concentration in g/L. Induction with IPTG at 16.5h. .... 60

Figure 5.1.1 Schematic view of the bioreactor experimental set-up..... 73

Figure 5.1.2 Schematic representation of the approach used to design the AI system . 74

Figure 5.1.3 NN training flowchart. Exp refers to experimental data ..... 77

Figure 5.1.4 NNC construction and test configuration..... 78

Figure 5.1.5 Exp 2: Experimental data for *rE. coli* expressing PspA3 cultivation. (a) Online signals: medium permittivity (Perm), feed flow rate (F); State variables: optical density (OD, measured at-line), and glycerol concentration (C<sub>S</sub>, measured off-line through HPLC); (b) online signals: agitation (Agit) and carbon dioxide evolution rate (CER); (c) Zoom around the start-up of feeding; HEX is the human expert dummy variable (the desired output of the AI system)..... 80

Figure 5.1.6 Validation of the AI systems: Time delay ( $\Delta t < 0$ ) or advance ( $\Delta t > 0$ ) of the start-up of the supplementary medium pump. (a) eight GNNs and NNCs with better performance; (b) five best AI systems ..... 82

Figure 5.1.7 Validation results for committee NNC\_34. a) Exp5; b) Exp6 c) Exp1\_Agit\_0; d) Exp1\_CO2\_pulse; e) Exp1\_CO2\_0; f) Exp1\_Agit\_pulse. Outputs above 0.72 turned on the pump (NNC\_output in the plots). HEX are the human expert decisions to start the feeding, and NNC\_class are the AI decisions (both TRUE when equal to 1) ..... 84

Fig 5.2.1 Pseudocode of the moving average filters implemented in the control and monitoring software of the bioreactor. Code implemented in MatLab 6.5 ..... 96

Fig 5.2.2 Pseudocode for specific growth rate estimation implemented in the control and monitoring software of the bioreactor. Code implemented in MatLab 6.5 ..... 97

Fig 5.2.3 Supplementary medium feed control: $C_{X0}$ , $\mu$ , $\mu_{SET}$ , $V_0$ , $C_{S0}$ , $C_{SR}$ re-tuned online; $Y_{XS}$ , and $m$ corrected at-line .....	98
Fig 5.2.4 Correlation between permittivity [Perm (pF/cm)] and cellular concentration [Cx (g/L)]. Symbols are experimental data, while the line is the linear fit .....	99
Fig 5.2.5 Data obtained using filters with lack of fit ( $S - pF^2/cm^2$ ) and roughness ( $R - pF^2/cm^2$ ), a) Classical moving average (CMA) and b) Smoothed moving average (SMA) .....	100
Fig 5.2.6 Smoothed experimental data; <i>E. coli</i> cultivation. The blue line is the rough signal and the red one is the smoothed signal. a) CMA method, $n=30$ , $R=47.7$ and $S=11882$ ; b) SMA method, $n=30$ , $R=1.2$ and $S=199332$ .....	100
Fig 5.2.7 Cultivation 1: Fast IPTG induction; $\mu_{SET}=0.3 \text{ h}^{-1}$ ; $Y_{XS}=0.5$ ; $m=0.02$ ; A) Cellular concentration measured by DCW (Cx - g/L) and by the smoothed permittivity signal [Cx(perm) - g/L]; B) Specific growth rate calculated from DCW ( $\mu^* - \text{h}^{-1}$ ), and from the smoothed permittivity signal ( $\mu - \text{h}^{-1}$ ); C) Substrate concentration (Cs - g/L) and feed flow rate (F - L/h) .....	101
Fig 5.2.8 Cultivation 2: Fast IPTG induction; $\mu_{SET}=\mu_{DYN}$ ; $Y_{XS}=0.5$ ; $m=0.02$ ; A) Cellular concentration measured by DCW (Cx - g/L) and by the smoothed permittivity signal [Cx(perm) - g/L]; B) Specific growth rate calculated from DCW ( $\mu^* - \text{h}^{-1}$ ), and from the smoothed permittivity signal ( $\mu - \text{h}^{-1}$ ); C) Substrate concentration (Cs - g/L) and feed flow rate (F - L/h) .....	103
Fig 5.2.9 Cultivation 3: Fast IPTG induction; $\mu_{SET}=\mu_{DYN}$ ; re-tuning yield and maintenance coefficients: $0.1 < Y_{XS} < 0.5$ ; $0.01 < m < 0.5$ ; A) Cellular concentration measured by DCW (g/L) and by the smoothed permittivity signal; B) Specific growth rate calculated from DCW ( $\mu^*$ ), and from the smoothed permittivity signal; C) Substrate concentration (g/L) and feed flow rate (L/h); D) Dissolved oxygen concentration (DOC - % of saturation) .....	104
Fig 5.2.10 Cultivation 4: Slow lactose induction; $\mu_{SET}=\mu_{DYN}$ ; $0.1 < Y_{XS} < 0.5$ ; $0.01 < m < 0.5$ ; A) Cellular concentration measured by DCW (g/L) and measured by the smoothed permittivity signal; B) Specific growth rate calculated from DCW ( $\mu^*$ ), and calculated from the smoothed permittivity signal; C) Substrate concentration (g/L)	

and feed flow rate (L/h); D) Yield coefficient ( $Y_{XS} - g_X/g_S$ ), and maintenance ( $m - h^{-1}$ )	105
.....	105
Fig 5.2.11 Cell productivity profiles, g/(L.h): 1) using $\mu_{SET} = 0.3$ ; 2) using $\mu_{DYN}$ ; 3) using $\mu_{DYN}$ , and re-tuning $Y_{XS}$ and $m$ ; 4) same as in (3) but with induction with lactose.	107
Fig 5.2.12 Image of a gel from Cultivation 3 (induction with IPTG). The first lane corresponds to the standard, the second lane to the sample immediately before induction, the third one is a sample immediately after the addition of IPTG, and the other are in chronological order after induction.....	107
Fig 5.2.13 Profiles of PspA3 protein expression ( $C_p$ in g/L <sub>reactor</sub> ) in Cultivation 3 (induced with IPTG) and 4 (induced with lactose). Error bars are standard deviations of triplicate measurements.....	108
Figure 5.3.1. Correlation between permittivity [Perm (pF/cm)] and cellular concentration [ $C_x$ (g <sub>DCW</sub> /L)] for <i>B. megaterium</i> , <i>S. cerevisiae</i> and <i>Pichia pastoris</i> . Symbols correspond to experimental data, lines correspond to linear fit.....	119
Figure 5.3.2. Correlation between permittivity [Perm (pF/cm)] and cellular concentration [ $C_x$ (g <sub>DCW</sub> /L)] for different recombinant <i>Escherichia coli</i> strains. Symbols correspond to experimental data, lines correspond to linear fit.....	120
Figure 5.3.3. In the left $\mu$ ( $h^{-1}$ ) calculated online from smoothed permittivity (pF/cm), in the right permittivity (continuous line) and smoothed permittivity (dotted line). A and B: <i>Pichia pastoris</i> (Montaño, 2010). C and D: <i>Escherichia coli</i> "b". E and F: <i>Bacillus megaterium</i> (Suárez, 2010). G and H: <i>Saccharomyces cerevisiae</i> . ....	123
Figura A-1. Telas principais do programa de computador SUPERSYS_HCDC. A) tela inicial de escolha das configurações dos instrumentos a serem utilizados; B) tela de visualização dos gráficos e operação manual das vazões de gases e da agitação; C) tela de configurações e ativação dos controles automáticos. ....	144
Figura A-2. Esquema dos instrumentos e sistema de monitoramento e controle dos cultivos. A: computador com o programa de aquisição e controle SUPERSYS_HCDC; B: pHmetro; C: analisador de gases; D: cFP (Compact Field Point – National Instruments); E e F: bombas de ácido e Base; G: sensor de oxigênio dissolvido; H: fluxômetro de ar; I: fluxômetro de oxigênio; J: sensor de	

biomassa (Fogale nanotech); K: controlador da velocidade de agitação do reator; L: bomba de alimentação, M: banho termostatizado. ....	145
Figura A-3. Arquitetura da RN utilizada para reconhecimento das fases do cultivo. ...	148

## LISTA DE QUADROS

Quadro 2.2.1. Densidades celulares atingidas com diferentes estratégias em cultivos de <i>E. coli</i> .....	38
Quadro A-1. Pseudocódigo do controle automático das vazões de ar e oxigênio.....	150
Quadro A-2. Pseudocódigo do filtro de média móvel e cálculo da velocidade específica de crescimento. ....	151
Quadro A-3. Pseudocódigo do cálculo do volume do reator.....	152

## LISTA DE TABELAS

Tabela 2.1.1. Valores médios de diferentes sistemas de medição de biomassa.....	31
Tabela 3.2.1. Composição do meio de cultivo .....	42
Tabela 3.2.2. Identificação dos principais cultivos .....	42
Table 4.1.1. Medium composition .....	49
Table 5.1.1. Heterologous antigen, expression system, operational conditions and media composition used in bioreactor cultivations .....	71
Table 5.1.2. Tested NN architectures .....	76
Table 5.1.3 Best AI system configurations after validation.....	83
Table 5.2.1. Medium composition .....	94
Table 5.2.2. Feeding strategy definition.....	94
Table 5.3.1. Optimal length of SMA filter smoothing vectors for the different microorganisms. ....	121



## LISTA DE ABREVIações E SÍMBOLOS

C: Capacitância	(F)
$C_X$ : Concentração celular	(g/L)
$C_{X0}$ : Concentração celular no início da alimentação	(g/L)
$C_S$ : Concentração de substrato	(g/L)
$C_{SR}$ : Concentração de substrato residual	(g/L)
$C_P$ : Concentração de produto	(g/L)
$CO_2$ : Gás carbônico	
COD: Concentração de Oxigênio Dissolvido	(% da saturação)
CADC: Cultivo de alta densidade celular	
DO: Densidade óptica	
F: Vazão de alimentação de meio suplementar	(L/h)
F': Vazão de retirada de meio	(L/h)
G: Condutância	(S)
IA: Inteligência Artificial	
m: Coeficiente de manutenção celular	(h <sup>-1</sup> )
$O_2$ : Oxigênio	
PS: Polissacarídeo capsular	
RN: Rede neural	
RQ: Quociente respiratório	
t: tempo	(h)
T: temperatura	(°C)
V: volume do biorreator	(L)
$V_0$ : volume do biorreator no início da alimentação	(L)
$Y_{PS}$ : Fator de formação de produto a partir do substrato	(g <sub>p</sub> /g <sub>s</sub> );
$Y_{XS}$ : Fator de conversão de substrato a células	(g <sub>x</sub> /g <sub>s</sub> );
$\mu$ : Velocidade específica de crescimento	(h <sup>-1</sup> )
$\mu_M$ : Velocidade específica máxima de crescimento	(h <sup>-1</sup> )
$\mu_{DYN}$ : Velocidade específica dinâmica de crescimento	(h <sup>-1</sup> )
$\mu_{CRIT}$ : Velocidade específica crítica de crescimento	(h <sup>-1</sup> )
$\mu_{SET}$ : Velocidade específica de crescimento ajustada	(h <sup>-1</sup> )

# SUMÁRIO

<b>RESUMO</b> .....	<b>5</b>
<b>ABSTRACT</b> .....	<b>7</b>
<b>1. Introdução</b> .....	<b>19</b>
1.1. Objetivos .....	21
1.2. Estrutura Da Tese .....	21
<b>2. Revisão bibliográfica</b> .....	<b>23</b>
2.1. Monitoramento da concentração de biomassa .....	23
2.1.1 Método ótico .....	25
2.1.2 Espectroscopia de infravermelho .....	26
2.1.3 Fluorescência .....	27
2.1.4 Espectroscopia dielétrica .....	27
2.2. Estratégias de controle da alimentação .....	32
2.3. Programas de monitoramento e controle de biorreatores .....	39
<b>3. Materiais e Métodos</b> .....	<b>41</b>
3.1. Microrganismos .....	41
3.2. Meio de cultivo .....	41
3.3. Preparação do inóculo .....	42
3.4. Cultivos .....	42
3.5. Métodos analíticos .....	43
3.5.1. Concentração celular .....	43
3.5.2. Concentração de açúcares e ácidos orgânicos .....	44
<b>4. Sistema de automação para cultivos de alta densidade celular: SUPERSYS_HCDC</b> .....	<b>45</b>
4.1. A supervision and control tool based on artificial intelligence for high cell density cultivations .....	46
4.1.1. Introduction .....	47

4.1.2. Materials and Methods.....	48
4.1.3. Results and Discussion.....	55
4.1.4. Conclusions.....	61
Nomenclature.....	62
References.....	63
<b>5. Monitoramento e controle de cultivos de alta densidade celular.....</b>	<b>66</b>
<b>5.1. Robust artificial intelligence tool for automatic start-up of the supplementary medium feeding in recombinant <i>E. coli</i> cultivations .....</b>	<b>68</b>
5.1.1 Introduction .....	69
5.1.2 Materials and Methods.....	71
5.1.3 Methodology.....	73
5.1.4 Results and discussion .....	79
5.1.5 Conclusion .....	85
References.....	86
<b>5.2. Intensification of high cell density cultivations of rE. coli for production of <i>S. pneumoniae</i> antigenic surface protein, PspA3, using model-based adaptive control ...</b>	<b>89</b>
5.2.1 Introduction .....	90
5.2.2 Materials and Methods.....	93
5.2.3 Results and Discussion.....	99
5.2.4 Conclusion .....	108
References.....	109
<b>5.3. On-line monitoring of biomass concentration based on a capacitance sensor: assessing the methodology for different bacteria and yeast high cell density fed-batch cultures.....</b>	<b>112</b>
5.3.2. Introduction .....	113
5.3.3. Materials and methods .....	115
5.3.4. Results and Discussion .....	118
5.3.5. Conclusion .....	124
References.....	125
<b>6. Conclusão .....</b>	<b>128</b>
<b>7. Sugestões para trabalhos futuros.....</b>	<b>130</b>
<b>Referências .....</b>	<b>132</b>

**Apêndice A - Relatório encaminhado para registro do programa de computador  
SUPERSYS\_HCDC..... 140**

**Anexo A – Certificado de Registro do Programa SUPERSYS\_HCDC ..... 170**

## 1. INTRODUÇÃO

A Indústria biotecnológica é uma das responsáveis pela crescente evolução em tecnologias de fármacos e melhorias na qualidade de vida da população mundial. Novas vacinas são lançadas, novos antibióticos, suplementos, enzimas, medicamentos e até mesmo novos processos para produção de combustíveis estão continuamente surgindo graças aos avanços da biologia molecular e ao desenvolvimento de bioprocessos para cultivo dos organismos recombinantes em altas concentrações celulares (SHILOACH; FASS, 2005; ROCHA, 2003; VOJINOVIC *et al.*, 2006).

A bactéria *Escherichia coli* é o organismo mais usado na expressão de proteínas heterólogas, por apresentar capacidade de crescer em alta velocidade, atingindo altas densidades celulares em meios de cultivo simples e baratos (BARAZZONE *et al.*, 2011; BANEYX, 1999). Para alcançar altas concentrações celulares, as condições de cultivo devem permitir o crescimento balanceado. A fonte de carbono escolhida, a composição do meio de alimentação, o pH, a temperatura e a disponibilidade de oxigênio dissolvido são fatores determinantes da velocidade de crescimento, da conversão de nutrientes em células e, conseqüentemente, da produtividade em biomassa.

Entretanto, a execução de um cultivo de alta densidade celular (CADC) é um desafio até para pesquisadores experientes, já que variáveis como a concentração de oxigênio dissolvido e o fornecimento de substrato são difíceis de controlar. A dificuldade de controlar essas variáveis pode levar à inibição do crescimento e da expressão da proteína de interesse devido à ocorrência de dois problemas comumente observados em CADC: 1) acúmulo de nutrientes como glicose, amônia, ferro, atingindo concentrações tóxicas e 2) acúmulo de subprodutos metabólicos como acetato, formiato, propionato, lactato e etanol (LEE, 1996; ZHANG; GREASHAM, 1999; SUÁREZ; KILIKIAN, 2000). Para evitar a inibição por substrato, a concentração dos mesmos deve ser mantida em níveis inferiores aos inibitórios durante todo o cultivo. Por isso, o modo de operação do biorreator em batelada alimentada é estratégia frequentemente empregada.

Desde 2007, cultivos em biorreator de *E. coli* e outros recombinantes para produção de diferentes proteínas heterólogas vêm sendo realizados no Laboratório de

Desenvolvimento e Automação de Bioprocessos em função da diversificação dos projetos de pesquisa envolvendo biologia molecular e engenharia genética. O primeiro cultivo de *E. coli* recombinante no LaDABio foi realizado pela aluna de Doutorado Ana María Vélez utilizando o plasmídeo pT101/D-TOPO com o gene PAC, capaz de super-expressar a enzima penicilina G acilase (PGA), obtido por doação, junto ao Laboratório de Biocatálisis do ICP-CSIC-Madri, Espanha, que tem um papel fundamental na produção de penicilinas semi-sintéticas pela indústria farmacêutica e está inserido no projeto de pesquisa da FINEP, coordenado pela Profa. Dra. Raquel de Lima Camargo Giordano (VELEZ *et al.*, 2010).

Em 2008, foi adquirido o sensor de capacitância (Biomass System da Fogale Nanotech) para medida em tempo real da concentração de biomassa e utilização nos experimentos do projeto de pesquisa "Produção de vacinas recombinantes contra erisipela suína". No âmbito desse projeto, foram iniciados os cultivos de outro clone (*E. coli* BL21(DE3) pET28a\_spaA) por Adilson Silva expressando a proteína de superfície SpaA (Surface protective antigen A), a qual é reconhecida como um dos principais agentes antigênicos de *E. rhusiopathiae*, bactéria causadora da erisipela suína, apresentando potencial de aplicação veterinária como vacina de sub-unidade antigênica contra essa enfermidade (SILVA *et al.*, 2009; IMADA, 1999). O sensor de biomassa utiliza as propriedades dielétricas das células vivas, e tem a vantagem de possibilitar a medida da concentração de células viáveis.

Finalmente, em 2009, com a aprovação do Projeto Temático FAPESP 08/05207-4, "Vacina conjugada anti-Pneumocócica: Estudos sobre a viabilidade de uma vacina Polissacarídeo – Proteína de superfície de pneumococo A", coordenado pela Dra. Martha Tanizaki do Instituto Butantan, foram iniciados, por Antonio Carlos Luperni Horta e Cintia Regina Sargo, os cultivos dos clones BL21(DE3) PspA3/pET37b+, PspA245/pET37b e PspA4Pro/pET37b que super-expressam fragmentos de proteínas de superfície PspA presentes nas células de *S. pneumoniae*, bactéria causadora de várias infecções do trato respiratório em crianças e adultos portadores de imunodeficiências (KORNELISSE *et al.*, 1995; BEREZIN *et al.*, 2002; BRANDILEONE *et al.*, 2003; YOTHER, 2004; BARAZZONE *et al.*, 2011). O referido projeto temático conta com a participação de vários grupos de pesquisa, cabendo ao LaDABio a responsabilidade de cultivar os clones de *rE. coli* expressando proteínas dos diferentes clados gerados pelo grupo de Biologia Molecular, com o objetivo de fornecer biomassa contendo uma quantidade significativa de proteína para ser

processada nas etapas posteriores de purificação e conjugação, as quais são conduzidas por outros colaboradores no âmbito do projeto.

Apesar da diversidade de proteínas sendo produzidas pelos diferentes clones de *E. coli* recombinante cultivados no LaDABio, todos os clones apresentavam três importantes características em comum: a linhagem BL21(DE3) de *E. coli*, o promotor T7/lac e o vetor pET foram empregados em todas as construções (SARGO, 2011; SILVA, 2011; VELEZ *et al.*, 2010).

## 1.1. OBJETIVOS

O desenvolvimento do presente trabalho ocorreu de forma integrada aos projetos descritos anteriormente, contribuindo para a transferência de conhecimento e de soluções de um cultivo para o outro. Os principais objetivos desta tese de doutorado foram:

- 1) Estudar estratégias de alimentação que permitam alcançar altas concentrações celulares com máxima produtividade em biomassa e formação mínima de metabólitos;
- 2) Implementar estratégias de controle baseadas no sensor de capacitância e avaliar o desempenho do mesmo no monitoramento da concentração celular, da viabilidade celular e da velocidade específica de crescimento ao longo de todas as etapas do bioprocessamento;
- 3) Desenvolver uma ferramenta de supervisão preparada para auxiliar o operador humano (ou até substituí-lo) na tomada de decisões relacionadas ao controle de variáveis críticas para o desempenho do bioprocessamento.

## 1.2. ESTRUTURA DA TESE

A tese foi organizada em sete seções, incluindo a presente introdução e os objetivos como seção 1. A seção 2 contém uma revisão bibliográfica sobre os principais assuntos tratados ao longo do trabalho. A seção 3 detalha os materiais e métodos. A seção 4 apresenta, em forma de artigo, o programa de computador desenvolvido e

registrado no INPI sob o no. 11008-6. A seção 5 apresenta três artigos que detalham o desenvolvimento de algumas ferramentas especiais do programa de computador e o funcionamento do sensor de biomassa. O primeiro artigo foi publicado na revista *Bioprocess and Biosystems Engineering* (DOI 10.1007/s00449-011-0540-0) e descreve o uso de inteligência computacional para reconhecimento do final da batelada e início da fase de alimentação, permitindo o acionamento automático da bomba de alimentação. O segundo artigo (submetido a *Bioprocess and Biosystems Engineering* em novembro de 2011) detalha a avaliação de diferentes estratégias de alimentação com perfil exponencial, incluindo a estimativa online da velocidade específica de crescimento e o ajuste *em tempo real* dos coeficientes de manutenção e de rendimento em biomassa. O terceiro artigo está em fase de conclusão e descreve a utilização do sensor de biomassa para estimativa da concentração celular e inferência da atividade fisiológica em cultivos de alta densidade celular de diferentes microrganismos. A seção 6 reúne as principais conclusões sobre o trabalho desenvolvido, e a seção 7 relaciona as principais sugestões para trabalhos futuros.



## 2. REVISÃO BIBLIOGRÁFICA

Esta seção apresenta uma revisão bibliográfica sintetizando os assuntos mais relevantes sobre as tecnologias utilizadas neste trabalho, que são: o monitoramento de biomassa, as estratégias de controle de alimentação em batelada alimentada, bem como os programas de operação de biorreatores.

### 2.1. MONITORAMENTO DA CONCENTRAÇÃO DE BIOMASSA

A maioria dos autores concorda que a concentração de biomassa é uma das medidas mais críticas em processos biotecnológicos. Muitos métodos vêm sendo desenvolvidos ao longo dos anos e alguns se tornaram protótipos, mas ainda não são muito utilizados industrialmente (JUNG *et al.*, 1999; KIVIHARJU *et al.*, 2008).

Os métodos de quantificação da biomassa podem ser classificados como diretos e indiretos. Os métodos diretos medem as propriedades físicas das células e seus componentes. Este grupo de método pode ainda ser subdividido nos que utilizam sensores *in-situ* que são instalados diretamente dentro do fermentador, e sistemas cujos sensores são externos ao biorreator, com operação contínua ou descontínua (VOJINOVIC *et al.*, 2006). Por outro lado, os métodos de estimativa indireta de biomassa são baseados em técnicas para determinação de parâmetros relacionados às células e à atividade metabólica das mesmas. Estes parâmetros incluem a quantidade de dióxido de carbono liberado ou o oxigênio consumido pelo metabolismo, dados utilizados para estabelecimento de balanços materiais, relações estequiométricas ou modelos empíricos (MATANGUIHAN *et al.*, 1994).

As medidas feitas por detectores em contato diretamente com o processo, como os sensores *in-situ*, são chamadas de *in-line*, enquanto que as medidas *on-line* são feitas em corrente desviada (*by-pass*) automaticamente do reator para o detector. Em relação às medidas feitas em amostras retiradas ao longo do processo, estas são chamadas *off-line* quando o instrumento está distante do biorreator ou o tempo de análise é longo. Já quando o instrumento fica próximo ao processo, a análise é denominada *at-line* (também conhecido como quase *on-line*). As análises *in*, *on* e *at-line* fornecem informações em tempo real (VOJINOVIC *et al.*, 2006).

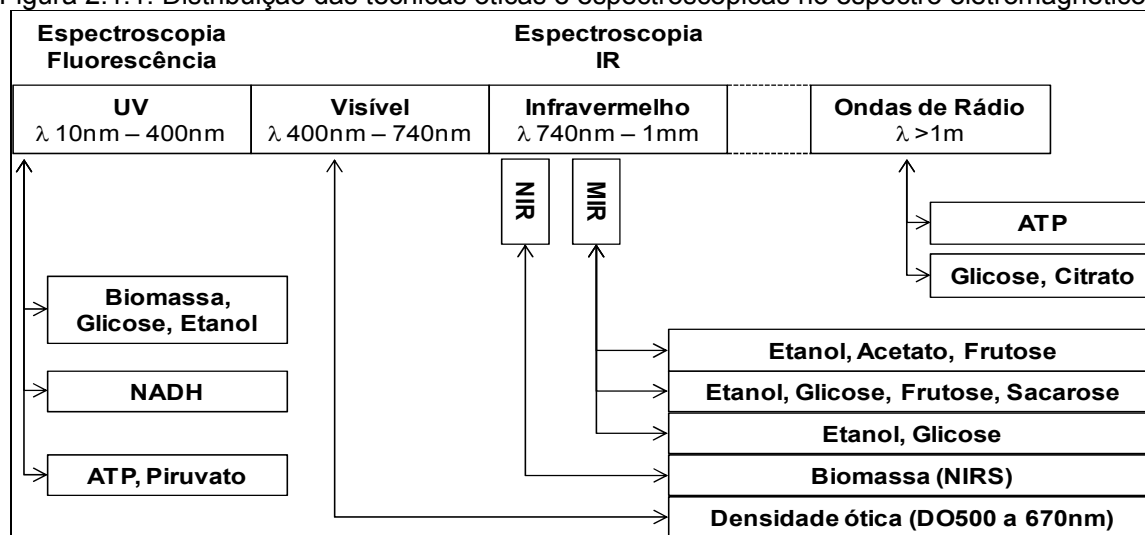
Em geral, as sondas de medida *on-line* e *in-situ* (*in-line*), além de serem esterilizáveis, devem apresentar outras características fundamentais para serem utilizadas em biorreator: i) possibilidade de calibração *on-line*; ii) dependência linear e precisão tanto em baixas quanto em altas concentrações celulares; iii) estabilidade térmica e barométrica de longo prazo (até 20 dias); iv) estabilidade à corrosão; v) inativas biologicamente; vi) não sofrer interferências de células não viáveis ou de sólidos não biológicos presentes em suspensão, vii) apresentar pouca sensibilidade às variações nas condições operacionais como pH, agitação, aeração, concentração de oxigênio dissolvido e composição do meio de cultivo (DAVEY; KELL, 1998; OLSSON; NIELSEN, 1997; BEUTEL; HENKEL, 2011).

Existem no mercado diferentes sensores de biomassa, a maioria baseada em técnicas espectroscópicas, que podem ser classificados de acordo com as propriedades medidas em:

- i) Óticos
- ii) de Infravermelho
- iii) de Fluorescência
- iv) Dielétricos

A Figura 2.1.1 apresenta o espectro eletromagnético e a extensão das técnicas óticas e espectroscópicas. Podemos observar que as técnicas de monitoramento de biomassa se estendem ao longo do espectro, além disso, os instrumentos que utilizam uma determinada banda podem detectar, além da biomassa, outras variáveis importantes, como é o caso do UV e do infravermelho.

Figura 2.1.1. Distribuição das técnicas óticas e espectroscópicas no espectro eletromagnético.



Adaptado de Beutel e Henkel (2011).

### 2.1.1 MÉTODO ÓTICO

O método indireto mais usual, rápido e preciso de medida de biomassa para organismos unicelulares é a medida *off-line* da densidade ótica (DO) (VOJINOVIC *et al.*, 2006). A determinação da DO é mais rápida que o método gravimétrico de massa seca, mas mesmo assim ainda apresenta como desvantagens o risco de contaminação do cultivo a cada alíquota retirada do biorreator. Além disso, o método não é contínuo (com algumas exceções apresentadas a seguir), exige manipulação significativa das amostras e é propenso a erros sistemáticos, necessitando de diluição volumétrica precisa para amostras com altas concentrações celulares (DEBRECZENY *et al.*, 2003). Alguns sistemas tentam automatizar o processo de amostragem e leitura em espectrofotômetro (JUNG *et al.*, 1999; LEE; LIM, 1980; YANO *et al.*, 1993, DEBRECZENY *et al.*, 2003). Para manter a relação linear entre a DO e a concentração em massa seca, os sistemas desenvolvidos apresentam recursos automáticos para diluir as amostras ou mudar o comprimento de onda. Entretanto, a complexidade e os riscos de contaminação impediram a disseminação destes sistemas (DEBRECZENY; DAVIES, 2008; DEBRECZENY *et al.*, 2003).

Existem ainda alguns dispositivos disponíveis comercialmente que oferecem medidas contínuas de densidade ótica usando o sensor imerso no meio de cultivo. Infelizmente estes dispositivos apresentam sinais sujeitos a flutuações, particularmente devido ao crescimento de microrganismo no corpo do sensor e ainda são sensíveis à agitação, à aeração (KIVIHARJU *et al.*, 2008), à presença de sólidos suspensos e não diferenciam células não viáveis e debris celulares de células viáveis (VOJINOVIC *et al.*, 2006). De acordo com Debreczeny e colaboradores (2003), o intervalo de concentração de biomassa que pode ser medido é limitado pelo uso de um único comprimento de onda. Muitos cultivos alcançam altas concentrações celulares (maiores que 50 g/L), valores que estão acima da capacidade de medida destes dispositivos baseados em medida de transmitância óptica.

Os sistemas óticos mais recentes são baseados em medidas de reflectância, apresentando a grande vantagem de instalação externa ao reator, eliminando os problemas de contaminação. Como não necessita ser esterilizado, o encapsulamento do sensor apresenta menor custo do que para os sensores esterilizáveis, que ficam imersos no biorreator. Além disso, múltiplos fermentadores podem ser monitorados simultaneamente pelo mesmo sensor (DEBRECZENY *et al.*, 2003).

Debreczeny; Davies (2008) apresentam uma comparação entre o desempenho de dois sensores óticos de biomassa: um sensor de imersão (FSC402 Turbidity Analyzer da Mettler-Toledo AG) e um sensor não invasivo (BugEye 100 Biomass Monitor da BugLab LLC). O sensor imersível contém dois pares de fibras ópticas, uma que emite luz de excitação dentro do meio de cultivo e outra que quantifica a luz refletida pelo meio de cultivo. A reflectância é assim quantificada e disponibilizada no monitor. O sensor de biomassa não invasivo é montado exteriormente ao fermentador na parede de vidro ou em alguma porta de visão. Este sistema possui um conjunto vertical de luzes de excitação em 850nm e detectores arranjados em múltiplas distâncias para detectar a luz refletida pelas células (DEBRECZENY *et al.*, 2003). Os autores compararam o desempenho dos dois sensores no monitoramento de cultivos de *E. coli* e *S. cerevisiae*, conduzidos em diferentes fermentadores e sob diferentes condições de temperatura, agitação e aeração. Os resultados apresentados mostraram que o sistema não invasivo apresentou uma maior região de linearidade com as medidas *off-line*, na faixa de 0,1 a 130 unidades de DO. Além disso, verificou-se que ambos são sensíveis à velocidade de agitação.

#### 2.1.2 ESPECTROSCOPIA DE INFRAVERMELHO

A vantagem da espectroscopia (varredura de medidas de absorbância ou transmitância) em relação às medidas simples de DO é a possibilidade de obter mais informações dos componentes do processo além da concentração de biomassa. Espectroscopia de infravermelho tem sido utilizada de diferentes formas: infravermelho próximo (NIR) e infravermelho médio (MIR). Este método pode ser aplicado *at-line* ou *on-line* através de sistemas de circulação de meio (KIVIHARJU *et al.*, 2008). O método *on-line* usando NIR foi aplicado por Gonzalez-Vara e colaboradores (2000) no controle do processo de produção de ácido láctico com *L. casei*, fornecendo informações sobre as concentrações de biomassa, glicose e ácido láctico para o sistema de controle automático. De acordo com Vojinovic e colaboradores (2006), é possível medir concentrações de glicose, glutamato, frutose, glutamina, prolina, amônia e CO<sub>2</sub> em meio de cultivo por espectroscopia NIR e MIR. Porém, estes métodos apresentam alto custo e calibração complexa (VOJINOVIC *et al.*, 2006).

### 2.1.3 FLUORESCÊNCIA

O método é baseado na excitação com luz ultravioleta e medição da emissão dos componentes do cultivo em outro comprimento de onda. O método é adequado para medir concentrações de células viáveis, na presença de concentrações intracelulares constantes de NAD(P)H. Além disso, mudanças nas condições de cultivo e a fluorescência de fundo podem interferir nas medidas de biomassa em fluorescência de 1D (ie, um comprimento de onda de excitação e outro de medição) (KIVIHARJU *et al.*, 2008).

Alguns problemas de interferência podem se resolvidos utilizando múltiplas medidas de excitação, que caracteriza a espectroscopia 2D. Esta técnica habilita a detecção de muitos fluoróforos (por exemplo, triptofano, piridoxina, FAD, FMN, NAD(P)H, e riboflavina) que podem ser correlacionados com a concentração de biomassa. O método requer ferramentas computacionais e por isso pode apresentar um atraso de medida de mais de 1 min (KIVIHARJU *et al.*, 2008).

### 2.1.4 ESPECTROSCOPIA DIELÉTRICA

Esta tese de doutorado utilizou as medidas de biomassa com sistema baseado nas propriedades capacitivas, por isso esta técnica será apresentada em mais detalhes.

A permissividade dielétrica de uma suspensão de células é um indicador da concentração de células viáveis, pois a membrana plasmática sofre polarização sobre a influência de um campo elétrico na frequência entre 0,1 a 10 MHz (ABI *et al.*, 2010; NEVES *et al.*, 2000). A capacitância resultante é correlacionada com a fração volumétrica de células, ou seja, com a concentração de biomassa. A variação da permissividade dielétrica da suspensão de células (dispersão  $\beta$ ) não depende apenas da concentração de células, mas também da condutividade do meio e do citoplasma celular, do tamanho das células e da capacitância da membrana celular, sendo este último dependente do estado fisiológico da célula (ABI *et al.*, 2010; MATANGUIHAN *et al.*, 1994).

A capacitância é uma propriedade física da célula e de acordo com Matanguihan e colaboradores (1994) mudanças nas características estruturais e morfológicas durante o crescimento ou formação de produto pode afetar o valor da capacitância. Abi e colaboradores (2010) confirmam esta hipótese utilizando o sensor para identificar a esporulação de *Bacillus subtilis*.

Em bactérias recombinantes, a ocorrência da expressão do plasmídeo durante o cultivo provoca alterações relevantes nas características superficiais da célula. Isto demonstra a importância de determinar a capacitância *in-vivo* durante os cultivos e examinar a relação entre as propriedades celulares e mudanças fisiológicas. Assim, conhecendo-se o estado fisiológico, torna-se possível identificar se o produto está sendo formado ou se as células são capazes de sintetizá-lo (MATANGUIHAN *et al.*, 1994).

Fisicamente falando, a capacitância (C) e a condutância (G) são variáveis macroscópicas e dependem das configurações e dimensões do eletrodo de medida. Por outro lado, a permissividade dielétrica ( $\epsilon$ ) e a condutividade ( $\kappa$ ) são propriedades intrínsecas das substâncias, independentes do eletrodo, como mostram as equações 2.1.1 e 2.1.2.

$$C = \epsilon \cdot \epsilon_0 \cdot \frac{A}{d} \quad (2.1.1)$$

$$G = \kappa \cdot \frac{A}{d} \quad (2.1.2)$$

Assumindo que o eletrodo consista de duas placas paralelas, “d” seria a distância entre as placas e “A” a área superficial, enquanto que  $\epsilon_0$  é a constante de permissividade do vácuo (MATANGUIHAN *et al.*, 1994).

Ainda de acordo com os autores já citados, assumindo uma suspensão de células contendo íons ou sais, quando uma corrente elétrica de baixa frequência passa entre os dois eletrodos, a capacitância é influenciada principalmente pelas células, enquanto que a variação na condutância é atribuída ao meio. Uma explicação para esse comportamento é a presença de membranas celulares intactas e não condutivas. A membrana separa a porção celular interna que também contém substâncias condutivas do meio de suspensão externo, que conduz a corrente elétrica. A separação de cargas resultante provoca um aumento na capacitância das células.

A capacitância da célula está relacionada com a fisiologia, a morfologia e com as estruturas celulares. Os parâmetros biológicos que influenciam diretamente a capacitância são: a fração volumétrica (P), a capacitância da membrana celular por unidade de área ( $C_m$ ), o tamanho da célula ou o raio celular (r) em caso de células esféricas (eq. 2.1.3).

$$C = f(P, C_m, r) \quad (2.1.3)$$

Quando o tamanho médio e as propriedades dielétricas das membranas celulares ficam constantes, a capacitância celular ( $C$ ) vira uma função da fração volumétrica de células ( $P$ ) e, conseqüentemente, da concentração celular ( $C_x$ ).

Neves e colaboradores (2000) compararam o monitoramento *in-line* de biomassa feito por medidas de capacitância com métodos padrões de determinação de biomassa (massa seca, viscosidade, concentração de DNA e produção total de  $\text{CO}_2$ ) em meio de cultivo complexo durante cultivos de *Streptomyces clavuligerus*. Os autores tiveram como principal objetivo estabelecer um monitoramento preciso em tempo real da concentração celular através de medidas de capacitância e, como objetivo secundário, correlacionar a capacitância com outras variáveis monitoradas relacionadas à concentração de biomassa. Em um reator com capacidade para 550 litros, os autores obtiveram relação linear entre medidas de capacitância e: produção de  $\text{CO}_2$  durante a fase de crescimento exponencial; concentração de DNA nas fases de crescimento lag e exponencial; viscosidade do meio nas fases de transição e estacionária. Os autores concluíram que o uso do sensor de capacitância fornece uma descrição precisa em tempo real do perfil de biomassa durante todo o processo. Além disso, enfatizaram a possibilidade de calcular a velocidade específica de crescimento *on-line*, e desta forma desenvolver estratégias de controle mais precisas e reprodutivas.

Yang e colaboradores (2003) utilizaram um sistema de três eletrodos para detectar a impedância durante o crescimento de *S. typhimurium*, separando as componentes resistividade e capacitância pela extensão da freqüência para um valor alto (1 MHz) e um valor baixo (1 Hz). Os autores analisaram a impedância em função do tempo de crescimento bacteriano, e identificaram que a capacitância deve ser medida em baixas freqüências de impedância (menor que 10 kHz) ao passo que altas freqüências de impedância (maior que 10 kHz) permitem coletar informações sobre a resistência elétrica do meio. Além disso, os autores concluem que ambas, capacitância e a resistividade, podem ser utilizadas como forma de medição da presença de bactérias.

Ferreira e colaboradores (2005) compararam o desempenho de um eletrodo de quatro pinos com um eletrodo anular para monitoramento da concentração de células viáveis em tempo real durante a produção de um fármaco por fermentação. Os autores monitoraram 4 fermentações de *Streptomyces sp* com o sensor de quatro pinos e 10 fermentações com o sensor anular. As medidas obtidas por ambos foram

comparadas entre si e com medidas *off-line* (viscosidade e volume micelial). Os autores afirmaram que, embora a viscosidade seja o método mais aceito de estimar biomassa em sistemas miceliais, discrepâncias entre medidas de viscosidade e de capacitância nos estágios finais das fermentações foram observadas devido à modificação na morfologia, causada pela lise celular ou fragmentação micelial, que alteram as propriedades reológicas do caldo fermentativo. Os autores observaram ainda que o volume micelial, apesar de ser um dos métodos mais comuns de medidas de biomassa, não é confiável visto que é influenciado por mudanças morfológicas no microrganismo e pela presença de matéria insolúvel. Tendo em vista a linearidade entre concentração celular ( $X$ ) e capacitância ( $C$ ), equação 2.1.4, os autores utilizaram os dados de capacitância para calcular a velocidade específica de crescimento  $\mu$ , equação 2.1.5 (Balanço de massa para um reator batelada).

$$X = \alpha C \quad (2.1.4)$$

$$\mu = \frac{1}{X} \frac{dX}{dt} = \frac{1}{C} \frac{dC}{dt} \quad (2.1.5)$$

O trabalho apresenta bons resultados no monitoramento de biomassa através das medidas de capacitância. Além disso, essas medidas viabilizaram o cálculo em tempo real da velocidade específica de crescimento, e desta forma o desenvolvimento de estratégias de controle mais precisas e reprodutivas.

Mas e colaboradores (2001) também utilizaram medidas de capacitância para monitoramento em tempo real de cultivos de *Saccharomyces cerevisiae*. Os autores desenvolvem uma técnica automática para determinar a velocidade de crescimento do fermento, que é validada através de métodos convencionais de balanço de massa. As concentrações de biomassa de referência foram medidas por massa seca convencional enquanto medidas de biomassa on-line foram feitas por um dispositivo de impedância de frequência dual. Os resultados do trabalho confirmam a potencialidade do monitoramento da concentração de biomassa através de medidas de capacitância.

Dabros e colaboradores (2010) implantaram um sistema de controle da velocidade específica de crescimento em cultivos batelada alimentada de *Kluyveromyces marxianus* através do controle da vazão de alimentação baseado em medidas de capacitância.

Além do monitoramento de biomassa por capacitância, estes sensores têm sido utilizados para diversos estudos, como a investigação da camada-S suportada



por membrana lipídica pela técnica modificada de Langmuir-Blodgett (GUFLER *et al.*, 2004). Bai e colaboradores (2006) também determinaram as propriedades dielétricas de vários componentes celulares através do modelo “*three-shell*” esférico pelo espectro dielétrico de suspensões celulares de *E. coli*.

A Tabela 2.1.1 permite a comparação dos diferentes sistemas de medição de biomassa em termos de custo. Nota-se que os sensores NIR e de permissividade apresentam um custo significativamente superior ao do sensor ótico externo. Porém, enquanto o último é capaz de monitorar apenas a concentração celular, o sensor de permissividade permite obter informações sobre o estado fisiológico das células e o sensor NIR consegue quantificar vários componentes do meio de fermentação, além da biomassa. Desta forma, a escolha do sensor mais adequado para o monitoramento da concentração celular em um bioprocessamento é uma decisão que deve levar em conta não só o preço, mas também a qualidade da informação disponibilizada e o uso pretendido para a mesma.

Tabela 2.1.1. Valores médios de diferentes sistemas de medição de biomassa.

SISTEMA	R\$	DATA
SENSOR NIR DE IMERSÃO (METTLER TOLEDO - LED COM FREQUÊNCIA DE 880NM)	33.000,00	AGOSTO DE 2011
SENSOR DE PERMISSIVIDADE (ABER INSTRUMENTS - INCLUI SOFTWARE)	32.000,00	NOVEMBRO DE 2011
SENSOR ÓTICO EXTERNO (BUGLAB - INCLUI SOFTWARE)	14.000,00	NOVEMBRO DE 2011

Destaca-se ainda que, apesar da grande quantidade de estudos disponíveis na literatura sobre cultivos de *E. coli* recombinante em batelada alimentada (BABAEIPOUR *et al.*, 2007; BÄCKLUND; MARKLAND; LARSSON, 2008; CASTAN; NÄSMAN; ENFORS, 2002; KIM, *et al.*, 2004; KORZ *et al.*, 1995), há pouca informação sobre o desempenho de sensores de biomassa nesses cultivos, particularmente em condições de alta densidade celular (KIM; PARK; JUNG, 2009; KIVIHARJU *et al.*, 2007; KIVIHARJU *et al.*, 2008; MATANGUIHAN; KONSTANTINOV; YOSHIDA, 1994).

## 2.2. ESTRATÉGIAS DE CONTROLE DA ALIMENTAÇÃO

Cultivos em batelada alimentada de *rE. coli* são comumente empregados para a obtenção de altas concentrações celulares. Nesses cultivos, o fornecimento do meio suplementar deve ser controlado de forma que os nutrientes sejam fornecidos em quantidade suficiente para sustentar o crescimento, porém sem ocorrência de acúmulo de nutrientes, que podem provocar inibição do crescimento ou a formação de metabólitos indesejados pelo metabolismo *overflow*. Além disso, tanto o crescimento celular quanto a síntese protéica nas micro-fábricas celulares são fortemente influenciados pelo perfil de alimentação e pela velocidade de crescimento, impactando a produtividade do processo (AULICINO *et al.*, 2010).

Existem várias estratégias de controle da vazão de alimentação para cultivos em batelada alimentada, inclusive combinações entre elas, sendo que as clássicas consistem em manter um perfil de alimentação pré-definido, mantendo a velocidade de crescimento limitada pela vazão de meio suplementar (AKESSON *et al.*, 2001).

Assim, os perfis de alimentação mais comumente usados em bateladas alimentadas podem ser classificados nas seguintes categorias (KIM *et al.*, 2004):

- i) sem *feedback* com vazão constante ou exponencial;
- ii) com *feedback* baseado no pH (pH-stat);
- iii) com *feedback* baseado na COD (DO-stat);
- iv) com *feedback* baseado na medição ou inferência de outras variáveis.

Nas estratégias de cultivo baseadas em perfis de alimentação constantes, os nutrientes são adicionados ao fermentador a uma vazão pré-determinada. Devido ao aumento no volume e na concentração celular, a velocidade específica de crescimento pode diminuir continuamente, e o aumento na concentração celular desacelerar ao longo do tempo. O aumento da vazão da alimentação em degrau permite que o fornecimento de nutrientes acompanhe melhor o aumento da biomassa e que concentrações celulares mais elevadas sejam alcançadas (LEE, 1996).

Dos perfis de alimentação sem *feedback* o exponencial é o mais elaborado, pois usa um modelo empírico do crescimento celular para regular a vazão de alimentação. Em condições de crescimento balanceado e não limitado pela falta de

nutrientes, as células crescem exponencialmente atingindo uma concentração celular elevada mais rapidamente (AULICINO *et al.*, 2010). No entanto, dependendo da fonte de carbono, a imposição de uma alta velocidade de crescimento nos cultivos de *E. coli* pode resultar na formação de acetato, que inibi o crescimento e a formação de produto de interesse. O desvio do metabolismo que leva à formação do acetato pode ser evitado controlando a velocidade específica de crescimento abaixo de um valor crítico (LEE *et al.*, 1999).

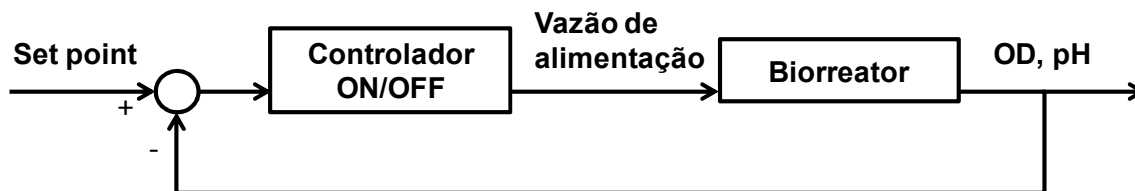
A vazão de alimentação que permite crescimento exponencial com velocidade específica de crescimento constante pode ser obtida a partir de um balanço de massa simples, assumindo que  $\mu$  só dependa da concentração de substrato, que o rendimento ( $Y_{XS}$  em  $g_X/g_S$ ) e a manutenção ( $m$  em  $h^{-1}$ ) sejam constantes e que a concentração residual de substrato limitante seja constante e igual a zero, equação 2.2.1 (LEE, 1996).

$$F(t) = \left( \frac{\mu}{Y_{XS}} + m \right) \frac{X_0 \cdot V_0}{S_F} \exp(\mu \cdot t) \quad (2.2.1)$$

Na equação 2.2.1  $F$  representa a vazão de meio suplementar (L/h),  $\mu$  a velocidade específica de crescimento ( $h^{-1}$ ),  $X_0$  a concentração celular no início da alimentação,  $V_0$  o volume no início da alimentação e  $S_F$  a concentração de substrato no meio de alimentação. Porém esta estratégia sem *feedback* não evita o acúmulo de substrato caso haja uma superalimentação (KIM *et al.*, 2004). Assim, esta estratégia pode falhar caso haja algum outro fator limitando o crescimento celular, como a baixa concentração de oxigênio dissolvido (LIN *et al.*, 2005).

Existem esquemas simples de controle com *feedback* indireto que acoplam a alimentação com medidas de pH (pH-stat) ou de OD (OD-stat), Figura 2.2.1 De acordo com Lin e colaboradores (2005), estas medidas são indicadores do estado fisiológico do cultivo e são facilmente monitoradas. Nestas estratégias, a vazão de adição de nutrientes é manipulada para manter o pH ou o OD no valor desejado. Quando o pH ou o OD se tornam maior que o *setpoint* o controlador ON/OFF alimenta nutrientes numa vazão predeterminada. O método de pH-stat é baseado na elevação do pH devido à excreção de íons de amônio quando ocorre o esgotamento da principal fonte de carbono. Similarmente, o OD-stat é baseado no aumento rápido na concentração de oxigênio dissolvido quando ocorre o esgotamento da fonte de carbono (LEE *et al.*, 1999; LIN *et al.*, 2005).

Figura 2.2.1. Diagrama do método de *feedback* (loop simples) indireto.



Adaptado de Lee e colaboradores (1999)

Geralmente o método de DO-stat funciona bem em meio definido, que contém apenas uma fonte de carbono bem definida. Este método pode falhar em meios complexos (suplementados com extrato de levedura, triptona, peptona etc), pois estes meios podem suportar o crescimento celular pelo catabolismo dos nutrientes ricos em aminoácidos, mantendo o consumo de oxigênio elevado mesmo em condições de baixa concentração da fonte principal de carbono (LIN *et al.*, 2005).

O método de pH-stat é mais indicado para cultivos com meio complexo. Kim e colaboradores (2004) utilizaram o método pH-stat combinado com alimentação exponencial em cultivos de *E. coli*. Os autores alcançaram 101 g/L de células em 52 h de cultivo, mantendo a velocidade específica de crescimento em  $0,1 \text{ h}^{-1}$ . Os autores notaram que, no entanto, a limitação da velocidade de crescimento através da alimentação afetou negativamente a síntese do produto recombinante, neste caso, PHB [poly(3-hydroxybutyrate)].

É importante destacar que as estratégias DO-stat e pH-stat impõe um grande estresse às células, tendo em vista as grandes oscilações no pH, na OD e na concentração residual de substrato, inerentes à forma como o controle da alimentação é realizada.

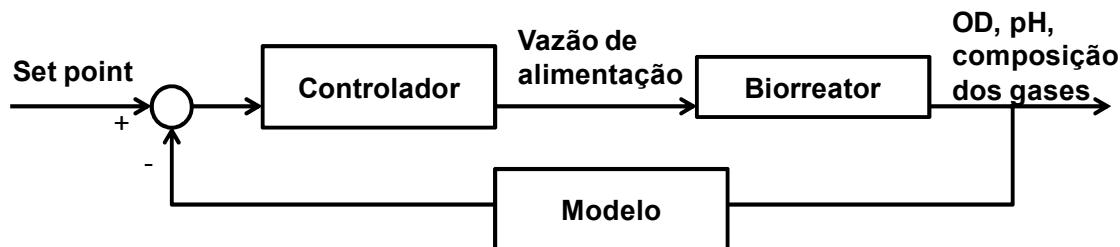
A medida *online* da concentração de fonte de carbono no reator, se disponível, poderia ser utilizada para manipular a vazão de alimentação. Kim e colaboradores (2004) usaram um analisador de glicose em cultivo em batelada alimentada de *Alcaligenes eutrophus* para produção de poly(3-hydroxybutyrate). A glicose foi controlada entre 10 e 20 g/L. Apesar de apresentar resultados satisfatórios, alcançando uma concentração celular de 164 g/L em 50h, o método apresentou algumas deficiências devido à instabilidade nas medidas de glicose. De maneira geral, existe uma preferência por estratégias de controle baseadas na estimativa da velocidade de consumo de substrato, pois as análises online de concentração de substrato são de alto custo e implementação complexa (LEE *et al.*, 1999).

Konstantinov e colaboradores (1990) apresentaram um método baseado na estimativa da velocidade de consumo de glicose, no qual a vazão de alimentação de meio suplementar é modulada a partir de medidas da composição dos gases de saída do biorreator. O trabalho apresenta um algoritmo capaz de garantir suprimento suficiente de oxigênio de forma a evitar a superalimentação com glicose. Konstantinov e colaboradores (1990-b) introduziram o conceito de balanceamento entre a alimentação de glicose e o consumo da mesma, para evitar acúmulo ou falta de nutrientes, manipulando a vazão de alimentação juntamente com a velocidade de agitação para evitar a limitação por oxigênio.

Kleman e colaboradores (1991) desenvolveram uma metodologia sistemática para determinar  $F$  durante os cultivos em batelada alimentada de *E. coli*. O esquema de controle apresenta dois componentes: a estimativa da demanda futura de glicose (GD) baseada no valor recente de  $F$  (componente de controle *feedforward*); e a correção dos desvios (componente de controle *feedback*). Os autores utilizaram um analisador de glicose como entrada (*input*) do sistema de controle, com o qual obtiveram 65 g/L de biomassa em 9 h de cultivo. Uma observação importante é que o suprimento de oxigênio se torna um problema em alta densidade celular, provocando instabilidade em todas as outras variáveis (LEE *et al.*, 1999).

A Figura 2.2.2 apresenta o diagrama da estratégia de alimentação baseada em modelo.

Figura 2.2.2. Diagrama da estratégia de controle da vazão de alimentação baseada em demanda ou consumo de substrato.



Adaptado de Lee e colaboradores (1999)

Como visto até aqui, existem muitas opções de controle disponíveis na literatura (LIU *et al.* 1993; WIPF *et al.* 1994; MASSIMO *et al.* 1991; YAMANE *et al.* 1992), sendo que algumas opções são baseadas em modelos. Através de medidas da concentração de oxigênio dissolvido, composição dos gases e/ou da concentração de nutrientes pode-se utilizar um modelo para calcular a demanda por substrato e controlar a vazão de alimentação. Quanto aos modelos, podem ser caixa branca, com

equações de balanço e de crescimento (como a equação de Monod), caixa preta (com redes neurais), ou caixa cinza (uma combinação do caixa preta com caixa branca) (NELLES, 2001).

Zhou e colaboradores (1995) utilizaram a relação estequiométrica entre o consumo de oxigênio e de glicose. Hosobuchi e colaboradores (1993) utilizaram análises *online* por HPLC para determinar as concentrações de fonte de carbono como entrada para o sistema de controle. Já os autores Oh e colaboradores (1998) utilizaram apenas o valor de F e da concentração de oxigênio dissolvido como entrada.

Além das estratégias convencionais, diversos trabalhos utilizam lógica fuzzy (ALFAFARA *et al.*, 1993; HONDA; KOBAYASHI, 2000; KARAKUZU *et al.*, 2006; KITSUTA; KISIMOTO, 1994; NYTTLE; CHIDAMBARAM, 1993; REIS *et al.*, 2009; ZHANG *et al.*, 1994) e redes neurais (HORTA *et al.*, 2011; MÉSZÁROS *et al.*, 2004; THOLUDUR; RAMIREZ, 1996; SHI; SHIMIZU, 1992) como partes integrantes dos sistemas de controle da vazão de alimentação. Geralmente as redes neurais são utilizadas para estimar uma variável que não pode ser medida, enquanto que o fuzzy é utilizado para mesclar informações medidas com conhecimento acumulado de especialistas humanos (LEE *et al.*, 1999).

Dabros e colaboradores (2010), publicaram recentemente um sistema de controle com *feedback* dos dados *in-line* de biomassa (medidas baseadas em capacitância, seção 2.1.4) estimando  $\mu$  de acordo com a equação 2.2.2, sendo X a biomassa (g/L) medida *online* e t o tempo (h) e  $\mu_{EST}$  a velocidade específica de crescimento estimada.

$$\mu_{EST} = \frac{\ln(X_t/X_{t-1})}{t_t - t_{t-1}} \quad (2.2.2)$$

Os autores utilizaram um filtro de média móvel para alisar os dados de biomassa antes da estimativa de  $\mu$ , sendo este valor fornecido como entrada para três sistemas de controle. O primeiro sistema consistiu num *feed-forward*, equação 2.2.3, cuja vazão de alimentação segue um perfil predeterminado de  $\mu_{SET}$ . O segundo e o terceiro sistema, equação 2.2.4 e 2.2.5, apresentaram controle com *feedback* (proporcional e proporcional-integral, respectivamente) incluso diretamente no termo exponencial da equação de alimentação.

$$F(t) = \frac{\mu_{SET}}{Y_{XS}} \frac{X_0 \cdot V_0}{S_F} \exp(\mu_{SET} \cdot t) \quad (2.2.3)$$

$$F(t) = \frac{\mu_{SET}}{Y_{XS}} \frac{X_0 \cdot V_0}{S_F} \exp\{[\mu_{SET} + K_p \cdot (\mu_{SET} - \mu_{EST})]t\} \quad (2.2.4)$$

$$F(t) = \frac{\mu_{SET}}{Y_{XS}} \frac{X_0 \cdot V_0}{S_F} \exp\left\{\left[\mu_{SET} + K_p(\mu_{SET} - \mu_{EST}) + K_i \int_0^t (\mu_{SET} - \mu_{EST}) dt\right]t\right\} \quad (2.2.5)$$

Nestas equações  $\mu_{SET}$  é o *setpoint* igual à velocidade específica de crescimento desejada,  $X_0$  é a concentração celular,  $V_0$  o volume do reator no momento inicial do estágio de alimentação,  $S_F$  a concentração de substrato no meio suplementar,  $Y_{XS}$  o coeficiente de rendimento,  $K_p$  o ganho proporcional e  $K_i$  o ganho integral. O trabalho apresentou melhores resultados com o terceiro sistema, incluindo o ganho proporcional e integral no controle da vazão de alimentação (DABROS et al., 2010).

Esta tese de doutorado utiliza uma técnica de controle semelhante à empregada por Dabros e colaboradores (2010). Porém ao invés de limitar a velocidade específica de crescimento em um *setpoint*,  $\mu_{SET}$  foi atualizada a partir de medidas on-line de  $\mu$ , estimada a partir dos dados de permissividade,  $\mu_{EST}$ , o que permitiu a manutenção do crescimento nos limites estabelecidos pela própria célula (HORTA et al., 2011b).

O principal objetivo de manipular a vazão de alimentação é controlar o crescimento e assim alcançar altas concentrações celulares. Não há uma definição clara do que seja um cultivo de alta densidade celular. Alguns trabalhos encontrados na literatura consideram cultivos de alta densidade celular aqueles com concentrações de biomassa superiores a 50 g/L de massa seca de células, enquanto outros estabelecem 100 g/L como limite (CHOI et al., 2006; KIM et al., 2004). O Quadro 2.2.1 traz alguns exemplos das maiores concentrações celulares atingidas com diferentes estratégias de cultivo.

Observa-se que os melhores resultados, com concentração celulares superiores a 150 g/L, foram obtidos em biorreatores acoplados a sistemas de perfusão, o qual envolve um modo de operação bastante complexo. Para cultivos sem perfusão, a máxima concentração celular atingida foi ~ 150 g/L, com estratégia de controle da alimentação baseada em  $\mu_{SET}$  pré-definido. Em termos de produtividade em biomassa, os valores variam entre 3 e 7,2 g/L.h.

É importante ressaltar que a maioria dos trabalhos que envolvem estratégias de controle não aborda a fase de indução, focando apenas na fase de crescimento. Porém, além de garantir um crescimento elevado, a estratégia de controle

deve garantir ainda as condições ideais para a máxima produção da proteína de interesse na fase de indução (AULICINO *et al.* 2010).

Quadro 2.2.1. Densidades celulares atingidas com diferentes estratégias em cultivos de *E. coli*

Estratégia de alimentação	Meio básico	Fonte de Carbono	Massa seca atingida (g/L)	produtividade (g/L.h)	Referência
<i>Batelada alimentada</i>					
1. Crescimento exponencial	Meio semi-definido com extrato de levedura	Glicose	54	-	Shiloac; Bauer, 1975
2. Crescimento linear limitado pela fonte de C	Meio definido	Glicose sólida	134	-	Neidhardt <i>et al.</i> , 1996 Matsui <i>et al.</i> , 1989
	Meio definido	Glicose Ácido cítrico	104	-	Riesenberg <i>et al.</i> , 1991
	Meio definido	Glicerol	148	3	Korz <i>et al.</i> , 1995
			Glicose	128	5,3
	Hidrolisado protéico e extrato de levedura	Glicerol	84	-	Macaloney <i>et al.</i> , 1996
3. Crescimento linear lento para manter concentração de acetato próxima a zero	Meio definido	Baixa glicose Glicerol	145	4,4	Horn <i>et al.</i> , 1996
<i>Diálise</i>					
1. Reator com membrana de diálise	Meio complexo	Glicerol	174	-	Märkl <i>et al.</i> , 1993
	Meio basal	Glicose Glicerol	190 180	6,3 7,2	Nakano <i>et al.</i> , 1997
2. Alimentação "nutrient-split"	Contra solução tampão de sais	Glicerol	150	-	Ogbonna e Märkl, 1993

Fonte: Adaptado de SHILOACH; FASS, (2005).

Em cultivos de alta densidade de *E. coli* recombinante, a produção da proteína de interesse está relacionada a diversos fatores, entre eles o momento da indução e a natureza do indutor utilizado para promover sua expressão. As células de *E. coli* BL21(DE3) possuem em seu cromossomo uma cópia do gene que codifica a



enzima T7 RNA polimerase sob controle do promotor lacUV5. A adição de IPTG ou lactose (que é intracelularmente convertida à alolactose) provoca o deslocamento da proteína repressora LacI do *operon* dos promotores que controlam a transcrição, tanto do gene da enzima T7 RNA polimerase no cromossomo como do gene alvo clonado no vetor pET, e a proteína de interesse passa a ser expressa (NOVAGEN, 2005).

A super-expressão da proteína heteróloga causa um grande estresse à célula, que é forçada a redirecionar o seu metabolismo para a síntese dos constituintes da proteína, afetando o seu crescimento. O nível de estresse depende da proteína recombinante, da construção, das condições de cultivo (temperatura, composição do meio) e, portanto, a estratégia de alimentação a ser adotada na fase de indução é uma questão específica do sistema em estudo e deve ser analisada para cada caso.

### 2.3. PROGRAMAS DE MONITORAMENTO E CONTROLE DE BIORREADORES

Tendo em vista as considerações levantadas nos tópicos anteriores sobre a importância do monitoramento nos cultivos de *rE. coli* e da adoção de estratégias de controle mais sofisticadas nos mesmos, este último item aborda justamente os ambientes computacionais disponíveis para a aquisição de dados, supervisão e controle de biorreatores.

Um levantamento preliminar na Internet permite identificar inúmeros fabricantes de programas para monitoramento e controle de biorreatores, dentre eles:

- i) New Brunswick (<http://www.nbsc.com/files/BioCommand.pdf>);
- ii) DASGIP (<http://www.pharmaceutical-int.com/suppliers/dasgip-ag.html>);
- iii) Electrolab ([http://www.electrolab.biz/Fermentation\\_Software.aspx](http://www.electrolab.biz/Fermentation_Software.aspx));
- iv) Applikon ([http://www.applikon-bio.com/index.php?option=com\\_content&view=article&id=206:specifications&catid=74:software&Itemid=144&Itemid=235](http://www.applikon-bio.com/index.php?option=com_content&view=article&id=206:specifications&catid=74:software&Itemid=144&Itemid=235));
- v) Finesse Solutions (<http://finesse.com/trubio-overview>);
- vi) INFORS HT (<http://infors-ht.com/index.php/en/products/software/software-for-bioreactors/iris>).

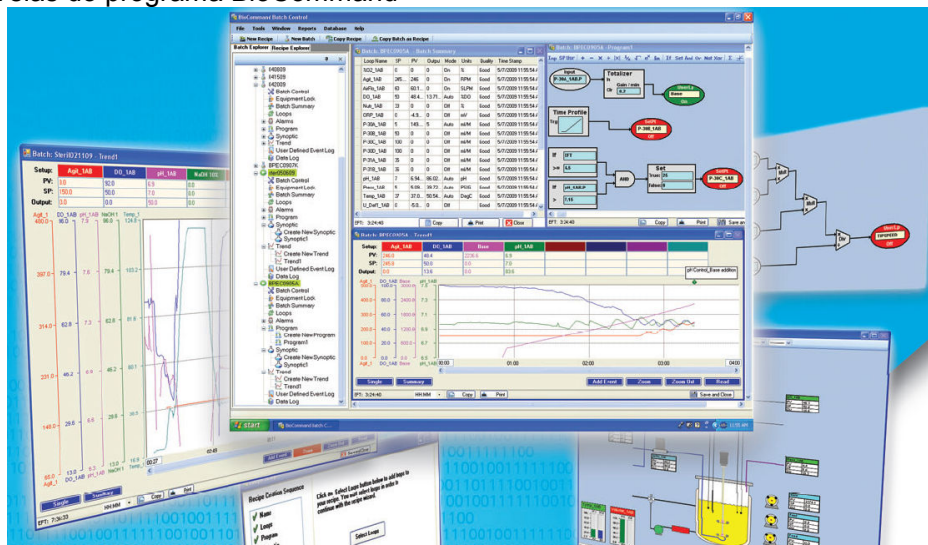
Em geral todos os programas apresentam as mesmas funcionalidades básicas de controle de alguns dos principais parâmetros como pH, COD, temperatura, agitação, espuma e vazão de líquidos (alimentação, ácidos, bases) e de gases.

Nesta última década, por volta de 2008, os programas de controle passaram a incorporar inovações como o acesso remoto pela rede, permitindo que o

biorreator seja monitorado a distância (HORTA *et al.*, 2010; JESUS *et al.*, 2007). Foram ainda incluídas funções tecnológicas como alarmes sonoros e mensagens para celular (SMS). Os programas também se tornaram mais flexíveis, permitindo que o operador construa sua própria rotina de controle, com os novos sistemas SCADA (*supervisory control and data acquisition* – supervisão controle e aquisição de dados), que podem ser adaptados às especificidades de cada processo.

A Figura 2.3.1 ilustra o programa de controle da New Brunswick. O operador pode definir a variável controlada, as variáveis que alimentam o controle e as relações entre elas, sendo que a programação pode apresentar ação baseada em tempo ou em eventos.

Figura 2.3.1. Telas do programa BioCommand®



Fonte: New Brunswick

### 3. MATERIAIS E MÉTODOS

Esta seção apresenta as principais técnicas utilizadas no desenvolvimento da pesquisa envolvida neste trabalho.

#### 3.1. MICRORGANISMOS

Os principais microrganismos utilizados foram:

- a) *Escherichia coli* BL21(DE3) PspA3;
- b) *Escherichia coli* BL21(DE3) PspA245;
- c) *Escherichia coli* BL21(DE3) PspA4-pro;

Os clones de *E. coli* foram gentilmente cedidos pelo Centro de Biotecnologia, Instituto Butantan, São Paulo, Brasil.

Outros microrganismos também foram testados e utilizados como foco principal de outros trabalhos do grupo, são eles:

- a) *Saccharomyces cerevisiae* (REIS, 2009);
- b) *Escherichia coli* XL1-blue (PspA) (SILVA, 2011);
- c) *Escherichia coli* pT101/D-TOPO (VELEZ, 2009);
- d) *Pichia pastoris* (CAVALCANTI\_MONTANO, 2010);
- e) *Bacillus megaterium* (SUAREZ, 2010)

#### 3.2. MEIO DE CULTIVO

Durante o decorrer do trabalho foram realizados 3 cultivos em biorreator para a otimização da composição do meio definido, dado na Tabela 3.2.1. Preparou-se uma solução contendo EDTA, ácido cítrico, ácido bórico e os sais descritos na Tabela 3.2.1, com exceção do sulfato de magnésio ( $MgSO_4 \cdot 7H_2O$ ). Ajustou-se o pH para 6,3 e esterilizou-se esta solução a 121°C e 2 atm por 20 minutos. Tanto a solução de glicerol quanto a solução de  $MgSO_4 \cdot 7H_2O$  foram esterilizadas separadamente também a 121°C e 2 atm por 20 minutos e adicionadas posteriormente à primeira solução. As soluções de canamicina, tiamina e IPTG foram esterilizadas por filtração em membranas de 0,22

µm. As soluções foram misturadas assepticamente em temperatura ambiente para evitar possíveis reações e precipitação (SARGO, 2011).

Tabela 3.2.1. Composição do meio de cultivo

Componentes	Batelada	Alimentação
Glicerol (g/L)	40	800
Kanamycin (mg/L)	50	50
MgSO <sub>4</sub> .7H <sub>2</sub> O (g/L)	1.6	20
Antiespumante PPG 30 % (mL/L)	1	1
(NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub> (g/L)	5.3	6.4
KH <sub>2</sub> PO <sub>4</sub> (g/L)	17.7	21.8
Citric acid (g/L)	2,27	2,27
Fe(III)citrate (mg/L)	133.3	40
CoCl <sub>2</sub> .6H <sub>2</sub> O (mg/L)	3.33	4
MnCl <sub>2</sub> .4H <sub>2</sub> O (mg/L)	20	23.5
CuCl <sub>2</sub> .2H <sub>2</sub> O (mg/L)	2	2.5
H <sub>3</sub> BO <sub>3</sub> (mg/L)	4	5
Na <sub>2</sub> MoO <sub>4</sub> .2H <sub>2</sub> O (mg/L)	2.8	4
Zn(CH <sub>3</sub> COOH) <sub>2</sub> .H <sub>2</sub> O (mg/L)	33.8	16
EDTA (mg/L)	18.8	13
Tiamina (mg/L)	45	45

Adaptado de Sargo (2011).

### 3.3. PREPARAÇÃO DO INÓCULO

Antes de iniciar o cultivo em biorreator, foram necessárias duas etapas de crescimento celular em frascos agitados chamadas de Pré-Inóculo e Inóculo. O Pré-Inóculo foi realizado em um volume total de 60 mL, divididos em 2 frascos de 500 mL com inoculação de duas colônias (uma para cada frasco) obtidas em uma placa de LB-Ágar/kan que foi estriada com a suspensão de células do estoque congelado e, em seguida, incubada por 24 horas a 37°C.

Ao atingir uma DO<sub>600nm</sub> de aproximadamente 2,5, transferiu-se uma alíquota de 10 mL para cada um dos 3 frascos de 1 L contendo 100 mL de meio fresco. O Inóculo foi incubado a 250 rpm e 30° C até que se obtivesse uma densidade ótica entre 2,5 e 3,0, podendo, então, inocular os 300 mL de meio no reator.

### 3.4. CULTIVOS

Os experimentos foram realizados em biorreator de 5L, de montagem própria no LADABio, (Laboratório de Desenvolvimento e Automação de Bioprocessos - Departamento de Engenharia Química da Universidade Federal de São Carlos). O

aparato experimental para aquisição, monitoramento e controle do reator está detalhado na seção 4.1.

A maioria dos experimentos realizados ao longo do desenvolvimento da ferramenta e as respectivas condições de cultivo adotadas estão resumidos na Tabela 3.2.2. A descrição detalhada das condições de cultivo empregadas nos experimentos pode ser encontrada na dissertação de mestrado de Cíntia Regina Sargo (2011), na tese de Doutorado de Adilson José da Silva (2011) ou ainda em trabalhos apresentados em congressos (VÉLEZ et al., 2011; SARGO et al., 2011; SILVA et al., 2010; REIS et al., 2009).

Tabela 3.2.2. Identificação dos principais cultivos

id	Microorganism	Medium	Maximum cellular concentration (g/L)	Reference
A1	pET37b+/PspA3	Defined	76	Horta et al., 2010
A2	pET37b+/PspA3	Defined	86	Horta et al., 2010
A3	pET37b+/PspA3	Defined	123	Sargo et al., 2011
A4	pET37b+/PspA245	Defined	108	Sargo et al., 2011
A5	pET37b+/PspA245	Defined	81	
A6	pET37b+/PspA245	Defined	68	
A7	pET37b+/PspA245	Defined	21	
A8	pET37b+/PspA245	Defined	102	
A9	pET37b+/PspA245	Defined	10	
A10	pET37b+/PspA245	Defined	45	
A11	pET37b+/PspA3	Defined	105	
A12	pET37b+/PspA3	Defined	123	Sargo et al., 2011
A13	pET37b+/PspA4Pro	Defined	116	Sargo, 2011
A14	pET37b+/PspA3	Defined	134	
A15	pET37b+/PspA3	Defined	20	
A16	pET37b+/PspA3	Defined	155	Horta et al, 2012
B1	pET28a/SpaA	Complex	21	Silva et al., 2011a
B2	pET28a/SpaA	Complex	24	Silva et al., 2011a
B3	pET28a/SpaA	Complex	65	Silva et al., 2011a
B4	pET28a/SpaA	Complex	66	Silva et al., 2011a
C1	pT101/D-TOPO	Complex	101	Vélez et al.,2011
C2	pT101/D-TOPO	Defined	125	Vélez et al.,2011

### 3.5. MÉTODOS ANALÍTICOS

#### 3.5.1. CONCENTRAÇÃO CELULAR

A concentração de biomassa foi medida por três métodos diferentes:

- Turbidimetria: as leituras de Densidade Ótica (DO) foram feitas no comprimento de onda de 600nm, com espectrofotômetro Pharmacia

Biotech modelo Ultrospec 2000. As amostras foram diluídas de modo que a leitura da absorbância não ultrapassasse o valor de 0,8, para garantir que a medida fosse realizada dentro da faixa de linearidade da relação entre densidade óptica e massa seca;

- b) Gravimétrico: as medidas de massa seca (Cx - g/L) foram determinadas através da filtração de um volume conhecido da amostra em uma membrana com poro de 0,22 $\mu$ m previamente pesada. A amostra foi submetida à secagem em estufa a 60°C até obter massa constante.
- c) Permissividade (pF/cm): As medidas de permissividade foram feitas online através do sensor de biomassa (Biomass System – Fogale Nanotech).

### 3.5.2. CONCENTRAÇÃO DE AÇÚCARES E ÁCIDOS ORGÂNICOS

A concentração de açúcares (glicose, lactose e galactose), glicerol e ácidos orgânicos nas amostras retiradas durante os cultivos foi determinada por HPLC (sistema Waters Co; Bombas HPLC 510, Injetor W717, refratômetro W410 e leitor de UV PDA W996) utilizando a coluna Aminex HPX-87H (Bio-Rad) como fase estacionária e uma solução de H<sub>2</sub>SO<sub>4</sub> 5 mM como fase móvel, a uma vazão de 0,6 L/min e 60°C (SARGO, 2011).

#### 4. SISTEMA DE AUTOMAÇÃO PARA CULTIVOS DE ALTA DENSIDADE CELULAR: SUPERSYS\_HCDC

Esta seção apresenta o desenvolvimento do *software* de controle SUPERSYS\_HCDC através do artigo “A SUPERVISION AND CONTROL TOOL BASED ON ARTIFICIAL INTELLIGENCE FOR HIGH CELL DENSITY CULTIVATIONS” (em finalização para submissão ao Brazilian Journal of Chemical Engineering) e do relatório de registro de software no apêndice A.

O programa SUPERSYS\_HCDC foi utilizado para monitoramento e controle de diversos experimentos com diferentes microrganismos, como *Saccharomyces cerevisiae*, *Pichia pastoris*, *Bacillus subtilis*, *Bacillus megaterium*, além de diferentes clones de *Escherichia coli* recombinante.

O software SUPERSYS\_HCDC foi desenvolvido para controlar e supervisionar cultivos de alta densidade celular (CADC). O sistema foi desenvolvido empregando os programas Labview e Matlab e inclui algumas funcionalidades especiais, como o controle híbrido hierárquico automático da concentração de oxigênio dissolvido, atuando na velocidade de agitação e nas vazões de ar e de oxigênio; o acionamento automático da bomba de alimentação, pelo reconhecimento do final da batelada por um comitê de redes neurais; além dos demais controles, como o de pH e temperatura.

Utilizando o sistema com todas as suas funcionalidades chegou-se a uma produtividade de biomassa em cultivos de *E. coli* de mais de 9 g(massa seca).L<sup>-1</sup>.h<sup>-1</sup> (sistemas mais complexos usando membranas de perfusão relatam produtividade máxima inferior a 8 g.L<sup>-1</sup>.h<sup>-1</sup> (NAKANO et al., 1997)), com significativa diminuição de problemas operacionais e de erros de operação pelos usuários.

#### 4.1. A SUPERVISION AND CONTROL TOOL BASED ON ARTIFICIAL INTELLIGENCE FOR HIGH CELL DENSITY CULTIVATIONS

Antonio Carlos Luperni Horta<sup>1</sup>, Adilson José da Silva<sup>2</sup>, Cíntia Regina Sargo<sup>1</sup>, Ana Maria Velez<sup>1</sup>, Marina de Carvalho Gonzaga<sup>1</sup>, Roberto de Campos Giordano<sup>1</sup>, Viviane Maimoni Gonçalves<sup>3</sup>, Teresa Cristina Zangirolami<sup>1</sup>

<sup>1</sup>Universidade Federal de São Carlos, Departamento de Engenharia Química  
e-mail: horta@ufscar.br

<sup>2</sup>Universidade Federal de São Carlos, Biotecnologia  
e-mail: adiljs@gmail.com

<sup>3</sup>Instituto Butantan, Centro de Biotecnologia  
e-mail: vivimaig@butantan.gov.br

**Abstract** – High cell density cultivations of recombinant *E. coli* have been increasingly used for the production of heterologous proteins. However, it is a challenge to maintain these cultivations within the desired conditions, given that some variables such as dissolved oxygen concentration (DOC), feed flow rate and temperature are difficult to control. SUPERSYS\_HCDC is a tool specially developed to oversee high cell density cultivations of *rE. coli*. The tool was implemented in Labview 8.0 and Matlab 6.5 and includes some special features such as: automatic control of the DOC by integrated action at the stirrer speed as well as at the air and oxygen flow rates; automatic start-up of the feed flow of fresh medium (system based on neural network committee); automatic slow down of feeding when oxygen consumption exceeds the maximum capacity of oxygen supply. The tool allows remote access and warnings indicating specific faults in the system by voice alarms or SMS messages can be activated by the user. Twenty-two fed batch cultures with five different clones of recombinant *E. coli* were carried out for continuous upgrade of SUPERSYS\_HCDC full functionalities. Due to the improved control and supervision environment provided by the tool, cell productivities up to 9 g.L<sup>-1</sup>.h<sup>-1</sup> were achieved, which is one of the highest values ever reported in the literature for *E. coli* fed-batch cultures.

**Keywords** – High cell density cultivations; monitoring and control system; recombinant *E. coli*; fed-batch; Artificial intelligence; Computational intelligence



#### 4.1.1. INTRODUCTION

Fed-batch cultivations of recombinant *Escherichia coli* (*rE. coli*) are used in the production of numerous products of the pharmaceuticals industry like insulin, growth hormones, vaccines and interferons (Liljeqvist and Stahl, 1999; Eiteman and Altman, 2006; Tripathi et al., 2009). Regardless the product of interest, the cultures are carried out with the initial goal of achieving high cell density and then proceed to the induction of the desired protein.

The operation mode most widely employed in high cell density cultivations (HCDC) of *rE. coli* is the fed-batch culture with exponentially increasing feed flow rate (Shiloach and Fass, 2005). However, when running a HCDC, operational difficulties are commonly found in both growth and protein expression phases. Throughout the cultivation, substrate concentration must be kept at the desired level, because its accumulation can cause inhibition of cell growth and protein expression. On the other hand, an excessive shortage of substrate supply can hamper growth and protein synthesis. Furthermore, it can lead to decreasing viability and even to cell death if the rate of carbon supplied is not enough to meet the maintenance requirements. For this reason, a precise control of the feed flow rate of supplementary medium (F) is required and F must be adjusted to the demand at every stage of cultivation, which can vary considerably after the induction due to metabolic stress associated to the expression of heterologous proteins (Dabros et al., 2010).

Besides preventing substrate accumulation, F should restrict the growth rate in order to avoid the formation of metabolic by-products (Lee, 1996; Kilikian et al., 2000). The formation of acetate, the main by-product of *E. coli* metabolism, is triggered whether by the so-called overflow metabolism or by oxygen depletion. The mechanism known as overflow occurs mainly in cultures with glucose as carbon source, when growth is maintained above the critical specific growth rate ( $\mu_C$ ) and a rapid assimilation of the carbon source takes place (Rocha, 2003). In order to avoid overflow metabolism, the specific growth rate is limited below a critical value limiting the replenishment of the carbon source (Rocha, 2003; Demain and Vaishnav, 2009). This procedure ensures no accumulation of substrate or acetate. But, in turn, it decreases the productivity of the process by limiting the growth rate.

Other problems can occur frequently in HCDC conducted under conditions of exponential growth, such as excessive heating of the reactor, maintenance of dissolved oxygen below desirable levels, early (or late) start-up of the feed pump, foam formation, overpressure, among others.

Due to all above mentioned reasons, a full-time, well trained operator is required to oversee a high cell density cultivation, which can last 30 hours or more. Alternatively, an artificial intelligence based tool could be used to assist or even replace the human operator in the process supervision.

The objective of this study was to develop, test and validate the tool called SUPERSYS\_HCDC, a supervisory system for monitoring and control of *rE. coli* HCDC.

#### 4.1.2. MATERIALS AND METHODS

##### 4.1.2.1. MICROORGANISMS

Five recombinant *E. coli* BL21(DE3) strains, modified with the following plasmids, were cultivated under the supervision of SUPERSYS\_HCDC: a) pET37b+/PspA3, PspA245 and PspA4Pro, kindly provided by Dr. Eliane Miyaji (Centro de Biotecnologia, Instituto Butantan, São Paulo, Brazil), expressing a fragment of the membrane protein of *Streptococcus pneumoniae* (Carvalho et al., 2011); b) pET28a/SpaA, producing a recombinant fragment of the SpaA protein from *Erysipelothrix rhusiopathiae* (Silva et al., 2011b); c) pT101/D-TOPO, kindly donated by Laboratorio de Biocatálisis, ICP-CSIC-Madri, Spain, with *pac* gene codifying for penicillin G acylase (PGA) production (Montes et al., 2007).

##### 4.1.2.2. MEDIUM

All the cultivations were carried out in fed-batch mode. A modified defined medium composition (Seeger et al., 1995; Sargo et al., 2011), containing glycerol as carbon source, was used in 17 cultures of *rE. coli* strains pET37b+/PspA3, PspA245, PspA4Pro and pT101/D-TOPO. Five cultures of *rE. coli* strains pT101/D-TOPO and pET28a/SpaA were performed using modified complex media (Silva et al, 2011a), based on LB or ZYM-50552 auto-induction media compositions (Studier, 2005), containing glucose,

glycerol or a mixture of both as main carbon sources and yeast extract and tryptone as nutrients. The medium composition is described in Table 4.1.1.

Table 4.1.1. Medium composition

Component	<i>Escherichia coli</i> SpaA and TOPO		<i>Escherichia coli</i> PspA3, PspA245, PspA4Pro and TOPO	
	Batch	Fed solution	Batch	Fed solution
Glucose (g/L)	10	240		
Glycerol (g/L)			40	800
Tryptone (g/L)	10	10		
Yeast extract (g/L)	5	5		
(NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub> (g/L)			5.33	6.4
K <sub>2</sub> HPO <sub>4</sub> (g/L)	7	7		
KH <sub>2</sub> PO <sub>4</sub> (g/L)			17.73	21.28
MgSO <sub>4</sub> ·7H <sub>2</sub> O (g/L)	2.5	20	1.6	20
NaCl (g/L)	5	5		
CoCl <sub>2</sub> ·6H <sub>2</sub> O (mg/L)			3.33	4
CuCl <sub>2</sub> ·2H <sub>2</sub> O (mg/L)			2	2.3
Fe(III)citrate (mg/L)			133.33	40
MnCl <sub>2</sub> ·4H <sub>2</sub> O (mg/L)			20	23.5
Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O (mg/L)			2.8	4
Zn(CH <sub>3</sub> COOH) <sub>2</sub> ·H <sub>2</sub> O (mg/L)			33.8	16
Citric acid·H <sub>2</sub> O (g/L)			2.27	2.27
H <sub>3</sub> BO <sub>3</sub> (mg/L)			4	4.7
EDTA (mg/L)			18.8	13
Thiamine (mg/L)			45	45
Kanamycin (mg/mL)	30	30	50	50
Polipropilenoglicol 30% (mL/L)	1	1	1	1
pH	7.0	7.0	6.7	6.9

#### 4.1.2.3. BIOREACTOR OPERATION AND INSTRUMENTATION

The twenty-two fed-batch cultures were carried out in a 5.0 L home-made bioreactor and monitored by SUPERSYS\_HCDC<sup>®</sup> (Horta et al., 2011). The complete experimental set-up is shown in Figure 4.1.1. The pH was controlled (on/off) at 6.7 (pHmeter GLI PRO) by addition of NH<sub>4</sub>OH (30%) in the batch phase and at 6.9 in the feeding phase. Temperature was set at 25, 30, 35 or 37°C (depending on the strain being cultivated). The dissolved oxygen concentration (DOC) was monitored using a dissolved oxygen sensor (Mettler Toledo Inpro 6800 probe, connected to a CE O<sub>2</sub> 4050 transmitter), and maintained at 30% by a hybrid controller (PID + Heuristic) which automatically changed both agitation speed (between 200 and 900 rpm) and the composition of the gas stream supplied to the bioreactor (by mixing pure oxygen with

air). The total gas flow rate was maintained at 4 - 6 L/min by two mass flow controllers (GFC AALBORG). The broth permittivity and conductivity were monitored by a Biomass Sensor (FOGALE® Nanotech). The exhaust gas composition was assessed by a Sick/Maihak S.710 CO<sub>2</sub> and O<sub>2</sub> analyzer. On-line data acquisition as well as monitoring/control of all instruments via a compact field point 2020 (cFP-2020, National Instruments) was performed by the software SuperSys\_HCDC<sup>R</sup> developed in LabView® and MatLab®. Supplementary feed flow rate was controlled by SuperSys\_HCDC<sup>R</sup> to match the growth requirements through on-line estimation of specific growth rate based on permittivity data (Horta et al., 2011c). All listed instruments presented data acquisition in real time, with acquisition intervals selected by the user, which was every 10 s for the experiments described here.

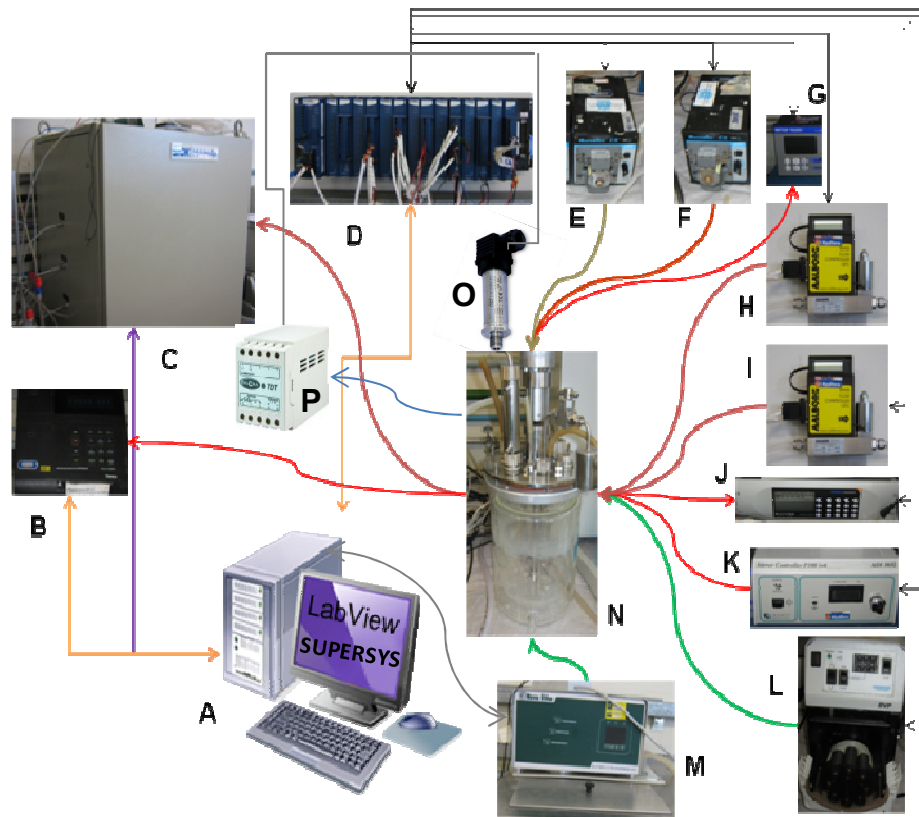


Figure 4.1.1. Experimental set-up: "A" - the computer program SUPERSYS\_HCDCR, "B"- pH meter, "C" - exhaust gas analyzer, "D" - cFP; "E" and "F" - acid and base pumps, "G" - dissolved oxygen sensor, "H" - mass flow controller of air supply, "I" - mass flow controller of oxygen supply, "J"- biomass sensor; "K" - stirring speed controller, "L" - feed pump, "M" - cooling/heating water bath, "N"- bioreactor, "O" - pressure transducer and "P" – temperature transducer.

For all experiments, the exponential feeding flow rate was calculated using Equation 4.1.1 (Nielsen et al., 2002), and it was automatically controlled by the supervisory system.

$$F = \left( \frac{\mu_{\text{SET}}}{Y_{\text{XS}}} + m \right) \cdot \frac{C_{\text{X0}} \cdot V_0}{C_{\text{S0}} - C_{\text{SR}}} \cdot e^{(\mu_{\text{SET}} \cdot t)} \quad (4.1.1)$$

In Equation 4.1.1,  $F$  ( $\text{Lh}^{-1}$ ) is the feed flow rate,  $\mu_{\text{SET}}$  ( $\text{h}^{-1}$ ) is the desired specific growth,  $Y_{\text{XS}}$  ( $\text{g}_{\text{biomass}} \cdot \text{g}_{\text{glycerol}}^{-1}$ ) is the biomass yield coefficient on glycerol,  $m$  ( $\text{g}_{\text{biomass}} \cdot \text{g}_{\text{glycerol}}^{-1} \text{h}^{-1}$ ) is the maintenance coefficient,  $C_{\text{X0}}$  [ $\text{g}$  (dry cell weight)  $\text{L}^{-1}$ ] and  $V_0$  ( $\text{L}$ ) correspond to the cellular concentration and volume, respectively, at the beginning of the fed-batch phase,  $C_{\text{S0}}$  ( $\text{g}_{\text{glycerol}} \text{L}^{-1}$ ) is the carbon source concentration in the supplementary medium and  $C_{\text{SR}}$  ( $\text{g}_{\text{glycerol}} \text{L}^{-1}$ ) represents the residual glycerol concentration.

The strategy for controlling the exponential feed flow rate changed accordingly the cultivation (Silva et al, 2011; Sargo et al., 2011). It evolved from the classical control strategy, where  $\mu_{\text{SET}}$ ,  $Y_{\text{XS}}$  e  $m$  are taken as constants, based on prior knowledge, to a more sophisticated approach where, these parameters were continuously retuned at each 10 min using the values of  $\mu$  obtained online from the permittivity measurements provided by the capacitance biomass sensor (Horta et al, 2011c).

#### 4.1.2.4. ANALYTICAL PROCEDURES

Cellular growth was followed by culture broth optical density reading (OD,  $\lambda = 600$  nm), dry cell weight (DCW) measurements ( $C_{\text{x}}$ :  $\text{g}_{\text{DCW}} \cdot \text{L}^{-1}$ ), counting CFU. $\text{mL}^{-1}$  and on-line estimation by measuring broth permittivity ( $\text{pF} \cdot \text{cm}^{-1}$ ).

Metabolite concentrations were assessed by HPLC (Waters Corp. system), using an Aminex HPX-87H column (Bio-Rad) and a 5 mM sulfuric acid solution as mobile phase (flow rate of  $0.6 \text{ mL min}^{-1}$ ), at  $60^\circ\text{C}$ . Organic acids were detected at 210 nm (Waters 486 UV detector), while glycerol, glucose, lactose and galactose were measured with a refractive index detector (Waters 410).

#### 4.1.2.5. SUPERVISORY AND CONTROL TOOL

The SUPERSYS\_HCDC software comprises modules that perform from basic functions to even real time inference of important variables. Carbon dioxide evolution rate (CER), oxygen uptake rate (OUR) and respiratory quotient (RQ) are calculated taking into account the on-line molar fractions of CO<sub>2</sub> (Y<sub>CO2</sub>) and O<sub>2</sub> (Y<sub>O2</sub>) in the exhaust gas, the air (Q<sub>air</sub>) and oxygen (Q<sub>O2</sub>) inlet flow rates (Eqs. 4.1.2, 4.1.3 and 4.1.4).

$$\text{CER} = \frac{P}{R \cdot T} \cdot Q_{\text{air}} \cdot \frac{0.79}{1 - Y_{\text{CO}_2} - Y_{\text{O}_2}} \cdot Y_{\text{CO}_2} \quad (4.1.2)$$

$$\text{OUR} = \frac{P}{R \cdot T} \cdot \left( 0.21 \cdot Q_{\text{air}} + Q_{\text{O}_2} - Y_{\text{O}_2} \cdot Q_{\text{air}} \cdot \frac{0.79}{1 - Y_{\text{O}_2} - Y_{\text{CO}_2}} \right) \quad (4.1.3)$$

$$\text{RQ} = \frac{n_{\text{CO}_2}}{n_{\text{O}_2}} \quad (4.1.4)$$

In Eqs 4.1.2 to 4.1.4, R is the ideal gas constant, P the atmospheric pressure and T the temperature, both at STP.

On-line cell concentration (Cx\_perm) data was generated after treating permittivity signal with a smoothed moving average (SMA) filter. Cx\_perm was further used for inference of growth rate (r<sub>x</sub>) and specific growth rate (μ), as described in Horta et al, (2011c). Bioreactor volume (V) was also continuously updated as function of samples withdrawal and feeding medium supplied.

The on-line estimated μ values enabled the implementation of an automatic control of the feeding flow rate (F) defined by Eq. 4.1.1 through the dynamic updating of the specific growth rate. For this purpose, each 10 minutes μ<sub>SET</sub> was replaced by μ<sub>DYN</sub>, which was obtained after processing a vector containing 50 samples with the SMA filter (Horta et al, 2011c).

To cope with the changes in cell metabolism taking place after induction, which strongly impact the specific growth rate as well as the maintenance and biomass yield coefficients, a module for automatic fitting of Y<sub>xs</sub> and m at Eq. (4.1.1) was set-up and integrated to the control program. The parametric fitting system is based on the global

(Simulated Annealing, Particle Swarm, Differential Evolution) and local search (Levenberg-Marquadt) algorithms to estimate, inside a confidence region, the optimal values for  $Y_{xs}$  and  $m$  (Horta et al, 2011c).

Besides the automatic updating of key parameters in Eq. (4.1.1), the feed flow rate control also incorporated a modulation on the growth rate to maintain DOC at the desired set-point. Thus,  $F$  control only followed Eq. 4.1.1 if the DOC was  $\sim 30\%$  of saturation. When OUR exceeded the maximum bioreactor oxygen transfer rate (OTR), at the maximum stirring speed and maximum oxygen flow rate, the  $F$  control combined Eq. 4.1.1 to the logic described at Figure 4.1.2, preventing oxygen depletion and the formation of the undesirable metabolic by-products, such as acetic acid.

**Procedure:** DOC - Feeding restrictions

**Start**

if  $DOC < DOC_{min}$  Then  $\alpha = 0.9$

else  $\alpha = 1$

$$C1 = (\alpha \cdot \mu / Y_{xs} + m) \cdot C_{x0} \cdot V_0 / (C_{S0} - C_{SR})$$

$$C2 = \alpha \cdot \mu_{DYN}$$

$$F = C1 \cdot \exp(C2 \cdot t)$$

**End**

Figure 4.1.2. Pseudocode for combined control of feed flow rate

To complete the automation of feed supply, a special module for automatic start-up of the feed pump was developed and implemented. The identification of the end of the batch phase was performed by a neural network based softsensor, which produced a command signal to start the feed pump (Horta et al., 2011a).

To match the typically high OUR of *E. coli* HCDC, the DOC control was accomplished by developing a hybrid system, which combined the classical PID (proportional, integral and derivative) controller directly actuating on the stirrer speed to a heuristic algorithm modulating the flow rates of air and oxygen. The PID controller is described at Figure 4.1.3, where “Agit” refers to the stirrer speed (rpm), “Kc” to the proportional constant, “Td” the derivative constant, “Ti” to the integral constant, “SetP” to the DOC set-point and “t” to time (s).

$$\begin{aligned}
 & \text{Agit}(t) = \text{Agit}(t-1) + \\
 & + K_c \cdot \left\{ \begin{aligned} & \left[ \begin{aligned} & (\text{SetP} - \text{OD}(t)) + \\ & - (\text{SetP} - \text{OD}(t-1)) \end{aligned} \right]^+ \\ & + \frac{\Delta t}{T_i} \cdot (\text{SetP} - \text{OD}(t)) + \\ & + \frac{T_d}{\Delta t} \cdot \left( \begin{aligned} & (\text{SetP} - \text{OD}(t)) + \\ & - 2 \cdot (\text{SetP} - \text{OD}(t-1)) + \\ & + (\text{SetP} - \text{OD}(t-2)) \end{aligned} \right) \end{aligned} \right\}
 \end{aligned}$$

Figure 4.1.3. Pseudo-code for the dissolved oxygen concentration control module.

The control of air (QAIR) and oxygen (QO2) flow rates is performed by an independent algorithm based on the heuristic logics presented at Figure 4.1.4.

```

Procedure: oxygen enrichment control
Start
  If Agit < 0,7*AgitSL and QO2.1 > 0
    QAIR = QAIR + Δ
    QO2 = QO2.1 - Δ
  If Agit ≥ 0,99*AgitSL and OD < ODIL
    If QAIR ≥ (QMAX - QO2.1)
      QAIR = QAIR - Δ
      QO2 = QO2 + Δ
    If not
      QAIR = Δ + QAIR
      QO2 = QO2.1
End

```

Figure 4.1.4. Pseudo-code for the control of air and oxygen flow rates

Another important feature developed and implemented SUPERSYS\_HCDC is a set of specific alarms for fault detection and warning, whose pseudo-code is given in Figure 4.1.5. The alarm system warns the operator via the computer loudspeakers as well sending text messages via SMS.



```

Procedure: alarms
Start
If pH>1.05*pHsetPoint then Message= fault at acid pump;
If pH<0.95*pHsetPoint then Message= fault at base pump;
If T>1.05*TsetPoint then Message= decrease the bath temperature;
If T<0.95*TsetPoint then Message= raise the bath temperature;
If OD<0.5*ODsetPoint then Message= increase the total gas flow rate or the stirrer speed;
If QO2<0.95*QO2setPoint then Message= raise the pressure of oxygen supply line;
If QAir<0.95*QAirsetPoint then Message= raise the pressure of air supply line;
If P>PSL then Message= Excessive overpressure in the reactor;
End

```

Figure 4.1.5. Pseudo-code for the set of alarms.

#### 4.1.3. RESULTS AND DISCUSSION

SUPERSYS\_HCDC was upgraded, tested and validated by using it as supervisory and control system in 22 fed-batch cultures of recombinant *E. coli*, 19 of them being high cell density cultures (final biomass concentration over 50 g<sub>DCW</sub>/L). The most important results showing the performance of the developed supervisory system concerning the dynamic control of the supplementary medium flow rate and the automatic start-up of the feed pump have been already detailed (Horta et al, 2011 a, b). The results concerning the on-line estimation of CER, OUR and RQ as well as the DOC control are highlighted in the followings.

##### 4.1.3.1. ON-LINE ESTIMATION OF CER, OUR, RQ AND $\mu$

Figure 4.1.6b exhibits the results of CER, OUR and RQ (respiratory quotient) on-line calculation from the gas analyzer measurements of CO<sub>2</sub> and O<sub>2</sub> mole fractions in the exhaust gas leaving the bioreactor (Eqs. 4.1.2, 4.1.3 and 4.1.4). These data can be compared to the trends in permittivity signal, biomass experimental points and  $\mu$  inferred from de biomass sensor, which are displayed in Figure 4.1.6a, b, c.

In this experiment, cells were exposed to the inducer (lactose) from the beginning, but protein expression was actually intensified 4h after the onset of cultivation, when the glucose present in the batch medium was totally consumed. Figure 4.1.6a shows that the on-line signs (perm and CER) and the off-line (Cx) showed similar increasingly

profiles within ~ 4 h of culture. On the other hand, after induction intensification, while the measures of Cx (Figure 4.1.6a) indicated that growth continued at a slow specific growth rate (decreasing from ~ 0.1 to 0.05 h<sup>-1</sup>), CER, OUR, RQ and permittivity on-line data (Figure 4.1.6b) suggested severe growth stagnation and even decline in metabolism. This mismatch happens because optical density readings or dry cell weight measurements do not distinguish viable from non-viable cells. Conversely, CER, OUR and permittivity data reflect the physiological state of the cells. Thus, the stressful condition imposed by the prolonged induction impact cell viability leading to the growth cessation after 10 h of cultivation, as reproduced by the specific growth rate (Figure 4.1.6c) estimated from permittivity data.

These results highlight the importance of permittivity, CER and OUR data for physiological state identification as well as for setting up suitable control strategies of feed supply during recombinant *E. coli* cultivations, since these signs indicate a decrease in metabolism while the dry weight continued to rise.

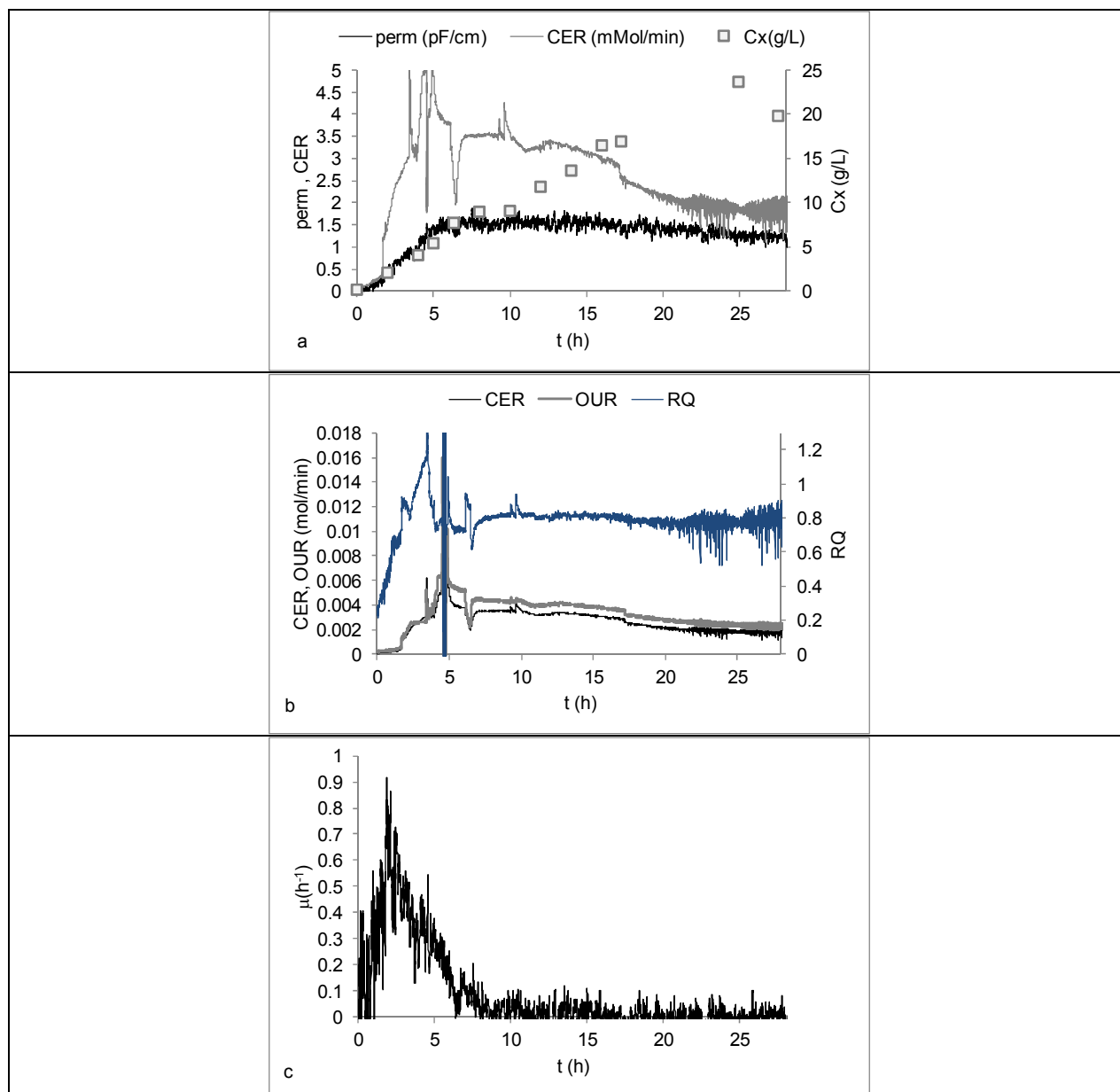


Figure 4.1.6. Fed-batch cultivation of *E. coli* BL21(D3) expressing a fragment of SpaA antigenic protein. Culture carried out with auto-induction complex medium, containing glucose, glycerol and lactose as carbon sources, at 37 °C (Silva et al, 2011). a) permittivity (perm – pF/cm), carbon evolution rate (CER – mmol/min), and cellular concentration (Cx – g/L); b) CER (mol/min), oxygen uptake rate (OUR – mol/min) and respiratory quotient (RQ); c) specific growth rate ( $\mu$  –  $h^{-1}$ ).

#### 4.1.3.2 HYBRID CONTROLLER OF DOC

*E. coli* cultures are characterized by their high oxygen demand. Kuprijanov et al. (2009), emphasize the importance of fine controlling the DOC and proposed a very efficient feedforward/feedback controller based on a gain schedule approach. The controller was set-up with air flow rate and stirring speed as manipulated variables and validated in fermentations of *rE. coli* releasing a green fluorescent protein.

However, to meet the OUR requirements, the control strategy must include not only the usual actuators (stirrer speed and total gas flow rate) but also be combined with new actuators (oxygen and air flow rates) to promote the gradual oxygen enrichment of inlet air. For this purpose, a hybrid controller actuating on the stirrer speed, air and oxygen flow rates was implemented as previously described. The performance of the DOC hybrid controller can be seen in Figure 4.1.7 A and B. The desired levels of DOC were kept throughout the cultivation (Figure 4.1.7A) by the automatic changes in  $Q_{air}$ ,  $Q_{O_2}$  and Agit (Figure 4.1.7B).

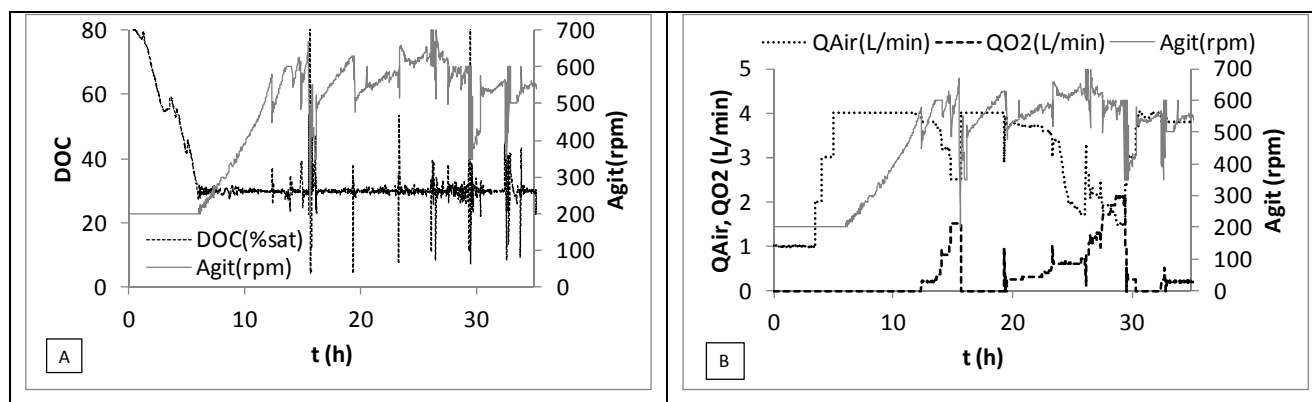


Figure 4.1.7. Fed-batch cultivation 1 of *rE. coli* BL21(DE3) expressing a fragment of PspA3 antigenic protein. Culture carried out with modified HDF medium using glycerol as carbon source, at 30°C (Sargo et al, 2011); A) Stirrer speed (Agit -rpm), dissolved oxygen concentration (DOC - %saturation); B) Air flow rate ( $Q_{air}$  – L/min), oxygen flow rate ( $Q_{O_2}$  – L/min) and stirrer speed. Culture carried out with auto-induction complex medium, containing glucose, glycerol and Lactose.

The occasional oscillations observed at DOC data are due to the variations at the air and oxygen flow rates which are related to the mass flow controllers operation ranges. Mass flow controllers specified to operate at maximum flow rates of  $\sim 5$  L/min have accuracy limits in the range of  $\pm 0.3$  L/min. The accuracy of the flow rates supplied

by the mass flow controllers was also affected by the head loss increase due to the biomass build-up in the broth that also led to higher viscosity.

#### 4.1.3.3. DOC CONTROL BY GROWTH RATE MODULATION

Actually, the maximum achievable biomass concentration in a fed-batch culture is limited by the highest oxygen transfer rate (OTR) which can be supplied by the own bioreactor system. For *rE. coli* HCDC special care must be taken to prevent carbon source accumulation, which can easily happen if growth is limited by OTR. Furthermore, when OUR exceeds OTR, low levels of DOC are established, driving cell metabolism towards the production of undesirable fermentative products. Thus, an additional module was included in the DOC controller so that the feeding flow rate was manipulated to keep the DOC at the set point when the bioreactor was operated at its maximum OTR.

Figure 4.1.8 shows the performance of DOC controller when the modulation by the feeding flow rate was activated.

Shortly before 18 h of cultivation, Figure 4.1.8A and C shows that both oxygen flow rate and the stirring speed were at their top limits. In spite of that, the DOC (Figure 4.1.8D) presented a decreasing profile. At this moment, the controller automatically reduced the feeding flow rate to decrease the growth rate and, consequently, oxygen consumption. The effects of the control action can also be seen at Figure 4.1.8H and J: substrate and acetate accumulation profiles (between 14 and 18 h) were reverted after the feeding was slowed down. The combined controller performance was satisfactory and DOC was kept between 10 to 40 % of saturation throughout the cultivation. But the performance could be improved if the ON/OFF control of feed pump is replaced by a smoother stepwise reduction of the flow rate to promote a less oscillatory DOC control.

Analysing the Figure 4.1.8F, it's observed that the permittivity estimated concentration matched perfectly the off-line biomass concentration points (dry weight measurements) up to induction moment. After the inducer is added, the viability decrease is depicted by the permittivity data but not by the off-line biomass concentration measures, as discussed previously. The cumulative alkali consumption (Figure 4.1.8I) for pH control also shows a close correlation to biomass concentration. The pH remained at the desired set-point throughout the experiment, with slightly higher values after induction (Figure 4.1.8E).

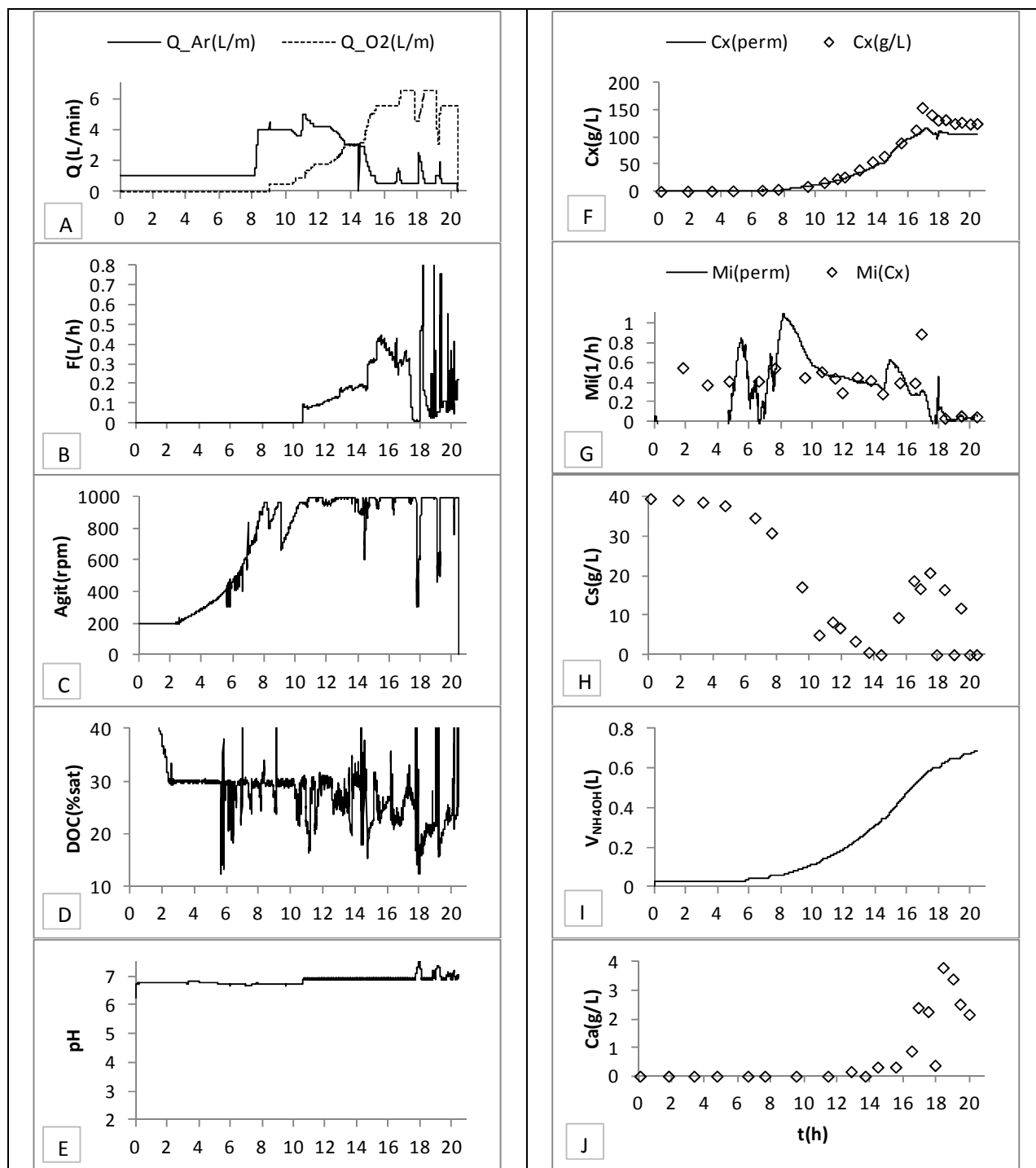


Figure 4.1.8. Typical time profiles of stirring speed, feed flow rate, air and oxygen flow rates and other variables for the combined DOC control. Fed-batch cultivation 2 of *rE. coli* BL21(DE3) expressing a fragment of PspA3 antigenic protein. Culture carried out with modified HDF medium using glycerol as carbon source, at 35°C (Sargo et al, 2011). A) air and oxygen flow rates in L/min; B) feed flow rate in L/min; C) stirrer speed in rpm; D) dissolved oxygen concentration as % of saturation; E) pH; F) biomass concentration on-line estimated from permittivity data and measured off-line by dry weight method; G) on-line specific growth rate inferred from both permittivity data and dry weight measurements; H) substrate concentration (glycerol); I) cumulative volume of  $\text{NH}_4\text{OH}$  added; J) acetate concentration in g/L. Induction with IPTG at 16.5h.

Concerning the specific growth rate ( $\mu$ ) estimated from permittivity data (Figure 4.1.8G), the oscillatory behavior registered up to  $\sim 7.5$  h is related mainly to the low resolution of the permittivity measures and to the high polarization state of the cell membrane (Abi et al., 2010). After the induction, a sharp decrease of  $\mu$  and a gradual increase of substrate concentration are observed as consequences of the cell metabolic effort to produce the heterologous protein.

#### 4.1.4. CONCLUSIONS

The adequate control, monitoring and supervision of a bioprocess is an important step to approach batch to batch reproducibility, quality control high standards as well as more robust and cost effective processes, less prone to human errors. These issues are crucial for the production of biopharmaceuticals and are present in both FDA and GMP regulations.

In this work, the development of the automatic supervision and control tool SUPERSYS\_HCDC is presented and discussed. The tool includes several up-to-date functionalities, such as: automatic combined control of the DOC; automatic start-up of the feed pump; remote access and warnings system by voice alarms or SMS messages. It was set-up in LabView environment, enabling a friendly user interface, with flexibility of choices by the operator for all tasks.

The tool was designed to be used as support in HCDC of *rE. coli* and so far it has been extensively tested in 22 fed-batch cultures of 5 different recombinant strains of this bacterium. But it has also been used to supervise batch and fed-batch cultures of other industrially important microorganisms like *Pichia pastoris*, *Bacillus subtilis*, *B. megaterium* and *Saccharomyces cerevisiae*.

The tool revealed to be extremely useful for correctly running long lasting ( $\sim 40$  h) HCDC, when the bioreactor is operated by several users with different practice skill. Important decisions such as the right moment for starting the feed pump, oxygen and air flow rates to be used, suitable feeding profiles could be left to the tool.

In the specific case of HCD cultures of *rE. coli*, the tool showed a reliable performance when controlling the DOC and the nutrient supply even considering the high level of difficulty imposed by the high biomass concentrations and the metabolic

drifts triggered by protein expression. It certainly contributed to achieve a biomass concentration of 155 gDCW/L in only 17 h of culture, leading to a maximum cell productivity of  $9.2 \text{ g}_{\text{DCW}} \cdot \text{L}^{-1} \cdot \text{h}^{-1}$  without metabolite formation or substrate accumulation.

The SUPERSYS\_HCDC tool will be available as free software at [www.ladabio.deq.ufscar.br](http://www.ladabio.deq.ufscar.br).

ACKNOWLEDGEMENTS: The authors would like to thank the São Paulo Research Foundation (FAPESP) and Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) for funding this work, and Tiago Martins Pereira, Amadeus Gomes de Azevedo and Eudoro Lemos de Oliveira Neto for technical support.

#### NOMENCLATURE

HCDC	High cell density cultivation
$C_S$	Substrate Concentration (g/L);
$C_{S0}$	Concentração de substrato no meio de alimentação (g/L);
$C_{SR}$	Residual Substrate Concentration (g/L);
$C_X$	Cellular Concentration (g/L);
$C_{X0}$	Cellular Concentration (g/L) in the beginning of the feeding stage;
OD	Optical Density (600nm);
F	Feeding rate (L/h)
G	Grams
H	Hour
AI	Artificial Intelligence
L	Liter
M	Maintenance coefficient ( $\text{g}_x/\text{g}_s \cdot \text{h}$ )
$\mu$ or $M_i$	Specific growth rate ( $\text{h}^{-1}$ )
$\mu_{\text{CRIT}}$	Critical specific growth rate ( $\text{h}^{-1}$ ), from which occurs overflow
$\mu_{\text{MAX}}$	Maximum specific growth rate ( $\text{h}^{-1}$ )
$\mu_{\text{SET}}$	Specific growth rate limited by the fed flow rate ( $\text{h}^{-1}$ )
$\mu_{\text{DYN}}$	Dynamic specific growth rate coupled in the fed flow rate ( $\text{h}^{-1}$ )
S	Second
DOC	Dissolved Oxygen Concentration (% of saturation)
T	Time (h)
$V_0$	Volume of cultivation medium at the beginning of feeding stage (L)
$Y_{XS}$	Yield coefficient ( $\text{g}_x/\text{g}_s$ )



## REFERENCES

- Abi, A.; Sarrafzadeh, M.H.; Mehrnia, M.R.; Ghommidh, C., Application of Dielectric Permittivity Measurements in Physiological State Monitoring of *Bacillus subtilis* Culture. 2010 2nd International Conference on Chemical, Biological and Environmental Engineering (ICBEE 2010) (2010).
- Babaeipour, V.; Shojaosadati, S.A.; Robatjazi, S.M.; Khalilzadeh, R.; Maghsoudi, N., Over-production of human interferon- $\gamma$  by HCDC of recombinant *Escherichia coli*. Process Biochemistry, 42:112-117 (2007).
- Carvalho, R.J.; Cabrera-Crespo, J.; Tanizaki, M.M.; Gonçalves, V.M., Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. Appl Microbiol and Biotechnol. DOI: 10.1007/s00253-011-3649-9 (2011).
- Dabros, M.; Schuler, M.M.; Marison, I.W., Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. Bioprocess Biosyst Eng 33:1109-1118 (2010).
- Demain, A.L.; Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. Biotech. Advances, 27, 297-306 (2009).
- Eiteman, M.A.; Altman, E., Overcoming acetate in *Escherichia coli* recombinant protein fermentations. Trends in Biotechnology, v. 24 (11), 530-536 (2006).
- Horta, A.C.L.; Sargo, C.R.; Silva, A.J.; Gonzaga, M.C.; Santos, M.P.; Gonçalves, V.M.; Zangirolami, T.C.; Giordano, R.C., Intensification of high cell density cultivations of rE. coli for production of *S. pneumoniae* antigenic surface protein, PspA3, using model-based adaptive control. Bioprocess and Biosystems Engineering, 2012.
- Horta, A.C.L.; Zangirolami, T.C.; Giordano, R.C.; Cruz, A.J.G.; Reis, G.B.; Jesus, C.D.F., Supervisory system for bioreactor high cell density cultivations, Registered software, Process number 11008-6, RPI 2115, INPI, Brazil (2011).
- Horta, A.C.L.; Silva, A.J.; Sargo, C.R.; Gonçalves, V.M.; Zangirolami, T.C.; Giordano, R.C., Robust artificial intelligence tool for automatic start-up of the supplementary medium feeding in recombinant *E. coli* cultivations. Bioprocess Biosyst Eng;34(7):891-901 (2011a).
- Horta, A.C.L.; Gonzaga, M.C.; Sargo, C.R.; Possedente, M.; Silva, A.J.; Gonçalves, V.M.; Zangirolami, T.C.; Giordano, R.C., Software para Estimativa em Tempo Real de Parâmetros Cinéticos para Cultivos em Alta Densidade Celular. XVIII Simpósio Nacional de Bioprocessos, 24 a 27 de julho de 2011, Caxias do Sul-RS. (In Portuguese) (2011b).
- Horta, A.C.L.; Sargo, C.R.; Silva, A.J.; Gonzaga, M.C.; Santos, M.P.; Gonçalves, V.M.; Zangirolami, T.C.; Giordano, R.C., Intensification of high cell density cultivations of rE.

*coli* for production of *S. pneumoniae* antigenic surface protein, PspA3, using model-based adaptive control. Submitted to Bioprocess Biosyst Eng, December (2011c).

Horta, A.C.L.; Silva, A.J.; Sargo, C.R.; Giordano, R.C.; Gonçalves, V.M.; Zangirolami, T.C., A Supervision and Control Tool Based on Artificial Intelligence For High Cell Density Cultivations (Ferramenta de Supervisão e Controle Baseada em Inteligência Artificial Para Cultivos de Alta Densidade Celular). XVIII Congresso Brasileiro de Engenharia Química, Foz do Iguaçu/PR, 2010.

Kilikian B.V; Suárez I.D; Liria C.W; Gombert A.K, process strategies to improve heterologous protein production in *Escherichia coli* under lactose or iptg induction. *Process biochemistry* 35:1019–1025 (2000).

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

Kuprijanov, A.; Gnoth, S.; Simutis, R.; Lübbert, A., Advanced control of dissolved oxygen concentration in fed batch cultures during recombinant protein production. *Appl Microbiol Biotechnol*, 82:221-229 (2009).

Lee, S.Y. High cell density culture of *Escherichia coli*. *TIBTECHT* 14, 98-105 (1996).

Liljeqvist, S.; Stahl, S., Production of recombinant subunit vaccines: protein immunogens, live delivery systems and nucleic acid vaccines. *J. Biotechnol.*, v. 73, p. 1-33 (1999).

Montes, T.; Grazú, V.; López-Gallego, F.; Hermoso, J.A.; Garcia, J.L.; Manso, I.; Galán, B.; González, R.; Fernandez-Lafuente, R.; Guisán, J.M., Genetic modification of the Penicillin G Acylase surface to improve its reversible immobilization on ionic exchangers. *Applied and environmental microbiology*, 312-319 (2007).

Nielsen, J.; Villadsen, J.; Lidén, G., *Bioreaction Engineering Principles*, 2nd edition. New York, Kluwer Academic/Plenum Publishers (2002).

Rocha, I.C.A.P., Model-based strategies for computer-aided operation of recombinant *E. coli* fermentation. Dissert. Para Doutor. Em Eng. Quím. e Biológ., Escol. De Eng. Universidade de Minho (2003).

Sargo, C.R.; Horta, A.C.L.; Silva, A.J.; Possedente, M.; Machado, M.M.P.; Giordano, R.L.C.; Giordano, R.C.; Gonçalves, V.M.; Zangirolami, T.C., Improved *rE. coli* High Cell Density Cultivation Conditions for Glycerol as Carbon Source. Submitted to *Applied Microbiology and Biotechnology*, December, (2011).

Seeger, A.; Schneppe, B.; McCarthy, J.E.G.; Deckwer, W.D.; Rinas, U., Comparison of temperature- and isopropyl-P-D-thiogalacto-pyranoside-induced synthesis of basic fibroblast growth factor in high-cell-density cultures of recombinant *Escherichia coli*. *Enzyme and Microbial Technology*, 17:947-953 (1995).

Shiloach, J.; Fass, R., Growing *E. coli* to high cell density: a historical perspective on method development. *Biotech Adv.*, v. 23, p. 345-357 (2005).

Silva, A.J.; Horta, A.C.L.; Escalón, A.M.V.; Iemma, M.R.C.; Sargo, C.R.; Giordano, R.L.C.; Novo, M.T.M.; Giordano, R.C.; Zangirolami, T.C., Non-conventional induction strategies for production of subunit swine erysipelas vaccine antigen in *rE. coli* fed batch cultures. Submitted to Microbial Cell Factories, December (2011a).

Silva, A.J.; Iemma, M.R.C.; Horta, A.C.L.; Sargo, C.R.; Giordano, R.L.C.; Giordano, R.C.; Zangirolami, T.C.; Novo, M.T.M., Cloning, expression, purification and auto-induction production of rSpaA swine erysipelas antigen. Submitted to Current Microbiology, December (2011b).

Studier, F.W., Protein production by auto-induction in high-density shaking culture. *Protein Express Purif* 41:207-234 (2005).

Suárez, D.C.; Killikian, B.V., Acid acetic accumulation in aerobic growth of recombinant *E. coli*. *Process Biochem* 35, 1051-1055 (2000).

Tripathi, N.K.; Sathyaseelan, K.; Jana, A.M.; Rao, P.V.L., High Yield Production of Heterologous Proteins with *Escherichia coli*, *Defence Science Journal*, 59 (2):137-146 (2009).

Vélez A.M.; Horta, A.C.L.; Silva, A.J.; Sargo, C.R.; Mateo, E.C.; Zangirolami, T.C.; Giordano, R.L.C., Produção em frascos agitados e em batelada alimentada de Penicilina G Acilase (PGA) de *Escherichia coli* Recombinante com o meio sintético HDF. Anais do XVIII Simpósio Nacional de Bioprocessos, Caxias do Sul/RS, 2011.(In Portuguese)

Zhang, J.; Greasham, R., Chemically defined media for commercial fermentations. *Appl. Microbiol. Biotechnol.* 51:407-421 (1999).

## 5. MONITORAMENTO E CONTROLE DE CULTIVOS DE ALTA DENSIDADE CELULAR

Esta seção é composta por artigos que detalham o desenvolvimento de duas funcionalidades especiais inclusas no software de controle SUPERSYS\_HCDC e avaliam o desempenho do sensor de capacitância no monitoramento de cultivos de diferentes microrganismos.

O primeiro artigo trata do sistema de partida automática da bomba de alimentação, intitulado “Robust artificial intelligence tool for automatic start-up of the supplementary medium feeding in recombinant *E. coli* cultivations” (publicado em *Bioprocess and Biosystems Engineering*, em março de 2011). Este trabalho apresenta uma metodologia racional para seleção da arquitetura de um sistema de inteligência artificial, baseado num comitê de redes neurais que define o fim da batelada e aciona a bomba de alimentação. O sistema foi treinado e validado em cultivos de *E. coli* recombinante, e a configuração da rede neural mais robusta apresentou as seguintes entradas: velocidade de agitação, vazão de ar e de oxigênio, produção de CO<sub>2</sub> e fração molar de CO<sub>2</sub> nos gases de saída, com uma camada oculta com 10 neurônios e função log-sigmoide nas camadas oculta e de saída.

O segundo artigo (submetido a *Bioprocess and Biosystems Engineering* em novembro de 2011), intitulado “Intensification of high cell density cultivations of r*E. coli* for production of *S. pneumoniae* antigenic surface protein”, PspA3, using model-based adaptive control” avalia uma metodologia inovadora para controlar cultivos de alta densidade de *E. coli* em batelada alimentada baseada em medidas *on-line* de concentração de oxigênio dissolvido e concentração celular, e na estimativa da velocidade específica de crescimento (*on-line*), do coeficiente de rendimento ( $Y_{xs}$  - *at-line*) e do coeficiente de manutenção ( $m$  - *at-line*). A idéia principal é permitir que o microrganismo cresça de acordo com sua capacidade metabólica, sem as limitações inerentes das estratégias de crescimento pré-determinadas. Utilizando esta estratégia conseguiu-se uma produtividade de 9,2 g.L<sup>-1</sup>.h<sup>-1</sup> em massa seca, mantendo-se elevada produção de proteína recombinante, 26,1 g/L (produtividade de proteína: 1,3 g.L<sup>-1</sup>.h<sup>-1</sup>),

mais de oito vezes superior a concentração de proteína de 3,5 g/L (produtividade de proteína: 0,12 g.L<sup>-1</sup>.h<sup>-1</sup>), relatada em estudos anteriores que cultivaram o mesmo clone de *E. coli* expressando a mesma proteína (BARAZZONE *et al.*, 2011; CARVALHO *et al.*, 2011).

O terceiro artigo apresenta um estudo do comportamento do sinal de permissividade em diferentes condições de cultivos e para diferentes microrganismos, intitulado “On-Line Inference of Biomass Concentration In Bioreactors Using A Capacitance Sensor: Assessing The Methodology For Different Bacteria And Yeast Cultures” (em finalização). O monitoramento *online* pelo sensor de biomassa é avaliado para alguns dos microrganismos mais importantes da indústria biotecnológica: *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris* e *Bacillus megaterium*, comprovando a validade da correlação linear entre permissividade e biomassa medida por massa seca e por DO na fase de crescimento em todas as situações estudadas.

## 5.1. ROBUST ARTIFICIAL INTELLIGENCE TOOL FOR AUTOMATIC START-UP OF THE SUPPLEMENTARY MEDIUM FEEDING IN RECOMBINANT *E. COLI* CULTIVATIONS

Antônio Carlos Luperni Horta<sup>1</sup>, Adilson José da Silva<sup>1</sup>, Cíntia Regina Sargo<sup>1</sup>, Viviane Maimoni Gonçalves<sup>2</sup>, Teresa Cristina Zangirolami<sup>1</sup>, Roberto de Campos Giordano<sup>1</sup>.

<sup>1</sup> *Departamento de Engenharia Química, Universidade Federal de São Carlos, Brazil*

<sup>2</sup> *Centro de Biotecnologia, Instituto Butantan, São Paulo, Brazil*

**Corresponding author:** Roberto de Campos Giordano

P.O. Box 676, São Carlos, SP

Zip code: 13565-905

Phone: 55 16-3351-8708; FAX: 55 16-3351-8266

**Abstract.** One of the most important events in fed-batch fermentations is the definition of the moment to start the feeding. This paper presents a methodology for a rational selection of the architecture of an artificial intelligence (AI) system, based on a neural network committee (NNC), which identifies the end of the batch phase. The AI system was successfully used during high cell density cultivations of recombinant *Escherichia coli*. The AI algorithm was validated for different systems, expressing three antigens to be used in human and animal vaccines: fragments of surface proteins of *Streptococcus pneumoniae* (PspA), clades 1 and 3, and of *Erysipelothrix rhusiopathiae* (SpaA). Standard feedforward neural networks (NNs), with a single hidden layer, were the basis for the NNC. The NN architecture with best performance had the following inputs: stirrer speed, inlet air and oxygen flow rates, carbon dioxide evolution rate, and CO<sub>2</sub> molar fraction in the exhaust gas.

**Keywords:** *fed-batch bioreactor, recombinant Escherichia coli, neural networks committee, high cell density cultivation.*

### 5.1.1 INTRODUCTION

The increasing competition in the biotechnology industry poses several challenges nowadays for bioprocess engineers: lower production costs, higher productivities and fine-tuned control of product quality [1]. Within this scope, automated process monitoring and control is a key factor for the successful operation of modern bioreactors. In batch and fed-batch processes, small changes in the operating conditions may impact severely the quality of the desired product, which is frequently a bio-molecule whose concentration is analyzed off-line [2].

The cultivation of microorganisms, usually carried out in batch or fed-batch bioreactors, is a very complex dynamic process. The microorganisms' metabolic state changes during the process, following non-linear dynamics. Frequently, complex cultivation media are used, which are mixtures of several nutrients whose composition/characteristics may change from run to run. There is a strong dependence on previous process stages (strain selection, inoculum preparation and adaptation), which may randomly affect the bioreactor performance. Data sets are usually multidimensional and complex [3, 4].

*E. coli* is a usual host for the production of recombinant proteins, and its cultivation has been extensively studied for decades [1, 5], and fed-batch operation of the bioreactor is a common choice [6]. *E. coli* high cell density cultures (HCDC) usually comprise three stages: a) batch phase, where growth is supported by the nutrient present in the initial medium; b) fed-batch phase, where additional medium is supplied to the bioreactor, usually in an exponential feed rate; c) induction phase, where the heterologous protein production takes place after the addition of the inducer [7]. The extent of a HCDC can exceed 50 h of operation and constant process surveillance by skillful operators is often required to adjust the operational conditions in response to changes in cell metabolism occurring at each phase. The right moment to shift from batch to fed-batch phase is particularly critical. If the feed supply is started too early, growth inhibition by the substrate and/or production of organic acids due to the overflow metabolism [8] may occur. On the other hand, if the feed supply begins too late, the shortage of nutrients may lead to loss of viability. In both cases, the overall process

productivity regarding biomass formation as well as recombinant protein expression will be greatly affected [9].

"Computational intelligence" (or "artificial intelligence") is a field that comprises computational algorithms inspired by the physiology of the nervous system (artificial neural networks) or by linguistic models (fuzzy logic), which can be used for decision making, fault detection, state inference and prediction. The use of artificial intelligence (AI) tools to control, monitor and supervise bioprocesses was frequently reported by the literature of the last decade. Some works addressing this topic are [10-24].

A neural network (NN) can be informally defined as a network of simple, interlinked, processors ("nodes" or "neurons"), with weights associated to each connection. A detailed review about NN theory and NN committees can be found in Haykin [25]. There are several works available in the literature reporting the application of NNs to bioreactors, which include data classification, noise filtration and reaction rates inference [26-28].

In this work, an AI-based decision system was applied to identify the optimum time to start the feeding phase during high density recombinant *Escherichia coli* cultivations aiming at the expression of immunogenic fragments of three surface proteins of infectious bacteria, which will be used for vaccines manufacturing.

Fragments of the *pneumococcal* surface protein A (PspA) from two clades (1 and 3) of *Streptococcus pneumoniae* were produced to be used in a new human vaccine formulation [29, 30]. *S. pneumoniae* is a leading biological agent responsible for many human infections [31], especially in developing countries, and recombinant PspA fragments are key components of a new conjugate vaccine (linked to fragments of the capsular polysaccharide of the bacterium, [32]).

Swine erysipelas is responsible for great economic losses worldwide. Attenuated or inactivated cells of *Erysipelothrix rhusiopathiae* have been used for vaccine preparations, but it has been reported that animal exposure to pathogen whole cells can aggravate arthritic lesions. Here, *E. coli* was the host to express a fragment of the antigenic portion of *E. rhusiopathiae* surface protein A (SpaA) to be used in the formulation of an animal subunit vaccine.





Experiments 1 to 4 correspond to Cultivations A1 to A4 in Table 5.1.4.3; While Experiment 5 and 6 correspond to id B1 and B2 from Table 3.2.2.

#### 5.1.2.2 EXPERIMENTAL PROCEDURE:

The cultivations were conducted in a 5 L in-house bioreactor monitored by SuperSys\_HCDC<sup>R</sup> [37]. The complete experimental set-up is showed in Figure 5.1.1. The pH was controlled (on/off) at 6.7 (pHmeter GLI PRO) by addition of NH<sub>4</sub>OH (30%). Temperature was set at 30 °C (experiments 1 to 4) or 37 °C (experiments 5 and 6) and dissolved oxygen concentration (DOC) was monitored by a Mettler Toledo probe Inpro 6800, connected to CE O<sub>2</sub> 4050 transmitter, and kept at 30% of saturation by a PID controller which automatically changed the stirrer speed between 200 and 900 rpm. The gas stream supplied to the bioreactor was a mixture of air and pure oxygen and its total flow rate was maintained at 4 L/min by two mass flow controllers (GFC AALBORG). The broth permittivity and conductivity were monitored by a Biomass Sensor (FOGALE® Nanotech). The exhaust gas composition was assessed by a Sick/Maihak S.710 CO<sub>2</sub> and O<sub>2</sub> analyzer. On-line data acquisition as well as monitoring/control of all instruments via a compact field point 2020 (cFP-2020, National Instruments) was performed by the software SuperSys\_HCDC<sup>R</sup> developed in LabView®. Supplementary feed supply was controlled by SuperSys\_HCDC<sup>R</sup> to match the growth requirements through on-line estimation of specific growth rate based on permittivity data [34].

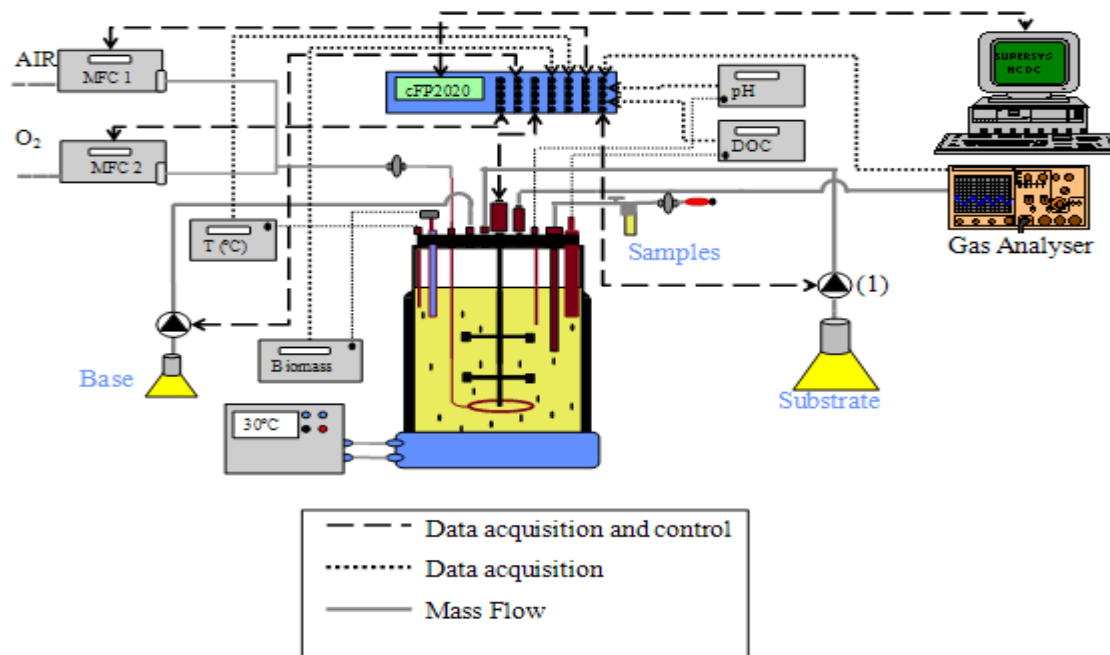


Figure 5.1.1 Schematic view of the bioreactor experimental set-up

#### 5.1.2.3 ANALYTICAL METHODS:

Culture samples were collected at each two or three hours for off-line measurement of key variables. Biomass formation was estimated by the optical density (OD,  $\lambda = 600$  nm) and dry cell weight method. Metabolites concentration and carbon sources consumption were measured by HPLC as described in Silva et al. [38].

### 5.1.3 METHODOLOGY

#### 5.1.3.1 CONSTRUCTION OF THE AI SYSTEM:

Six different HCDCs of *rE. coli* were used in this work, separated in a training dataset (Exp 1 to Exp 4) and in a validation dataset (Exp 5 and Exp 6). The training and validation datasets were purposely very different experiments (see Table 5.1.1). They encompassed distinct expression systems, operational conditions and medium compositions. In particular, complex media were used in the validation runs, whereas the training dataset was generated with defined media. This is a very important feature, since experiments with complex media are more susceptible to faults of the AI system, due to its greater variability of composition. Hence, using these experiments for validation (Exp 5 and Exp 6) implied a strong validation test.

Figure 5.1.2 is a schematic representation of the approach used to design the AI system. This problem has several dimensions. First of all, two conceptions were tested for the system. The first one, called “generic neural network” (GNN) consisted in training one single neural network with all the training data set, i.e., using all the four different experiments. The second approach consisted in training “expert neural networks” (ENN), i.e., one NN for each experiment, and then seeking for generalization using a committee of these ENNs (the Neural Network Committee, NNC). In practice, the second approach would be preferred for routine applications, since the knowledge added after a new experiment can be easily aggregated to the NNC. Indeed, it is much easier to train an ENN for each new dataset and then include it in the NNC, than to start retraining the GNN all over again (including the optimization of its architecture) with the expanded databank.

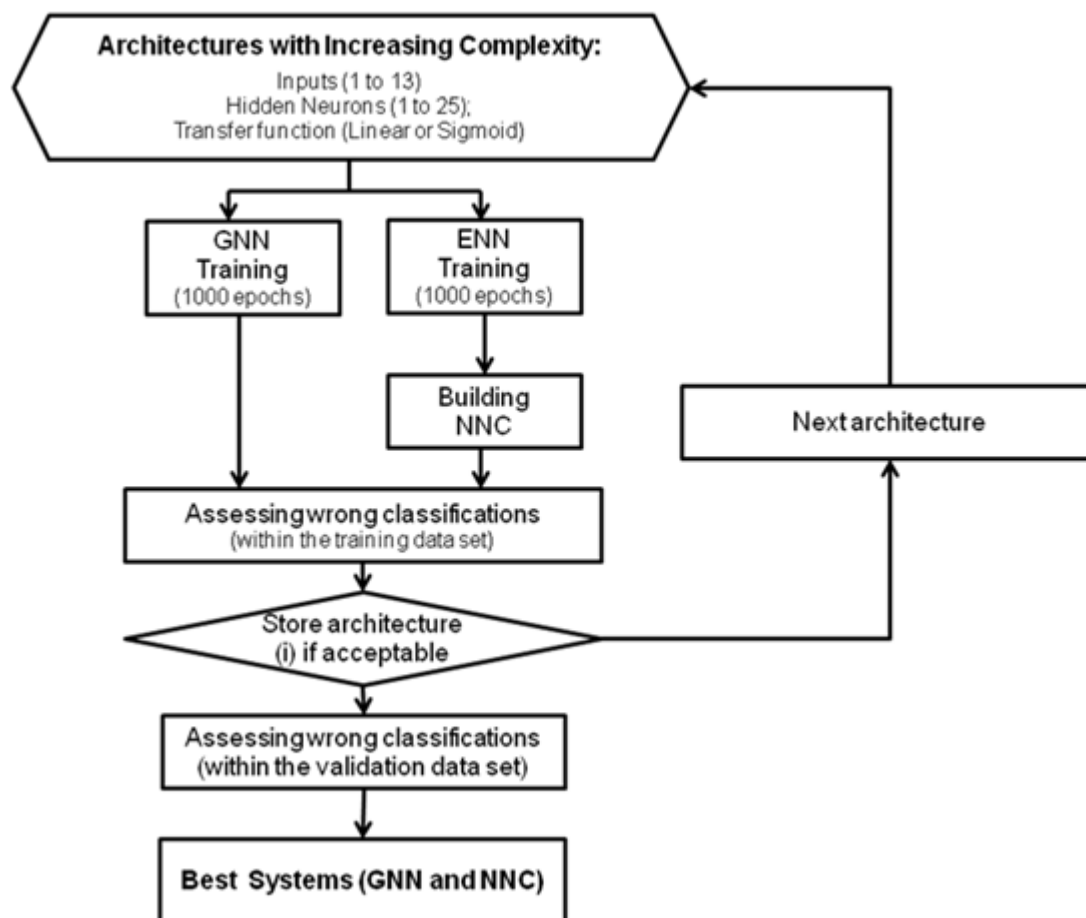


Figure 5.1.2 Schematic representation of the approach used to design the AI system

As it is shown in Figure 5.1.2, the two approaches (GNN and NNC) were compared and the best NN architecture was sought for each of them. The architecture of the NN is defined by the number of inputs, number of neurons in the hidden layer, and by the transfer functions of the neurons. Another crucial issue is the choice of what attributes (inputs) should be presented to the NNs [39].

The definition of the NN architecture is an optimization problem, concerning the model complexity. Its solution will reflect the tradeoff between the model accuracy of prediction (which increases with the complexity, reducing the bias error) and its capacity for generalization (which decreases with the complexity, since a model with more parameters is more prone to reflect the noise contained in the training data, as discussed by Nelles, [39]).

Several combinations of inputs were tested. The input variables were: stirring speed (Agit, rpm), the time derivative of Agit (dAgit), Agit derivative at the preceding time step (dAgit1), Agit derivative at two preceding steps (dAgit2), and at three preceding steps (dAgit3), dissolved oxygen concentration (DOC) (% of saturation), air flow rate (Q<sub>air</sub>) (L.min<sup>-1</sup>), oxygen flow rate (Q<sub>O<sub>2</sub></sub>), carbon dioxide evolution rate (CER)(mol/min), CO<sub>2</sub> molar fraction in the exhaustion gas (CO<sub>2</sub>), CO<sub>2</sub> derivative until three previous steps (dCO<sub>2</sub>, dCO<sub>21</sub>, dCO<sub>22</sub> and dCO<sub>23</sub>), as well as permittivity (Perm) (pF/cm) and conductivity (Cond) (mS/cm) of the culture medium, given by a capacitance sensor.

The signals coming from the instrumentation were transformed from mA to the respective units using calibration curves. The only exceptions were the permittivity and conductivity, which signals (in A×10<sup>4</sup>) were straight inputs to the NNs. Table 5.1.2 shows the range of architectures covered during this procedure, in a total of 52 different options. Derivatives were approximated by differences as given in equation 5.1.1. Carbon dioxide evolution rate (CER) was calculated on line using equation 5.1.2, where Y is the molar fraction of the specific gas, M indicate the measured variable, n (mol/min) is the molar flow rate.

$$\frac{d\text{Variable}}{dt}(i) \approx \frac{\Delta\text{Variable}}{\Delta t}(i) = \frac{\text{Variable}(i) - \text{Variable}(i-1)}{t(i) - t(i-1)} \quad (5.1.1)$$

$$\text{CER} = Y_{\text{CO}_2}^{\text{M}} \cdot \dot{n}_{\text{Air}} \cdot \frac{1 - Y_{\text{CO}_2}^{\text{Air}} - Y_{\text{O}_2}^{\text{Air}}}{1 - Y_{\text{CO}_2}^{\text{M}} - Y_{\text{O}_2}^{\text{M}}} - Y_{\text{CO}_2}^{\text{Air}} \cdot \dot{n}_{\text{Air}} \quad (5.1.2)$$

The sampling period was 10s. Analyzing these data, a group of human experts consensually determined the best moment to start the supply of supplementary medium (the qualitative reasoning behind this choice may be found, for instance, in Yamane et al., [40]).

Table 5.1.2 Tested NN architectures

Input	number of neurons of the hidden layer	transfer function
Agit, Q_Air, Q_O2	1, 5 or 10	Linear or log-sigmoid
Perm	1, 5 or 10	Linear or log-sigmoid
CER, CO2	5 or 10	Log-sigmoid
CER, CO2, dCO2	5 or 10	Log-sigmoid
Agit, Q_Air, Q_O2, dAgit	5 or 10	Log-sigmoid
Agit, Q_Air, Q_O2, dAgit, dAtit1, dAgit2	5, 10 or 15	Log-sigmoid
CER, CO2, dCO2, (plus successively adding dCO21, dCO22 dCO223)	5, 10 or 15	Log-sigmoid
Agit, Q_Air, Q_O2, dAgit, (plus successively adding dAtit1, dAgit2, dAgit3)	5, 10 or 15	Log-sigmoid
Agit, Q_Air, Q_O2, CER, CO2	5, 10 or 15	Log-sigmoid
Agit, Q_Air, Q_O2, CO2 (plus successively adding CER, Perm, Cond)	5, 10 or 15	Log-sigmoid
Agit, Q_Air, Q_O2, CER, CO2, dAgit, dAgit1, dAgit2, dAgit3	5, 10 or 15	Log-sigmoid
Agit, Q_Air, Q_O2, CER, CO2, dCO2, dCO21, dCO22, dCO23	5, 10 or 15	Log-sigmoid
Agit, Q_Air, Q_O2, CER, CO2, dAgit, dAgit1, dAgit2, dAgit3, dCO2, dCO21, dCO22, dCO23	5, 10, 15 or 20	Log-sigmoid

Each feed-forward NN was created by the function “newff” of Matlab 6.5, and trained with the function “train”, using the Levenberg-Marquardt algorithm for estimating the NN weights. The reader is referred to the Matlab manual for implementation details. The NNs were trained with fixed 1000 epochs (number of times that the data set is presented to the network). To reinforce the learning, each NN was trained five times and the NN with the lowest training error was chosen.

All data instances underwent a class assignment process. The human experts, identified as HEX, classified each instance of the datasets into one of the phases,

introducing a dummy response variable [28]. This new vector was formed by “zeros”, for the batch phase and “ones” for the feeding phase. Figure 5.1.3 shows the flowchart of the NN training procedure.

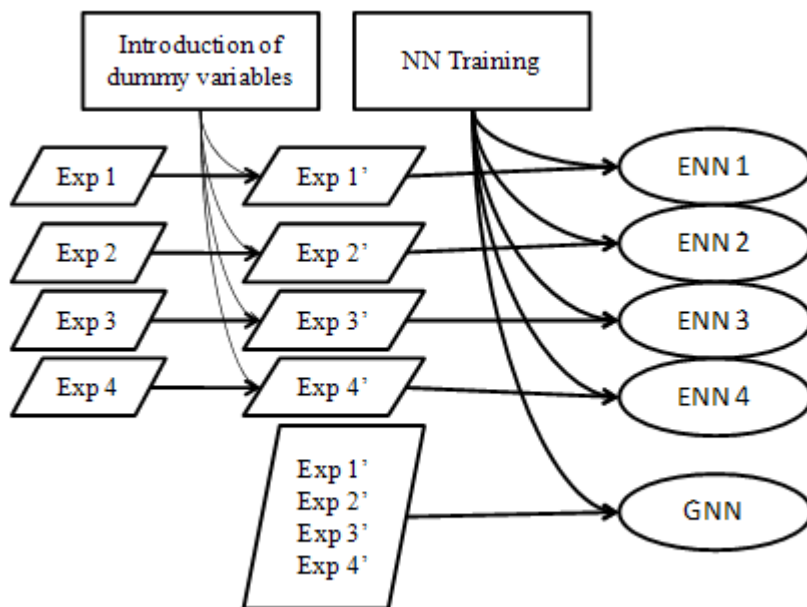


Figure 5.1.3 NN training flowchart. Exp refers to experimental data

Figure 5.1.4 shows the concept of the NNC. Each of the  $N$  expert neural networks (ENN) that integrate the committee (in the present case, four ENNs) delivers one output. A mediator module calculates the system final output using a simple average of the individual ENN outputs. It should be noticed that other criteria might be used by the mediator to negotiate among the ENN outputs. For instance, the center of gravity formulation [41] is a criterion that uses the inverse of the norm of the distance between the online signal and the input data sets of each ENN to weigh its output. This approach implicitly includes the time as a variable for the committee. But, in fact, the shape of the signal profiles is the important feature to be learnt by the intelligent system, since delays (or anticipations) are frequent in real fermentations. The lag phase and/or the kinetics of cellular growth may change from one experiment to the other, especially if complex media are used. The use of the distance from the online data and the dataset of old experiments would imply that these samples were taken at the same time. For one specific experiment, it may happen that, at this time, the growth has restarted after the

beginning of the feed. In this way, one might have contradictory signals to the mediator of the NNC.

Figure 5.1.4 illustrates how the NNC and the generalist NN, GNN, were tested, too. Note that both AI systems were submitted to the same testing procedure.

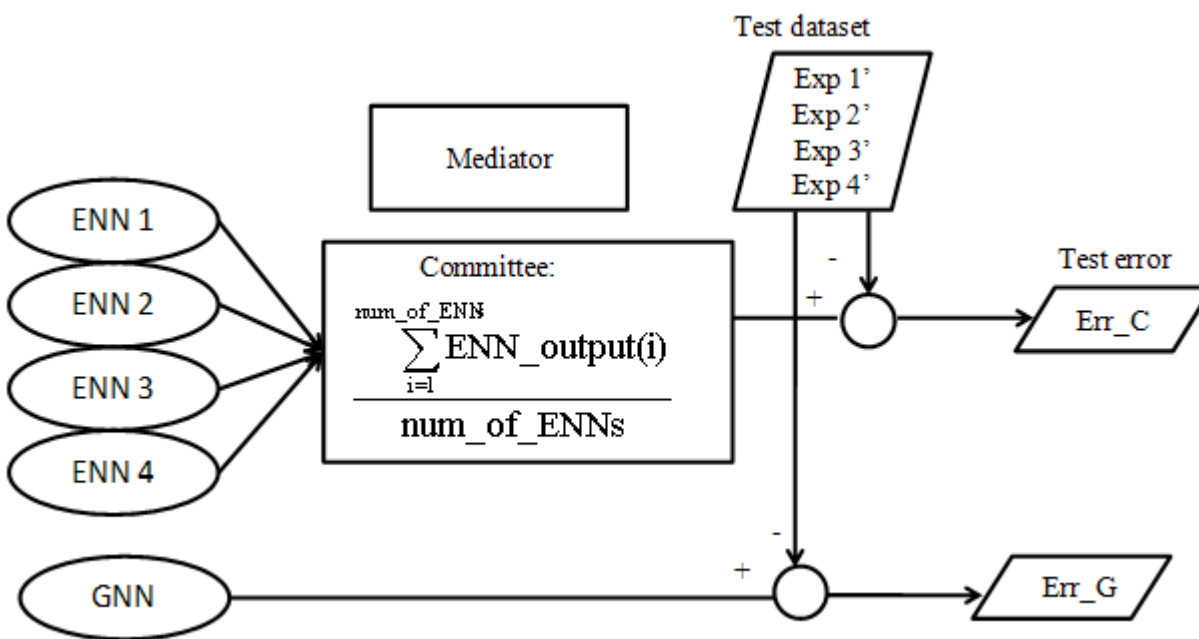


Figure 5.1.4 NNC construction and test configuration

Feedforward NNs outputs are real numbers, rather than binary logic operators (0 or 1). The NNs in this work have normalized output, between 0 and 1. For both AI systems, GNN and NNC, an output equal to 0.72 was selected for the threshold to the feeding phase – i.e., the supplementary medium pump was turned on when the output of the AI system was equal or higher 0.72. This rule of thumb showed to be very effective for generalization of the results.

### 5.1.3.2 GNN AND NNC VALIDATION

The eight configurations of the AI system with smaller test errors were validated in two different ways. A validation test was done using two new experiments (Exp 5 and 6); and the robustness of the system was assessed using simulations of faults in the data acquisition system.

In the new experiments used for validation, only the host was the same, *E. coli* BL21(DE3). The heterologous protein, the plasmid and the cultivation media were



different (*spaA* protein of *E. rhusiopathiae*, pET28a and complex LB medium, respectively). Exp5 used a mixture of glucose and glycerol as carbon sources during the batch phase, and glycerol in the feeding medium. Exp6 had only glycerol as carbon source during the batch, and a mixture of glycerol and lactose in the feeding medium.

The robustness tests were run offline, based on simulated faults in the data acquisition system. The dataset of Exp1 was modified for this purpose, as described in the following:

- 1) Exp1\_Agit\_0 - Agit vector was set equal to zero, to simulate a complete failure of the reading of the stirrer speed, during the entire experiment.
- 2) Exp1\_Agit\_pulse - a temporary failure of the agitation reading was simulated, including a time window of zeros in Agit.
- 3) Exp1\_CO2\_0 - CO<sub>2</sub> vector was set to zero, to simulate a complete failure in the signal coming from the carbon dioxide analyzer.
- 4) Exp1\_CO2\_pulse – similar to Exp1\_CO2\_0, but for simulating just a temporary failure of the CO<sub>2</sub> reading.

#### 5.1.4 RESULTS AND DISCUSSION

Six experiments were carried out using the experimental set-up described previously. Data from experiment 2 are depicted in Figure 5.1.5, illustrating the trends of some important process variables. The HEX data (the human experts classification), which were the basis for training the NNs, are shown in Figure 5.1.5, too. As it can be seen, during the batch phase the release of CO<sub>2</sub> increased exponentially, following the cellular growth. Simultaneously, the stirrer speed increased continuously up to 750 rpm, as a response to the increasing oxygen demand.

Thirteen hours after the onset of the culture, the substrate was almost depleted and consequently the release of CO<sub>2</sub> decreased sharply. Oxygen consumption was also significantly reduced, causing an increase in the DOC and a sharp drop in the agitation speed. These are the key indications for the start of the feeding phase. From this moment on, the slope of optical density data shifted from exponential to an almost constant rate of biomass accumulation modulated by the feed rate of supplementary medium. To train the NNs, the human experts (HEX) consensually defined a time window when the feed pump should be turned on (corresponding to a value of one in the

dataset, whose elements are otherwise zero). The AI system classification is considered 100% correct when it points out a time within this interval.

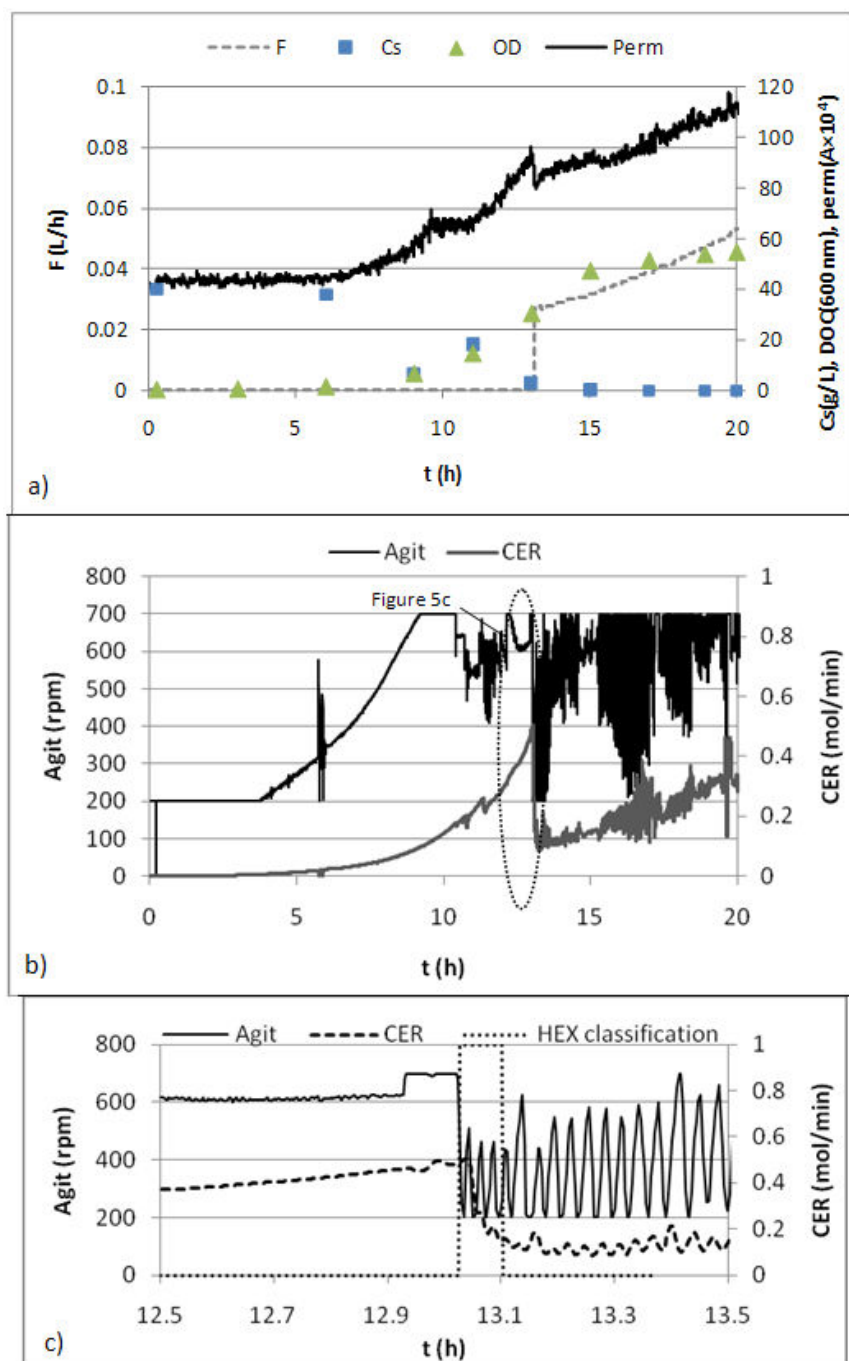


Figure 5.1.5 Exp 2: Experimental data for *rE. coli* expressing PspA3 cultivation. (a) Online signals: medium permittivity (Perm), feed flow rate (F); State variables: optical density (OD, measured at-line), and glycerol concentration ( $C_s$ , measured off-line through HPLC); (b) online signals: agitation (Agit) and carbon dioxide evolution rate (CER); (c) Zoom around the start-up of feeding; HEX is the human expert dummy variable (the desired output of the AI system)

Experimental data of four independent fed-batch runs (Exp1 to Exp 4) were employed for training the two AI machines, GNN and NNC. A total of 104 different combinations of inputs and configurations for these two systems were tried.

The eight AI systems with best performance were pre-selected for validation tests, always comparing the GNN and NNC configurations with the same architecture. As described previously, these systems had their capacity of generalization challenged in two kinds of tests: i) online validation, in completely new experiments, which were not used for training (Exp 5 and 6), and ii) offline validation, against simulated failures of the data acquisition system.

Figure 5.1.6a presents the results of these tests, for the eight NNCs and GNNs with better performance. Every experiment had a window in which it was assumed correct to turn on the feeding pump, as illustrated in Figure 5.1.5c (the HEX classification). The time delay or advance of the fed-batch phase classification by the AI systems was chosen as their performance index. It should be noticed, however, that a small bias towards advancing the moment to turn on the feed pump ( $\Delta t > 0$ ) is probably less harmful than a similar time delay ( $\Delta t < 0$ ), since the stress imposed by a depletion of carbon source that occurs in this case may decrease cell viability. Figure 5.1.6a indicates that positive errors were more important than the negative ones a promising result. The simulation of a complete failure in the reading of the rotation rate (Exp1\_Agit\_0) caused the greater advance errors ( $\Delta t > 0$ ) for almost all AI systems.

Figure 5.1.6 shows clearly that the most relevant attributes for phase identification are Agit and CER, which caused the largest number of misclassifications when faults were simulated for these signals.

The five best configurations of Figure 5.1.6a are highlighted in Figure 5.1.6b. From these top five, three are committees and two are generalist networks. From these results, NNC\_34 was finally selected as the best architecture, because: i) it correctly classified experiments 5 and 6 (passed a validation test), showing negligible error; GNN\_41, NNC\_39 and NNC\_48 also passed this test; ii) it had the lowest classification error for the Agit\_0 dataset, the fault for which the AI systems had the worst performance; iii) differently from NNC\_39, it does not use time derivatives, which amplify the signals noise, as inputs; iv) contrarily from NNC\_39 and GNN\_41, it does not use

the permittivity signal; the capacitance sensor for biomass is not a widespread detector in industrial (and laboratory) bioreactors, and an AI system that does not need this information is certainly in advantage for practical applications.

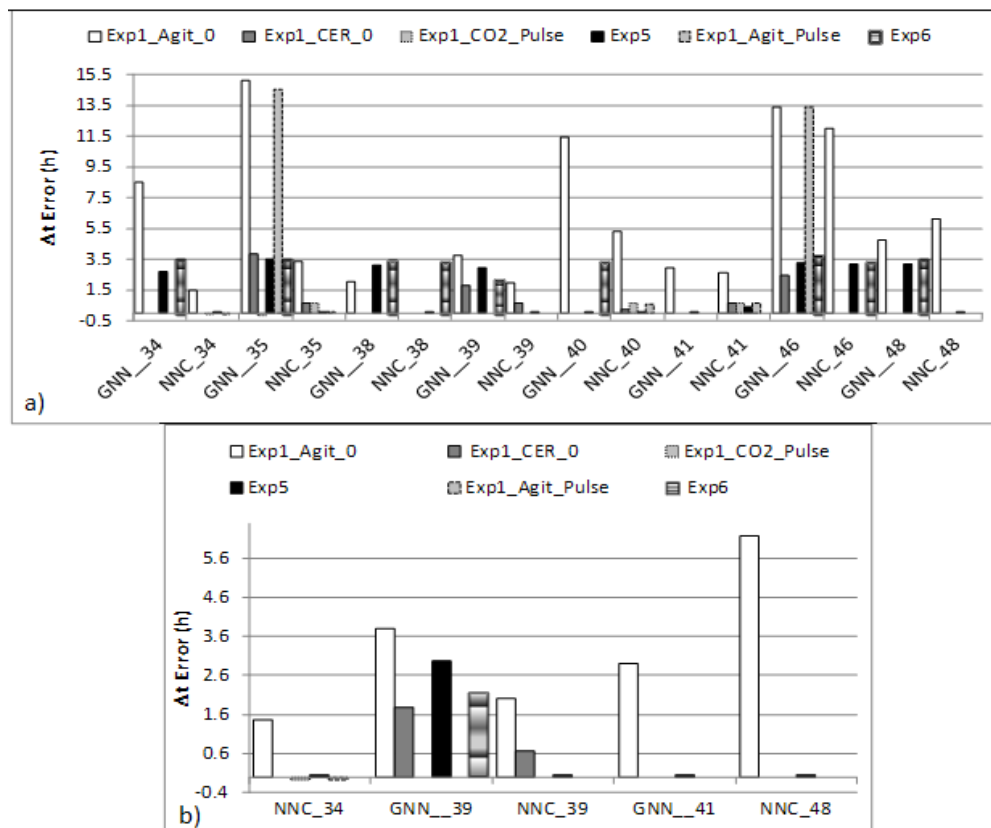


Figure 5.1.6 Validation of the AI systems: Time delay ( $\Delta t < 0$ ) or advance ( $\Delta t > 0$ ) of the start-up of the supplementary medium pump. (a) eight GNNs and NNCs with better performance; (b) five best AI systems

In fact, as Figure 5.1.5a shows, permittivity data follow the slow dynamics of cell growth, while the metabolic changes that occur when the carbon source is exhausted are much faster. Thus, these metabolic changes are better described by fast response variables directly related to the catabolic bioreactions, such as  $\text{CO}_2$  production and  $\text{O}_2$  consumption, which reflects on the stirring speed.

Table 5.1.3 details the five best systems, shown in Figure 5.1.6b. In conclusion, the selected AI system was composed by a NNC with 10 neurons in the hidden layer, log-sigmoid as activation function in the hidden and output layer and five inputs (Agit,  $F_{\text{air}}$ ,  $F_{\text{O}_2}$ , CER and  $\text{CO}_2$ ).

Table 5.1.3. Best AI system configurations after validation

ID	Inputs	Number of neurons in the hidden layer	Advance of the command to turn on the feed
NNC_34	<i>Agit, Far, FO2, CER, CO2</i>	10	0.06 h* 0.00 h** 0.33 h***
GNN_39	Agit, Far, FO2, nCO2, CO2, Perm, Cond	5	2.95 h* 2.14 h** 1.39 h***
NNC_39	Agit, Far, FO2, nCO2, CO2, Perm, Cond	5	0.06 h* 0.00 h** 0.66 h***
GNN_41	Agit, Far, FO2, nCO2, CO2, Perm, Cond	15	0.06 h* 0.00 h** 0.73 h***
NNC_48	Agit, Far, FO2, CER, CO2, dAgit, dAgit1, dAgit2, dAgit3, dCO2, dCO21, dCO22, dCO23	5	0.04 h* 0.00 h** 1.54 h***

\* Error for validation experiment #5 (Exp5)

\*\* Error for validation experiment #6 (Exp6)

\*\*\* Average error for simulated faults (Agit\_0, Agit\_pulse, CER\_0, CO2\_pulse)

One advantage of the AI system final configuration is that it could be used in bioreactors equipped with a conventional instrumentation: dissolved oxygen probe, the usual PID control to maintain the dissolved oxygen concentration at the set-point, mass flow meters for air and oxygen and a carbon dioxide gas analyzer, the less usual accessory in common practice, but still not a very expensive one.

Figure 5.1.7 shows the classification results of committee NNC\_34 for all datasets used in validation.

Figure 5.1.7 shows that NNC\_34 correctly classified all experiments, except the dataset with a null agitation vector. Based on its excellent performance, committee NNC\_34 was the recommended, in the present case, for recognizing the end of the batch and start of the feeding phase. Moreover, failures in the bioreactor control system keeping the stirring rate at zero for long periods of time are improbable, which increases the reliability of the system.

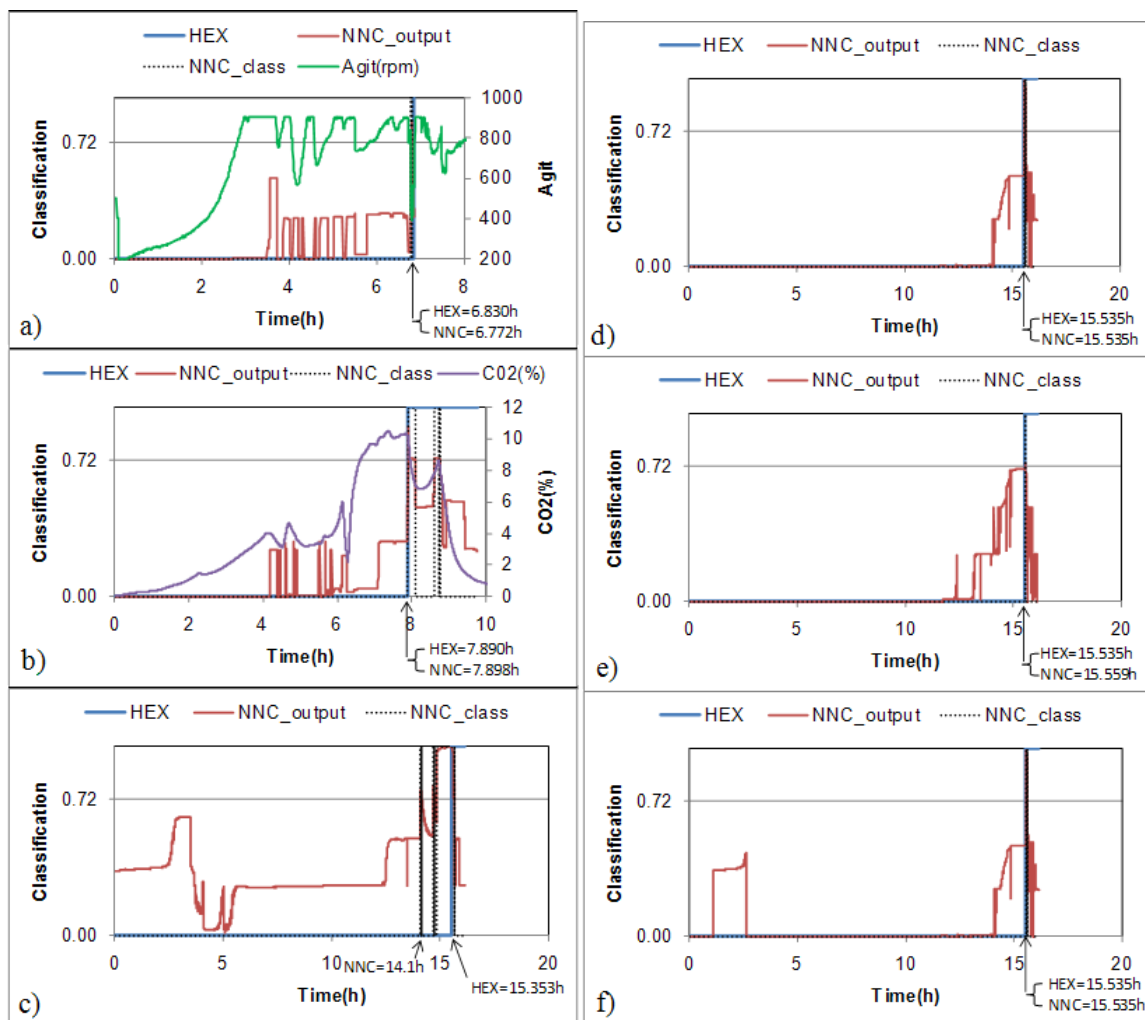


Figure 5.1.7 Validation results for committee NNC\_34. a) Exp5; b) Exp6 c) Exp1\_Agit\_0; d) Exp1\_CO2\_pulse; e) Exp1\_CO2\_0; f) Exp1\_Agit\_pulse. Outputs above 0.72 turned on the pump (NNC\_output in the plots). HEX are the human expert decisions to start the feeding, and NNC\_class are the AI decisions (both TRUE when equal to 1)

Besides presenting a good performance, the NNC has one great advantage over the generalist networks. Its modular nature enables the easy incorporation of new NNs as new experimental data becomes available, without any additional effort other than training the new network using one new data set. This is much simpler than retraining the GNN all over again, using an expanded data bank with all experiments, including the new one. It should be pointed out that Experiments 5 and 6, used for validation, were carried out in complex media, with *rE. coli* expressing a fragment of *E. rhusiopathiae* surface protein SpaA, while a *rE. coli* clone expressing another heterologous protein (a

fragment of *S. pneumoniae* surface protein PspA) cultivated in defined medium was used for training. This means that NNC\_34 was robust enough to cope with significant changes in the system, including medium composition and the *rE. coli* clone itself.

#### 5.1.5 CONCLUSION

Fed-batch high cell density cultivations of *rE. coli* are required for achieving high concentrations of the desired heterologous protein. The moment to start the feeding phase is very important to establish the pattern of cell growth, with impacts on product quality, yield and productivity. This paper presents a methodology to systematize the design of the architecture of an artificial intelligent system which can automatically decide the correct moment to turn on the feeding pump. That implies the on-line identification of the substrate depletion at the end of the batch phase.

The proposed system, a feedforward neural networks committee was able to classify the cultivation phase (batch or fed-batch) correctly, recognizing substrate depletion for both minimal (Exp1 until Exp4) and complex media (Exp5 and Exp6), having glycerol, mixtures of glucose and glycerol, and of glucose and lactose as carbon sources. Different faults in the instrumentation were simulated, and the system was robust enough to avoid false positives. The inputs for the NNC were four on-line measured variables and one calculated variable: the stirrer rotation speed (measured), the inlet flow rate of air and, when required, of supplementary oxygen (both measured), the molar fraction of carbon dioxide in the outlet gas (measured) and the carbon dioxide evolution rate (calculated). All four on-line data used by the proposed AI system are generally available in standard bioreactors, both in industrial and in lab scale. This fact makes it a tool with high potential for implementation in *rE. coli* and other relevant microorganism cultivations.

**Acknowledgements.** The authors thank the Brazilian research-funding agencies: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Financiadora de Estudos e Projetos (FINEP).

## REFERENCES

- [1] Choi, J.H.; Keum, K.C.; Lee, S.Y. (2006) Production of recombinant proteins by high cell density culture of *Escherichia coli*. Chem Eng Sci 61:876-885.
- [2] Xu Z, Weiwu Y, Huihe S (2006) Monitoring and Fault Diagnosis for Batch Process in Feature Extract in Fisher Subspace. Chin J Chem Eng 14(6):759-764.
- [3] Lopes JA, Menezes JC (2004) Multivariate monitoring of fermentation process with non-linear modelling methods. Anal Chim Acta 515:101-108.
- [4] Nucci, ER, Cruz, AJG, Giordano, RC (2009) Monitoring bioreactors using principal component analysis (PCA): Production of penicillin G acylase as a case study. Bioprocess Biosyst Eng (print).
- [5] Shojaosadati, S. A.; Kolaeil, S. M. V.; Babaeipour, V.; Farnoud, A. M. (2008) Recent advances in high cell density cultivation for production of recombinant protein. Iran J Biotechnol v 6, 2:3-84.
- [6] Shiloach, J. and Fass, R. (2005) Growing *E. coli* to high cell density: a historical perspective on method development. Biotechnol Adv 23:345-357.
- [7] Seeger A, Schbeppe B, Mccarthy JEG, Deckwer WD, Rinas U. (1995) Comparison of temperature and isopropyl- $\beta$ -D-thiogalacto-pyranoside-induced synthesis of basic fibroblast growth factor in high-cell density cultures of recombinant *Escherichia coli*. Enzym Microb Tech 17:947-953.
- [8] Rocha, I.C.A.P, (2003) Model-Based strategies for computer-aided operation of a recombinant *E. coli* fermentation. Dissertação para Doutoramento em Engenharia Química e Biológica, Univ. Minho.
- [9] Kim B.S.; Lee, S.C.; Lee, S.Y.; Chang, Y.K.; Chang, H.N. (2004) High Cell density fed-batch cultivation of *Escherichia coli* using exponential feeding combined with pH-stat. Bioprocess Biosyst Eng 26:147-150.
- [10] Alfafara, C.G., Miura, K., Shimizu, H., Shioya, K.-I.S., Suzuli, K., (1993) Fuzzy control of ethanol concentration and its application to maximum glutathione production in yeast-batch culture, *Biotechnol Bioeng* 41:493-501.
- [11] Shiba, S., Nishida, Y., Parque, Y.S., Ijima, S., Kobayashi, T. (1994) Improvement of cloned  $\alpha$ -amylase gene-expression in fed-batch culture of recombinant *Saccharomyces cerevisiae* by regulation both glucose and ethanol concentrations using a Fuzzy controller. Biotechnol Bioeng 44:1055-1063.
- [12] Lee J, Lee SY, Park S, Middelberg APJ (1999) Control of fed-batch fermentations. Biotechnol Adv 17:29-48.
- [13] Sousa Jr, R., Almeida, P.I.F., (2001) Design of fuzzy system for the control of a biochemical reactor in fed-batch culture. Process Biochem 461-469.
- [14] Hisbullah, M.A.H., Ramachandran, K.B., (2003) Design a fuzzy logic controller for regulating substrate feed to fed-batch fermentation. Inst Chem Eng 81:138-146.
- [15] Karim MN, Hodge D, Simon L. (2003) Data-based modeling and analysis of bioprocesses: some real experiences. Biotechnol Prog 19:1591-1605.
- [16] Komives C, Parker RS. (2003) Bioreactor state estimation and control. Curr Opin Biotechnol 14:468-474.



- [17] Mandenius C-F. (2004) Recent developments in the monitoring, modeling and control of biological production systems. *Bioprocess Biosyst Eng* 26:347-351.
- [18] Chen LZ, Nguang SK, Li XM, Chen XD. (2004) Soft sensors for on-line biomass measurements. *Bioprocess Biosyst Eng* 26:191-195.
- [19] Chen LZ, Nguang SK, Chen XD, Li XM. (2004) Modeling and optimization of fed-batch fermentation processes using dynamic neural networks and genetic algorithms. *Biochem Eng J* 22:51-61.
- [20] Alford JA. (2006) Bioprocess control: Advances and challenges. *Comput Chem Eng* 30:1464-1475.
- [21] Karakuzu, C., Turker, M., Ozturk, S. (2006) Modelling, on-line state estimation and fuzzy control of production scale fed-batch baker's yeast fermentation. *Control Eng Pract* 14:959-974.
- [22] Gnoth S, Jenzsch M, Simutis R, Lübbert A. (2008) Control of cultivation processes for recombinant protein production: a review. *Bioprocess Biosyst Eng* 31:21-39.
- [23] Nicoletti, MC, Giordano RC, Jain LC. Computational Intelligence Techniques for Bioprocess Modelling, Supervision and Control. In: Nicoletti MC, Jain LC. (2009) Computational intelligence techniques as tools for bioprocess modeling, optimization, supervision and control. Springer Verlag 218:1-23.
- [24] Mantovanelli ICC, Rivera EC, da Costa AC, Maciel Filho R. (2007) Hybrid neural network model of an industrial ethanol fermentation process considering the effect of temperature, *Appl Biochem Biotech*, 137-140: 817-833.
- [25] Haykin S. (1999) *Neural Networks: A Comprehensive Foundation*. Prentice Hall.
- [26] Simon L, Karim, NM (2001) Probabilistic neural networks using Bayesian decision strategies and a modified Gompertz model for growth phase classification in the batch culture of *Bacillus subtilis*. *Biochem. Eng. J.* 7(1):41-48.
- [27] Silva RG, Pinotti LM, Cruz AJG, Giordano RLC, Giordano RC. (2008) Artificial neural networks to infer biomass and product concentration during the production of penicillin G acylase from *Bacillus megaterium*. *J. of Chem. Technol Biotechnol* 83:739-749.
- [28] Horta, A.C.L.; Zangirolami, T.C.; Nicoletti, M. do C.; Montera, L.; Carmo, T.S.; Gonçalves, V.M. (2008) An Empirical Investigation of the Use of a Neural Network Committee for Identifying the *Streptococcus pneumoniae* Growth Phases. N.T. Nguyen et al. (Eds.): IEA/AIE 2008, LNAI 5027:215-224.
- [29] Singh R, Singh S, Sharma PK, Singh UP, Briles DE, Hollingshead SK, Lillard JW Jr. (2010) Helper T cell epitope-mapping reveals MHC-peptide binding affinities that correlate with T helper cell responses to pneumococcal surface protein A. *PLoSOne*, 5, e9432.
- [30] Ferreira DM, Darrieux M, Silva DA, Leite LC, Ferreira JM Jr, Ho PL, Miyaji EN, Oliveira ML. (2009) Characterization of protective mucosal and systemic immune responses elicited by pneumococcal surface protein PspA and PspC nasal vaccines against a respiratory pneumococcal challenge in mice. *Clin Vaccine Immunol* 16:636-645.
- [31] Bricks, L.F.; Berezin, E. (2006) Impact of pneumococcal conjugate vaccine on the prevention of invasive pneumococcal diseases. *J Pediatr (Rio J)* 82(3):S67-74 (in Portuguese).

- [32] Csordas F.C.L.; Perciani C T., Darrieux M, Gonçalves V M., Cabrera-Crespo J, Takagi M, Sbrogio-Almeida M E, Leite L C C, Tanizaki M M. (2008) Protection induced by pneumococcal surface protein A (PspA) is enhanced by conjugation to a *Streptococcus pneumoniae* capsular polysaccharide. *Vac* 26:2925-2929.
- [33] Silva, M. ; Cabrera-Crespo, J.; Sbrogio-Almeida, M.E.; Miyaji, E.N., Ho, P.L., Leite, L.CC., Lopes, A. P. Y. (2007) Optimizing expression of *Streptococcus pneumoniae* surface protein a, PspA: serocross-reactivity within families of antisera induced against clades 1 and 3. *Mol Biotechnol* 37:146-154.
- [34] Horta, ACL; Carvalho, RJ, Silva, AJ; Sargo, CR; Giordano, RC; Gonçalves, VM ; Zangirolami, TC (2010) Glicerol como fonte de carbon para cultivos de alta densidade de *E. coli*. Anais do XVIII Congr. Bras de Eng. Quim, Foz do Iguaçu-PR; (in Portuguese).
- [35] Studier, F.W. (2005) Protein production by auto-induction in high-density shaking culture. *Protein Expr Purif* 41:207-234.
- [36] Silva, AJ; Horta, ACL; Novo, MTM; Zangirolami, TC; Giordano, RC (2010) Estudo de diferentes estratégias de indução para produção do antígeno rSpaA de *E. rhusiopathiae* em cultivos de *E. coli* recombinante. Anais do XVIII Congr. Bras. de Eng. Quim. Foz do Iguaçu-PR; (in Portuguese).
- [37] Horta, ACL; Zangirolami, TC; Giordano, RC; Cruz, AJG; Reis, GB; Jesus, CDF (2010) Supervisory system for bioreactor high cell density cultivations, Application for software registration n 0000271004222441 INPI, Brazil.
- [38] Silva, A.J.; Baptista-Neto, A.; Cilento, M.C.; Giordano, R. C.; Zangirolami, T. C., (2008). Bioreactor aeration conditions modulate growth and antigen expression during *Erysipelothrix rhusiopathiae* cultivation. *Appl Microbiol Biotechnol* 79:23-31.
- [39] Nelles, O. (2001), Nonlinear system identification: from classical approaches to neural networks and fuzzy models. Springer, Germany.
- [40] Yamane, T.; Hibino, W.; Ishihara, K.; Kadotani, Y.; Kominami, M. (1991) Fed-Batch Culture Automated by uses of Continuously Measured Cell Concentration and Culture Volume. *Biotechnol Bioeng* 39:550-555.
- [41] Giordano, RC; Bertini Jr, JR; Nicoletti, MC; Giordano, RLC (2007) Online filtering of CO<sub>2</sub> signals from a bioreactor gas outflow using a committee of constructive neural networks. *Bioprocess Biosyst Eng* 31(2):101-109.

## 5.2. INTENSIFICATION OF HIGH CELL DENSITY CULTIVATIONS OF *E. COLI* FOR PRODUCTION OF *S. PNEUMONIAE* ANTIGENIC SURFACE PROTEIN, PSPA3, USING MODEL-BASED ADAPTIVE CONTROL

Antonio Carlos Luperni Horta<sup>1</sup>, Cíntia Regina Sargo<sup>1</sup>, Adilson José da Silva<sup>1</sup>, Marina de Carvalho Gonzaga<sup>1</sup>, Maurício Possedente dos Santos<sup>1</sup>, Viviane Maimoni Gonçalves<sup>2</sup>, Teresa Cristina Zangirolami<sup>1</sup>, Roberto de Campos Giordano<sup>1</sup>

<sup>1</sup>*Departamento de Engenharia Química, Universidade Federal de São Carlos, Brazil.*

<sup>2</sup>*Centro de Biotecnologia, Instituto Butantan, São Paulo, Brazil*

Corresponding author: Roberto de Campos Giordano

roberto@ufscar.br

P.O. Box 676, São Carlos, SP

Zip code: 13565-905

Phone: 55 16-3351-8708; FAX: 55 16-3351-8266

**Short running title:** Fed-batch system control based on on-line cellular growth estimation

**Abstract.** This work proposes an innovative methodology to control high density fed-batch cultures of *E. coli*, based on measurements of the concentration of dissolved oxygen and on estimations of the cellular specific growth rate ( $\mu$ ), of the biomass/limiting substrate yield ( $Y_{xs}$ ) and of the maintenance coefficient ( $m$ ). The underlying idea is to allow cells to grow according to their metabolic capacity, without the constraints inherent to pre-set growth rates. Cellular concentration was assessed on-line through a capacitance probe. Three configurations of the control system were compared: i) pre-set value for the three control parameters; ii) continuously updating  $\mu$ ; iii) updating  $\mu$ ,  $Y_{xs}$  and  $m$ . Implementation of an efficient filter for the noise of the capacitance probe was essential for a good performance of the control system. The third control strategy, within the framework of an adaptive model-based control, led to the best results, with biomass productivity reaching  $9.2 \text{ g}_{\text{cell}} \cdot \text{L}^{-1} \cdot \text{h}^{-1}$ .

**Keywords:** *adaptive fed-batch bioreactor control, inline biomass growth rate estimation, high cell density cultivation, recombinant E. coli.*

### 5.2.1 INTRODUCTION

Fed-batch is by far the most widespread operation mode for industrial and lab scale bioreactors, comprising from baker's yeast production to heterologous protein expression in high cell density cultures of genetically modified organisms [1]. The success of the fed-batch strategy is mainly due to three of its features: i) large amounts of limiting substrate can be added without triggering substrate inhibition; ii) a convenient specific growth rate ( $\mu$ ) can be imposed by controlling the limiting substrate supply, *i.e.*, the rate of addition of supplementary culture medium [2-4]; iii) and its operation does not have the complications that are common when coupled with microfiltration membranes.

The productivity of a bioprocess and the quality of the product are frequently related to growth conditions, and are sensitive to deviations from the optimal growth rate profile. Indeed, the biosynthesis of many products is closely related not only with medium composition, but also depends on the history of the system, *i.e.*, of the cell growth profile. In processes involving strains prone to metabolic bottlenecks, such *S. cerevisiae* and, under some conditions, *E. coli*, controlling the growth rate below a critical value is also necessary to prevent the formation of overflow metabolites, which may cause a decrease of biomass density and/or of the recombinant proteins productivity.

Feeding strategies for fed-batch processes are often based on open-loop dynamics [5], which may not reflect the evolution of the system during closed-loop operation. Thus, on-line information concerning the specific growth rate during the fed-batch bioreactor operation is extremely valuable. Furthermore, the availability of efficient on-line estimators for the specific growth rate of different microorganisms is very important for the implementation of robust automation strategies. Real time values for  $\mu$  are usually inferred from on-line measured variables such as carbon dioxide evolution rate (CER); oxygen uptake rate (OUR); heat generation; alkali volume consumed for pH control; optical methods like multi-wavelength fluorescence (MWF) or near-infrared spectroscopy (NIRS) and by-product formation rate, obtained from at-line HPLC measurements. These data are often input for a softsensor which would estimate biomass concentrations [6-9].

Devices that can measure properties directly correlated to cell concentration are being increasingly used in this area. The capacitance sensor [5, 10] falls into this category. With this sensor, the concentration of viable cells may be inferred on-line from measurements of the medium capacitance [11-17].

When submitted to low radio frequencies (0.1-10 MHz), the viable cell membrane behaves like an electric capacitor. The magnitude of the membrane polarization is directly proportional to the volume that it encapsulates. Hence, the dielectric permittivity is directly proportional to the viable biomass concentration, since dead cells do not present polarization in their membranes [11].

Since dielectric permittivity measurements are sensitive to cell viability, they can be used for recognition of metabolic changes during the cultivation [11, 16 and 18]. It has also been reported that some morphological changes, such as sporulation, can be detected measuring the medium permittivity, since these changes will alter the capacitance of the cells [19, 20].

Once having on-line capacitance measurements, the probe signal can be correlated to biomass concentration and the evolution of biomass concentration with the time, inferred on-line, allows estimation of the specific growth rate ( $\mu$ ). The use of this technique in industrial and lab-scale bioreactors, nevertheless, demands the implementation of efficient signal-treatment and post-processing algorithms, since the calculation of  $\mu$  requires mathematical manipulation, including the differentiation of noisy data series [21]. Thus, the inference algorithm must provide accurate estimates, smooth enough to be used by the control system.

With this purpose, a suitable filter should be designed. Moving average filters are the most commonly used algorithm, due to their simplicity. Basically, the method is to average a certain number of measured points. Throughout the process, the window of points is updated and the noisy data are smoothed [5, 22].

For *E. coli* high density cultures, it is crucial to control the feed flow rate of supplementary medium, which is frequently calculated by Eq. 5.2.1, 5.2.2 and 5.2.3. Many works use previously fixed, constant values for the parameters  $\mu_{SET}$ ,  $Y_{XS}$  and  $m$  [3, 4, 9 and 23]. More recently, some studies have addressed the use of on-line estimated specific growth rate for automatic control of the feed flow rate [5, 6].

$$F = \left( \frac{\mu}{Y_{XS}} + m \right) \cdot \frac{C_{X0} \cdot V_0}{C_{S0} - C_{SR}} \cdot e^{(\mu_{SET} \cdot t)} \quad \text{eq. 5.2.1}$$

$$\mu = \frac{dC_X}{dt} \cdot \frac{1}{C_X} + \frac{F}{V} \quad \text{eq. 5.2.2}$$

$$\mu_{SET} = \frac{dC_X}{dt} \cdot \frac{1}{C_X} \quad \text{eq. 5.2.3}$$

In eq 5.2.1,  $F$  is the feed flow rate (L/h) of fresh medium,  $\mu_{SET}$  the preset specific growth rate ( $\text{h}^{-1}$ ),  $Y_{XS}$  the biomass/substrate yield ( $\text{g}_{DCW}/\text{g}_S$ ),  $m$  the maintenance coefficient ( $\text{g}_{DCW}/\text{g}_S \cdot \text{h}$ ),  $C_{X0}$  the cellular concentration in the start of the feeding ( $\text{g}_{DCW}/\text{L}$ ),  $V_0$  the volume in the start of the feeding (L),  $t$  the feeding time (h),  $C_{S0}$  the substrate concentration in the feeding solution (g/L),  $C_{SR}$  the residual substrate concentration (g/L). In equations 5.2.2 and 5.2.3,  $\mu$  represents the specific growth rate ( $\text{h}^{-1}$ ),  $\mu_{SET}$  the apparent specific growth rate ( $\text{h}^{-1}$ ),  $t$  the time (h),  $C_X$  the cellular concentration (g/L) and  $V$  the volume (L) [21].

The underlying assumption when using equations 5.2.1-3 with constant parameters is that the concentration of the limiting substrate alone defines the specific growth rate. This is obviously a severe oversimplification of the dynamic interactions between cell metabolism/physiology and the environment. Dabros et al. [5] adopted a strategy of upgrading  $\mu_{SET}$  values, including a PI control action in equation 5.2.1, but the parameters  $m$  and  $Y_{XS}$  continued to be regarded as constants. Nevertheless, these parameters are defined within a very simplified framework for kinetic modeling of the biological system and, consequently, are known to present important changes during the evolution of the cultivation, and especially during the induction phase. Babaeipour et al. [9] showed that  $Y_{XS}$  decreases dramatically during the process, gradually falling from 0.5, before induction, to 0.1 after induction.

The parameters  $Y_{XS}$  and  $m$  can be estimated by adjusting the simplified growth model coupled to mass balances for the ideal isothermal reactor (eq 5.2.4 to 5.2.6) to experimental data. The parameters are obtained minimizing an objective function with a search algorithm. A widely used algorithm is the Levenberg–Marquardt's, and the most popular objective function is the squared norm of the deviation of the model output

variables from the experiments, taking into account an estimate of the experimental variance matrix [24] – the weighted least squares method.

$$\mu = \frac{\mu_{\text{MAX}} \cdot C_S}{K_S + C_S} \quad \text{eq. 5.2.4}$$

$$\frac{dC_X}{dt} = \left( \mu - \frac{F}{V} \right) \cdot C_X \quad \text{eq. 5.2.5}$$

$$\frac{dC_S}{dt} = \left( \frac{F}{V} \right) \cdot (C_{SF} - C_S) - \frac{1}{Y_{XS}} \mu \cdot C_X - m \cdot C_X \quad \text{eq. 5.2.6}$$

In these equations,  $F$  is the feeding rate (L/h) of fresh medium,  $\mu_{\text{MAX}}$  the maximum specific growth rate ( $\text{h}^{-1}$ ),  $Y_{XS}$  the biomass/substrate yield ( $\text{g}_{\text{DWX}}/\text{g}_S$ ),  $m$  the maintenance coefficient ( $\text{g}_{\text{DWX}}/\text{g}_S \cdot \text{h}$ ),  $C_X$  the cellular concentration ( $\text{g}_{\text{DW}}/\text{L}$ ),  $V$  the volume (L),  $t$  the time (h),  $C_{SF}$  the substrate concentration in the feeding solution (g/L),  $C_S$  the substrate concentration (g/L),  $\mu$  represents the specific growth rate ( $\text{h}^{-1}$ ), and  $K_S$  the affinity constant for substrate (g/L).

This work presents a control strategy that uses the simplified growth model (Eq. 5.2.4) but takes into account changes in some important parameters/variables ( $\mu$ ,  $Y_{XS}$  and  $m$ ) during the cultivation. The feeding flow rate is the manipulated variable, aiming to supply the real metabolic demands of the cells, in order to improve the productivity of fed-batch cultures of recombinant *E. coli*. The proposed strategy was evaluated during high cell density cultivations (HCDC) of *rE. coli* expressing Pneumococcal surface protein A (PspA), recognized as one of the most important pneumococcal virulence factors and a possible vaccine candidate, and further compared to the classical HCDC feeding approaches.

## 5.2.2 MATERIALS AND METHODS

### 5.2.2.1 EXPERIMENTAL PROCEDURE OUTLINE

#### Microorganism and cultivation protocols

Recombinant *E. coli* BL21(DE3) cells expressing a fragment of the surface protein PspA from *Streptococcus pneumoniae* under lacUV5-T7 control [25] were used in this work. The defined medium used in all bioreactor cultivations is presented in Table 5.2.1.

Table 5.2.1. Medium composition

Component	Batch phase	Feeding solution
Glycerol (g/L)	40	800
Kanamycin (mg/L)	50	50
MgSO <sub>4</sub> ·7H <sub>2</sub> O (g/L)	1.6	20
Antifoam PPG 30 % (mL/L)	1	1
(NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub> (g/L)	5.3	6.4
KH <sub>2</sub> PO <sub>4</sub> (g/L)	17.7	21.8
Citric acid (g/L)	2,27	2,27
Fe(III)citrate (mg/L)	133.3	40
CoCl <sub>2</sub> ·6H <sub>2</sub> O (mg/L)	3.33	4
MnCl <sub>2</sub> ·4H <sub>2</sub> O (mg/L)	20	23.5
CuCl <sub>2</sub> ·2H <sub>2</sub> O (mg/L)	2	2.5
H <sub>3</sub> BO <sub>3</sub> (mg/L)	4	5
Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O (mg/L)	2.8	4
Zn(CH <sub>3</sub> COOH) <sub>2</sub> ·H <sub>2</sub> O (mg/L)	33.8	16
EDTA (mg/L)	18.8	13
Thiamine (mg/L)	45	45

The feed flow of fresh medium followed an exponential profile according to equation 5.2.1, whose parameters are shown in Table 5.2.2.

Table 5.2.2. Feeding strategy definition

Id from Table 3.2.2	Cultivation	$\mu_{SET}$	Y <sub>xs</sub>	m	Inductor
A3	1	0.3 h <sup>-1</sup>	0.5	0.02 h <sup>-1</sup>	IPTG
A4	2	$\mu_{DYN}$	0.5	0.02 h <sup>-1</sup>	IPTG
A16	3	$\mu_{DYN}$	0.1 < Y <sub>xs</sub> < 0.5	0.01 < m < 0.5	IPTG
A12	4	$\mu_{DYN}$	0.1 < Y <sub>xs</sub> < 0.5	0.01 < m < 0.5	Lactose

### *Bioreactor operation, instrumentation and automation*

The cultivations were conducted in a 5 L (in-house) bioreactor monitored by SuperSys\_HCDCR [26]. The pH was controlled (on/off, pHmeter GLI PRO) at 6.7 (during the batch phase) and 6.9 (during the fed-batch phase) by addition of NH<sub>4</sub>OH (30%). Temperature was set at 35°C. Dissolved oxygen concentration (DOC) was monitored by a Mettler Toledo probe Inpro 6800, connected to CE O2 4050 transmitter, and kept at 30% of saturation by a PID controller which automatically changed the stirrer speed between 200 and 1000 rpm. The gas stream supplied to the bioreactor was a mixture of air and pure oxygen and its total flow rate was maintained at 6 L/min by two mass flow controllers (GFC AALBORG). The broth permittivity and conductivity were monitored by a biomass sensor from FOGALE<sup>®</sup> Nanotech. The exhaust gas composition



was assessed by a Sick/Maihak S.710 CO<sub>2</sub> and O<sub>2</sub> analyzer. The feeding supply was provided by an Ismatec – BVP pump, following the profile defined by equation 5.2.1. On-line data acquisition as well as monitoring/control of all instruments via a compact field point 2020 (cFP-2020, National Instruments) was performed by the software SuperSys\_HCDCR<sup>®</sup>, developed in LabView<sup>®</sup> [26].

#### *Off-line and at-line measurements*

Cell mass quantification was done by three methods: i) correlation with permittivity [online: perm(pF/cm)] measured by the capacitance probe, providing viable cells densities; ii) correlation with optical density [at-line: OD ( $\lambda$  =600nm)]; iii) dry cell weight (off-line: C<sub>x</sub>(g/L)] [27]. Metabolites formation and carbon sources consumption were measured by HPLC as described in Silva et al. [28]. The carbon source was also measured (at-line) by an enzymatic kit (Laborlab).

#### 5.2.2.2 DATA TREATMENT AND CALCULATIONS

##### *On-line biomass inference*

During all cultivations, online biomass monitoring by a capacitance sensor (BIOMASS System - Viable cell density monitoring - FOGALE Nanotech) was carried on. A reference frequency of 10 MHz was used as blank signal, to compensate disturbances caused by changes in medium composition and on the overall bioreactor environment (agitation, aeration, presence of debris etc). The frequency of measurement was set at 2.984 MHz. The response signal at the reference frequency was automatically subtracted from the signal at the frequency of measurement.

The probe response was transferred to the computer database via its analog output, with a sampling rate of 1.0 s<sup>-1</sup>.

### *Protein production*

Protein synthesis was checked by SDS-PAGE, and the gels were photographed in order to estimate protein production by the intensity of protein bands using the software ImageJ [29].

### *Permittivity data treatment*

Data collected from the capacitance probe presented high noise. A filter similar to the classical moving average (CMA) was implemented, aiming at reducing the noise in order to make permittivity response data suitable for growth rate estimation. The CMA filter operates by averaging a number of data points from the input signal to produce each data point in the output signal [30]. The CMA assumes the value of the average of the rough vector data. The array size ( $n$ ) can vary depending on the degree of smoothing required. A variation of this technique is to use the average of the smoothed data array instead of the rough data vector, what is known as smoothed moving average filter (SMA).

Both filters (CMA and SMA) were implemented in accordance to the pseudocode shown in Fig 5.2.1. Permittivity data was acquired as a 0.004-0.02 A analogical signal, but in the acquisition software it was multiplied by a calibration factor proportional to the instrument output range to convert the readings in A to pF/cm. This analogical signal was used as input to the SMA filter.

```

Procedure: Classical moving average (CMA)
begin
  Actual_perm=permittivity
  t=actual_time

  if t=0 Perm_Vector_R(1:n)=Actual_perm

  Perm_Vector_R(n)=Actual_perm
  Perm_Vector_Smooth(n)=average[Perm_Vector_R(1:n)]
  Perm_Vector_R(1:n-1)= Perm_Vector_R(2:n)
  Perm_Vector_Smooth(1:n-1)=Perm_Vector_Smooth(2:n)
End

```

Fig 5.2.1 Pseudocode of the moving average filters implemented in the control and monitoring software of the bioreactor. Code implemented in MatLab 6.5

The selection of the filters was evaluated by using the lack-of-fit (S) and the roughness (R) metrics, given by equations 5.2.7 and 5.2.8, respectively.

$$S = \sum_n [a(n) - z(n)]^2 \quad \text{eq. 5.2.7}$$

$$R = \sum_n [z(n) - z(n-1)]^2 \quad \text{eq. 5.2.8}$$

In equations 5.2.7 and 5.2.8,  $a(n)$  is the  $n$ -th crude data point and  $z(n)$  is the  $n$ -th smoothed data point.

The filtering process attempted to satisfy two criteria: smoother data imply higher S and lower R. If S is too high, however, relevant information regarding the trends of the data may be discarded by the filter. On the other hand, if R is too high, the noise will pass through and the filter will be ineffective [31]. Since the flow rate  $F$  is a function of the permittivity signal [ $C_x$ ,  $\mu$  and  $\mu_{SET}$  (Eq. 5.2.1)], the values of R and S need to be correctly chosen to provide a smooth trajectory for the control variable (supplementary medium/inducer flow rates), without loss of relevant information.

### *Specific growth rate estimation*

The smoothed permittivity data obtained as described in Fig 5.2.1 were used to estimate the growth rate according to the pseudocode presented in Fig 5.2.2.

```

Procedure: Growth rate
begin
  Actual_Cx = Cx(permittivity)
  T(n) = actual_time
  F = actual_feed
  V = actual_volume

  if (t=0) then
    Cx_Vector(1:n) = Actual_Cx

  Cx_Vector(n) = Actual_Cx
  ΔCx = Cx_Vector(n) - Cx_Vector(1)
  Δt = T(n) - T(1)
  μ = (ΔCx/Δt)/[ΣCx_Vector(1:n)/n]+F/V
  μSET = (ΔCx/Δt)/ΣCx_Vector(1:n)/n
  Cx_Vector(1:n-1)= Cx_Vector(2:n)
  T(1:n-1)= T(2:n)
End

```

Fig 5.2.2 Pseudocode for specific growth rate estimation implemented in the control and monitoring software of the bioreactor. Code implemented in MatLab 6.5

The data calculated by the pseudocode described in Fig 5.2.2 was further treated by a SMA filter to improve its smoothness.

### 5.2.2.3 AUTOMATIC CONTROL OF FEED FLOW RATE

The feeding profile following the actual growth rate was implemented using the biomass and the specific growth rate estimated by the filtered permittivity signal. The specific growth rate values were also smoothed by the same SMA filter (with the same vector size).

The control parameters  $\mu$ ,  $\mu_{SET}$ ,  $C_{X0}$  and  $V_0$  from equation 5.2.1 (feeding profile) were automatically updated each 10 minutes.

To identify changes in the parameters  $Y_{XS}$  and  $m$ , a tool for at-line re-tuning of the control parameters was incorporated to the bioreactor control program. Online information from the bioreactor and surplus data of substrate concentrations, measured at-line, were used for parameter update throughout the cultivation. The parameters  $Y_{XS}$  and  $m$  were updated at intervals of one hour, by fitting the model, eq. 5.2.4 to 5.2.6, to the online and at-line data; the parameter re-calibration procedure was done using the Matlab<sup>®</sup> Levenberg–Marquardt algorithm (MathWoks<sup>®</sup>) [24]. Fig 5.2.3 presents the configuration of the control system.

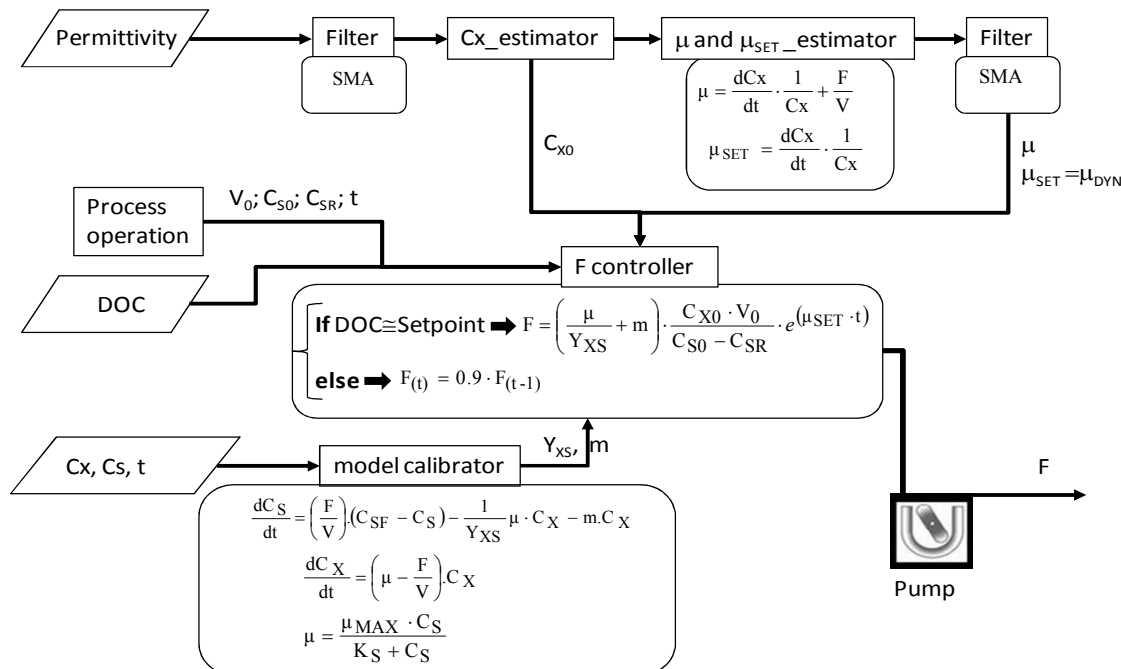


Fig 5.2.3 Supplementary medium feed control:  $C_{X0}$ ,  $\mu$ ,  $\mu_{SET}$ ,  $V_0$ ,  $C_{S0}$ ,  $C_{SR}$  re-tuned online;  $Y_{XS}$ , and  $m$  corrected at-line

With this control strategy, all parameters of the feeding equation can be adjusted according to metabolic changes during the cultivation, comprising also the bioreactor capacity of supplying the demand for oxygen.

## 5.2.3 RESULTS AND DISCUSSION

### 5.2.3.1 CORRELATION BETWEEN PERMITTIVITY AND CELL CONCENTRATION – OBTAINING C<sub>X</sub>

A typical correlation between permittivity data and cellular concentration from three independent cultivations can be seen in Fig 5.2.4.

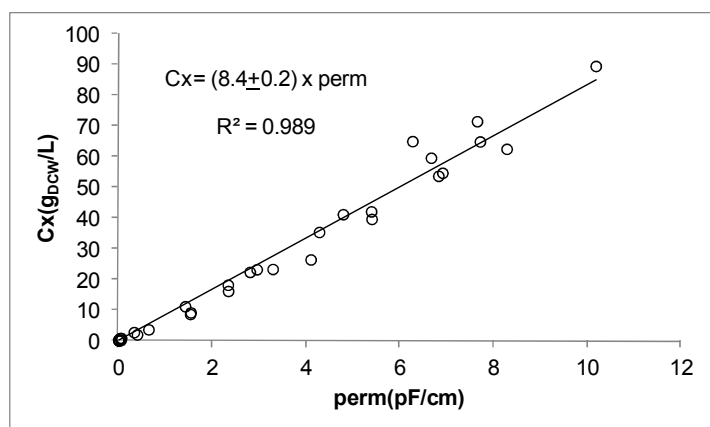


Fig 5.2.4 Correlation between permittivity [Perm (pF/cm)] and cellular concentration [C<sub>x</sub> (g/L)]. Symbols are experimental data, while the line is the linear fit

Fig 5.2.4 illustrates the linear correlation of permittivity (perm – pF/cm) and cellular concentration (C<sub>x</sub> - dry cell weight), in a wide range of cell densities (0-100 g/L).

### *SMA filter configuration and performance*

One *E. coli* cultivation data set was used to compare the two tested filters, CMA and SMA. For both, the smoothing window size (n) was varied from 20 to 1500, from 10 to 10 data, and compared. The R and S as function of the n length are shown in Fig 5.2.5.

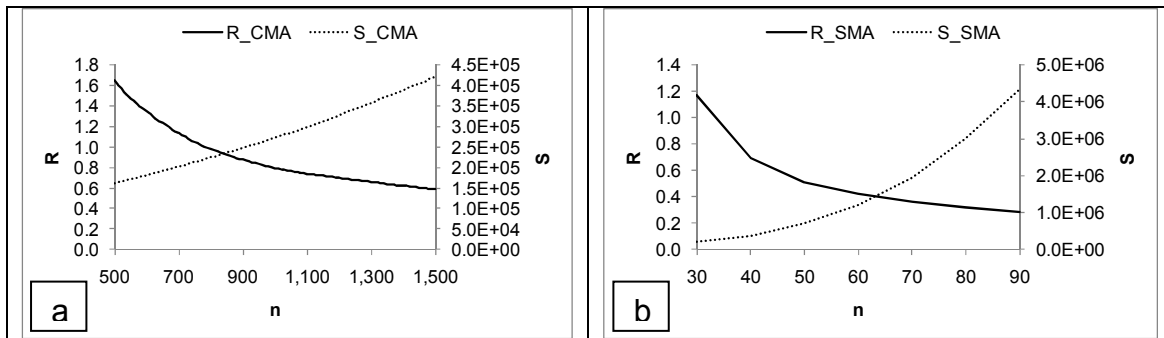


Fig 5.2.5 Data obtained using filters with lack of fit ( $S - pF^2/cm^2$ ) and roughness ( $R - pF^2/cm^2$ ), a) Classical moving average (CMA) and b) Smoothed moving average (SMA)

Fig 5.2.5 shows that SMA performance is clearly superior. For instance, to achieve similar values of the performance indexes, CMA demanded  $n=1000$  sampling points (giving  $R=0.79$  and  $S=27 \cdot 10^4 \text{ pF}^2/cm^2$ ), while SMA used only  $n=30$  ( $R=1.2$  and  $S=20 \cdot 10^4 \text{ pF}^2/cm^2$ ). The higher the value of "n", the greater the delay caused by the filter. The performance of each filter is shown in Fig 5.2.6, for  $n=30$ .

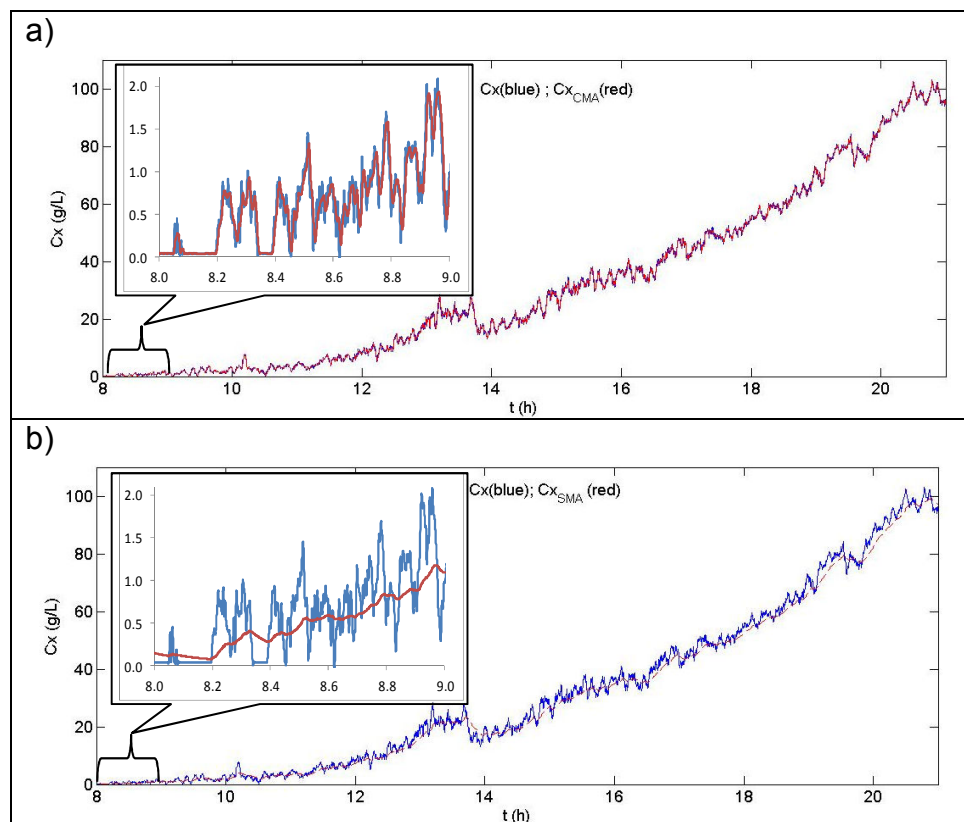


Fig 5.2.6 Smoothed experimental data; *E. coli* cultivation. The blue line is the rough signal and the red one is the smoothed signal. a) CMA method,  $n=30$ ,  $R=47.7$  and  $S=11882$ ; b) SMA method,  $n=30$ ,  $R=1.2$  and  $S=199332$

The acquisition and processing time is 1s, so the delay of the SMA filter is 30 seconds, compared to 16 minutes for CMA with  $n=1000$ .

The SMA method with  $n = 30$  was chosen as the standard for the automation system employed in this work, because with a smaller smoothing vector the control system can respond more accurately to metabolic changes during the cultivation.

### Comparing feed strategies

Cultivation 1 was done using the conventional approach of maintaining a predetermined  $\mu_{SET}$ , in this case of  $0.3\text{h}^{-1}$ . This value was based in previous experiments. Results of Cultivation 1 are depicted in Fig 5.2.7.

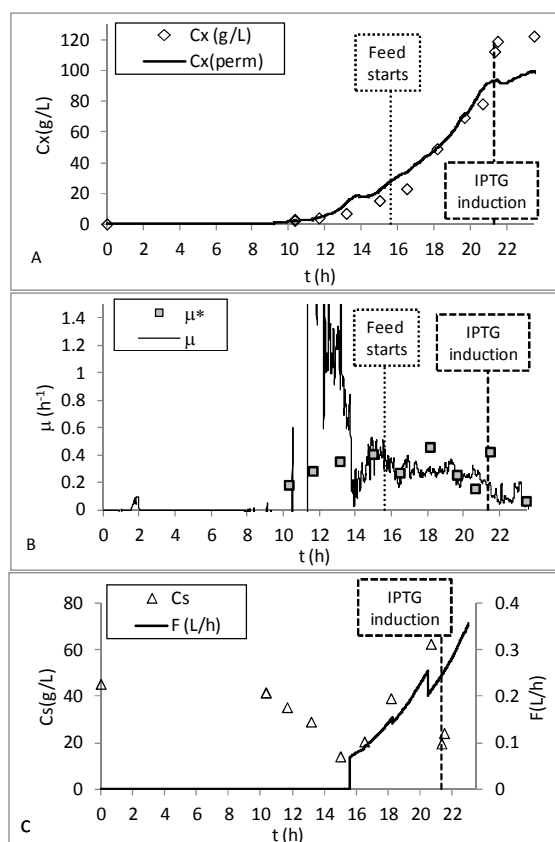


Fig 5.2.7 Cultivation 1: Fast IPTG induction;  $\mu_{SET}=0.3\text{ h}^{-1}$ ;  $Y_{XS}=0.5$ ;  $m=0.02$ ; A) Cellular concentration measured by DCW ( $C_x$  - g/L) and by the smoothed permittivity signal [ $C_x(\text{perm})$  - g/L]; B) Specific growth rate calculated from DCW ( $\mu^*$  -  $\text{h}^{-1}$ ), and from the smoothed permittivity signal ( $\mu$  -  $\text{h}^{-1}$ ); C) Substrate concentration ( $C_s$  - g/L) and feed flow rate ( $F$  - L/h)

Fig 5.2.7\_A shows a good correlation between the cellular concentration estimated from the capacitance probe and the one measured as DCW until the moment of induction. After IPTG addition, permittivity data detected a drop in cellular viability, which was not observed in the dry cell weight data: DCW is not sensitive to cell viability. In this experiment, a biomass production of 120 g/L (dry weight) was reached after 21.5 h of cultivation. We may notice in Fig 5.2.7\_B that after the start of the feeding phase, in general, the  $\mu$  calculated online represented well the one calculated offline; but even filtering the permittivity, the specific growth rate is noisy and the crude  $\mu$  is not a good sign for the control of the feeding pump. This result shows the necessity of applying the SMA filter on  $\mu$  before using it as an input of the feeding control.

Fig 5.2.7\_C shows that maintaining the feeding in a constant rate is not a good strategy because changes in metabolism can lead to substrate accumulation, as noted, or to the lack of substrate, limiting growth and decreasing productivity. In this case, the maximum  $\mu$  was  $0.45 \text{ h}^{-1}$ , that is higher than the value used in the feeding profile, showing that  $\mu_{\text{SET}}$  of  $0.3 \text{ h}^{-1}$  was not oversized. Discontinuities observed in the flow rate are caused by withdrawals of large volumes of culture medium, which altered the value of  $V_0$ , updated after every sample withdrawal (see Eq 5.2.1).

For the second cultivation the same SMA filter was used for calculating  $\mu$ , to obtain the  $\mu_{\text{DYN}}$  used in the feeding control; results are presented in Fig 5.2.8.

Fig 5.2.8\_A shows that the biomass estimation using the permittivity followed closely the DCW measurements. A cell concentration around 100 g/L was reached after 25.5 h of cultivation.

Fig 5.2.8\_B replicated a phenomenon observed in Fig 5.2.7\_B: a maximum of  $\mu$  during the batch phase, as a consequence of the availability of substrate to sustain a high metabolic activity of the cells. The instantaneous drop of the feed rate at 20 h (Fig 5.2.8\_C) was once again due to the withdrawal of culture medium, thus reducing  $V_0$ .

It is important to notice that, using  $\mu_{\text{DYN}}$  instead of  $\mu_{\text{SET}}$ , the accumulation of substrate was avoided, but glycerol was almost depleted much before induction (see Fig 5.2.8\_C, after app. 17h).



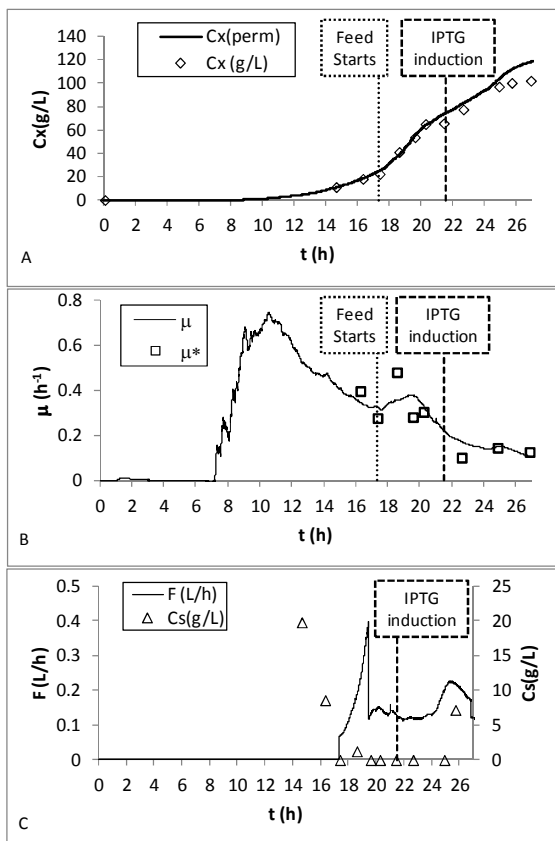


Fig 5.2.8 Cultivation 2: Fast IPTG induction;  $\mu_{\text{SET}} = \mu_{\text{DYN}}$ ;  $Y_{\text{XS}} = 0.5$ ;  $m = 0.02$ ; A) Cellular concentration measured by DCW ( $C_x$  - g/L) and by the smoothed permittivity signal [ $C_x(\text{perm})$  - g/L]; B) Specific growth rate calculated from DCW ( $\mu^*$  -  $\text{h}^{-1}$ ), and from the smoothed permittivity signal ( $\mu$  -  $\text{h}^{-1}$ ); C) Substrate concentration ( $C_s$  - g/L) and feed flow rate ( $F$  - L/h)

The experiment shown in Fig 5.2.9 used  $\mu_{\text{DYN}}$ , coupled to the (at-line) re-tuning of  $Y_{\text{XS}}$  and of  $m$ , control parameters that were invariant in the previous experiments.

Fig 5.2.9\_A shows that, once again, the permittivity signal followed closely the DCW profile, and that there is a peak of  $\mu$  in the middle of the batch phase – despite the presence of a higher level of noise in this estimate. In this experiment, the supplementary feed started before the depletion of glycerol (at  $C_s \approx 5$  g/L), thus avoiding a sudden drop of the specific growth rate at the end of the batch phase.

It should be noticed that the feed rate was often cut down during this cultivation (Fig 5.2.9\_C), due to the override of the control action by the DOC supervisor that was implemented (see the scheme in Figure 5.2.3). In this way, the dissolved oxygen was controlled close to its set point after the end of the lag phase, throughout the cultivation (Fig 5.2.9\_D).

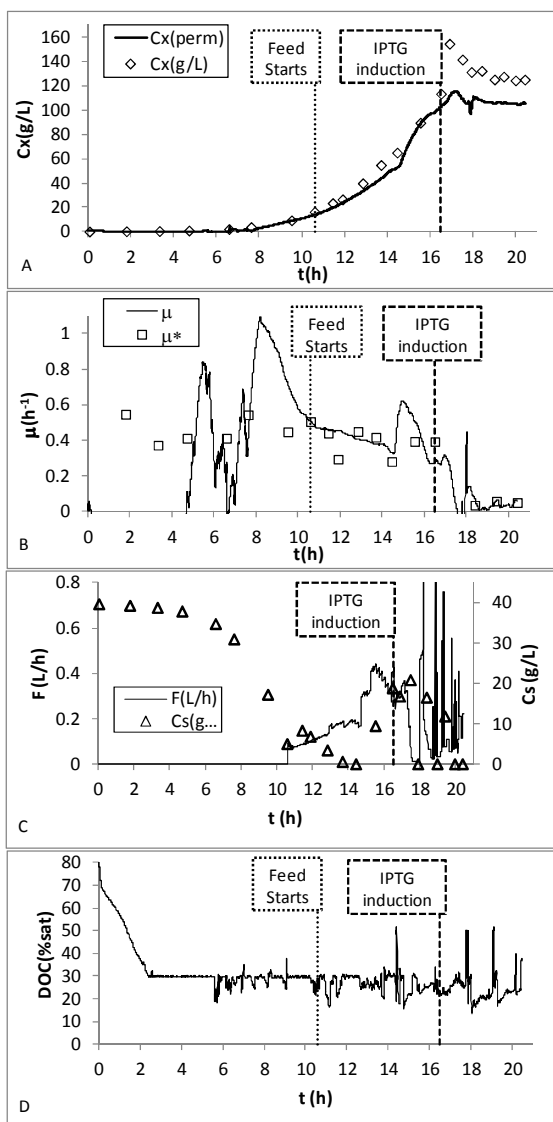


Fig 5.2.9 Cultivation 3: Fast IPTG induction;  $\mu_{\text{SET}} = \mu_{\text{DYN}}$ ; re-tuning yield and maintenance coefficients:  $0.1 < Y_{\text{XS}} < 0.5$ ;  $0.01 < m < 0.5$ ; A) Cellular concentration measured by DCW (g/L) and by the smoothed permittivity signal; B) Specific growth rate calculated from DCW ( $\mu^*$ ), and from the smoothed permittivity signal; C) Substrate concentration (g/L) and feed flow rate (L/h); D) Dissolved oxygen concentration (DOC - % of saturation)

The previous experiments used the inducer IPTG, which has a fast action – the induction phase endured 4-6 h. Fig 5.2.10 shows a cultivation where lactose was used for induction of the expression of the heterologous protein. Typically, with lactose the induction phase is longer, 6-12h.

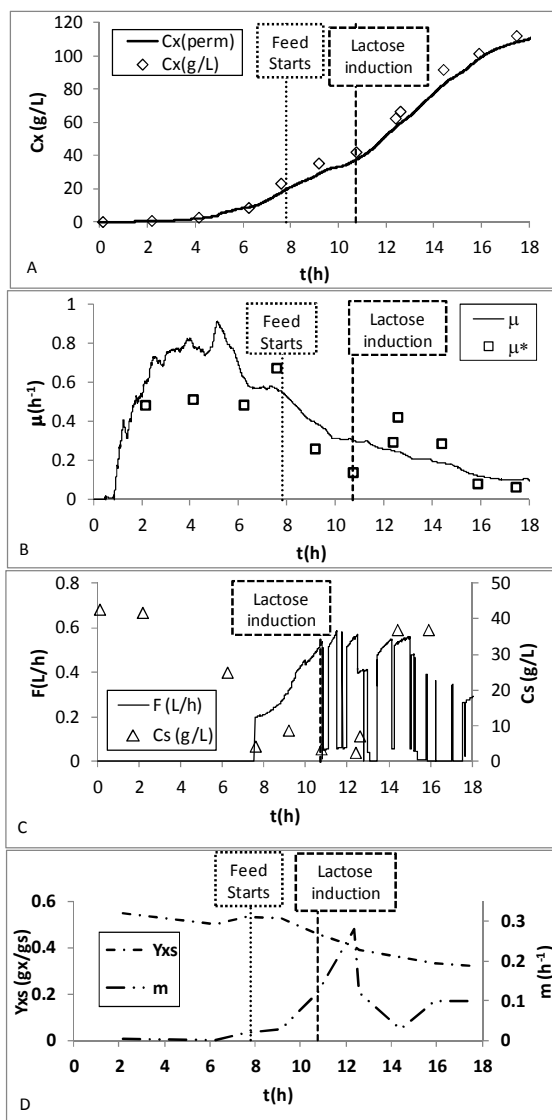


Fig 5.2.10 Cultivation 4: Slow lactose induction;  $\mu_{\text{SET}} = \mu_{\text{DYN}}$ ;  $0.1 < Y_{\text{XS}} < 0.5$ ;  $0.01 < m < 0.5$ ; A) Cellular concentration measured by DCW (g/L) and measured by the smoothed permittivity signal; B) Specific growth rate calculated from DCW ( $\mu^*$ ), and calculated from the smoothed permittivity signal; C) Substrate concentration (g/L) and feed flow rate (L/h); D) Yield coefficient ( $Y_{\text{XS}} - \text{gX/gS}$ ), and maintenance ( $m - \text{h}^{-1}$ )

Fig 5.2.10\_A and Fig 5.2.10\_B show that permittivity data were again consistent with DCW measurements. The shorter lag phase in Cultivation 4 is noteworthy, although all experimental protocols were the same as before – this is an indication that run to run variability may always be present, even when using defined medium and the same frozen stock culture as inoculum. In this case, the use of adaptive control strategies and of feedback information becomes even more important, in order to properly handle unexpected growth profiles – such the one of Cultivation 4.

The peak of  $\mu$  at the middle of the batch phase, when glycerol concentration reaches app. 30 g/L, is present, as in the previous experiments. In this cultivation cell concentration at the beginning of the induction phase was smaller than in the previous ones. Even though, a high cell density was achieved, since the bacterial growth was sustained even after induction started. Since lactose does not have the toxicity of IPTG, no sudden drop of cell viability was detected by the capacitance sensor after induction, as it has been observed in the previous assays.

Fig 5.2.10\_C shows that the feed flow was occasionally reduced, in order to control the DOC. In all experiments, this intervention of the supervisory system, overriding the control action, was essential to avoid accumulation of organic acids, produced by anaerobic fermentative pathways. Indeed, acetate concentrations were never higher than 3 g/L, a negligible value when compared to the high cell density of the cultivations (above 100 g/L). For the sake of comparison, Kim et al. [32] reported reaching 31 g/L of acetate (achieving 101 g<sub>cell</sub>/L), and observed inhibitory effects on cell growth. The authors' carbon source was glucose, a substrate whose catabolism is prone to deviate towards acetic fermentation in *E. coli* (including aerobic fermentation via the glucose overflow mechanism). In the present work, using glycerol and a modified formulation of HDF feed medium [4] supplemented with phosphate and other salts (Sargo et al, submitted), this problem could be avoided.

Fig 5.2.10\_D shows an interesting profile for the evolution of  $Y_{XS}$  along the cultivation. After some soft oscillations in batch and feeding phases, during the induction phase the cell yield consistently decreases, as expected: the processing of glycerol by the microorganism is strongly affected by the metabolic changes caused by the induction of the expression of the foreign protein. These results support the importance of considering the influence of these variations on the feed flow control parameters, which must be re-tuned for obtaining a proper control action.

Fig 5.2.11 shows cellular productivity profiles of the four cultivations.

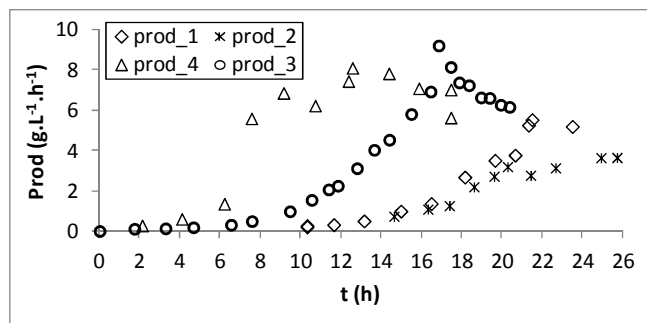


Fig 5.2.11 Cell productivity profiles, g/(L.h): 1) using  $\mu_{SET} = 0.3$ ; 2) using  $\mu_{DYN}$ ; 3) using  $\mu_{DYN}$ , and re-tuning  $Y_{XS}$  and  $m$ ; 4) same as in (3) but with induction with lactose

The higher cell productivities were reached in Cultivations 3 and 4, using  $\mu_{DYN}$  and re-tuning  $Y_{XS}$  and  $m$ . The exceptionally shorter lag phase in Cultivation 4 caused a sharp increase in cell productivity during the first hours. The productivity of Cultivation 1, using constant  $\mu_{SET}$ , was higher during the final phase because the limiting substrate concentration was not depleted throughout the feeding phase.

Protein expression was checked by SDS-PAGE, as illustrated in Fig 5.2.12. Fig 5.2.13 compares results of Cultivations 3 and 4. In both cases (induction with IPTG and lactose), the high specific growth rates attained with the aid of the control system did not hinder the expression of the protein of interest at high rates, too.

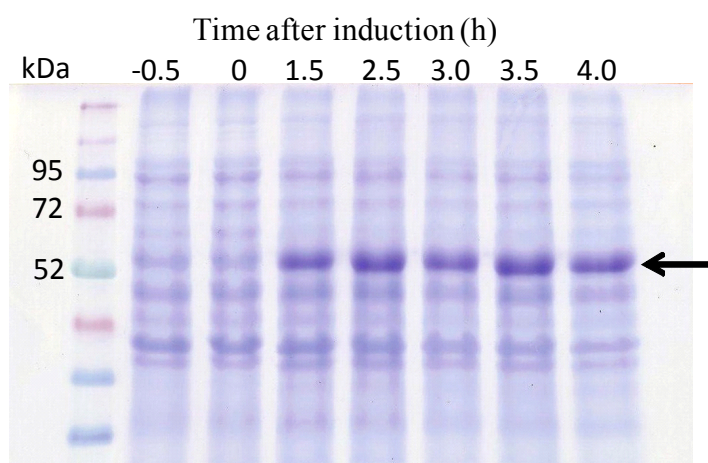


Fig 5.2.12 Image of a gel from Cultivation 3 (induction with IPTG). The first lane corresponds to the standard, the second lane to the sample immediately before induction, the third one is a sample immediately after the addition of IPTG, and the other are in chronological order after induction

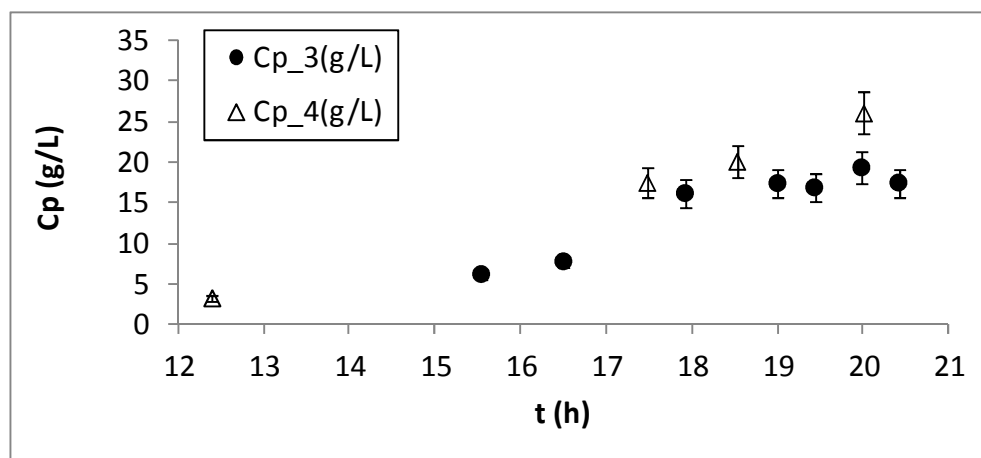


Fig 5.2.13 Profiles of PspA3 protein expression ( $C_p$  in  $\text{g/L}_{\text{reactor}}$ ) in Cultivation 3 (induced with IPTG) and 4 (induced with lactose). Error bars are standard deviations of triplicate measurements

Fig 5.2.13 shows that there was a higher protein expression with lactose as inducer (Cultivation 4), despite the higher cellular productivity at the first stages of Cultivation 3, i.e., IPTG causes faster synthesis, but lactose leads to higher yield.

#### 5.2.4 CONCLUSION

The control strategy presented in this study proved very effective, sustaining intense cell growth under glycerol without limitation or accumulation of substrate and without sensible detour towards acetic fermentation. Avoiding reduction of substrate concentration to low levels proved to be a key factor for the productivity of the process when glycerol is used as main carbon source. The concentration of dissolved oxygen could be effectively controlled within an adequate range, in order to provide the demands of the growing cellular population, thus avoiding fermentative routes. Whenever direct manipulation of stirring and air+O<sub>2</sub> flows was not enough to avoid oxygen limitation, the control system reduced the feeding of supplementary medium, but the exponential feed was automatically restored when DOC was back to its set point. With this strategy, good volumetric productivities in cell mass and in product were achieved.

The online re-tuning of the specific growth rate, based on the signal of a capacitance probe, was essential for a good performance of the bioreactor. For this

purpose, an adequate filter of the signal noise was demanded. The control strategy was model-based, but using a very simplistic picture of this complex system. Consequently, important further improvements were achieved when the cell/limiting substrate yield ( $Y_{X/S}$ ) and the maintenance coefficient ( $m$ ) were re-tuned using at-line data (basically, concentrations of glycerol, easily measured with an enzymatic kit). After induction, both with IPTG and with lactose, the metabolic stress imposed by the activation of the heterologous protein synthesis caused a reduction of growth. But lactose provided higher protein productivities than IPTG, because the toxicity of the last one imposed an extra restriction to cellular growth.

The cellular maximum productivity was higher than  $9 \text{ gL}^{-1}\text{h}^{-1}$ , achieving 154 g/L of dry mass in less than 17h of cultivation. Concentrations of PspA3 in the bioreactor reached 26.1 g/L (after 20 h of cultivation).

**Acknowledgments.** The authors would like to thank the São Paulo Research Foundation (FAPESP) and Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) for funding this work, and Tiago Martins Pereira and Amadeus Gomes de Azevedo for technical support.

## REFERENCES

- [1] Shiloach J, Fass R (2005) Growing *E. coli* to high cell density: a historical perspective on method development. *Biotechnol Adv* 23:345–357
- [2] Knabben I, Regestein L, Grumbach C, Steinbusch S, Kunze G, Büchs J (2010) Online determination of viable biomass up to very high cell densities in *Arxula adenivorans* fermentations using an impedance signal. *Journal of Biotechnology* 149: 60–66
- [3] Lee J, Lee SY, Park S, Middelberg APJ (1999) Control of fed-batch fermentations. *Biotechnology Advances* 17:29-48
- [4] Seeger A, Schbeppe B, McCarthy JEG, Deckwer WD, Rinas U (1995) Comparison of temperature and isopropyl- $\beta$ -D-thiogalacto-pyranoside-induced synthesis of basic fibroblast growth factor in high-cell-density cultures of recombinant *Escherichia coli*. *Enz. and Microb. Tech.* 17:947-953
- [5] Dabros M, Schuler MM, Marison IW (2010) Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. *Bioprocess Biosyst Eng* 33:1109-1118
- [6] Warth B, Rajkai G, Mandenius CF (2010) Evaluation of software sensors for on-line estimation of culture conditions in an *Escherichia coli* cultivation expressing a recombinant protein. *Journal of Biotechnology* 147:37-45

- [7] Ödman P, Johansen CL, Olsson L, Gernaey KV, Lantz AE (2009) On-line estimation of biomass, glucose and ethanol in *Saccharomyces cerevisiae* cultivations using in-situ multi-wavelength fluorescence and software sensors. *Journal of Biotechnology* 144:102–112
- [8] Gnoth S, Jenzsch M, Simutis R, Lubbert A (2008) Control of cultivation processes for recombinant protein production: a review. *Bioprocess Biosyst Eng* 31:21–39
- [9] Babaeipour V, Shojaosadati SA, Robotjazi SM, Khalilzadeh R, Maghsoudi N (2007) Over-production of human interferon- $\gamma$  by HCDC of recombinant *Escherichia coli*. *Process Biochemistry*, 42:112-117
- [10] Ansorge S, Esteban G, Schmid G (2009) Multifrequency Permittivity Measurements Enable On-Line Monitoring of Changes in Intracellular Conductivity Due to Nutrient Limitations During Batch Cultivations of CHO Cells. *Biotechnol Prog*, 26(1)272-283
- [11] Markx GH, Davey CL (1999) The dielectric properties of biological cells at radiofrequencies: Applications in biotechnology. *Enzym Microb Technol* 25:161–171
- [12] Neves AA, Pereira DA, Vieira LM, Menezes JC (2000) Real time monitoring biomass concentration in *Streptomyces clavuligerus* with industrial media using a capacitance probe. *J Biotechnol* 84:45-52
- [13] Ferreira AP, Vieira LM, Cardoso JP, Menezes JC (2005) Evaluation of a new annular capacitance probe for biomass monitoring in industrial pilot-scale fermentations. *J Biotechnol* 116:403-409
- [14] Bai W, Zhao KS, Asami K (2006) Dielectric properties of *E. coli* cell as simulated by the three-shell spheroidal model. *Biophysical Chemistry* 122:136–142
- [15] Xiong ZQ, Guo MJ, Guo YX, Chu J, Zhuang YP, Zhang SL (2008) Real-Time Viable-Cell Mass Monitoring in High-Cell-Density Fed-Batch Glutathione Fermentation by *Saccharomyces cerevisiae* T65 in Industrial Complex Medium. *J Biosci bioeng* 105(4):409–413
- [16] Kiviharju K, Salomen K, Moilanen U, Eerikainen T (2008) Biomass measurement online: the performance of in situ measurements and software sensors. *J Ind Microbiol Biotechnol* 35:657-665
- [17] Reis GB, Horta ACL, Zangirolami TC, Giordano RC, Cruz A.JG (2009) Control of Fed-Batch Yeast Cultivation Using a Capacitance Sensor. 10th Int Symp Process Syst Eng – PSE 2009
- [18] Maskow T, Röllich A, Fetzer I, Yao J, Harms H (2008) Observation of non-linear biomass–capacitance correlations: Reasons and implications for bioprocess control. *Biosens Bioelectron* 24:123-128
- [19] Abi A, Sarrafzadeh MH, Mehrnia MR, Ghommidh C (2010) Application of Dielectric Permittivity Measurements in Physiological State Monitoring of *Bacillus subtilis* Culture. 2010 2<sup>nd</sup> International Conference on Chemical, Biological and Environmental Engineering (ICBEE 2010)
- [20] Sarrafzadeh M H, Belloy L, Esteban G, Navarro JM, Ghommidh C (2005) Dielectric monitoring of growth and sporulation of *Bacillus thuringiensis*. *Biotechnol Lett* 27:511–517
- [21] Nielsen J, Villadsen J, Lidén G (2002), *Bioreaction Engineering Principles*, 2nd Ed., Kluwer Academic/Plenum Publishers, New York
- [22] Eilers PHC (2003) A Perfect Smoother, *Analytical Chemistry*, 75(14):3631-3636.
- [23] Korz DJ, Rinas U, Hellmuth K, Sanders EA, Deckwer WD (1995) Simple fed batch technique for high cell density cultivation of *Escherichia coli*. *Journal of Biotechnology* 39:59-65



- [24] Nelles O (2001), Nonlinear system identification: from classical approaches to neural networks and fuzzy models. Springer, Germany.
- [25] Carvalho RJ, Cabrera-Crespo J, Tanizaki MM, Gonçalves VM (2011) Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. Applied Microbiology and Biotechnology, 1-12.
- [26] Horta ACL, Zangirolami TC, Giordano RC, Cruz AJG, Reis GB, Jesus CDF (2010) Supervisory system for bioreactor high cell density cultivations. Software registration proc. 11008-6 INPI, Brazil.
- [27] Olsson L, Nielsen J (1997) On-line and in situ monitoring of biomass in submerged cultivations. TIBTECH, 15:517-522.
- [28] Silva AJ, Baptista-Neto A, Cilento MC, Giordano RC, Zangirolami TC (2008) Bioreactor aeration conditions modulate growth and antigen expression during *Erysipelothrix rhusiopathiae* cultivation. Appl Microbiol Biotechnol 79:23–31
- [29] Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image processing with ImageJ. Biophotonics International, 11:36-42
- [30] Smith SW (1997) The scientist and engineering guide to digital signal processing. California Technical Production, San Diego
- [31] Giordano RC, Bertini Jr JR, Nicoletti MC, Giordano RLC (2007) Online filtering of CO<sub>2</sub> signals from a bioreactor gas outflow using a committee of constructive neural networks. Bioprocess Biosyst Eng 31(2):101-109
- [32] Kim BS, Lee SC, Lee SY, Chang YK, Chang HN (2004) High cell density fed-batch cultivation of *Escherichia coli* using exponential feeding combined with pH-stat. Bioprocess Biosyst Eng, 26:147–150

### 5.3. ON-LINE MONITORING OF BIOMASS CONCENTRATION BASED ON A CAPACITANCE SENSOR: ASSESSING THE METHODOLOGY FOR DIFFERENT BACTERIA AND YEAST HIGH CELL DENSITY FED-BATCH CULTURES

Antonio Carlos Luperni Horta<sup>1</sup>, Adilson José da Silva<sup>1</sup>, Cíntia Regina Sargo<sup>1</sup>, Inti Doraci Cavalcanti-Montaño<sup>1</sup>, Carlos Alberto Galeano-Suarez<sup>1</sup>, Ana Maria Velez<sup>1</sup>, Maurício Possedente dos Santos<sup>1</sup>, Viviane Maimoni Gonçalves<sup>2</sup>, Roberto de Campos Giordano<sup>1</sup>, Teresa Cristina Zangirolami<sup>1</sup>

<sup>1</sup>*Universidade Federal de São Carlos, Departamento de Engenharia Química.*

<sup>2</sup>*Centro de Biotecnologia, Instituto Butantan, São Paulo, Brazil*

P.O. Box 676, São Carlos, SP

Zip code: 13565-905

Phone: 55 16-3351-8264; FAX: 55 16-3351-8266

E-mail: horta@ufscar.br

**Short running title:** on-line estimation of biomass concentration by a capacitance sensor

**Abstract.** In this work, the performance of an *in-situ* capacitance sensor for on-line monitoring of biomass concentration was evaluated for some of the most important microorganisms in the biotechnology industry: *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris* and *Bacillus megaterium*. A total of 33 batch and fed-batch cultures were carried out in bench-scale bioreactor and biomass formation trends were followed by the dielectric measurements during growth phase as well as induction phase, for 4 recombinant *E. coli* strains. Permittivity measurements and viable cellular concentrations presented a linear correlation for all the studied conditions. In addition, permittivity signal pre-treated by a smoothed moving average (SMA) filter was further used for inference of cellular growth rate. The estimated specific growth rates mirrored the main trends of the metabolic states of the different cells and they can be further used for setting-up control strategies in fed-batch cultures.

**Keywords:** capacitance probe, permittivity, recombinant *E. coli*, *S. cerevisiae*, *P. pastoris*, *B. megaterium*, on-line specific growth rate estimation.

### 5.3.2. INTRODUCTION

Undoubtedly reliable, on-line measurements of biomass concentration are the most important information for running a fed-batch culture. Continuous measurement of cell concentration is technically feasible by means of several techniques as laser turbidimeter, suspended solid concentration analyzer, in situ microscope, near-infrared spectroscopy, laser light scattering analysis, fluorescence, dielectric spectroscopy and capacitance measurement as well as optical density scanners (Yamane et al., 1991; Macaloney et al., 1997; Xiong et al 2008; Kiviharju et al, 2008; Dabros et al., 2009).

Among the above mentioned alternatives, dielectric spectroscopy is one of the most popular biomass sensors and has been applied so far to monitor biomass formation in submerged cultures of different organisms like stem cells, animal cells, plant cells, *Escherichia coli*, *Saccharomyces cerevisiae*, *Lactobacillus casei*, *Streptomyces sp.*, among others (Cannizzaro et al., 2003; Arnoux et al, 2005; Ferreira et al, 2005; Maskov et al 2008; Reis et al 2009; Knabben et al 2010; Justice et al., 2011; Horta et al 2011; Bryant et al 2011; Justice et al 2011).

Capacitance probe working principle is based on the polarization of cell membrane when placed in an ionic solution and subjected to an alternating electric field. Ion movement within the highly conducting cellular cytoplasm is restricted by the nonconductive cellular membrane and each cell will act as a capacitor. The polarization charge is measured by the capacitance of the solution, and the cells polarization is a function of the electric field frequency. At low frequencies (< 0.1 MHz), cell membranes are completely polarized and the capacitance of the solution is high. As the excitation frequency increases up to 10 MHz, the capacitance will decrease due to the incomplete cellular polarization. At high electric field frequencies (>10 MHz), there is not enough time for the complete cell polarization and only small particles are polarized. The magnitude of the membrane polarization is directly proportional to the volume that it encapsulates. Hence, the dielectric permittivity is directly proportional to the viable biomass concentration, since dead cells do not present polarization in their membranes (Davey and Kell, 1998; Markx and Davey, 1999).

The possibility of assessing the physiological state of the cells through the dielectric properties of cell suspensions is a very important feature when capacitance probes are used for monitoring genetically modified organisms (Matanguihan et al., 1994). Usually strong promoters are used for heterologous protein production in host cells, exposing them to a stressful condition after induction takes place. This problem has been widely described for recombinant protein expression in *E. coli* under the control of T7/*lac* and  $\lambda$  promoters (Khalilzadeh et al., 2003; Shiloach and Fass, 2005; Norsyahida et al. 2009) and the main consequences are: viability loss, growth cessation or even cell death, acetate production and substrate accumulation.

In addition, since dielectric permittivity measurements are sensitive to cell viability, they can be used for recognition of physiological changes during the cultivation (Markx and Davey, 1999; Kiviharju et al., 2008; Maskow et al, 2008), like Crabtree effect in yeast cultures (Davey et al, 1996); overflow metabolism in *E. coli* fermentations (Knabben et al., 2010) and spore formation in *Bacillus* or fungi cultivations (Sarrafzadeh et al, 2005).

Once having on-line capacitance measurements, the probe signal can be correlated to viable cell concentration. Furthermore, once on-line inferred biomass concentration data are available, it can be used in several applications, including the on-line estimation of the specific growth rate and the control of feeding flow rate in fed-batch cultures (Dabros et al., 2010; Horta et al., 2011).

This work presents a discussion of the use of the permittivity biomass sensor in batch and fed-batches cultures of 4 different microorganisms, with the following objectives: i) evaluate the performance of a capacitance sensor for on-line estimation of biomass concentrations of four different microorganisms; ii) estimate on-line values of the specific growth rates from the raw permittivity signal provided by the capacitance sensor.

### 5.3.3. MATERIALS AND METHODS

#### 5.3.3.1. MICROORGANISMS AND CULTIVATION MEDIA

Different wild and recombinant microorganisms were used in the studies:

- i) *Bacillus megaterium*: Strain PV361 was gently donated by Prof. Patricia S. Vary, Northern Illinois University. PV361 is a variant of the strain QMB1551, where the genes of its seven constitutive megaplasmids were deleted (Silva, 2001).
- ii) Four different clones of *Escherichia coli* BL21(DE3) were studied. The first one, called *E. coli* “a”, expresses a fragment of the antigenic surface protein SpaA from *Erysipelothrix rhusiopathiae* (Silva et al., 2011) (Cultivations B1 to B4 from Table 3.2.2). The second and third clone, called *E. coli* “b” and “c”, produces fragment of the antigenic surface proteins PspA clade 1 and 3 from *Streptococcus pneumoniae* (Carvalho et al., 2011) (Cultivations A1 to A16 from Table 3.2.2) and were kindly provided by Dr. Eliane Miyaji (Centro de Biotecnologia, Instituto Butantan, São Paulo, Brazil). The fourth clone, called *E. coli* “d”, transformed with the plasmid pT101/D-TOPO with *pac* gene codifying for penicillin G acylase (PGA) production (Montes et al., 2007) (Cultivations C1 and C2 from Table 3.2.2), was kindly donated by Laboratorio de Biocatálisis, ICP-CSIC-Madri, Spain.
- iii) Recombinant cells of *Pichia pastoris* GS115 (Invitrogen USA) expressing  $\alpha$ -amylase from *B. subtilis* were kindly provided by Prof. Fernando Torres (University of Brasilia) (Macauley-Patrick et al., 2005; Araújo, 2008; Burke et al., 2000).
- iv) *Saccharomyces cerevisiae*: batches and fed-batches were carried out with fresh baker’s yeast, using a complex medium (Reis et al., 2009).

Modified complex auto-induction media formulations (Studier, 2005, Silva et al., 2011), containing yeast extract and tryptone as nutrients and glycerol, glucose and a mixture of both as carbon sources, was used in cultivations of *E. coli* “a”; a modified defined HDF medium composition (Seeger et al., 1995; Sargo et al., 2011) was used in

cultivations of *E. coli* “b” and “c” and both media were used for cultivations of *E. coli* “d”. Commercial *S. cerevisiae* was cultivated in complex medium containing yeast extract as nutrient and glucose as main carbon source (Reis et al., 2009). Recombinant *P. pastoris* was cultivated in complex medium containing yeast extract and peptone as nutrient and glucose as main carbon source (Montano, 2010). *B. megaterium* was also cultivated in complex SNB and LB media, supplemented or not with cheese whey (Suárez, 2010).

#### 5.3.3.2. BIOREACTOR OPERATION, INSTRUMENTATION AND AUTOMATION

The cultivations were conducted in a 5 L (in-house) bioreactor monitored by SuperSys\_HCDC® (Horta et al, 2011). The pH was controlled (on/off) at 6.7 (*E. coli*), 6.5 (*Pichia pastoris*) and 7.5 (*Bacillus megaterium*) (pHmeter GLI PRO) by addition of H<sub>3</sub>PO<sub>4</sub> and NH<sub>4</sub>OH (30%). Temperature was set at 27 °C for *Saccharomyces cerevisiae*, 30 °C for *Bacillus megaterium*, *E. coli* “b” and “c” as well as for *Pichia pastoris*. *E. coli* “a” was cultivated at 37 °C. Dissolved oxygen concentration (DOC) was monitored by a Mettler Toledo probe Inpro 6800, connected to CE O<sub>2</sub> 4050 transmitter, and kept at 30% (20% for *Bacillus megaterium*) of saturation by a hybrid controller which automatically changed both agitation speed (between 200 and 900 rpm) and the composition of the gas stream supplied to the bioreactor (by mixing pure oxygen with air). The gas stream supplied to the bioreactor was a mixture of air and pure oxygen and its total flow rate was maintained at 6 L/min by two mass flow controllers (GFC AALBORG). The exhaust gas composition was assessed by a Sick/Maihak S.710 CO<sub>2</sub> and O<sub>2</sub> analyzer. The broth permittivity and conductivity were monitored by the biomass sensor from FOGALE® Nanotech. On-line data acquisition as well as monitoring/control of all instruments via a compact field point 2020 (cFP-2020, National Instruments) was performed by the software SuperSys\_HCDC<sup>R</sup> developed in LabView®.

The feeding supply was provided by Ismatec – BVP pump following the exponential profile given by equation (5.3.1) (Nielsen et al., 2002).

$$F = \left( \frac{\mu}{Y_{XS}} + m \right) \cdot \frac{C_{X0} \cdot V_0}{C_{S0} - C_{SR}} \cdot e^{(\mu_{SET} \cdot t)} \quad (5.3.1)$$

In Equation (5.3.1),  $F$  ( $L \cdot h^{-1}$ ) is the feed flow rate,  $\mu_{SET}$  ( $h^{-1}$ ) is the desired specific growth,  $Y_{XS}$  ( $g_{DCW} \cdot g_{glycerol}^{-1}$ ) is the biomass yield coefficient on glycerol,  $m$  ( $g_{DCW} \cdot g_{glycerol}^{-1} \cdot h^{-1}$ ) is the maintenance coefficient,  $C_{X0}$  [ $g_{DCW} \cdot L^{-1}$ ] and  $V_0$  (L) correspond to the cellular concentration and volume, respectively, at the beginning of the fed-batch phase,  $C_{S0}$  ( $g_{glycerol} \cdot L^{-1}$ ) is the carbon source concentration in the supplementary medium and  $C_{SR}$  ( $g_{glycerol} \cdot L^{-1}$ ) represents the residual glycerol concentration.

The exponential feed profile changed accordingly the cultivated microorganism. For *B. megaterium* and *P. pastoris*, a pre-defined feeding profile with  $\mu_{SET}$  equal to  $0.12 h^{-1}$  was used. For *E. coli*, different feeding strategies were studied, including pre-defined feeding profiles at  $\mu_{SET}$  equal to  $0.13 h^{-1}$  (*E. coli* "a") and  $0.3 h^{-1}$  (*E. coli* "b"). In addition, cultivations of *E. coli* "a", "b" and "c" were also carried out using an advanced controller for the feeding flow rate given by Eq. (1), where the parameters  $\mu_{SET}$ ,  $Y_{XS}$  e  $m$  were continuously and automatically retuned at each 10 min using the values of  $\mu$  obtained online from the permittivity measurements provided by the capacitance biomass sensor (Horta et al., 2011c).

#### 5.3.3.3. ANALYTICAL METHODS

Biomass concentration in the samples withdrawal during the cultivations was measured by optical density reading at  $\lambda = 600$  nm. Biomass concentration was also assessed by dry cell weight method. Metabolites concentration and carbon sources consumption were measured by HPLC as described in Silva et al., (2008). Glycerol and glucose were also measured by enzymatic kits (Silva et al., 2008; Sargo et al., 2011).

#### 5.3.3.4. ON-LINE BIOMASS INFERENCE AND SPECIFIC GROWTH RATE ESTIMATION

During all cultivations, online biomass monitoring by a capacitance sensor (BIOMASS System - Viable cell density monitoring - FOGALE nanotech) was carried out. A reference frequency of 10 MHz was used as blank signal, to compensate disturbances caused by changes in medium composition and on the overall bioreactor environment (agitation, aeration, presence of debris etc). The frequency of measurement was set at 2.076 MHz for yeast and 2.984 MHz for bacterial cultivations.

The response signal at the reference frequency was automatically subtracted from the signal at the frequency of measurement (Fogale, 2008).

The probe response was transferred to the computer database via its analog output, in an acquisition rate of one second. These data were used for on-line estimation of biomass concentration and specific growth rate as shortly described in the followings. A detailed description of data treatment, including pseudocodes equations and calculations, is available at Horta et al, 2011.

Permittivity data was acquired as 0.004-0.02 Ampere (A) analogical signal from the Fogale Nanotech Instrument and multiplied by a calibration factor proportional to the instrument output range to convert the readings in A to pF/cm. This analogical signal was noisy and it had to be firstly treated by a smoothed moving average (SMA) filter to be used as input in the calibration curve relating the smoothed permittivity to cellular concentration ( $g_{DCW}/L$ ). The smoothed permittivity output from SMA filter, combined with mass balance equations for batch and fed-batch cultivations, was also used to estimate the specific growth rate ( $\mu$ ). The estimated raw  $\mu$  data were further treated by a SMA filter to improve its smoothness (Horta et al., 2011).

#### 5.3.4. RESULTS AND DISCUSSION

##### 5.3.4.1. CORRELATION BETWEEN PERMITTIVITY AND CELL CONCENTRATION – CX ON-LINE ESTIMATION

Growth of microorganisms with different sizes, shapes and physiological states was on-line monitored by permittivity measurement.

The correlation between permittivity and cellular concentration can be observed in Figure 5.3.1, for *Bacillus megaterium*, *Saccharomyces cerevisiae* and *Pichia pastoris* and in Figure 5.3.2 for the recombinant *Escherichia coli* strains.

It can be seen in Figure 5.3.1 and Figure 5.3.2 that the correlation of permittivity and cellular concentration changes accordingly to each microorganism, as expected. The good fittings and very similar correlations were obtained for yeast cells, whose average diameters are in the ranges (5 to 10  $\mu\text{m}$  for *S. cerevisiae* and 1 to 2.2  $\mu\text{m}$  for *P. pastoris*). Yet, the correlation parameters for *B. megaterium* and *E. coli* were rather



different, reflecting the differences between these two bacteria in cell size (*E. coli* has a total volume of  $\sim 1 \mu\text{m}^3$ ; *B. megaterium* cell volume is  $\sim 60 \mu\text{m}^3$ ) and cell wall composition (*B. megaterium* is Gram positive, *E. coli* is Gram negative).

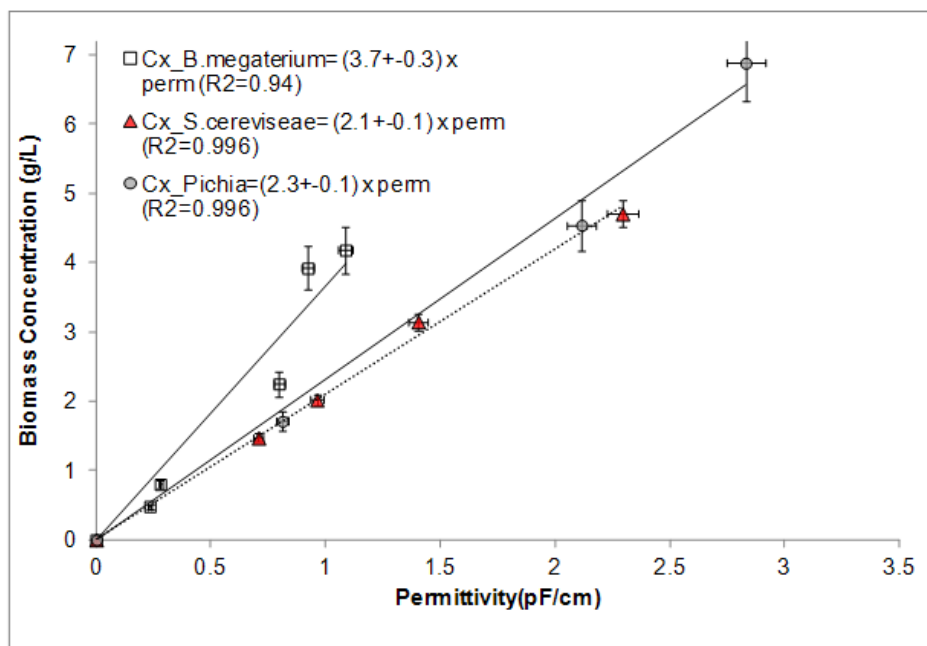


Figure 5.3.1. Correlation between permittivity [Perm (pF/cm)] and cellular concentration [Cx ( $g_{DCW}/L$ )] for *B. megaterium*, *S. cerevisiae* and *Pichia pastoris*. Symbols correspond to experimental data, lines correspond to linear fit.

From the huge amount of available data for the 4 different *E. coli* clones grown in different media, for concentrations ranging from 0.5 to 150  $g_{DCW}/L$ , a deep analysis of the capacitance sensor performance can be made. Thus, comparing *E. coli* “a”, “b”, “c” and “d” cultures, it can be noticed from Figure 5.3.2 that changes in temperature and medium composition do not cause significant deviation between permittivity and cellular concentration. In fact, the unique correlation reproduced in Eq. 5.3.2 could be used to estimate biomass concentration for all strains, during the growth phase. We can also observe significant scattering for biomass concentration lower than 20  $g_{DCW}/L$  due to the lack of accuracy of permittivity measurements within this range.

On the other hand, 50 % of the data belonging to the induction phase could not be described by Eq. 5.3.2 and the new correlation given at Eq. 5.3.3 was generated to fit

these data. Comparing the coefficients of Eqs. 5.3.2 and 5.3.3, it can be noticed that the slope increases during the induction phase. This phenomenon, not mentioned in the literature, is caused by the loss of cell viability associated to the stress caused by heterologous protein synthesis. Since non-viable cells do not contribute to the permittivity, the capacitance probe precisely reflected the change in cell physiology taking place during the induction phase. This change is not detected by conventional optical density or dry weight measurements, which can not distinguish viable and non viable cells. Because of this mismatch between the two methodologies, a new correlation relating dry cell weight measurements to permittivity is required.

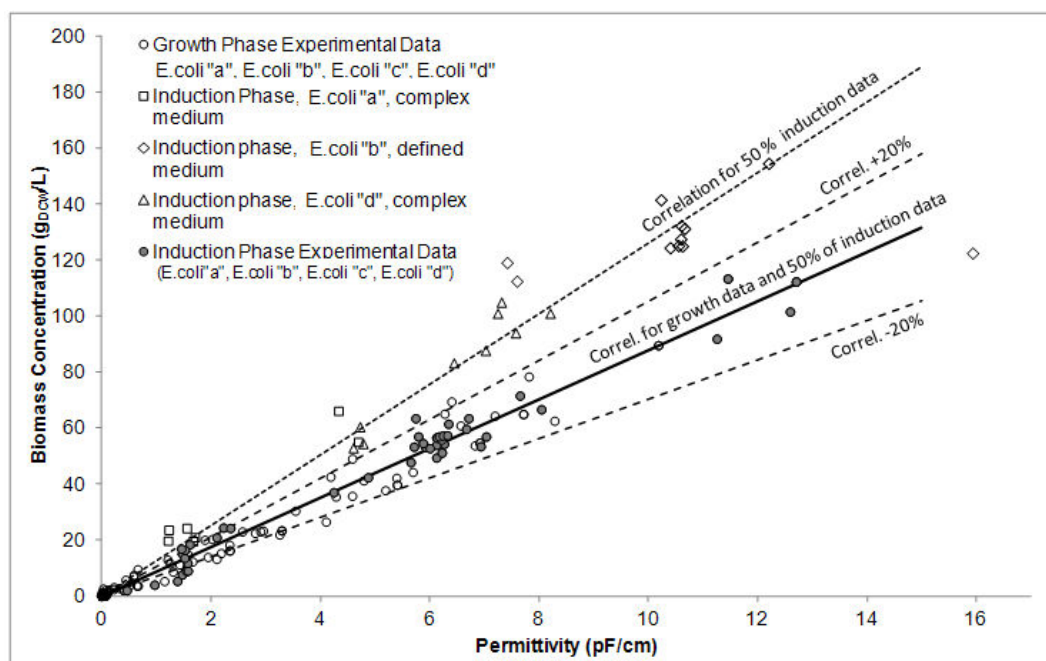


Figure 5.3.2. Correlation between permittivity [Perm (pF/cm)] and cellular concentration [Cx (gDCW/L)] for different recombinant *Escherichia coli* strains. Symbols correspond to experimental data, lines correspond to linear fit.

$$Cx = (8.8 \pm 0.1) \cdot \text{Perm} \quad (R^2 = 0.989) \quad (5.3.1)$$

$$Cx = (12.6 \pm 0.2) \cdot \text{Perm} \quad (R^2 = 0.984) \quad (5.3.2)$$

It is also interesting to notice that half of the induction phase data is described by Eq. (5.3.1 (Group 1) while the other half is described by Eq. (5.3.2 (Group 2). Group 1 contains the experimental points belonging the beginning of induction. For the all cases

discussed here, the production of heterologous protein was triggered by inducer addition, which can be IPTG or lactose. Shortly after inducer addition, the cells have not undergone drastic metabolic changes, which explain the validity of Eq. (5.3.1). Group 1 also includes 85 % of the experimental points from the cultivations carried out with *E. coli* “a” and “d” in complex media. It has been suggested in the literature that medium supplementation with aminoacids or proteins can partially relieve the stress caused by the recombinant protein production (Shiloach and Fass, 2005; Shojaosadati et al., 2008; Tripathi et al., 2009). On the other hand, Group 2 mainly consists of data from *E. coli* “b” and “c” cultures, which were carried out in defined media. In these cases, the metabolic effort to produce the recombinant protein from basic medium components (glycerol, ammonia and salts) was significantly higher, impairing cell viability.

#### 5.3.4.2. ON-LINE GROWTH RATE ESTIMATION

As shown previously, reliable correlations between viable cell concentrations and permittivity data can be set-up for all cases studied. On-line permittivity signals can also be employed for inferring the specific growth rate, after subjected to the appropriate treatment with SMA filter for improving their smoothness (Horta et al, 2011).

Table 5.3.1 shows the optimal values of the SMA length vector for each microorganism studied in this work.

Table 5.3.1. Optimal length of SMA filter smoothing vectors for the different microorganisms.

Microorganism	n
<i>Escherichia coli</i>	40
<i>Saccharomyces cerevisiae</i>	50
<i>Pichia pastoris</i>	50
<i>Bacillus megaterium</i>	110

Figure 5.3.3 compares both raw and smoothed permittivity data, and presents the on-line  $\mu$  values estimated from the smoothed permittivity for microorganisms studied. The estimated  $\mu$  curve followed closely cell growth pattern, but the  $\mu$  data were still too noisy, even after being treated by a SMA filter.

The amplitude of the oscillations was generally higher in the first 5-10 h of cultivation. As explained previously, a significant scattering on the estimated biomass concentration is caused by the low precision of the capacitance probe for measuring permittivity values below  $\sim 1$  pF/cm. The mathematical manipulation to estimate  $\mu$  also contributes to amplify the noise. Time course oscillation of  $\mu$  was also reported by other researchers (Xiong et al., 2008; Davey et al., 1996; Ferreira et al., 2005) and according to Davey et al (1996) this oscillatory behaviour can also be caused by continuous readout.

This noise in the  $\mu$  signal presents no problems for the analysis of growth profiles, but it can not be used to control the feed flow rate. Nevertheless, feed supply is usually started at higher biomass concentrations, where the inferred  $\mu$  tends to be more stable (Figure 5.3.3 A, C and G for  $t > 20, 14$  and  $12$  h, respectively)

Besides the on-line inference of specific growth rate, that can be used for monitoring or updating  $\mu_{SET}$  at Eq. (5.3.1), the metabolic changes well reproduced by permittivity data can also be applied to improve the tuning of other parameters in the same equation. During the induction phase, cell metabolism is driven to the foreign protein production. Therefore, the assumption of approximately constant values for the maintenance ( $m$ ) and biomass yield ( $Y_{xs}$ ) coefficients in Eq. (5.3.1) does not hold anymore. In fact,  $Y_{xs}$  decreases and  $m$  increases along the induction phase and these changes need to be incorporated in Eq. (5.3.1) for a better control of the bioprocess (Horta et al., 2011).

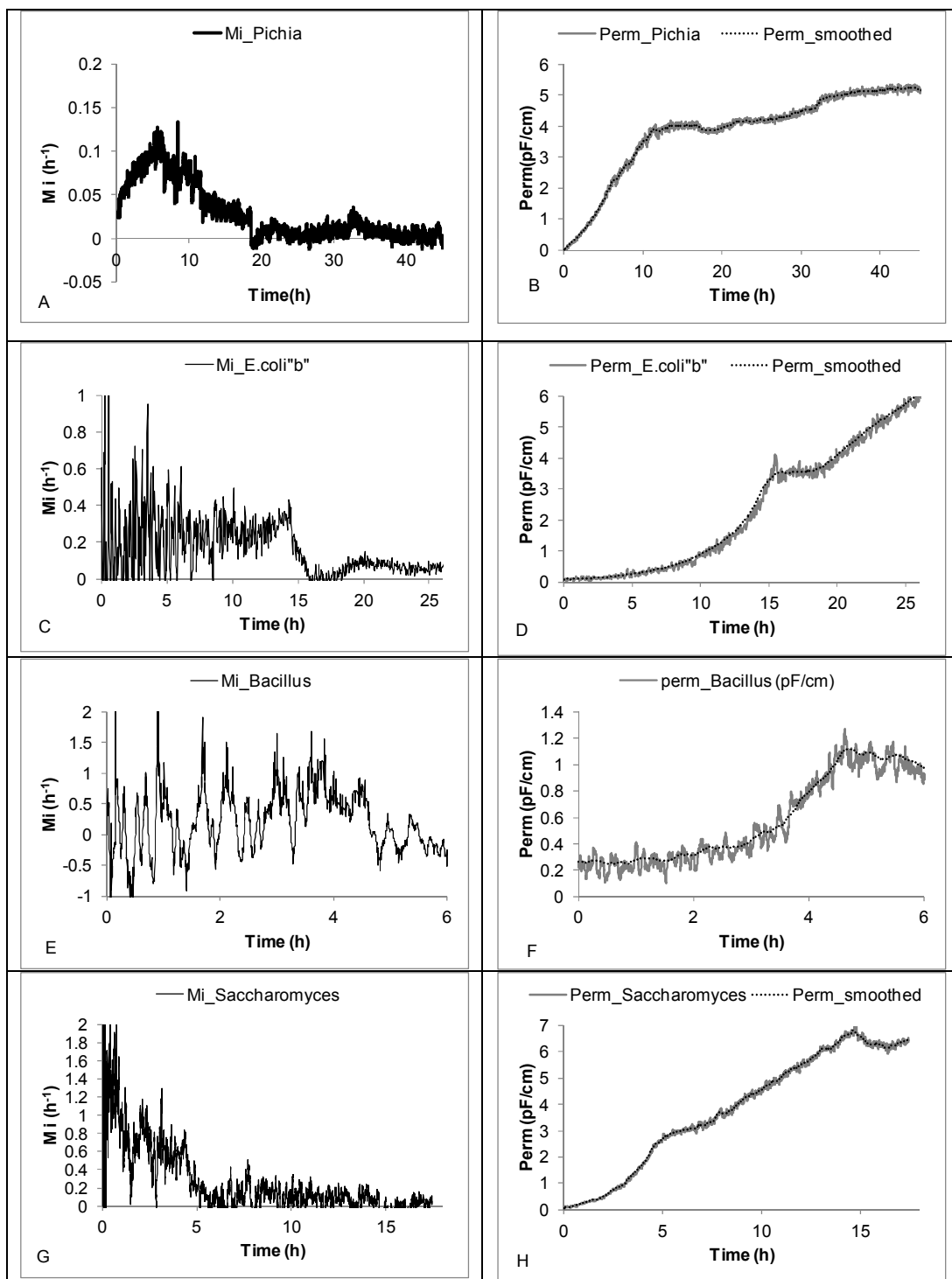


Figure 5.3.3. In the left  $\mu$  ( $h^{-1}$ ) calculated online from smoothed permittivity (pF/cm), in the right permittivity (continuous line) and smoothed permittivity (dotted line). A and B: *Pichia pastoris* (Montaño, 2010). C and D: *Escherichia coli* "b". E and F: *Bacillus megaterium* (Suárez, 2010). G and H: *Saccharomyces cerevisiae*.

### 5.3.5. CONCLUSION

The capacitance sensor showed to be a reliable tool for biomass on-line monitoring during cultivation of different kinds of microorganisms, including Gram-positive and Gram-negative bacteria and the two yeasts mainly used nowadays in the biotechnology industry, *S. cerevisiae* and *P. pastoris*. After filter processing steps, the permittivity data collected allowed to on-line follow the biomass formation and to estimate the specific growth rate during the cultivation process.

For recombinant *E. coli* cultures, the drastic changes occurring during induction phase were clearly indicated by permittivity measurements. The analysis of permittivity data helped to understand the influence of media composition and induction period on the cell response to the stress caused by protein expression. They also could be used to assess the influence of other factors, such as temperature and type of inducer, during the protein production phase.

The inference of on-line specific growth rate can certainly contribute for improving the control of the growth rate and of the bioprocess as a whole. Based on the present results, a dynamic control of the exponential growth rate by maintaining the feeding rate at optimal levels in fed-batch processes was developed and implemented for *rE. coli* high cell density cultivations (Horta et al, 2011). This technology is part of the monitoring and control software SUPERSYS\_HCDC<sup>®</sup> developed and registered (Horta et al, 2011b) by the researchers from the Laboratory of Development and Automation of Bioprocesses, Department of Chemical Engineering of Federal University of São Carlos, Brazil.

### **Acknowledgements**

The authors thank the Brazilian research-funding agencies: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Financiadora de Estudos e Projetos (FINEP).

## REFERENCES

- Araújo, J., (2008) Production of chymosin B from *Bos taurus* in *Pichia pastoris*. University of Brasilia. Dissertation
- Burke, D.; Dawson, D.; Stearns, T., (2000) Methods in Yeast Genetics: A Laboratory Course Manual (Cold Spring Harbor Lab. Press, Plainview, NY)
- Carvalho, R.J.; Cabrera-Crespo, J.; Tanizaki, M.M.; Gonçalves, V.M., (2011) Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. Appl Microbiol and Biotechnol. DOI: 10.1007/s00253-011-3649-9.
- Dabros, M.; Schuler, M.M.; Marison, I.W. (2010) Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. Bioprocess Biosyst Eng 33:1109-1118
- Davey, H.M., Davey, C. L., Woodward, A.M., Edmonds, A. N., Lee, A.W., and Kell, D. B. (1996) Oscillatory, stochastic and chaotic growth rate fluctuations in permissively controlled yeast cultures. Biosyst 39:43–61
- Ferreira, A.P., Vieira, L.M., Cardoso, J.P., Menezes, J.C. (2005) Evaluation of a new annular capacitance probe for biomass monitoring in industrial pilot-scale fermentations. J Biotechnol 116:403-409
- Fogale Nanotech (2008) Biomass System: User Manual, [www.fogalebiotech.com](http://www.fogalebiotech.com). Accessed 20 Sept 2008
- Horta, A.C.L; Sargo, C.R.; Silva, A.J.; Gonzaga, M.S.; Gonçalves, V.M.; Zangirolami, T.C.; Giordano, R.C. (2011) Intensification of high cell density cultivations of *rE. coli* for production of *S. pneumonia* antigenic surface protein, PspA3, using model-based adaptive control. Bioprocess Biosyst. Eng., Submitted: 23 Nov. 2011.
- Horta, ACL; Zangirolami, TC; Giordano, RC; Cruz, AJG; Reis, GB; Jesus, CDF (2011b) Supervisory system for bioreactor high cell density cultivations. Software registration proc. 11008-6 INPI, Brazil.
- Kiviharju, K.; Salomen, K.; Moilanen, U.; Eerikainen, T. (2008) Biomass measurement online: the performance of in situ measurements and software sensors. J Ind Microbiol Biotechnol 35:657-665
- Macauley-Patrick, S.; Fazenda, M.L.; Mcneil, B.; Harvey, L.M., (2005) Heterologous protein production using the *Pichia pastoris* expression system. Yeast 22:249–270
- Markx, G.H.; Davey, C.L. (1999) The dielectric properties of biological cells at radiofrequencies: Applications in biotechnology. Enzym Microb Technol 25:161–171

Maskow, T.; Röllich, A.; Fetzer, I.; Yao, J.; Harms, H. (2008) Observation of non-linear biomass–capacitance correlations: Reasons and implications for bioprocess control. *Biosens Bioelectron* 24:123-128

Montaño, I.D.C. (2010) Otimização Dinâmica do Cultivo Semi-contínuo de *Pichi pastoris* recombinante para produção de enzimas heterólogas. Dissertação de mestrado, PPG-EQ, UFSCar. (in Portuguese)

Montes, T.; Grazú, V.; López-Gallego, F.; Hermoso, J.A.; Garcia, J.L.; Manso, I.; Galán, B.; González, R.; Fernandez-Lafuente, R.; Guisán, J.M. (2007) Genetic modification of the Penicillin G Acylase surface to improve its reversible immobilization on ionic exchangers. *Applied and environmental microbiology*, 312-319.

Nielsen, J.; Villadsen, J.; Lidén, G. (2002), *Bioreaction Engineering Principles*, 2nd Ed., Kluwer Academic/Plenum Publishers, New York

Reis, G.B.; Horta, A.C.L., Zangirolami, T.C., Giordano, R.C., Cruz, A.J.G, (2009) Control of Fed-Batch Yeast Cultivation Using a Capacitance Sensor. *10th Int Symp Process Syst Eng: Part A Computer Aided Chemical Eng*, 27:1323-1328

Sargo, C.R.; Horta, A.C.L.; Silva, A.J.; Possedente, M.; Machado, M.M.P.; Giordano, R.L.C.; Giordano, R.C.; Gonçalves, V.M.; Zangirolami, T.C. (2011) Improved *rE. coli* High Cell Density Cultivation Conditions for Glycerol as Carbon Source. Submitted to *Applied Microbiology and Biotechnology*, December

Sarrafzadeh, M. H., Belloy, L., Esteban, G., Navarro, J.M., and Ghommidh, C. (2005) Dielectric monitoring of growth and sporulation of *Bacillus thuringiensis*. *Biotechnol Lett* 27:511–517

Seeger, A.; Schbeppe, B.; McCarthy, J.E.G.; Deckwer, W.D.; Rinas, U. (1995) Comparison of temperature and isopropyl- $\beta$ -D-thiogalacto-pyranoside-induced synthesis of basic fibroblast growth factor in high-cell-density cultures of recombinant *Escherichia coli*. *Enz. and Microb. Tech.* 17: 947-953

Shiloach, J., Fass, R. (2005) Growing *E. coli* to high cell density: a historical perspective on method development. *Biotechnol Adv* 23:345–357

Shojaosadati, S.A.; Kolaei, S.M.V.; Babaeipour, V.; Farnoud, A.M. (2008) Recent advances in high cell density cultivation for production of recombinant protein. *Iranian Journal of Biotechnology*, 6:63-84

Silva, A.J.; Baptista-Neto, A.; Cilento, M.C.; Giordano, R.C.; Zangirolami, T.C. (2008) Bioreactor aeration conditions modulate growth and antigen expression during *Erysipelothrix rhusiopathiae* cultivation. *Appl Microbiol Biotechnol* 79:23–31

Silva, A.J.; Iemma, M.R.C.; Horta, A.C.L.; Sargo, C.R.; Giordano, R.L.C.; Giordano, R.C.; Zangirolami, T.C.; Novo, M.T.M. (2011) Cloning, expression, purification and auto-induction production of rSpaA swine erysipelas antigen. Submitted to *Current Microbiology*, December.

Silva, M. ; Lopes, A. P. Y. ; Gamberini, M. ; Schenkman, R. P. ; Ramos, C. R. ; Ho, P.L. ; Leite, L. C. C. (2001) Improved expression of PsaA from *Streptococcus pneumoniae* in *E.coli* and its purification. In: *Reunião Anual da Sociedade Brasileira de Bioquímica*, 2001, Caxambu



Suárez, C.A.G. (2010) Otimização do cultivo de *Bacillus megaterium* recombinante em bateladas alimentadas. Dissertação de mestrado, PPG-EQ, UFSCar (in Portuguese)

Studier, F.W. (2005) Protein production by auto-induction in high-density shaking culture. *Protein Express Purif* 41:207-234

Tripathi, N.K.; Sathyaseelan, K.; Jana, A.M.; Rao, P.V.L., (2009) High Yield Production of Heterologous Proteins with *Escherichia coli*, *Defence Science Journal*, 59 (2):137-146.

Xiong, Z.Q.; Guo, M.J.; Guo, Y.X.; Chu, J.; Zhuang, Y.P.; Zhang, S.L. (2008) Real-Time Viable-Cell Mass Monitoring in High-Cell-Density Fed-Batch Glutathione Fermentation by *Saccharomyces cerevisiae* T65 in Industrial Complex Medium. *J Biosci bioeng* 105(4):409–413

## 6. CONCLUSÃO

Os objetivos propostos inicialmente para o presente trabalho foram atingidos tanto em termos do desenvolvimento do software como da aplicação do mesmo na otimização dos cultivos de alta densidade de *E. coli* recombinante.

O programa de computador SUPERSYS\_HCDC evoluiu à medida que os cultivos em batelada alimentada de *E. coli* recombinante e outros microrganismos eram realizados, sendo um atrativo software de supervisão e controle de biorreatores disponível no mercado atualmente. A versão mais recente dispõe de controle automático da concentração de oxigênio dissolvido e das vazões de ar e de oxigênio (visando ao menor consumo possível de oxigênio) e realiza automaticamente a interrupção da alimentação para o controle de temperatura e da concentração de oxigênio dissolvido. Técnicas de inteligência computacional foram incorporadas para execução de algumas funcionalidades, tais como o acionamento automático da bomba de alimentação ao término da fase de batelada e a atualização dinâmica da vazão de alimentação em função da velocidade de crescimento estimada a partir dos dados de permissividade obtidos com o sensor de biomassa.

A automatação da maior parte das decisões de controle por meio desta ferramenta contribuiu para a diminuição de custos (menor consumo de oxigênio e de nutrientes) e também para tornar o processo mais reprodutível, deixando-o menos dependente da tomada de decisões de especialistas humanos. Além disso, a presença do software e de seus recursos de comunicação (alarmes sonoros, mensagens via e-

mail, SMS e acesso remoto) tornou a execução dos cultivos em batelada alimentada, que podem durar 40 h ou mais, menos exaustiva para a equipe de operadores.

Por outro lado, o aumento na autonomia do software e no número de funcionalidades disponíveis foi acompanhado pelo aumento na complexidade técnica de operação da ferramenta. Portanto, apesar de estar disponível como software livre, a realização de um treinamento é imprescindível para a instalação, a configuração e a operação do SUPERSYS\_HCDC. Pelo fato de ter sido selecionado no Edital do Programa SEARCH SEBRAE de 2011, o “produto” SUPERSYS\_HCDC será submetido a uma análise de viabilidade mercadológica e econômica pela Fundação Biominas em 2012, o que permitirá uma melhor caracterização dos pontos fortes e das limitações do mesmo.

Após quatro anos de estudos da literatura relacionada a cultivos de alta densidade e observações experimentais, conseguiu-se definir uma estratégia ótima de operação de CADC para glicerol como fonte de carbono, mantendo crescimento acelerado, sem acúmulo inibitório de acetato e com boa expressão de proteína recombinante. A estratégia mais bem sucedida consistiu em manter a velocidade de crescimento próxima à capacidade de crescimento das próprias células em cada fase do cultivo; garantir uma concentração residual de glicerol no meio em torno de 5 g/L; e usar todos os recursos disponíveis para manter a concentração de oxigênio disponível em ~ 30% da saturação. Graças aos melhoramentos introduzidos no bioprocessamento, a produtividade em biomassa nos cultivos de alta densidade celular de *E. coli* recombinante evoluiu de 2,1 para 9,2 gL<sup>-1</sup>h<sup>-1</sup>. Parte significativa deste aumento (de 2,1 para 6 gL<sup>-1</sup>h<sup>-1</sup>) é resultado do melhoramento na formulação do meio de alimentação contendo glicerol como fonte de carbono. No entanto, o ganho de 50 % em

produtividade para alcançar o valor  $9,2 \text{ gL}^{-1}\text{h}^{-1}$  é resultado da utilização do software como ferramenta de supervisão e controle do bioprocesso e da implementação de estratégias de controle da velocidade de crescimento baseadas em inteligência computacional embarcada no software. Em consequência das melhores condições de cultivo estabelecidas, a concentração de proteína recombinante também evoluiu de 3 para mais de 25 g/L em reator de 5L.

## 7. SUGESTÕES PARA TRABALHOS FUTUROS

Para dar continuidade ao processo de melhoria e otimização de CADC algumas ferramentas podem ser desenvolvidas:

- i) Implementação de um sistema que substitua o sensor de biomassa para o cálculo do  $\mu$ , e de Cx baseado em redes neurais artificiais, que possam inferir a evolução de Cx ao longo do cultivo, tendo como dados de entrada o volume de base adicionado, a liberação de  $\text{CO}_2$ , entre outras variáveis.
- ii) Implementação de um sistema de ajuste de modelos que funcione *on-line*, corrigindo os parâmetros de alimentação. Os parâmetros podem ser corrigidos por redes neurais previamente treinadas, eliminando a necessidade de medidas da concentração de substrato em tempo real.
- iii) Evolução do controle de oxigênio dissolvido utilizando lógica fuzzy e/ou RN.

- iv) Desenvolvimento de sistema inteligente, baseado em lógica fuzzy e no sensor de biomassa para ativar a bomba de alimentação ainda com uma concentração residual elevada de glicerol, evitando desaceleração da velocidade específica de crescimento ao final da batelada.
- v) Mapear o comportamento das variáveis adquiridas *on-line* durante a fase de indução, visando ao desenvolvimento de uma ferramenta baseada em inteligência artificial para reconhecer o momento de máxima expressão da proteína e interromper o cultivo diminuindo a temperatura para 5°C e avisando o operador.
- vi) Explorar com mais profundidade o comportamento do sensor de biomassa na batelada, identificando as condições de maior atividade metabólica, procurando correlacionar a concentração de glicerol com a atividade fisiológica.
- vii) Adaptar o sistema desenvolvido para cultivos de bactérias em reator tipo tanque agitado, SUPERSYS\_HCDC, para controlar outros tipos de biorreatores ou cultivos de outros organismos.

## REFERÊNCIAS

ABI, A.; SARRAFZADEH, M.H.; MEHRNIA, M.R.; GHOMMIDH, C. Application of dielectric permittivity measurements in physiological state monitoring of *Bacillus subtilis* culture. 2010 2ND INTERNATIONAL CONFERENCE ON CHEMICAL, BIOLOGICAL AND ENVIRONMENTAL ENGINEERING, ICBE 2010, p. 26-29.

AKESSON, M.; HAGANDER, P.; AXELSSON, J.P. Probing control of fed-batch cultivations: analysis and tuning. **Control Engineering Practice**, v. 9, p. 709-723, 2001

ALFAFARA, C.G.; MIURA, K.; SHIMIZU, H.; SHIOYA, S.; SUGA, K.; SUZUKI, K. Fuzzy control of ethanol concentration and its application to maximum glutathione production in yeast fed-batch culture. **Biotechnol Bioeng**, v. 41, p. 493-501, 1993.

ARISTIDOU, A.; SAN KA-Y; BENNETT, G.N. Improvement of biomass yield and recombinant gene expression in *Escherichia coli* by using fructose as the primary carbon source. **Biotechnology Progress** v. 15, p. 140-145, 1999.

AULICINO, J.; HERMIDA, M.; MEDINA, R.; SABAROTS, L.; DUCREY, L.; ORTI, E.; ROGERS, F. Developing an automatically controlled feeding process in an *E. coli* fermentation process for recombinant protein production. **Biopharm International**, June, 2010.

BABAEIPOUR, V.; SHOJAOSADATI, S.A.; ROBATJAZI, S.M.; KHALILZADEH, R.; MAGHSOUDI, N. Over-production of human interferon-g by HCDC of recombinant *Escherichia coli*. **Process Biochemistry**, v. 42, p. 112-117, 2007.

BÄCKLUND, E.; MARKLAND, K.; LARSSON, G. Cell engineering of *Escherichia coli* allows high cell density accumulation without fed-batch process control. **Bioprocess Biosyst. Eng.**, v. 31, p. 11-20, 2008.

BAI, W., ZHAO, K.S., ASAMI, K. Dielectric properties of *E. coli* cell as simulated by the three-shell spheroidal model. **Biophysical Chemistry**, v. 122, p. 136-142, 2006.

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

BARAZZONE, G.C.; CARVALHO, R.J.R.; KRASCHOWETZ, S.; HORTA, A.C.L.; SARGO, C.R.; SILVA, A.J.; ZANGIROLAMI, T.C.; GOULART, C.; LEITE, L.C.C.; TANIZAKI, M.M.; GONÇALVES, V.M.; CABRERA-CRESPO, J. Production and purification of recombinant fragment of pneumococcal surface protein A (PspA) in *Escherichia coli*. **Procedia in Vaccinology**, v. 4, p. 27-35, 2011.

BEREZIN, EM, CARVALHO, LH, LOPEZ, CR, SANAJOTTA, AT, BRANDILEONE, MCC, MENEGATTI, S. SAFADI, MA, GUERRA, MLCS. **J. Pediatr**, Rio, v. 78, p. 19-23, 2002.

BEUTEL, S.; HENKEL, S. In situ sensor techniques in modern bioprocess monitoring. **Appl. Microbiol. Biotechnol.**, 91:1493–1505, 2011.

BRANDILEONE, M.C., DE ANDRADE, A.L., DI FABIO, J.L., GUERRA, M.L. AND AUSTRIAN, R. Appropriateness of a pneumococcal conjugate vaccine in Brazil: potential impact of age and clinical diagnosis, with emphasis on meningitis. **J Infect Dis**, v. 187, p. 1206-1212, 2003.

CARVALHO, R.J.; CABRERA-CRESPO, J.; TANIZAKI, M.M.; GONÇALVES, V.M. Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. **Appl. Microbiol. Biotechnol.**, 1-12, 2011.

CASTAN, A.; NÄSMAN, A.; ENFORS, S.O. Oxygen enriched air supply in *Escherichia coli* processes: production of biomass and recombinant human growth hormone. **Enzyme and Microbial Technology**, v. 30, p. 847–854, 2002.

CAVALCANTI\_MONTANO, I.D. Otimização dinâmica do cultivo semi-contínuo de *Pichia pastoris* recombinante para produção das enzimas heterólogas alfa amilase e penicilina G acilase. Dissertação de mestrado apresentada ao programa de Engenharia Química da Universidade Federal de São Carlos, 2010.

CHOI, J.H.; KEUM, K.C.; LEE, S.Y. Production of recombinant proteins by high cell density culture of *Escherichia coli*. **Chemical Engineering Science** v. 61, p. 876-885, 2006.

DAVEY, C.L.; KELL, D.B. The influence of electrode polarisation on dielectric spectra, with special reference to capacitive biomass measurements: (II) reduction in the contribution of electrode polarisation to dielectric spectra using a two-frequency method. **Bioelectrochemistry and Bioenergetics**, v. 46, p. 105–114, 1998.

DABROS, M.; SCHULER, M.M.; MARISON, I.W. Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. **Bioprocess Biosyst. Eng.**, v. 33, p. 1109–1118, 2010.

DEBRECZENY, M.P; DAVIES, E.T. **Non-Invasive Biomass Monitor with Wide Linear Range**. Disponível em: <<http://www.buglab.com/docs/OnLineODMeasurement.pdf>>; Acesso em 10 nov. 2011. 2008

DEBRECZENY, M.P; O'NEIL, M.P; KASAPI, A. Self-Compensating Radiation Sensor With Wide Dynamic Range. **Patent No.: US 6,573,991 B1**. Date of Patent: Jun. 3, 2003.

DEMAIN, A.L.; VAISHNAV, P. Production of recombinant proteins by microbes and higher organisms. **Biotechnology Advances**, v. 27, p. 297-306, 2009.

FAULKNER, E.; BARRETT, M.; OKOR, S.; KIERAN, P.; CASEY, E.; PARADISI, F.; ENGEL, P.; GLENNON, B. Use of fed-batch cultivation for achieving high cell densities

for the pilot-scale production of a recombinant protein (phenylalanine dehydrogenase) in *Escherichia coli*. **Biotechnol. Progr.**, v. 22, p. 889-897, 2006.

FERREIRA, A.P., VIEIRA, L.M., CARDOSO, J.P., MENEZES, J.C. Evaluation of a new annular capacitance probe for biomass monitoring in industrial pilot-scale fermentations. **Journal of Biotechnology**, v. 116, p. 403-409, 2005.

GONZALEZ-VARA, R.A.; VACCARI, G; DOSI, E.; TRILLI, A.; ROSSI, M.; MATTEUZZI, D. Enhanced production of L-(+)-lactic acid in chemostat by *Lactobacillus casei* DSM 20011 using ion-exchange resins and cross-flow filtration in a fully automated pilot plant controlled via NIR. **Biotechnol Bioeng.**, v. 67, p. 147–156, 2000.

GUFLER, P.C., PUM, D., SLEYTR, U.B., SCHUSTER, B. Highly robust lipid membranes on crystalline S-layer supports investigated by electrochemical impedance spectroscopy. **Biochimica et Biophysica Acta**, v. 1661, p. 154-165, 2004.

HONDA, H.; KOBAYASHI, T. Fuzzy control of bioprocess. **Journal of Bioscience and Bioengineering**, v. 89, n. 5, p. 401-408, 2000.

HORTA, A.C.L.; SILVA, A.J.; SARGO, C.R.; GONÇALVES, V.M.; ZANGIROLAMI, T.C.; GIORDANO, R.C. Robust artificial intelligence tool for automatic start-up of the supplementary medium feeding in recombinant *E. coli* cultivations. **Bioprocess Biosyst. Eng.**, v. 33, n.7, p. 891-901, 2011.

HORTA, A.C.L.; SARGO, C.R.; SILVA, A.J.; GONZAGA, M.S.; GONÇALVES, V.M.; ZANGIROLAMI, T.C.; GIORDANO, R.C. Intensification of high cell density cultivations of *rE. coli* for production of *S. pneumonia* antigenic surface protein, PspA3, using model-based adaptive control. **Bioprocess Biosyst. Eng.**, Submitted: 23 Nov. 2011, 2011b.

HORTA, A.C.L.; ZANGIROLAMI, T.C.; GIORDANO, R.C.; CRUZ, A.J.G.; REIS, G.B.; JESUS, C.D.F. SUPERSYS-HCDC - Programa de computador para monitoramento e controle de biorreator para cultivos de alta densidade celular (Supervisory system for bioreactor high cell density cultivations). Registro de software **proc. 11008-6 INPI**, Brasil, 2010.

HOSOBUCHI, M.; KUROSAWA, K.; YOSHIKAWA, H. Application of computer to monitoring and control of fermentation process: Microbial conversion of ML-236B Na to pravastatin. **Biotechnol Bioeng.**, v. 42, p. 815–820, 1993.

IMADA, Y.; GOJI, N.; ISHIKAWA, H.; KISHIMA, M.; SEKIZAKI, T. Truncated surface protective antigen (SpaA) of *Erysipelothrix rhusiopathiae* serotype 1a elicits protection against challenge with serotypes 1a and 2b in pigs. **Infection and Immunity**, v. 67, p. 4376-4382, 1999.

JESUS, C.D.; GIORDANO, R.C.; CRUZ, A.J.G.; ALLIET-GAUBERT, M.; JOULIA, X.; ROUZINEAU, D.; ALBERT, J.; COUFORT, C.; ROUX, G.C.L. Weblab in chemical engineering between France and Brazil: validation of the methodology. **International Conference on Engineering Education – ICEE**, 3-7 set. 2007.



JUNG, Y.H.; LEE, J.D.; CHANG, Y.K. Development of a new device for on-line measurement of high cell concentrations. **Biotech. Tech.**, v. 13, n. 7, p. 459-462, 1999.

KARAKUZU, C.; TÜRKER, M.; OZTURK, S. Modelling, on-line state estimation and fuzzy control of production scale fed-batch baker's yeast fermentation. **Control Engineering Practice**, v. 14, p. 959-974, 2006.

KILIKIAN B.V.; SUAREZ, I.D.; LIRIA, C.W.; GOMBERT, A.K. Process strategies to improve heterologous protein production in *E. coli* under lactose or IPTG induction. **Process Biochemistry**, v. 35, p. 1019-1025, 2000.

KIM B.S.; LEE, S.C.; LEE, S.Y.; CHANG, Y.K.; CHANG, H.N. High cell density fed-batch cultivation of *Escherichia coli* using exponential feeding combined with pH-stat. **Bioprocess Biosyst. Eng.**, v. 26, p. 147-150, 2004.

KIM, Y.H.; PARK, J.S.; JUNG, H.I. An impedimetric biosensor for real-time monitoring of bacterial growth in a microbial fermentor. **Sensors and Actuators B**, v. 138, p. 270-277, 2009.

KITSUTA, Y.; KISHIMOTO, M. Fuzzy supervisory control of glutamic acid production. **Biotechnol. Bioeng.**, v. 44, p. 87-94, 1994.

KIVIHARJU, K.; SALONEN, K.; MOILANEN, U.; EERIKAINEN, T. Biomass measurement online: the performance of in situ measurements and software sensors. **Journal of Industrial Microbiology & Biotechnology**, v. 35, n. 7, p. 657-665, 2008.

KIVIHARJU, K.; SALONEN, K.; MOILANEN, U.; MESKANEN, E.; LEISOLA, M.; EERIKÄINEN, T. On-line biomass measurements in bioreactor cultivations: comparison study of two on-line probes. **J. Ind. Microbiol. Biotechnol.**, v. 34, p. 561-566, 2007.

KLEMAN, G.L.; CHALMERS, J.J.; LULI, G.W.; STROHL, W.R. A predictive and feedback control algorithm maintains a constant glucose concentration in fed-batch fermentations. **Appl. Environ. Microbiol.**, v. 57, n. 4, p. 910-917, 1991.

KONSTANTINOV, K.; KISHIMOTO, M.; SEKI, T.; YOSHIDA, T. A balanced DO-stat and its application to the control of acetic acid excretion by recombinant *Escherichia coli*. **Biotechnology and Bioengineering**, v.36, 1990.

KONSTANTINOV, B.K.; NISHIO, N.; YOSHIDA, T. Glucose feeding strategy accounting for the decreasing oxidative capacity of recombinant *Escherichia coli* in fed-batch cultivation for phenylalanine production. **J. Ferm. Bioeng.**, v. 70, n. 4, p. 253-60, 1990-b.

KORNELISSE, R.F.; WSTERBECK, C.M.L.; SPOOR, A.P.; SPANGAARD, H.J. Pneumococcal meningitis in children : prognostic indicators and outcome. **Clin. Infect. Dis.**, v. 21, p. 1390-1397, 1995.

KORZ, D.J.; RINAS, U.; HELLMUTH, K.; SANDERS, E.A.; DECKWER, W.D. Simple fed batch technique for high cell density cultivation of *Escherichia coli*. **Journal of Biotechnology**, v. 39, p. 59-65, 1995.

LEE, C.; LIM, H. New Device for continuously monitoring the optical density of concentrated microbial cultures. **Biotech. Bioeng.**, v. 22, 639-642, 1980.

LEE, S.Y. High cell density culture of *Escherichia coli*. **TIBTECHT** 14, p. 98-105, 1996.

LEE, J.; LEE, S.Y.; PARK, S.; MIDDELBERG, A.P.J. Control of fed-batch fermentations. **Biotechnology Advances**, v. 17, p. 29-48, 1999.

LEHNINGER, A.L. **Princípios de Bioquímica**. Forth Ediction. Sarvier, 2006.

LIN, W.R.; MISTRY, F.R.; NARAYANASWAMY, A.; SORENSEN, E.T.; KIM, W.S.; PERRIN, D. Novel feeding processes for fermentation. **United States Patent Application 20050202525**, 2005.

LIU, Y.C.; WU, W.T.; TSAO, J.H. Fed-batch culture for L-lysine production via on-line state estimation and control. **Bioproc. Eng.**, v. 9, n. 4, p. 135-139, 1993.

MARKX, G.H.; DAVEY, C.L. The dielectric properties of biological cells at radiofrequencies: Applications in biotechnology. **Enzyme and Microbial Technology**, v. 25, p. 161-171, 1999.

MASSIMO, C.D.; WILLIS, M.J.; MONTAGUE, G.A.; THAM, M.T.; MORRIS, A.J. Bioprocess model building using artificial neural networks. **Bioproc. Eng.**, v. 7, p. 77-82, 1991.

MAS, S., OSSARD, F., GHOMMIDH, C. On-line determination of flocculating *Saccharomyces cerevisiae* concentration and growth rate using a capacitance probe. **Biotechnology Letters**, v. 23, p. 1125-1129, 2001.

MATANGUIHAN, R.M.; KONSTANTINOV, K.B.; YOSHIDA, T. Dielectric measurement to monitor the growth and the physiological states of biological cells. **Bioprocess Engineering**, v. 11, p. 213-222., 1994.

MÉSZÁROS, A.; ANDRÁSIK, A.; MIZSEY, P.; FONYÓ, Z. Computer controlo f pH and DO in laboratory fermenter using neural network technique. **Bioprocess Biosyst. Eng.**, v. 26, p. 331-340, 2004.

MONTANO, I.D.C.; SUAREZ, C.A.G.; SILVA, A.J.; HORTA, A.C.L.; ARARIPE, F.; IEMMA, M.R.C.; GIORDANO, R.L.C.; GIORDANO, R.C. Cultivo de *Pichia pastoris* para Expressão Constitutiva de  $\alpha$ -Amilase de *Bacillus subtilis*. IX Simpósio de hidrólise Enzimática de Biomassas, Maringá, 2009.

NEIDHARDT, F.C.; UMBARGER, H.E. ***Escherichia coli* and Salmonella**: cellular and molecular biology. Second edition. ASM press, Washington, D.C. 1996.

NELLES, O. **Nonlinear system identification**: from classical approaches to neural networks and fuzzy models. Springer, Germany, 2001.

NEVES, A.A., PEREIRA, D.A., VIEIRA, L.M., MENEZES, J.C. Real time monitoring biomass concentration in *Streptomyces clavuligerus* with industrial media using a capacitance probe. **Journal of Biotechnology**, v. 84, p. 45-52, 2000.

NIELSEN J; VILLADSEN, J; LIDÉN, G. **Bioreaction engineering principles**. Second edition. Kluwer Academic / Plenum Publishers. New York, 2002.

NOVAGEN. **pET system Manual**, 11th edition, p. 3, 2005.

NYTTLE, V.G.; CHIDAMBARAM, M. Fuzzy logic control of a fed-batch fermentor. **Bioproc Eng.**, v. 9, p. 115–118, 1993.

OH, G.; MOO-YOUNG, M.; CHISTI, Y. Automated fed-batch culture of recombinant *Saccharomyces cerevisiae* based on on-line monitored maximum substrate uptake rate. **Biochem. Eng. J.**, v. 1, p. 211–217, 1998.

OLSSON, L.; NIELSEN, J. On-line and in situ monitoring of biomass in submerged cultivations. **Trends Biotechnol.**, v. 15, p. 517-522, 1997.

REIS, G.B. Simulação e controle do processo de produção de levedura. Dissertação de mestrado apresentada ao programa de Engenharia Química da Universidade Federal de São Carlos, 2009.

REIS, G.B.; HORTA, A.C.L.; ZANGIROLAMI, T.C.; GIORDANO, R.C.; CRUZ, A.J.G. Control of fed-batch yeast cultivation using a capacitance sensor. **Anais do 10th International Symposium on Process Systems Engineering: Part A. Computer Aided Chemical Engineering**, v. 27, p.1323-1328, 2009.

RIESENBERG, D. High-cell-density cultivation of *Escherichia coli*. **Current Opinion in Biotechnology**, v. 2, p. 380-384, 1991.

SARGO, C.R. Aperfeiçoamento de Cultivos de Alta Densidade Celular de *rE.coli* Utilizando Glicerol Como Fonte de Carbono. Dissertação de mestrado apresentada ao programa de Engenharia Química da Universidade Federal de São Carlos, 2011.

SHILOACH, J.; FASS, R. Growing *E. coli* to high cell density- A historical perspective on method development. **Biotechnology Advances**, v. 23, p. 345-357, 2005.

SHI, Z.; SHIMIZU, K. Neuro-fuzzy control of bioreactor systems with pattern recognition. **J. Ferm. Bioeng.**, v. 74, n. 1, p. 39–45, 1992.

SILVA, A.J., Vacinas recombinantes contra erisipela suína: desenvolvimento integrado de bioprocesso, da biologia molecular ao biorreator. **Tese de doutorado** apresentada ao programa de Biotecnologia da Universidade Federal de São Carlos, 2011.

SILVA, A.J.; HORTA, A.C.L.; NOVO, M.T.M.; GIORDANO, R.C.; ZANGIROLAMI, T. C. Cultivos de *E. coli* recombinante em batelada alimentada para produção do antígeno rSpaA da vacina de subunidade contra erisipela suína. **Anais do XVII Simpósio Nacional de Bioprocessos SINAFERM 2009**, Natal/RN. Mídia digital. 2009.

SILVA, A.J.; HORTA, A.C.L.; NOVO, M.T.M.; ZANGIROLAMI, T. C.; GIORDANO, R.C. Estudo de Diferentes Estratégias de Indução para Produção do Antígeno rSpaA de *E. rhusiopathiae* em Cultivos de *E. coli* Recombinante. XVIII Congresso Brasileiro de Engenharia Química, Foz do Iguaçu-PR, 2010.

SUAREZ, C.A.G. Otimização do cultivo de *Bacillus megaterium* recombinante em bateladas alimentadas. Dissertação de mestrado apresentada ao programa de Engenharia Química da Universidade Federal de São Carlos, 2010.

SUÁREZ, D.C.; KILIKIAN, B.V. Acid acetic accumulation in aerobic growth of recombinant *E. coli*. **Process Biochemistry**, v. 35, p. 1051-1055, 2000.

THOLUDUR, A.; RAMIREZ, W.F. Optimization of fed-batch bioreactors using neural network parameter function models. **Biotechnol. Prog.**, v. 12, p. 302–309, 1996.

VELEZ, A.M.E. Estudo da expressão de lipase BTL2 de *Bacillus thermocatenulatus* em *E. coli* recombinante. Dissertação de mestrado apresentada ao programa de Engenharia Química da Universidade Federal de São Carlos, 2009.

VELEZ, *et al.*, Influência da temperatura e da concentração de indutores na expressão de penicilina G Acilase (PGA) por *E. coli* recombinante. **Anais do COBEQ**, 2010, Foz do Iguaçu. 2010.

VOJINOVIC, V.; CABRAL, J.M.S.; FONSECA, L.P. Real-time bioprocess monitoring. Part I: In situ sensors. **Sensors and Actuators B** v. 114, p. 1083-1091. 2006.

WIPF, B.; WIEBEL, E.K.; VOGEL, S. Computer controlled large scale production of a-interferon by *E. coli*. **Bioproc. Eng.**, v. 10, p. 145–53, 1994.

XIONG, Z.Q.; GUO, M.J.; GOU, Y.X.; CHU, J.; ZHUANG, Y.P; ZHANG, S.L. Real-Time Viable Cell Mass Monitoring in High-Cell-Density Fed-Batch Glutathione Fermentation by *Saccharomyces cerevisiae* T65 in Industrial Complex Medium. **Journal of Bioscience and Bioengineering**, v. 105, n. 4, p. 409-413, 2008.

YAMANE, T.; HIBINO, W.; ISHIHARA, K.; KADOTANI, Y.; KOMINAMI, M. Fed-batch culture automated by uses of continuously measured cell concentration and culture volume. **Biotechnol Bioeng.**, v. 39, p. 550–555, 1992.

YANG, L., RUAN, C., LI, Y. Detection of viable *Salmonella typhimurium* by impedance measurement of electrode capacitance and medium resistance. **Biosensors and Bioelectronics**, v. 19, p. 495-502, 2003.

YANO, T.; OKANO, K.; MASDUKI, A.; NISHIZAWA, Y. Measurement of Cell Density in Cultures of Aggregative Organisms by Continuous-Dilution-Photometric-Assay". **J. Ferm. Bioeng.**, v. 76, n. 2, p. 123-127, 1993.

YOTHER, J. CAPSULES. IN TUOMANEN, E.I., MITCHELL, T.J., MORRISON, D.A. AND SPRATT, B.G. **The Pneumococcus**. ASM Press, Washington, D.C., p. 30-48, 2004.

ZHANG, J.; GREASHAM, R. Chemically defined media for commercial fermentations. **Appl. Microbiol. Biotechnol.**, v. 51, p. 407-421, 1999.

ZHANG, X.C.; VISALA, A.; HALME, A.; LINKO, P. Functional state modeling and fuzzy control of fed-batch aerobic baker's yeast process. **J. Biotechnol.**, v. 37, p. 1-10, 1994.

ZHOU, W.; REHM, J.; HU, W.S. High viable cell concentration fed-batch cultures of hybridoma cells through online nutrient feeding. **Biotechnol. Bioeng.**, v. 46, p. 579-587, 1995.

## APÊNDICE A - RELATÓRIO ENCAMINHADO PARA REGISTRO DO PROGRAMA DE COMPUTADOR SUPERSYS\_HCDC

O software SUPERSYS\_HCDC, depois de registrado no Brasil com o número 11008-6, foi eleito como um dos 11 projetos para financiamento e criação de plano de negócio pelo programa Search SEBRAE Biominas Brasil, protocolo número 130.

### DADOS DO TITULAR

Nome da Instituição	CNPJ
Universidade Federal de São Carlos	45.358.058/0001-40

### DADOS DO(S) CRIADOR(ES)

<b>Nome civil completo:</b> Antonio Carlos Luperni Horta
<b>Nome civil completo:</b> Teresa Cristina Zangirolami
<b>Nome civil completo:</b> Roberto de Campos Giordano
<b>Nome civil completo:</b> Antonio José Gonçalves da Cruz
<b>Nome civil completo:</b> Giann Braune Reis
<b>Nome civil completo:</b> Charles Dayan Farias de Jesus

### INFORMAÇÕES SOBRE O PROGRAMA DE COMPUTADOR

#### 1. TÍTULO DO PROGRAMA

#### **SUPERSYS\_HCDC**

**Programa de computador para monitoramento e controle de biorreator para cultivos de alta densidade celular (Supervisory System for Bioreactor High Cell Density Cultivations).**

## 2. PALAVRAS-CHAVE RELACIONADAS AO PROGRAMA DE COMPUTADOR:

### Palavras-chave em língua portuguesa inglesa

1.	Sistema supervisor
2.	Cultivo de alta densidade celular
3.	Batelada alimentada
4.	Controle automático da vazão exponencial de alimentação
5.	Início automático da alimentação
6.	Inferência em tempo real de biomassa
7.	Inferência em tempo real da velocidade específica de crescimento
8.	<i>E. coli</i>

### Palavras-chave em língua

1.	Supervisory System
2.	High density cell cultivation
3.	Fed-batch
4.	Exponential feed flow rate automatic control
5.	Automatic feed startup
6.	Online biomass estimation
7.	Online specific growth rate estimation
8.	<i>E. coli</i>

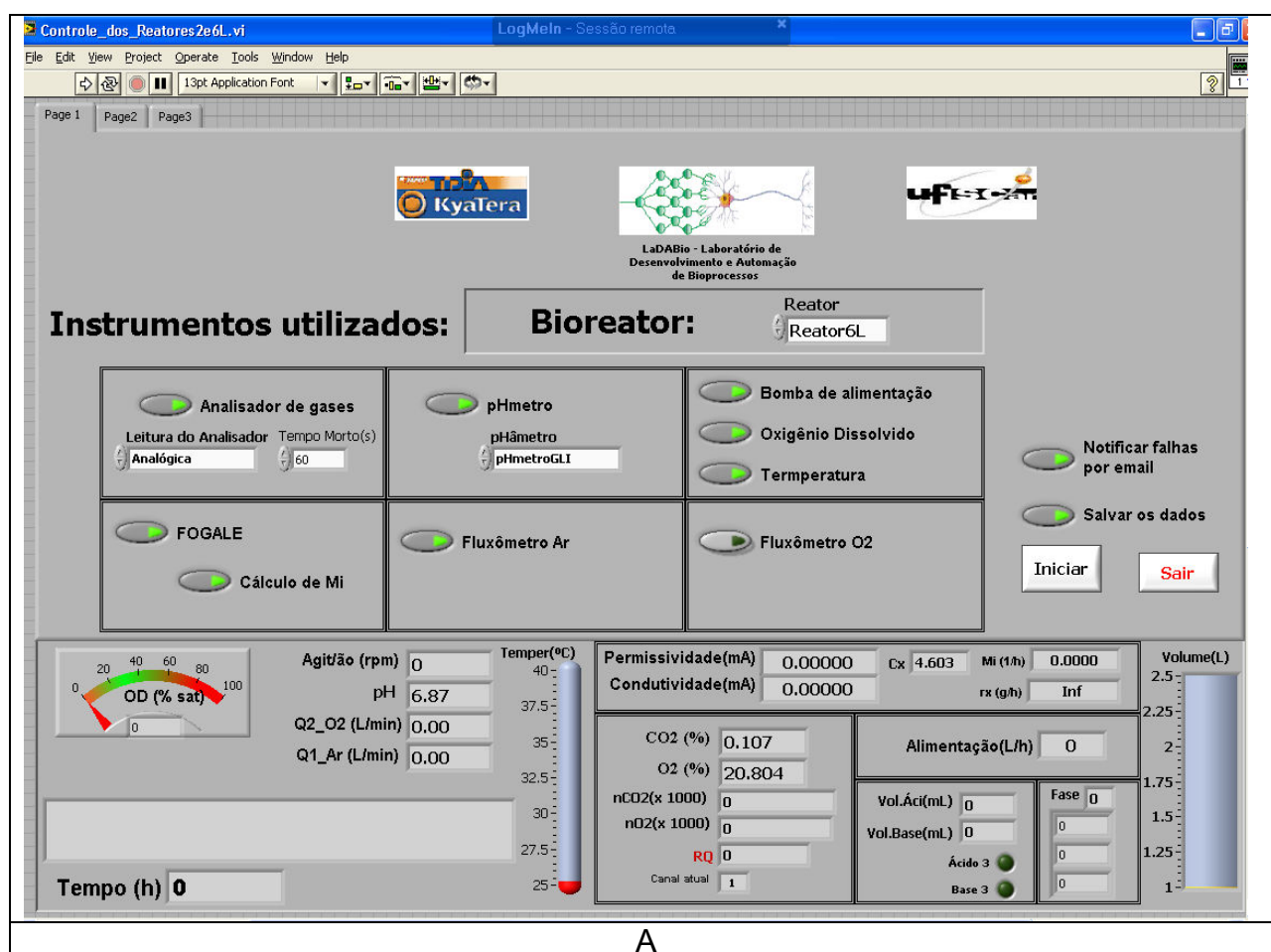
## 3. DESCRIÇÃO FUNCIONAL DO PROGRAMA DE COMPUTADOR

Cultivos em biorreator são fundamentais para a produção de inúmeros produtos da indústria farmacêutica (Liljeqvist e Stahl, 1999). Quaisquer que sejam os produtos de interesse, os cultivos são realizados com o objetivo inicial de atingir alta densidade celular para então proceder à indução, fazendo com que as células cultivadas produzam a proteína de interesse.

O modo de operação mais amplamente empregado nos cultivos de alta densidade celular (CADC) é a batelada alimentada com vazão exponencial para fornecimento do meio suplementar (Shiloach e Fass, 2005). Além disso, o microrganismo mais utilizado para a expressão de proteínas heterólogas é a *Escherichia coli* recombinante (Demain e Vaishnav, 2009). No entanto, durante a realização do CADC dificuldades operacionais são comumente encontradas tanto na fase voltada à formação de biomassa como na fase de expressão da proteína. Em ambas as fases é preciso evitar tanto a inibição do crescimento quanto da expressão da proteína de interesse devido ao acúmulo de substratos. Isso requer um controle preciso da vazão de alimentação do meio suplementar, adequando-a à demanda em cada momento do cultivo, a qual pode variar consideravelmente após a indução devido ao estresse metabólico associado à expressão das proteínas heterólogas.

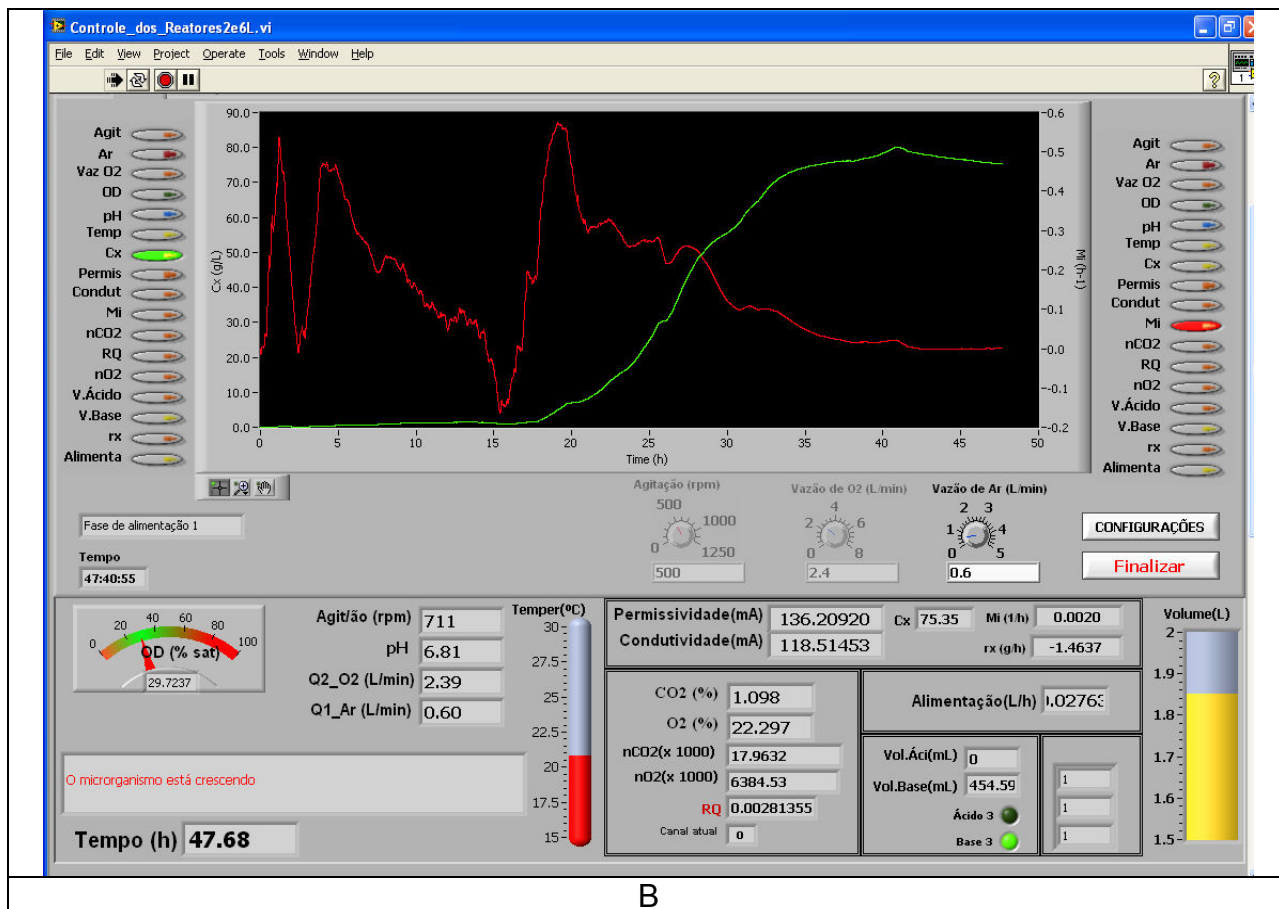
Este programa de computador foi desenvolvido inicialmente para realizar a aquisição de dados e alguns cálculos em cultivos em alta densidade de *E.coli* em biorreator. O programa evoluiu a ponto de exercer funções de supervisão e controle automaticamente, sendo capaz de conduzir uma batelada alimentada de *E. coli* recombinante até altas concentrações celulares, superiores à 100g/L.

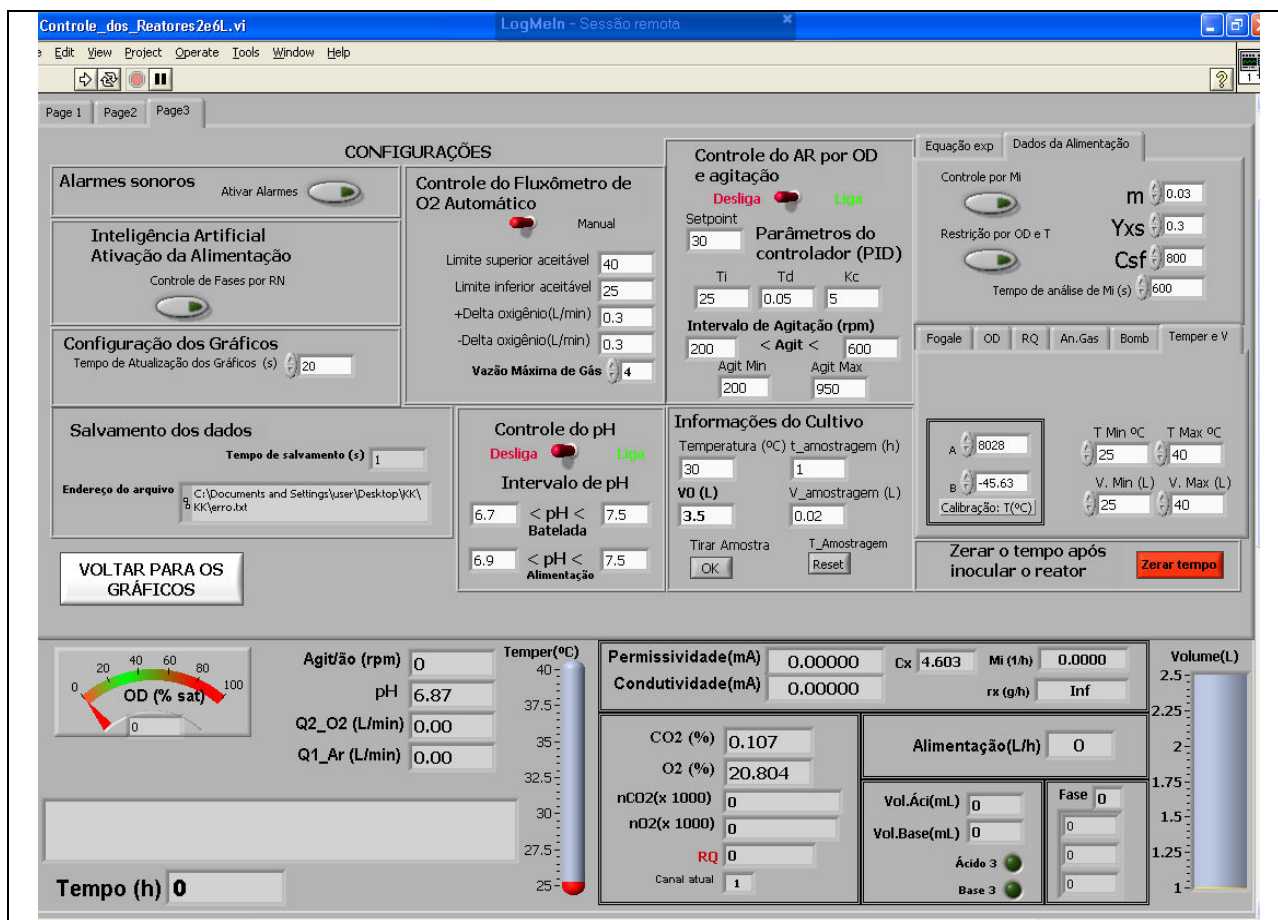
O programa deve ser instalado em um computador, juntamente com os equipamentos controlados ou de aquisição de dados. A Figura A-1 apresenta o programa de computador SUPERSYS\_HCDC.



A







C

Figura A-1. Telas principais do programa de computador SUPER SYS\_HCDC. A) tela inicial de escolha das configurações dos instrumentos a serem utilizados; B) tela de visualização dos gráficos e operação manual das vazões de gases e da agitação; C) tela de configurações e ativação dos controles automáticos.

A Figura A-2 apresenta os equipamentos de aquisição e monitoramento do programa: biorreator de 5,0 L (montagem própria na UFSCar) equipado com sensor de capacitância Biomass System (FOGALE Nanotech), transmissor de oxigênio dissolvido (eletrodo Metler Toledo, modelo CE, O2 4050); pHmetro GLI; sensor de temperatura PT100; 2 fluxômetros de massa (GFC AALBORG) para suprimento de ar e oxigênio puro; analisador de gases (Sick/MAIHAK S710); banho termostaticado; bombas de ácido e base (Masterflex) e bomba de alimentação (ISMATEC – BVP). Para o monitoramento em tempo real das variáveis pH, temperatura, concentração de oxigênio dissolvido, fração molar de oxigênio e gás carbônico no gás de saída e concentração de biomassa, foi implementada a aquisição em LabView 8.0, National Instruments). Todos

os equipamentos se comunicam com o computador via rede através do compact FieldPoint (cFP) 2020 (National Instruments), com exceção do banho térmico que apresenta comunicação via interface NOVUS.

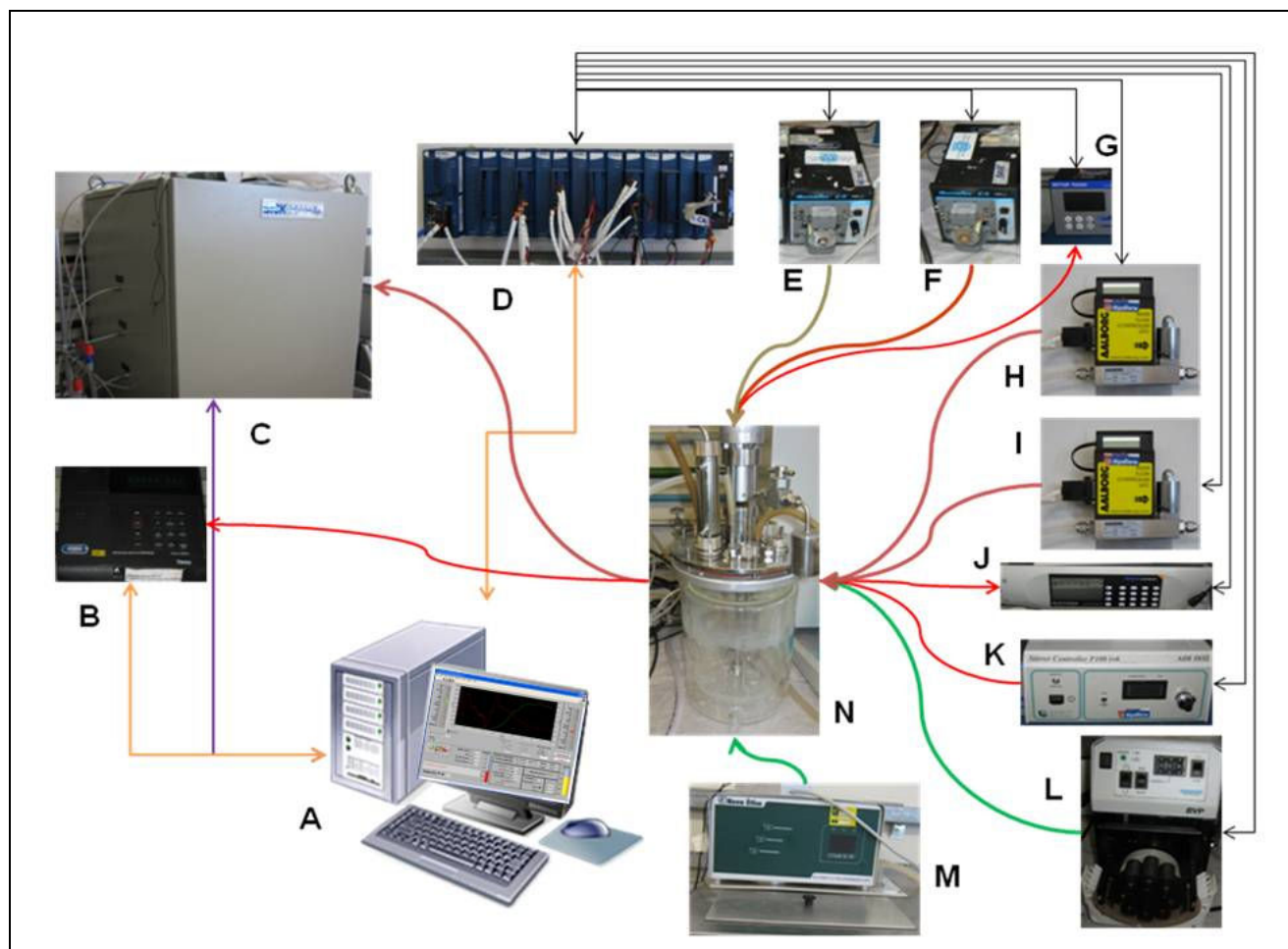


Figura A-2. Esquema dos instrumentos e sistema de monitoramento e controle dos cultivos. A: computador com o programa de aquisição e controle SUPERSYS\_HCDC; B: pHmetro; C: analisador de gases; D: cFP (Compact Field Point – National Instruments); E e F: bombas de ácido e Base; G: sensor de oxigênio dissolvido; H: fluxômetro de ar; I: fluxômetro de oxigênio; J: sensor de biomassa (Fogale nanotech); K: controlador da velocidade de agitação do reator; L: bomba de alimentação, M: banho termostatizado.

As medidas de capacitância (feitas com o sensor de biomassa – Fogale nanotech) são realizadas em duas frequências, sendo uma para medir as propriedades do meio, de 10 MHz e a outra para medir as propriedades do meio em conjunto com as células, que

neste caso é de 2984 KHz. O equipamento fornece como resultado as medidas de permissividade descontando a influência do meio. Além disso, o valor fornecido passa por um filtro de média móvel de 90 pontos para a minimização de ruídos. O programa lê os valores de permissividade e condutividade através da saída analógica do transmissor, sinal cujo valor varia de 4 a 20 mA. Este sinal é então multiplicado por  $10^4$  para ser processado pelo SUPERSYS\_HCDC.

Para utilizar o programa, o usuário deve preparar uma batelada alimentada. O programa possui vários controles automáticos com a opção do modo manual. Após conectar todas as mangueiras (alimentação e de base) o usuário deve ajustar as configurações necessárias no programa, antes de inocular o biorreator. Estas configurações são:

- 1) Equação de calibração da bomba de alimentação, que varia de acordo com a mangueira utilizada;
- 2) Equação de calibração da bomba de base, que varia de acordo com a mangueira utilizada e com as configurações da bomba;
- 3) Equação de calibração entre a permissividade e a concentração celular;
- 4) Escolha do setpoint de oxigênio dissolvido;
- 5) Escolha da temperatura, faixa de pH, tempo de retirada de amostra, volume médio de cada amostra retirada, bem como o intervalo de tempo de cada amostragem, e o volume inicial da batelada.
- 6) Intervalo em segundos e o nome do arquivo em que os dados serão salvos.

Após a calibração dos sensores de oxigênio dissolvido e pH o reator estará pronto para ser inoculado. Ao inocular o biorreator o programa deve ser “notificado” para reiniciar a contagem de tempo e zerar os gráficos. A partir deste momento o usuário pode ativar os controles automáticos para que o programa controle o cultivo, são eles:

- 1) Ativar alarmes sonoros;
- 2) Ativação automática da vazão de alimentação por sistema de inteligência artificial;
- 3) Controle automático das vazões de ar e oxigênio;
- 4) Controle automático da concentração de oxigênio dissolvido (COD);
- 5) Controle automático do pH;
- 6) Controle automático da vazão de alimentação através do  $\mu_{DIN}$ ;
- 7) Controle automático da vazão de alimentação com restrição por temperatura e COD.
- 8) Envio de e-mails automáticos alertando para falhas no sistema.

### 3.1 Detalhando os sistemas automáticos: Ativar alarmes sonoros;

Os alarmes sonoros são alertas gravados em voz que o sistema ativa para cada problema detectado. Estão programados em uma estrutura orientada a evento, são eles:

- i) Mal funcionamento da bomba de ácido: acionado se o pH fica 5% acima do limite superior de setpoint;
- ii) Mal funcionamento da bomba de base: acionado se o pH fica em menos de 95% do limite inferior de setpoint;
- iii) Diminua a temperatura do banho: acionado se a temperatura do cultivo ficar 5% acima do setpoint;
- iv) Aumente a temperatura do banho: acionado se a temperatura do cultivo ficar em menos de 95% do setpoint;
- v) Aumente a vazão de gases e/ou a agitação: acionado se a COD fica a menos de 50% do setpoint;
- vi) Aumente a pressão da linha de oxigênio: acionado se a vazão linha estiver em menos de 95% do setpoint;
- vii) Aumente a Pressão da linha de Ar: acionado se a vazão lida estiver em menos de 95% do setpoint;
- viii) O microrganismo está morrendo: acionado se  $\mu$  for negativo;
- ix) O microrganismo voltou a crescer: acionado se  $\mu$  for positivo;

### 3.2 Detalhando os sistemas automáticos: Controle de fases por Sistema de inteligência artificial;

O controle do acionamento automático da bomba de alimentação está vinculado ao reconhecimento da fase do cultivo, fase de batelada ou de alimentação. Este reconhecimento é feito por três algoritmos diferentes:

- i) Comitê de redes neurais;
- ii) Lógica Fuzzy;
- iii) Derivada do CO<sub>2</sub>.

Foi observado que ao esgotar a fonte de carbono em cultivos com meio definido, ocorre um súbito aumento na concentração de oxigênio dissolvido, seguido por uma diminuição na velocidade de agitação e uma queda na vazão molar de CO<sub>2</sub> produzido no reator. Para que o programa de computador pudesse assimilar este conhecimento e utilizá-lo para identificar o momento de iniciar a alimentação, foram incluídas três técnicas diferentes, de forma redundante, para evitar falsos positivos. Estes falsos positivos podem ocorrer pois os sinais medidos *online* podem apresentar ruídos e oscilações.

A teoria de redes neurais artificiais (RN) foi escolhida como algoritmo principal de reconhecimento de fases, no modo de comitê de especialistas. A vantagem de se utilizar um comitê é a de poder agregar uma nova RN ao comitê após a realização de um novo cultivo, sem a necessidade de re-treinar uma rede com os dados de todos os cultivos anteriores. A arquitetura de RN adotada está descrita na Figura A-3.

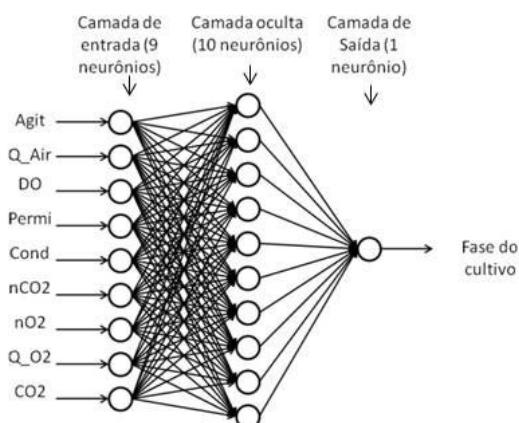


Figura A-3. Arquitetura da RN utilizada para reconhecimento das fases do cultivo.

Foram realizados 3 cultivos em batelada alimentada para aquisição de dados e treinamento das RN. Cada cultivo realizado foi utilizado para treinar uma RN. A resposta do comitê foi feita pela média da saída de classificação de cada RN.

Para contribuir com o comitê de RN na tarefa de reconhecimento das fases do cultivo, foi utilizada a lógica nebulosa (Fuzzy). A vantagem deste algoritmo é que ele dispensa o treinamento, requerido pelas redes neurais. A lógica fuzzy, por outro lado, deve ser configurada com os conhecimentos do especialista do processo. A Figura A-4 apresenta as funções de pertinência ajustadas com o conhecimento adquirido. TO corresponde ao tempo ótimo, que ocorre entre 13 e 20 horas de cultivo, TC ao tempo cedo, TT ao tempo tarde; CM corresponde à fração molar média de CO<sub>2</sub>, região mais provável de ocorrer o fim das fontes de carbono da batelada, CB à fração baixa e CA à fração alta; DCN corresponde à derivada negativa da fração de CO<sub>2</sub>, valor provável de ocorrência do fim da batelada, e DCZP à derivada positiva; NP identifica a fase de batelada e PC identifica o início da alimentação.

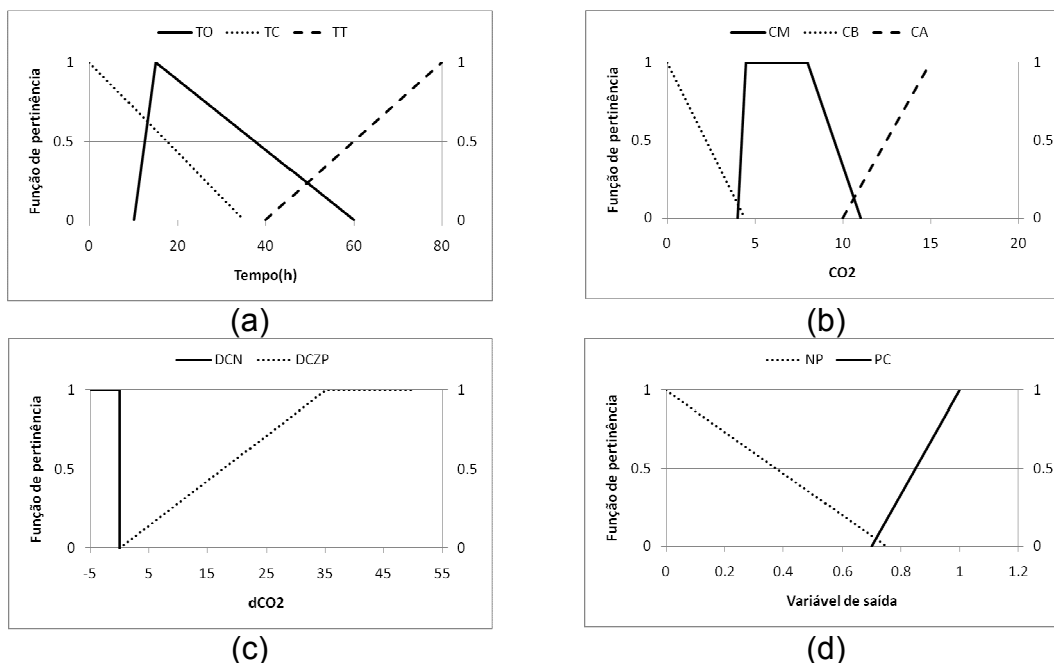


Figura A-4. Funções de pertinência ajustadas. (a) tempo, (b) fração molar de CO<sub>2</sub>, (c) derivada da fração molar de CO<sub>2</sub>, e (d) Saída.

Além destes dois métodos de inteligência artificial, adotou-se no programa uma lógica pura de decisão: Se a derivada da fração molar de CO<sub>2</sub> for menor que -2 aciona a bomba de alimentação (fim da batelada e início da fase de alimentação).

A decisão final em ativar ou não a bomba de alimentação foi realizada através da média entre os três métodos utilizados: Se média das saídas do comitê, do fuzzy e da lógica for maior que 0,6, a bomba de alimentação é acionada, marcando o início da fase de alimentação. Estes algoritmos foram implementados em MatLab 6.5. O acionamento da bomba ocorre quando o tempo do cultivo se torna superior à variável “t<sub>0</sub> alimentação”. O sistema de IA modifica a variável “t<sub>0</sub> alimentação” para um valor de tempo menor ou igual ao tempo do cultivo ao reconhecer o fim da fonte de carbono da batelada.

### 3.3 Detalhando os sistemas automáticos: Controle automático das vazões de ar e de oxigênio;

O controle automático das vazões de ar e oxigênio no gás de entrada foi implementado por árvore de decisão e está apresentado no Quadro A-1, no qual  $Agit_{Média}$  corresponde à agitação média,  $LS_{agit}$  ao limite superior da agitação,  $LI_{agit}$  ao limite inferior da agitação,  $LS_{OD}$  ao limite superior do OD,  $LI_{OD}$  ao limite inferior do OD,  $QO_{2\text{anterior}}$  à

vazão de oxigênio do passo anterior, QAR à vazão de ar,  $\Delta O_2^-$  ao delta negativo de oxigênio,  $\Delta O_2^+$  ao delta positivo de oxigênio.

Quadro A-1. Pseudocódigo do controle automático das vazões de ar e oxigênio.

<p><b>Procedimento:</b> Controle enriquecimento de oxigênio</p> <p><b>Início</b></p> <p><b>Se</b> <math>Agit_{Média} &lt; 0,7 \cdot LS_{agit}</math> e <math>QO2_{anterior} &gt; 0</math></p> <p>    <math>QAR = QAR + \Delta O_2^-</math></p> <p>    <math>QO2 = QO2_{anterior} - \Delta O_2^-</math></p> <p><b>Se</b> <math>Agit_{Média} \geq 0,99 \cdot LS_{agit}</math> e <math>OD &lt; LI_{OD}</math></p> <p>    <b>Se</b> <math>QAR \geq (Q_{MAX} - QO2_{anterior})</math></p> <p>        <math>QAR = QAR - \Delta O_2^+</math></p> <p>        <math>QO2 = QO2 + \Delta O_2^+</math></p> <p>    <b>Senão</b></p> <p>        <math>QAR = \Delta O_2^+ + QAR</math></p> <p>        <math>QO2 = QO2_{anterior}</math></p> <p><b>Fim</b></p>
--

Ao ativar o controle automático o sistema passa a aumentar ou diminuir as vazões de ar e oxigênio gradativamente, de acordo com a necessidade e as configurações fornecidas pelo usuário.

### 3.4 Detalhando os sistemas automáticos: Controle automático da concentração de oxigênio dissolvido (COD);

O controle automático responsável por manter as concentrações de OD é realizado pela equação A-1, na qual “Agit” corresponde à velocidade de agitação, “t” ao tempo (s), “Kc” à constante proporcional, “SetP” ao setpoint de OD (oxigênio dissolvido), “Ti” à constante integral e “Td” à constante derivativa.

$$Agit(t) = Agit(t-1) + Kc \cdot \left\{ \begin{array}{l} [(SetP - OD(t)) - (SetP - OD(t-1))] + \\ + \frac{\Delta t}{Ti} \cdot (SetP - OD(t)) + \frac{Td}{\Delta t} \cdot \left( \begin{array}{l} (SetP - OD(t)) + \\ - 2 \cdot (SetP - OD(t-1)) + \\ + (SetP - OD(t-2)) \end{array} \right) \end{array} \right\} \quad A-1$$

### 3.5 Detalhando os sistemas automáticos: Controle automático do pH;

O controle automático de pH é feito através de uma árvore de decisão simples: Se pH lido for maior que o limite superior aceitável aciona a bomba de ácido; se o pH lido for menor que o Limite inferior aceitável aciona a bomba de base.

### 3.6 Detalhando os sistemas automáticos: Controle automático da vazão de alimentação através do $\mu_{DIN}$ ;

Este controle foi implementado a partir da velocidade específica de crescimento calculada *online* utilizando o sinal de permissividade fornecido pelo sensor de biomassa



(Fogale nanotech). Para possibilitar o uso do sinal de permissividade para cálculo do  $\mu_{DIN}$  foi necessário o uso de dois filtros de média móvel, o primeiro alisando os dados da permissividade e o segundo os dados de  $\mu$ .

O Filtro de média móvel utilizado para alisar o sinal de permissividade e calcular a velocidade específica de crescimento pode ser observado no pseudocódigo do Quadro A-2, cuja implementação foi feita em MatLab. No pseudocódigo  $n$  representa o tamanho do vetor (neste caso, 100),  $t$  o tempo (h),  $Cx$  a concentração celular medida *online* pelo sensor de permissividade,  $CX$  o vetor de tamanho  $n$  de  $Cx$ ,  $Mi$  a velocidade específica de crescimento,  $Cxm$  a concentração celular média dos últimos  $n$  pontos;  $T$  o vetor de tempo,  $V$  o volume (L),  $F$  a vazão de alimentação (L/h).

Quadro A-2. Pseudocódigo do filtro de média móvel e cálculo da velocidade específica de crescimento.

**Procedimento:** Filtro de Média Móvel

**Início**

**Se**  $t=0$

$CX(1:n)=Cx$

$Mi(1:n)=0$

$T(1:n)=t$

$F(1:n)=0$ ;

$V(1:n)=V0$ ;

**Se**  $t < 0,03$

$CX(n)=Cx$

$Cxm=Média(CX)$

$CX(n)=Cxm$

$Mi(n)=0$

$T(n)=t$

**Senão**

$CX(n)=Cx$

$Cxm=Média(CX)$

$CX(n)=Cxm$

$dcx=CX(n)-CX(1)$

$CX(1:n-1)=CX(2:n)$

$T(n)=t$

$dt=T(n)-T(1)$

$F(n)=f$

$V(n)=v$

**Se**  $dt=0$

$dcxdt=0$

**Senão**  $dcxdt=dcx/dt$

**Se**  $Cxm=0$

$Mi(n)=0$

**Senão**  $Mi(n)=dcxdt/Cxm+F(n)/V(n)$

$Mim=Média(Mi)$

$Mi(n)=Mim$

$Mi(1:n-1)=Mi(2:n)$

$F(1:n-1)=F(2:n)$

$V(1:n-1)=V(2:n)$

$T(1:n-1)=T(2:n)$

**Fim**

O cálculo do volume do instantâneo do cultivo foi implementado em MatLab e pode ser observado no Quadro A-3, considerando as retiradas de amostras e a alimentação. C1 e C2 representam constantes da equação exponencial de alimentação.

Quadro A-3. Pseudocódigo do cálculo do volume do reator.

<p><b>Procedimento:</b> Volume</p> <p><b>Início</b></p> <p><b>Se</b> t=0  V(1:n)=V0;  T(n)=t  V(n)=v</p> <p><b>Se</b> amostragem=0  Vretirado=0</p> <p><b>Senão</b> Vretirado =Vamostra</p> <p><b>Se</b> t&lt;Tbatelada  V(n)=V(n-1)-Vretirado;</p> <p><b>Ou se</b> t≥Tbatelada  <b>Se</b> C2=0  V(n)=V(n-1)+C1*[T(n)-T(n-1)]-Vretirado;  <b>Senão</b> V(n)= V(n-1)+(C1/C2)*{exp[C2*(T(n)-Tbatelada)]-exp[C2*(T(n-1)-Tbatelada)]}-Vretirado</p> <p>V(1:n-1)=V(2:n)  T(1:n-1)=T(2:n)</p> <p><b>Fim</b></p>
---

O valor calculado de  $\mu$  foi então utilizado para alimentar automaticamente, a cada 10 minutos, o  $\mu_{DIN}$ , e então recalculer a vazão de alimentação, equação A-2; na qual F representa a vazão de alimentação de meio fresco (L/h),  $\mu_{DIN}$  a velocidade específica de crescimento adotada ( $h^{-1}$ ),  $Y_{XS}$  o coeficiente de rendimento ( $g_X/g_S$ ), m a manutenção ( $g_X/g_S \cdot h$ ),  $C_{X0}$  a concentração celular no início da alimentação (g/L),  $V_0$  o volume no início da alimentação (L), t o tempo (h),  $C_{S0}$  a concentração de substrato no meio de alimentação (g/L),  $C_{SR}$  a concentração de substrato residual (g/L).

$$F = \left( \frac{\mu_{DIN}}{Y_{XS}} + m \right) \cdot \frac{C_{X0} \cdot V_0}{C_{S0} - C_{SR}} \cdot e^{(\mu_{DIN} \cdot t)} \quad A-2$$

### 3.7 Detalhando os sistemas automáticos: Controle automático da vazão de alimentação com restrição por temperatura e COD.

Ao utilizar o sistema de alimentação automático descrito anteriormente, com o  $\mu_{DIN}$ , o crescimento celular pode ser tão intenso que supera a capacidade do biorreator em suprir oxigênio dissolvido. Neste caso, a estratégia deixa de ser aumentar a velocidade de crescimento, e passa a ser manter o crescimento sob controle, desligando a alimentação automaticamente, sempre que o consumo de oxigênio superar a reposição. Além de restringir a alimentação em função da deficiência em manter a concentração

de oxigênio dissolvido, a mesma lógica foi adotada para a restrição da temperatura, pois em altas concentrações celulares o biorreator chega a emitir até 1000W, aquecendo muito além da capacidade do banho em remover calor.

### **3.8 Detalhando os sistemas automáticos: Envio de e-mails automáticos indicativos de falhas no sistema.**

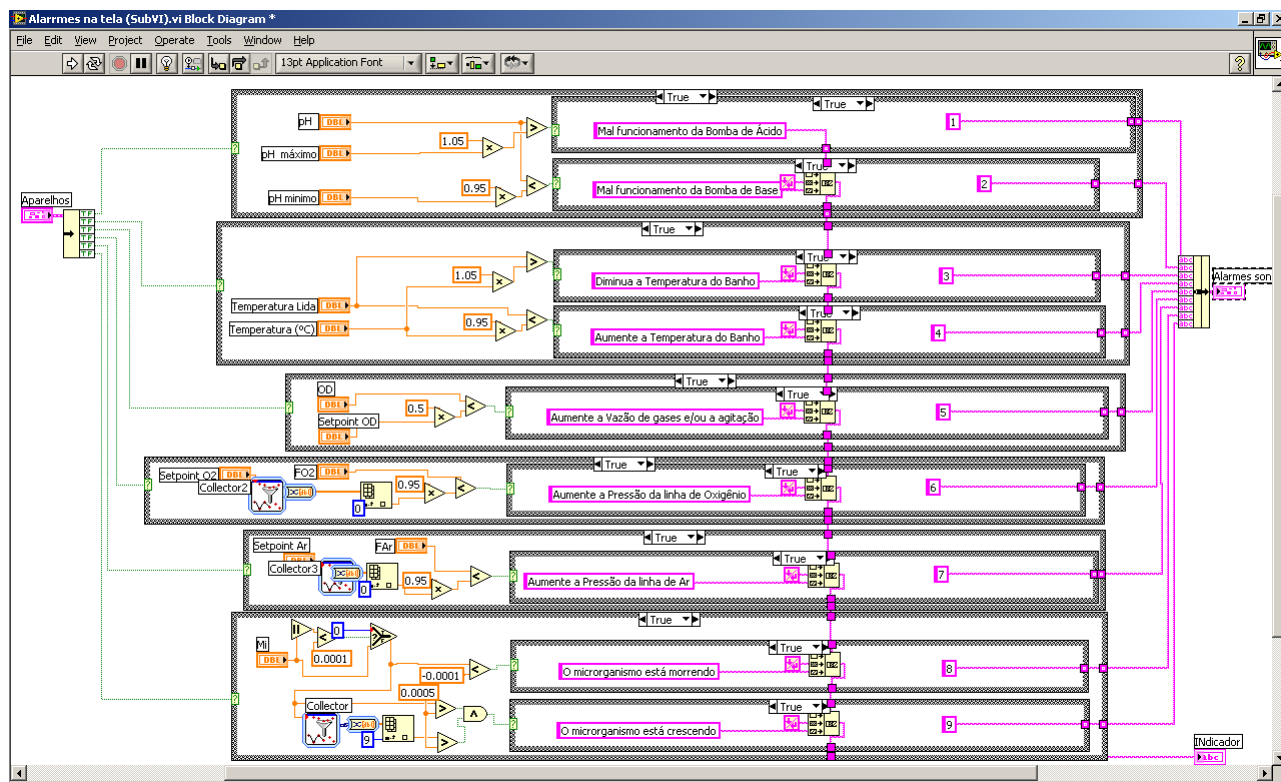
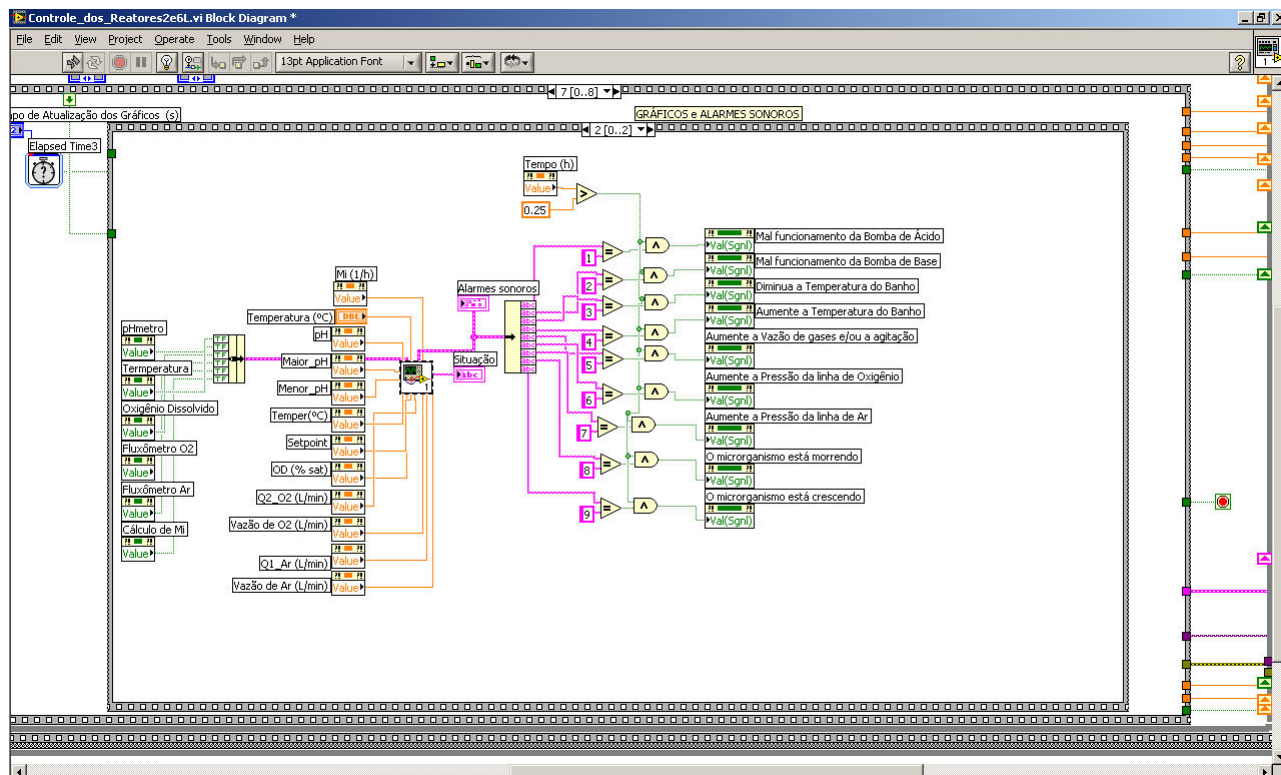
Os alertas de falhas por e-mail são gerados sempre que houver um alerta sonoro de falha. Ao detectar uma falha o sistema sinaliza na tela do computador uma mensagem indicando a falha, reproduz uma gravação de voz com a respectiva falha e envia um e-mail ao operador do programa relatando a ocorrência da falha. As operadoras de celular podem ser programadas para alertar via SMS (mensagem para celular) sempre que um email de falha for detectado na caixa de entrada do operador; desta forma, o programa de computador acaba alertando as falhas por quatro modos distintos: i) mensagens de falhas na tela do computador; ii) mensagens sonoras; iii) mensagens via e-mail; iv) mensagens para celular.

## **4. CÓDIGO FONTE**

(Apresente o código fonte completo. Caso o código fonte supere três páginas, por favor, anexar arquivo com código fonte completo, indicando os trechos que caracterizem a autenticidade/originalidade de criação, sem que possam permanecer dúvidas a esse respeito.)

A maior parte do código fonte está em anexo, porém as partes que caracterizam inovação estão apresentadas a seguir.

## 4.1 Alarmes sonoros;



## 4.2 Controle de fases por Sistema de inteligência artificial;

```
function fcn(n)
V(1:n-1)=V(2:n);
rx(1:n-1)=rx(2:n);
T(1:n-1)=T(2:n);

path(path, pa);
if islgada==1
classificar1=[Agit Far OD permis Conduz nCO2 nO2 FO2 CO2];
load Net1fases.mat;
net1=BestNet;
clasrede1 = sim(net1,classificar1');
Fase1n=clasrede1;

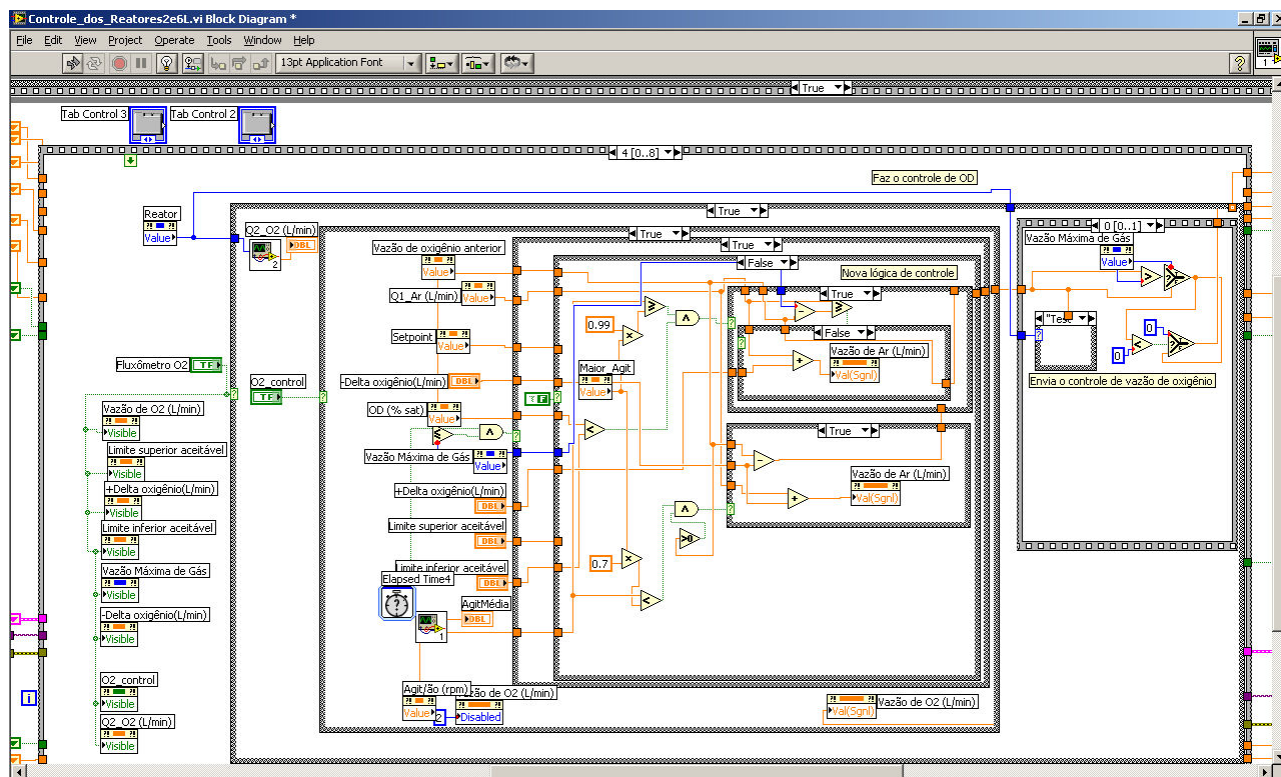
load Net2fases.mat;
net2=BestNet;
clasrede2 = sim(net2,classificar1');
Fase2n=clasrede2;

load Net3fases.mat;
net3=BestNet;
clasrede3 = sim(net3,classificar1');
Fase3n=clasrede3;
Fase1=(Fase1n+Fase2n+Fase3n)/3;

VetCO2(10)=CO2;
dCO2=VetCO2(10)-VetCO2(1);
VetCO2(1:9)=VetCO2(2:10);
if dCO2<-1
Fase2=1;
else
Fase2=0;
end
fis = readfis('fuzzy_fase2');
Fase3 = evalfis([ CO2 dCO2], fis);
Fase=(Fase2+Fase3+Fase1)/3;

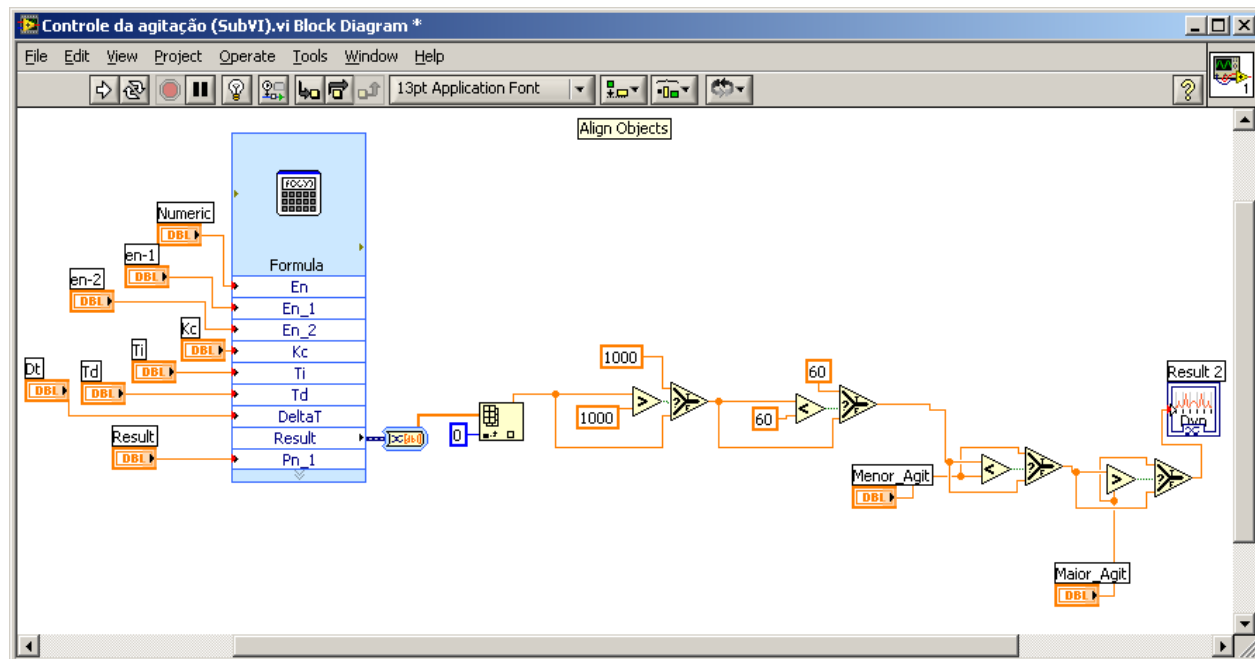
else
Fase2=1;
Fase3=1;
Fase1=1;
Fase=(Fase2+Fase3+Fase1)/3;
end
end
```

### 4.3 Controle automático das vazões de ar e oxigênio;

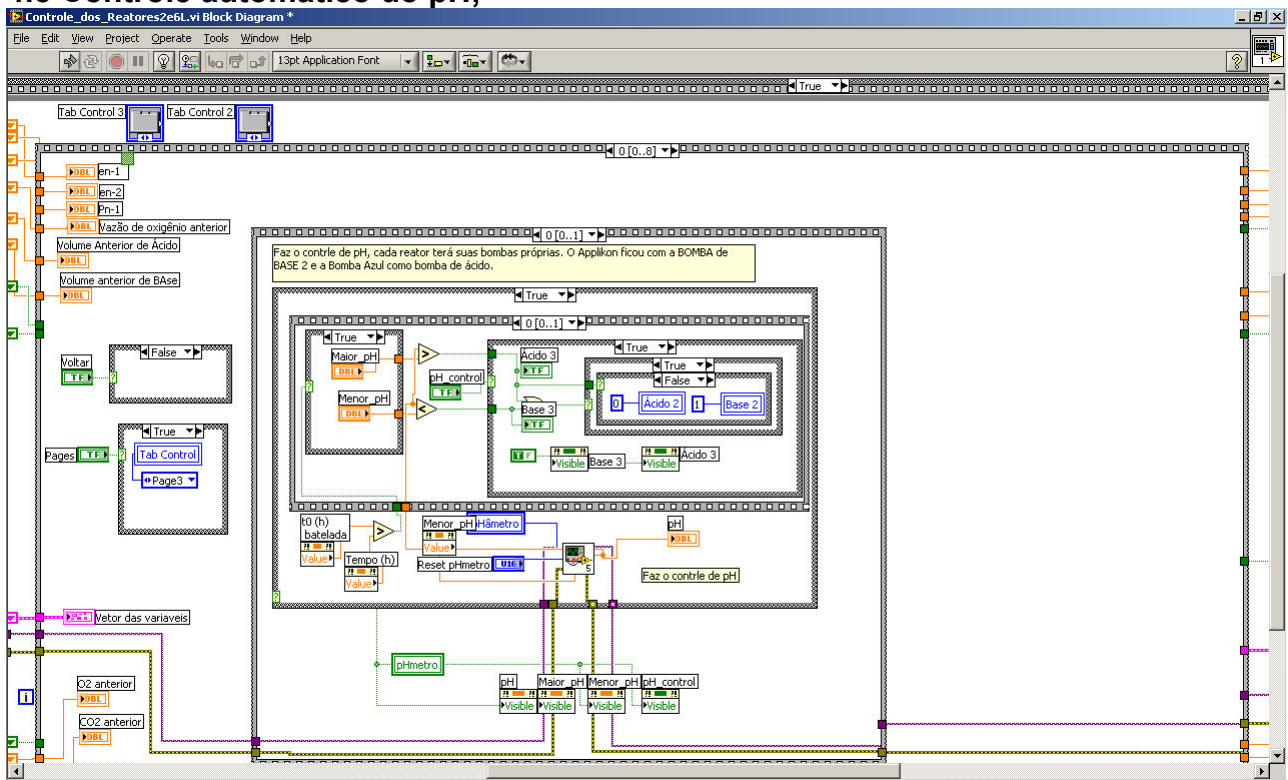


#### 4.4 Controle automático da concentração de oxigênio dissolvido (OD);

Formula= $Pn\_1 + Kc * ((En - En\_1) + \Delta T / Ti * En + Td / \Delta T * (En - 2 * En\_1 + En\_2))$

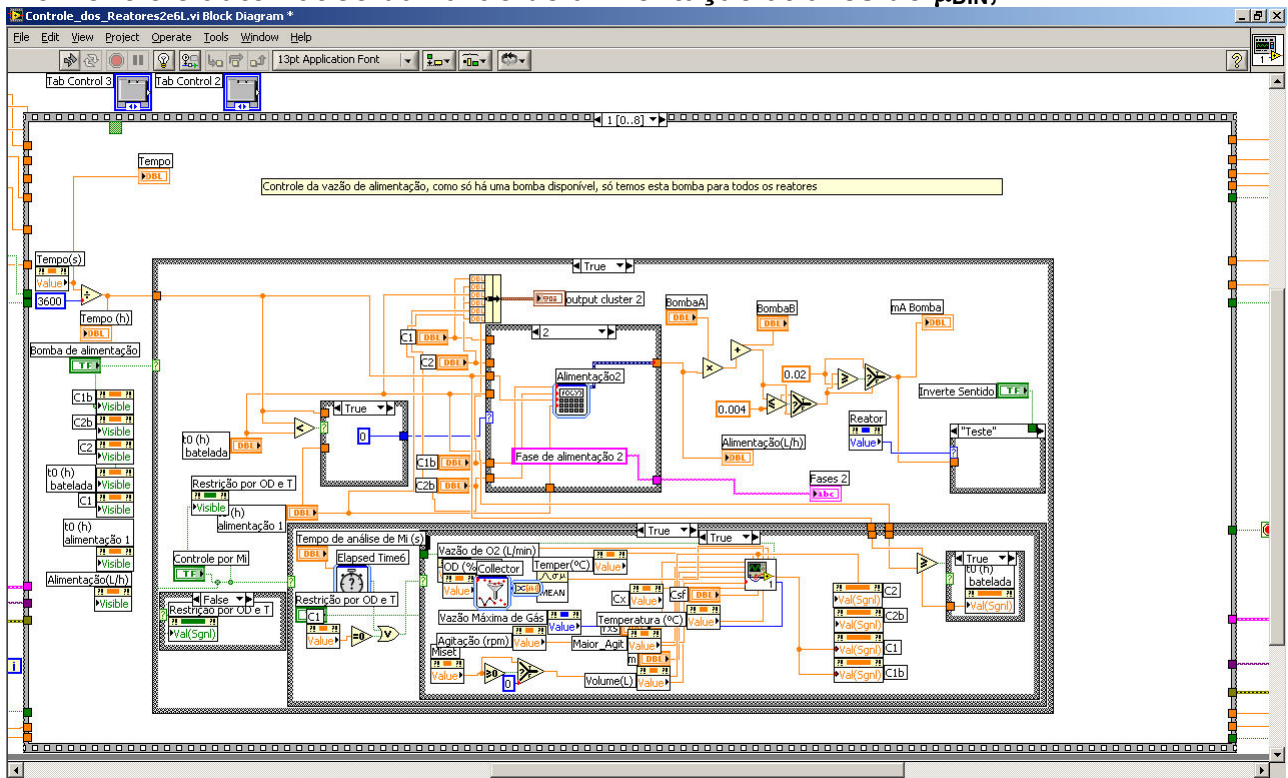


## 4.5 Controle automático do pH;

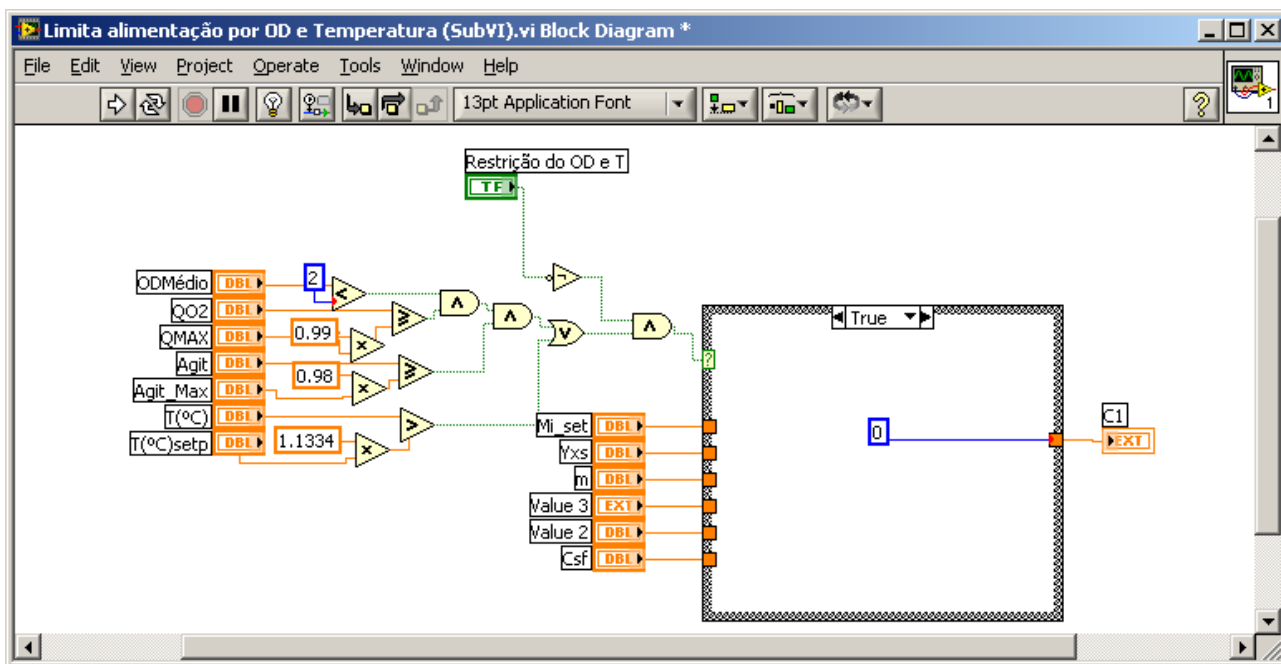
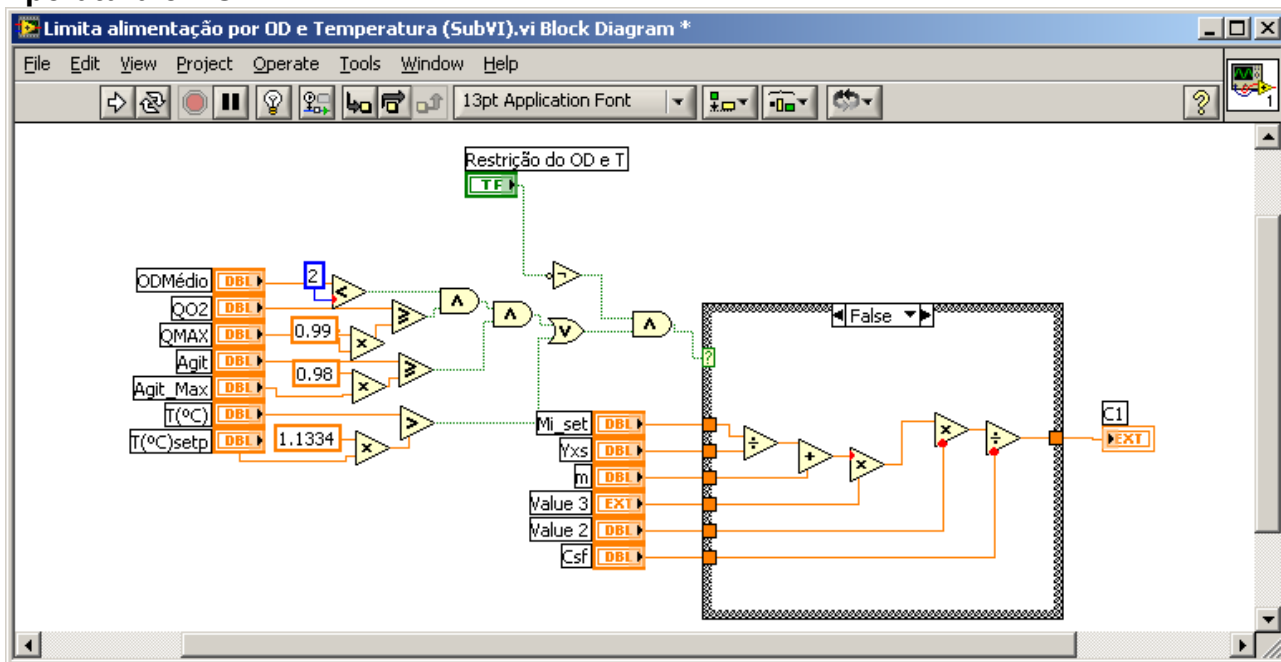




## 4.6 Controle automático da vazão de alimentação através do $\mu$ DIN;



#### 4.7 Controle automático da vazão de alimentação com restrição por temperatura e COD.





## INFORMAÇÕES SOBRE O ESTADO DA ARTE

### 1. IDENTIFICAR PROGRAMAS DE COMPUTADOR PRÉ-EXISTENTES, OBTIDOS ATRAVÉS DE BUSCAS, REFERÊNCIAS BIBLIOGRÁFICAS E/OU CONHECIMENTO DOS CRIADORES.

O programa desenvolvido evoluiu de um programa simples de operação do biorreator, feito pelo grupo do LaDABio do Departamento de Engenharia Química da UFSCar.

Programas similares já existem e são fornecidos pelos fabricantes de biorreatores, juntamente com os biorreatores. Porém, nenhum dos softwares oferecidos atualmente conta com um sistema capaz de conduzir um cultivo em batelada alimentada automaticamente até altas densidades, iniciando a alimentação e modificando a vazão de acordo com o crescimento celular.

Os programas existentes no mercado baseiam a alimentação em função da concentração de oxigênio dissolvido, do tempo e/ou do pH.

Alguns fabricantes de biorreatores e de programas de computador semelhantes ao que estamos desenvolvendo estão listados abaixo:

Fermentec: [www.ifermentec.com.br](http://www.ifermentec.com.br).

Marconi: <http://www.marconi.com.br/home.jsp>.

Fairport: <http://www.fairport.com.br/a-empresa>.

New Brunswick Scientific: <http://www.nbsc.com/default.aspx>.

### 2. DETALHAR A INOVAÇÃO DO PROGRAMA DE COMPUTADOR, EXPLICANDO QUE PROBLEMAS SÃO RESOLVIDOS E/OU QUAIS AS VANTAGENS EM RELAÇÃO ÀQUELES PRÉ-EXISTENTES.

O programa SUPERSYS\_HCDC além de realizar a aquisição de dados e controlar o pH, a COD e a temperatura, como qualquer programa de monitoração de biorreatores, apresenta algumas funcionalidades inovadoras exclusivas, são elas:

- i) *Startup* (Partida) da alimentação de forma automática por reconhecimento do final da batelada por Redes Neurais, Lógica Fuzzy e Derivada da vazão molar de gás carbônico:

Esta funcionalidade garante que a alimentação seja iniciada no momento correto, eliminando a necessidade do acompanhamento integral por um especialista ao longo do cultivo (ie. elimina a necessidade de retiradas de amostras de hora em hora e análises das concentrações de substrato durante o cultivo). Esta funcionalidade influencia diretamente na diminuição de custo, e aumento da produtividade do processo. Esta função está habilitada para meios definidos contendo glicerol ou glicose como fonte de carbono. Um novo módulo do programa para a partida automática da bomba de alimentação em cultivos com meios complexos está em desenvolvimento.

- ii) Correção automática da vazão de alimentação de acordo com o crescimento celular, garantindo manutenção de crescimento máximo, sem ocorrência de excesso de substrato;

A escolha da vazão de alimentação geralmente é feita por meio de simulações de modelos matemáticos antes da realização do experimento, e seguem um perfil pré-estabelecido mesmo se ocorrerem variações no perfil de crescimento, que podem acontecer. A abordagem de controle que é proposta em nosso sistema corrige a vazão de alimentação de acordo com a velocidade específica de crescimento, medida *online*, garantindo que a alimentação siga estritamente o perfil de crescimento do microrganismo e de assimilação da fonte de carbono. Esta funcionalidade influencia diretamente na redução de subprodutos, levando ao aumento da qualidade do produto, diminuição dos custos e aumento da produtividade. Esta funcionalidade está habilitada para meio suplementar contendo glicerol como fonte de carbono na versão atual do sistema de supervisão. Um novo módulo do programa para o controle automático da vazão de meio suplementar contendo glicose como substrato está em desenvolvimento.

- iii) Enriquecimento automático da alimentação de gases com oxigênio puro, aumentando ou diminuindo as vazões de ar e de oxigênio de acordo com os limites definidos pelo usuário, utilizando uma lógica de árvores de decisão;

O controle da concentração de oxigênio dissolvido em CADC não é uma tarefa trivial, o qual é em parte realizado por um algoritmo de controle (PID) atuando diretamente na velocidade de agitação. Porém, as vazões de ar e oxigênio não podem ser facilmente controladas por um PID, (conhecimento adquirido por testes anteriores), sendo que o controle por árvores de decisão apresentou ótimo desempenho, garantindo que haja enriquecimento com oxigênio automaticamente à medida que a demanda aumenta. Esta funcionalidade influencia diretamente na redução de custos, aumento da qualidade do produto, e contribui para a reprodutibilidade do processo.

- iv) Reconhecimento de falhas nos sistemas periféricos, como nas linhas de gás, banho termostático, bombas de adição de base e ou de ácido e início da fase de morte;

Os cultivos em biorreator dependem do funcionamento de vários acessórios, que podem apresentar alguma falha ou mesmo serem mal ajustados (por exemplo, o operador pode se esquecer de abrir a válvula de ar ou de oxigênio, ou então acabar o estoque de gás durante o experimento) durante o processo. Uma destas variáveis pode também não estar no valor desejado e o operador pode não notar. Nos casos mencionados, o programa de computador alerta o usuário com uma mensagem na tela e com uma mensagem de voz, indicando a ocorrência da falha. Esta funcionalidade contribui para a redução de custos, identificando problemas de fácil solução para que sejam sanados antes de comprometerem todo o cultivo.

- v) Notificação das falhas por email, por mensagens sonoras e até mesmo por mensagens para celular (SMS);

Como relatado anteriormente o sistema reconhece falhas nos periféricos e sinaliza por mensagens na tela e por mensagens sonoras, caso o operador esteja desatento. Além disso, as mesmas mensagens são enviadas ao email do operador, que pode estar ausente do laboratório e ao receber o email, entrar na internet e intervir no processo. Além disso, o operador pode cadastrar, junto à operadora de celular, a conta de email para que sempre que receber uma mensagem do software, automaticamente encaminhar a mensagem via torpedo SMS, recebendo qualquer notificação de falha no celular. Esta funcionalidade contribui para a redução de custo, aumento na reprodutibilidade do processo, à medida que dispensa o operador especialista da obrigação de estar o tempo todo ao lado do processo.

- vi) Flexibilidade do sistema de supervisão SUPERSYS\_HCDC

O sistema de supervisão foi construído em módulos. Assim, a ativação ou não das diferentes funcionalidades disponíveis é de livre escolha do usuário. Da mesma forma, graças à estrutura modular, é possível expandir as funcionalidades por meio da inclusão de novos módulos, o que garante a constante atualização das ferramentas e a adaptação da mesma às especificidades de cada cultivo.

Exemplos de novos módulos em implantação na ferramenta são fornecidos a seguir:

- i) Monitoração da pressão do biorreator, tornando o processo mais seguro, de forma que o programa possa identificar algum entupimento na saída de gás e prevenir acidentes por meio da interrupção automática da alimentação de ar e/ou oxigênio e alertando o operador sobre o problema.
- ii) Inclusão de um módulo para cultivos com glicose;
- iii) Inclusão de um sistema de inferência do crescimento celular independente do sensor de permissividade.
- iv) Inclusão de um módulo de intervenção inteligente na temperatura do banho termostaticado, de forma a antecipar as mudanças no setpoint do banho para atender as necessidades do cultivo;
- v) Salvamento de todo o histórico dos controles ao longo de todo tempo de funcionamento;
- vi) Salvamento de todos os dados colhidos ao longo do experimento a cada iteração do software (tempo aproximado de até 2s). Os dados são salvos em txt, podendo ser exportados e abertos em qualquer outro programa.

### **3. DESCREVA POSSÍVEIS DESVANTAGENS/LIMITAÇÕES QUE O PROGRAMA DE COMPUTADOR POSSUI, COMPARANDO-O, SE POSSÍVEL, ÀQUELES PRÉ-EXISTENTES.**

Este programa foi construído para evitar as limitações, possibilitando a inserção de novos módulos para suprir cada limitação que surgir. Em geral o programa de computador pode ser usado para operar cultivos de qualquer microrganismo, com as seguintes ressalvas:

- i) O reconhecimento de fases e acionamento automático da bomba de alimentação só foi testado com cultivos de *E. coli* com meio definido na batelada inicial. Para ser utilizado em outros microrganismos e outros meios, é necessário um experimento prévio, similar ao que será reproduzido futuramente, para gerar informações que serão utilizadas no treinamento das redes neurais, de forma que as redes aprendam o novo processo.
- ii) a ferramenta oferece muitas opções para o usuário e por isso mesmo sua operação é de relativa complexidade, sendo recomendado um treinamento antes da sua utilização e configuração correta dos parâmetros.

## INFORMAÇÕES RELATIVAS AO DESENVOLVIMENTO E INVESTIMENTO

### 1. DATAS RELACIONADAS À CRIAÇÃO E DESENVOLVIMENTO DO PROGRAMA DE COMPUTADOR.

EVENTO	DATA APROXIMADA
Concepção/Idéia	Novembro de 2008
Data de criação	Janeiro de 2009
Primeira versão funcionando	Fevereiro de 2009
Versão atual funcionando	Mai de 2010

### 2. QUAL O ESTÁGIO DE DESENVOLVIMENTO DO PROGRAMA DE COMPUTADOR EM RELAÇÃO À COMERCIALIZAÇÃO?

- Estado embrionário (Precisa muito trabalho para levá-lo ao mercado)  
 Parcialmente desenvolvido (Poderá ser levado ao mercado com um investimento razoável)  
 Desenvolvido (Pode ser levado ao mercado com um mínimo investimento)

Observações:

O programa já está em uso e pode ser instalado em biorreatores em outros laboratórios, necessitando apenas da instalação (a instalação é complexa e deve ser feita por um técnico bem treinado para configurar os canais de comunicação com os equipamentos).

### 3. ESPECIFIQUE O INVESTIMENTO JÁ REALIZADO NO DESENVOLVIMENTO DO PROGRAMA DE COMPUTADOR.

O investimento constitui-se em horas de programação e pesquisa, pagas com bolsa de pós-graduação, além das licenças dos softwares de programação LabView e MatLab. Além destes investimentos o programa de computador foi testado inúmeras vezes em cultivos de diferentes microrganismos: *E.coli*, *Bacillus megaterium*, *Bacillus subtilis*, *Pichia pastoris* e *Saccharomyces cerevisiae*. A Tabela a seguir apresenta a estimativa de custos envolvidos com referido programa de computador.



Tabela de estimativa de custos.

Gastos	Valor
Bolsa Doutorado	R\$ 45.000,00
Cultivos	R\$ 2.500,00
Licença LabView	R\$ 5.000,00
Licença MatLab	R\$ 20.000,00
Outros	100.000,00
<b>TOTAL</b>	<b>R\$ 172.500,00</b>

**4. EXISTEM ÓRGÃOS DE FOMENTO OU FUNDOS DE APOIO ENVOLVIDOS? SE EXISTEM, INDIQUE QUAIS E SE FORAM INFORMADOS A RESPEITO DO PROGRAMA DE COMPUTADOR?**

Sim, existem. A CAPES financiou a bolsa de doutorado do aluno responsável pela programação, implementação e validação da ferramenta. A FAPESP financiou equipamentos e insumos para a pesquisa e desenvolvimento do sistema. O desenvolvimento deste programa está vinculado a um subprojeto que se insere no projeto temático da FAPESP: “Vacina conjugada anti-pneumocócica: Estudos sobre a viabilidade de uma vacina Polissacarídeo – Proteína de superfície de pneumococo A”, proc. FAPESP 2008/05207-4; sob coordenação: Martha Massako Tanizaki. A CAPES não foi notificada sobre o programa, porém a FAPESP sim; foi notificada através do relatório anual do projeto, porém, até o momento não houve manifestação a esse respeito.

**5. EXISTE ALGUM INSTRUMENTO LEGAL (CONTRATO OU SIMILAR) ACORDADO COM ALGUMA INSTITUIÇÃO PARCEIRA QUE DEFINE DIREITOS E DEVERES EM MATÉRIA DE PROPRIEDADE INTELECTUAL, RELATIVOS AO PROGRAMA DE COMPUTADOR?**

Não, o programa foi desenvolvido todo no LaDABio, sem envolvimento de qualquer instituição externa à universidade.

**INFORMAÇÕES RELATIVAS À DIVULGAÇÃO E COMERCIALIZAÇÃO**

O Programa de computador em questão, com as funcionalidades aqui especificadas já está em funcionamento; além disso, uma versão com novas funcionalidades já está

sendo desenvolvida. Portanto, já podemos instalar o programa para usuários interessados.

O grupo LaDABio decidiu que o software desenvolvido na universidade, com recursos públicos, deveria ter licença aberta, e que usuários interessados no produto, pagariam pelos serviços de instalação, configuração e treinamento.

### **1. O PROGRAMA DE COMPUTADOR E SUAS FUNCIONALIDADES JÁ FORAM REVELADAS ANTERIORMENTE? SE SIM, ESPECIFICAR A QUEM E QUANDO?**

Sim, foi escrito um trabalho para um congresso de engenharia química relatando as vantagens do programa, sem mencionar o código fonte nem mesmo o pseudo-código. O trabalho foi aceito e será apresentado em setembro de 2010.

As funcionalidades implementadas foram relatadas para pesquisadoras do Instituto Butantan, que também realizam cultivos em biorreator.

### **2. SE REVELADO, FOI MANIFESTADO INTERESSE COMERCIAL? EM CASO AFIRMATIVO, INFORME NOME, CONTATO E TELEFONE DO INTERESSADO.**

As pesquisadoras do instituto Butantan demonstraram muito interesse e esperam que o programa esteja pronto o quanto antes para poderem utilizar.

Uma pesquisadora interessada no programa é a Doutora Viviane M. Gonçalves (vivimaig@butantan.gov.br), e também a Doutora Mickie Takagi (mtakagi@butantan.gov.br).

A divulgação do programa em congressos da área irá potencializar o número de interessados.

### **3. EM QUE MERCADOS POTENCIAIS EXISTE A POSSIBILIDADE DE COMERCIALIZAÇÃO DO PROGRAMA DE COMPUTADOR?**

O programa desenvolvido pode ser usado para operar biorreatores com quaisquer organismos. Sendo assim, pode ser utilizado em centros de pesquisa e desenvolvimento que utilizem com biorreatores, empresas que cultivam microrganismos para quaisquer fins assim como em indústrias farmacêuticas como a Vallée ou a Pfizer, dentre outras, que também possuem etapas envolvendo cultivo de microrganismos em seus processos produtivos. O programa foi especialmente elaborado para supervisão e monitoramento de cultivos de alta densidade de *E. coli* recombinante para a produção de proteínas e, portanto, tem nesse tipo de bioprocessos seu maior potencial de

aplicação e desempenho. Pode ainda ser adquirido por fabricantes de biorreatores que queiram incluir o programa como parte do produto. Centros de pesquisa e de produção de vacinas como o Instituto Butantan e o Instituto de Tecnologia em Imunológicos Biomanguinhos da FIOCRUZ são mercados potenciais. Outros institutos de pesquisa como IPT, CTBE, CTC, Instituto Pasteur etc, localizados no Brasil ou no exterior, também são.

Existe ainda a possibilidade de em um futuro próximo, mais ou menos dois anos, o programa venha a fazer parte de um produto, biorreator e programa de controle, que queremos desenvolver, provavelmente iniciando o processo de desenvolvimento em 2011.

## **Referências**

DEMAIN, A.; VAISHNAV, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances* 27, 297-306, 2009.

LILJEQVIST, S.; STAHL, S. Production of recombinant subunit vaccines: protein immunogens, live delivery systems and nucleic acid vaccines. *J. Biotechnol.*, v. 73, p. 1-33, 1999.

SHILOACH, J.; FASS, R. Growing *E. coli* to high cell density: a historical perspective on method development. *Biotech Adv.*, v. 23, p. 345-357, 2005.

# ANEXO A – CERTIFICADO DE REGISTRO DO PROGRAMA SUPERSYS\_HCDC



REPÚBLICA FEDERATIVA DO BRASIL  
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR  
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

## CERTIFICADO DE REGISTRO DE PROGRAMA DE COMPUTADOR

## RS 11008-6

O INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL expede o presente Certificado de Registro de Programa de Computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de criação indicada, em conformidade com o art. 3º da Lei Nº 9.609, de 19 de Fevereiro de 1998, e arts. 1º e 2º do Decreto 2.556 de 20 de Abril de 1998.

Título: **SUPERSYS\_HCDC - PROGRAMA DE COMPUTADOR PARA  
MONITORAMENTO E CONTROLE DE BIORREATOR PARA CULTIVOS**

Data de Criação 02 de Agosto de 2010

Titular: 45.358.058/0001-40 UNIVERSIDADE FEDERAL DE SÃO CARLOS

Criadores: 709.153.931-15 ANTONIO CARLOS LUPERNI HORTA  
066.705.805-20 TERESA CRISTINA ZANGIROLAMI  
013.100.148-50 ROBERTO DE CAMPOS GIORDANO  
159.873.418-09 ANTONIO JOSÉ GONÇALVES DA CRUZ  
305.748.218-05 GIANN BRAUNE REIS  
644.424.381-53 CHARLES DAYAN FARIAS DE JESUS

Linguagem: G

Campos de Aplicação: BL-07, IN-03, IN-05, SD-09

Tipos de Programa: AT-06, CD-04, IA-01, IA-02, SO-07

Documentação Técnica em depósito SOB SIGILO até 10/09/2020.

A exclusividade de comercialização do programa de computador objeto deste Certificado não tem a abrangência relativa à exclusividade de fornecimento estabelecida pelo art. 25, inciso I da Lei Nº 8.666, de 21 de Junho de 1993, para fins de inexigibilidade de licitação para compras pelo poder público.

Expedido em 19 de Julho de 2011.

**Rodrigo Moerbeck de Almeida Rego**  
Chefe da Divisão de Registro de Programas de  
Computador e Topografia de Circuitos Integrados

**Breno Bello de Almeida Neves**  
Diretor de Contratos, Indicações Geográficas e  
Registros

