SPECIFIC GROWTH RATE REGULATION IN A SIMULATED FED-BATCH *E. COLI* FERMENTATION

I. Rocha and E. C. Ferreira

Centro de Engenharia Biológica, Universidade do Minho
4710-057 Braga, PORTUGAL

e-mail: {irocha, ecferreira}@deb.uminho.pt

Abstract: The specific growth rate is one of the most important process variables characterizing the state of microorganisms during fermentations mainly because the biosynthesis of many products of interest is often related with the values assumed by this parameter. In the particular case of the fed-batch operation of *Escherichia coli* for the production of recombinant proteins, it is often argued that both pre- and the post-induction specific growth rates should be closely controlled in order to achieve maximum productivities on the desired recombinant protein.

In this work a feedforward-feedback controller was developed with the purpose of regulating the global specific growth rate during a fed-batch fermentation of *E. coli*. The developed algorithm allows to maintain the cells in two different metabolic regimens (simultaneous oxidative and fermentative growth on glucose or oxidative growth on glucose), depending on the selected setpoint for the controlled variable.

The pure open-loop version of the controller revealed a relatively poor performance when dealing with process noise. However, the introduction of on-line measurements of fermenter weight and biomass estimation obtained from asymptotic observers allowed a better approximation between the desired setpoints and the simulated values of the specific growth rates. Finally, the introduction of a proportional action in the controller equation allows an improved robustness against variations in model parameters.

Key-words: Specific growth rate estimators, Biomass observers, Asymptotic Observer, Fed-batch fermentation, *E. coli*, Specific growth rate control, feedback-feedforward controller.

1. INTRODUCTION

The specific growth rate is one of the most important process variables characterizing the state of microorganisms during fermentations mainly because the biosynthesis of many products of interest is often related with the values assumed by this variable. As opposed to batch fermentation, in fed-batch cultures it is possible to manipulate the specific growth rate at an appropriate value providing a desirable metabolic condition, resulting in maximum productivity.

Additionally, for certain types of measurements, of vital importance in the post-genomic era (Bro and Nielsen, 2004), like mRNA abundance or the analysis of the fluxome and proteome, microbial cultures have to be sampled at a pseudo steady-state condition that can be obtained by imposing a fixed specific growth rate, either in continuous or fed-batch cultures.

In the particular case of the fed-batch operation of *Escherichia coli* for the production of recombinant proteins, the specific growth rate should be kept

below a certain threshold if the goal is to avoid the accumulation of acetic acid (Jana and Deb, 2005) and, additionally, it is often argued that both pre- and the post-induction specific growth rates should be closely controlled in order to achieve maximum productivities on the desired recombinant protein (Curless *et al.,* 1990; Sanden *et al*, 2003; Lim and Jung, 1998).

In order to keep the specific growth rate at a pre-determined value, the most common approach is to apply a feed-forward exponential feeding strategy, where the nutrients required by the culture for achieving the desired growth rate are pre-determined and satisfied at any moment. However, the inherent features of a feed-forward method limit the application of this feeding scheme, due to the likely occurrence of external perturbations or variations on culture parameter.

Therefore, the development of reliable algorithms for the feedback automatic control of the specific growth rate in fed-batch systems is of paramount importance in fermentation technology. However, the performance of such algorithms is critically dependent on a reliable determination of the specific growth rate, which cannot be obtained directly from common fermentation measurements mainly due to the lack of reliable sensors for the determination of biomass concentration. A combination of a reliable model of the process and on-line data is therefore often necessary for the estimation of both biomass concentration and specific growth rates.

Some algorithms used for on-line estimation of reaction rates using biomass concentration or other correlated variables measurements have been proposed. Pomerleau and Perrier (1990) and Pomerleau and Perrier (1992) proposed and validated experimentally an on-line estimation algorithm for multiple reaction rates. This procedure was applied to baker's yeast fermentation, and the algorithm required the on-line measurement of two or three state variables. Also, in Lubenova (1999) the author describes a methodology for the design of a new parameter estimator of biomass growth rate and yield coefficient for oxygen consumption on the basis of the theory of adaptive estimation, for a class of aerobic bioprocesses in fed-batch or continuous mode. In Lubenova *et al* (2003), the authors proposed an approach for on-line growth rate estimation for a class of aerobic batch processes with dissolved oxygen control in the culture medium. The only required on-line measurement is the oxygen consumption rate. An adaptive model-based algorithm for the on-line estimation of reaction rates is described by Oliveira *et al.* (2002), considering the yield coefficients invariable and known, based on the approach of Bastin and Dochain (1990) to stirred tank reactors.

Regarding the development of control algorithms for the specific growth rate, there are very few examples where closed loop control approaches have been applied to *E. coli* or to other relatively complex bioprocesses. In Levisauskas (2001) and Levisauskas *et al.* (1996) the authors have developed a strategy for automatic control of the specific growth rate in fed-batch cultivation that requires the on-line monitoring of OUR or CER and feed rate and volume of the fermentation broth. This strategy was applied to fed-batch fermentation of *E. coli*.

The main purpose of this work is to develop reliable algorithms for controlling the specific growth rates in *E. coli* fermentation. The performance of open-loop controllers is evaluated and compared with controllers with incorporation of on-line weight measurements and biomass estimation obtained using an asymptotic observer derived form the mathematical model of the process. Finally, a proportional feedback controller is constructed that uses the values of the specific growth rates obtained by estimators for the determination of the controller error.

2. PROCESS MODELLING

The development of the mathematical model for the fed-batch fermentation of recombinant *E. coli* was based on the assumption that the aerobic growth of the microorganism can follow three main different metabolic pathways (Rocha and Ferreira, 2002): oxidative growth on glucose, fermentative growth on glucose, and oxidative growth on acetate, the corresponding dynamical model for fed-batch fermentation being represented as follows:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ A \\ O \\ C \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -y_1 & -y_2 & 0 \\ 0 & y_3 & -y_4 \\ -y_5 & -y_6 & -y_7 \\ y_8 & y_9 & y_{10} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} X - \frac{F}{W} \begin{bmatrix} X \\ S \\ A \\ O \\ C \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{F_{in}}{W} S_{in} \\ 0 \\ OTR \\ -CTR \end{bmatrix} \quad (1)$$

where *X*, *S*, *A*, *O*, and *C* represent biomass, glucose, acetate, dissolved oxygen, and dissolved carbon dioxide concentrations, respectively; μ_1, μ_2, and μ_3 are the specific growth rates; y_i are the yield (stoichiometric) coefficients; F_{in} and S_{in} are the substrate feed rate and the influent glucose concentration, respectively; *W* is the culture medium weight. *CTR* is the carbon dioxide transfer rate from liquid to gas phase, and *OTR* is the oxygen transfer rate from gas to liquid phase.

The variation of the culture medium weight with the time is given by:

$$\frac{dW}{dt} = F \tag{2}$$

where F includes weight variations due to the substrate feed rate, the amount of culture removed or added during sampling, base and acid additions, evaporation and mass taken from the reactor due to gas exchanges, that can not be considered negligible in small-scale high-cell density reactors.

The specific growth rates μ_1, μ_2, and μ_3 are represented by nonlinear functions of the Monod kinetics type with non-competitive inhibition caused by acetate. The oxidative bottleneck exhibited by this microorganism is accounted for by imposing a threshold in the oxygen uptake rate that will limit the fraction of the glucose flux that will be directed to the oxidative pathway (μ_1). After that threshold is reached, the microorganism will also follow the fermentative pathway. A full description of the kinetic model used for the fed-batch growth of *E. coli* can be found in Rocha (2003).

As a consequence of the characteristics of the oxidative bottleneck, the three metabolic pathways represented in the mathematical model do not occur simultaneously in the cell, originating four partial models corresponding to different metabolic regimens:
- Simultaneous oxidative and fermentative growth on glucose (μ_1, $\mu_2 > 0$)
- Oxidative growth on glucose ($\mu_1 > 0$)
- Simultaneous oxidative growth on acetate and glucose (μ_1, $\mu_3 > 0$)
- Oxidative growth on acetate ($\mu_3 > 0$)

3. BIOMASS OBSERVER AND SPECIFIC GROWTH RATE ESTIMATORS

Biomass concentration is an essential variable for the design of controllers for the specific growth rate in *E. coli* because it has necessarily to be incorporated in the controller equation and also because it is critical for the estimation of the specific growth rates. However, due to the inexistence of reliable sensors for this determination, its values have to be estimated. This estimation can be performed from equation (1) by integrating the corresponding mass balance using initial values for the state variable and the feeding rate profile. However, this estimation usually originates great discrepancies because it is highly dependent on the accuracy of the initial values and it assumes that the process does not suffer any deviation from the proposed model.

An interesting alternative comes from the use of software sensors, where the process model is used together with a subset of on-line measured variables to obtain an estimation of the biomass concentration values.

Although in the authors' lab there exists the possibility of measuring on-line all the state variables from equation 1 except for biomass (Rocha and Ferreira, 2002), for the sake of more general applicability, the algorithms described in this paper assume the availability of on-line data for dissolved oxygen and carbon dioxide, *CTR*, *OTR* and culture weight. Additionally, both the biomass observer and the specific growth rates estimators assume that the yield coefficient matrix Y is known.

Therefore, using the information available on-line and according to Veloso *et al.* (2005) it is possible to obtain biomass estimation using either Extended Kalman Observers (EKO) or Asymptotic Observers (AO) only if a partial model is considered. Taking into account that acetate consumption is very often negligible in these processes, the "simultaneous oxidative and fermentative growth on glucose" and the "oxidative growth on glucose" regimens are the only ones considered. Additionally, and although the performance of the EKO was shown to be superior for this process, the AO was selected for biomass estimation, as it requires no knowledge on the kinetic structure of the model and also due to the inexistence of tuning parameters.

The AO allows reconstructing the missing state variables even when the process is not exponentially observable and the kinetics are unknown (Bastin and Dochain, 1990). Biomass estimation requires the integration of one differential equation associated with the auxiliary variable Z where its dynamics is independent of growth rate vector $r(\xi, t)$:

$$\frac{d\hat{Z}}{dt} = -D\hat{Z} - \alpha_1 OTR - \alpha_2 CTR \tag{3}$$

The variables α_i are a function of the yield coefficients matrix as follows:

$$\begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} = \frac{1}{y_5 y_9 - y_6 y_8} \begin{bmatrix} -y_9 & -y_6 \end{bmatrix} \tag{4}$$

The biomass estimate is then given by

$$\hat{X} = \hat{Z} + \alpha_1 O + \alpha_2 C \tag{5}$$

The design of specific growth rate estimators is based on the formulation proposed by Bastin and Dochain (1990) reformulated by Pomerleau and Perrier (1990) so that a decoupling of the dynamic model from the growth rate is achieved. Estimation of the three specific growth rates in equation 1 will come:

$$\frac{d\hat{\psi}_1}{dt} = \hat{\mu}_1 \hat{X} - D\hat{\psi}_1 + \frac{1}{y_5 y_9 - y_6 y_8}(y_9 OTR + y_6 CTR) + \omega_{11}(\psi_1 - \hat{\psi}_1)$$

$$\frac{d\hat{\psi}_2}{dt} = \hat{\mu}_2 \hat{X} - D\hat{\psi}_2 + \frac{1}{y_5 y_9 - y_6 y_8}(y_8 OTR - y_5 CTR) + \omega_{12}(\psi_2 - \hat{\psi}_2)$$

$$\frac{d\hat{\mu}_1}{dt} = \omega_{21}(\psi_1 - \hat{\psi}_1)$$

$$\frac{d\hat{\mu}_2}{dt} = \omega_{22}(\psi_2 - \hat{\psi}_2)$$

(6)

ψ is obtained from the transformation $\psi = Y_a^{-1}\xi_a$ where ξ_a is the partition of the vector ξ that includes the measured state variables, and (Y_a, F_a, Q_a) the corresponding parts of (Y, Q, F):

$$\begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} \begin{bmatrix} O \\ C \end{bmatrix}$$

(7)

With:

$$\begin{bmatrix} \alpha_3 & \alpha_4 \end{bmatrix} = \frac{1}{y_5 y_9 - y_6 y_8} \begin{bmatrix} y_8 & y_5 \end{bmatrix}$$

(8)

The calculation of the gains ω_{ij} for each instant i is made such that a second order dynamics (Oliveira *et al.*, 1996) is obtained:

$$\omega_{11,i} = 2\zeta_1/\tau_1 - \frac{X_i - X_{i-1}}{TX_i}$$

$$\omega_{12,i} = 2\zeta_2/\tau_2 - \frac{X_i - X_{i-1}}{TX_i}$$

$$\omega_{21,i} = (X_i \tau_1^2)^{-1}$$

$$\omega_{22,i} = (X_i \tau_2^2)^{-1}$$

(9)

where T is the integration step. Therefore, the implementation of these algorithms requires the tuning of 4 parameters (ζ_1, τ_1, ζ_2, and τ_2).

4. CONTROL ALGORITHM

In any controller for microbial growth the exponential growth of the microorganisms has to be taken into consideration. Using a feedforward–feedback controller these characteristics can be kept in view in the feedforward contribution. For the metabolic regimens assumed to be active in this study, the deduction of the feedforward component of the control equation is obtained from the mass balance for the substrate S of equation 1:

$$\frac{dS}{dt} = -y_1\mu_1(t)X(t) - y_2\mu_2(t)X(t) - \frac{F(t)}{W(t)}S(t) + \frac{F_{in}(t)}{W(t)}S_{in}$$

(10)

In the fed-batch phase the cultivation is operated under glucose limitation. Therefore, it can be assumed that $S=0$ and $dS/dt=0$ and rearranging equation 10 we obtain:

$$F_{in} = \frac{y_1\mu_{1,set}X(t)W(t)}{S_{in}} + \frac{y_2\mu_{2,set}X(t)W(t)}{S_{in}}$$

(11)

where $\mu_{1,set}$ and $\mu_{2,set}$ are the desired growth rates for the oxidative growth and the fermentative growth on glucose, respectively. This equation, however, is only valid when the desired specific growth rate is greater than the oxidative capacity of the microorganism. For smaller values, then $\mu_{2,set}$ has necessarily to be equal to zero and the second term is eliminated. This is the case when the objective is to keep the microorganism in the pure oxidative growth on glucose regimen

The pure feedforward equation can be improved with the incorporation of on-line estimations of biomass obtained from equation 5. Also, natural weight variations caused for example by evaporation are important sources of errors and therefore the incorporation of on-line weight measurements obtained using a balance to measure the bioreactor's weight is another alternative to obtain a better performance.

However, the application of this methodology may still not be sufficient for obtaining adequate control of bioprocesses because model parameters, such as the yield coefficients, must be given to compute the feeding rate and these values may change as the cultivation proceeds.

Therefore, the exponential feeding strategy is sometimes compensated for by incorporating some appropriate feedback control action. In Lee *et al.* (1997), the authors developed a simple method to adjust feeding rate based on the calculation of the actual specific growth rate, and correcting the feed rate using a proportional action. A similar approach was proposed by Smets *et al.* (2004). In Arndt *et al.* (2005), a Proportional Integral action is used in a similar way.

For this process, the feedforward-feedback controller equation will then come as:

$$F_{in} = \frac{\left(y_1\mu_{1,set} + y_2\mu_{2,set}\right)\hat{X}(t)W(t)}{S_{in}} - \tau_\mu \frac{\hat{\mu}(t) - \mu_{set}}{S_{in}}W(t)$$

(12)

where $\hat{\mu}(t)$ and μ_{set} are, respectively, the estimated (from equation 6) and the setpoint for the sum of the specific growth rates 1 and 2, corresponding to a global specific growth rate.

5. RESULTS AND DISCUSSION

For assessing the applicability of the developed algorithms, several simulated experiments were conducted. For that purpose, model simulations were performed by integrating equations 1 using the MATLAB version 7.1 subroutine ODE23s. The implementation of the observer and the estimators was conducted using the Euler integration method. Most of the mathematical operations behind the

design of the observer and the estimators were performed using the Symbolic Math toolbox running in MATLAB 7.1.

For validating the developed algorithms "real" values of the state variables were obtained by integration of equation 1. The "real" values of the variables that can be obtained on-line, i. e., *O*, *C*, *OTR*, *CTR* and *W* were then corrupted with white noise, according to the standard deviations typically found in this process at the authors' lab, originating "experimental" values. When the observer and estimator algorithms were used, "estimated" values were obtained from the "experimental" data corresponding to the measured variables.

In a first experiment, the pure feedforward controller was evaluated without incorporating on-line weight measurements and biomass estimation. For this case, those variables were obtained from integration of the corresponding dynamic equations (equations 1 and 2), by considering constant specific growth rates. These results are represented in figure 1 for a setpoint value of 0.15 h^{-1}. It is clear that both estimated and real values of the specific growth rates diverge from the desired setpoint, while real biomass concentration is lower than what is predicted from estimation.
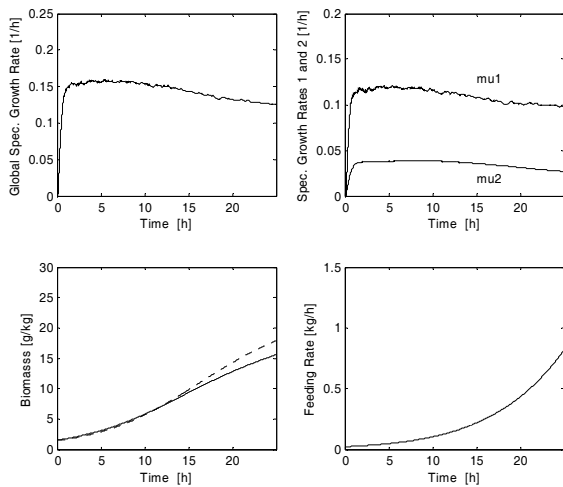

Figure 1 Performance of the pure feedforward controller for a setpoint of 0.15 h^{-1} on the global specific growth rate without incorporation of weight measurements and with estimation of biomass based on initial values of this variable. For biomass values, dotted lines correspond to "estimated" values while full lines are the simulated data. mu1 and mu2 mean μ_1 and μ_2.

The incorporation of the on-line estimation of biomass and weight measurements improves this performance (results not shown) but still do not force a convergence of the specific growth rate to the desired setpoint under the presence of perturbations in the model parameters.

In another experiment, the feedback-feedforward controller of equation 12 was implemented. The results obtained are shown in Figure 2, where it is clear that, after a trial-and-error approach for the selection of the tunning parameters, the controller is able to keep the specific growth rate very close to the setpoint, despite the noise associated to "experimental" variables.

In additional experiments the robustness of the feedback-feedforward controller to 5% variations on the model parameters k_1 and k_2 was verified and the controller was able to keep the desired specific growth rate after re-tuning of the parameters.

The performance of the controller is also similar when the objective is to keep the cells in the pure oxidative growth on glucose regimen, i. e., when $\mu_2=0$.
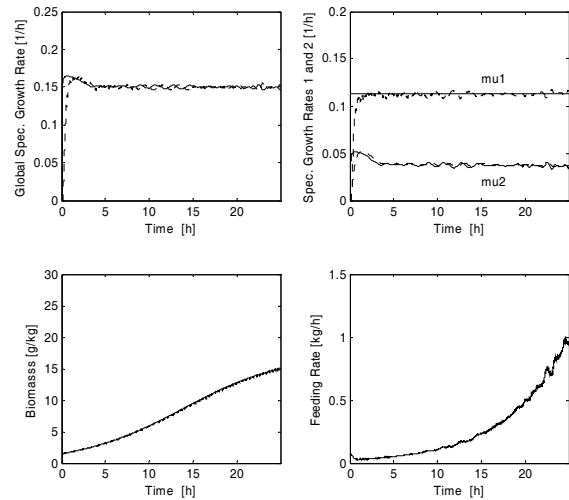

Figure 2 Performance of the feedback-feedforward controller for a setpoint of 0.15 h^{-1} on the global specific growth rate with incorporation of weight measurements and using the developed Asymptotic Observer for biomass estimation. The estimation of specific growth rates is used in the feedback part of the controller equation. Dotted lines correspond to "estimated" values while full lines are the simulated data. mu1 and mu2 mean μ_1 and μ_2.

6. CONCLUSIONS AND FUTURE WORK

In this work a controller was developed with the purpose of regulating the global specific growth rate during a fed-batch fermentation of *E. coli*. The developed algorithm allows to maintain the cells in two different metabolic regimens (simultaneous oxidative and fermentative growth on glucose or oxidative growth on glucose), depending on the selected setpoint for the controlled variable.

The pure open-loop version of the controller revealed a relatively poor performance when dealing with process noise. However, the introduction of on-line measurements of fermenter weight and biomass estimation obtained from asymptotic observers allowed a better approximation between the desired setpoints and the "experimental" values of the specific growth rates. Finally, the introduction of a proportional action in the controller equation allows an improved robustness against variations in model parameters.

Future work involves the experimental validation of the developed algorithms.

ACKNOWLEDGMENTS

REFERENCES

Arndt, M., Kleist, S., Miksch, G., Friehs, K., Flaschel, E., Trierweiler, J., and Hitzmann, B. (2005). A feedforward-feedback substrate controller based on a Kalman filter for a fed-batch cultivation of *Escherichia coli* producing phytase. *Computers and Chemical Engineering* **29**, 1113-1120.

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

Bro, C. and Nielsen, J. (2004). Impact of 'ome' analyses on inverse metabolic engineering. *Metabolic Engineering, 6*, 204-211.

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

Jana, S. and Deb, J. K. (2005). Strategies for efficient production of heterologous proteins in *Escherichia coli. App. Microb. Biotech., 67*, 289-298.

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

Levisauskas, D. (2001). Inferential control of the specific growth rate in fed-batch cultivation processes. *Biotechnology Letters, 23*, 1189-1195.

Levisauskas, D., Simutis, R., Borvitz, D. and Lübert, A. (1996). Automatic control of the specific growth rate in fed-batch cultivation processes based on an exhaust gas analysis. *Bioprocess Engineering* **15**, 145-150.

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

Lubenova, V. N. (1999). Stable adaptive algorithm for simultaneous estimation of time-varying parameters and state variables in aerobic bioprocesses. *Bioprocess Engineering* **21**, 219-226.

Lubenova, V., Rocha, I. and Ferreira, E. C. (2003). Estimation of multiple biomass growth rates and biomass concentration in a class of bioprocesses. *Bioprocess and Biosystems Engineering, 25*, 395-406.

Oliveira, R., Ferreira, E. C., Feyo de Azevedo, S. , (2002). Stability, dynamics of convergence and tuning of observer-based kinetics estimators. *Journal of Process Control, 12*, 311–323.

Oliveira, R., Ferreira, E.C., Oliveira, F, Feyo de Azevedo, S. (1996). A Study on the Convergence of Observer-Based Kinetic Estimators in Fed-Batch Fermentations. *J. Process Control*, **6**, , 367-371.

Pomerleau, Y. and Perrier, M. (1990). Estimation of multiple specific growth rates in bioprocesses. *AIChE Journal, 36*, 207-215.

Pomerleau, Y. and Perrier, M. (1992). Estimation of multiple specific growth rates: experimental validation. *AIChE Journal, 38*, 1751-1760.

Rocha, I. *Model-based strategies for computer-aided operation of recombinant E. coli fermentation*. 2003. Universidade do Minho. (PhD Thesis).

Rocha, I. and Ferreira, E. C. (2002). On-line simultaneous monitoring of glucose and acetate with FIA during high cell density fermentation of recombinant *E. coli. Analytica Chimica Acta.* **462**, 293-304.

Sanden, A. M., Prytz, I., Tubulekas, I., Forberg, C., Le, H., Hektor, A., Neubauer, P., Pragai, Z., Harwood, C., Ward, A., Picon, A., de Mattos, J. T., Postma, P., Farewell, A., Nystrom, T., Reeh, S., Pedersen, S., and Larsson, G. (2003). Limiting factors in *Escherichia coli* fed-batch production of recombinant proteins. *Biotechnology and Bioengineering, 81*, 158-166.

Smets, I. Y., Claes, J. E., November, E. J., Bastin, G. P. and Van Impe, J. F. (2004). Optimal adaptive control of (bio)chemical reactors: past, present and future. *Journal of Process Control, 14*, 795-805.

Veloso, A. C. A., Rocha, I. and Ferreira, E. C. (2005). Design of on-line state estimators for a recombinant *E. coli* fed-batch fermentation. Horacek, P., Simandl, M, and Zitek, P. *Preprints of the 16th IFAC World Congress. 2005. Prague - Czech Republic.*