AN INTEGRATED SYSTEM FOR ADVANCED MONITORING AND CONTROL OF FED-BATCH FERMENTATIONS OF RECOMBINANT *E. coli*

I. Rocha, E.C. Ferreira

Centro de Engenharia Biológica – IBQF, Universidade do Minho
4710–057 Braga, PORTUGAL
e-mail: {irocha, ecferreira}@deb.uminho.pt

Abstract: A system for the on-line monitoring of the major state variables during the production of recombinant proteins with *E. coli* was developed. Substrate (glucose) and the major by-product (acetate) were measured with a Flow Injection Analysis system, while the carbon dioxide and oxygen transfer rates were calculated from data obtained with exhaust gas analysis. The fermentation culture weight was also continuously assessed with a balance placed below the fermenter, allowing the use of more precise mass-based concentrations. The graphical programming environment LabVIEW was used to acquire and integrate these variables in a supervisor computer, allowing to perform integrated monitoring and control of the process. This framework is illustrated with different experiments for model calibration, state estimation by software sensors and process performance optimisation by exponential feed control. *Copyright 2001 IFAC*

Key-words: *Escherichia coli*, monitoring, estimation, process control, flow injection analysis, LabVIEW

1 INTRODUCTION

Successful operation and control of recombinant fermentation processes require accurate on-line knowledge about the real process state, including information about the main state variables. However, due to the existing diversity among the different sensors available, it is often a difficult task to integrate all the information in one single computer system in order to carry out supervisory process control. The graphical programming environment LabVIEW has been widely used in the last years (Claes and Van Impe, 1998; Kellerhals, *et al*., 1999) to accomplish this task. In this paper, a system composed of several equipments (including flow injection analysis, mass spectrometer, pumps, and balances) for the on-line monitoring during the production of recombinant proteins from *E. coli* and the programs developed to integrate these data are described.

2 SYSTEM DESCRIPTION

2.1 Experimental Conditions and Equipment

The bioreactor is a Biostat MD fermenter from *B. Braun Biotech* (Germany) with 5 L of capacity connected to a Digital Control Unit (DCU in Fig. 3). The *E. coli* fermention is operated in fed-batch mode, with an initial weight of 3 kg and the addition of 2 kg of feeding solution. A modified M9 medium with addition of trace metals and vitamins is used for the batch cultivation of the microorganism and addition of glucose and ammonia during the fed-batch phase (Riesenberg, *et al.*, 1990). The set-points for the environmental properties are 37ºC for temperature, pH 7, and 30% of saturation for dissolved oxygen. A mass spectrometer (Bioquad from Ledamass, UK) is connected to the exhaust gas line of the fermenter and also to the inlet aeration line for the on line measurement of the composition of those two channels. It was reassembled and adapted according

to Ferreira, *et al*. (1998). Two balances are used, one for the measurement of the total weight of the fermenter and the other for weighting the feeding solution reservoir.

A Flow Injection Analysis (FIA) system (Ismatec, Switzerland) is employed for the on-line measurement of glucose and acetate. Retrieval of the liquid medium samples free of biomass and suspended particles needed for this measurement is obtained by on-line filtration of the fermentation culture using an external unit, composed of a peristaltic pump and a A-SEP tangential filtration device (Applikon, The Netherlands). The FIA system for the measurement of acetate was adapted from Forman, *et al*. (1991) and is illustrated in Fig. 1. It is based on the diffusion of acetate through a gas-diffusion chamber into a stream containing an acid-base indicator. The subsequent decrease in the absorbance is detected with an incorporated photometer. Glucose is measured amperometrically by the presence of hydrogen peroxide produced from glucose by Glucose Oxidase (Fig. 2).

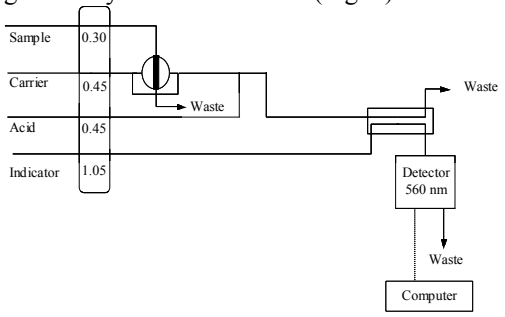

Fig. 1. Schematic representation of the system used for the analysis of acetate with FIA.

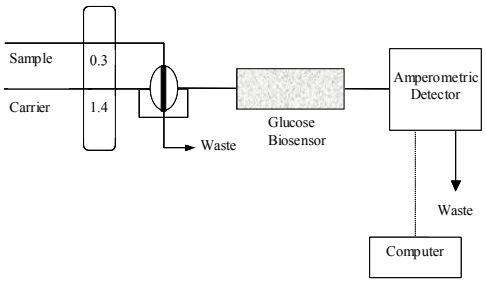

Fig. 2. Representation of the method used for the analysis of glucose with FIA.

2.2 Hardware and Software

LabVIEW programs consist of three parts: the front panel, which is the interface between the program and the user; the block diagram (the program itself) and the icon connector, which is responsible for data flow between routines. Together, these three elements integrate a Virtual Instrument (VI), the basic element of a LabVIEW program. Several VIs were constructed to acquire data from each specific instrument, and then called from a superior hierarchic level in the supervisor program. All the control routines were also implemented in this main program.

RS-232 protocol is used for the communication with the fermenter's DCU. During a fermentation, dissolved oxygen concentration, temperature, pH, stirrer speed and air flow rate values are converted into ASCII strings and received by the correspondent VI, that continuously decodes the information, stores it and leaves it ready for being used for control purposes. RS-232 and string interpretation are also used in the communication with the balances for the acquisition of the weight measurements.

Information from the mass spectrometer is acquired by the software RGA for Windows (Spectra, UK), running in multitask mode in the same computer. A VI is used to communicate with that program through the standard Windows Dynamic Data Exchange (DDE) protocol.

Data from the FIA system (glucose and acetate concentrations) are acquired in one dedicated computer with the ASIA software (ISMATEC, Switzerland) that also controls the entire system (testing and recalibrating when necessary). In order to obtain the results in the main computer and to integrate them with the remaining state variables, a VI was developed for the communication between both computers. That program identifies the last measurement, reads it, saves it and leaves it ready for being used in any algorithm.

These variables are integrated in the supervisor VI as input information for control algorithms, being the controller output the set-point for the feeding pump.

The experimental set-up showing all the equipments described and the correspondent connections is shown in Fig. 3.

2.3 Process Model

For advanced monitoring and control purposes, a mathematical model was derived, based on the general state space dynamical model described by Bastin and Dochain (1990). Accordingly, the dynamics of a reaction network in a stirred tank bioreactor can be described by the following mass balance equations written in matrix form as

$$\frac{d\xi}{dt} = Kr(\xi) - D\xi + F - Q \qquad (1)$$

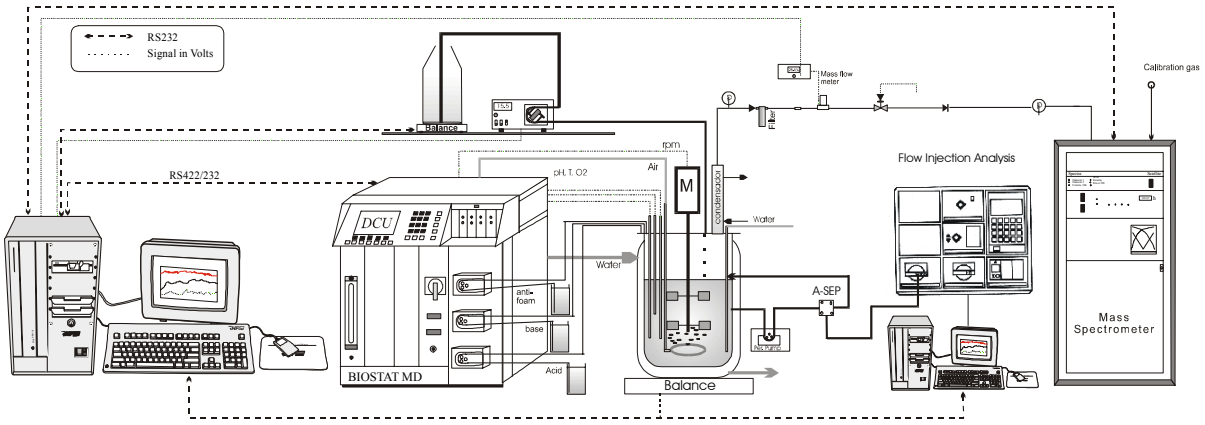

Fig. 3. Experimental set-up.

where $\boldsymbol{\xi}$ is the vector of the n component concentrations ($\xi \in R^n$), **r** is the vector of reaction rates ($\mathbf{r} \in R^m$), **K** is the matrix of yield coefficients ($\mathbf{K} \in R^{nxm}$), **F** is the vector of feed rates and **Q** is the vector of gaseous outflow rates ($\mathbf{F}, \mathbf{Q} \in R^n$), **D** is the dilution rate (D^{-1} is the residence time) a scalar defined as

$$D(t) = \frac{F_{in}(t)}{W(t)} \tag{2}$$

being F_{in} the influent flow rate (Kg h^{-1}), and W the culture medium weight (Kg).

For the growth of *Escherichia coli* several models are available in the literature (Cockshott and Bogle, 1999, Levisaukas, *et al.*, 1999, Galvanauskas, *et al*., 1998). These authors usually consider three main metabolic pathways for this process:

Respiratory growth on glucose:

$$k_1 S + k_5 O \xrightarrow{\mu_1} X + k_7 C \tag{3}$$

Fermentative growth on glucose:

$$k_2 S \xrightarrow{\mu_2} X + k_8 C + k_3 Ac \tag{4}$$

Respiratory growth on Acetate:

$$k_4 Ac + k_6 O \xrightarrow{\mu_3} X + k_9 C \tag{5}$$

However, since high cell densities are achieved in this process, cellular maintenance has also to be considered:

Maintenance:

$$S + k_{m1} Ac + k_{m2} O \xrightarrow{m} X + k_{m3} C \tag{6}$$

where S, O, X, C, Ac represent sugar (glucose), oxygen, biomass, carbon dioxide, and acetate respectively; μ_1, μ_2, and μ_3 are the specific growth rates, and m is the maintenance coefficient. In the sequel S, O, X, C, Ac mean concentrations.

The associated dynamical model is represented as follows:

$$\frac{d}{dt}\begin{bmatrix} X \\ S \\ Ac \\ O \\ C \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -k_1 & -k_2 & 0 & -1 \\ 0 & k_3 & -k_4 & -k_{m1} \\ -k_5 & 0 & -k_6 & -k_{m2} \\ k_7 & k_8 & k_9 & k_{m3} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ m \end{bmatrix} - D \begin{bmatrix} X \\ S \\ Ac \\ O \\ C \end{bmatrix} + \begin{bmatrix} 0 \\ DS_{in} \\ 0 \\ OTR \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ CTR \end{bmatrix} \tag{7}$$

where *CTR* is the carbon dioxide transfer rate from liquid to gas phase, *OTR* is the oxygen transfer rate from gas to liquid phase, and S_{in} is the influent glucose concentration.

It is commonly accepted (Han, *et al.*, 1992; Rothen, *et al.*, 1998; Walle and Shiloach, 1998; Kleman and Strohl, 1994) that the growth of *Escherichia coli* is very similar to the growth of *Saccharomyces cerevisiae* in the manner that the undesirable production of acetate (like the production of ethanol in *S. cerevisiae*) is only taken under two circumstances: the excess of available sugar for the bacteria or insufficient amounts of oxygen in the medium. The second case is easily explainable (acetate is produced anaerobically) and the explanation for the first is on the limited oxidative capacity of the bacteria.

This theory allows the elaboration of a restriction to the model: acetate cannot be consumed and produced at the same time. That is to say: the model can be divided in two sub-models, each one with three reactions: one with reactions 3, 4, and 6 (fermentative and oxidative growth on glucose, and maintenance) and the other with reactions 3, 5, and 6 (fermentative growth on glucose and acetate, and maintenance).

3. RESULTS

With the framework described above several fermentations were conducted with different objectives. In this section experimental results are reported to illustrate the capacities of the integrated system for advanced monitoring and control.

Process performance was first evaluated by running fed-batch fermentations to assess the dynamical behaviour. As a requisite for model-based operation fermentations were planned for identification of yield coefficients. Thereafter, state estimation was performed in new fermentations using software sensors integrating the identified yield parameters. Finally the maximization of biomass concentration at the end of fermentation was envisaged.

To evaluate the process performance fed-batch fermentations were conducted. This allowed us to assess the dynamical behaviour and also to validate all the experimental configuration setup.

As a first step towards the use of software sensors for state estimation and model-based control an identification procedure was implemented. The main objective of this identification task was the computation of yield coefficients through complete measurements of the state. The experimental design strategy was based in heuristic rules to program the substrate feed trajectories. Yield coefficients were computed by a linear regression technique without any knowledge about the kinetics being necessary, by a reparametrization (Chen, 1992) on the yield coefficients matrix.

Results from a validation of the identified yield coefficients are shown in Fig. 4. Estimated values of biomass and glucose concentrations were computed using the obtained yield coefficients and regressors calculated from the remaining state variables and transfer rates available from on-line measurements. These estimates are compared with off-line measurements.

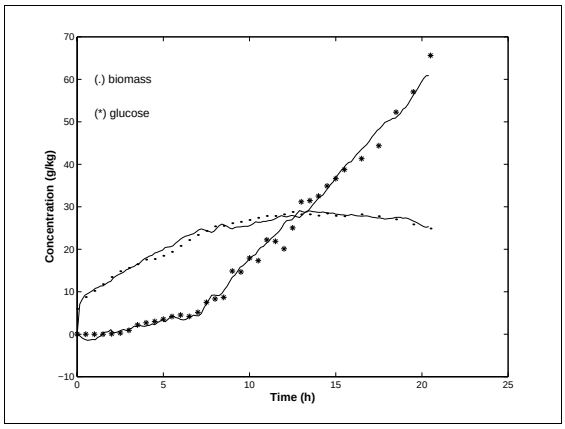

Fig. 4 Validation of a parameter set in yield coefficients identification. (.) biomass off-line measurements; (*) glucose off-line measurements; lines are the corresponding estimated values.

The assessment of biomass concentrations for control purposes at higher sampling frequency than the one obtained by off-line measurements lead us to design a software sensor to reconstruct biomass values from on-line measured variables. This state estimation was performed using an asymptotic observer (Bastin & Dochain, 1990) that integrates the identified yield parameters. Estimated biomass values are computed as follows:

$$\hat{X}_t = \hat{Z}_t - \frac{1}{k_3\left(k_7 k_{m2} - k_5 k_{m3}\right)}\Big[\left(k_8 k_{m2} - k_7 k_{m2} + k_5 k_{m3}\right) S_t - k_3 k_{m3} Ac_t - k_3 k_{km2} O_t\Big] \tag{8}$$

The caret (^) means on-line estimated value; the subscript t is a time index. In eq. (8) Z is an auxiliary variable resulting from a suitable state transformation without any knowledge of the growth rates being necessary. Z is computed in a discrete-time form obtained from a forward Euler discretization of a continuous-time differential equation.

$$\hat{Z}_{t+1} = \hat{Z}_t - TD_t\hat{Z}_t + \frac{T}{k_3\left(k_7 k_{m2} - k_5 k_{m3}\right)} \times \Big[\left(k_8 k_{m2} - k_7 k_{m2} + k_5 k_{m3}\right) D_t\, S_{in} - k_3 k_{km2} OTR_t\Big] \tag{9}$$

where T is the sampling period.

Fig. 5 compares on-line estimated biomass (doted line) with off-line values available from optical density measurements.

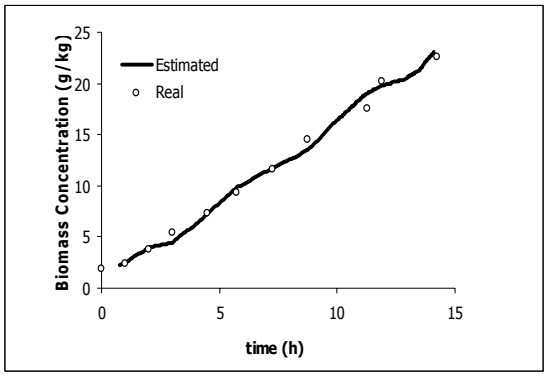

Fig. 5 Software sensor for biomass concentration estimation: comparison between estimated values (dotted line) and off-line data (points).

Since protein production is related to biomass growth, the maximization of biomass concentration at the end of fermentation was envisaged to maximize the final product concentration. To achieve this a controller was implemented to maintain the specific growth rate below its critical value while keeping the metabolic state in oxidative growth on glucose – eq. (3). The controller was designed taking into account the on-line information from estimated values of biomass concentration. This was achieved with exponential feeding as:

$$F_{in,t} = k_1 \mu_1^{crit} \hat{X}_t W_t \qquad (10)$$

being μ_1^{crit} a critical value of the specific growth rate for oxidative growth on glucose obtained experimentally.

The results obtained are shown in figures 6, and 7, where accumulation of acetate was kept under trace levels during 12 hours. Afterward this metabolite has raised but never reached toxic levels. This happened because the feeding strategy was able to maintain the glucose levels always very close to zero, and consequently, the specific growth rate was kept below its critical value. This controller made possible to reach a biomass concentration of 40 g/Kg a value that represents more than 53 g/L.

4. CONCLUSIONS

With the flexible graphical programming environment LabVIEW several data from different instruments were integrated during the fermentations of *E. coli* in a supervisory program. This framework permits the easily implementation of software sensors and control algorithms to improve process performance in a recombinant protein production.

Improvements are now being achieved in aspects related to model-based operation. The model calibration may be enhanced by optimising the richness of data coming out from the experiments, quantified by indexes related to the Fisher information matrix. Subsequently optimal experimental design can be performed in terms of the programming of input trajectories. Optimal adaptive control is aimed as a control law to further increase biomass and protein production.

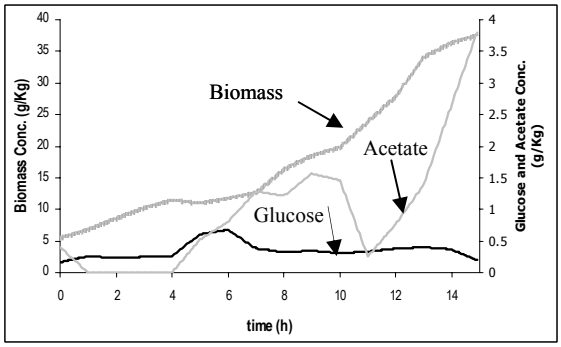

Fig. 6 Biomass, glucose and acetate concentrations during the course of the controlled fed-batch fermentation.

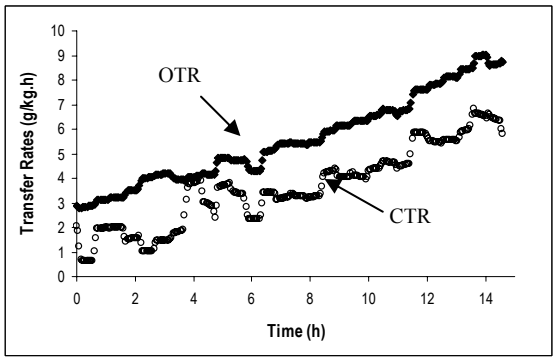

Fig. 7 Oxygen and Carbon Dioxide transfer rates during the course of the controlled fed-batch fermentation.

ACKNOWLEDGEMENTS

This work was supported by the PROTEXPRESS project (*Agência de Inovação*). *Fundação para a Ciência e a Tecnologia* provided financial support for I. Rocha through a doctoral research grant (PRAXIS XXI/BD/16961/98).

REFERENCES

Bastin, G., and D. Dochain (1990). *On-line Estimation and Adaptive Control of Bioreactors*. Elsevier Science Publishers, Amsterdam.

Chen, L. (1992) *Modelling, Identifiability and Control of Complex Biotechnological Systems*, Ph.D. thesis, Université Catholique de Louvain, Belgium.

Claes, J. E. and J. F. Van Impe (1998). On-line Monitoring and Optimal Adaptive Control of the Fed-Batch Baker's Yeast Fermentation. *7th International Conference on Computer Applications in Biotechnology*, Osaka 1998, 405-410.

Cockshott, A. and I. Bogle (1999). Modelling the effect of glucose feeding on a recombinant *E. coli* fermentation. *Bioprocess Engineering*, **20**, 83-90.

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

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

Galvanauskas, V., R. Simutis, N. Volk and A. Lübert (1998). Model Based Design of a Biochemical cultivation process. *Bioprocess Engineering*, **18**, 227-234.

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

Kellerhals, M., B. Kessler and B. Witholt (1999). Closed-Loop Control of Bacterial High-Cell-Density Fed-Batch Cultures: Production of mcl-PHAs by *Pseudomonas putida* KT2442 Under Single Substrate and Cofeeding Conditions. *Biotechnology and Bioengineering*, **65**, 306-315.

Kleman, G., and W. Strohl (1994). Acetate Metabolism by *Escherichia coli* in High-Cell-Density Fermentation. *Applied and Environmental Microbiology*, **60,** 3952-3958.

Levisaukas, D., V. Galvanaukas, R. Simutis and A. Lübert (1999). Model-based calculation of substrate/inducer feed-rate profiles in fed-batch processes for recombinant protein production. *Biotechnology Techniques*, **13**, 37-42.

Riesenberg, D., K. Menzel, V. Schulz, K. Schumann, G. Veith, G. Zuber, and W.A. Knorre (1990). High Cell Density Fermentation of Recombinant Escherichia coli Expressing Human Interferon Alpha1. *Applied Microbiology and Biotechnology*, **34**, 77-82.

Rothen, S. A., M. Sauer, B. Sonnleitner and B. Witholt (1998). Growth Characteristics of *Escherichia coli* HB101[pGEc47] on Defined Medium. *Biotechnology and Bioengineering*, **58**, 92-100.

Walle, M. and J. Shiloach (1998). Proposed Mechanism of Acetate Accumulation in Two Recombinant *Escherichia coli* Strains during High Density Fermentation. *Biotechnology and Bioengineering*, **57,** 71-78.