A general artificial neural network for the modelization of culture kinetics of different CHO strains

T. Marique¹, M. Cherlet¹, V. Hendrick¹, F. Godia², G. Kretzmer³ & J. Wérenne^{1*}

¹ Animal Cell Biotechnology CP 160/17, Université Libre de Bruxelles, avenue F. D. Roosevelt 50, 1050 Brussels, Belgium; ² Universitat Autònoma de Barcelona, Departament Enginyeria de Química, Spain; ³ University of Hannover, Institut für Technische Chemie, Germany (* Author for correspondence; E-mail: tmarique@skynet.be; Fax: +32-2-650.00.00)

Received 4 January 2001; accepted 20 August 2001

Key words: animal cells, artificial neural network, CHO, growth kinetics, modelization

Abstract

Animal cell cultures are characterized by very complex nonlinear behaviors, difficult to simulate by analytical modeling. Artificial Neural Networks, while being black box models, possess learning and generalizing capacities that could lead to better results. We first trained a three-layer perceptron to simulate the kinetics of five important parameters (biomass, lactate, glucose, glutamine and ammonia concentrations) for a series of CHO K1 (Chinese Hamster Ovary, type K1) batch cultures. We then tried to use the same trained model to simulate the behavior of recombinant CHO TF70R. This was achieved, but necessitated to synchronize the time-scales of the two cell lines to compensate for their different specific growth rates.

Introduction

Animal cell cultures are now widely used, either in research laboratories or for industrial productions (vaccines, transgenic proteins, etc.). Determining which culture conditions are best suited for a particular cell strain application is not an easy task, indeed animal cells grow slowly and often depend on complex and expensive media. Reliable simulations of animal cell culture kinetics would therefore be of great interest, as they could provide ways to save both time and money.

But such simulations are not easy to develop, because animal cell cultures display complex and highly nonlinear behaviors. We thus have to choose between two very different approaches. The first one uses classical analytical models. These offer the advantage that the model parameters have clear physical, chemical or biological meaning. Simplified equations easy to compute should be used, but they should not be too simplistic and are therefore limited to describe particular situations only. These problems have been particularly well explored by Bastin and Dochain (1990) and

applied with some success for animal cell cultures by Chotteau et al. (1992, 1991).

Another approach uses Artificial Neural Networks (ANN) models. These are 'black box' models, but endowed with learning and generalizing capacities which are best suited to our needs (Schalkoff, 1997; Willis et al., 1992).

We wanted to build an ANN that could predict the evolution of five important parameters in our CHO (Chinese Hamster Ovary) batch cultures, knowing only their initial values at the start of the culture. More precisely, we were interested in two continuously decreasing parameters (glucose and glutamine concentrations), in two continuously growing parameters (lactate and ammonia concentrations) and in one transiently growing parameter (biomass). Such an ANN would be valuable both to simulate the kinetics of culture for initial conditions still untried, and to predict future events (e.g. time of glucose or glutamine exhaustion, or time of maximum biomass production) in the course of an already running batch culture.

We developed simulations for two very different CHO strains. CHO K1 needs to be anchored on a

Figure 1. Three layers perceptron network with five inputs X , n_n neurons in the hidden layer and five outputs Y .

Figure 2. Structure of the simulation model. $\xi(k)_0$: initial inputs. $\xi(k)$ inputs. $\xi(k+1)$: outputs. NN: neural net. T: time-delay. Z^{-1} : conversion of outputs into next inputs.

Figure 3. Experimental data (symbols) and NN simulation (lines) for CHO K1 batch cultures, adimensional relative concentrations *versus* days in culture. A: validation data. B, C and D: learning data. Circles and thick line: biomass; triangles and interrupted line: glucose; diamonds and continuous line: glutamine; crosses and interrupted line: lactate; squares and continuous line: ammonia.

substrate to sustain good growth and viability. CHO TF70R are more easy to grow directly in suspension and they also produce recombinant t-PA. We wanted to know if an ANN trained with the data of the first strain could also provide correct simulation of the data of the second strain. This would of course greatly enhance the value of such an ANN if it could be valid for different cell strains.

Few cases of ANN application to microbial or animal cell culture modeling have been described up to now (di Massimo et al., 1992; Tholudur and Ramirez, 1996; Zhang et al., 1996). However, a radial basis three layers ANN has been demonstrated to give correct prediction for CHO K1 batch cultures (Graefe et al., 1999; Hanomolo et al., 2000). We choose to use a more classical sigmoid activation function, which allows simpler network training.

Material and methods

CHO cell cultures

We used two cell strains: CHO K1 (CCL 61, ATCC) and CHO TF70R: t-PA producing recombinant CHO, also capable of growth in suspension, and supplied to us by Dr. T. Björing (Pharmacia-Upjohn).

The CHO K1 were grown in Bellco spinner flasks and the CHO TF70R in Techne spinner flasks (500 ml total volume, 250 ml of medium) with a revolution speed of 50 rpm (30 rpm during the first 4 hours of culture). Microcarrier beads (Cytodex 3, 5 g l⁻¹) were used for CHO K1 only, CHO TF70R growing without attachment. Spinners were seeded at the density of 2×10^5 cells ml⁻¹ coming from standard TC flasks cultures. Standard medium for CHO K1 was DMEM/HamF12 (1:1) buffered with bicarbonate at 3.7 g l⁻¹, supplemented with 5% (v/v) FCS, 4 mM glutamine, 1% (V/V) non essential amino

Figure 4. Experimental data (symbols) and NN simulation (lines) for CHO TF70R batch cultures A, B and C, adimensional relative concentrations versus days in culture. Circles and thick line: biomass; triangles and interrupted line: glucose; diamonds and continuous line: glutamine; crosses and interrupted line: lactate; squares and continuous line: ammonia.

acids and 1% (v/v) antibiotics (stock solution 10000 U ml⁻¹ penicillin and 10000 µg ml⁻¹ streptomycin), all obtained from GIBCO-LIFE Technology, and 20 mM glucose (Sigma). For CHO TF70R, we used a serum-free medium (BIOPRO, Bio Whittaker). Spinners were incubated at 37 °C, without CO₂/O₂ control. Different initial glucose or glutamine concentrations were used for spinner cultures.

During culture, analysis of the culture medium were performed using the Sigma kits 315–100 for glucose and 735–10 for lactate, and the Boehringer Mannheim kit 139–092 for glutamine. Ammonium ions were determined using a HPLC column Pico-Tag from Waters or a specific electrode. Biomass was estimated by microscopic counting with a hemacytometer, using trypan blue to detect dead cells. We used living cells data for simulation with both CHO cell lines.

Artificial neural network

Raw data from the batch cultures were converted into percentage, relative to the highest value obtained for each parameter. We used data from 7 different batch cultures for both CHO K1 and CHO TF70R, differing in initial glucose or glutamine concentrations only. Initial glucose concentrations ranged between 9 and 30 mM. Initial glutamine concentrations ranged between 1 and 6.75 mM.

We used a three layers feed-forward perceptron ANN with sigmoid activation functions (Figure 1) for neurons of the hidden layer and linear functions for neurons of the output layer. The network inputs were the five initial values of the culture parameters. The outputs were the five values of the time derivatives of these same parameters. These derivatives were then simply multiplied by the value of the time-step and the result was added to the previous values of the parameters, to give the new input values (Figure

Figure 5. Experimental data with time-scale correction (symbols) and NN simulation (lines) for CHO TF70R batch cultures A, B and C, adimensional relative concentrations *versus* days in culture. Circles and thick line: biomass; triangles and interrupted line: glucose; diamonds and continuous line: glutamine; crosses and interrupted line: lactate; squares and continuous line: ammonia.

Table 1. Error value after learning, as a function of the number of neurons in the hidden layer

Number of hidden neurons	Error value
17	2.81
18	2.39
19	2.02
20	2.75

2). A Levenberg-Marquardt algorithm was applied for weights and biases determination.

Results and discussion

We first optimized our ANN for the CHO K1 cells grown on Cytodex 3, using five data series for training and two data series for cross-validation. All these

data series were different concerning initial glucose or glutamine concentrations. The minimum value of the cross-validation error function is obtained for 19 neurons in the hidden layer (Table 1). Introducing more neurons will lead to better fit with the learning data but only to decreasing performance in cross-validation (overfitting). Satisfying results are obtained (Figure 3), even for data series starting with low glutamine or glucose concentrations. These have biomass kinetics going through a maximum value and make them particularly difficult to learn by the neural network.

We then used the seven data series from CHO K1 to train the network, and performed cross-validation on the seven data series from CHO TF70R. The results were unsurprisingly disappointing, with poor fitting obtained (Figure 4).

When looking closer, it appeared that biomass and lactate concentration were systematically overestimated, and glucose and glutamine concentrations were systematically underestimated. This could result

from a slower metabolism of CHO TF70R compared to CHO K1, so they would grow, produce lactate and consume glucose and glutamine more slowly. We checked this assumption and found that the specific growth rate of CHO TF70R is indeed smaller than the one of CHO K1 by a factor of 0.72.

We thus tried again our network on the CHO TF70R data, after adjustment of the time scale. Precisely, the same model was run as before, now using a time-step value smaller by a factor of 0.72. The results were this time much better (Figure 5). It is worth noticing that the best correlation is obtained precisely for a time correction factor of 0.72.

Conclusions

Our results show that it is possible to obtain good modelization for CHO K1 batch culture parameters, using a simple perceptron ANN with one hidden neuron layer and sigmoid activation functions. Useful predictions can be gained, such as time of exhaustion of a medium constituent, or time of maximum biomass achievement.

Moreover, the same ANN can also be used to predict the behavior of another cell line, CHO TF70R, endowed with different characteristics (slower metabolism, free suspension growth, recombinant protein production). This can be achieved provided the time-step has been accurately adjusted, so that the two cell lines now share the same internal time-scale. The two cell lines obviously share the same metabolic schemes because they can be simulated using the same model with only a time scale correction, which happens to correspond precisely to the ratio of the specific growth rates of the two CHO cell lines.

In the future, it would be interesting to know if this ANN simulation model can be further extended to still other CHO strains, or even to other cell lines. This approach could also be used to determine which parameters are necessary for a good prediction, and thus which would be of fundamental importance to understand and control cell growth and metabolism.

Acknowledgements

This work was supported in part by the European Union Biotechnology Programme Framework IV 'Cell Factory'. We thank Dr R. Hanus for his critical comments.

References

- Bastin G & Dochain D (1990) *On Line Monitoring and Adaptive Control of Bioreactors*. Elsevier, New York.
- Chotteau V, Fabry L, Lowagie L & Wérenne J (1991) Biomass estimation in anchorage dependent animal cell culture. In: Spier RE, Griffith JB and Maignier B (eds.) *Production of Biologicals from Animal Cell Culture* (pp. 580–585) Butterworth Heinemann, Oxford.
- Chotteau V, Bastin G, Wérenne J & Fabry L (1992) Experimental validation of reaction mechanism for a class M animal cell culture. In: Spier RE, Griffith JB and Mc Donald C (eds.) *Animal cell technology: Developments, Processes and Products* (pp. 292–297) Butterworth Heinemann, Oxford.
- di Massimo C, Montague GA, Willis MJ, Tham MT & Morris AJ (1992) Towards improved penicillin fermentation via artificial neural networks. *Comp Chem Eng* 16(4): 283–291.
- Graefe J, Bogaerts P, Castillo J, Cherlet M, Wérenne J, Marenbach P & Hanus R (1999) A new training method for hybrid models of bioprocesses. *Bioprocess Engineering* 21: 423–429.
- Hanomolo A, Bogaerts Ph, Graefe J, Cherlet M, Wérenne J & Hanus R (2000) Maximum likelihood parameter estimation of a hybrid neural-classic structure for the simulation of bioprocesses. *Mathematics and Computers in Simulation* 51: 375–385.
- Schalkoff RJ (1997) *Artificial Neural Networks*. McGraw-Hill, New York.
- Tholudur A & Ramirez WF (1996) Optimization of fed-batch bioreactors using neural network parameter function model. *Biotechnol Prog* 12: 302–309.
- Willis MJ, Montague JA, Di Massimo C, Tham MT & Morris AJ (1992) Artificial neural networks in process estimation and control. *Automatica* 28: 1181–1187.
- Zhang BS, Tang R, Leigh JR, Dixon K & Hinge RD (1996) Improved quality and productivity in secondary metabolite fermentation through estimation, control and scheduling. In: *Proceedings IFAC, the 13th Triennial World Congress*, (pp. 437–442), San Francisco, USA.