Bioprocess Biosyst Eng (2006) 29:315–321
DOI 10.1007/s00449-006-0080-1

ORIGINAL PAPER

Improving the batch-to-batch reproducibility in microbial cultures during recombinant protein production by guiding the process along a predefined total biomass profile

Marco Jenzsch · Stefan Gnoth · Martin Kleinschmidt · Rimvydas Simutis · Andreas Lübbert

Received: 15 July 2006 / Accepted: 2 August 2006 / Published online: 6 September 2006

Abstract In industry *Escherichia coli* is the preferred host system for the heterologous biosynthesis of therapeutic proteins that do not need posttranslational modifications. In this report, the development of a robust high-cell-density fed-batch procedure for the efficient production of a therapeutic hormone is described. The strategy is to guide the process along a predefined profile of the total biomass that was derived from a given specific growth rate profile. This profile might have been built upon experience or derived from numerical process optimization. A surprisingly simple adaptive procedure correcting for deviations from the desired path was developed. In this way the batch-to-batch reproducibility can be drastically improved as compared to the process control strategies typically applied in industry. This applies not only to the biomass but, as the results clearly show, to the product titer also.

Keywords PAT · *E. coli* fed batch · Reproducibility · Recombinant proteins

M. Jenzsch · S. Gnoth · A. Lübbert (✉)
Zentrum für Bioverfahrenstechnik,
Martin-Luther-Universitaet, Weinbergweg 22,
c/o Biozentrum, 06120 Halle (Saale), Germany
e-mail: andreas.luebbert@biochemtech.uni-halle.de

M. Jenzsch
e-mail: marco.jenzsch@biochemtech.uni-halle.de

M. Kleinschmidt
Protein Research, Probiodrug AG, Weinbergweg 22,
06120 Halle (Saale), Germany

R. Simutis
Institute of Automation and Control Systems,
Kaunas University of Technology,
Studentu g. 48, 3028 Kaunas, Lithuania

Introduction

Biologics are known to be rather complex products. Apparently small changes in the manufacturing processes can cause significant differences in their clinical properties. Hence, production processes for biologics are approved by authorities only with clearly defined constraints on their manufacturing procedures. Consequently, reproducibility is of utmost importance. Additionally, reproducibility is very important as it affects the downstream processing and thus quality of the final product.

From the engineering point of view there are two challenges in guaranteeing batch-to-batch reproducibility. First of all, within the given constraints, the operational procedure, most robust with respect to typically appearing process fluctuations, must be found. And, secondly, while running the process along this robust path, the remaining randomly appearing disturbances must be eliminated by means of feedback control.

With respect to batch-to-batch reproducibility, production processes for recombinant proteins are lagging far behind most other industrial processes. Figure 1 shows a typical example of the repeatability of biomass concentration profiles in a recombinant protein production process. The variability is quite high.

This fact was recognized by the FDA. The agency responded with a couple of measures. One essential reaction is FDA's PAT initiative [2]. With PAT, improvements in pharmaceutical and biologics production processes with respect to real time automated process monitoring and control are demanded. A rigorous science-based approach to manufacturing is demanded, as better understanding is thought to

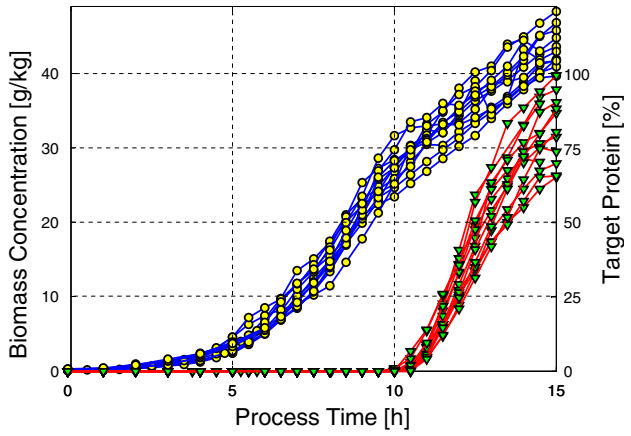

Fig. 1 Biomass (*open circles*) and target protein (*open triangles*) concentration profiles from 13 fed-batch fermentations for the production of a recombinant protein. Typically the batch-to-batch reproducibility of these production processes is rather low

lead to more efficient process control, lower process variability, thus high product quality and finally patient safety. For biologics it is particularly important to keep the processes under control early in the product synthesis process (e.g., [1]).

In bioprocess engineering, process monitoring and control is being developed since many years. A review was given by Lee et al. [8]. The objective was to keep the process on trajectories that finally lead to a high value of the desired mass m_P of the product. Within a more or less well-defined production time t_P, this can be related to biomass by the following expression:

$$m_P = \int_0^{t_P} \pi x \, dt, \tag{1}$$

where π is the specific product formation rate and x the total biomass. Both arguments of the integral are functions of time, but more importantly, both are primarily dependent on the specific biomass growth rate μ in most industrial production systems. The growth rate that a specific medium supports determines the physiological state of the cells, and particularly the cell's protein-synthesizing machinery that is important to recombinant protein production is also under growth rate control [12]. Hence, the variable that rules the final outcome of the process is the specific biomass growth rate μ. Thus, design of effective cultivation processes should be based on an optimal or at least a quasi-optimal profile of the specific growth rate. This can be obtained by means of numerical optimization procedures (e.g., [10]) or simply by deriving a profile from data records and experience with the production system under consideration. In order to make sure that the process follows this profile it is straightforward to control μ in the engineering sense (e.g., [5–7, 11, 13, 14, 16–18]).

Direct control of μ works perfectly as long as there are no severe disturbances in the process. When, however, some disturbances lead to a significant deviation of the biomass from its desired path, one must correct it before one can proceed with the desired optimal or quasi-optimal μ profile. Otherwise reproducible process trajectories cannot be obtained.

The decisive innovation in this paper is to show that the batch-to-batch reproducibility of the production processes can be significantly enhanced when the process is controlled to a predefined profile of the biomass x. This does not mean to stay away from keeping the μ profiles found to be optimal from the physiological point of view. The x-setpoint profiles are simply derived from predefined μ profiles. Using start biomass and desired specific biomass growth rate profile it is easy to estimate total biomass profile during the cultivation:

$$\frac{dx}{dt} = \mu_{set} x. \tag{2}$$

A given biomass profile $x_{set}(t)$ is then in close relationship with the specific biomass growth rate μ. Hence, controlling the process to an x profile should satisfy the corresponding specific growth rate profile $\mu_{set}(t)$ as well. In this case the cultivation process is more robust because the deviations in biomass concentration can be eliminated directly by controlling the integral variable x.

As x cannot be measured directly with sensors that work reliably at a production fermenter and provide biomass values representative for the culture, it should be measured indirectly. This is a further advantage as compared to the μ-control because we are able to estimate x much more reliably than the specific growth rate μ [4]. In production environments, where sufficiently many data records are available from the process under consideration, artificial neural networks (ANNs) yield very accurate estimates. Hence we use ANNs to estimate biomass x. The ANN was trained on 26 data sets measured during a process development project with the strain used in this work.

Materials and methods

Experiments were performed with *Escherichia coli* BL21(DE3) as the host cell. The recombinant target

protein was coded on the plasmid pET 28a and expressed under the control of the T7 promoter after induction with isopropyl-thiogalactopyranosid (IPTG). The strain was resistant to kanamycin. The product appears as inclusion body within the cytoplasm. The particular strain used did not produce measurable amounts of acetate under the cultivation conditions adjusted in the experiments reported. This was tested in the beginning by means of appropriate test kits.

The main substrate was glucose. It was fed at a concentration of 600 g/kg. The other components are compiled in Table 1.

All experiments were performed within BBI Sartorius System's BIOSTAT® ED 15 L bioreactor with 8 L working volume (Fig. 2). The fermenter was equipped with standard 6-blade Rushton turbines that could be operated at up to 1,400 rpm. The aeration rate could be increased up to 24 L/min. Aeration rate and then stirrer speed were increased one after the other in order to keep the dissolved oxygen concentration at 25% saturation. In order to suppress foam formation, increase oxygen solubility and reduce the risk of contamination, the fermenter head pressure was kept at 0.5 bar above the ambient pressure in the laboratory.

The fermentations were operated in the fed-batch mode from the early beginning with an initial volume of 5 L at pH 7 and a temperature of 35°C. All fermentations were started in the night by automatic transfer of the inoculation biomass from a refrigerator into the reactor. Substrate feeding starts immediately after inoculation with a fixed exponential profile. According to Jenzsch et al. [3], the specific growth rate setpoint was chosen as 0.5 h^{-1}. Thus the glucose concentration appeared to be at a very low value just from the beginning on. Its value is in the order of the K_s value and thus below the values that can be measured during the fermentation. After the biomass concentration reached values of about 35 g/kg, the culture was induced with 1 mM IPTG. From the seventh fermentation hour on, the growth rate was kept under feedback control along an x profile that was derived from a predefined profile of the specific growth rate. The corresponding profile of the substrate feed rate was determined from the profiles of μ and x. This was taken as reference feeding profile F_{ref} for all the fermentations described.

Temperature was measured with a Pt-100, pH with an Ingold-pH-probe, pO_2 with an Ingold pO_2-Clark-electrode, CO_2 in the vent line with MAIHAK®'s Unor 610 and O_2 there with MAIHAK®'s Oxor 610. Further, the total ammonia consumption during pH control was recorded by means of a balance beneath the base reservoir. All these quantities were measured online. Additionally, enhanced foam levels could be detected with a foam sensor and, if the critical level

Table 1 Composition of the mineral medium

Mineral salt solution		Trace element solution	
Component	Concentration (g/kg)	Component	Concentration (g/kg)
K_2HPO_4	14.60	Na_2-EDTA	20.10
$NaH_2PO_4 \cdot H_2O$	3.60	$FeCl_3 \cdot 6H_2O$	16.70
$(NH_4)_2SO_4$	2.46	$CaCl_2 \cdot 2H_2O$	0.74
Na_2SO_4	2.00	$CoCl_2 \cdot 6H_2O$	0.21
$MgSO_4 \cdot 7H_2O$	1.20	$ZnSO_4 \cdot 7H_2O$	0.18
$(NH_4)_2$-H-citrate	1.00	$CuSO_4 \cdot 5H_2O$	0.10
NH_4Cl	0.50	$MnSO_4 \cdot H_2O$	0.10
Kanamycin	0.10		
Thiamin	0.10		
Trace element solution	2 mL/kg		

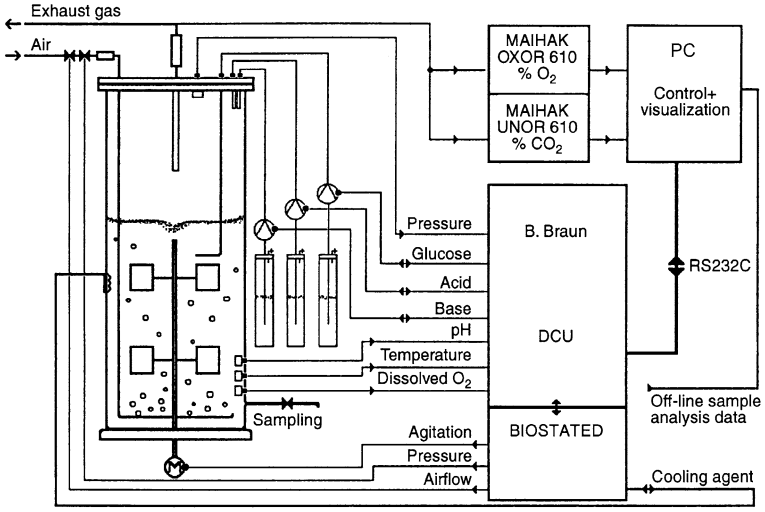

Fig. 2 Experimental setup of the cultivation equipment

was reached, a silicone antifoaming emulsion (ROTH®) was added.

Biomass concentrations were measured offline (via the optical density at 600 nm) with a Shimadzu® photo-spectrometer (UV-2102PC). In preceding experiments a correlation was established between these values and the biomass dry weights, which were determined with the standard drying and weighting technique. Glucose was determined enzymatically with a YSI 2700 Select Bioanalyzer. The product was measured with SDS-PAGE after separation of the inclusion bodies and their solubilization.

Results

Preliminary simulations and experiments showed that the total biomass $x(t)$ is better suited as the controlled variable than the biomass concentration $X(t)$ itself. The control can then be performed with a simple adaptive control algorithm comprising the following steps:

1. Estimate the total biomass x_{est} at the current time instant t using the artificial neural network trained before on the available data records [4]. A simple feedforward artificial neural network with a single hidden layer was used which makes use of the online measured signals OUR, CPR, total base consumption resulting from pH control. With an additional bias node in the input layer the network has four input nodes. Five hidden layer nodes were used where the nonlinear response function of these nodes was chosen to be tangens hyperbolicus (tanh) functions. A single output, namely the biomass x, was taken, where the weighted sum of its input signals was directly used as the output signal.
2. Compute the deviation of estimated biomass x_{est} from its setpoint x_{set}:

$$\Delta x = x_{\text{set}} - x_{\text{est}}. \tag{3}$$

3. The deviation Δx in x was used to correct the growth yield value Y_{xs} in the feed function, which itself was limited to the interval given in order to avoid too big variations in cases where the measurements are disturbed in an unusual way:

$$F = \frac{\mu_{\text{set}} x_{\text{est}}}{(Y_{\text{xs}} - \alpha) S_{\text{f}}} \quad \text{with } 0.7F_{\text{ref}} \le F \le 1.3F_{\text{ref}}. \tag{4}$$

The concrete limits to F were chosen by experience in control practice.

4. In order to determine the controller variable α by which the growth yield is adapted to the current state of the process we found that it is better to make use not only of the deviation Δx but also the smoothing action of an integral correction term to the deviation:

$$\alpha = k_1 \Delta x + k_2 \int_{t_S}^{t} \Delta x \, \mathrm{d}t \quad \text{with} -0.15 \le \alpha \le 0.15. \tag{5}$$

Again, the correction was limited. Here the limitation was chosen by experience in such a way that the resulting yield values are kept within reasonable limits. The parameters k_1 and k_2 were initially determined in simulation studies. Later, during preliminary control experiments, they were slightly adapted. Their final values are $k_1 = 0.1$ (kg(S))$^{-1}$ and $k_2 = 0.02$ (kg(S))$^{-1}$ h^{-1}.

The controlled biomass profiles depicted in Fig. 3 show that this simple control approach leads to a very good reproducibility of the total biomass profiles.

As one is usually not interested in the total biomass, the corresponding biomass concentrations are depicted in Fig. 4. As can easily be seen from both plots, the batch-to-batch variability of the trajectories is rather small, i.e., the process total biomass control works well.

In order to get a better impression of the controller action, the deviations from the desired profile, i.e., the relative deviations between the total biomass and its setpoint, are plotted in Fig. 5. In all the experiments, the controller was switched on at $t = 7$ h. In the first 4 h thereafter, the relative deviations remain within a 5% interval. The controller action then improves so that

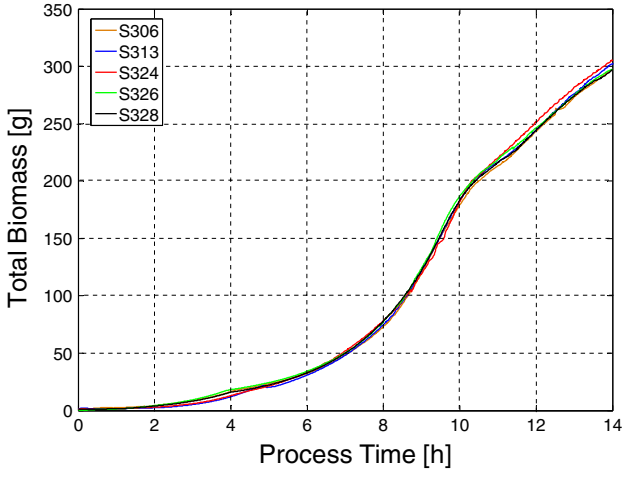

Fig. 3 Total biomass signals from five fermentations performed sequentially using the same setpoint profile

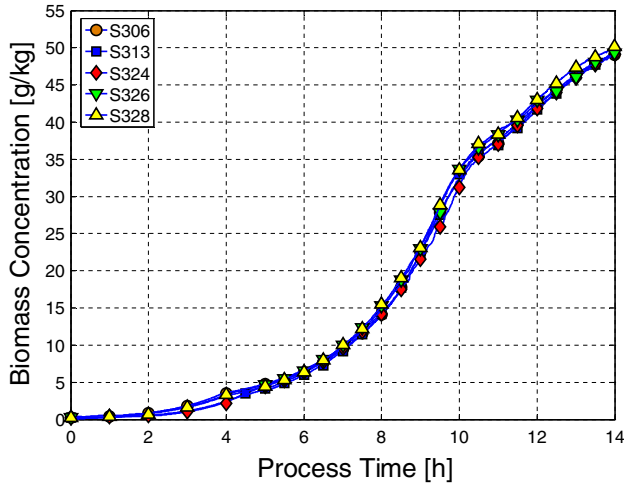

Fig. 4 Measurements of the biomass concentration during five experiments in which the total biomass was controlled. The induction time was t_{ind} = 11 h in all cases

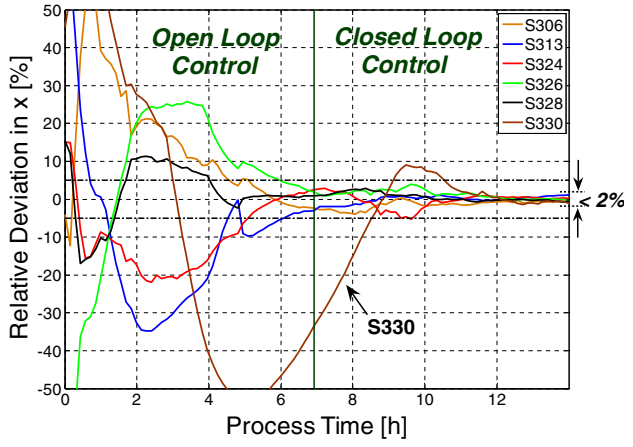

Fig. 5 Relative deviations of the total biomass from the mean. The controller was switched on 7 h after the cultivation was inoculated

finally the relative deviation of the total biomass remains within a 2% interval around the mean. The feed rate profiles $F(t)$ applied in the experiments are shown in Fig. 6.

There is one exception. In the last experiment (S330) feed pump was switched off from 3 to 5 h in order to test the controller performance under process conditions with an extremely hard disturbance. The controller appeared to be robust enough to cope with this disturbance. Again in the end the relative deviation from the mean remained in the 2% interval.

The next question is what does this mean for the profiles of the specific biomass growth rate, which was initially chosen to determine the setpoint profiles for

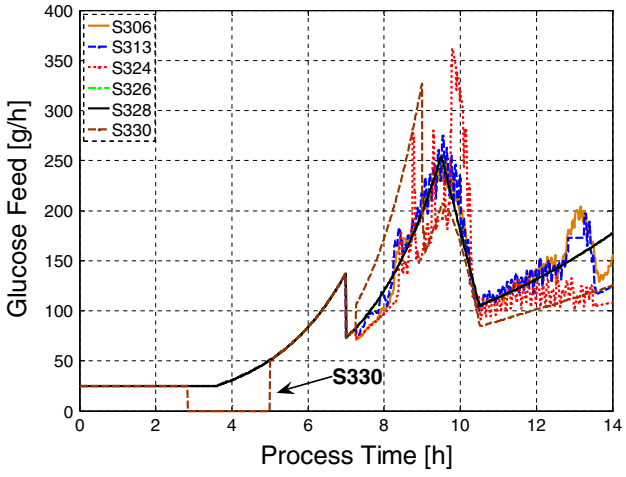

Fig. 6 Substrate feed rate profiles for the six x-controlled fed-batch experiments

the total biomass. This comparison essentially should demonstrate data consistency. The specific growth rate μ was determined from a Luedeking–Piret type relationship between the biomass concentration, carbon dioxide production or oxygen uptake rate and the specific growth rate:

$$\mathrm{OUR} = Y_{\mathrm{OX}}\mu X + m_{\mathrm{O}}X, \tag{6}$$

$$\mathrm{CPR} = Y_{\mathrm{CX}}\mu X + m_{\mathrm{C}}X. \tag{7}$$

Both equations can be resolved for μ. The parameters were already known from a fit of the models to the fermentation data.

Hence, with the biomass concentration profiles estimated and the measured OUR(t) and CPR(t) signals, μ can easily be computed. The result is depicted in Fig. 7. In the decisive phase of the process, the product formation phase, the resulting trajectories of the specific biomass formation rate $\mu(t)$ resemble quite well the original setpoint profiles. It should be recalled that rather high noise is to be expected when computing $\mu(t)$. However, the results depicted in Fig. 7 nevertheless clearly show that the data are consistent.

Finally the question arises what does a high batch-to-batch reproducibility in the biomass profiles mean for the variance in the product formation profiles. Results corresponding to the data shown before are depicted in Fig. 8.

The corresponding protein concentration profiles are rather close together saying that the improved reproducibility in biomass profiles by means of x control also leads to an improved batch-to-batch reproducibility in the product concentration profiles. All profiles stay within the confidence interval of the

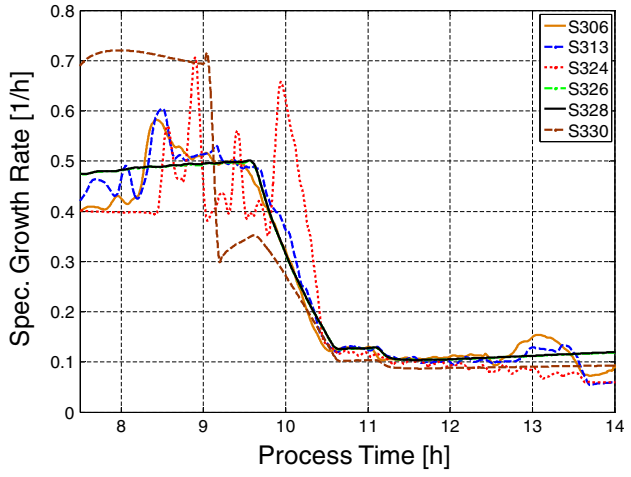

Fig. 7 Specific biomass growth rate profiles determined from the measurement profiles of OUR, CPR as well as from the estimates for biomass concentration $X(t)$. Note that the study 330 was the test case where the substrate supply was interrupted for some period during the biomass growth phase

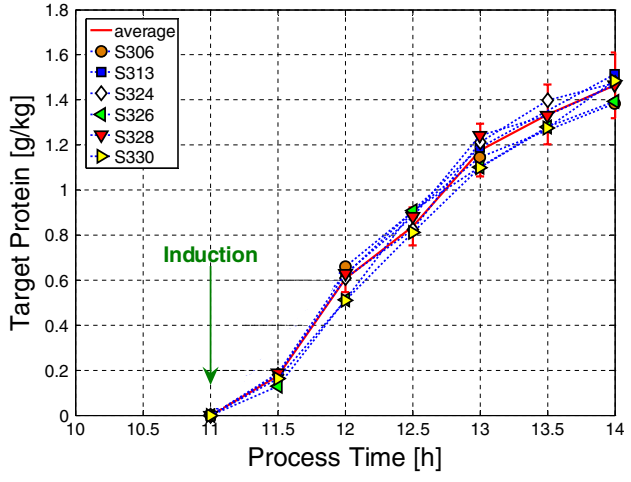

Fig. 8 Product concentration profiles for the cultivation processes already mentioned. All protein data stay within the error bar ranges representing the confidential interval of the protein analysis method

analytical protein detection method. The relative error of ±15% of target protein analysis by SDS-PAGE was estimated from a fivefold analysis using the same fermentation samples.

Discussion

Previous work on controlling the fermentation processes performed to manufacture recombinant proteins focused on keeping the cultures on tracks that guaranteed a high productivity or yield. The obviously first quantity influencing the amount of protein that is produced in a cultivation run is the biomass employed, hence high-cell-density cultures are required (e.g., [9, 15]). As the cells' protein-synthesizing machinery that is particularly important to recombinant protein production is under growth rate control [12], the performance of the cells is ruled by the specific biomass growth rate μ. Hence, much work has been put into closed loop control of fermentation processes along appropriate profiles $\mu_{set}(t)$ of the specific biomass growth rate [5–7, 11, 13, 16–18]. While advantageous from the cell physiological point of view, controlling the specific growth rate directly has the disadvantage of leading to a relatively low reproducibility of the fermentations [6]. This is a significant disadvantage from the process quality point of view.

We wished to extend the work on fermentation control towards quality assurance of process and thus product formation. This first of all requires improving the batch-to-batch reproducibility of the processes. There are two motivations for this. First, the product quality in recombinant protein manufacturing processes can be affected by changes in the fermentation operational procedure, hence the authorities link process approval with tight constraints on the process trajectories. Thus, good reproducibility increases product quality. Secondly the downstream processing can work much more efficiently when the cultivation results are highly reproducible. Therefore, having the same culture each time should be beneficial to the overall product yield as well.

The results of the new control procedure discussed below show that the batch-to-batch reproducibility is significantly improved. The procedure controls the biomass along a profile that is directly derived from the specific growth rate profile that was found to be desirable. The first decisive advantage of this approach is that the control procedure is robust to distortions usually appearing in practice. This is opposite to control along a predefined profile of the specific biomass growth rate. Practically all distortions lead to changes in the biomass and this is the controlled variable. A second advantage is that the actual biomass values can be estimated with higher accuracy than the specific growth rate from the online measurement signals available at most fermenters. As shown previously [4] x can be representatively estimated with a root mean square error of about 0.5 g/kg by means of simple artificial neural networks.

The reason for the robustness of the x control is easy to understand. The x-setpoint profile for the control was determined from a μ profile. This μ profile leads to a feed rate profile $F(t)$ and a biomass profile $x(t)$. The

$F(t)$ profile is taken as the feed forward component of the process and programmed in the programmable controller at the fermenter. If due to distortion the biomass becomes smaller as expected, the cells will see more substrate than expected and respond temporarily by increasing their growth rate. If, on the other hand, the distortion leads to a higher biomass concentration, the cells see less substrate than expected and reduce their growth rate, thus correcting for their deviation in the total biomass. The feedback component of the process has thus only to correct for larger deviations.

Control of the specific biomass growth rate or qualitatively similar control technique that uses the substrate feed rate F as the manipulated variable is currently not generally used in production practice. Its adoption is now supported by recent initiatives of the FDA in its PAT initiative [2], but will critically depend on the ease of the control algorithm. The algorithm used in this work is an extremely simple adaptive control approach which only corrects the biomass-on-substrate yield in the basic feed rate expression. The control algorithm proved to be stable in many fermentation runs with many different organisms and several different fermenter scales up to the large production scale.

It is very important to note again that the new x control approach does not require refraining from taking the μ profile of the process as the basic process control variable. This basic concept can be maintained and it is easy to derive the x profile from a given μ profile. As clearly shown in the examples, not only does this x-based control keep biomass and specific growth rate in tight limit, but also the batch-to-batch reproducibility with respect to the desired product concentration is high.

Acknowledgments This work has been supported financially by the Ministry of Cultural Affairs of the state ''Sachsen-Anhalt'', Germany. We gratefully acknowledge this support.

References

1. DePalma A (2004) PAT: taking process monitoring to next level. Gen Eng News 24(9):46–47
2. FDA (2004) Guidance for Industry: PAT—a framework for innovative pharmaceutical manufacturing & quality assurance. http://www.fda.gov/cvm/guidance/published.html
3. Jenzsch M, Gnoth S, Beck M, Kleinschmidt M, Simutis R, Lübbert A (2006) Open loop control of the biomass concentration within the growth phase of recombinant protein production processes. J Biotechnol (in press)
4. Jenzsch M, Simutis R, Eisbrenner G, Stückrath I, Lübbert A (2006) Estimation of biomass concentrations in fermentation processes for recombinant protein production. Bioprocess Biosyst Eng 29(1):19–27
5. Jenzsch M, Simutis R, Lübbert A (2005) Application of model predictive control to cultivation processes for protein production with genetically modified bacteria. In: Pons MN, van Impe JFM (eds) Computer application in biotechnology 2004 (CAB9). IFAC/Elsevier, Amsterdam, pp 511–516, ISBN 0-08-044251-X
6. Jenzsch M, Simutis R, Lübbert A (2006) Generic model control of the specific growth rate in recombinant *Escherichia coli* cultivations. J Biotechnol 122(4):483–493
7. 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
8. Lee J, Lee SY, Park S, Middelberg APJ (1999) Control of fed-batch fermentations. Biotechnol Adv 17:29–48
9. Lee SY (1996) High cell-density culture of *Escherichia coli*. Trends Biotechnol 14:98–105
10. Levisauskas D, Galvanauskas V, Henrich S, Wilhelm K, Volk N, Lübbert A (2003) Model-based optimization of viral capsid protein production in fed-batch culture of recombinant *Escherichia coli*. Bioprocess Biosyst Eng 25:255–262
11. Levisauskas D, Simutis R, Borvitz D, Lübbert A (1996) Automatic control of the specific growth rate in fed-batch cultivations processes based on exhaust gas analysis. Bioprocess Eng 15(3):145–150
12. Neidhardt FC, Ingraham JL, Schaechter M (1990) Physiology of the bacterial cell, a molecular approach. Sinauer, Sunderland
13. Picó-Marco E, Picó J, De Battista H (2005) Sliding mode scheme for adaptive specific growth rate control in biotechnological fed-batch processes. Int J Control 78(2):128–141
14. Pirt SJ (1975) Principles of microbe and cell cultivation. Blackwell, London, pp 115–117
15. Riesenberg D, Guthke R (1999) High-cell-density cultivation of microorganisms. Appl Microbiol Biotechnol 51:422–430
16. Shioya S (1992) Optimization and control in fed-batch bioreactors. Adv Biochem Eng Biotechnol 46:111–142
17. Soons ZITA, Voogt JA, van Straten G, van Boxtel AJB (2006) Constant specific growth rate in fed-batch cultivation of *Bordetella pertussis* using adaptive control. J Biotechnol, published online 10-07-2006
18. Yoon SK, Kang WK, Park TH (1994) Fed-batch operation of recombinant *Escherichia coli* containing Trp promoter with controlled specific growth rate. Biotechnol Bioeng 43:995–999