

Microbial Cell Factories

**Open Access**

Research

Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris*

Michael Maurer¹, Manfred Kühleitner², Brigitte Gasser¹ and Diethard Mattanovich*^{1,3}

Address: ¹University of Natural Resources and Applied Life Sciences Vienna, Department of Biotechnology, Institute of Applied Microbiology, Vienna, Austria, ²University of Natural Resources and Applied Life Sciences Vienna, Department of Integrative Biology, Institute of Mathematics, Vienna, Austria and ³School of Bioengineering, University of Applied Sciences FH-Campus Vienna, Austria

Email: Michael Maurer - michael.maurer@boku.ac.at; Manfred Kühleitner - manfred.kuehleitner@boku.ac.at; Brigitte Gasser - brigitte.gasser@boku.ac.at; Diethard Mattanovich* - d.mattanovich@iam.boku.ac.at

* Corresponding author

Published: 11 December 2006

Received: 25 October 2006

Microbial Cell Factories 2006, **5**:37 doi:10.1186/1475-2859-5-37

Accepted: 11 December 2006

This article is available from: <http://www.microbialcellfactories.com/content/5/1/37>

© 2006 Maurer et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Secretion of heterologous proteins depends both on biomass concentration and on the specific product secretion rate, which in turn is not constant at varying specific growth rates. As fed batch processes usually do not maintain a steady state throughout the feed phase, it is not trivial to model and optimize such a process by mathematical means.

Results: We have developed a model for product accumulation in fed batch based on iterative calculation in Microsoft Excel spreadsheets, and used the Solver software to optimize the time course of the media feed in order to maximize the volumetric productivity. The optimum feed phase consisted of an exponential feed at maximum specific growth rate, followed by a phase with linearly increasing feed rate and consequently steadily decreasing specific growth rate. The latter phase could be modeled also by exact mathematical treatment by the calculus of variations, yielding the explicit shape of the growth function, however, with certain indeterminate parameters. To evaluate the latter, one needs a numerical optimum search algorithm. The explicit shape of the growth function provides additional evidence that the Excel model results in correct data. Experimental evaluation in two independent fed batch cultures resulted in a good correlation to the optimized model data, and a 2.2 fold improvement of the volumetric productivity.

Conclusion: The advantages of the procedure we describe here are the ease of use and the flexibility, applying software familiar to every scientist and engineer, and rapid calculation which makes predictions extremely easy, so that many options can be tested *in silico* quickly. Additional options like further biological and technological constraints or different functions for specific productivity and biomass yield can easily be integrated.

Background

Modeling of bioprocesses has been pursued since the 1970s, with the aim to rationally optimize processes. While the mathematical description of processes like growth and product formation have been fairly well achieved, it is still not routine practice to design biotechnological production processes based on model prediction. An especially difficult case in this respect is fed-batch as a dynamic system usually not reaching steady state. General attempts to model fed-batch processes have been described (for an overview see [1]). Based on these modeling approaches, optimization of fed batch processes has been attempted using Pontryagin's Maximum Principle [2,3], Green's Theorem [4], or Dynamic Programming [5]. These approaches are rather complex, and they did not find their way in routine application.

A typical case of fed-batch process is the production of recombinant proteins with microorganisms or mammalian cells. While the description of product concentration in the cell mass is rather straight forward (in the case of an intracellular product), it is more complex to predict the kinetics of a secreted product. A typical case for secretion systems are recombinant yeasts [6]. As the production of many proteins in yeasts is quite cost sensitive, it will be highly desirable to have a tool available that allows a simple yet reliable prediction of productivity, process time and product titers. Approaches to optimize fed batch processes for the methylotrophic yeast *Pichia pastoris* have been described [3,5]. The latter employ dynamic programming by dividing the total process time into a discrete number of intervals, and assigning a value of the specific growth rate μ selected from a discrete set of values. The major drawback of this approach is that the process time is fixed and not an issue of optimization. The algorithm used by this group is complex, and not readily available to others. Zhang and coworkers [3] present an approach based on Pontryagin's Maximum Principle. A general applicability seems hampered by the complex calculation, complicating a simple recalculation with modified data or calculation procedures.

With this work we aimed at the development of an optimization tool for fed batch processes using calculation tools available for every PC. MS Excel allows the approximation of a model by numerically solving equations describing the system, and the optimization of an objective function by modifying defined fields (the decision variables), while different constraints can be defined which have to be complied with. The calculation is based on the generalized reduced gradient method described in [7]. While the general concept of calculation is similar to the approaches above, the definition of the optimization objective – while obviously a crucial step – is not consistently resolved in the existing literature. The variable costs of a bioprocess correlate with the volumetric capacity of

the required fermentation unit, and the process time this unit is required to produce a defined amount of the product [8]. Thus the volumetric productivity Q_p is the most plausible target for optimization. At a given process time point t , Q_p is defined as:

$$Q_p = \frac{P}{V \cdot t} \quad (1)$$

Expanding this concept to total manufacturing costs is feasible but depends on a profound and reliable cost calculation. As outlined below, Q_p can be calculated from the specific growth rate μ and the specific production rate q_p . μ should be one of the decision variables of the optimization (defining the feed rate profile to be developed), while q_p depends on μ . The exact function, $q_p = f(\mu)$, of this dependence for secreted recombinant proteins has been subject to discussion [9]. These authors provide some evidence that secreted protein productivity is saturated at high μ , but a clear experimental solution of this function and its biological basis has not been achieved yet. Zhang et al. approximate this relation by an empirical 3rd order function [3], while Ohya and coworkers model it by a two step linear function [10]. Both groups base their model functions on rather few experimental data. To improve the accuracy of the model $q_p = f(\mu)$, we examined the entire space of μ of a *P. pastoris* strain in chemostat cultures for the respective values of q_p , as well as the observed biomass yield coefficient Y'_{XS} in order to calculate the substrate needed for each increment of biomass increase. It has been discussed whether data derived from steady state are applicable to model transient situations as they usually occur in fed batch. The parameters determining the accuracy of a steady state model are the relaxation time constants of environmental changes and of biological processes based on the change in environment, which become critical at highly transient situations like a shift to growth limiting conditions at the end of batch [11]. However, substrate limited fed batch cannot be considered as highly transient, so that the steady state model should be applicable.

A *P. pastoris* strain expressing the Fab fragment of the anti-HIV antibody 2F5 [12] was employed as a model. As expression is based on the glyceraldehyde phosphate dehydrogenase (GAP) promoter, glucose is used as a substrate for growth. Modeling of the fed batch process and optimization of Q_p was used to predict an optimal feed protocol, which was then evaluated experimentally. The model optimization was also solved analytically in order to prove the accuracy of the Excel approximation.

Results and discussion

Chemostat

The 2F5 expression strain was cultivated in chemostat at dilution rates D between 0.0086 h⁻¹ and 0.2 h⁻¹. Steady

state samples were taken after 5 volume changes each. The setpoints were passed through once from high to low dilution rates, and once from low to high dilution rates. Specific production rates q_p and observed biomass yield coefficients Y'_{XS} are plotted against $D = \mu$ in figure 1. The constants of eq (22), describing q_p , were derived by the method of least squares as $q_{pmax} = 0.0735 \text{ mg g}^{-1} \text{ h}^{-1}$ and $k_q = 0.116 \text{ h}^{-1}$. $Y_{XS} = 0.559$ and $m_S = 0.0161 \text{ h}^{-1}$ were derived according to eq (26). The estimated standard deviation of q_p is $s_q = 0.0048 \text{ mg g}^{-1} \text{ h}^{-1}$, and that of Y'_{XS} is $S_Y = 0.023$.

Standard fed batch

Two independent fed batch cultures using a standard protocol with constant feed rate [12] were performed. The final Fab titer was $p = 46 \text{ mg L}^{-1}$, and the final biomass concentration $x = 96 \text{ g L}^{-1}$, both at a total process time $t = 117 \text{ h}$ (92 h feed). Fig. 2A shows the development of these parameters over time, while Q_p and q_p are plotted in Fig. 2B. Apparently Q_p has a maximum of $0.31 \text{ mg L}^{-1} \text{ h}^{-1}$ at $t = 94 \text{ h}$ (69 h feed).

Optimized fed batch – model and experimental

Using the optimization algorithm described in Materials and Methods, a feed protocol leading to maximum volumetric productivity was determined. As the maximum final biomass concentration was set to 100 g L^{-1} , and the feed medium was identical to the standard fed batch, almost the same biomass concentration and total feed volume was to be expected for both processes. Optimal μ and feed rate is plotted over time in Fig. 3. The feed starts with an exponential phase of 3.6 h, followed by a 16 h phase with more slowly increasing feed, which was approximated by a linearly increasing feed rate following the function

$$F_L = 0.012 \text{ g} \cdot \text{h}^{-2} \cdot t_L + 30.672 \text{ g} \cdot \text{h}^{-1} \quad (2)$$

All values were calculated for a unit batch volume of 1 L and needed to be adjusted to the respective batch volume of 1.2 L. The modeled process was verified experimentally in two independent fed batch cultures. The resultant plots of biomass and product concentrations, as well as the predicted values, are displayed in Fig. 4A, while Q_p and q_p and their model prediction are shown in Fig. 4B. A final product concentration of 45 mg L^{-1} was reached after 21 h feed at a biomass concentration of 94 g L^{-1} . The maximum volumetric productivity Q_p was $0.67 \text{ mg L}^{-1} \text{ h}^{-1}$ (slightly below the predicted value of $0.77 \text{ mg L}^{-1} \text{ h}^{-1}$), which is a 2.2 fold improvement over the standard fed batch.

Modeling of the standard fed batch

Using the same equations as for optimization we also attempted to model the standard fed batch process. However, the predicted values of biomass and product concentrations deviated significantly from the experimental data. Therefore we reconsidered the data source used to obtain the functions of Y'_{XS} and q_p based on μ . The obvious difference between the standard and optimized fed batch protocol is that the standard culture is performed at very low μ from 0.05 h^{-1} decreasing to 0.005 h^{-1} , while the optimized culture starts at $\mu = 0.2 \text{ h}^{-1}$, decreasing to 0.05 h^{-1} . As saturation functions like the Monod function are more susceptible to low values of the x-coordinate, and the majority of the data from chemostat were naturally obtained at higher μ values, we remodeled the function $q_p = f(\mu)$ for low μ based on data derived from previous fed batch cultures. The best approximation in the range of $\mu \leq 0.05 \text{ h}^{-1}$ was a linear function

$$q_p = 0.2051 \cdot \mu + 0.002 \quad (3)$$

Similarly, Y_{XS} and m_S were remodeled in this range of $\mu \leq 0.05 \text{ h}^{-1}$.

Based on these refined approximations, the predictions of biomass and product concentration, volumetric productivity and specific product formation rate fit well to the experimental values (Fig. 5).

Apparently the model based on chemostat data fits well at higher μ , while it needed adjustment at values below $\mu = 0.05 \text{ h}^{-1}$. Most importantly, it was valid for the optimized feed protocol derived from the model, which led to a 2.2 fold increase of volumetric productivity. The sensitivity of the model to the accuracy of the function $q_p = f(\mu)$ stresses the importance of an accurate experimental determination of q_p both at low and high μ . Importantly, this function can be refined in future utilizing additional data from fed batch (and chemostat) cultures, so that the model acquires features of a self learning model.

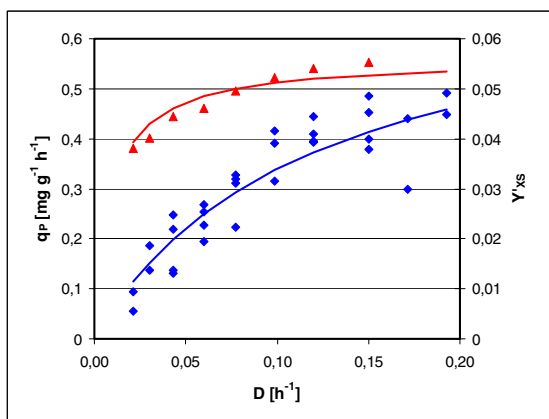


Figure 1
 q_p and Y'_{XS} in chemostat cultures. Specific product formation rate q_p (blue diamonds) and observed biomass yield coefficient Y'_{XS} (red triangles), as well as the respective approximations.

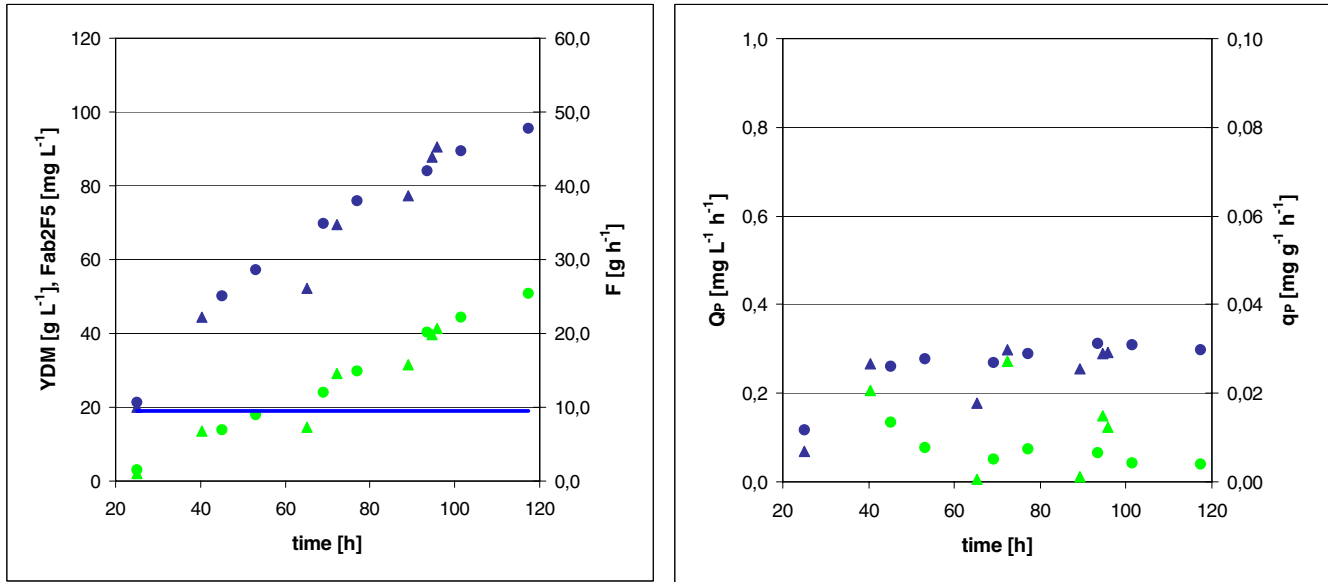


Figure 2
Kinetics of standard fed batch cultures. Values of two cultures (triangles and circles) are shown. A: biomass x (blue), product concentration p (green), and feed rate F (blue line) B: volumetric productivity Q_p (blue) and specific product formation rate q_p (green).

Analytic approach

The calculus of variations yields a method to derive an analytic formula (containing indeterminate parameters) for our optimization problem. Let $X = X(t)$ be the amount of biomass, $P = P(t)$ the amount of product and $\mu = \mu(t)$ the specific growth rate of biomass at the point of time t . The process to be controlled is described by the following equations:

The growth of biomass is modeled by the equation

$$X'(t) = \mu(t) \cdot X(t) \quad (4)$$

with the initial value $X(0) = X_0$.

The yield of product is modeled by the equation

$$P'(t) = q_p(t) \cdot X(t) \quad (5)$$

with a Monod like formula

$$q_p(t) = q_{p\max} \frac{\mu(t)}{k_q + \mu(t)} \quad (6)$$

and the initial value $P(0) = P_0$.

In a first step we maximize the cumulative yield of product in a fixed time interval $[0, T]$. Therefore, $\frac{1}{T} P(T) \rightarrow \text{Max}$ is equivalent to $P(T) \rightarrow \text{Max}$. We therefore consider

$$P(t) = \int_0^T P'(t) \cdot dt = q_{p\max} \int_0^T \frac{\mu(t)}{k_q + \mu(t)} X(t) \cdot dt \quad (7)$$

by formulas (5) and (6). Inserting $X'(t)/X(t)$ for $\mu(t)$ results in maximizing the integral

$$I(X) = \int_0^T \frac{X(t) \cdot X'(t)}{k_q \cdot X(t) + X'(t)} dt \quad (8)$$

This integral has an extremum only if the Euler-Lagrange differential equation

$$\frac{\partial F}{\partial x} - \frac{d}{dt} \frac{\partial F}{\partial X'} = 0 \quad (9)$$

$$\text{is satisfied with } F = F(X, X') = \frac{X \cdot X'}{k \cdot X + X'} \quad (10)$$

Evaluating the Euler Lagrange equation (9) yields

$$\frac{X'^2}{(k_q \cdot X + X')^2} - \frac{d}{dt} \frac{k_q \cdot X^2}{(k_q \cdot X + X')^2} = 0 \quad (11)$$

Inserting $\mu(t) \cdot X(t)$ for $X'(t)$ (from eq. 4) yields

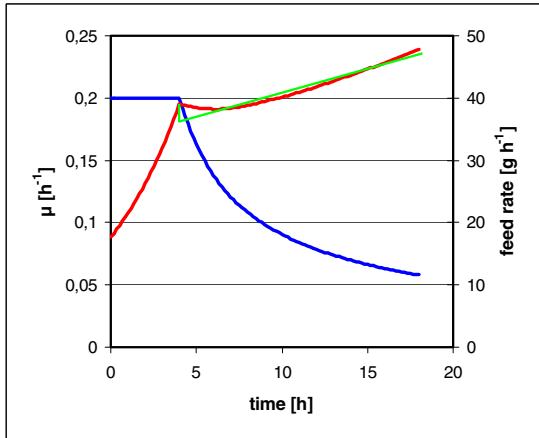


Figure 3
Optimum time course of specific growth rate and medium feed. μ (blue) and F (red) as obtained with the Solver. The approximated linearly increasing feed rate is shown in green.

$$\frac{\mu^2}{(k_q + \mu)^2} - \frac{d}{dt} \frac{k_q}{(k_q + \mu)^2} = 0 \quad (12)$$

Calculating the differential and reducing to a common denominator gives

$$2 \cdot k_q \cdot \mu'(t) + \mu^3(t) + k_q \cdot \mu^2(t) = 0 \quad (13)$$

We solve this differential equation and get the following equation for $\mu(t)$:

$$t + c = \frac{2}{\mu} + \frac{2}{k_q} \ln \frac{\mu}{k_q + \mu} \quad (14)$$

Eq (14) defines only a necessary condition for the optimum trajectory of μ over time, with indetermined parameter c and indetermined optimal value for the total feed time T . Here the parameter c and the optimal value for the total feed period T depend on the constraints μ_{\min} , μ_{\max} and $q_{P\max}$. Since we know a numerical optimal solution for the growth function, we calculate the value c by fitting the analytic curve (eq 14) to the optimal solution calculated by Excel with the method of least squares. We get as numeric value $c = -1.49967812$ h. Figure 6 shows the excellent correspondence of the analytic solution and the solution calculated by the Excel Solver. This proves that the Solver solution obeys the necessary condition of the maximization problem, as defined by the Euler-Lagrange equation.

Conclusion

We have developed a modeling and optimization algorithm for fed batch cultures of secreted products based on MS Excel. The validity of this iterative calculation, which is highly flexible and versatile, was proven by analytic solution of the equations forming the basis of the fed batch model. While the analytic solution fits exactly to the phase of decreasing specific growth rate of the Excel Solver solution, it is not possible to calculate the duration of the initial μ_{\max} phase. As the optimum feed profiles obviously

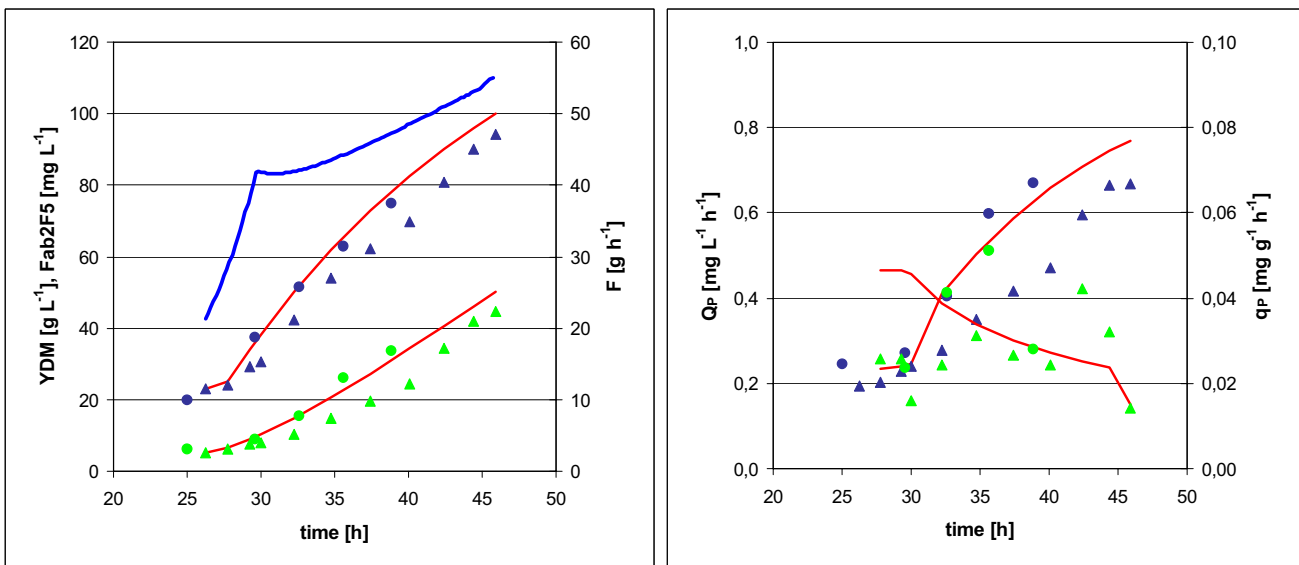


Figure 4
Kinetics of optimized fed batch cultures. Values of two cultures (triangles and circles) are shown. The model approximations are indicated by red lines. A: biomass x (blue), product concentration p (green), and feed rate (blue line) B: volumetric productivity Q_p (blue) and specific product formation rate q_p (green).

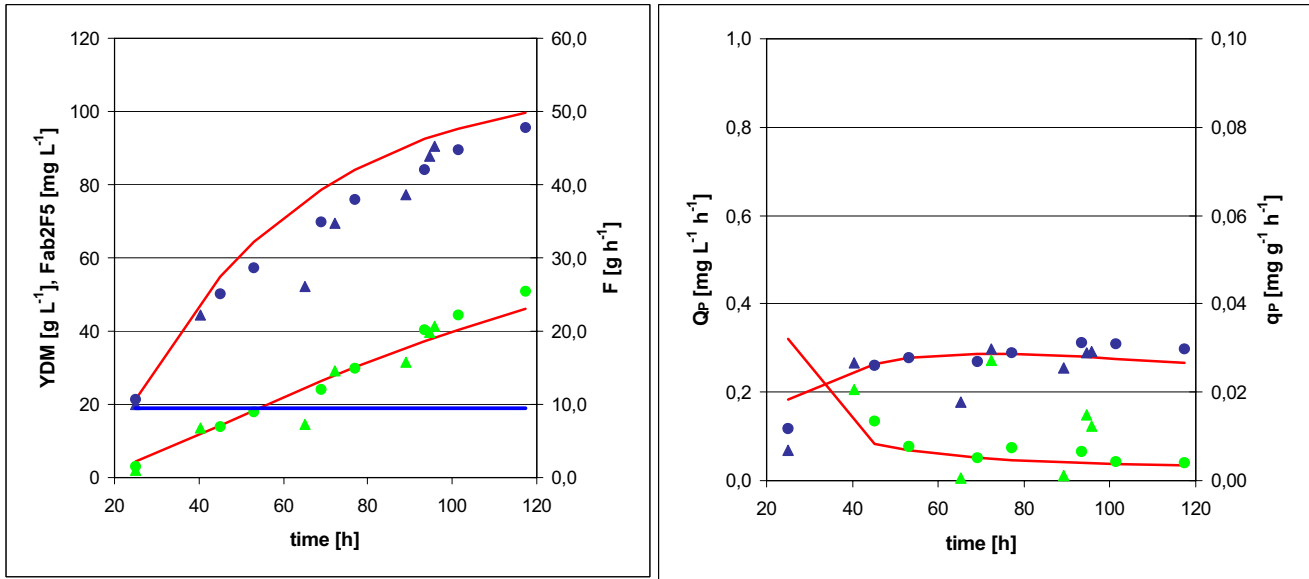


Figure 5
Kinetics of standard fed batch cultures with adapted model approximations indicated by red lines. Values of two cultures (triangles and circles) are shown. A: biomass x (blue), product concentration p (green), and feed rate (blue line) B: volumetric productivity Q_p (blue) and specific product formation rate q_p (green).

consist of an exponential phase followed by a phase of steadily decreasing μ , the analytic approach could only serve as evidence of the correct solution of the optimization problem obtained with Excel Solver. Both the Euler-Lagrange approach used here and Pontryagin's Maximum Principle depend on data fitting to obtain a numeric solution. Given the perfect match of the two approaches presented here, we consider it much more straight forward to apply a numeric data fitting approach directly to the equations of growth and product formation.

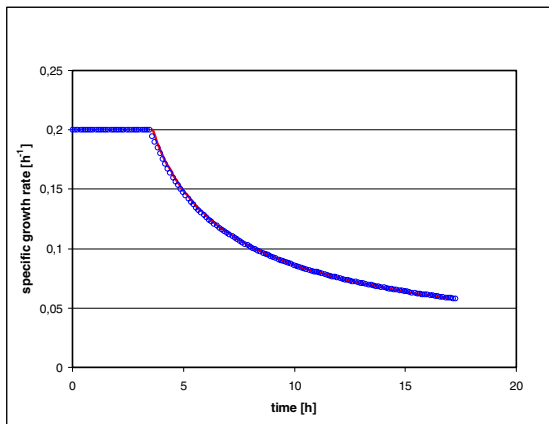


Figure 6
Optimized time course of specific growth rate. Overlay of the solution obtained by Solver (blue circles) and analytic solution (red line).

The advantages of the procedure we describe here are the ease of use, applying software familiar to every scientist and engineer, and rapid calculation which makes predictions extremely easy, so that many options can be tested *in silico* quickly. Additional options like further biological and technological constraints or different functions for specific productivity and biomass yield can easily be integrated.

We could prove that the experimental data basis for the functions behind the algorithm is very important. Different to previous work this was taken into account, and especially the sensitivity at very low specific growth rates needs to be highlighted.

The Excel file containing the model and optimization procedure is provided as accompanying file [see additional file 1].

Materials and methods

Unless stated otherwise, all chemicals were purchased from Merck Eurolab and all antisera were from Sigma.

Strain

A *P. pastoris* strain X33 (wild type strain) expressing extracellularly the Fab fragment of the anti-HIV antibody 2F5 under control of the GAP promoter was used in this study. The development of this strain has been described elsewhere [12]. A cell bank of the strain was prepared, divided in 1.8 mL aliquots and stored at -80°C.

Fermentation

A shake flask containing 100 mL of YPG medium (per liter: 10 g yeast extract, 10 g peptone, 10 g glycerol) was inoculated with one cryovial from the *P. pastoris* cell bank, and incubated at 28 °C for approximately 24 hours and agitated at 180 rpm.

This culture was used to inoculate the starting volume in the bioreactor to a starting optical density (OD₆₀₀) of 1.0. Depending on the operation mode the starting volume was either 1.2 L for fed batch or 1.4 L for chemostat process.

Fermentations were carried out in a 2.0 L working volume bioreactor (MBR; Wetzikon, Switzerland) with a computer based process control (ISE; Vienna, Austria). Fermentation temperature was controlled at 25 °C, pH was controlled at 5.0 with addition of 25% ammonium hydroxide and the dissolved oxygen concentration was maintained above 20% saturation by controlling the stirrer speed between 600 and 1200 rpm, whereas the airflow was kept constant at 100 L h⁻¹.

The media were as follows:

Batch medium contained per liter:

2.0 g citric acid, 12.4 g (NH₄)₂HPO₄, 0.022 g CaCl₂ · 2H₂O, 0.9 g KCl, 0.5 g MgSO₄ · 7H₂O, 40 g glycerol, 4.6 ml PTM₁ trace salts stock solution. The pH was set to 5.0 with 25% HCl.

Glucose fed batch solution contained per liter:

550 g glucose · 1H₂O, 10 g KCl, 6.45 g MgSO₄ · 7H₂O, 0.35 g CaCl₂ · 2H₂O and 12 ml PTM₁ trace salts stock solution.

Chemostat medium contained per liter:

55 g glucose · 1H₂O, 2.5 g KCl, 1.0 g MgSO₄ · 7H₂O, 0.035 g CaCl₂ · 2H₂O, 21.8 g (NH₄)₂HPO₄ and 2.4 ml PTM₁ trace salts stock solution, furthermore the pH was set to 5.0 with 25% HCl.

PTM₁ trace salts stock solution contained per liter:

6.0 g CuSO₄ · 5H₂O, 0.08 g NaI, 3.0 g MnSO₄ · H₂O, 0.2 g Na₂MoO₄ · 2H₂O, 0.02 g H₃BO₃, 0.5 g CoCl₂, 20.0 g ZnCl₂, 65.0 g FeSO₄ · 7H₂O, 0.2 g biotin and 5.0 ml H₂SO₄ (95%–98%). All chemicals for PTM₁ trace salts stock solution were from Riedel-de Haën (Seelze, Germany), except for biotin (Sigma, St. Louis, MO, USA), and H₂SO₄ (Merck Eurolab).

After approximately 24 hours the batch was finished and – depending on the fermentation strategy – the feed and if required the harvest was started.

The continuous fermentation was initiated at a $D = 0.15$ h⁻¹ and performed at least for 5 resident times τ to reach steady state conditions.

$$\tau = \frac{1}{D} = \frac{V}{F} \quad (15)$$

Then the dilution rate was decreased stepwise, always achieving steady state conditions before the next change of the dilution rate. At $D = 0.0086$ h⁻¹ the procedure was reversed and the dilution rate was increased stepwise up to the critical dilution rate $D_{\text{crit}} = 0.2$ h⁻¹. Samples were taken after 3 and 5 τ and analyzed as described below.

The standard fermentation strategy was a fed batch with a constant feed, this means that the batch phase was followed by the glucose fed batch with a feed rate $F = 8.925$ g h⁻¹. The fermentations were terminated at appr. $t = 120$ h. Samples were taken frequently and processed as described below.

The optimized fermentation strategy consists of different phases to perform the calculated growth kinetic. The batch phase was followed by an exponential feed phase with a growth rate of 0.2 for 3.6 hours, followed by a linearly increasing feed rate calculated by equation (16), where $k = 0.0144$ g h⁻² and $d = 36.8064$ g h⁻¹ for 16.0 hours.

$$F_L = k \cdot t_L + d \quad (16)$$

Analytical methods

Optical density

The samples were diluted in ddH₂O up to 1:500 to measure the OD at 600 nm.

Biomass determination

2 × 5 ml culture were centrifuged and the supernatants frozen for further analysis. The pellets were resuspended in ddH₂O, recentrifuged, and the pellets again resuspended in ddH₂O, transferred to a weighed beaker, dried at 105 °C until constant weight.

Product quantification (ELISA)

To determine the Fab content, 96 well microtiter plates (MaxiSorb, Nunc, Denmark) were coated with anti-hIgG (Fab specific) overnight at RT (1:1000 in PBS, pH 7.4), before serially diluted supernatants of *P. pastoris* cultures secreting 2F5 Fab (starting with a 1:100 dilution in PBS) were applied and incubated for 2 h at RT. Fab of normal IgG (Nordic) was used as a standard protein at a starting concentration of 200 ng/ml. After each incubation step the plates were washed four times with PBS containing 1% Tween 20 adjusted to pH 7.4. 100 μ l of anti-kappa light chain – AP conjugate as secondary antibody (1:1000 in PBS/Tween + 2% BSA) were added to each well, and incubated for 1 h at RT. After washing, the plates were

stained with pNPP (1 mg/ml p-nitrophenyl phosphate in coating buffer, 0.1 N Na₂CO₃/NaHCO₃; pH 9.6) and read at 405 nm (reference wavelength 620 nm).

Method of calculation

1. Setup of calculations

We divide the total feed period in equal intervals $[t_n, t_{n+1}]$ ($1 \leq n \leq N$) of length dt . Therefore,

$$t_{n+1} = t_n + dt \quad (17)$$

We start with an initial value $dt = 1$ [h]. The best value for dt is determined within the optimization process.

At every point of time t_n we denote by $X_n = X(t_n)$ the amount of biomass and by $P_n = P(t_n)$ the amount of product in the bioreactor. At the beginning of the fed-batch process the initial values are $X(0) = X_0$ and $P(0) = P_0$, as achieved at the end of the batch phase.

First we have to describe the growth of the biomass. We use the simplest model, the exponential growth model,

$$\frac{dX}{dt} = \mu \cdot X \quad (18)$$

Since the specific growth rate μ of the biomass depends on time, we calculate (eq. 18) in discrete time steps

$$X_{n+1} = X_n \cdot e^{\mu_n dt} \quad (19)$$

where μ_n is the specific growth rate during the interval $[t_n, t_{n+1}]$. The initial values for μ_n are chosen arbitrarily, for instance $\mu_n \equiv \mu_{\max}$. The optimal values for all of the μ_n 's are determined within the optimization process.

Second we have to describe the accumulation of the product. We simply calculate the total product yield during the interval $[t_n, t_{n+1}]$ by the following formula

$$P_{n+1} = P_n + dP_n \quad (20)$$

with

$$dP_n = q_{Pn} \cdot X_n \cdot dt \quad (21)$$

The relationship between the specific rate q_p of product formation and the specific growth rate μ was experimentally determined in chemostat cultures. The dependence of q_p on μ was described analogous to Monod equation:

$$q_{Pn} = q_{P\max} \cdot \frac{\mu_n}{k_q + \mu_n} \quad (22)$$

The values for $q_{P\max}$ and k_q are derived from the experimental data by the method of least squares, i.e. the parameters $q_{P\max}$ and k_q are chosen that the sum of the deviations from the experimental data squared is minimal.

Next we have to calculate the amount of substrate dS which we must feed in the time interval $[t_n, t_{n+1}]$. To do this, let S_n be the amount of substrate added to the bioreactor until the time point t_n . Then the substrate consumption rate depends on the amount and on the increase of biomass, i.e.

$$\frac{dS}{dt} = -(m_S \cdot X + \frac{1}{Y_{XS}} \frac{dX}{dt}) \quad (23)$$

where m_S is the maintenance coefficient and Y_{XS} is the true yield coefficient of biomass from substrate. Inserting formula (18) in (23) the amount of substrate feed in the interval $[t_n, t_{n+1}]$ calculates as

$$dS_n = (\frac{\mu_n}{Y_{XS}} + m_S) \cdot X_n \cdot dt \quad (24)$$

To calculate the parameters Y_{XS} and m_S from experimental data of chemostat cultures by the method of least squares, we use the observed biomass yield coefficient Y'_{XS} depending on the specific growth rate μ . This is done by $dX = -Y'_{XS} \cdot dS$ and inserting formula (18) and the formula for the whole substrate consumption which implies

$$Y'_{XS} = \frac{\mu}{\frac{\mu}{Y_{XS}} + m_S} \quad (25)$$

Formula (25) can be transformed to

$$\frac{1}{Y'_{XS}} = \frac{1}{Y_{XS}} + \frac{m_S}{\mu} \quad (26)$$

From this double reciprocal plot Y_{XS} and m_S were determined by linear regression.

Last but not least we need the total volume for the calculation of the volumetric productivity. The model process starts with a batch volume of $V_0 = 1$ L. The total volume at each time interval is then

$$V_{n+1} = V_n + \frac{dS_n}{s_f \cdot \rho_f \cdot 1000} \quad (27)$$

with the substrate concentration in the feed medium s_f and the density of the feed medium ρ_f . Due to the high biomass concentrations achieved in *P. pastoris* fermentations, the cells occupy a significant fraction of the total volume, while the product is secreted to the liquid phase,

the culture supernatant. In order to calculate the product concentration, the available liquid volume V_l is calculated at each time interval with the specific volume of wet biomass, which is derived from dry biomass as the specific volume per dry biomass $v_{YDM} = 0.0033 \text{ L g}^{-1}$.

$$V_{ln} = V_n - X_n \cdot v_{YDM} \quad (28)$$

Finally, we calculate the biomass and the product concentrations. The product concentration p at the time point t_n is calculated as

$$p_n = \frac{P_n}{V_{ln}} \quad (29)$$

and the biomass concentration x at the same time point is

$$x_n = \frac{X_n}{V_n} \quad (30)$$

The medium feed rate F_n at each time point is

$$F_n = \frac{dS_n}{s_f} \cdot \frac{1}{dt} \quad (31)$$

These values are used to determine the feed rate profile of the optimized fed batch process.

2. Optimization

The goal of our optimization problem is to find the best values for the specific growth rates μ_n and the best value for dt (which implies that the total feed period undergoes the optimization process too) such that the volumetric productivity Q_p calculated at the point of time t_{N+1} as

$$Q_{PN+1} = \frac{P_{N+1}}{(t_0 + t_{N+1}) \cdot V_{N+1}} \quad (32)$$

is maximized under the following constraints:

$$\mu_{\min} \leq \mu_n \leq \mu_{\max} \text{ for } (1 \leq n \leq N) \quad (33)$$

and

$$X(t_{N+1}) = X_{\max} \quad (34)$$

Here $\mu_{\max} = 0.2 \text{ h}^{-1}$ is the maximum specific growth rate at just below washout in chemostat cultures. Since below $\mu = 0.02 \text{ h}^{-1}$ significant product degradation appeared, the lower boundary was set at $\mu_{\min} = 0.03 \text{ h}^{-1}$. Also the biomass concentration needs to be limited. The upper limit is mainly defined by the cell separation step, which is practically limited with approximately 100 g L^{-1} dry mass.

Remark

Additional constraints may be entered, e.g. the final product concentration may be set at a minimum level.

In the Excel sheet we set $N = 150$. The values t_n, μ_n, X_n, \dots are organized in columns, with each time point t_n, \dots a row. The values of X_n, P_n, V_n, \dots are calculated from the respective previous row using the equations provided above. The optimization process is performed by the Excel Solver as a black box. It maximizes the final Q_p field by varying the μ fields within the boundaries and the dt field.

The Excel file used for this work is provided as an additional file.

Analytic approach

To verify the Excel Solver solution, the exact solution of the optimization problem was determined with calculus of variation.

List of symbols

The symbols, their definitions and units are summarized in table 1.

Abbreviations

ddH₂O double distilled water

Table 1: List of symbols

Symbol	Definition	unit
c	model parameter	h
D	dilution rate	h^{-1}
d	axis intercept	g h^{-1}
F	flow rate	g h^{-1}
F_L	flow rate of linear feed	g h^{-1}
k	slope	g h^{-2}
k_q	Monod constant for q_p	h^{-1}
m_s	maintenance coefficient	h^{-1}
p	product concentration	mg L^{-1}
P	product mass	mg
q_p	specific product formation rate	$\text{mg g}^{-1} \text{h}^{-1}$
Q_p	volumetric productivity	$\text{mg L}^{-1} \text{h}^{-1}$
$q_{p\max}$	maximum specific productivity	$\text{mg g}^{-1} \text{h}^{-1}$
S	substrate mass	g
s_f	substrate concentration	g L^{-1}
s_q	estimated standard deviation of q_p	$\text{mg g}^{-1} \text{h}^{-1}$
s_Y	estimated standard deviation of Y'_{XS}	-
t	time	h
t_L	time of linear feed	h
T	total feed time	h
V	volume	L
V_l	volume of liquid supernatant	L
x	dry biomass concentration	g L^{-1}
X	dry biomass	g
Y'_{XS}	observed biomass yield coefficient	-
Y_{XS}	theoretical biomass yield coefficient	-
<i>Greek symbols</i>		
μ	specific growth rate	h^{-1}
v_{YDM}	specific volume of biomass	L kg^{-1}
ρ_f	density of feed medium	kg L^{-1}
τ	average residence time	h

Fab2F5 Fab fragment of antibody 2F5

GAP glyceraldehyde-3-phosphate dehydrogenase

OD optical density

YDM yeast dry mass

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

MM supported designing the model, planned and performed the fermentations, and helped in drafting the manuscript. MK developed the model and the optimization procedure, and prepared part of the manuscript. BG developed the expression strain and supported in fermentations and analytics. DM formulated the project, supported the development of the model and drafted the manuscript. All authors read and approved the final manuscript.

Additional material

Additional File 1

Model of *P. pastoris* fed batch process for secreted heterologous protein production. This file contains the model as described in Material and Methods, with the optimized values for μ and dt. By adjusting the specific parameters, other fed batch processes can be modeled and optimized with the Excel Solver.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1475-2859-5-37-S1.xls>]

Acknowledgements

The authors wish to thank Prof. Werner Nowak, Institute of Mathematics, University of Natural Resources and Applied Life Sciences Vienna, for support and valuable discussions. This work was supported by the Austrian Research Promotion Agency (Program FHplus), and Polymun Scientific GmbH, Vienna, Austria. Part of the results presented here have been communicated at the 4th Recombinant Protein Production Meeting (Barcelona, 2006).

References

- Sinclair CG, Kristiansen B, Bu'Lock JD: *Fermentation kinetics and modelling* Milton Keynes: Open University Press; 1987.
- Modak JM, Lim HC, Tayeb YJ: **General characteristics of optimal feed rate profiles for various fed batch fermentation processes.** *Biotechnol Bioeng* 1986, **28**:1396-1407.
- Zhang W, Sinha J, Smith LA, Inan M, Meagher MM: **Maximization of production of secreted recombinant proteins in *Pichia pastoris* fed-batch fermentation.** *Biotechnol Prog* 2005, **21**:386-393.
- Ohno H, Nakanishi E, Takamatsu T: **Optimal control of a semi-batch fermentation.** *Biotechnol Bioeng* 1976, **18**:847-864.
- Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Tomomitsu K: **High level secretion of recombinant human serum albumin by fed-batch fermentation of the methylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy.** *J Biosci Bioeng* 2000, **90**:280-288.
- Porro D, Sauer M, Branduardi P, Mattanovich D: **Recombinant protein production in yeasts.** *Mol Biotechnol* 2005, **31**:245-259.
- Lasdon LS, Waren AD, Jain A, Ratner M: **Design and testing of a generalized reduced gradient code for nonlinear programming.** *ACM T Math Software* 1978, **4**:34-49.
- Werner RG: **Economic aspects of commercial manufacture of biopharmaceuticals.** *J Biotechnol* 2004, **113**:171-182.
- Hensing MCM, Rouwenhorst RJ, Heijnen JJ, van Dijken JP, Pronk JT: **Physiological and technological aspects of large-scale heterologous-protein production with yeasts.** *Antonie van Leeuwenhoek* 1995, **67**:261-279.
- Ohya T, Morita M, Miura M, Kuwae S, Kobayashi K: **High-level production of prourokinase-annexin V chimeras in the methylotrophic yeast *Pichia pastoris*.** *J Biosci Bioeng* 2002, **94**:467-473.
- Roels JA: *Energetics and kinetics in biotechnology* Amsterdam, Elsevier Biomedical Press; 1983.
- Gasser B, Maurer M, Gach J, Kunert R, Mattanovich D: **Engineering of *Pichia pastoris* for improved production of antibody fragments.** *Biotechnol Bioeng* 2006, **94**:353-361.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

