ORIGINAL PAPER

B. Zelić · Đ. Vasić-Rački · C. Wandrey · R. Takors

Modeling of the pyruvate production with *Escherichia coli* in a fed-batch bioreactor

Abstract A family of 10 competing, unstructured models has been developed to model cell growth, substrate consumption, and product formation of the pyruvate producing strain *Escherichia coli* YYC202 *ldhA::Kan* strain used in fed-batch processes. The strain is completely blocked in its ability to convert pyruvate into acetyl-CoA or acetate (using glucose as the carbon source) resulting in an acetate auxotrophy during growth in glucose minimal medium. Parameter estimation was carried out using data from fed-batch fermentation performed at constant glucose feed rates of $q_{VG} = 10$ mL h^{-1}. Acetate was fed according to the previously developed feeding strategy. While the model identification was realized by least-square fit, the model discrimination was based on the model selection criterion (MSC). The validation of model parameters was performed applying data from two different fed-batch experiments with glucose feed rate $q_{VG} = 20$ and 30 mL h^{-1}, respectively. Consequently, the most suitable model was identified that reflected the pyruvate and biomass curves adequately by considering a pyruvate inhibited growth (Jerusalimsky approach) and pyruvate inhibited product formation (described by modified Luedeking–Piret/Levenspiel term).

Keywords *Escherichia coli* · Production of pyruvate · Unstructured cell model · Fermentation of recombinant *E. coli*

List of symbols

c_A	acetate concentration (g L^{-1})
$c_{A,0}$	acetate concentration in the feed (g L^{-1})
c_G	glucose concentration (g L^{-1})
$c_{G,0}$	glucose concentration in the feed (g L^{-1})
c_P	pyruvate concentration (g L^{-1})
$c_{P,max}$	critical pyruvate concentration above which reaction cannot proceed (g L^{-1})
c_X	biomass concentration (g L^{-1})
K_I	inhibition constant for pyruvate production (g L^{-1})
K_I^A	inhibition constant for biomass growth on acetate (g L^{-1})
K_P	saturation constant for pyruvate production (g L^{-1})
K_P	inhibition constant of Jerusalimsky (g L^{-1})
K_S^A	Monod growth constant for acetate (g L^{-1})
K_S^G	Monod growth constant for glucose (g L^{-1})
m_A	maintenance coefficient for growth on acetate (g g^{-1} h^{-1})
m_G	maintenance coefficient for growth on glucose (g g^{-1} h^{-1})
n	constant of extended Monod kinetics (Levenspiel) (–)
q_V	volumetric flow rate (L h^{-1})
q_{VA}	volumetric flow rate of acetate (L h^{-1})
q_{VG}	volumetric flow rate of glucose (L h^{-1})
r_A	specific rate of acetate consumption (g g^{-1} h^{-1})
r_G	specific rate of glucose consumption (g g^{-1} h^{-1})
r_P	specific rate of pyruvate production (g g^{-1} h^{-1})
$r_{P,max}$	maximum specific rate of pyruvate production (g g^{-1} h^{-1})
t	time (h)
V	reaction (broth) volume (L)
$Y_{P/G}$	yield coefficient pyruvate from glucose (g g^{-1})
$Y_{X/A}$	yield coefficient biomass from acetate (g g^{-1})
$Y_{X/A,max}$	maximum yield coefficient biomass from acetate (g g^{-1})
$Y_{X/G}$	yield coefficient biomass from glucose (g g^{-1})

B. Zelić (✉) · Đ. Vasić-Rački
Faculty of Chemical Engineering and Technology,
University of Zagreb, Marulicev trg 19, 10000 Zagreb, Croatia
E-mail: bzelic@marie.fkit.hr
Tel.: +385-1-4597146
Fax: +385-1-4597133

C. Wandrey · R. Takors
Institute of Biotechnology 2, Research Center Jülich,
52425 Jülich, Germany

$Y_{X/G,max}$	maximum yield coefficient biomass from glucose (g g^{-1})
α	growth associated product formation coefficient (g g^{-1})
β	non-growth associated product formation coefficient (g g^{-1} h^{-1})
μ	specific growth rate (h^{-1})
μ_{max}	maximum specific growth rate (h^{-1})

Introduction

Pyruvic acid and its salts are important chemicals used in the pharmaceutical, food, agrochemical, and cosmetic industry [1, 2, 5]. Pyruvate represents one of the most important metabolites in the central metabolism of living cells because of its role in the glucose uptake (via carbohydrate phosphoenolpyruvate:phosphotransferase system (pts)), its impact as a precursor for amino acid synthesis, its relevance as an intermediate of glycolysis, etc., thus making the metabolite one of the most widely used reactants in the *E. coli* metabolic network [3].

Basically, there are two different approaches known for the production of pyruvate: (i) the classical chemical route considering the energy-intensive pyrolysis of tartaric acid [4] and (ii) the biotechnological access using purified enzymes, non-growing, immobilized or living cells [1, 2, 5, 7]. Currently, the fermentation is regarded as one of the most promising routes for the production of pyruvate. As an example, the conversion of glucose to pyruvate with non-growing, acetate auxotrophic cells of *E. coli* YYC202 *ldhA::Kan* was published, which achieved a maximum volumetric productivity (Q_P) of 145 g$_{Pyruvate}$/L/d, a maximum pyruvate/glucose yield ($Y_{P/G}$) of 1.78 mol$_{Pyruvate}$/mol$_{Glucose}$, and maximum product titers (c_P) of about 720 mM (approx. 65 g L^{-1}), thus offering good starting conditions for subsequent downstream processing [1, 6, 7].

For further process development (optimization) and scale-up, the quantitative understanding of the basic cellular characteristics, such as the cellular demand for the sole carbon source glucose and for the auxotrophic substance acetate, are of outstanding importance. Therefore, modeling studies were performed to identify simple, easy-to-use, and robust models that are suitable to support the engineering tasks of process optimization and design.

For this purpose, "black-box" unstructured cell models were chosen because they offer a significant simplicity at the same time providing a general understanding of the dominant metabolic processes in the production strain. Although these type of models require very rigorous constraints such as "balanced growth" (i.e. "pseudo" steady-state metabolic conditions) or "one-component cell systems", they nevertheless showed their applicability and suitability for a variety of different modeling tasks and they additionally represent a valuable basis for subsequently establishing

structured models that aim at describing the intracellular metabolism in detail [8, 9]. Unstructured models usually use simple Michaelis–Menten type kinetic equations with few variables, each (more or less) contributing a physical meaning while structured models consist of a wide number of variables thus trying to reflect the microbial complexity on a biochemical level. Because of the simplified model structure "black-box" models usually have problems in describing lengthy cell-lag phases [10], which necessitates their careful application for highly unstable process conditions. However, they can legitimately be applied to systems in which "balanced growth" occurs [11, 12].

The estimation of kinetic parameters in unstructured growth models with high parameter accuracy is essential for successful model validation. Parameter estimation in unstructured growth models is often performed with the aid of continuous fermentation [13]. These experiments are usually time-consuming and their practical realization is not always feasible because of the (relatively) complex experimental set-up (e.g. complex cell recycle system, complex pump–balance system). In contrast, simpler experiments can be achieved using batch fermentations, but the kinetic parameters of the Michaelis–Menten type models cannot be uniquely identified from noisy batch measurements [11, 14]. The extension of the batch experiment by a fed-batch phase with time-variant feed rate leads to a higher accuracy of the parameter estimates [15]. However, it is noteworthy that feed rate profiles obtained for optimal process performance are not necessarily optimal for parameter estimation [16]. For instance, Baltes et al. [11] emphasized that changes in feeding rates should be as small as possible to avoid fast dynamic system responses, which are not described by unstructured models. Nevertheless, our study was based on fed-batch experiments for the estimation of kinetic parameters. This decision was motivated by the aim to identify kinetic parameters of a process that resembles the optimal production conditions as much as possible. Therefore, the glucose-feeding profile kept constant and the acetate feed followed previous findings [2].

Thus the aim of this work was to develop an unstructured, mathematical model of the pyruvate production process considering the acetate auxotrophic strain *E. coli* YYC202 *ldhA::Kan*. Because of primary uncertainties regarding the microbial kinetics (and the underlying mechanism) several models for microbial growth and product formation were formulated as a "competing" set of model candidates. To our knowledge, this represents the first approach to model *E. coli* based pyruvate formation with a production strain.

Modeling

In our previous experimental studies [1, 7] we observed that pyruvate formation followed a typical pattern, which can be described as a mixture of growth and non-

growth associated product formation. Glucose was assumed to be the only carbon source for cell growth and product formation. However, the acetate auxotrophy had to be considered as well. Furthermore, high extracellular pyruvate levels were identified as an inhibiting factor for cell growth and/or pyruvate formation. Taking these "vague" basic characteristics into account, the following constraints were formulated and used to define primary, verbal models:

– Glucose and acetate are limiting substrates. All other nutrients such as nitrogen, phosphate, and growth factors are ample;
– Cell growth occurs simultaneously using glucose and acetate, however acetate is essential for growth;
– There was no oxygen effect on either biomass growth or bioconversion of glucose to pyruvate—at least at the dissolved oxygen range investigated (from 40 to 100%);
– Product formation kinetics combined growth associated and non-growth associated characteristics;
– For simplification, the bioconversion of glucose to pyruvate is regarded as a one-step enzymatic reaction:

$$C_6H_{12}O_6 \xrightarrow{O_2} 2C_3H_4O_3 + 2H_2O;$$
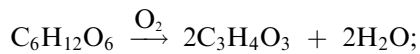
– Pyruvate production and/or biomass growth are inhibited by high pyruvate concentrations;
– The suspension viscosity in the reactor remains constant during the experiment;
– Potential mixing effects of the highly concentrated feeds with the cultivation medium are neglected for the sake of the model simplicity.

Microbial growth

To model the microbial specific growth rate various, well-known approaches were tested. The "multiple substrates Monod kinetics" was used in models 1 and 2 [17, 18, 19]. However, it is known that this kinetic term has an inherent disadvantage if many essential substrates are considered. Even if the concentrations of, for example, five essential substrates are so high that their contribution of the growth rate reaches 90% of the saturation value, the total growth rate will be limited to 60% of the maximum possible value.

In models 3 and 4, growth inhibition by-product formation was considered according to the Levenspiel approach [20]. This model presumes a critical inhibitor (product) concentration, $c_{P,max}$, as an upper limit for cell growth. As a consequence, the parameters of the Monod equation are dependent on the product concentration.

In models 5, 6, and 9 inhibition of growth by-product formation was described by the model of Jerusalimsky [21]. This model represents an approximate analogy to the non-competitive substrate inhibition, which is often used in pure enzyme kinetic models.

In models 7 and 8, acetate was assumed to be only limiting substrate for growth, which, strictly speaking, was not experimentally observed. However, this approach was motivated by the strain genotype, namely its acetate auxotrophy.

Finally, in model 10, the well-known Andrews kinetics [22] were considered with respect to a potential acetate (substrate) inhibition of growth at high acetate levels. In models 7, 8, and 10, glucose was only used for pyruvate production and not for growth.

Product formation

In models 1, 3, 5, 7, and 10 the kinetics of product formation was based on the Luedeking–Piret equation [23] to cope with the experimental observations mentioned in the introductory part of this section. As

Table 1 Kinetic equations used for description of biomass growth and product formation in the developed models

Model	Biomass growth	Product formation
1	$\mu = \mu_{max} \cdot \frac{c_G}{K_S^G + c_G} \cdot \frac{c_A}{K_S^A + c_A}$	$r_P = \alpha \cdot \mu + \beta$
2	$\mu = \mu_{max} \cdot \frac{c_G}{K_S^G + c_G} \cdot \frac{c_A}{K_S^A + c_A}$	$r_P = r_{P,max} \cdot \frac{c_P}{\left(K_P + c_P + \frac{c_P^2}{K_I}\right)}$
3	$\mu = \mu_{max} \cdot \frac{c_G}{K_S^G + c_G} \cdot \frac{c_A}{K_S^A + c_A} \cdot \left(1 - \frac{c_P}{c_{P,max}}\right)^n$	$r_P = \alpha \cdot \mu + \beta$
4	$\mu = \mu_{max} \cdot \frac{c_G}{K_S^G + c_G} \cdot \frac{c_A}{K_S^A + c_A} \cdot \left(1 - \frac{c_P}{c_{P,max}}\right)^n$	$r_P = r_{P,max} \cdot \frac{c_P}{\left(K_P + c_P + \frac{c_P^2}{K_I}\right)}$
5	$\mu = \mu_{max} \cdot \frac{c_G}{K_S^G + c_G} \cdot \frac{c_A}{K_S^A + c_A} \cdot \frac{K_P}{(c_P + K_P)}$	$r_P = \alpha \cdot \mu + \beta$
6	$\mu = \mu_{max} \cdot \frac{c_G}{K_S^G + c_G} \cdot \frac{c_A}{K_S^A + c_A} \cdot \frac{K_P}{(c_P + K_P)}$	$r_P = r_{P,max} \cdot \frac{c_P}{\left(K_P + c_P + \frac{c_P^2}{K_I}\right)}$
7	$\mu = \mu_{max} \cdot \frac{c_A}{K_S^A + c_A}$	$r_P = \alpha \cdot \mu + \beta$
8	$\mu = \mu_{max} \frac{c_A}{K_S^A + c_A}$	$r_P = r_{P,max} \cdot \frac{c_P}{\left(K_P + c_P + \frac{c_P^2}{K_I}\right)}$
9	$\mu = \mu_{max} \cdot \frac{c_G}{K_S^G + c_G} \cdot \frac{c_A}{K_S^A + c_A} \cdot \frac{K_P}{(c_P + K_P)}$	$r_P = \left(\alpha \cdot \frac{dc_X}{dt} + \beta \cdot c_X\right) \cdot \left(1 - \frac{c_P}{c_{P,max}}\right)$
10	$\mu = \mu_{max} \cdot \frac{c_A}{K_S^A + c_A + \frac{c_A^2}{K_I}}$	$r_P = \alpha \cdot \mu + \beta$

an alternative, pyruvate formation was modeled using a modified Michaelis–Menten equation for non-competitive product inhibition [24] (see models 2, 4, 6, and 8). Finally, model 9 includes a modified Luedeking–Piret equation together with the Levenspiel term to describe pyruvate formation and product inhibition. The proposed models are summarized in Table 1.

Substrate uptake

The biomass dependent substrate consumption rates are expressed in Eqs. 1 and 2. Yields such as biomass/glucose $Y_{X/G}$ and biomass/acetate $Y_{X/A}$ were assumed to be functions of the biomass growth and maintenance energy demand (Eqs. 3 and 4).

$$r_G = \frac{\mu}{Y_{X/G}} \tag{1}$$

$$r_A = \frac{\mu}{Y_{X/A}} \tag{2}$$

$$Y_{X/G} = \frac{Y_{X/G,\,max} \cdot \mu}{Y_{X/G,\,max} \cdot m_G + \mu} \tag{3}$$

$$Y_{X/A} = \frac{Y_{X/A,\,max} \cdot \mu}{Y_{X/A,\,max} \cdot m_A + \mu} \tag{4}$$

A simplified overview of the basic model characteristics is given in Table 2.

The mass balance model for the fed-batch fermentation

The mathematical model (Eqs. 5, 6, 7, 8, and 9) for the fed-batch fermentation of recombinant *E. coli* based on the mass balance of the components in a fed-batch reactor is represented by the following differential equations:

$$\frac{dc_X}{dt} = -\frac{q_V}{V} \cdot c_X + \mu \cdot c_X \tag{5}$$

$$\frac{dc_G}{dt} = -\frac{q_V}{V} \cdot c_G + \frac{q_{VG}}{V} c_{G,0} - r_G \cdot c_X - r_P \cdot c_X \tag{6}$$

$$\frac{dc_A}{dt} = -\frac{q_V}{V} \cdot c_A + \frac{q_{VA}}{V} c_{A,0} - r_A \cdot c_X \tag{7}$$

$$\frac{dc_P}{dt} = -\frac{q_V}{V} \cdot c_P + r_P \cdot c_X \cdot Y_{P/G} \tag{8}$$

$$\frac{dV}{dt} = q_{VG} + q_{VA} = q_V \tag{9}$$

where V is the biosuspension volume, q_V is the time-dependent overall volumetric flow rate, q_{VG} and q_{VA} are the analogues for glucose (G) and acetate (A), and $c_{G,0}$ and $c_{A,0}$ are glucose and acetate concentration in the feed.

Experimental

Strain and medium

E. coli YYC202*ldhA::Kan* [1, 6, 7] was stored at −80°C in LB medium containing glycerol (50%). The culture was reactivated by inoculating a frozen aliquot (500 μL) into an Erlenmeyer flask with 200 mL mineral medium, which was incubated at 37°C on a rotary shaker (3033, GFL GmbH, Burgwedel, Germany) for 15 h.

The fermentation medium contained per liter: 1.50 g $NaH_2PO_4 \cdot H_2O$, 3.25 g KH_2PO_4, 2.50 g K_2HPO_4, 0.20 g NH_4Cl, 2.00 g $(NH_4)_2SO_4$, 0.50 g $MgSO_4$, 1 mL trace element solution, 11 g glucose monohydrate, and 0.79 g potassium acetate. Trace element solution contained per liter: 10.00 g $CaCl_2 \cdot 2H_2O$, 0.50 g $ZnSO_4 \cdot 7H_2O$, 0.25 g $CuCl_2 \cdot 2H_2O$, 2.50 g $MnSO_4 \cdot H_2O$, 1.75 g $CoCl_2 \cdot 6H_2O$,

Table 2 Kinetics used for description of biomass growth and product formation in the developed models. Microbial specific growth rate approaches: multiple substrate Monod kinetics [17, 18, 19], growth inhibition by product formation according to the Levenspiel [20], growth inhibition by product formation described by the Jerusalimsky [21], Monod kinetics and acetate growth inhibition described by Andrews kinetics [22]. Product formation kinetics: product formation based on the Luedeking–Piret [23], product formation according to the Michaelis–Menten equation for non-competitive inhibition with product [24] and product formation and product inhibition described by the modified Luedeking–Piret equation together with the Levenspiel term (L–P Levenspiel)

Model	Biomass growth					Product formation		
	Multiple substrate Monod	Levenspiel	Jerusalimsky	Monod	Andrews	Luedeking–Piret	Michaelis–Menten	L–P Levenspiel
1	+					+		
2	+						+	
3			+			+		
4			+				+	
5		+				+		
6		+					+	
7				+		+		
8				+			+	
9		+						+
10					+	+		

0.12 g H_3BO_3, 2.50 g $AlCl_3 \cdot 6H_2O$, 0.50 g $Na_2MoO_4 \cdot 2-H_2O$, 10.00 g $FeSO_4$. Shake flask growth medium contained per liter: 3.00 g Na_2HPO_4, 1.50 g KH_2PO_4, 0.25 g NaCl, 0.50 g NH_4Cl, 0.05 g $MgSO_4$, 0.01 g $CaCl_2$, 2.00 g glucose monohydrate, 0.10 g potassium acetate. The fed-batch fermentation feed medium consisted of 700 g L^{-1} glucose monohydrate and 109 g L^{-1} potassium acetate.

Fed-batch fermentation

All experiments were carried out in a 7.5 L bioreactor (INFORS AG, Bottmingen, Switzerland) containing 2.25 L fermentation medium, which was equipped with standard control units for pH, pressure, temperature, aeration, stirrer speed, etc.

After sterilization of the bioreactor and peripheral equipment, fermentation medium was filled into the bioreactor through a sterile microfiltration unit (0.2 μm cut-off, Sartobran, Sartorius AG, Göttingen, Germany) and pH was adjusted at 7.0 by 25% ammonia titration, also during the later fermentation process. All experiments were carried out at the temperature of 37°C. Sufficient aeration (DO (dissolved oxygen) ≥ 40%) was obtained by vigorous stirring (200–1,800 rpm), airflow rate (1–10 L min^{-1}), and reactor overpressure (0.2–0.8 bar). A reflux cooler condensed the outlet gas stream. The condensate was returned to the medium. Exhaust gases were analyzed by the gas analyzer (Binos 100 2 M, Rosemont Analytical/Process Analytical Division, Orville, OH, USA). The bioreactor was inoculated with 10% of the working volume of a preculture.

The fed-batch experiments were performed under different glucose-feeding conditions. All considered a constant feed at 10, 20, and 30 mL h^{-1}, respectively. The acetate was fed by applying the previously developed indirect acetate control ensuring acetate saturating conditions. According to the heuristic approach, acetate consumption was calculated based on the on-line estimated CO_2 production rate. As a result of our experimental studies an optimum equimolar ratio between acetate consumption rate and CO_2 production rate was identified [1, 2].

Analytical methods

The optical density (OD) was measured in a double beam spectrophotometer (UV-160, Shimadzu, Kyoto, Japan) at 600 nm. Cell dry weight was measured by filtration of 2.5 mL fermentation broth using pre-weighed microfilters (0.2 μm cut-off, Schleicher & Schuell, Dassel, Germany). After drying for 48 h at 80°C, the filters were cooled in an exsiccator for another 48 h. After filter weighing, the cell dry mass was calculated. Glucose concentration was measured off-line by the enzyme-based biosensor appliance Accutrend (F. Hoffmann–La Roche Ltd, Basel, Switzerland).

The concentrations of pyruvate and acetate in the fermentation supernatant were measured using the HPLC. Determination of the organic acid concentration involves the use of two Aminex HPX-87H (Bio-Rad Laboratories GmbH, Munich, Germany) columns in series. The separation was performed with 0.2 M H_2SO_4 solution at a flow rate of 0.5 mL min^{-1} (PUMP S1000, Sykam Chromatografie Vertriebs GmbH, Eresing, Germany) and the detection was at a wavelength of $\lambda = 254$ nm (diode array detector, DAD). The signal was analyzed with an integrator (C-R3A, Sykam Chromatografie Vertriebs GmbH) and the sample volume used was 100 μL.

Data handling

The model parameters were estimated by non-linear regression analysis and they were optimized using the Nelder–Mead algorithm [25]. The numerical values of the parameters were evaluated by fitting the model to the experimental data with the "Scientist" [26] software (MicroMath Inc., St. Louis, MO, USA). The model equations were solved numerically by the fourth order Runge–Kutta algorithm, which is also offered in the same software. The set of estimated parameters has been used for the simulation (Figs. 1, 2, and 3).

The calculated data were compared with the experimental data, recalculated in the optimization routine and fed again to the integration step until minimal errors between experimental and integrated values was achieved (built-in Scientist). The residual sum of squares was defined as the sum of the squares of the differences between experimental and calculated data. For discrimination of various models, the minimal value of the residual sum of squares and the model selection criterion (MSC) [27] were used as trial functions. The MSC is defined as (Eq. 10):

$$\text{MSC} = \ln \left(\frac{\sum_{i=1}^{n} w_i \cdot \left(Y_{\text{obs}_i} - \overline{Y}_{\text{obs}}\right)^2}{\sum_{i=1}^{n} w_i \cdot \left(Y_{\text{obs}_i} - Y_{\text{cal}_i}\right)^2} \right) - \frac{2 \cdot p}{n} \qquad (10)$$

where n is the number of points, w_i is the weight applied to each points, $\overline{Y}_{\text{obs}}$ is the weighted mean of the observed data, Y_{obs_i} is the weighted value of observed data, Y_{cal_i} represents the weighted value of calculated data and p is the level of significance of the simulation. The MSC attempts to represent the "information content" of a given set of parameter estimates by relating the coefficient of determination to the number of parameters (or equivalently, the number of degrees of freedom) that were required to obtain the fit. When comparing two models with different numbers of parameters, this criterion imposes a burden on the model with more parameters. The most appropriate model will be the one with the largest MSC.

The "Episode" algorithm for a stiff system of differential equations, implemented in the "Scientist"
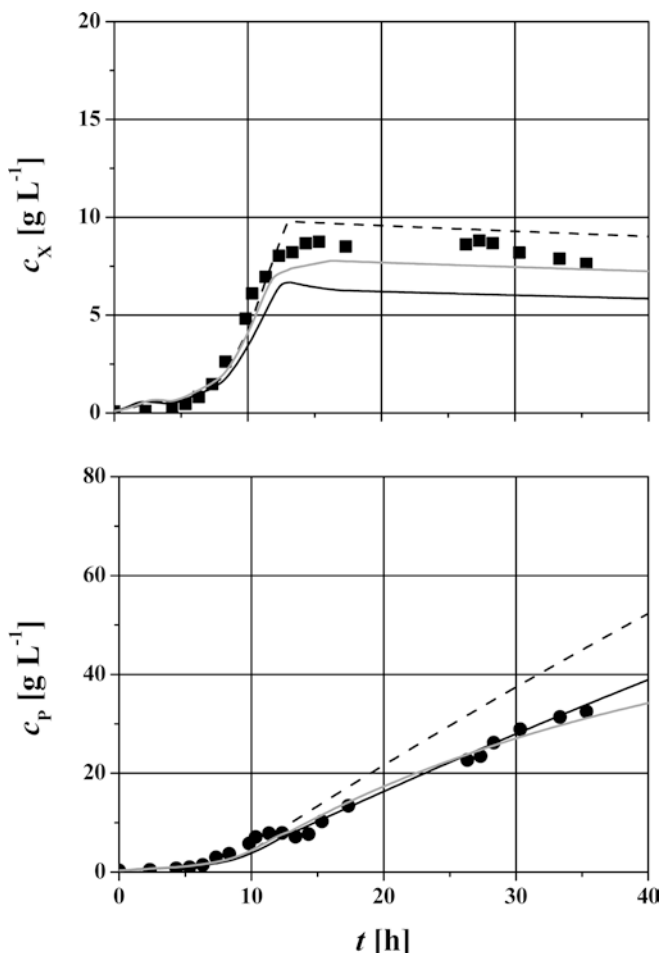
Fig. 1 Data obtained by model simulation (model 5, *solid line*; model 7, *dashed line*; model 9, *gray line*), and experimental data for biomass (*square*) and pyruvate (*circle*) concentration in fed-batch process at constant glucose volumetric flow rate of $q_{VG} = 10$ mL h^{-1}
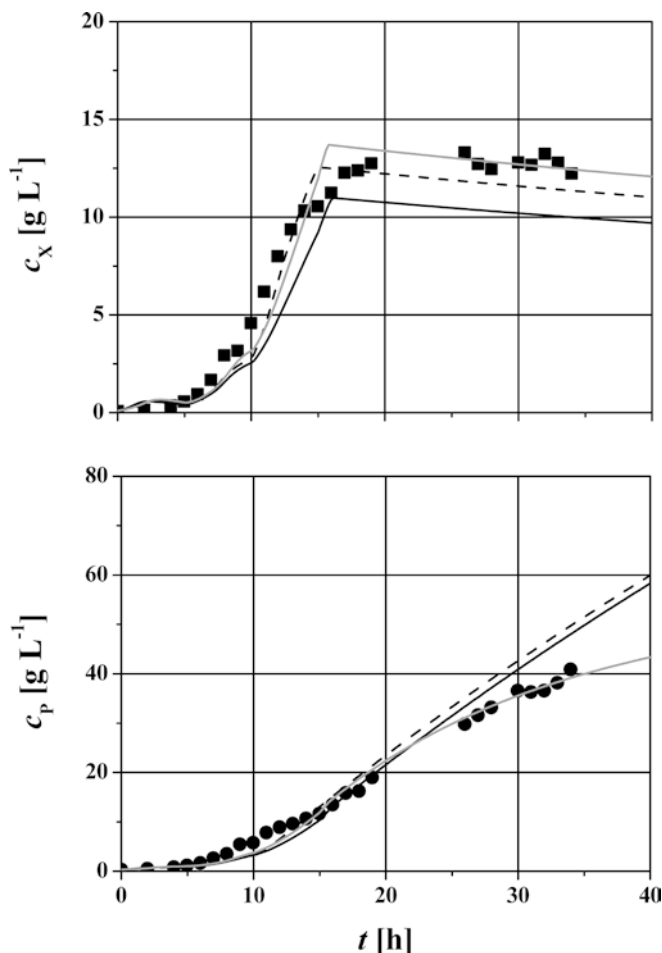
Fig. 2 Data obtained by model simulation (model 5, *solid line*; model 7, *dashed line*; model 9, *gray line*), and experimental data for biomass (*square*) and pyruvate (*circle*) concentration in fed-batch process at constant glucose volumetric flow rate of $q_{VG} = 20$ mL h^{-1}

software package, was used for the simulations. It uses variable coefficient Adams–Moulton and backward differentiation formula methods in the Nordsieck form, treating the Jacobian matrix as full or banded.
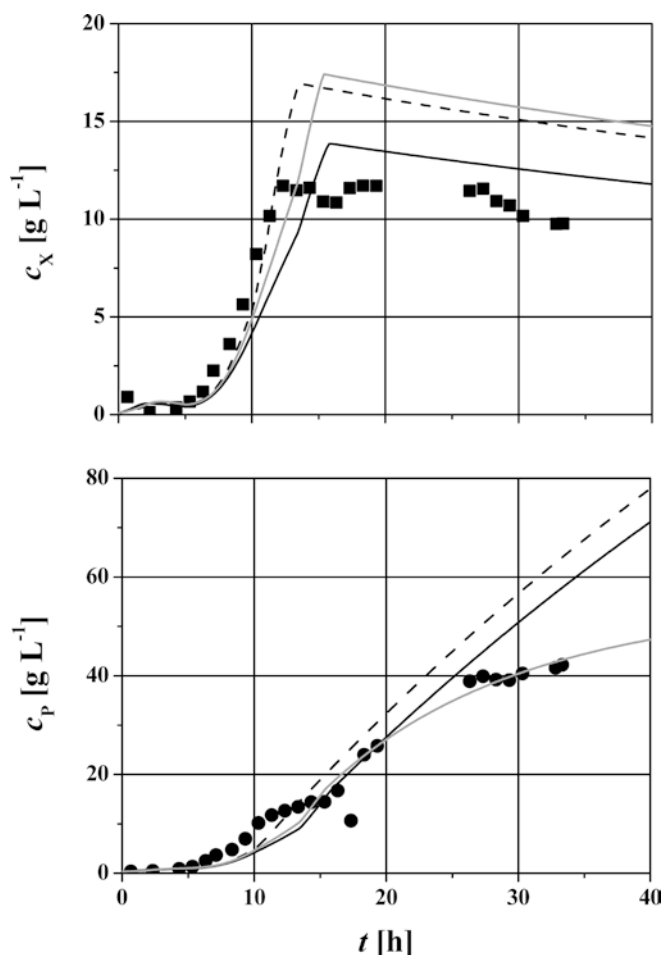
Results and discussion

Estimation of parameters

As shown in the preceding section, all models consisted of a set of five differential equations (Eqs. 5, 6, 7, 8, and 9) thus representing five dependent state variables and up to 13 parameters. State variables such as biomass concentration c_X, glucose concentration c_G, acetate concentration c_A, pyruvate concentration c_P, and reaction volume were taken from the experimental studies presented elsewhere [1, 2]. At the beginning, a primary model identification was performed using the data obtained from the fed-batch fermentation with $q_{VG} = 10$ mL h^{-1}. As already mentioned, the model

identification based on the least-square method to minimize difference between experimental and calculated values of state variables. Model validation was performed using two alternative data sets, namely those of the fed-batch fermentations when glucose feed rates of $q_{VG} = 20$ and 30 mL h^{-1} were installed. In each case, the MSC criterion was used to qualify the model suitability for reflecting the experimental findings.

Initial values of parameters $Y_{X/G,max}, Y_{X/A,max}, K_S^G$, and K_S^A were those previously identified for growth of *E. coli* K12 on a single substrate [28]. The initial value of the parameter $Y_{P/G}$ was calculated from the experimental data (fed-batch fermentation performed at $q_{VG} = 10$ mL h^{-1}). In all models $c_{P,max}$ was set to be 63.6 g L^{-1}, which represents the experimentally observed value [1, 2]. All other model parameters were initially estimated so that negative estimates of simulated concentration curves were avoided.

Table 3 provides an overview of the residual sum of squares and MSC values obtained after the modeling of the first data set. As indicated, the models, 5, 7, and 9

Table 3 Residual sum of squares and model selection criterion (MSC) for different models

Model	Residual sum of squares	MSC
1	$3.97 \cdot 10^2$	2.18
2*	$1.15 \cdot 10^3$	1.09
3	$5.58 \cdot 10^2$	1.82
4*	$1.39 \cdot 10^3$	0.88
5	$2.46 \cdot 10^2$	2.68
6	$8.37 \cdot 10^2$	1.39
7	$2.06 \cdot 10^2$	2.59
8	$4.31 \cdot 10^2$	2.14
9	$1.99 \cdot 10^2$	2.77

*Negative values of estimated parameters.

Fig. 3 Data obtained by model simulation (model 5, *solid line*; model 7, *dashed line*; model 9, *gray line*), and experimental data for biomass (*square*) and pyruvate (*circle*) concentration in fed-batch process at constant glucose volumetric flow rate of $q_{VG} = 30$ mL h^{-1}

have the similar and the lowest sum of square errors. Additionally, the MSC values are the highest for models 5, 7, and 9 thus stressing their convenience. As a consequence, these models were favored for further analysis.

Interestingly, models 2, 4, 6, and 8, in which pyruvate formation was described according to the modified Michaelis–Menten equation with non-competitive product inhibition, achieved the lowest accuracy level. However, the consideration of the Luedeking–Piret approach (models 1, 3, 5, 7, and 9) enabled better model predictions. Furthermore, some parameter estimates of the models 2 and 4 were negative, which is meaningless in a biological sense (data not shown). It is also noteworthy that the multiple substrate Monod kinetics (models 1 and 2) and multiple substrate Monod kinetics with the Levenspiel term for product inhibition were not able to describe the behavior of the process adequately. Furthermore, it is known that high acetate concentrations in the fermentation medium lead to inhibition of biomass growth. This was the reason for considering Andrews kinetics [22] in model 10. Unfortunately, the

kinetics were not able to describe experimental results acceptably (Table 3).

For the sake of brevity, only the parameter estimates (together with the confidence intervals) of the most favored models, namely models 5, 7, and 9, are listed in Table 4. The calculated values obtained by simulation compared to the experimental data points, for biomass and pyruvate concentrations, are shown in Fig. 1.

As it can be seen, all models are able to mirror the dynamic behavior of the process. The simulated biomass concentration curves obtained from model 5 were significantly lower than the measured values. Additionally, the simulated pyruvate curves of model 7 were higher than the experimental data (Fig. 1). In general, relatively large error bars were estimated for almost all parameters of model 5 (except for the parameters α and β of the Luedeking–Piret approach), which are probably caused by inadequate kinetic assumptions. On the contrary, relatively high parameter accuracies were achieved using the models 7 and 9.

Despite the high parameter accuracy and the relatively high model predictive quality of model 7 this model should be not favored because acetate was assumed to be the only limiting substrate for growth. As already stated, this simplifying assumption motivated the acetate auxotrophy of the *E. coli* strain. However, the strain was not capable of growing without glucose, which, basically, would necessitate considering additional intracellular biochemical reactions and thus obviously contradict the initial motivation to solely use unstructured models.

The predicted biomass and pyruvate curves of model 9 reflect well the experimental data. On the basis of a single criterion, the residual sum of squares, model 9 should thus be favored because it achieved the lowest residual sum of squares. Additionally, the estimated parameters in this model have acceptable confidence interval and the model is mechanistically correct. Furthermore, model 9 is mechanistically the most accurate one because it contains all experimentally observed effects: growth inhibition by pyruvate and pyruvate inhibited product (pyruvate) formation.

Some parameter values given in Table 4, can be qualified as biologically meaningless and are the result of the numerical Nelder–Mead parameter identification

Table 4 Estimated values and confidence intervals of parameters for the models 5, 7, and 9

Parameter	Model 5	Model 7	Model 9
$Y_{X/G,max}$	1.45 ± 0.74	–	$9.18\ 10^{-1} \pm 1.01\ 10^{-1}$
$Y_{X/A,max}$	$9.21\ 10^{-1} \pm 2.18\ 10^{-1}$	$8.91\ 10^{-1} \pm 5.14\ 10^{-1}$	$1.52 \pm 7.52\ 10^{-1}$
$Y_{P/G}$	$6.78\ 10^{-1} \pm 3.12\ 10^{-1}$	$5.49\ 10^{-1} \pm 4.90\ 10^{-2}$	$9.32\ 10^{-1} \pm 3.80\ 10^{-1}$
μ_{max}	2.35 ± 2.27	$1.31 \pm 9.04\ 10^{-1}$	$1.69 \pm 2.87\ 10^{-1}$
K_S^G	$1.78\ 10^{-2} \pm 4.76\ 10^{-2}$	–	$1.10\ 10^{-1} \pm 1.05\ 10^{-1}$
K_S^A	$2.93\ 10^{-1} \pm 9.59\ 10^{-2}$	$2.46\ 10^{-1} \pm 5.49\ 10^{-3}$	$3.22\ 10^{-1} \pm 1.94\ 10^{-1}$
m_G	$4.41\ 10^{-2} \pm 6.14\ 10^{-2}$	–	$4.90\ 10^{-2} \pm 1.77\ 10^{-2}$
m_A	$1.26\ 10^{-1} \pm 3.44\ 10^{-2}$	$1.13\ 10^{-1} \pm 8.30\ 10^{-2}$	$1.47\ 10^{-1} \pm 6.60\ 10^{-2}$
α	$7.91\ 10^{-1} \pm 3.53\ 10^{-1}$	$5.34\ 10^{-1} \pm 2.50\ 10^{-1}$	$6.10\ 10^{-1} \pm 3.34\ 10^{-1}$
β	$2.03\ 10^{-1} \pm 8.00\ 10^{-2}$	$1.79\ 10^{-1} \pm 3.93\ 10^{-2}$	$2.14\ 10^{-1} \pm 2.89\ 10^{-2}$
K_P	1.77 ± 2.87	–	3.93 ± 1.32

considering partially non-appropriate macrokinetic models. For example, the yield coefficients $Y_{X/G,max}$ and also μ_{max} appear to be too high. However, if large confidence intervals of these parameters (μ_{max}, $Y_{X/G,max}$, $Y_{X/A,max}$, m_G, m_A, etc.) are taken into account the parameter values are fairly reasonable (Table 4). Preferably, these parameters should have been estimated directly from the experiments which—unfortunately—was not possible because of the properties of the *E. coli* YYC202 *ldhA::Kan* used. For example, cell growth is realized by the simultaneous consumption of the C sources glucose and acetate, thus making independent growth-parameter estimation in fed-batch experiments impossible. Nevertheless, a simple, easy-to-use, and robust model has successfully been identified, which can be used for typical process engineering applications such as process optimization and design. This model (model 9) and the estimated parameter are the first step in the quantitative understanding of the basic cellular characteristics of the genetically modified production strain used in this study.

Validation of the models

Kinetic modeling is still a difficult task for bioprocesses, where a succession of decisions must be made in respect of which systematic methodologies are still lacking. Therefore it is important to validate a model by using data other than those applied to identify parameters. In order to further validate the selected models (models 5, 7, and 9), further fed-batch experiments were performed at constant volumetric flow rate of glucose of $q_{VG} = 20$ and $30\ \mathrm{mL\ h^{-1}}$ as mentioned before. For model validation experimental data were compared to data obtained by model simulation for biomass and pyruvate concentrations (Figs. 2 and 3).

The good fitting quality of the model 9 to the biomass concentration and especially to the pyruvate concentration is remarkable for both volumetric glucose flow rates. The models 5 and 7 were not able to predict biomass and pyruvate concentrations with acceptable accuracy. Pyruvate levels predicted by those models were approximately 50% higher than experimental observations.

Unfortunately, all developed models were not able to predict glucose concentration profiles adequately especially when high glucose feeds were used ($q_{VG} = 30\ \mathrm{mL\ h^{-1}}$). Rising glucose levels were suggested by all models for the pyruvate production period starting after 20 h fermentation time (data not shown). This was in accordance with the primary assumption that glucose is significantly used for biomass growth, which obviously did not occur during the pyruvate production. However, no glucose accumulation was observed experimentally indicating that a higher amount of glucose was consumed for pyruvate production, by-product formation and maintenance as primarily expected. In this regard, the working hypothesis is remarkable in that high extracellular pyruvate concentrations (which occur during the production phase) could cause futile cycling owing to active pyruvate export and diffusive pyruvate import (see [2]). Obviously, as yet, this significant impact has not been covered by the models, nor has the fact that approximately 50% of the acetate used is converted into CO_2 production. Both aspects would necessitate a structured model formulation and could not be covered by the current modeling approaches used.

Sensitivity analysis

In order to qualify the sensitivity of model predictions with respect to estimated parameter errors, a sensitivity analysis has been performed [29]. The parameter values of model 9 (Table 4) were taken as reference values. The solutions of the model equations are then calculated with relative parameter errors ranging from −50% to +50% of the reference values. The final concentration of biomass c_X and the final pyruvate concentration c_P were than compared to the reference concentrations. All simulations were performed for the period of 35 h. The results are shown in Fig. 4.

Both, calculated biomass and pyruvate concentration, are sensitive to yield coefficients, $Y_{X/G}$, $Y_{X/A}$, $Y_{P/G}$. This effect was expected because decreasing yields thus indicate reduced final titers provided that the same amount of substrate is used. Interestingly, there was almost no effect on the investigated concentrations
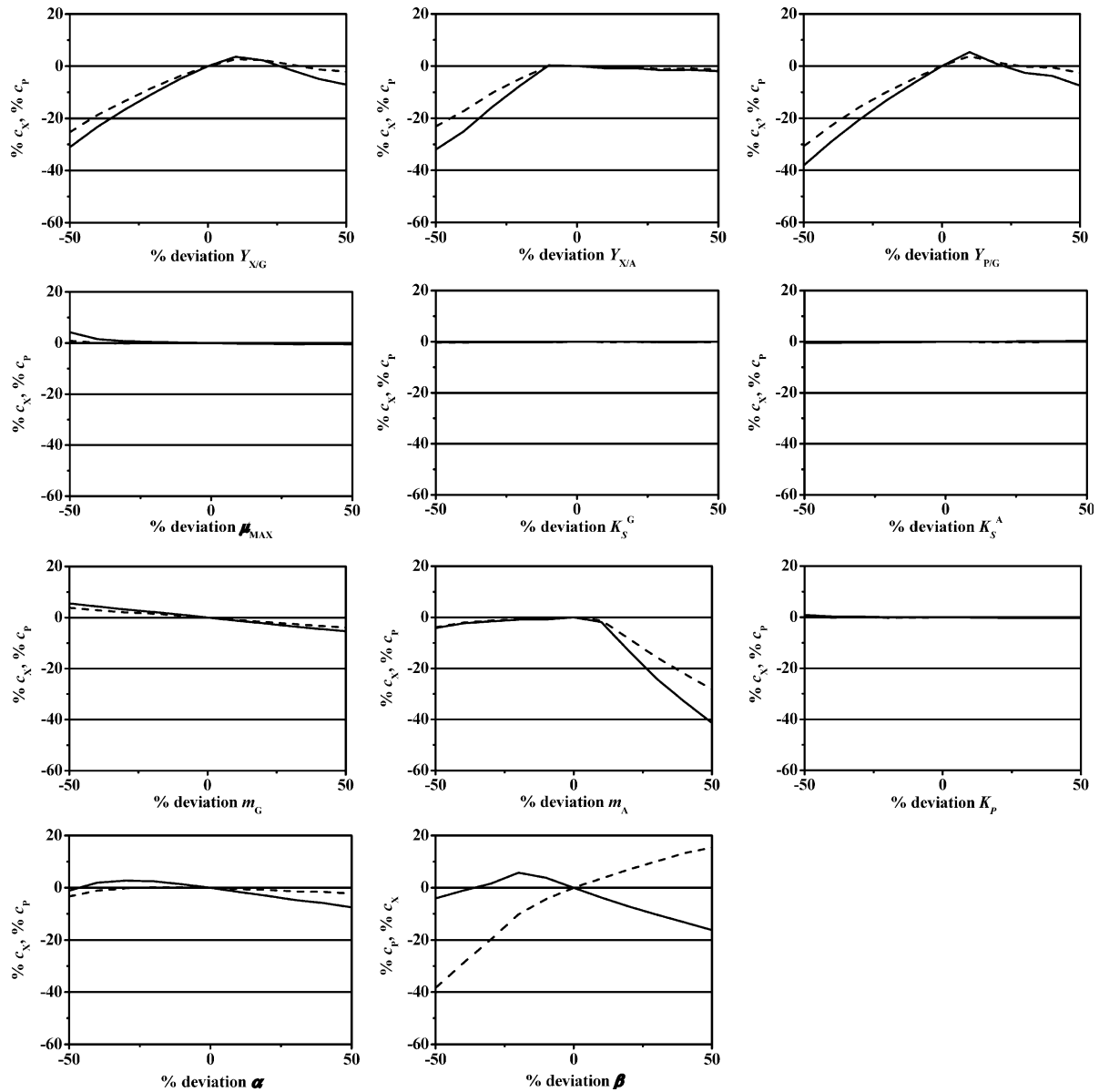
Fig. 4 Sensitivity analysis for the kinetic parameters of the model. The changes in the final value of the biomass (*solid line*) and pyruvate (*dashed line*) concentrations are represented with respect to the deviation of the nominal value of considered parameter

when yield coefficients are higher than their nominal value.

The maintenance coefficient for growth on acetate, m_A, was found to be the most sensitive parameter for the achievable biomass concentration. If m_A was higher than the reference value, the final biomass concentration reduced linearly. On the contrary, there was no effect on the biomass titer if m_A values were lower than the reference. m_A also significantly affects the pyruvate concentrations (as shown in Fig. 4). The non-growth-associated product formation coefficient β was found to be most important parameter for product formation. Increasing β caused linear pyruvate concentration rises.

Both, calculated biomass and pyruvate concentrations are slightly sensitive to maintenance coefficient for growth on glucose m_G and growth associated product formation coefficient α.

Finally, the model predictions are almost insensitive to the parameters μ_{max}, K_S^G, K_S^A, and K_P. Unfortunately, this clearly indicates the necessity to improve the parameter accuracy in further studies. For this, the optimal experimental design method focusing on continuous (or fed-batch) experiments should be applied [11, 13].

Conclusions

Ten simple, unstructured models have been developed for the bioconversion of glucose to pyruvate in a fed-batch process by genetically modified *E. coli* strain. All

models were identified by the least-square fit and qualified by the MSC thus allowing the estimation of corresponding model parameters. It is noteworthy that the model identification was based on a different data set than the model validation. Within the investigated experimental range the model, which combines growth inhibition by pyruvate (Jerusalimsky approach), and pyruvate inhibited product formation (described by modified Luedeking–Piret/Levenspiel term) should be favored. Using this modeling approach, an acceptable model prediction for cell growth and pyruvate formation was achieved, which are both essential variables to model process alternatives and scale-ups. In the case of glucose, further studies must be performed to increase the model predictive quality.

In principle, even the favored model 9 represents only the first step in the modeling of the pyruvate production process, strictly speaking, by uncovering the limits of the unstructured model used. The need to incorporate additional aspects of energy demand or by-product formation was identified with necessities to extend the existing model by structured modeling terms. Nevertheless, the identified model 9 should be qualified as a promising tool for modeling studies as well as for further, more detailed, kinetic modeling approaches.

References

1. Zelić B (2003) Study of the process development for *Escherichia coli* based pyruvate production. Ph Dissertation, Faculty of Chemical Engineering and Technology, University of Zagreb, Croatia (in English)
2. Zelić B, Gerharz T, Bott M, Vasić-Rački Đ, Wandrey C, Takors R (2003) Fed-batch process for pyruvate production by recombinant *Escherichia coli* YYC202 strain. Eng Life Sci 3:299–305
3. Fell D, Wagner A (2000) The small world of metabolism. Nature Biotechnol 18:1121–1122
4. Howard JW, Fraser WA (1932) Pyruvic acid. Org Synth Coll 1:475–476
5. Li Y, Chen J, Lun S-Y (2001) Biotechnological production of pyruvic acid. Appl Microbiol Biotechnol 57:451–459
6. Gerharz T, Zelić B, Takors R, Bott M (2002) Processes and microorganisms for microbial production of pyruvate from carbohydrates and alcohols (in German). German Patent Application 10220234.6
7. Zelić B, Gostović S, Vuorilehto K, Vasić-Rački Đ, Takors R (2004) Process strategies to enhance pyruvate production with recombinant *Escherichia coli*: from repetitive fed-batch to in situ product recovery with fully integrated electrodialysis. Biotechnol Bioeng 85:638–646
8. Nielsen J, Emberg C, Halberg K, Villadsen J (1989) Compartment model concept used in the design of fermentation with recombinant microorganisms. Biotechnol Bioeng 34:478–486
9. Nielsen J, Pederson AG, Strudsholm K, Villadsen J (1991) Modeling fermentations with recombinant microorganisms: formulation of a structured model. Biotechnol Bioeng 37:802–808
10. Daae EB, Ison A P (1998) A simple structured model describing the growth of *Streptomyces lividans*. Biotechnol Bioeng 58:263–266
11. Baltes M, Schneider R, Sturm C, Reuss M (1994) Optimal experimental design for parameter estimation in unstructured growth models. Biotechnol Prog 10:480–488
12. Canovas M, Maiquez JR, Obon JM, Iborra JL (2002) Modeling of the biotransformation of crotonobetaine into L-(-)-carnitine by *Escherichia coli* strains. Biotechnol Bioeng 77:764–775
13. Takors R, Wiechert W, Weuster-Botz D (1997) Experimental design for the identification of macrokinetic models and model discrimination. Biotechnol Bioeng 56:564–576
14. Munack A (1989) Optimal feeding strategy for identification of Monod-type models by fed-batch experiments. In: Computer applications in fermentation technology modelling and control of biotechnological processes. Elsevier, Amsterdam, pp 195–204
15. Nihtila M, Virkkunen J (1977) Practical identifiability of growth and substrate consumption models. Biotechnol Bioeng 19:1831–1850
16. Versyck KJ, Claes JE, Van Impe JF (1998) Optimal experimental design for practical identification of unstructured growth models. Math Comp Sim 46:621–629
17. Bailey JE, Ollis DF (1977) Biochemical engineering fundamentals. McGraw-Hill, New York
18. Dunn IJ, Heinzle E, Ingham J, Prenosil JE (1992) Biological reaction engineering: principles, applications and modelling with PC simulation. VCH, Weinheim
19. Bae W, Rittman BE (1996) A structured model of dual-limitation kinetics. Biotechnol Bioeng 49:683–689
20. Han K, Levenspiel O (1988) Extended Monod kinetics for substrate, product and cell inhibition. Biotechnol Bioeng 32:430–437
21. Roels JA (1983) Energetics and kinetics in biotechnology. Elsevier, Amsterdam
22. Andrews JF (1968) A mathematical model for the continuous culture of microorganisms utilizing inhibitory substrates. Biotechnol Bioeng 10:707–723
23. Luedeking R, Piret EL (2000) A kinetic study of the lactic acid fermentation. Batch process at controlled pH. Biotechnol Bioeng 67:633–644
24. Blanch HW, Clark DS (1997) Biochemical engineering. Marcel Dekker, New York
25. Nelder JA, Mead R (1965) A simplex method for function minimisation. Comput J 7:308–331
26. Scientist handbook (1986–1995). MicroMath, Salt Lake City, UT
27. Akaike H (1976) An information criterion (AIC). Math Sci 14:5–9
28. Venkatesh KV, Doshi P, Rengaswamy R (1997) An optimal strategy to model microbial growth in a multiple substrate environment. Biotechnol Bioeng 56:635–644
29. Bernard O, Bastin G, Stentelaire C, Lesage-Meessen L, Asther M (1999) Mass balance modeling of vanillin production from vanillic acid by cultures of the fungus *Pycnoporus cinnabarinus* in bioreactors. Biotechnol Bioeng 65:558–571