

A. Emmanuel, D. C. López Cárdenas, P. Neubauer, M. N. Cruz-Bournazou

# Modelling overflow metabolism in *Escherichia coli* by acetate cycling

Journal article | Accepted manuscript (Postprint)

This version is available at <https://doi.org/10.14279/depositonce-8272>



Anane, E.; López Cárdenas, D. C.; Neubauer, P.; Cruz-Bournazou, M. N. (2017). Modelling overflow metabolism in *Escherichia coli* by acetate cycling. *Biochemical Engineering Journal*, 125, 23–30. <https://doi.org/10.1016/j.bej.2017.05.013>

## Terms of Use

Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.

WISSEN IM ZENTRUM  
UNIVERSITÄTSBIBLIOTHEK

Technische  
Universität  
Berlin

Anane E, López C DC, Neubauer P, Cruz-Bournazou MN. 2017. Modelling overflow metabolism in *E. coli* by acetate cycling. *Biochem Engin J.* 125, 23-30. <https://doi.org/10.1016/j.bej.2017.05.013>

**Modelling overflow metabolism in *Escherichia coli* by acetate cycling**

Emmanuel Anane\*, Diana C López C, Peter Neubauer, M Nicolas Cruz Bournazou

Chair of Bioprocess Engineering, Institute of Biotechnology, Technische Universität Berlin, Berlin, Germany

**\*Corresponding Author:**

Emmanuel Anane, Chair of Bioprocess Engineering, Institute of Biotechnology, Technische Universität Berlin, Ackerstraße 76, ACK24, 13355 Berlin, Germany.

Email: [anane@tu-berlin.de](mailto:anane@tu-berlin.de)

## **Abstract**

A new set of mathematical equations describing overflow metabolism and acetate accumulation in *E. coli* cultivation is presented. The model is a significant improvement of already existing models in the literature, with modifications based on the more recent concept of acetate cycling in *E. coli*, as revealed by proteomic studies of overflow routes. This concept opens up new questions regarding the speed of response of the acetate production and its consumption mechanisms in *E. coli*. The model is formulated as a set of continuous differentiable equations, which significantly improves model tractability and facilitates the computation of dynamic sensitivities in all relevant stages of fermentation (batch, fed-batch, starvation). The model is fitted to data from a simple 2 L fed-batch cultivation of *E. coli* W3110M, where twelve (12) out of the sixteen (16) parameters were exclusively identified with relative standard deviation less than 10%. The framework presented gives valuable insight into the acetate dilemma in industrial fermentation processes, and serves as a tool for the development, optimization and control of *E. coli* fermentation processes.

## **Keywords**

*Escherichia coli*; overflow metabolism; acetate; modelling; fed-batch

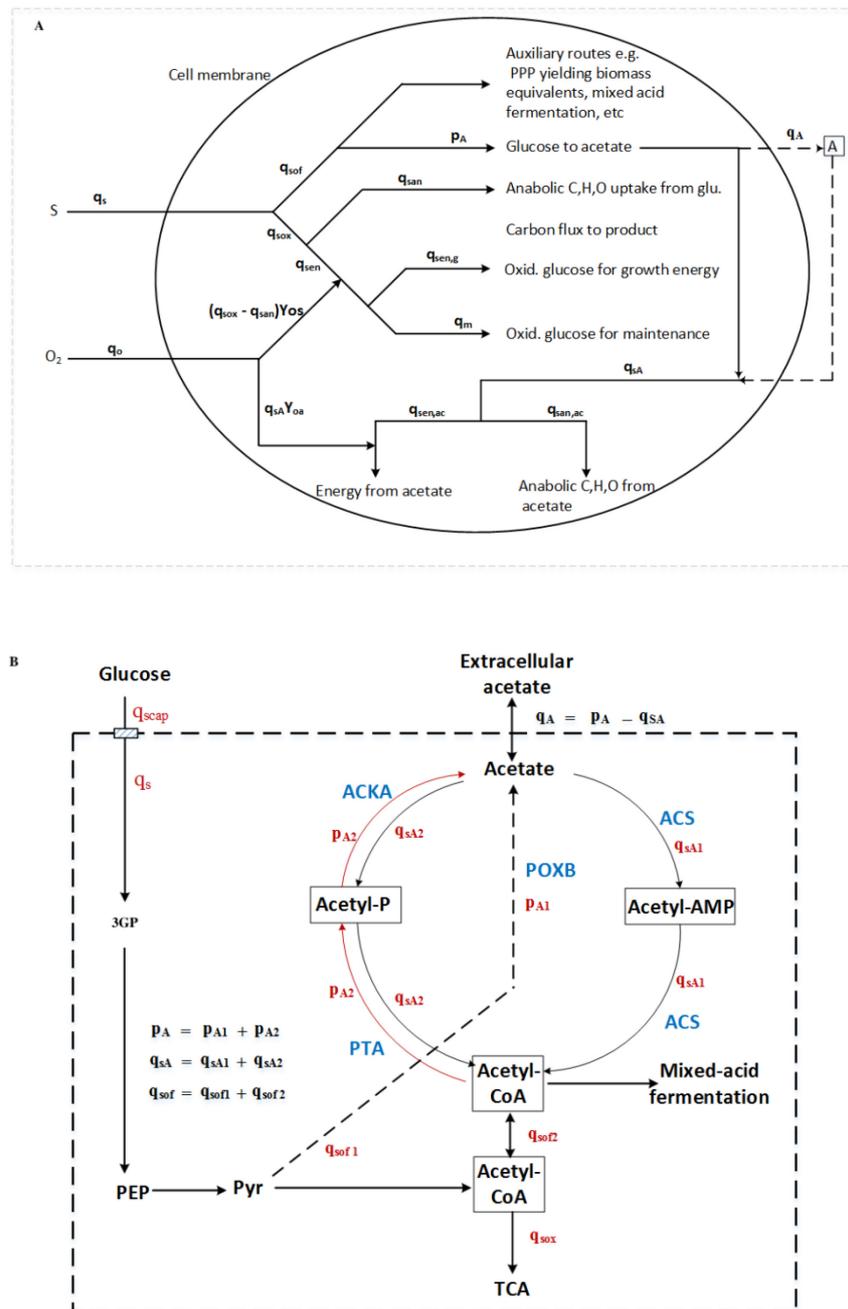
## 1.0 Introduction

One of the most embattled physiological phenomena in industrial scale *E. coli* cultivation is overflow metabolism and the associated excretion of acetate into the broth. Apart from the fact that extracellular acetate inhibits the growth of *E. coli* [1], redirection of the carbon source to acetate through overflow metabolism is wasteful in recombinant protein production processes [2]. The concept of overflow metabolism in *E. coli* is not new and a number of models in the literature have attempted to describe acetate production mechanistically [3–7]. In all these models, the acetate profile is presented in two distinct phases: an initial batch (overflow) phase during which acetate is produced, followed by a substrate limited phase during which acetate is consumed. Most of these models assume saturation of TCA cycle enzymes and are mainly built on discrete *conditional* statements in the metabolic routes [4,8]. Although these models suffice in describing acetate profiles during *E. coli* cultivation, a major limitation is the discontinuous nature of the functions (e.g, if  $q_O < q_{Omax}$ , then  $q_A = 0$  [8]), which makes further mathematical development and the use of sophisticated simulation and optimization programmes difficult. Furthermore, new evidence suggests that acetate conversion is indeed a continuous process. Peebo et al. [9] and Valgepea et al [10] used advanced proteomic analysis [9,11] and systems biology approaches [10] to show that intracellular production and re-assimilation of acetate (acetate cycling) is a continuous process in *E. coli* metabolism, even under non-overflow conditions. Two intermediates of this cycling process (Acetyl-AMP and Acetyl-P) were shown to play vital roles in *E. coli* motility and osmoregulation [1,12]. Thus, acetate excretion into the extracellular medium only results from an offset of the equilibrium between its production and re-assimilation, which is triggered by carbon catabolite repression, either at higher specific substrate uptake rates [11,13] or under anoxic conditions [11]. Therefore, a new mathematical representation is needed for model-based process development, which i) copes better with the real acetate conversion process in *E. coli* and ii) fulfils the requirements for gradient based algorithms and solution of partial differential equation systems.

### 1.1 Glucose partitioning and acetate cycling in *E. coli*

The mechanistic models of Xu [4], Lin [5] and Neubauer [8] are among the most widely used macro-kinetic models for description of *E. coli* fermentations. These were built on the concept

of glucose partitioning which was initially developed for *Saccharomyces cerevisiae* [14], and was shown to be equally applicable to the *E. coli* system [15].



**Figure 1 (A)** Glucose partitioning and oxygen usage in *E. coli*. Scheme for flux distribution used to model glucose (S) and oxygen ( $O_2$ ) usage in *E. coli* [4,5] with modifications showing acetate (A) excretion and re-assimilation across the cell membrane. **(B)** Enzyme-mediated acetate cycling in *E. coli*, with summaries of the glucose uptake system and upper glycolytic routes, adapted from Valgepea et al. 2010 [10], Li et al. 2014 [40] and Enjalbert et al. 2016 [32]. POXB—pyruvate oxidase, ACS—acetyl-CoA synthetase, PTA—phosphotransacetylase, ACKA—acetate kinase.

The structured model covers the intracellular partitioning of glucose for various physiological demands of the cell and the associated oxygen consumption (Figure 1 A). Additionally, the subsequent conversion of glucose to acetate through the overflow route is based on the ethanol process in *S. cerevisiae*. The large set of metabolic routes are approximated by lumped kinetic parameters, the values of which are estimated from measurable extracellular species in typical cultivation set-ups. Nevertheless, Valgepea et al. [10], Peebo et al. [9], and Martinez-Gomez et al [16] showed that *E. coli* uses a continuous acetate cycling system for acetate conversion (Figure 1 B). Under *typical* cultivation conditions and at low glucose uptake rates, acetate production is in equilibrium with its re-assimilation ( $p_A = q_{sA}$ ), hence there is neither net accumulation nor release of acetate into the extracellular medium. As the inflow of glucose into the cell increases, acetate consumption through the ACS pathway (acetate consumption) becomes insufficient, hence there is an offset in the cycling equilibrium ( $p_A > q_{sA}$ ), which results in intracellular accumulation of acetate followed by its excretion into the medium. This cycling system can be represented by a set of continuous equations that describe overflow as an off-set of the equilibrium between  $p_A$  and  $q_{sA}$ .

In the current paper, the concept of acetate cycling is used to derive a set of tractable and continuously differentiable equations at macro-kinetic scale to describe acetate production in *E. coli*. This new set of equations allow a cheap computation of both first and second order sensitivity functions, which are needed in gradient-based methods for optimal experimental design [17,18] and process optimization [19]. Finally, the model is validated with data from fed-batch cultivation of *E. coli* W3110M to show its applicability in real processes.

## 2.0 Materials and methods

### 2.1 Strain and cultivation conditions

The data for model calibration was obtained from a fed-batch culture of non-recombinant *E. coli* W3110M. The cultivation was conducted at a temperature of 37 °C and at pH of 7.0 in a 3.7 L BioEngineering® bench top bioreactor fitted with a polarographic dissolved oxygen probe. The medium consisted of mineral salts and trace elements, as described by Glazyrina et al [20], with 5 g/L glucose in the batch phase. To start the batch phase, 2 L of medium was inoculated with *E. coli* W3110 to OD<sub>600</sub> of 0.1. The fed-batch phase was initiated with a 300 g/L glucose feed at the exhaustion of the batch phase glucose, signalled by a sudden rise in dissolved oxygen tension. Two fed-batch regimes were implemented: an exponential feed followed by a constant feed. The exponential feed was implemented to maintain a set-point

specific growth rate ( $\mu_{set}$ ) of  $0.22 \text{ h}^{-1}$ . Using the biomass concentration ( $X$ , g/L) and broth volume ( $V$ , L) at the end of the batch phase, the exponential feed rate ( $F$ , L/h) was calculated as

$$F(t) = \frac{\mu_{set}}{Y_{x/s} S_i} (XV) e^{\mu_{set} t} \quad (1)$$

where  $S_i$  represents the glucose concentration in the feed solution (300 g/L) and  $t$  is the feed time. After 3 hours, the feed was switched to a constant feed, beginning at a specific growth rate of  $0.11 \text{ h}^{-1}$  for a period of 17 hours. To test the robustness of the acetate equations, intermittent glucose pulses were given in the constant feeding phase, and the corresponding response in all profiles were modelled. Cellular growth was monitored by measuring the optical density of samples at 600 nm in a UV-vis spectrophotometer (Novaspec III, Amersham Biosciences, Amersham, UK). Conversion factors that were developed with the same spectrophotometer and the same *E. coli* strain and given in Glazyrina et al [20] were used to convert  $OD_{600}$  values to cell dry weight.

## 2.2 Analyses

To determine the concentration of residual glucose and acetate, hourly samples were analysed on an Agilent 1200 HPLC system, equipped with a HyperRez<sup>TM</sup> XP Carbohydrate H+ column (Fisher Scientific, Schwerte, Germany) and a refractive index detector. As eluent, five mM  $\text{H}_2\text{SO}_4$  was used at a flow rate of  $0.5 \text{ mL min}^{-1}$ . In total,  $N_m = 19$  data points were collected and analysed for each observable variable i.e., biomass ( $X$ ), glucose ( $S$ ) and acetate ( $A$ ), whereas  $N_m = 11336$  data points were logged online for measured dissolved oxygen (DOT).

## 2.3. Parameter estimation

The *E. coli* model in Section 3 was solved using Matlab R2015a<sup>®</sup> with the CVODE integrator from SUNDIALS TB [21]. The parameter estimates (PE) were computed using *lsqnonlin* with the trust region reflective algorithm. The initial parameter values for PE were based on literature [4,5,17]. The initial values for the search in the parameter space were generated using the Minimum bias Latin hypercube design (MBLHD) [22]. The dynamic model can be written in its general form as follows,

$$\dot{x}(t) = f(x(t), z(t), u(t); \theta) \quad (2)$$

$$0 = g(x(t), z(t), u(t); \theta) \quad (3)$$

$$y_{of}(t) = A x(t) \quad (4)$$

$$y_{on}(t) = B x(t) \quad (5)$$

$$x(t_0) = x_0 \quad (6),$$

where the set of differential equations  $f$  corresponds to Equations 12, 14, 18, 22 and 24 whereas the algebraic equation set  $g$  refers to Equations 13, 15-17, 19-21 and 23,  $t \in [t_0, t_{end}] \subseteq \mathbb{R}$  is the independent time variable,  $x(t) \in \mathbb{R}^{N_x}$  and  $z(t) \in \mathbb{R}^{N_z}$  are the differential and algebraic state variables, respectively;  $u(t) \in \mathbb{R}^{N_u}$  are the time-varying inputs or experimental design variables and  $\theta \in \mathbb{R}^{N_p}$  is the unknown parameter vector. The vector  $y_{of}(t) \in \mathbb{R}^{N_{y_{of}}}$  are the predicted offline response variables (variables corresponding to sampled measurements) whose elements are defined by the selection matrix  $A \in \mathbb{R}^{N_{y_{of}} \times N_x}$ . The vector  $y_{on}(t) \in \mathbb{R}^{N_{y_{on}}}$  are the predicted response variables measured online (i.e., oxygen) whose elements are defined by the selection matrix  $B \in \mathbb{R}^{N_{y_{on}} \times N_x}$ . Note that not all states were measured, therefore  $N_{y_{of}} + N_{y_{on}} < N_x$ .

The model parameters were estimated by solving the optimization problem

$$\hat{\theta} := \arg \min_{\theta} \Phi(U, \theta) \quad (7)$$

where the cost function  $\Phi(U, \theta)$ , which is the weighted nonlinear least-squares criterion between the model predictions  $Y(U, \theta)$  and the experimental data  $Y^m$  was calculated as

$$\Phi(U, \theta) := \frac{1}{2} (Y(U, \theta) - Y^m)^T (C_y)^{-1} (Y(U, \theta) - Y^m) \quad (8)$$

All the measured data were collected in the vector  $Y(U, \theta)$ ,

$$\begin{aligned} Y(U, \theta) &:= ((y_{of_1}(t_1, U, \theta), \dots, y_{of_1}(t_{Nm_{of}}, U; \theta))^T, \dots, (y_{of_{N_{y_{of}}}}(t_1, U, \theta), \dots, y_{of_{N_{y_{of}}}}(t_{Nm_{of}}, U; \theta))^T, \\ &\quad (y_{on_1}(t_1, U, \theta), \dots, y_{on_1}(t_{Nm_{on}}, U; \theta))^T, \dots, (y_{on_{N_{y_{on}}}}(t_1, U, \theta), \dots, y_{on_{N_{y_{on}}}}(t_{Nm_{on}}, U; \theta))^T) \in \\ &\mathbb{R}^{N_{y_{of}} \cdot Nm_{of} + N_{y_{on}} \cdot Nm_{on}} \end{aligned} \quad (9)$$

The weighting matrix,  $C_y \in \mathbb{R}^{N_{y_{of}} \cdot Nm_{of} + N_{y_{on}} \cdot Nm_{on} \times N_{y_{of}} \cdot Nm_{of} + N_{y_{on}} \cdot Nm_{on}}$  was obtained from the measurement errors, which were assumed to be unbiased, independent and normally distributed. Therefore  $C_y$  is diagonal matrix with entries given by the variance  $\sigma_{y,i}^2$  of each measurement  $i$ . Consequently, the observed measured responses in  $Y^m$  are normally-

distributed random variables, i.e.,  $Y^m \sim \mathcal{N}(\mathbb{E}(Y^m), \text{Var}(Y^m))$ , hence the expectation  $\mathbb{E}(Y^m)$  is equal to the model output at the unknown true parameters values  $\theta^*$ , i.e.,  $\mathbb{E}(Y^m) = Y(U, \theta^*)$ . The previous assumption ensures that the solution  $\hat{\theta}$  of the optimization (Equation 8) is equivalent to the maximum likelihood estimation solution [23].

#### 2.4. Parameter uncertainty quantification

To quantify the uncertainty associated with the PE, the precision of parameter values was assessed with the variance, relative standard deviation and the confidence interval calculated using a Monte Carlo (MC) method. A total of  $L$  –replications of the experimental data  $Y_j^m, j = 1, \dots, L$  were generated, drawing  $L$  – random MC samples from the normal distribution  $\mathcal{N}(Y^m, C^m)$ . For each data set  $Y_j^m$ , the PE was repeated to obtain the point estimates  $\hat{\theta}_1, \dots, \hat{\theta}_L$ . Then the parameter covariance matrix was computed as [23,24]

$$\text{Cov}(\hat{\theta}) = \frac{1}{(L-1)} \sum_{j=1}^L (\hat{\theta}_j - \mathbb{E}(\hat{\theta})) (\hat{\theta}_j - \mathbb{E}(\hat{\theta}))^T \quad (10)$$

where  $\mathbb{E}(\hat{\theta}) \approx \frac{1}{L} \sum_{j=1}^L \hat{\theta}_j$  represents the mean of the parameter distribution. The variances of the parameters,  $\sigma_{\theta_i}^2$  were equal to the diagonal entries of the covariance matrix  $\text{Cov}(\hat{\theta})$  from which the standard deviations  $\sigma_{\theta_i}$  were also calculated. The 95% confidence intervals were estimated using  $\sigma_{\theta_i}$  and a Student  $t$ -distribution.

### 3.0 Macro-kinetic model formulation

A significant improvement to the referenced models is the conversion of the inherently discontinuous systems into a continuous one that is more mathematically stable. The model comprises a set of ordinary differential equations (ODEs) describing six state variables, namely biomass  $X$  and extracellular concentrations of substrate  $S$ , acetate  $A$ , dissolved oxygen DOTa and DOT, and the feed  $F$  (which can be constant or a function of time). A set of auxiliary algebraic equations describing intracellular interactions relating substrate (glucose), oxygen and acetate consumption as well as biomass formation are coupled to the ODEs to form the kinetic model.

The state variables  $X, S, A$  and  $F$  are modelled as in a conventional fed-batch fermentation process. The dissolved oxygen, on the other hand, is modelled with two ordinary differential

equations (Equations 22 and 24). The difference between the two dissolved oxygen profiles  $DOT_a$  and  $DOT$  is due to the response lag of the sensor, which is approximated by first order response kinetics for the DO-probe [25].

The general form of the governing mass balance is expressed as follows,

$$\frac{dx}{dt} = \frac{F}{V}(x_i - x) + rX \quad (11)$$

where  $x \in \{X, S, A\}$  represents the state variable in [g/l] and the subscript  $i$  represents the inlet concentration,  $F$  the feed,  $V$  the volume and  $r$  is the corresponding specific rate. Considering that inlet concentration of biomass is zero (sterile feed), we obtain the following expression for biomass balance in the fed-batch case:

$$\frac{dX}{dt} = \frac{F}{V}(0 - X) + \mu X \quad (12)$$

In equation 12,  $X$  represents the concentration (cell dry weight) of cells and  $\mu$  ( $\text{h}^{-1}$ ) is the non-inhibited Monod-type specific growth rate, given as

$$\mu = (q_{sox} - q_m)Y_{em} + q_{sof}Y_{xsof} + q_{sA}Y_{xa} \quad (13)$$

where  $q_{sox}$ ,  $q_{sof}$ ,  $q_{sA}$ , represent the uptake rates of substrate for oxidation, substrate metabolized through the overflow route and acetate respectively, the constants  $Y_{..}$  define the respective yield coefficients, whereas  $q_m$  represents the glucose expended for cell maintenance. Thus, according to Equation 13, the overall growth of the culture results from usage of glucose through the oxidative ( $q_{sox}$ ), acetate uptake ( $q_{sA}$ ) and re-use of other products from the overflow route ( $q_{sof}$ ), which all contribute energy equivalents for cell growth [14].

The mass balance for substrate (glucose) in the fed-batch process is given as

$$\frac{dS}{dt} = \frac{F}{V}(S_i - S) - q_s X \quad (14)$$

The substrate concentration  $S$  is modelled taking acetate inhibition into account, as reviewed by Shiloach and Fass [26]. The specific substrate uptake rate is therefore modelled with Mono-type kinetics with non-competitive inhibition:

$$q_s = \frac{q_{smax} \cdot S}{1 + \frac{A}{K_{ia}} + S + K_s} \quad (15)$$

where  $K_{ia}$  and  $K_s$  are the acetate inhibition and the substrate affinity constants respectively. Not all the substrate consumed is metabolized in the TCA cycle ( $q_{sox}$ ), but a portion goes to the overflow path  $q_{sof}$ .

$$q_{sox} = (q_s - q_{sof}) \cdot \frac{DOT}{DOT + K_o} \quad (16)$$

$$q_{sof} = \frac{P_{Amax} q_s}{q_s + K_{ap}} \quad (17)$$

where  $K_o$  is a dimensionless constant set to 0.1 to increase the stability of the numeric simulation,  $P_{Amax}$  and  $K_{ap}$  are the maximum acetate production and the production affinity constants. Acetate production/consumption is a cyclic process and considering no addition of acetate in the feed, the mass balance yields

$$\frac{dA}{dt} = \frac{F}{V} (0 - A) + q_{sA} X \quad (18)$$

The equilibrium  $q_{sA} = 0$  is reached when the acetate produced through the overflow route  $p_A$  is equal to the acetate consumed  $q_{sA}$ .

$$q_A = p_A - q_{sA} \quad (19)$$

with

$$p_A = q_{sof} Y_{as} \quad (20)$$

where  $Y_{as}$  is the gram of acetate per gram of substrate consumed through the overflow route.

The specific acetate consumption rate is modelled as

$$q_{sA} = \frac{q_{Amax}}{1 + \frac{q_s}{K_{is}}} \cdot \frac{A}{A + K_{sa}} \quad (21)$$

where  $q_{Amax}$ ,  $K_{is}$ , and  $K_{sa}$  are constant parameters representing the maximum acetate uptake rate, the acetate uptake inhibition, and acetate affinity constant. The uptake of acetate, however, is inhibited in a non-competitive way by the presence of glucose in the medium due to *E. coli*'s higher preference for glucose over acetate [27]. Thus, glucose and acetate exhibit a counter inhibition effect on their respective uptake capacities, as presented in Equations 15 and 21.

Finally, the actual dissolved oxygen ( $DOT_a$ ) is calculated in % of saturation with the assumption that the feed solution in the fed-batch phase is fully saturated with dissolved oxygen. The oxygen profile is described with the standard equation

$$\frac{dDOT_a}{dt} = K_{La} (DOT^* - DOT_a) - q_o X H \quad (22)$$

with  $DOT^*$  being the saturation value of dissolved oxygen in the medium,  $K_{La}$  the volumetric mass transfer coefficient,  $H$  the Henry equilibrium constant, and  $q_o$  the oxygen uptake rate described by

$$q_o = (q_{sox} - q_m)Y_{os} + q_{sA}Y_{oa} \quad (23)$$

where  $Y_{os}$ ,  $Y_{oa}$  are the yield coefficients for substrate and acetate to oxygen consumption respectively. With the probe response, the measured DOT is modelled as:

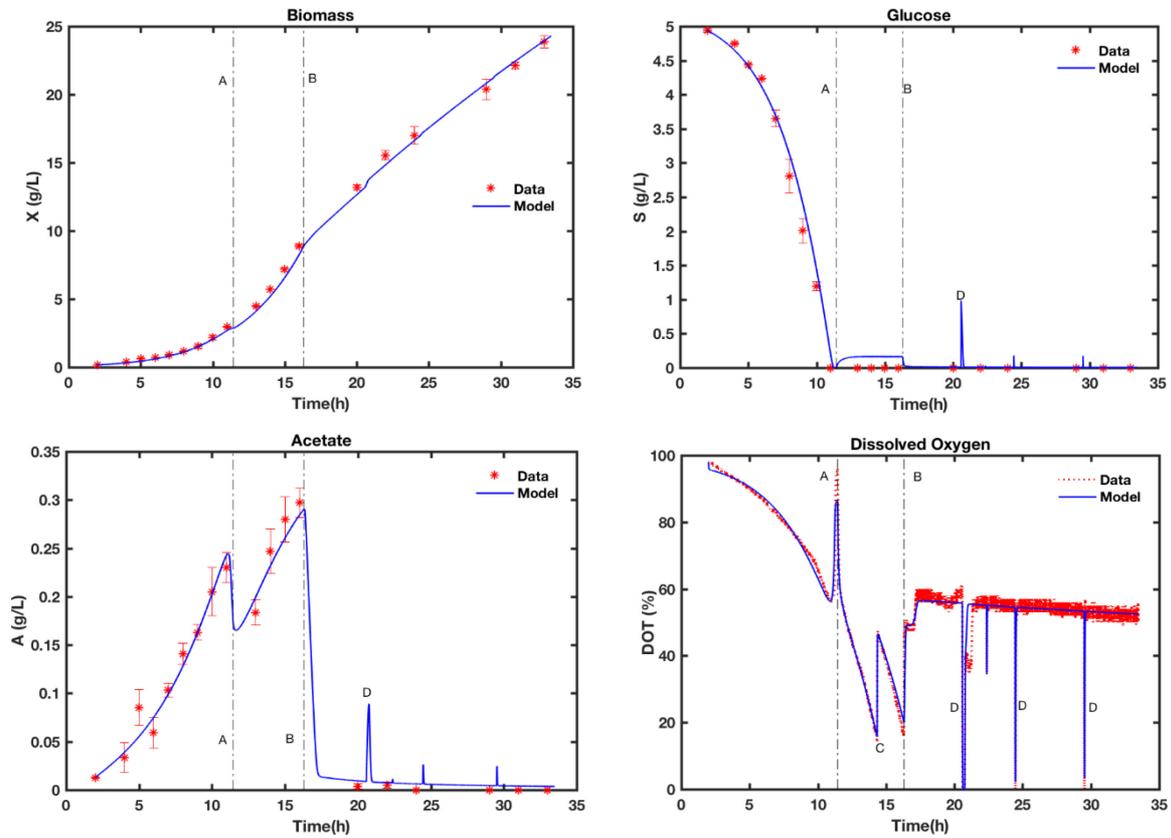
$$\frac{dDOT}{dt} = K_p(DOT_a - DOT) \quad (24)$$

where  $K_p = \frac{1}{\tau}$  is the static gain of the sensor given as the inverse of the probe response time,  $\tau$ . This is important when dealing with pulses in the system and situations where fast changes in the system are expected. In an experimental set-up,  $\tau$  is measured as the time required to reach 63.2% of the final response after a step change in the DOT profile [28]. At the calculated  $q_o$  values for the current *E. coli* strain, when the probe response time is greater than 8 seconds, it is important to include it in the dynamics of the oxygen profile.

#### 4.0 Results and Discussion

Acetate accumulation remains one of the biggest challenges in recombinant protein production using the *E. coli* expression system. In the past, the inclusion of acetate profiles in mechanistic models of *E. coli* meant that the system became inherently discontinuous because acetate was thought to be produced only during certain specific growth regimes in *E. coli* cultivation. This property of the system inhibited further mathematical development, especially in the fields of control theory and model-based optimization. In the current paper, we have explored the possibility of developing continuous differential equations for acetate production based on the concept of acetate cycling. The data to validate the new set of continuous equations describing acetate production in *E. coli* was obtained from a fed-batch cultivation of non-recombinant *E. coli* W3110M. The cultivation was done to cover all possible growth regimes in *E. coli* [29], to study the kinetics of acetate production and its consumption. From an excess substrate (glucose) environment in the batch phase to chronic starvation conditions at the end, the acetate profile evolved according to predictable kinetics (Figure 2). A maximum specific growth rate ( $\mu_{max}$ ) of  $0.31 \text{ h}^{-1}$  was recorded during the batch phase, in the presence of excess glucose. Growth at  $\mu_{max}$  was associated with a steady increase in extracellular acetate concentration. In the exponential feed fed-batch phase, acetate concentration increased further to a maximum of

0.3 g/L due to the higher  $\mu_{\text{set}}$  value ( $0.22 \text{ h}^{-1}$ ), which corresponds to a specific substrate uptake rate higher than that which would allow fully oxidative growth.



**Figure 2** Comparison of *E. coli* model with experimental data from fed-batch cultivation of wild type *E. coli* W3110. A = start of exponential feed fed-batch phase at  $\mu_{\text{set}} = 0.25 \text{ h}^{-1}$  B = start of constant feed fed-batch phase, C = change in  $K_{L}a$ , D = glucose pulses.

This  $\mu_{\text{set}}$  was about 75% of  $\mu_{\text{max}}$  and was higher than the threshold for the acetate switch at the given  $\mu_{\text{max}}$  [1,30]. During the constant feed phase the specific growth rate decreased to values below  $0.05 \text{ h}^{-1}$ , when the glucose supply fell to values below  $0.1 \text{ g}_{\text{glu}} \text{ g}_{\text{x}}^{-1} \text{ h}^{-1}$  (Figure 3). In the constant, slow feeding phase, acetate consumption was greater than its production, which represented a shift in the cycling equilibrium that favoured the consumption of extracellular acetate (point B, Figure 2). The profiles of all measurable state variables (X, S, A, DOT) in the model are plotted in Figure 2, with the corresponding model predictions after parameter estimation. There was an adequate fit between the model and the experimental data. The estimated model parameters based on the experimental data are given in Table1 together with

the uncertainty in the estimated parameter values calculated from the Monte Carlo method. From the measure of the relative standard deviation, it is apparent that the PE gave unique parameter values that describe the *E. coli* system under the given experimental conditions. Twelve (12) out of the sixteen (16) parameters had a  $\sigma_\theta$  values less than 10%. Although some of the parameters directly related to acetate cycling had somewhat higher  $\sigma_\theta$  values, the acetate profile in the cultivation was sufficiently described by the model.

Table 1. Summary of results of parameter estimation and quantification of the uncertainty associated with the parameter estimators. LB—lower bound, UB—upper bounds, CI—95% confidence interval.

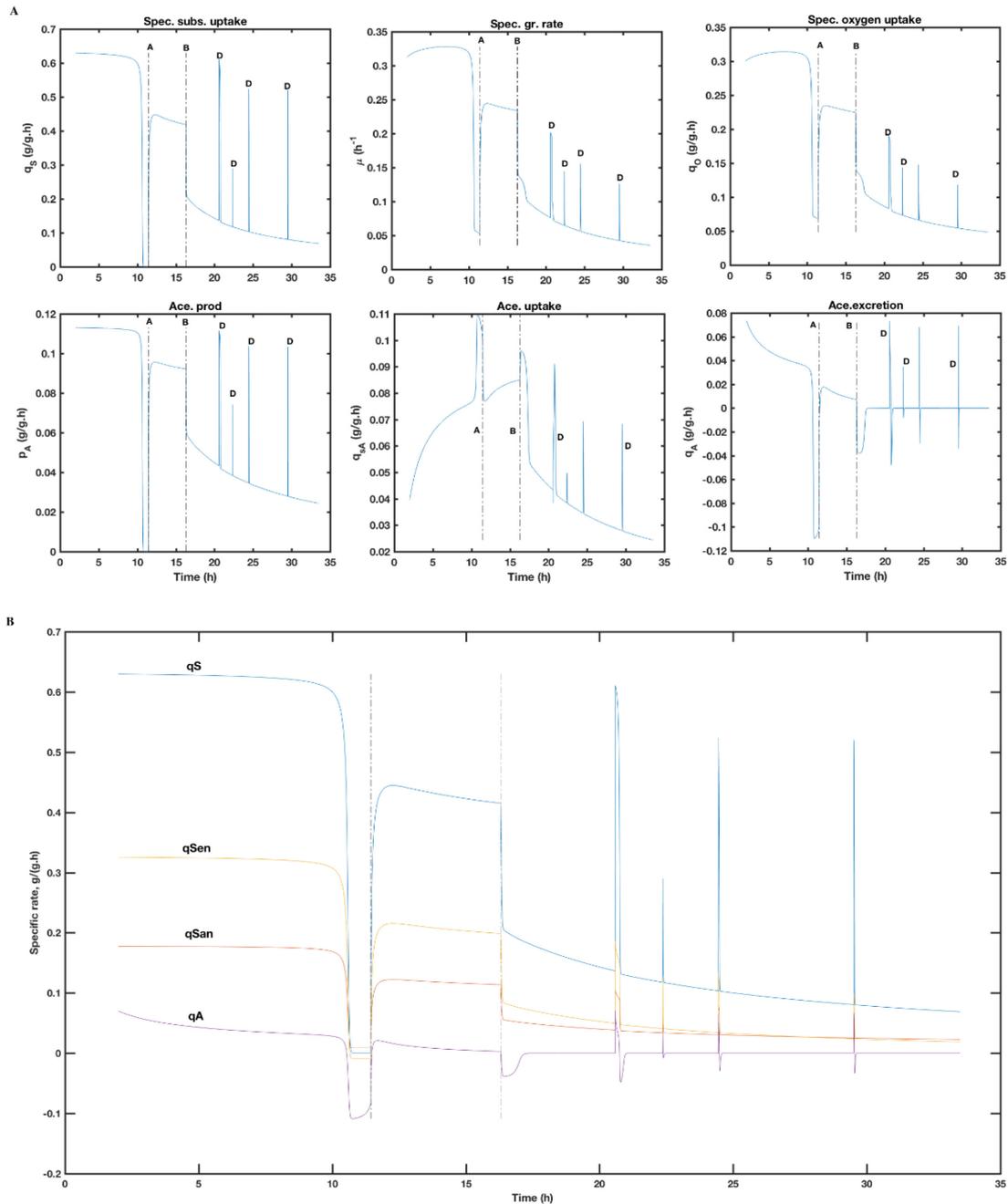
Parameter <sup>+</sup>	Units	Initial guess (literature)*	Estimate	PE uncertainty quantification		
				% $\sigma_\theta$	LB-CI	UB-CI
$K_{ap}$	$g L^{-1}$	0.10	0.5052	15.2	0.3539	0.6565
$K_{sa}$	$g L^{-1}$	0.05	0.0134	22.0	0.0076	0.0192
$K_o$	$g L^{-1}$	10.0	0.0001	0.0	0.0001	0.0001
$K_s$	$g L^{-1}$	0.05	0.0370	8.9	0.0305	0.0435
$K_{ia}$	$g L^{-1}$	5.00	1.2399	9.6	1.0062	1.4737
$K_{is}$	$g L^{-1}$	10.0	2.1231	27.3	0.9788	3.2673
$p_{Amax}$	$g g^{-1} h^{-1}$	0.17	0.2268	6.5	0.1977	0.2558
$q_{Amax}$	$g g^{-1} h^{-1}$	0.15	0.1148	6.1	0.1009	0.1287
$q_m$	$g g^{-1} h^{-1}$	0.04	0.0129	7.0	0.0111	0.0147
$q_{Smax}$	$g g^{-1} h^{-1}$	1.37	0.6356	0.3	0.6320	0.6392
$Y_{as}$	$g g^{-1}$	0.80	0.9097	4.5	0.8283	0.9911
$Y_{oa}$	$g g^{-1}$	1.06	0.5440	9.5	0.4425	0.6455
$Y_{xa}$	$g g^{-1}$	0.70	0.5718	9.9	0.4604	0.6833
$Y_{em}$	$g g^{-1}$	0.50	0.5333	2.4	0.5085	0.5580
$Y_{os}$	$g g^{-1}$	1.06	1.5620	5.4	1.3941	1.7298
$Y_{xsof}$	$g g^{-1}$	0.15	0.2268	12.0	0.1730	0.2807

<sup>+</sup> Parameter descriptions in nomenclature. \* References: [4,5,8,16]

#### 4.1 Dynamic acetate production and consumption rates

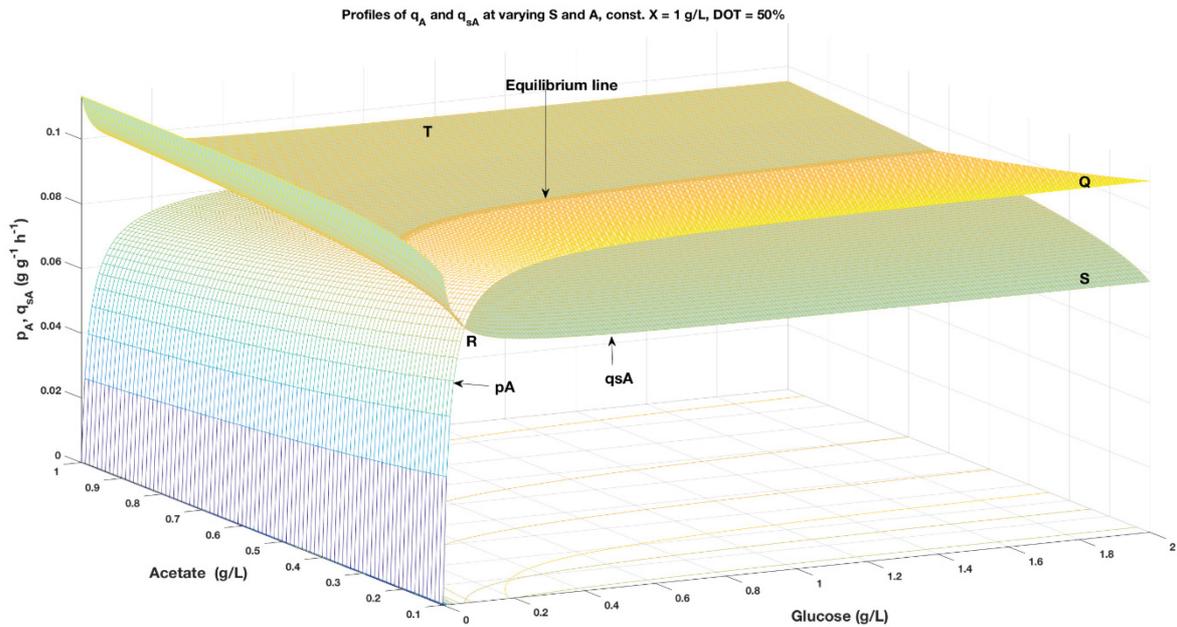
The estimation of model parameters was used to identify parameter values (Table 1) to fully describe the *E. coli* system. Therefore, the partitioning of the carbon source (Figure 3B), the rate of acetate cycling and other *specific* rates along the different phases of the cultivation could

be dynamically simulated from the estimated parameter values (Figure 3A). The simulated profiles and the model fit (Figure 2) reveal a rapid response of *E. coli* to acetate production/re-consumption after each glucose pulse.



**Figure 3** Simulated specific rates of the (A) acetate cycling pathways and (B) overall glucose partitioning during the cultivation of *E. coli*, with dynamic response to intermittent glucose pulses. Negative acetate excretion rates imply that extracellular acetate is taken up by the cell. A = end of batch phase, B = end of exponential feed, start of constant feed, C = change in KLa, D = glucose pulses.

The absence of a diauxic delay during the consumption of acetate after each pulse and after the shift to a lower  $\mu$  ( $0.11 \text{ h}^{-1}$ ) suggests that acetate is continuously being re-assimilated in the background alongside glucose [31–33], which further confirms the acetate cycling concept in *E. coli*. Due to the lack of an active transport mechanism for acetate across the cell membrane in *E. coli* [34,35] (unlike glucose which uses the PTS system),  $q_A$  also incorporates the net acetate influx by facilitated transport across the cell membrane by various permease proteins, such as those encoded by the genes *yaaH* [36] and *yjcG* [37]. Re-assimilation of acetate results in lower intracellular acetate concentrations, which alters the membrane gradients and forces extracellular acetate into the cell by facilitated transmembrane balance. In effect, when intracellular acetate production through POXB (Figure 1) is non-functional (e.g. when  $q_s = 0$ ),  $p_A$  is derived from extracellular acetate by this principle of transmembrane balance. The inhibition of extracellular acetate uptake by glucose is evidenced by the sharply rising profile of  $q_{SA}$  in response to the depletion of glucose during the batch phase (Figure 3). The delicate equilibrium in the acetate cycling pathways is shown in the 3D plot in Figure 4, which was generated by solving the set of algebraic equations at constant biomass concentration ( $X = 1.3 \text{ g/L}$ ) and DOT (65%) and variable glucose and acetate concentrations. This corresponds to about 9 hours in the batch phase. The point labelled T in Figure 4 represents pure oxidative growth. As the residual glucose concentration increases, the metabolism gradually shifts towards the equilibrium line, until  $p_A > q_{SA}$  where acetate excretion begins. At this point, the total overflow flux that leads to extracellular acetate accumulation is proportional to the area enclosed by the envelope QRS and the equilibrium line. Thus, at any given glucose concentration, the overflow flux, and consequently, the excreted acetate concentration can be estimated from the algebraic relations. The Monod-type dependence of the specific acetate uptake rate on the residual acetate concentration as well as the minimal acetate production at low glucose concentrations are also shown in Figure 4. The formulation is an advanced version of a similar representation of overflow metabolism in *Saccharomyces cerevisiae*, as presented by Pham and co-workers [14]. This formulation of acetate profiles and overflow metabolism is closer to reality since *E. coli* does not have clear cut switching points in its metabolic routes on the onset of overflow, but rather slow or gradual switching systems from oxidative to overflow metabolism [38,39].



**Figure 4** Dependence of overflow metabolism on residual glucose and acetate concentrations, showing the dynamic equilibrium between intracellular acetate production and re-assimilation at varying glucose concentrations in the medium.

## 5.0 Conclusions

Model-based process development and optimization is becoming state-of-the-art in the biotechnology industry. Therefore, the mathematical functions used to describe fermentation systems should have the appropriate properties to enable their application in this fast-growing field. We have used the concept of acetate cycling to derive a set of continuously differentiable and tractable equations to describe acetate accumulation in cultivations of *E. coli*. By fitting this model to experimental data, we show that the new set of equations sufficiently describe growth profiles in *E. coli*, as well as the acetate production and re-assimilation rates. Due to its continuous characteristics, this model is suitable for simulations that require higher order gradient calculation such as Computational Fluid Dynamics (CFD) models, as well as in complex optimization problems involving the *E. coli* expression system.

## Nomenclature

C	Carbon content of (s) substrate, (x) biomass
DOT	dissolved oxygen tension. DOT* represents saturating value of DOT in the broth at the given operating conditions.

F	Flow rate (L/h)
Kap	Monod-type saturation constant, intracellular acetate production
Ksa	affinity constant, acetate consumption (g/L)
Ko	affinity constant, oxygen consumption (g./L)
Ks	affinity constant, substrate consumption (g./L)
Kia	inhibition constant, inhibition of glucose uptake by acetate
Kis	inhibition constant, inhibition of acetate uptake by glucose(g/L)
pA	specific acetate production rate (g/(g.h))
qA	specific acetate consumption rate (g/(g.h))
qm	specific maintenance coefficient (g/(g.h))
qS	max spec glucose uptake rate (g/(g.h))
Y	yield coefficient (g/g)
Yxsof	yield of biomass on substrate from auxiliary overflow routes, such as the mixed acid and pentose-phosphate pathways.
$\mu$	specific growth rate ( $h^{-1}$ )
$\tau$	dissolved oxygen probe response time (h)

### Subscripts

A	acetate
an	anabolic
c	consumption
en	energetic
glu	glucose
i	inlet concentration
m	maintenance
max	maximum
O	oxygen
of	overflow
ox	oxidative
S	substrate (glucose)
X	biomass
em	excluding maintenance

### Acknowledgement

The authors are grateful to Prof Emeritus Sven-Olof Enfors (rtd) for his valuable input on the development of the model. EA is grateful for financial support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie actions

grant agreement No. 643056 (Biorapid). DCLC gratefully acknowledges financial support from the Faculty for the Future Fellowship program (Schlumberger Foundation) for her post-doctoral scholarship.

## References

- [1] A.J. Wolfe, The Acetate Switch, *Microbiol. Mol. Biol. Rev.* 69 (2005) 12–50. doi:10.1128/MMBR.69.1.12-50.2005.
- [2] M. De Mey, S. De Maeseneire, W. Soetaert, E. Vandamme, Minimizing acetate formation in *E. coli* fermentations, *J. Ind. Microbiol. Biotechnol.* 34 (2007) 689–700. doi:10.1007/s10295-007-0244-2.
- [3] G. Insel, G. Celikyilmaz, E. Ucisik-Akkaya, K. Yesiladali, Z.P. Cakar, C. Tamerler, D. Orhon, Respiriometric evaluation and modeling of glucose utilization by *Escherichia coli* under aerobic and mesophilic cultivation conditions, *Biotechnol. Bioeng.* 96 (2007) 94–105. doi:10.1002/bit.21163.
- [4] B. Xu, M. Jahic, S.-O. Enfors, Modeling of Overflow Metabolism in Batch and Fed-Batch Cultures of *Escherichia coli*, *Biotechnol. Prog.* 15 (1999) 81–90. doi:10.1021/bp9801087.
- [5] H.Y. Lin, B. Mathiszik, B. Xu, S.O. Enfors, P. Neubauer, Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of *Escherichia coli*, *Biotechnol. Bioeng.* 73 (2001) 347–357. doi:10.1002/bit.1068.
- [6] Y.F. Ko, W.E. Bentley, W. a Weigand, A metabolic model of cellular energetics and carbon flux during aerobic *Escherichia coli* fermentation., *Biotechnol. Bioeng.* 43 (1994) 847–855. doi:10.1002/bit.260430903.
- [7] A. Cockshott, I. Bogle, Modelling the effects of glucose feeding on a recombinant *E. coli* fermentation, *Bioprocess Eng.* 20 (1999) 83–90. doi:10.1007/PL00009037.
- [8] P. Neubauer, H.Y. Lin, B. Mathiszik, Metabolic load of recombinant protein production: Inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in *Escherichia coli*, *Biotechnol. Bioeng.* 83 (2003) 53–64. doi:10.1002/bit.10645.
- [9] K. Peebo, K. Valgepea, A. Maser, R. Nahku, K. Adamberg, R. Vilu, Proteome reallocation in *Escherichia coli* with increasing specific growth rate, *Mol. Biosyst.* 11 (2015) 1184–1193. doi:10.1039/C4MB00721B.

- [10] K. Valgepea, K. Adamberg, R. Nahku, P.-J. Lahtvee, L. Arike, R. Vilu, Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase, *BMC Syst. Biol.* 4 (2010) 166. doi:10.1186/1752-0509-4-166.
- [11] M. Basan, S. Hui, H. Okano, Z. Zhang, Y. Shen, J.R. Williamson, T. Hwa, Overflow metabolism in *Escherichia coli* results from efficient proteome allocation, *Nature.* 528 (2015) 99–104. doi:10.1038/nature15765.
- [12] M. Matsubara, T. Mizuno, EnvZ-independent Phosphotransfer Signaling Pathway of the OmpR-mediated Osmoregulatory Expression of OmpC and OmpF in *Escherichia coli*, *Biosci., Biotechnol. Biochem.* 63 (1999). doi:dx.doi.org/10.1271/bbb.63.408.
- [13] S. Renilla, V. Bernal, T. Fuhrer, S. Castaño-Cerezo, J.M. Pastor, J.L. Iborra, U. Sauer, M. Cánovas, Acetate scavenging activity in *Escherichia coli*: Interplay of acetyl-CoA synthetase and the PEP-glyoxylate cycle in chemostat cultures, *Appl. Microbiol. Biotechnol.* 93 (2012) 2109–2124. doi:10.1007/s00253-011-3536-4.
- [14] H.T.B. Pham, G. Larsson, S.O. Enfors, Growth and energy metabolism in aerobic fed-batch cultures of *Saccharomyces cerevisiae*: Simulation and model verification, *Biotechnol. Bioeng.* 60 (1998) 474–482. doi:10.1002/(SICI)1097-0290(19981120)60:4<474::AID-BIT9>3.0.CO;2-J.
- [15] P. Millard, K. Smallbone, P. Mendes, Metabolic regulation is sufficient for global and robust coordination of glucose uptake, catabolism, energy production and growth in *Escherichia coli*, *PLOS Comput. Biol.* 13 (2017) e1005396. doi:10.1371/journal.pcbi.1005396.
- [16] K. Martínez-Gómez, N. Flores, H.M. Castañeda, G. Martínez-Batallar, G. Hernández-Chávez, O.T. Ramírez, G. Gosset, S. Encarnación, F. Bolivar, New insights into *Escherichia coli* metabolism: carbon scavenging, acetate metabolism and carbon recycling responses during growth on glycerol., *Microb. Cell Fact.* 11 (2012) 46. doi:10.1186/1475-2859-11-46.
- [17] M.N. Cruz Bournazou, T. Barz, D. Nickel, D. López Cárdenas, F. Glauche, A. Knepper, P. Neubauer, Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities for validation of macro-kinetic growth models using *E. coli* as an example, *Biotechnol. Bioeng.* 9999 (2016) 1–10. doi:10.1002/bit.26192.
- [18] D.B. Nickel, N.M. Cruz-Bournazou, T. Wilms, P. Neubauer, A. Knepper, Online bioprocess data generation, analysis and optimization for parallel fed-batch fermentations at mL scale, *Eng. Life Sci.* (2016) 1–7. doi:10.1002/elsc.201600035.

- [19] L.T. Biegler, Recent Advances in Chemical Process Optimization, *Chemie Ing. Tech.* 86 (2014) 943–952. doi:10.1002/cite.201400033.
- [20] J. Glazyrina, E.-M. Materne, T. Dreher, D. Storm, S. Junne, T. Adams, G. Greller, P. Neubauer, High cell density cultivation and recombinant protein production with *Escherichia coli* in a rocking-motion-type bioreactor., *Microb. Cell Fact.* 9 (2010) 42. doi:10.1186/1475-2859-9-42.
- [21] A.C. Hindmarsh, P.N. Brown, K.E. Grant, S.L. Lee, R. Serban, D.E. Shumaker, C.S. Woodward, SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers, *ACM Trans. Math. Softw.* 31 (2005) 363–396. doi:10.1145/1089014.1089020.
- [22] D.C. López C., T. Barz, M. Penula, V. Adriana, S. Ochoa, G. Wozny, Model-Based Identifiable Parameter Determination Applied to a Simultaneous Saccharification and Fermentation Process Model for Bio-Ethanol Production, *Biotechnol. Prog.* 29 (2013) 1064–1082. doi:10.1002/btpr.1753.
- [23] Y. Bard, Non-linear parameter estimation, Academic Press, New York, 1974.
- [24] D.C. López, G. Wozny, A. Flores-Tlacuahuac, R. Vasquez-Medrano, V.M. Zavala, A Computational Framework for Identifiability and Ill-Conditioning Analysis of Lithium-Ion Battery Models, *Ind. Eng. Chem. Res.* 55 (2016) 3026–3042. doi:10.1021/acs.iecr.5b03910.
- [25] L.A. Tribe, C.L. Briens, A. Margaritis, Determination of the volumetric mass transfer coefficient ( $k(L)a$ ) using the dynamic “gas out-gas in” method: Analysis of errors caused by dissolved oxygen probes., *Biotechnol. Bioeng.* 46 (1995) 388–392. doi:10.1002/bit.260460412.
- [26] J. Shiloach, R. Fass, Growing *E. coli* to high cell density - A historical perspective on method development, *Biotechnol. Adv.* 23 (2005) 345–357. doi:10.1016/j.biotechadv.2005.04.004.
- [27] H. Ying Lin, P. Neubauer, Influence of controlled glucose oscillations on a fed-batch process of recombinant *Escherichia coli*, *J. Biotechnol.* 79 (2000) 27–37. doi:10.1016/S0168-1656(00)00217-0.
- [28] A.C. Badino, M. Cândida, R. Facciotti, W. Schmidell, Improving  $k(L)a$  determination in fungal fermentation, taking into account electrode response time, *J. Chem. Technol. Biotechnol.* 75 (2000) 469–474. doi:10.1002/1097-4660(200006)75:6<469::AID-JCTB236>3.0.CO;2-4.
- [29] W. Chesbro, M. Arbige, R. Eifert, When nutrient limitation places bacteria in the domains of slow growth: Metabolic, Morphologic and Cell Cycle Behavior, *FEMS*

- Microbiol. Ecol. 74 (1990) 103–119.
- [30] E. Anane, E. van Rensburg, J.F. Görgens, Optimisation and scale-up of  $\alpha$ -glucuronidase production by recombinant *Saccharomyces cerevisiae* in aerobic fed-batch culture with constant growth rate, *Biochem. Eng. J.* 81 (2013) 1–7. doi:10.1016/j.bej.2013.09.012.
- [31] V. Bernal, S. Castaño-Cerezo, M. Cánovas, Acetate metabolism regulation in *Escherichia coli*: carbon overflow, pathogenicity, and beyond, *Appl. Microbiol. Biotechnol.* 100 (2016) 8985–9001. doi:10.1007/s00253-016-7832-x.
- [32] B. Enjalbert, M. Cocaign-Bousquet, J.C. Portais, F. Letisse, Acetate exposure determines the diauxic behavior of *Escherichia coli* during the glucose-acetate transition, *J. Bacteriol.* 197 (2015) 3173–3181. doi:10.1128/JB.00128-15.
- [33] S. Leone, F. Sannino, M.L. Tutino, E. Parrilli, D. Picone, Acetate: friend or foe? Efficient production of a sweet protein in *Escherichia coli* BL21 using acetate as a carbon source., *Microb Cell Fact.* 14 (2015) 106. doi:10.1186/s12934-015-0299-0.
- [34] D.D. Axe, J.E. Bailey, Transport of lactate and acetate through the energized cytoplasmic membrane of *Escherichia coli*, *Biotechnol. Bioeng.* 47 (1995) 8–19. doi:10.1002/bit.260470103.
- [35] R.R. Wright, J.E. Hobbie, Use of Glucose and Acetate by Bacteria and Algae in Aquatic Ecosystems, *Ecology.* 47 (1966) 447–464. doi:10.2307/1932984.
- [36] J. Sá-Pessoa, S. Paiva, D. Ribas, I.J. Silva, S.C. Viegas, C.M. Arraiano, M. Casal, SATP (YaaH), a succinate–acetate transporter protein in *Escherichia coli*, *Biochem. J.* 454 (2013) 585–595. doi:10.1042/BJ20130412.
- [37] R. Gimenez, M.F. Nunez, J. Badia, J. Aguilar, L. Baldoma, The Gene *yjcG*, Cotranscribed with the Gene *acs*, Encodes an Acetate Permease in *Escherichia coli*, *J. Bacteriol.* 185 (2003) 6448–6455. doi:10.1128/JB.185.21.6448-6455.2003.
- [38] A. Kayser, J. Weber, V. Hecht, U. Rinas, Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. I. Growth-rate-dependent metabolic efficiency at steady state, *Microbiology.* 151 (2005) 693–706. doi:10.1099/mic.0.27481-0.
- [39] T. Paalme, R. Elken, A. Kahru, K. Vanatalu, R. Vilu, The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach, *Antonie Van Leeuwenhoek.* 71 (1997) 217–230. doi:10.1023/A:1000198404007.
- [40] Z. Li, M. Nimtz, U. Rinas, The metabolic potential of *Escherichia coli* BL21 in defined and rich medium., *Microb. Cell Fact.* 13 (2014) 45. doi:10.1186/1475-2859-13-45.