A CMI (Cell Metabolic Indicator)-based Controller for Achieving High Growth Rate *Escherichia coli* Cultures*

Matthew E. Pepper, Li Wang, Ajay Padmakumar, Timothy C. Burg, Sarah W. Harcum, and Richard E. Groff

***Abstract*— A large fraction of biopharmaceuticals are produced in *Escherichia coli*, where each new product and strain currently requires a high degree of growth characterization in benchtop and industrial bioreactors to achieve economical production protocols. The capability to use a standard set of sensors to characterize a system quickly without the need to conduct numerous experiments to determine stable growth rate for the strain would significantly decrease development time. This paper presents a cell metabolic indicator (CMI) which provides better insight into the *E. coli* metabolism than a growth rate value. The CMI is the ratio of the oxygen uptake rate (OUR) of the culture and the base addition rate (BAR) required to keep pH at a desired setpoint. The OUR and BAR are measured using a off-gas sensor and pH probe, respectively, and thus the CMI can be computed online. Experimental results demonstrate the relationship between CMI and the different cell metabolic states. A previously published model is augmented with acid production dynamics, allowing for comparison of the CMI-based controller with an open-loop controller in simulation. The CMI-based controller required little *a priori* knowledge about the *E. coli* strain in order to achieve a high growth rate. Since many different types of cells exhibit similar behaviors, the CMI concept can be extended to mammalian and stem cells.**

I. INTRODUCTION

Many important biopharmaceutical products, including insulin and human growth hormone, are grown in computer-controlled bioreactors using recombinant strains of the bacteria *Escherichia coli*[1]. Most often, automated industrial and benchtop bioreactors use a very simple control methodology: closed loop PID control on environmental variables such as pH, temperature, and dissolved oxygen, and a preset exponential feeding schedule or other growth profile based on prior experience with the particular strain. An operator monitors sensors throughout the process to detect and minimize the effects of stresses to the culture. One common source of stress is overfeeding of glucose, which creates the waste product acetate, which inhibits growth. Large numbers of cultures must be run to determine a feeding protocol that maximizes the growth rate, final cell mass, and recombinant protein production [2], [3].

Bioreactors usually employ a standard sensor set, monitoring pH, temperature, dissolved oxygen (DO), and oxygen off-gas (oxygen uptake rate, OUR). Commonly available sensors do not measure biomass, glucose concentration, or acetate concentration. While online sensors for some of these variables have been developed, the sensors tend to be either too expensive or too unreliable [4]. The inability to sense these variables online is central to the problem of growth rate control. The focus over the last three decades [5] has been to design control algorithms which maintain a desired growth rate. Without online biomass measurements, adaptive observer-based estimators for growth rate and biomass were constructed based on ODE models using coefficients derived from previous experiments. Controllers based on these estimators were used to successfully maintain conservative growth rate setpoints [6], [7]. Another popular estimation method for biomass and growth rate is neural networks [8], [9]. Regardless of the type of estimator employed, accurate estimation requires numerous prior characterization experiments and/or many training sets.

The cell metabolic indicator (CMI) presented in this paper provides a more accurate picture of cell metabolism than growth rate alone. Theoretical and experimental results agree that the distinct metabolic phases can be identified by their CMI profile. Based on sensors commonly available on a benchtop bioreactor, the CMI is the ratio of the OUR and BAR signals, and can be tracked on-line in real time. A simulated culture was characterized using the CMI and a CMI-based controller was designed that maintained high growth rates and avoided harmful byproduct production. This CMI-based controller is compared with a traditional open loop controller and achieved a larger total biomass production.

II. E. COLI CULTURE & METABOLISM

A typical *E. coli* fermentation has two phases, batch and fed-batch. In the batch phase, the initial glucose is consumed and cells grow rapidly. Fed-batch begins with the start of an external glucose feed after the initial glucose is depleted. The feed rate often follows an exponential profile with growth rate setpoint μ_{set}, which represents the substrate-limited growth rate of the cells.

In order to understand how the CMI is a more accurate representation of *E. coli* metabolism than the growth rate value, a brief overview of *E. coli* metabolism is presented.

A. E. coli Metabolism

Fundamentally, *E. coli* metabolism uses oxygen and glucose to generate biomass and excrete carbon dioxide. The cells absorb glucose, which is processed via three main

*This research was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103444.

M.E.Pepper, L.Wang, A.Padmakumar, T.C.Burg, and R.E.Groff are with the Department of Electrical & Computer Engineering, Clemson University, Clemson, SC, 29634-0915 (regroff@clemson.edu)

S.W.Harcum is in the Department of Bioengineering, Clemson University, Clemson, SC, 29634-0915

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

reactions: glycolysis, pyruvate decarboxylation, and the Tricarboxylic Acid (TCA) cycle. These reactions consume oxygen and produce some acidic byproducts. When the glucose concentration is not too high, the culture is only in oxidative metabolism, with growth rate, μ_1. When the glucose concentration is too high, $\mu_1 \rightarrow \mu_{1max}$ and the TCA cycle is processing oxygen (OUR) and glucose at its maximum rate. The excess glucose is redirected to production of acetate (acetic acid), a byproduct. The production of acetate is known as overflow metabolism, with a growth rate, μ_2. If the glucose concentration lowers, acetate production stops $\mu_2 \rightarrow 0$ and the TCA cycle processes acetate alongside of glucose. The consumption of acetate is known as metabolic consumption metabolism, with growth rate, μ_3 [10]. The production and presence of acetate is detrimental because acetate inhibits overall oxygen capacity, lowering μ_{1max}.

B. The Cell Metabolic Indicator

The metabolic state of the culture is not apparent simply by measuring the growth rate, μ. The growth rate only reflects the combinations of the different metabolic phases, for instance, in oxidative metabolism, $\mu = \mu_1$; during overflow, $\mu = \mu_{1max} + \mu_2$; and in metabolite consumption, $\mu = \mu_1 + \mu_3$ until all the acetate is consumed. In contrast, the cell metabolic indicator (CMI), defined as

$$CMI = \frac{OUR}{BAR} \tag{1}$$

can indicate the metabolic phase of the culture. The OUR is closely related to the rate of the TCA cycle, directly indicating the level of oxidative metabolism and if applicable, acetate consumption. Since pH is controlled in the fermenter, BAR is closed related to the total growth rate from all three metabolic phases. Thus, CMI represents the ratio of the oxidative metabolism to the total metabolism. In oxidative metabolism, the ratio of OUR and BAR has some minimum, CMI_{min}, which represents $\mu_1 \rightarrow \mu_{1max}$. In overflow metabolism, CMI $<$ CMI_{min} as OUR saturates and BAR grows due to acetate production. If the glucose concentration lowers, metabolite consumption metabolism begins and the combined OUR from $\mu_1 + \mu_3$ drives CMI $>>$ CMI_{min}. When the acetate is totally consumed, the CMI decreases to CMI_{mid}, some value above CMI_{min}. The metabolic phases of a culture can be accurately identified given a profile of CMI and only the values of CMI_{mid} and CMI_{min} (see Fig. 1).

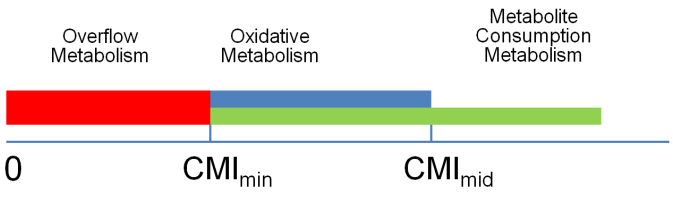

Fig. 1. Illustration of valid ranges for CMI and the oxidative (blue), overflow (red), and metabolite consumption (green) metabolism.

C. Demonstrating CMI

In order to test the validity of the assumptions behind the CMI, an *E. coli* MG1655 pTVP1GFP [11], [12] culture was grown and discrete glucose pulses were administered to push the culture into overflow metabolism. The full details of culture preparation can be found in [13].

The 5-L B.Braun Biostat B bioreactor (Sartorius-stedim North America Inc, Bohemia, NY) is an autoclavable glass vessel with steel head plate with attachments for a motor, sensors, and a digital control unit. The digital control unit (DCU) provides power, reads sensors, and provides control either onboard or from a separate PC. Temperature is controlled via a water jacket. The pH is controlled through the addition of base. The dissolved oxygen level is maintained via sparging in conjunction with varying the stir speed. These states are controlled using sensor data from temperature, pH, and dissolved oxygen probes. The base and feed solutions are added via peristaltic pumps, and the mass additions were monitored through balances (Ohaus Corp, Parsippany, NJ). The input gas was controlled by a mass flow controller (Omega Engineering Inc., Stamford, CT) and exhaust gas composition by an off-gas sensor (BlueSens gas sensor GmbH, Herten, Germany).

The fed-batch phase began around 5 hours. The Biostat B DCU onboard PID controllers maintained temperature and pH; the setpoints were 37 C and 6.95, respectively. A PC running a Simulink model with Matlab 2012a (Mathworks Inc, Natick, MA) provided feed rate commands using an open-loop feed rate profile with a growth rate setpoint of 0.25 h^{-1} and provided motor commands to keep the DO above 30%. The data sample rates for the DCU, off-gas sensor, and balances were 15, 10, and 5 seconds respectively.

During the fed-batch phase, two pulses with durations of 2 minutes and 5 minutes and with magnitudes five times and three times the current feed rates, respectively, were administered. The pulses were designed to help determine how much glucose is required to push the culture into overflow. By administering a discrete glucose pulse, the culture cycled through all metabolic phases, yielding the full range for CMI. The first pulse was clearly large enough to cycle the cell metabolism. The BAR increased causing the CMI to go down, indicating overflow metabolism. Minutes later, the glucose concentration declined, initiating metabolite consumption metabolism, driving up the CMI. After the acetate was consumed, the CMI returned to steady state (0.70), indicating only oxidative metabolism was active. The second spike was longer than the first, but with a smaller magnitude. The larger amount of biomass present at the time of the second pulse accounted for the lower magnitude of the CMI response; the cells were not sent as far into overflow. The CMI values during overflow for the first and second pulse, 0.365 and 0.408, appear to confirm this. The peak of the CMI response was lower after the second pulse (1.14 versus 1.65), which indicated a smaller amount of acetate had to be consumed before settling again to the steady state oxidative metabolism value of 0.70. The CMI_{mid} for this

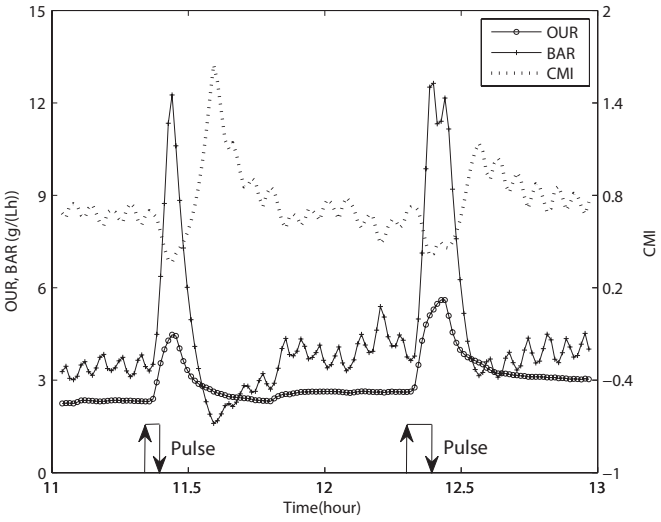

Fig. 2. This figure shows the results of the glucose pulse experiment. The arrows along the time axis indicate the location and duration of each pulse.

culture is between 0.70 and 1.14 and the CMI_{min} for this culture is between 0.408 and 0.70.

III. E.COLI BIOPROCESS SIMULATION MODEL

In order to further study the potential of the CMI as a metabolic indicator as well as explore its use for control, an appropriate bioprocess model was developed for simulation. The uptake rate model from Xu et al. [14] presented an accurate representation of *E. coli* metabolism and was used in many estimators and controllers [15], [16]. Two equations from the simulation model are shown below:

$$\begin{aligned}\mu(A,S) &= (qS_{ox} - q_m)Y_{X/S_{ox}} + qS_{of} + \\ &\quad qS_{of}Y_{X/S,of} + qA_cY_{X/A} \qquad (2)\\ \dot{H} &= qS_{ox}Y_{H/S_{ox}} + qS_{of}Y_{H/S_{of}} + qA_pY_{H/A} \\ &\quad +qA_cY_{H/A_{ox}} + BAR - \frac{F}{V}H \qquad (3)\end{aligned}$$

In Eq. (2) and Eq. (3), X, S, and A represent the concentrations of biomass, glucose, and acetate in the culture. The μ represents the growth rate, which shifts depending on the three metabolic phases. The F represents the input feed rate of glucose and V the culture volume. The q represents an uptake term, i.e. qS glucose uptake rate, qA_p acetate production rate, qA_c acetate consumption rate. A detailed representation of the Xu model can be found in [14].

The uptake rate model was augmented to form a new model (Xu^{+H} model) which includes acid production dyanamics, enabling simulation of BAR, see Eq. (3). Eq. (3) was built out of terms already present in the Xu model; the H represents the concentration of H^+ ions in the culture and the production and consumption of acidic byproducts according to each phase of *E. coli* metabolism.

IV. E.COLI BIOPROCESS SIMULATIONS

Using Simulink, the Xu^{+H} model was used to simulate the fed-batch culture pulse experiment. The reponse of the CMI in the different metabolic phases in reaction to simulated pulses matched that of the experiment. In another fed-batch simulation, with no acetate present, the CMI_{min} value was determined by examining the BAR curve, whose slope changes once overflow metabolism begins. The value for CMI_{mid} corresponds to some low growth rate, ($\mu < 0.15$ hr^{-1}).

The performance of an open-loop controller was compared with a CMI-based controller. Each simulation evaluated a controller's ability to keep the *E. coli* at a high rate of growth while not producing waste products, i.e. acetate. Each simulation starts in the fed-batch phase, with an initial level of acetate remaining from the overflow metabolism that occurred in the batch phase. Most *E. coli* experiments do not exceed 24 h in duration. Assuming 5 to 6 hours for batch phase, the following simulations show 18 hours of fed-batch culture. For the simulation, CMI_{min} is assumed to be known, either by online identification or CMI characterization from a previous experiment.

A. Open-Loop Controller

Open-loop controllers see wide use in industry and academia, the open loop controller uses known attributes about the strain, $Y_{X/S}$, and experimental setup, X_0V_0, to generate an exponential feeding curve at a rate set by the user, usually based on previous knowledge. Usually, an open-loop controller maintains the culture at a user-specified safe growth rate reasonably far away from overflow metabolism. For this simulation, a growth rate was chosen that would yield high biomass without producing acetate. The form of the open loop controller can been seen in eq (4).

$$F(t) = \frac{\mu_{set}X_0V_0}{Y_{X/S}(S_F - S_0)}\text{e}^{\mu_{set}t}, \qquad (4)$$

with μ_{set} representing the desired growth rate, X_0 the initial biomass concentration in the bioreactor, V_0 the initial volume, S_F the glucose concentration of the feed, $Y_{X/S}$ the glucose to biomass yield coefficient for this strain, and S_0 the initial glucose concentration [17].

B. A CMI-based Controller

With the ability to track metabolism using the CMI, this controller was designed to maintain the CMI close to CMI_{min}, ensuring a high growth rate without entering overflow metabolism. The CMI_{min} for this culture was identified to be 1.76 and CMI_d set to 1.78.

One issue with control for exponentially increasing processes is the lack of stability as the disturbances also exponentially increase [18]. Similarly, this also requires exponentially increasing control inputs. To overcome these constraints, a PID controller is moved into the exponential term of the simple exponential feed rate equation (eq (5),(6))

TABLE I

CONTROLLER SIMULATION RESULTS

	units	μ_{max}	CMI PID	Open Loop
setpoint		$\mu_{set} = 0.27\ h^{-1}$	CMI_d= 1.78	$\mu_{set} = 0.22\ h^{-1}$
$\bar{\mu}$	h^{-1}	0.27	0.24	0.21
$biomass, X$	g/L	126	106	78
$time, A \to 0$	hr	0	11	6.5

[18],

$$F = F_0 e^{\alpha t} \quad (5)$$

$$\alpha = K_p \left(CMI - CMI_d + K_i \int_0^t (CMI - CMI_d) \mathrm{d}t + K_d \frac{\mathrm{d}(CMI - CMI_d)}{dt} \right) \quad (6)$$

with the desired metabolic setpoint CMI_d, the proportional constant K_p, the integral constant K_i, and the differential constant K_d, with values of 1, 0.367, and 1.815, respectively. Initial gains were obtained using a 'no-overshoot' Zeigler-nichols tuning, and tuned to give the desired response.

V. RESULTS & DISCUSSION

The results of the simulations are found in table 1 and in Fig. 3. The maximum oxidative growth rate for the simulated *E. coli* bioprocess was 0.27 h^{-1} and maximum biomass yield of 126 g/L, seen in Table I under μ_{max}. The desired growth rates setpoints were set so that the TCA cycle could process the extracellular acetate remaining from batch phase while still achieving a high growth overall rate. Note that the goal of these tests was to achieve the highest biomass concentration using single setpoint control.

The highest average growth rate for the open loop controller was 0.21 h^{-1}, with a biomass production of 78 g/L. With a high growth rate setpoint, it took almost 6 hours for the extracellular acetate to be processed. Notice how the CMI value rises above CMI_{mid} as the acetate concentration rolls off in an exponential fashion, this shows the gradual reduction of the acetate influence in suppressing the rate of the TCA cycle. Once the acetate was processed, the TCA cycle could run at its full rate and the substrate concentration settled into a lower equilibrium . The metabolite consumption metabolism stops and the CMI represents only the oxidative metabolism.

The CMI controller attained the largest average growth rate of 0.24 h^{-1} and yielded the largest biomass with a value of 106 g/L. By controlling the CMI value to a setpoint of 101% of CMI_{min}, a high rate of oxidative metabolism was maintained while still having room to process the extracellular acetate. As the controller drives the CMI toward CMI_d, it is giving glycolysis and pyruvate decarboxylation more glucose to process, increasing the BAR to achieve the desired CMI. This action ensures that the gap between the current TCA rate and the desired TCA rate, as set by CMI_d, is filled with glucose as the acetate diminishes. Note the difference between the CMI profiles of the two controllers.

The results of the controller simulations demonstrated that choosing a CMI value close to CMI_{min} to control to yields 34% more biomass than an aggressively tuned open-loop controller. Experimental verification of these results would require numerous characterization runs and off-line processing to find a best guess for the maximum growth rate, μ_{1max}; however, determining the control parameters, CMI_{mid} and CMI_{min}, for the CMI PID controller could be found using only one experiment in which the CMI was cycled through it's entire range.

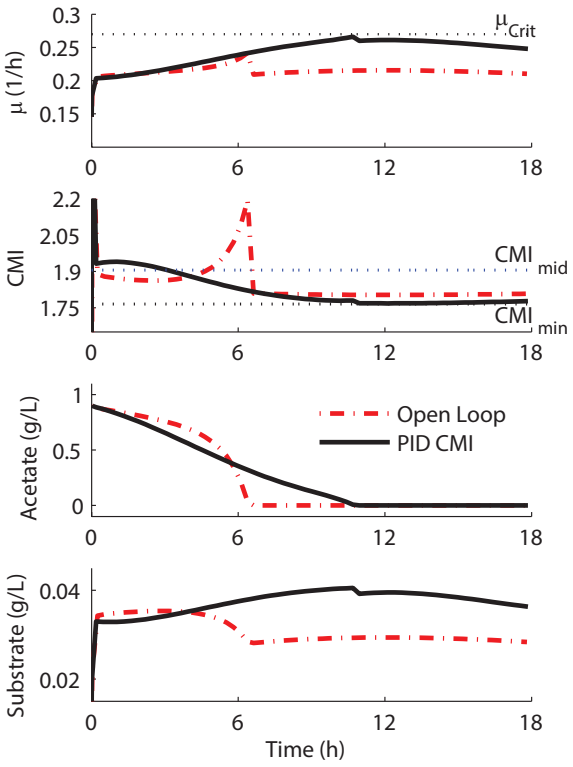

Fig. 3. Simulation Results of PID CMI (solid,black) & Open Loop (dashed,red) Controllers: A) Acetate B) Growth Rate C) CMI D)Substrate (glucose) concentration.

VI. CONCLUSION

In conclusion, this paper has presented a cell metabolic indicator (CMI) which gives more insight into the cells' metabolism than growth rate alone. Experimentally, the response of the CMI indicated the three distinct metabolic phases. The CMI consists of the ratio of two signals derived from the outputs of sensors common to most bioreactor systems. A bioprocess model was chosen that accurately reflected the metabolic behavior of an *E. coli* bioprocess, and uniquely augmented with the acid dynamics to allow for CMI simulation. Experimental response of *E. coli* cultures showed the CMI matched that seen in the simulated bioprocess, tracking the metabolic states of the *E. coli* . The simulations comparing a CMI controller with an open-loop controller demonstrated that the CMI controller maintained a higher growth rate yielding 34% more biomass. The fundamental advantage of using CMI is that the ratio of the oxidative metabolism to the total metabolism can be driven to a desired setpoint, rather than a growth rate, which is the

combination of the different metabolic phases and be difficult to determine closely. CMI_{min} is easier to identify empirically than μ_{1max}, a CMI-based controller can more reliably attain maximum output. Effective implementation of a CMI-based controller will rely on minimizing the time delay and phase shift effects,caused by sensor dynamics and pH controller dynamics in the OUR and BAR measurements. These effects will be compensated by using appropriately designed state estimators. Implementation of CMI-based control methods should enable researchers to achieve higher yield over conservative open-loop growth profiles as well as cut down on the characterization time of growth protocols. Many other cells types besides *E. coli* exhibit behavior similar to that characterized with the Xu model. The extension of the Xu^{+H} model to similar systems growing mammalian or pluripotent stem cells could enable the similar high growth rate control without excessive characterization.

REFERENCES

[1] U. S. Food and D. A. (USFDA), "Guidance for industry process analytical technology (pat) - a framework for innovative pharmaceutical development, manufacturing, and quality assurance," Sep 2004.

[2] R. G. Harrison, P. Todd, S. Rudge, and D. P. Petrides, *Bioseparations Science and Engineering*. New York, New York: Oxford University Press, Inc., 2003.

[3] N. S. Mosier and M. R. Ladisch, *Modern Biotechnology: Connecting Innovations in Microbiology and Biochemistry to Engineering Fundamentals*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2009.

[4] W. Johnston, R. Cord-Ruwisch, and M. Cooney, "Industrial control of recombinant e. coli fed-batch culture: new perspectives on traditional controlled variables," *Bioprocess and Biosystems Engineering*, vol. 25, no. 2, pp. 111–120, Jun 1 2002. [Online]. Available: http://dx.doi.org/10.1007/s00449-002-0287-8

[5] G. Bastin and D. Dochain, *On-line Estimation and Adaptive Control of Bioreactors*. Amsterdam, Netherlands: Elsevier Science, 1990.

[6] Y. Pomerleau and M. Perrier, "Estimation of multiple specific growth rates in bioprocesses," *AIChE Journal*, vol. 36, no. 2, pp. 207–215, 1990.

[7] V. N. Lubenova, "Stable adaptive algorithm for simultaneous estimation of time-varying parameters and state variables in aerobic bioprocesses," *Bioprocess Engineering*, vol. 21, no. 3, pp. 219–226, Sep 1 1999. [Online]. Available: http://dx.doi.org/10.1007/s004490050667

[8] D. C. Psichogios and L. H. Ungar, "Process modeling using structured neural networks," in *American Control Conference, 1992*, 1992, pp. 1917–1921.

[9] C. Karakuzu, M. Turker, and S. Ozturk, "Modelling, on-line state estimation and fuzzy control of production scale fed-batch baker's yeast fermentation," *Control Engineering Practice*, vol. 14, no. 8, pp. 959–974, Aug 2006.

[10] A. J. Wolfe, "The acetate switch," *Microbiology and Molecular Biology Reviews*, vol. 69, no. 1, pp. 12–50, Mar 2005. [Online]. Available: http://dx.doi.org/10.1128/MMBR.69.1.1250.2005

[11] Y. Liu, R. Hu, S. Zhang, L. Zhang, X. Wei, and L. Chen, "Expression of the foot-and-mouth disease virus vp1 protein using a replication competent recombinant canine adenovirus type 2 elicits a humoral antibody response in a porcine model," *Viral Immunology*, vol. 19, no. 2, pp. 202–209, Summer 2006.

[12] E. GarciaFruitos, A. Aris, and A. Villaverde, "Localization of functional polypeptides in bacterial inclusion bodies," *Applied and Environmental Microbiology*, vol. 73, no. 1, pp. 289–294, 2007.

[13] S. S. Sharma, J. W. Campbell, D. Frisch, F. R. Blattner, and S. W. Harcum, "Expression of two recombinant chloramphenicol acetyltransferase variants in highly reduced genome escherichia coli strains," *Biotechnology and Bioengineering*, vol. 98, no. 5, pp. 1056–1070, 2007.

[14] B. Xu, M. Jahic, and S. O. Enfors, "Modeling of overflow metabolism in batch and fed-batch cultures of escherichia coli," *Biotechnology progress*, vol. 15, no. 1, pp. 81–90, Jan-Feb 1999.

[15] I. Rocha, A. Veloso, S. Carneiro, E. C. Ferreira, and R. Costa, "Implementation of a specific rate controller in a fed-batch e.coli fermentation," in *Proceedings of the 17th World Congress of The International Federation of Automatic Control (IFAC)*, vol. 17, July 6-11,2008 2008, pp. 15 565–15 570.

[16] S. Tatiraju, M. Soroush, and R. Mutharasan, "Multi-rate nonlinear state and parameter estimation in a bioreactor," in *American Control Conference, 1998. Proceedings of the 1998*, vol. 4, 1998, pp. 2324–2328 vol.4.

[17] M. Jenzsch, S. Gnoth, M. Beck, M. Kleinschmidt, R. Simutis, and A. Lubbert, "Open-loop control of the biomass concentration within the growth phase of recombinant protein production processes," *Journal of Biotechnology*, vol. 127, no. 1, pp. 84–94, Dec 15 2006.

[18] M. Dabros, M. M. Schuler, and I. W. Marison, "Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass," *Bioprocess and Biosystems Engineering*, vol. 33, no. 9, pp. 1109–1118, Nov 1 2010. [Online]. Available: http://dx.doi.org/10.1007/s00449-010-0438-2