Optimal control and dynamic modulation
of the ATPase gene expression for enforced
ATP wasting in batch fermentations

Sebastián Espinel-Ríos∗∗∗, Bruno Morabito ∗∗
Johannes Pohlodek ∗∗, Katja Bettenbrock ∗, Steffen Klamt ∗
Rolf Findeisen ∗∗∗

Abstract: Competitive biotechnological processes need to operate over various conditions and adapt to changing economic contexts. Dynamic ATP turnover allows trading off declines in biomass formation and volumetric productivity for enhancements of product yields in fermentations where the product pathway is linked to ATP synthesis. To facilitate its practical implementation, we propose to dynamically manipulate the cellular ATP turnover by putting the ATPase enzyme, which hydrolyzes ATP into ADP, under the control of an optogenetic gene expression system. This allows achieving dynamic control of the ATP wasting online via a tunable external input. While light as control input is promising because it is easily tunable, generally non-invasive/non-toxic, and more affordable than classical chemical inducers, it makes the overall control task challenging. Thus, we derive an expanded version of dynamic enzyme-cost flux balance analysis that takes into account the dynamics of the optogenetic actuator. We then formulate a suitable optimal control problem to find optimal inputs for achieving the desired process performance. We test our approach in simulations using the batch anaerobic lactate fermentation of glucose by Escherichia coli as a case study.

Keywords: ATP wasting, ATPase, optimal control, gene expression, optogenetics, batch fermentation.

1. INTRODUCTION

Global challenges such as climate change and resource depletion are driving the move from a fossil-based to a circular bio-based economy (Yang et al., 2021). Biotechnologies can be key players in this transition by enabling the sustainable manufacturing of chemicals, materials and fuels from renewable resources (Clarke and Kitney, 2020). In this regard, metabolic engineers have developed a wide range of tools and methods to rewire metabolic networks towards enhancing the production of commercially relevant metabolites by microbial cell factories (Ko et al., 2020).

In the last years, enforced ATP wasting or turnover has gained attention as a useful metabolic design principle in fermentations where the product pathway is linked to ATP formation (e.g., Boecker et al. (2019); Zahoor et al. (2020); Boecker et al. (2021); Espinel-Ríos et al. (2022)). Under the latter conditions, introducing mechanisms that “waste” cellular ATP can lead to an increase in product yields (amount of product per amount of substrate) and specific productivities (amount of product per amount of biomass per time). Since the substrate is a shared resource in the metabolism, this also results in less biomass formation, hence lower volumetric productivities (amount of product per culture volume per time) in batch fermentations. Previous mathematical models of the ATP wasting effect, based on oversimplified and unstructured equations (Klamt et al., 2018), lack enough precision and predictability to enable efficient optimization and control.

Recently, we formulated optimal control problems to compute dynamic ATP turnover policies for maximizing batch fermentation efficiency (Espinel-Ríos et al., 2022) based on dynamic enzyme-cost flux balance analysis (deFBA) (Waldherr et al., 2015). We were able to predict different trade-offs between enhancement of product yield and decline in volumetric productivity. However, a remaining practical challenge is how to fine-tune the intracellular ATP turnover to make the implementation of this strategy possible in production setups. First, we propose to use the ATPase enzyme (F1-subunit) (Koebmann et al., 2002), which hydrolyses ATP into ADP, as the ATP wasting mechanism. This offers a degree of freedom to modulate the ATPase flux via fine-tuning of the ATPase gene expression. To achieve this, we consider to put the ATPase...
enzyme under the control of a CcaSR two-component optogenetic gene expression system. CcaS is a sensor histidine kinase that phosphorylates CcaR under green light. Phosphorylated CcaR binds to the promoter region, thereby inducing gene expression (Olson et al., 2014). Using light has many advantages as control input, e.g., it can be easily tuned online, it is generally non-toxic/non-invasive, and it is economically more affordable than chemical inducers (Carrasco-López et al., 2020).

To facilitate the model-based optimization of the described system, the considered fermentation model should be able to account for the dynamics of the CcaSR optogenetic activator. Unfortunately, the original deFBA modeling formulation does not cover this aspect. Therefore, the main contribution of this work is, we modify the original deFBA model to consider the CcaSR dynamics in batch fermentations (Section 2). We then formulate optimal control problems to find light input policies that maximize the batch fermentation efficiency in the ATP wasting context (Section 3). As a case study we use the anaerobic lactate fermentation from glucose by an Escherichia coli strain with the acetate and ethanol pathways blocked (Section 4). Note that, under these conditions, lactate production is linked to ATP synthesis, hence the ATP wasting strategy is applicable (Hädicke et al., 2015; Espinel-Ríos et al., 2022).

2. MODEL OF THE FERMENTATION WITH OPTOGENETIC REGULATION

First, we define the molar vector \( p(t) = [p_{\text{ATPase}}(t) \quad p_{\text{bc}}(t)]^T \) which contains all components that the cell is made up. The variable \( p_{\text{ATPase}} \) describes the concentration of the (regulated) ATPase enzyme and \( p_{\text{bc}} \) is a vector containing the remaining biomass components, including unregulated enzymes, ribosomes and quota elements (e.g., non-catalytic proteins, DNA, lipids, carbohydrates, etc.). By unregulated enzymes we mean that they are not controlled externally via an inducible gene expression system. Quota elements do not participate in catalytic reactions but they are necessary for vital functions such as cell structure, maintenance and reproduction. For simplicity, we assume that ribosomes catalyse the synthesis of all biomass components.

The dynamics of \( p_{\text{ATPase}} \) can be expressed as

\[
\frac{dp_{\text{ATPase}}}{dt} = F(p(t), I(t)) - D(p_{\text{ATPase}}(t)),
\]

where \( F: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R} \) and \( D: \mathbb{R} \rightarrow \mathbb{R} \) are functions of the ATPase enzyme production and degradation rates, respectively. The variable \( I(t) \) is the control input (in our case, the green light intensity) and \( p_{\text{ATPase}, 0} \) is the initial value of \( p_{\text{ATPase}} \). In the remainder of the paper, we will omit the explicit time dependency of \( I, p \) and \( p_{\text{ATPase}} \) when clear from the context. We assume that all cells receive the same light intensity and that there is homogeneous induction.

The CcaSR two-component system has been previously modeled in literature following the Hill equation (Olson et al., 2014). Normalized per biomass concentration, \( F \) can be described as

\[
F(p, I) = (b^T p) \left( \frac{\alpha_1 + \alpha_2 \frac{I^3}{\alpha_3 + I^3}}{1} \right),
\]

where \( b \) is the vector comprising the molecular weights of the elements in \( p \). Hence, \( b^T p \) is equal to the biomass concentration in g/L. The parameter \( \alpha_1 \) is the light-independent production rate, \( \alpha_2 \) is the light-dependent maximum rate of production, \( \alpha_3 \) is a saturation constant, and \( \beta \) is a coefficient that determines the steepness of the Hill function.

The function \( D \) is modeled as the product of the ATPase turnover degradation rate \( d_{\text{ATPase}} \) and \( p_{\text{ATPase}} \)

\[
D(p_{\text{ATPase}}) = d_{\text{ATPase}} p_{\text{ATPase}}.
\]

Note that dilution of \( p_{\text{ATPase}} \) by growth is implicitly considered in our model. Since we model all cellular components separately, the production of other compounds (e.g., other enzymes, ribosomes and quota elements) will have a dilution effect on the cellular ATPase.

Extracellular metabolites \( z \) and \( p_{\text{bc}} \) are collected in the molar vector \( x = [z \quad p_{\text{bc}}] \) which changes in time according to the product of the stoichiometric matrix \( S_x \) and the vector of fluxes \( V \) (in molar amount per time)

\[
\frac{dx(t)}{dt} = \left[ \frac{dz(t)}{dt} \quad \frac{dp_{\text{bc}}(t)}{dt} \right]^T = S_x V(t),
\]

where \( z_0 \) and \( p_{\text{bc}, 0} \) are the initial conditions. Here we assume that the degradation rate is negligible compared to the production rate, hence there is no degradation term in Eq. (4). However, the concentration of a \( p_i \) element will have a dilution effect related to the production of the remaining biomass components.

Intracellular metabolites \( m \) are assumed to be in quasi-steady state conditions

\[
0 = \frac{dm(t)}{dt} = S_m V(t),
\]

with \( S_m \) being the stoichiometric matrix of the species in \( m \).

Assuming for simplicity that the ATPase enzyme is working under substrate saturation conditions, and thus close to its maximum reaction rate, the ATPase flux is constrained by the amount of enzyme \( p_{\text{ATPase}} \) and its corresponding catalytic constant \( k_{\text{cat,ATPase}} \)

\[
\frac{V_{\text{ATPase}}(t)}{k_{\text{cat,ATPase}}} = p_{\text{ATPase}},
\]

where \( | \cdot | \) indicates the absolute value operator. Furthermore, the fluxes of the set of reactions \( \text{cat}_i \) catalyzed by an enzyme \( p_{\text{bc}} \), are constrained such that the maximum reaction rate must be less or equal than the product of the catalytic constant and the corresponding enzyme concentration.
The biomass dry weight must contain a fraction $\varphi_Q \in [0, 1]$ made of a quota compound $p_Q \in p$ that lumps all quota elements

$$\varphi_Q b^T p(t) \leq p_Q(t). \quad (8)$$

We also account for the metabolic cost to produce the ATPase enzyme via the CcaSR system

$$V_{\text{ATPase}}(t) = \frac{dp_{\text{ATPase}}(t)}{dt} = 0, \quad (9)$$

where $V_{\text{ATPase}}$ is the flux through the ATPase-producing reaction.

In addition, one can consider feasible lower and upper bounds for the fluxes

$$V_{\min}(t) \leq V(t) \leq V_{\max}(t). \quad (10)$$

Finally, we formulate an optimization problem to circumvent the fact that constraint-based models are often underdetermined

$$\max_{V(t)} \int_{t_0}^{t_{\text{deFBA}}} b^T p(t) \, dt \quad \text{s.t.} \quad \text{Eqs. (1) – (10)}. \quad (11)$$

In (11) we consider the maximization of the biomass integral over a time window $[t_0, t_{\text{deFBA}}]$ as the objective function of the cell. This objective function is frequently used in the frame of deFBA-based models for cells growing under non-starvation conditions (Waldherr et al., 2015; Reimers, 2017; Jabarivelisdeh and Waldherr, 2018; Jabarivelisdeh et al., 2020; Liu and Bockmayr, 2020; Espinel-Ríos et al., 2022).

### 2.1 Parameters of the lactate fermentation and the CcaSR gene expression system

As mentioned before, we focus on the anaerobic lactate fermentation from glucose by an *E. coli* strain with the acetate and ethanol pathways knocked-out. We followed a protocol for developing resource allocation models (Reimers et al., 2017) based on experimental data available for this specific lactate fermentation and strain (Hädicke et al., 2015). The resulting resource allocation model and parameters can be found in Espinel-Ríos et al. (2022). In Figure 1 we present the metabolic reactions considered in the network. We define a set of parameters in Table 1 to complement the latter model with the light-inducible CcaSR gene expression system dynamics.

In summary, the resulting model comprises 16 metabolic reactions and 18 biomass-producing reactions, thus a total of 34 fluxes. It contains 5 external metabolites (glucose, lactate, formate, succinate and carbon dioxide), 18 internal metabolites, and 18 cell components. Apart from the parameters in Table 1, there are also 34 catalytic constants, and 18 molecular weights for the biomass components. The value of $\varphi_Q$ for the referred model is constant and set to 0.67. Note that the biomass-producing reactions consider the cost of producing the biomass components in terms of amino acids and ATP.

### 3. MODEL-BASED OPTIMAL CONTROL OF LIGHT INTENSITY

We formulate an optimal control problem to find optimal green light inputs towards maximizing the batch efficiency. For the lactate fermentation we define the objective function to be the maximization of the lactate concentration at the end of the batch

$$\max_{I(t)} z_{\text{LAC}}(t_f) \quad \text{s.t.} \quad \text{Eq. (11)}. \quad (12)$$

It is expected that different $t_f$ values (fixed in the optimizations) will render different ATP turnover policies, and thus fermentation performance profiles. The selected batch time would constrain the maximum possible lactate yield enhancement achievable in the process through the ATP wasting strategy. Furthermore, any enhancement in the product yield should still allow for enough biomass formation and efficient substrate utilization towards maximizing the final lactate titer in the considered time frame.

The optimal control problem was solved using the collocation method based on Lagrange interpolation polynomials (Waldherr et al., 2015). The bilevel optimization problem was transformed into a single-level nonlinear program with complementary constraints by applying the Karush-Kuhn-Tucker conditions to (11), following an optimistic approach (Dempe and Franke, 2019). In the optimistic approach, using game theory terminology, the upper-level optimization (the “leader”) and the lower-level optimization (the “follower”) collaborate to satisfy the leader’s objective. This was deemed reasonable because the leader can directly manipulate the ATPase flux via the optogenetic activator, thereby already influencing the “best” $V$ distributions achievable by the follower. This helped to reduce the computational burden of the complementary slackness constraints and the other KKT conditions. The

---

1. For the derivation of the original model (Espinel-Ríos et al., 2022), the considered cell composition was 44% proteins (6% catalytic plus 38% non-catalytic), 27% ribosomes, and 29% other components (DNA, lipids, carbohydrates, etc.). Thus, $Q = (38 + 29)\%$.

2. Since the ultimate goal of the bilevel optimization is then to satisfy the leader’s objective, we assumed that the leader also influences the collocation points of the dynamics, hence the decision variables of the follower’s problem were limited to the unregulated fluxes.

---

### Table 1. Parameters of the CcaSR optogenetic system.

| Item | Value | Unit | Ref. |
|------|-------|------|------|
| $a_1$ | 0.02/2 | mmol/g/h | See note$^*$ |
| $a_2$ | 1 · 10^{-4} | mmol/g/h | See note$^*$ |
| $a_3$ | 0.138 | W/m² | (Olson et al., 2014) |
| $\beta$ | 2.49 | 1 | (Olson et al., 2014) |
| $d_{\text{ATPase}}$ | 6.3 · 10^{-2} | 1/h | (Benito et al., 1991) |

$^*$ Inferred from deFBA simulations.
3PG: 3-phospho-D-glycerate; AA: amino acid; AcCoA: acetyl-CoA; ADP: adenosine diphosphate; AKG: alpha-ketoglutarate; ATP: adenosine triphosphate; CO\textsubscript{2}: carbon dioxide; CoA: coenzyme A; DHAP: dihydroxyacetone phosphate; F6P: fructose 6-phosphate; FOR: formate; FUM: fumarate; G3P: glyceraldehyde 3-phosphate; G6P: glucose 6-phosphate; GLC: glucose; LAC: lactate; MAL: malate; NAD: nicotinamide adenine dinucleotide; NADH: NAD-reduced; OAA: oxaloacetic acid; PEP: phosphoenolpyruvate; PYR: pyruvate; SUCC: succinate.

Fig. 1. Metabolic pathway of the anaerobic lactate fermentation by E. coli. The ATPase enzyme (encoded by the \textit{atpAGD} gene) is regulated by the CcasSR optogenetic system. Production reactions of enzymes, ribosomes and the lumped quota compound are not depicted.

resulting optimization problem was solved using CasADi (Andersson et al., 2019) and IPOPT (Wächter and Biegler, 2006).

4. MODULATION OF THE ATPase GENE EXPRESSION

We solved the optimal control problem in (12) for three batch times, namely 13, 15 and 19 h. These scenarios were compared against the case with \( I = 0 \) at all control instances, i.e., no ATPase induction. The simulation results are shown in Figure 2. Relevant fermentation metrics were calculated to compare the process performance among the different scenarios. The average yield of product on substrate (\( Y_{PS} \)), the average yield of biomass on substrate (\( Y_{XS} \)), and the batch volumetric productivity (\( r_p \)) are presented in Table 2.

| \( t_f [h] \) | \( Y_{PS} [\text{mol/mol}] \) | \( Y_{XS} [\text{g/mol}] \) | \( r_p [\text{mmol/L/h}] \) |
|----------------|---------------------------|----------------|----------------|
| 7.5\*          | 1.72                      | 15.3           | 31.8           |
| 13             | 1.87                      | 7.2            | 20.0           |
| 15             | 1.90                      | 5.6            | 17.6           |
| 19             | 1.92                      | 4.5            | 14.0           |

\* Without induction of the ATPase gene expression.

Optimal optogenetic manipulation of the ATPase expression using light as control input allowed improving the product yield via exploitation of the ATP wasting concept. As expected, this happened at the expense of the biomass yield which in turn translated into lower volumetric productivity rates. The product yield was enhanced by 8.7, 10.3, and 11.7 % for batch times of 13, 15 and 19 h, respectively, compared to the case with no ATP wasting. That is, the longer the batch fermentation time, the higher the product yield enhancement because the ATP wasting effect can be applied for longer periods as we allow for lower volumetric productivity. In other words, the lower fermentation rates due to less biomass accumulation are compensated by the higher product yields achieved via ATP wasting.

In relation to the results, some may argue that running the 7.5 h-process (without ATPase induction) twice would achieve net higher production than running the 19 h-process once (with ATPase induction). Although this is in fact a valid point, the conclusion is not straightforward.
Fig. 2. Optimization results for the lactate fermentation considering different batch times. In the first row we show the case without induction of the ATPase enzyme expression, corresponding to a batch time of 7.5 h. The second to fourth rows belong to batch times of 13, 15 and 19 h, respectively.

because one should also consider that setting up a batch fermentation is labour-intensive, often involving a high proportion of unproductive time or down-time (Macauley-Patrick and Finn, 2008). For example, between batches, the operators would need to clean the equipment and connections, sterilize the bioreactor, allow time for cooling, charge/inoculate the bioreactor, discharge the bioreactor content, and restart the batch. Moreover, a bioprocess
generally involves not only the fermentation, but also upstream and downstream unit operations. Therefore, we recommend that the selection of a given yield and volumetric productivity trade-off should ideally come from the result of an overall optimization of the plant, considering scheduling, labour, upstream and downstream processes, and of course the fermentation. The outcome would be process-specific and not possible to generalize.

With regards to the fermentation dynamics under the effect of ATP wasting, one can see that the the ATPase enzyme expression is induced at a high rate at the beginning of the batch and then gradually decreased. As expected, with increasing ATPase accumulation, and thus ATP wasting level, the biomass growth rate starts to decrease even down to non-growth scenarios, e.g., by the end of the batch. The substrate flux is then redirected towards more product formation, explaining the enhanced product yields.

It is worth noting that the gene expression of the ATPase enzyme reached close to or slightly above 10% of the biomass dry weight in the considered scenarios. Over-expression of regulated genes can impose a considerable burden to the cell since there is an interplay between the usage of resources and cell growth (Santos-Navarro et al., 2021). That is, a very high protein expression can negatively affect cell growth and productivity. Therefore, the observed reduction of the growth rate with increasing ATPase expression should be in principle a combined effect of the ATP wasting mechanisms and any possible burden associated with the increased cost of protein synthesis. In fact, the higher the intensity of the ATPase induction, the less amino acids and energy co-factors such as ATP available for the synthesis of other enzymes and cell components required to sustain the cell viability. The optimizer should consider this potential resource burden into account when computing the optimal light input trajectories because this is an intrinsic aspect of our model (see e.g., Eq. (9)).

Another aspect to consider is that the implementation of optogenetics in large-scale bioreactors is not yet solved. In fact, the model assumes that all cells will receive the same light intensity and that the induction will be homogeneous. In small-scale bioreactors this assumption may hold, but it could be put into question for bigger bioreactors where gradients of all types might arise. For example, high cell densities in bioreactors may cause problems related to light penetration and the fermentation performance could become light-limited (Carrasco-López et al., 2020).

Furthermore, in our model we do not take into account the evolutionary stability of the CcaSR gene expression system. For instance, if the employed gene expression system imposes a heavy burden to the cell fitness, the cells may mutate to inactivate/reverse/modify the engineered elements driven by evolutionary principles. These effects could be reduced, e.g., by overlapping the costly gene expression system with essential genes (Blazejewski et al., 2019). We recommend that evolutionary stability should be assessed when considering experimental applications of the presented optimal control strategy.

Overall, in line with previous observations (Espinel-Ríos et al., 2022), these results support the idea that dynamic ATP turnover can be used as a way to find trade-offs between product yield and volumetric productivity (see Table 2). This provides flexibility to the batch-to-batch operation as the plant can easily adapt to changes in the economic context. For example, having a high product yield could be of particular interest in cases where the substrate cost has a significant share within the operational costs. The volumetric productivity might be more relevant if production time is a big constraint. In other scenarios, an economic analysis of the integrated plant, including upstream and downstream processing, might tell that the highest profit occurs at a given trade-off between product yield and volumetric productivity. With our approach, instead of developing a new strain every time that a given performance metric is required, one could keep the same production strain and achieve a different phenotype just by adjusting the process input online. Naturally, process development can be significantly shortened, making the biotechnology industry more flexible and competitive.

5. CONCLUSION AND OUTLOOK

We presented a model-based optimization strategy for modulating the ATP turnover in the cell via CcaSR-mediated optogenetic regulation of the ATPase enzyme expression in a batch anaerobic lactate fermentation. To do so, we first derived an expanded version of the deFBA model that takes into account the dynamics of the optogenetic gene expression system. After solving suitable optimal control problems, it was possible to improve the product yield of the process at the expense of the volumetric productivity by dynamically changing the light intensity throughout the fermentation. Different trade-offs between product yield and volumetric productivity were obtained. With the possibility of optogenetic modulation of the ATP turnover one could to easily adjust the batch-to-batch strain’s performance in production setups. This could provide the biotechnology industry with a higher degree of flexibility and adaptability to changing economic contexts.

Based on our results, we are working on the implementation of an automatic light delivery system to enable light-mediated ATP wasting applications. Since open-loop optimization can be affected by uncertainty (e.g., model-plant mismatch or disturbances), we are also considering advanced feedback control schemes such as model predictive control. Furthermore, we are developing state estimators for facilitating process monitoring and automatic control. Future work also includes the generalization of a modeling and predictive control framework to enable metabolic cybergenetic applications with focus on optogenetics and fed-batch fermentations.

REFERENCES

Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., and Diehl, M. (2019). CasADi: a software framework for nonlinear optimization and optimal control. Math. Program. Comput., 11(1), 1–36.

Benito, B., Moreno, E., and Lagunas, R. (1991). Half-life of the plasma membrane ATPase and its activating system in resting yeast cells. Biochim. Biophys. Acta - Biomembr., 1063(2), 265–268.
