Cloud-Orchestrated Autonomous Bioreactor Arrays for Closed-Loop Strain Characterization

Carlos Barajas, Justin Edaugal, Samuel McKey, Seneca Bessling
Johns Hopkins Applied Physics Laboratory
11100 Johns Hopkins Road
Laurel, MD 20723 USA
Carlos.Barajas@jhuapl.edu

Abstract

Biomanufacturing of biofuels, food, and energetic materials is increasingly important for defense, security, and disaster response operations, where supply-chain resilience and field-forward production are critical. Here, we present a cloud-orchestrated, edge-executed autonomous biomanufacturing platform built around an open-source, parallel mini-bioreactor array. The system integrates *in situ* sensing, state and growth-parameter estimation under uncertainty, and closed-loop Bayesian optimization to autonomously plan and execute experiments. We show that the platform identifies the optimal growth temperature of *Escherichia coli* in a single continuous run completed in under 24 hours, and enables parallel screening of *Rhodopseudomonas palustris* to identify favorable carbon sources and temperatures. This work demonstrates a deployable sense–reason–act autonomy loop on a real physical system.

Introduction

Operational autonomy research often focuses on uncrewed aerial systems operating under uncertainty, degraded communications, and tight resource constraints. A parallel challenge arises in biomanufacturing: autonomously discovering and maintaining cultivation conditions and strain configurations that maximize growth and/or product yield under limited time, instrumentation, and human supervision (Gargalo et al. 2020; Xie 2022). This research is especially relevant for defense, security, and disaster response, where field-forward biomanufacturing could produce fuels, energetic materials, food, or other critical materials on demand from locally available feedstocks (Hodgson, Maxon, and Alper 2022; Smanski et al. 2022).

Conventional strain and process development relies on manual sampling, repeated batch experiments, and sequential offline design-of-experiments, making iteration slow and difficult to deploy outside controlled labs (Stephanopoulos 2012). In addition, biological nonstationarity and measurement noise complicate both optimization and sustained operation (Siska et al. 2025). These issues map naturally to a sense-reason-plan-act autonomy loop: (i) sensing partial and noisy state (e.g., optical density, temperature, pump activity), (ii) reasoning under uncertainty to infer latent state and growth parameters, (iii) planning operational adjustments (e.g., temperature, dilution, media composition), and (iv) execution via closed-loop control for stable continuous operation. For field-forward use, this loop must run continuously and adapt without manual retuning.

In this short paper, we demonstrate an Artificial Intelligence (AI) enabled autonomy stack for bioprocess discovery and operation using an array of open-source mini-bioreactors (Steel et al. 2020) with cloud/ application programming interface (API) orchestration and edge-level control. The open-source platform enables rapid customization, replication, and scaling, while local microcontroller-based control supports continuous sensing and actuation and cloud connectivity enables longer-horizon coordination across reactors. Within this framework, the system autonomously identifies a productivity-sustaining culture-density operating point during batch growth and uses it to initialize downstream continuous operation (Hoffmann et al. 2017). Our implementation builds on the Chi.Bio platform, which provides an open-source, parallelized hardware and software foundation for *in situ* measurement and actuation, including continuous culture capabilities (Steel et al. 2020). We make the following contributions:

- A scalable, cloud-orchestrated 16-reactor platform enabling parallel strain and condition screening with edge-level closed-loop control.
- State and growth-parameter estimation under uncertainty to support autonomous decision-making during continuous operation.
- A closed-loop Bayesian optimization controller for continuous-mode growth-rate maximization using *in situ* optical density sensing and temperature actuation.
- Autonomous demonstrations on *Escherichia coli* and *Rhodopseudomonas palustris*, including single-run temperature optimization and multi-stage screening and refinement workflows.

Methods

Reactor Array and Cloud Control

Our experimental platform comprises 16 mini-bioreactors constructed from the Chi.Bio reactor concept (in situ optical density (OD) measurement, temperature actuation, and peristaltic pumping for turbidostat operation) (Steel et al. 2020).

Reactors are grouped under embedded microcontroller units (MCUs), each of which manages up to eight reactors and executes a periodic automation cycle for each reactor that(i) acquires OD and auxiliary telemetry, (ii) updates control actions (e.g., dilution, temperature), and (iii) logs time-series state for downstream inference. To enable scaling beyond a single controller, control is distributed across multiple MCUs, each exposing a REpresentational State Transfer (REST) API that supports remote monitoring and actuation of the reactors it manages.

A lightweight client periodically polls the control and measurement state of each controller via HTTP GET requests and issues HTTP POST requests to provide (i) centralized decision-making, (ii) reproducible logs, and (iii) the ability to run experiments unattended while retaining supervisory control (pause/resume, safety bounds).

State-space Growth-Rate Estimation With an Extended Kalman Filter

Estimating microbial growth rate during continuous operation is challenging because OD is actively regulated near a setpoint, suppressing the exponential growth dynamics typically used for inference. Prior work has shown that allowing small fluctuations around the setpoint enables growth-rate estimation using Kalman filtering frameworks (Hoffmann et al. 2017). Inspired by this approach, we explicitly model dilution by incorporating measured pump inputs into the OD dynamics. The resulting nonlinear state-space model captures actuator-driven biomass removal, enabling online estimation of growth rate in the presence of periodic dilutions, while recursive filtering provides uncertainty-aware estimates suitable for closed-loop decision-making.

Let y_k denote the measured optical density at time t_k with sampling interval $\Delta t_k = t_{k+1} - t_k$ (hours). We define the latent state

$$\mathbf{x}_k = \begin{bmatrix} \mathrm{OD}_k \\ \mu_k \\ a_k \end{bmatrix}, \tag{1}$$

where μ_k is the instantaneous specific growth rate (hr^{-1}) and a_k is an effective dilution/clearance gain mapping the commanded pump input U_k to an OD removal rate (hr^{-1} per unit U). The process model uses exponential biomass propagation:

$$\mathrm{OD}_{k+1} = \mathrm{OD}_k \exp\Big(\big(\mu_k - a_k U_k\big)\,\Delta t_k\Big) + w_k^{(\mathrm{OD})}, \tag{2}$$

$$\mu_{k+1} = \mu_k + w_k^{(\mu)}, \tag{3}$$

$$a_{k+1} = a_k + w_k^{(a)}, \tag{4}$$

and the measurement model is

$$y_k = \mathrm{OD}_k + v_k, \tag{5}$$

with zero-mean Gaussian process and measurement noise, $w_k \sim \mathcal{N}(0, Q)$ and $v_k \sim \mathcal{N}(0, R)$.

Because Eq. (2) is nonlinear, we apply EKF linearization. Denoting the deterministic state update by $\mathbf{x}_{k+1} = f(\mathbf{x}_k, U_k)$, we propagate the covariance using the Jacobian

$$F_k = \left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_k, U_k} = \begin{bmatrix} E_k & c_k & -c_k U_k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

$$E_k \equiv \exp\Big(\big(\mu_k - a_k U_k\big)\Delta t_k\Big), \quad c_k = \mathrm{OD}_k E_k \Delta t_k, \tag{7}$$

and the measurement Jacobian $H = [1 \quad 0 \quad 0]$. The EKF recursion is:

Predict:

$$\hat{\mathbf{x}}_{k+1}^{-} = f(\hat{\mathbf{x}}_k, U_k), \tag{8}$$

$$P_{k+1}^{-} = F_k P_k F_k^{\top} + Q, \tag{9}$$

Update:

$$K_{k+1} = P_{k+1}^{-} H^{\top}\big(H P_{k+1}^{-} H^{\top} + R\big)^{-1}, \tag{10}$$

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_{k+1}^{-} + K_{k+1}\big(y_{k+1} - H\hat{\mathbf{x}}_{k+1}^{-}\big), \tag{11}$$

$$P_{k+1} = (I - K_{k+1}H)P_{k+1}^{-}. \tag{12}$$

The filter yields time-resolved estimates $\hat{\mu}(t)$ and associated uncertainty $\sigma_\mu(t) = \sqrt{P_{\mu\mu}(t)}$ even during dilution, while explicitly accounting for pump actuation via U_k.

Initialization and robustness. We initialize μ_0 from a short early-time log-OD fit and set R from the residual OD variance after mild smoothing. Process noise Q is chosen to allow slow drift in μ and a while preventing unphysical overfitting. OD is constrained to remain positive in the propagation step to avoid numerical issues in the exponential model.

Autonomous OD Setpoint Selection via Biomass Production Maximization

In continuous biomanufacturing, the effective production rate is proportional to the amount of biomass removed from the system, which depends jointly on biomass concentration and growth rate. Operating at low optical density (OD) underutilizes reactor capacity, while excessively high OD can suppress growth and reduce net productivity. To autonomously select a productive operating point, we define a biomass production proxy

$$p(t) \equiv \mathrm{OD}(t)\,\mu(t), \tag{13}$$

and estimate $\hat{p}(t) = \hat{\mathrm{OD}}(t)\hat{\mu}(t)$ online using state and parameter estimates. This estimation is performed entirely during an initial batch-growth phase with no dilution. As nutrients are consumed and growth conditions evolve, $p(t)$ initially increases as biomass accumulates, but subsequently declines as growth slows under resource limitation, reflecting a tradeoff between increasing biomass concentration and declining growth rate. Rather than reacting to transient fluctuations, the controller monitors $\hat{p}(t)$ and identifies its maximum during a increase in OD. Although OD continues to increase beyond this point, the OD at which $\hat{p}(t)$ attains its maximum is recorded and selected as the continuous-culture operating point associated with high biomass production,

$$\mathrm{OD}_{\mathrm{sp}} \leftarrow \hat{\mathrm{OD}}(t^{\star}), \tag{14}$$

where $t^{\star}$ denotes the time at which $\hat{p}(t)$ attained its maximum. The system then transitions to turbidostat operation at this setpoint, anchoring continuous production to a data-driven optimum identified during batch growth rather than a manually chosen threshold.

Bayesian Optimization for Autonomous Experiment Planning

We cast temperature selection as sequential optimization of the (Kalman-filter) growth-rate response $g(T)$. After each trial j at temperature T_j, we compute a scalar observation from the filtered trajectory,

$$y_j \equiv \hat{g}(T_j) \ \text{(trial-averaged Kalman } \hat{\mu}), \tag{15}$$

with an associated variance estimate s_j^2 obtained by combining within-trial temporal variability of $\hat{\mu}(t)$ and the Kalman posterior variance.

GP surrogate with fixed noise. We model $g(T)$ with a Gaussian process surrogate,

$$\begin{aligned} g(\cdot) &\sim \mathcal{GP}\big(m(\cdot), k(\cdot,\cdot)\big), \\ y_j &= g(T_j) + \epsilon_j, \quad \epsilon_j \sim \mathcal{N}(0, s_j^2 + \sigma_0^2). \end{aligned} \tag{16}$$

where σ_0^2 is a small jitter term for numerical stability. In implementation we use a single-task GP with a fixed-noise Gaussian likelihood (Balandat et al. 2020), setting the likelihood noise per trial to $s_j^2 + \sigma_0^2$.

Acquisition and optimization (qLogEI). The next temperature is chosen by maximizing the logarithm of expected improvement (qLogEI) relative to the best observed value:

$$T_{n+1} = \arg\max_{T \in \mathcal{T}} \log \mathrm{EI}(T;\, f^\star), \qquad f^\star = \max_{j \le n} y_j + \beta, \tag{17}$$

where $\beta \geq 0$ is a small exploration bias. We optimize qLogEI over bounded temperature limits using multi-start optimization with quasi-random initial samples (e.g., N_{raw} seeds and N_{restart} restarts), and map between normalized $T \in [-1, 1]$ and physical units via $T = T_{\text{center}} + D_T x$.

Closed-loop trial structure. Each Bayesian optimization iteration proceeds as follows: the reactor temperature is set to a candidate value T_j, transients are allowed to settle, and time-series OD data are collected. An extended Kalman filter is applied online to estimate the instantaneous growth rate $\hat{\mu}(t)$ and its uncertainty, which are aggregated over the trial to obtain a scalar observation $(\hat{g}_j, \hat{\sigma}_j^2)$. This observation updates the GP posterior, and the next temperature T_{j+1} is selected by maximizing the qLogEI acquisition function. Together, this approach implements a closed autonomy loop of sensing → filtering/inference → planning → execution.

Results

We evaluate the proposed autonomy stack on an open-source, parallel mini-bioreactor platform (Fig. 1) by demonstrating fully autonomous batch-to-continuous operation and closed-loop optimization of microbial growth conditions, first for *E. coli* and then for *R. palustris*. We first validate the approach using *E. coli* by optimizing growth rate as a function of temperature in continuous culture. Experiments are executed on the parallel reactor platform shown in Fig. 1, with local microcontroller-based control and remote supervisory optimization.

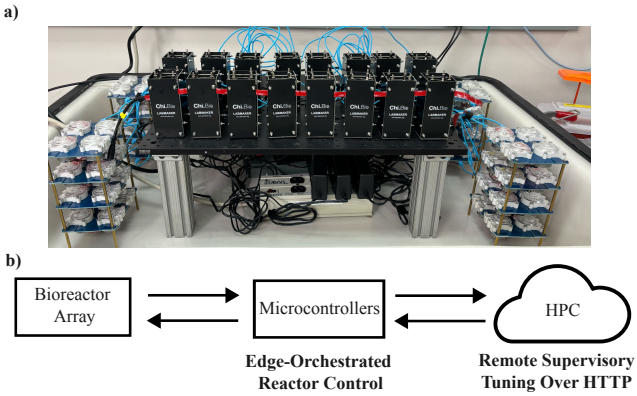

Figure 1: **Autonomous bioreactor platform and control architecture.** (a) Photograph of the parallel mini-bioreactor array used in this study, based on an open-source Chi.Bio-derived design. (b) Schematic of the control architecture: individual reactors are regulated by local microcontrollers enabling continuous sensing and actuation, while higher-level autonomy and supervisory optimization are performed remotely via cloud/HPC resources over HTTP.

E. coli: Autonomous Temperature Optimization

Each run begins in batch mode, during which the system autonomously infers growth rate and identifies the OD setpoint that maximizes biomass production. The OD corresponding to maximal biomass production is identified at time $t^\star$, after which batch growth is monitored to ensure sustained high productivity. The system then transitions to turbidostat operation and performs Bayesian optimization over temperature. Figure 2a shows the batch growth phase, autonomous OD setpoint identification, and subsequent continuous regulation. The estimated growth rate and biomass production proxy during batch growth are shown in Fig. 2b, with $t^\star$ marking the time at which the productivity-maximizing OD is identified.

In continuous operation, the system autonomously executes temperature trials (Fig. 3a) and updates a Gaussian process model of growth rate as a function of temperature (Fig. 3b). In a single fully autonomous run lasting approximately 25 hours, the optimizer converges to a growth-rate maximum near 37°C, consistent with the known optimal temperature for *E. coli* (Sezonov, Joseleau-Petit, and D'Ari 2007).

R. palustris: Parallel Carbon-Source Screening and Autonomous Temperature Optimization

We next apply the autonomy stack to *Rhodopseudomonas palustris*, a harder-to-grow organism whose performance depends strongly on media composition and temperature. Using the parallel reactor array, we screened eight carbon sources in parallel. For each carbon source, reactors were first operated in batch mode at a fixed initial temperature to autonomously identify a productivity-maximizing OD setpoint, following the same batch-phase logic used for *E. coli*. During this phase, temperature was held constant and no di-

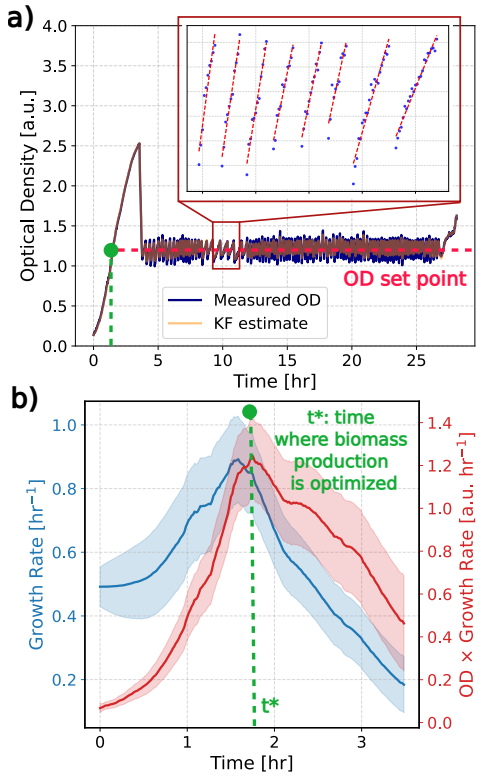

Figure 2: **Autonomous growth mode optimization for *E. coli*.** (a) Optical density (OD) trajectory along with its Kalman filter (KF) estimate, showing batch growth, autonomous identification of a productivity-maximizing OD setpoint, and subsequent transition to continuous (turbidostat) operation. (b) Estimated growth rate and biomass production proxy $p(t) = \mathrm{OD}(t)\mu(t)$ during batch growth; $t^\star$ denotes the time at which the OD corresponding to maximal biomass production is identified.

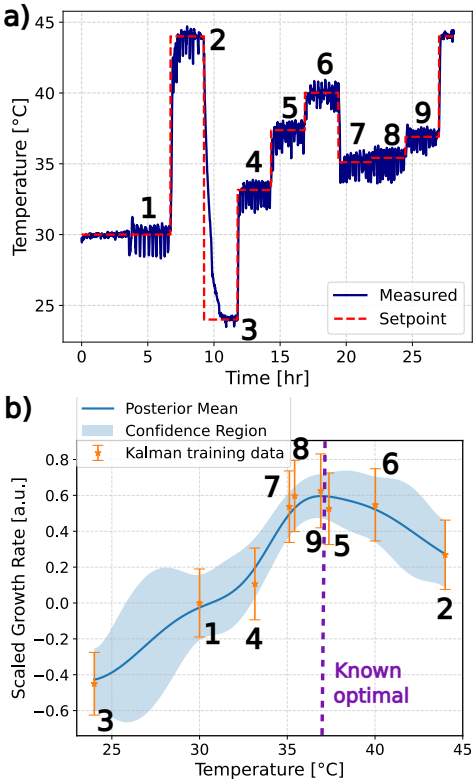

Figure 3: **Autonomous temperature optimization for *E. coli*.** (a) Temperature setpoints commanded autonomously during Bayesian optimization trials in continuous culture (growing around setpoint in Fig. 2a). (b) Gaussian process posterior over growth rate as a function of temperature, showing convergence to the known optimal temperature near 37°C.

lution was applied.

After the OD setpoint was identified, each reactor transitioned to continuous (turbidostat) operation at that setpoint. Bayesian optimization was then used to autonomously vary temperature and estimate steady-state growth rates, enabling direct comparison of carbon source–temperature combinations. Figure 4 summarizes the resulting growth rates across conditions. Among the tested substrates, fumarate/acetate (5 mM) yielded the highest growth rates, with optimal performance observed in the 34–37°C range. These results show how the platform enables rapid, fully autonomous screening of carbon source and temperature combinations in a single experimental campaign, providing an efficient first-pass characterization for non-model organisms.

Conclusion and Perspective

This short paper demonstrates an AI-enabled autonomy capability for bioprocess discovery and operation using a parallel, open-source Chi.Bio-derived reactor array with edge-level control, cloud/API orchestration, online state and parameter estimation, and Bayesian optimization. By focusing on simple, continuously observable quantities such as optical density and growth rate, the system enables robust closed-loop decision-making under realistic experimental constraints. These results align with the symposium's emphasis on bridging sensing to execution in autonomous physical systems and demonstrate how autonomy principles can translate to deployable biomanufacturing platforms.

Near-Term Extensions

- **Optimizing production, not just growth.** While we focus on online growth metrics, many applications target metabolite production; additional sensing modalities, such as inline nuclear magnetic resonance (NMR), can be incorporated to directly optimize yield.
- **Multi-variable optimization.** Although demonstrated here with single-variable Bayesian optimization over temperature, the framework naturally extends to multi-variable optimization over media composition, induction timing, or control parameters, enabling systematic exploration of higher-dimensional design spaces.
- **Scaling and transfer.** Transferring operating conditions

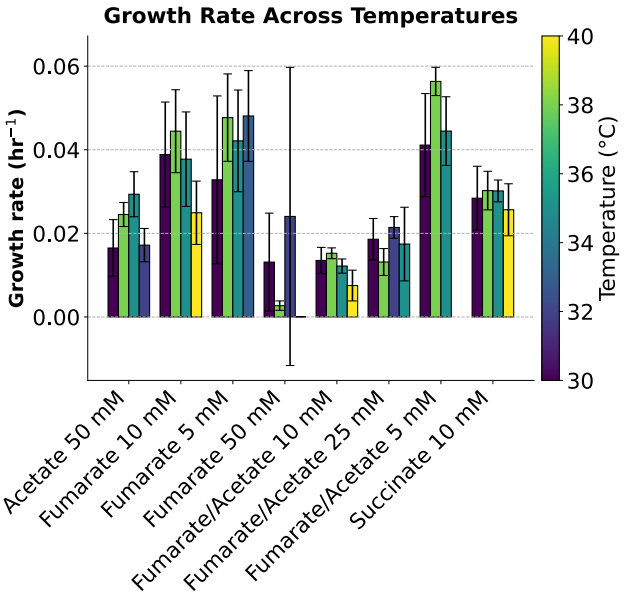

Figure 4: **Parallel autonomous screening of *R. palustris* growth conditions.** Growth rates measured across eight carbon sources following batch-phase identification of productivity-maximizing OD setpoints and subsequent continuous-culture Bayesian optimization over temperature. Error bars indicate variability across temporal averages. Fumarate/acetate (5 mM) exhibits the highest growth rates, with peak performance in the 34–37°C range.

across reactor scales remains a key challenge in biomanufacturing (Helleckes et al. 2022). This framework can support automated scale-up by coordinating experiments across reactors of different sizes and adaptively retuning operating points as system dynamics change. In addition, the platform enables parallel scaling and comparison across multiple microbial strains, facilitating systematic evaluation of strain-dependent process performance (Guarino et al. 2019).

- **Optogenetics and adaptive control.** The reactor platform supports fluorescence sensing, programmable light-emitting diode (LED) inputs, and ultraviolet (UV) actuation, enabling closed-loop optogenetic control (Pouzet et al. 2020) and autonomous adaptive evolution (Hirasawa and Maeda 2022). Fluorescence readouts can be integrated directly into the autonomy loop for real-time optimization, while UV perturbations enable systematic exploration of evolutionary adaptation.

Limitations and Robustness Considerations

Primary failure modes include biased state or parameter estimates during unmodeled transients, optimizer overconfidence when uncertainty is underestimated, and nonstationary objectives driven by biological adaptation. Addressing these challenges will require drift-aware estimation, change-point detection, and constraint-aware optimization to support reliable long-duration autonomous operation.

AI Statement

ChatGPT (version 5.2) was used to assist with streamlining and editing portions of this manuscript. The authors reviewed and validated all content.

Acknowledgments

This work was supported by internal research funding at the Johns Hopkins Applied Physics Laboratory.

References

Balandat, M.; Karrer, B.; Jiang, D. R.; Daulton, S.; Letham, B.; Wilson, A. G.; and Bakshy, E. 2020. BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. ArXiv:1910.06403.

Gargalo, C. L.; Udugama, I.; Pontius, K.; Lopez, P. C.; Nielsen, R. F.; Hasanzadeh, A.; Mansouri, S. S.; Bayer, C.; Junicke, H.; and Gernaey, K. V. 2020. Towards smart biomanufacturing: a perspective on recent developments in industrial measurement and monitoring technologies for bio-based production processes. *Journal of Industrial Microbiology & Biotechnology*, 47(11): 947–964.

Guarino, A.; Shannon, B.; Marucci, L.; Grierson, C.; Savery, N.; and di Bernardo, M. 2019. A low-cost, open-source Turbidostat design for *in-vivo* control experiments in Synthetic Biology. *IFAC-PapersOnLine*, 52(26): 244–248.

Helleckes, L. M.; Hemmerich, J.; Wiechert, W.; Lieres, E. v.; and Grünberger, A. 2022. Machine learning in bioprocess development: From promise to practice. ArXiv:2210.02200.

Hirasawa, T.; and Maeda, T. 2022. Adaptive Laboratory Evolution of Microorganisms: Methodology and Application for Bioproduction. *Microorganisms*, 11(1): 92.

Hodgson, A.; Maxon, M. E.; and Alper, J. 2022. The U.S. Bioeconomy: Charting a Course for a Resilient and Competitive Future. *Industrial Biotechnology*, 18(3): 115–136.

Hoffmann, S. A.; Wohltat, C.; Müller, K. M.; and Arndt, K. M. 2017. A user-friendly, low-cost turbidostat with versatile growth rate estimation based on an extended Kalman filter. *PLOS ONE*, 12(7): e0181923.

Pouzet, S.; Banderas, A.; Le Bec, M.; Lautier, T.; Truan, G.; and Hersen, P. 2020. The Promise of Optogenetics for Bioproduction: Dynamic Control Strategies and Scale-Up Instruments. *Bioengineering*, 7(4): 151.

Sezonov, G.; Joseleau-Petit, D.; and D'Ari, R. 2007. *Escherichia coli* Physiology in Luria-Bertani Broth. *Journal of Bacteriology*, 189(23): 8746–8749.

Siska, M.; Pajak, E.; Rosenthal, K.; Chanona, A. d. R.; Lieres, E. v.; and Helleckes, L. M. 2025. A Guide to Bayesian Optimization in Bioprocess Engineering. ArXiv:2508.10642.

Smanski, M. J.; Aristidou, A.; Carruth, R.; Erickson, J.; Gordon, M.; Kedia, S. B.; Lee, K. H.; Prather, D.; Schiel, J. E.; Schultheisz, H.; Treynor, T. P.; Evans, S. L.; Friedman, D. C.; and Tomczak, M. 2022. Bioindustrial manufacturing readiness levels (BioMRLs) as a shared framework for measuring and communicating the maturity of bioproduct manufacturing processes. *Journal of Industrial Microbiology and Biotechnology*, 49(5): kuac022.

Steel, H.; Habgood, R.; Kelly, C. L.; and Papachristodoulou, A. 2020. In situ characterisation and manipulation of biological systems with Chi.Bio. *PLOS Biology*, 18(7): e3000794.

Stephanopoulos, G. 2012. Synthetic biology and metabolic engineering. *ACS synthetic biology*, 1(11): 514–525.

Xie, D. 2022. Continuous biomanufacturing with microbes — upstream progresses and challenges. *Current Opinion in Biotechnology*, 78: 102793.