scientific reports

Check for updates

OPEN

A hybrid in silico/in-cell controller that handles process-model mismatches using intracellular biosensing

Tomoki Ohkubo[1]✉, Yuichi Sakumura[1,2], Fuzhong Zhang[3] & Katsuyuki Kunida[1,4]

The discrepancy between model predictions and actual processes, known as process-model mismatch (PMM), remains a substantial challenge in bioprocess optimization. We previously introduced a hybrid in silico/in-cell controller (HISICC) that combines model-based optimization with cell-based feedback to address this problem. Here, we extended this approach to regulate a key enzyme level using intracellular biosensing. The extended HISICC was implemented using an *Escherichia coli* strain engineered for fatty acid production (FA3). This strain contains a genetically encoded feedback controller that decelerates the expression of acetyl-CoA carboxylase (ACC) in response to malonyl-CoA synthesized through the enzymatic reaction. We modeled FA3 to allow the HISICC to optimize an inducer input that accelerates the enzyme expression. Simulations showed that the HISICC slowed the unexpectedly rapid accumulation of ACC resulting from PMMs before it reached cytotoxic levels, thereby improving fatty acid yields. These results highlight the potential of our approach, particularly in cases where monitoring intracellular biomolecules is required to handle PMMs.

Keywords Bioprocess control, Genetic circuit engineering, Model-based optimization, Synthetic biology, Negative feedback

Abbreviations

S	Extracellular glucose concentration (*Substrate*)
X	Normalized cell density
E	Intracellular ACC concentration (*Enzyme*)
M	Intracellular malonyl-CoA concentration (*interMediate*)
R	Intracellular LacI concentration (*Repressor*)
P	Extracellular fatty acid concentration (*Product*)
I	Extracellular IPTG concentration (*Inducer*)
u	Inducer 2 input for FA4
y_1	Measured relative cell density
y_2	Measured fatty acid concentration

Model parameters

k_X	Maximum growth rate
k_E	Maximum ACC expression rate
k_M	Malonyl-CoA synthesis rate coefficient
k_P	Fatty acid production rate coefficient
$k_{R,1}$	Rate of LacI expression from the chromosome
$k_{R,2}$	Maximum rate of LacI expression from plasmids (FA3)
$k_{R,3}$	Maximum rate of LacI expression from plasmids (FA4)
μ_d	Cell death rate
$T_{X,\max}$	Maximum toxicity of ACC on the cell growth
E_{tox}	Threshold concentration of ACC for toxicity on fatty acid production
d_E	Degradation rate of ACC

[1]Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Nara 8916-5, Japan. [2]Data Science Center, Nara Institute of Science and Technology, Ikoma, Nara 8916-5, Japan. [3]Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA. [4]School of Medicine, Fujita Health University, Toyoake, Aichi 470-1192, Japan. ✉email: okubo.tomoki.ou1@naist.ac.jp

d_R	Degradation rate of LacI
$K_{T,X}$	Saturation coefficient of ACC for toxicity on cell growth
$K_{T,P}$	Saturation coefficient of ACC for toxicity on fatty acid production
$K_{R,0}$	Dissociation constant of LacI
K_I	Dissociation constant of IPTG
K_M	Saturation coefficient of malonyl-CoA for transcription rate of P_{FR1}
$K_{S,P}$	Saturation coefficient of glucose for fatty acid production rate
$n_{T,X}$	Hill coefficient of ACC for toxicity on cell growth
$n_{T,P}$	Hill coefficient of ACC for toxicity on fatty acid production
n_M	Hill coefficient of malonyl-CoA for transcription rate of P_{FR1}
n_R	Hill coefficient of LacI
n_I	Hill coefficient of IPTG
H_X	Constant linking cell density measurement and X
H_P	Constant linking fatty acid measurement and P
Other variables	
T_X	Toxic effects on cell growth
T_P	Toxic effects of fatty acid production
Variables in parameter estimation	
$\boldsymbol{\theta}$	Vector of model parameters
$\hat{\boldsymbol{\theta}}$	Vector of model parameter estimates
V	Objective function for optimization
N	Number of measurement time points
M	Number of IPTG input conditions
i	Subscript for process outputs ($i = 1$ for relative cell density and $i = 2$ for fatty acid concentration)
j	Subscript for *E. coli* strains ($j = 1$ for FA2 and $j = 2$ for FA3)
k	Subscript for IPTG input conditions
l	Subscript for measurement time index
t_u	The time of addition of the inducer 2
t_N	25-H culture duration
FitPercent	The normalized root-mean-squared error expressed as a percentage

Engineered microorganisms show great promise as cell factories capable of producing a wide array of products, including foods[1], fuels[2,3], functional materials[4], and pharmaceuticals[5]. Achieving high product yields is crucial for ensuring the economic feasibility of this technology and is often accomplished by overexpressing bottleneck enzymes in target production pathways[6]. However, the excessive expression of enzymes can adversely affect microbial growth and/or viability[7]. For example, nutrients can be wasted in synthesizing excess enzymes, lowering the overall metabolite yield[8–12]. Furthermore, enzyme overexpression may deplete free ribozymes, decreasing the overall cell growth rate. In addition, the accumulation of intermediates and byproducts arising from enzymatic activity can lead to cytotoxicity[13–17]. Therefore, achieving high-yield production of the desired products necessitates stringent regulation of enzyme overexpression to maintain appropriate enzyme levels[18–21]. This can be achieved using two approaches: model-based process optimization (in silico feedforward controller) and autonomous feedback control using synthetic genetic circuits embedded in microbial cells (in-cell feedback controller)[22–24].

The in silico feedforward controller works in conjunction with genetically engineered microorganisms to change key enzyme expression levels in response to process inputs such as inducer feeds[8], temperature[25], and light[26,27], which are optimized using mathematical models before initiating the bioprocesses. This approach manages the expression of metabolic enzymes by predicting the future process state, thus maximizing the product yields. However, a challenge arises in the event of a significant process-model mismatch (PMM) between the model prediction and the actual process. When a PMM occurs, the input values predetermined using the mathematical model are suboptimal for the actual process. Model predictive control (MPC) is a method that addresses this issue by re-optimizing the process input using measurements from the ongoing process in real time[28–32]. Even though MPC provides both optimal control and robustness to PMMs, the simultaneous and frequent measurement of intracellular molecules (e.g., mRNAs, enzymes, and metabolites) using biochemical analysis techniques or fluorescent reporters presents its own set of challenges in terms of response time and reliability.

The in-cell feedback controller detects the concentrations of intracellular metabolites[33–35], extracellular nutrients[36], and cell density[9,37,38] to regulate enzyme expression accordingly, but it cannot optimize enzyme expression, unlike the in silico feedforward controller. Therefore, in silico feedforward and in-cell feedback controllers complement each other.

We previously proposed a hybrid control system (hybrid in silico/in-cell controller, HISICC) that combines an in silico feedforward controller with an in-cell feedback controller to overcome the PMM-associated difficulties inherent in model-based process optimizations[39]. This approach could be classified as prevailing hierarchical control consisting of a high-level open-loop model-based optimal controller and low-level proportional or PID feedback controllers[40]. We designed and evaluated an HISICC for resource allocation, which contained an in-cell feedback controller that autonomously triggered the overexpression of enzymes in an alcohol biosynthesis pathway concomitantly with the shutdown of a competing pathway for cell growth based on quorum sensing. The HISICC concept showed potential as a solution to PMM-associated difficulties. However, the in-cell feedback control of cell density, which relies on quorum sensing, can be substituted with electronic control using

a spectrophotometer. Hence, the dynamic adaptation of enzyme overexpression based on the in situ sensing of intracellular metabolites or intermediates remains to be demonstrated, which will be particularly useful when counterpart sensing techniques using fluorescence reporters or analytical instruments are not available due to low reliability or long turnaround time.

In this study, we aimed to demonstrate the advantage of our approach for achieving the regulated overexpression of rate-limiting enzymes using engineered *Escherichia coli* strains FA2 and FA3, originally developed by Liu et al. for fatty acid production[41], and strain FA4, which was conceptually designed based on FA2 and FA3. The overexpression of the bottleneck enzyme acetyl-CoA carboxylase (ACC) in these strains is amenable to external tuning using isopropyl-β-D-thiogalactopyranoside (IPTG). Prediction errors in the intracellular ACC concentrations constitute a critical PMM, leading to lower fatty acid yields. FA3 contains an in-cell feedback controller that indirectly detects an increase in ACC concentrations, thereby forming a HISICC when combined with an in silico feedforward controller (Fig. 1); the other two strains form no-feedback systems. First, the three strains were modeled to allow for the design of the corresponding in silico controllers. Next, the robustness of HISICC in the event of PMMs was evaluated by comparing the fatty acid yields obtained with HISICC versus its no-feedback counterparts in multiple rounds of simulation, assuming PMMs of various types and magnitudes. Our results indicate that this approach can mitigate the yield loss attributable to PMMs associated with regulating enzyme overexpression.

Results

Fatty acid production in engineered *E. coli* strains

We provided an overview of the two previously developed *E. coli* strains (FA2 and FA3[41]) and the strain conceptually designed for simulation (FA4) to help explain their mathematical modeling.

The first step in the fatty acid biosynthesis pathway in *E. coli* involves the conversion of acetyl-CoA to malonyl-CoA catalyzed by ACC, encoded by *accABCD*. This reaction is the rate-limiting step of the pathway[15]. Several studies have shown that ACC overexpression increases intracellular malonyl-CoA concentration and fatty acid yields[13,42]. However, ACC overexpression is toxic to cells and can decrease cell growth[13–15] and fatty acid production[41]. The exact molecular mechanisms underlying its toxicity remain unclear. Because suboptimal ACC concentrations reduce fatty acid yields, it is necessary to rapidly stabilize the ACC concentration to an optimal level to offer a balance between cytotoxicity and improvements in the reaction rate.

FA2 carries a genetic regulator device containing *accABCD* gene under the control of T7 RNA polymerase, allowing ACC overexpression to be externally controlled by IPTG (Fig. 2A and B). Therefore, the IPTG concentration was defined as the input variable to be optimized by the in silico feedforward controller for this strain. Unlike the other two strains, FA2 lacks mechanisms for reducing ACC overexpression during the mid-culture period. Therefore, to alleviate ACC toxicity, the in silico feedforward controller must maintain the initial ACC expression at a low level to ensure that ACC concentrations in the latter half of the culture period are not too high (Fig. 2C).

FA3 contains a genetically encoded feedback controller comprising the regulator device and a sensor device (Fig. 2D). The sensor device comprises FapR, a malonyl-CoA-responsive transcription factor, and FR1, a FapR-regulated promoter (Fig. 2E). In this design, if the ACC level becomes too high, the malonyl-CoA produced triggers the sensor device to express LacI, which in turn represses ACC overexpression. Meanwhile, the in-cell feedback controller can be externally initiated by IPTG, which alleviates the repression of ACC overexpression

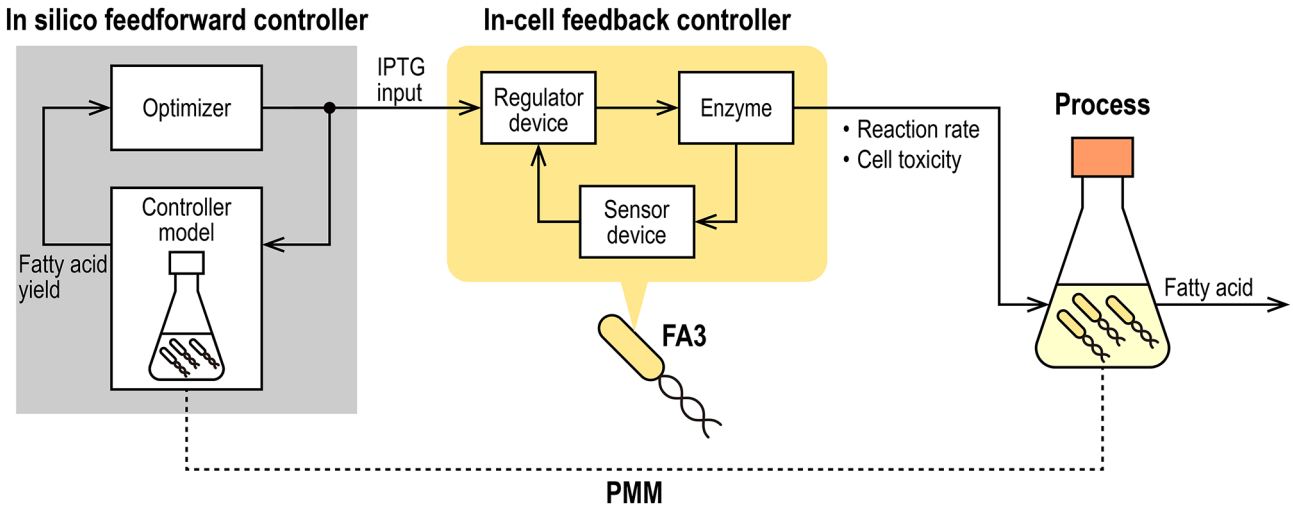

Fig. 1. Conceptual diagram of the Hybrid In Silico/In-Cell Controller (HISICC). The in silico feedforward controller optimizes the inducer input to maximize product yields by balancing the acceleration of the bottleneck reaction against the cytotoxicity caused by overexpression of the rate-limiting enzyme. Simultaneously, the in-cell feedback controller within the engineered *E. coli* strain FA3 autonomously regulates enzyme concentration in response to the optimized inducer input.

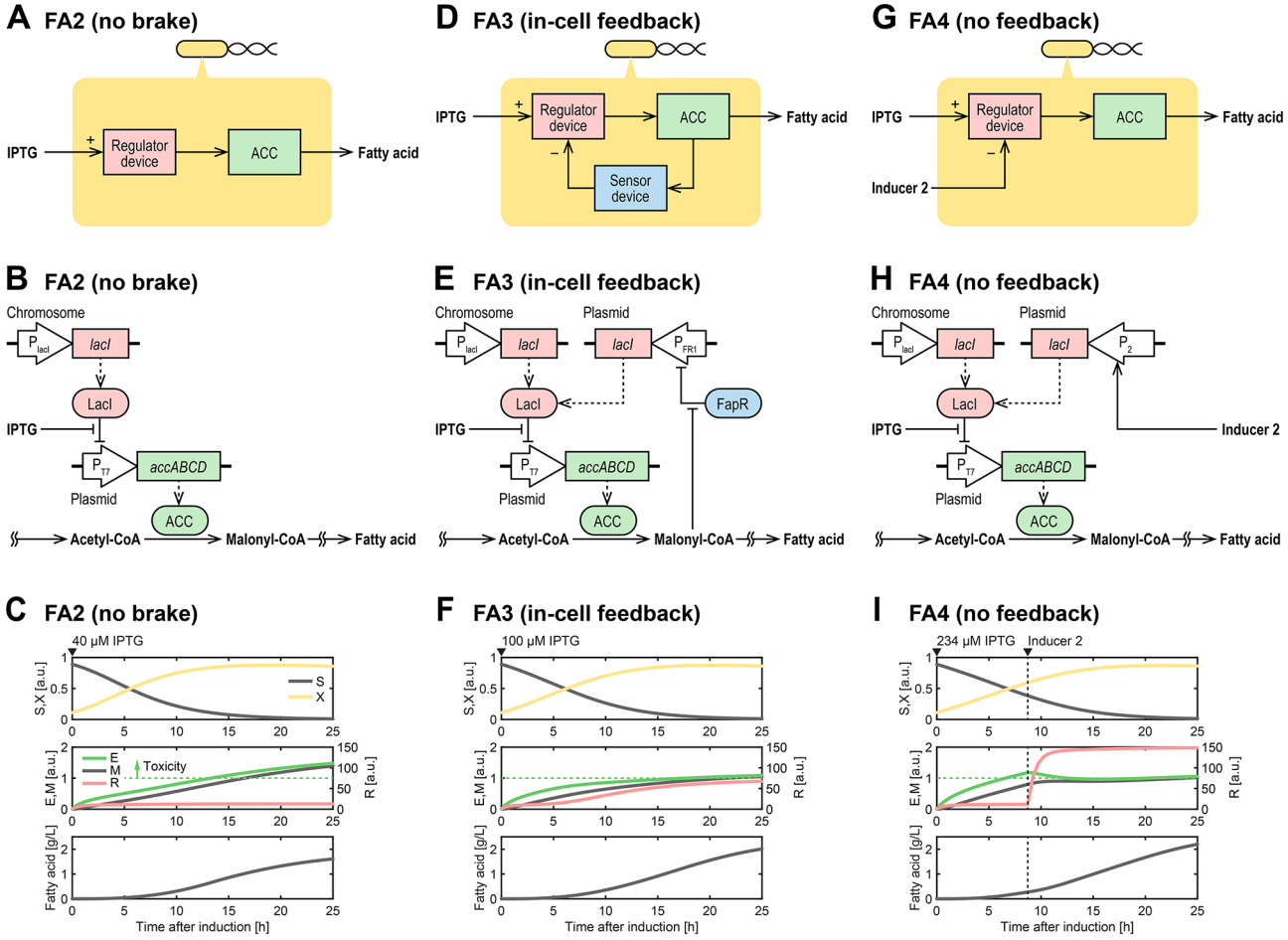

Fig. 2. Designs and functions of the engineered *Escherichia coli* strains. (Top row) Schematic representation of the control circuits in *E. coli* strains FA2 (**A**), FA3 (**D**), and FA4 (**G**). The regulator device allows tunable overexpression of ACC using IPTG, while the sensor device detects changes in ACC concentration by sensing malonyl-CoA levels. Arrows marked with plus (+) or minus (−) symbols indicate upregulation or downregulation, respectively. (Middle row) Synthetic genetic circuits of *E. coli* strains FA2 (**B**), FA3 (**E**), and FA4 (**H**). Rounded boxes represent proteins, while dotted arrows depict protein expression. Red and blue boxes correspond to the regulator and sensor devices, respectively. (Bottom row) Simulated dynamics of state variables in *E. coli* strains FA2 (**C**), FA3 (**F**), and FA4 (**I**). Variable symbols are listed in Table 1. Green dotted horizontal lines indicate the ACC toxicity threshold (E_{tox}). When ACC concentration (E) exceeds E_{tox}, toxicity reduces fatty acid secretion per unit cell density.

mediated by LacI. Overall, the in-cell feedback controller integrated into FA3 indirectly detects an increase in ACC concentration by sensing malonyl-CoA levels and responds by reducing ACC expression during the mid-culture period. Therefore, the in silico feedforward controller coupled to this strain can employ higher IPTG concentrations to induce cells, leading to a higher initial expression of ACC than the in silico controller coupled to FA2. This shortens the response time for achieving the optimal ACC concentration (Fig. 2F). In fact, a 27% higher yield of fatty acids was obtained with FA3 compared to FA2 under identical culture conditions (except for the IPTG concentration used for induction) in a previous study[41].

Because FA3 has an in-cell feedback controller, while FA2 does not, it forms the HISICC in combination with in silico feedforward controllers; FA2 forms the no-feedback system. This no-feedback system was compared to HISICC to demonstrate its robustness to PMMs. However, FA2 has no braking circuit to reduce ACC overexpression during bioprocessing. This makes the no-feedback system with FA2 an inadequate counterpart to HISICC for demonstration. Therefore, based on these strains, we conceptually designed another strain, FA4, which has a genetic circuit to slow ACC overexpression in response to a second external input instead of the output of the sensor device (Fig. 2G). The second input channel can be a chemical inducer or a light source. Here, we assumed an idealized inducer/promoter pair, inducer 2/promoter 2, which was orthogonal to induction with IPTG. Inducer 2 induces the expression of an additional *lacI* downstream of promoter 2 (Fig. 2H and I). Both IPTG and inducer 2 were defined as the optimizable variables for the bioprocess using this strain.

Mathematical modeling

FA2 model

This model forms the basis for modeling the other two strains, which are extensions of this model, with the addition of a term representing LacI expressed from plasmids to a differential equation for LacI dynamics. This model consists of two differential equations describing substrate consumption and cell growth and four differential equations describing the normalized concentrations of important compounds in the cell.

$$\mu = k_X S\left(1 - T_X\right) \tag{1}$$

$$\frac{dS}{dt} = -\mu X \tag{2}$$

$$\frac{dX}{dt} = \mu X - \mu_d X \tag{3}$$

$$\frac{dE}{dt} = k_E \frac{K_{R,0}^{n_R}}{K_{R,0}^{n_R} + \left(\frac{R}{1 + \left(\frac{I}{K_I}\right)^{n_I}}\right)^{n_L}} - \left(d_E + \mu\right) E \tag{4}$$

$$\frac{dM}{dt} = k_M E - k_P M - \mu M \tag{5}$$

$$\frac{dR}{dt} = k_{R,1} - \left(d_R + \mu\right) R \tag{6}$$

$$\frac{dP}{dt} = k_P M X \frac{S}{K_{S,P} + S}\left(1 - T_P\right) \tag{7}$$

The descriptions and initial values of the six state variables are summarized in Table 1, while the model parameters are listed in Table S1.

The equations for cell growth were derived based on Monod's model[43] with some modifications (Eqs. 1–3). First, the saturation of the specific growth rate μ at high substrate concentrations was ignored to simplify the model. Second, cell death was assumed to occur at a constant rate, μ_d. T_X represents a deceleration in cell growth owing to ACC cytotoxicity, as described below.

Equations 4–6 describe the dynamics of the intracellular concentrations of ACC (E), malonyl-CoA (M), and LacI (R). These three compounds were assumed to be diluted during cell growth. The concentrations of ACC and LacI were also assumed to decrease owing to protein degradation. The LacI degradation rate (d_R) was adopted from Basu et al.[44] LacI expression is regulated by the T7 promoter and is amenable to tuning using various IPTG concentrations. The mathematical expression for promoter output as a function of both LacI and IPTG concentrations was adopted from Gardner et al.[45] (Eq. 4). The saturation constant $K_{R,0}$ was defined as 1, indicating that R was normalized with respect to $K_{R,0}$. The production and consumption rates of malonyl-CoA were assumed to be proportional to the concentrations of ACC and malonyl-CoA, respectively (Eq. 5). By defining the proportionality constants k_M and k_P as equal, we equalized the ranges of the non-dimensionalized state variables M and E. Finally, LacI was assumed to be expressed at a constant rate ($k_{R,1}$) from the chromosomal *lacI* gene (Eq. 6). The model does not account for time delays due to gene transcription because the typical time scale for gene transcription in *E. coli* is less than 1 minute[46], which is sufficiently short and negligible compared to the time scale for ACC regulation.

Equation 7 describes the fatty acid concentration in the culture medium. The volumetric fatty acid production rate was assumed to be proportional to both the cell density X and the specific malonyl-CoA consumption rate $k_P M$. The saturation constant $K_{S,P}$ was introduced to reproduce the slowdown in fatty acid production per unit cell density due to the depletion of substrates in the culture medium. T_P represents the slowdown of fatty acid production due to the cytotoxicity associated with ACC overexpression, as discussed below.

Symbol	Initial value	Unit	Description
S	$1 - X\left(0\right)$	Dimensionless	Extracellular glucose concentration
X	Estimated	Dimensionless	Normalized cell density
E	0	Dimensionless	Intracellular ACC concentration
M	0	Dimensionless	Intracellular malonyl-CoA concentration
R	Estimated	Dimensionless	Intracellular LacI concentration
P	0	Dimensionless	Extracellular fatty acid concentration
I	—	μM	Extracellular IPTG concentration (input)
u	0	Dimensionless	Inducer 2 input for FA4, u is either 0 or 1 (unit step function)

Table 1. State and input variables for the models.

This model assumes that excessive ACC concentrations affect the specific rates of cell growth and fatty acid production.

$$T_X = T_{X,\max}\frac{E^{n_{T,X}}}{K_{T,X}^{n_{T,X}} + E^{n_{T,X}}} \tag{8}$$

$$T_P = \begin{cases} 0 & (E < E_{\text{tox}}) \\ \dfrac{K_{T,P}^{n_{T,P}}}{K_{T,P}^{n_{E,P}} + (E - E_{\text{tox}})^{n_{T,P}}} & (E \geq E_{\text{tox}}) \end{cases} \tag{9}$$

T_X represents the toxic effects on cell growth; an increase in E leads to an increase in T_X and a decrease in the specific cell growth rate μ (Eqs. 1 and 8). T_P represents the toxic effects on fatty acid production. We assumed that when E exceeded the threshold E_{tox}, the value of T_P increased, and fatty acid production slowed down (Eqs. 7 and 9). The value of E_{tox} was defined as 1, indicating that E was normalized with respect to E_{tox}. As noted earlier, the detailed molecular mechanisms underlying the cytotoxicity of ACC overexpression remain unclear. Besides the metabolic burden of excessive ACC biosynthesis, another possible cause of cytotoxicity could be the negative effects of excessive malonyl-CoA synthesis. However, we have abstracted the cytotoxicity of ACC overexpression by defining T_X and T_P as functions of E for the sake of model simplicity.

Equations 10 and 11 are the observation equations that link the model simulations to the experimental measurements. y_1 and y_2 represent the relative cell density calculated based on the optical density (OD_{600}) and fatty acid concentration, respectively. H_X and H_P are the proportionality constants.

$$y_1 = H_X X \tag{10}$$

$$y_2 = H_P P \tag{11}$$

FA3 model

FA3 contains an in-cell feedback controller that senses an increase in malonyl-CoA concentration and promotes LacI overexpression from plasmids to reduce ACC expression. A mathematical representation of this feedback was generated by extending the FA2 model; we added a term representing *lacI* expression from plasmids to the differential equation describing LacI dynamics in Eq. 6.

$$\frac{dR}{dt} = k_{R,1} + k_{R,2}\frac{M^{n_M}}{K_M^{n_M} + M^{n_M}} - (d_R + \mu)\,R \tag{6'}$$

The response of the FR1 promoter to malonyl-CoA was approximated using Hill's equation. The Hill coefficient n_M was defined as 4 following the modeling of another in-cell feedback controller for fatty acid production that employed the same FapR-based malonyl-CoA sensor as FA3[47,48]. To simplify the model, the concentration of FapR was assumed to be constant. The other equations and parameter values were shared with the FA2 model.

FA4 model

FA4 contains a brake circuit designed to decelerate ACC overexpression by promoting the overexpression of LacI from plasmids in response to inducer 2. A mathematical representation of this brake function was generated by adding a term representing *lacI* expression from the plasmids to Eq. 6 in the FA2 model.

$$\frac{dR}{dt} = k_{R,1} + k_{R,3}u - (d_R + \mu)\,R \tag{6''}$$

In this equation, u represents the normalized concentration of inducer 2. We then defined $u = 1$ as the concentration of inducer 2 at which saturation of the promoter 2 response is achieved. The equations and parameter values were shared with the FA2 model, except for the maximum lacI expression rate of promoter 2, $k_{R,3}$.

$k_{R,3}$ was arbitrarily defined to be equal to $k_{R,2}$ of the FR1 promoter in FA3. Since FA4 was idealized and conceptually designed at the device level[49] for comparison with FA3, there are many promoter candidates for promoter 2, including light-inducible promoters, in the part-level design of this strain[49] (Fig. 2H). Depending on the choice of promoter, the value of $k_{R,3}$ may be reduced, resulting in a weaker braking effect on ACC overexpression. In addition, chemical inducers may have unintended effects on the overall dynamics of bioprocesses through hierarchical carbon utilization. Consequently, a no-feedback system with FA4 may result in lower fatty acid yields in actual experiments compared to the simulations in this study. However, our argument still stands that a HISICC with FA3 is more robust to PMMs associated with the key enzyme than the no-feedback system with FA4.

Model simulation and validation

The FA2 and FA3 models were trained using experimental data from fatty acid-producing cultures from a previous study[41]. Experimental data from only one batch of FA3 were available for parameter estimation. This lack of experimental data for FA3 was overcome by exploiting the fact that many genetic circuits were common to FA2 and FA3. These two strains were modeled in such a way that most of the equations and parameters were shared, and trained collectively, using experimental data from both strains together as a single training data set.

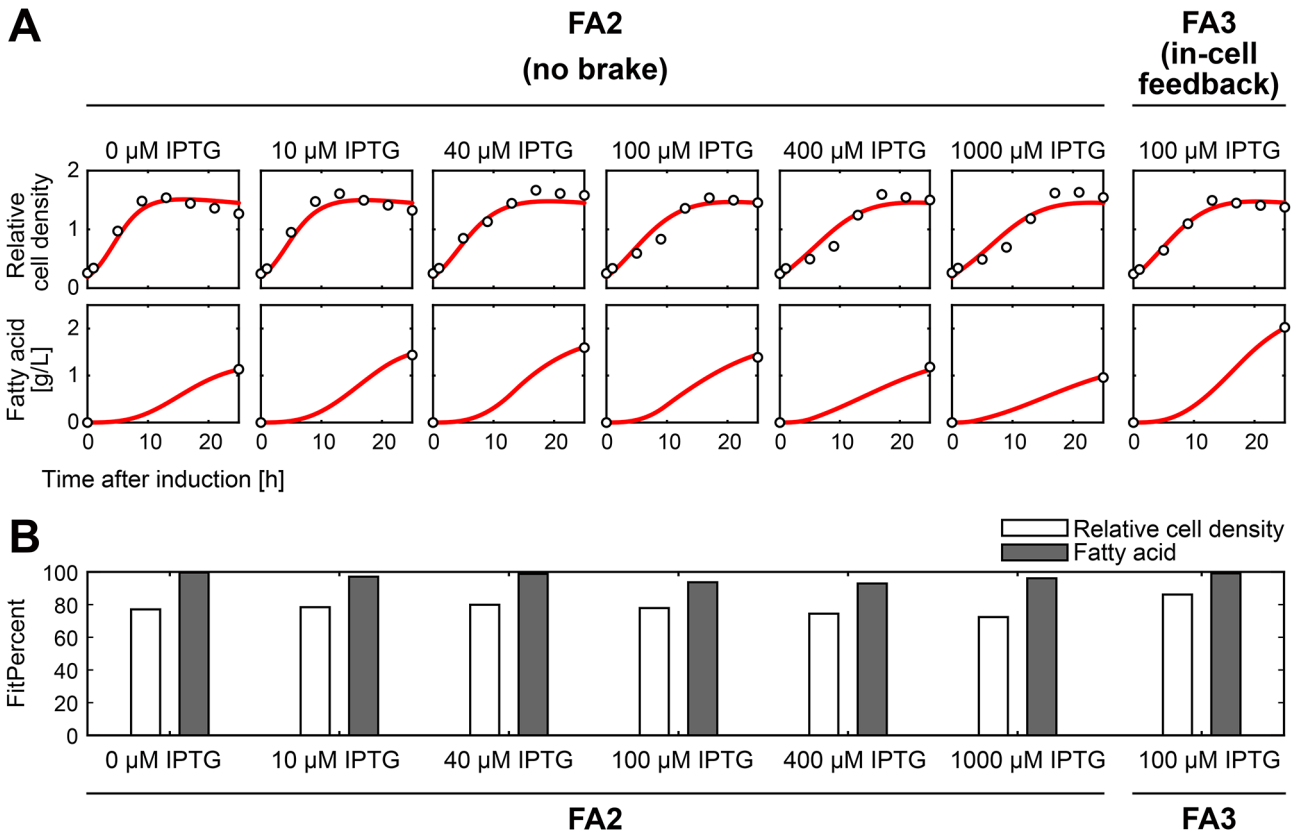

Fig. 3. Simulation of fatty acid production in *E. coli* strains FA2 and FA3 at various IPTG concentrations. (**A**) Simulated trajectories (red lines) compared with experimental data (white dots). (**B**) The normalized root-mean-squared error expressed as a percentage (FitPercent) is used to evaluate the model fit.

Namely, for each parameter common to both models, a single value was estimated so that both models could simultaneously fit the corresponding experimental data. Trained on all available experimental data, both models showed a close fit to the experimental data and consequently reproduced the improvement in fatty acid yields between FA2 and FA3, even though the FA3 model has only two unique estimated parameters ($k_{R,2}$ and K_M) (Fig. 3A and Table S1). FitPercent, the normalized root-mean-squared error expressed as a percentage[50], was used as an indicator of the fit of the model to the data. A FitPercent > 70% was obtained for all data sets (Fig. 3B). This suggests that these models successfully captured the dynamics of cell growth and fatty acid production by the two strains in response to different IPTG inputs. Furthermore, the holdout validation method (see the Methods section) was used to ensure that the trained models did not overfit the training dataset (Fig. 4). A FitPercent > 70% was obtained for all validation rounds, indicating that the model did not overfit the training data and had a good generalization performance within the range of IPTG concentrations applied in the experiments.

Two points should be noted regarding the experimental data. First, the small size of the training data set resulted in long confidence intervals relative to the estimated values for the parameters (Figure S1). However, in general, the predictive performance of the trained model is independent of the length of the confidence interval of the model parameters, and thus the demonstration of the HISICC. Another point to note is that fatty acid concentrations were only measured at the end of the culture period (Fig. 3). Therefore, our trained models may not accurately reproduce the detailed dynamics of fatty acid accumulation in culture media. However, our models predicted the final fatty acid concentrations resulting from different inducer inputs with acceptable accuracy.

Model-based input optimization

We subsequently sought to demonstrate the performance of the in silico feedforward controllers in the HISICC (FA3) or no-feedback systems (FA2 or FA4) by optimizing the IPTG input to obtain the maximum fatty acid yields at a fixed culture duration (25 h). The FA2 and FA3 models were trained using all the experimental datasets before input optimization (Fig. 3, Table S1). The FA4 model shares all parameters with the other two models except $k_{R,3}$, the maximum expression rate of *lacI* from plasmids. Since we assume that $k_{R,3}$ is equal to $k_{R,2}$, the counterpart parameter in the FA3 model, the FA4 model requires no training. The in silico feedforward controllers for FA2 and FA3 optimized the IPTG concentrations (Fig. 5A and B). In contrast, the controller for FA4 simultaneously optimized both the IPTG concentration and the time of addition of inducer 2 (Fig. 5C). To design FA4 as an idealized strain to compare with FA3, we assumed that LacI expression is induced at a maximum rate by a sufficient concentration of inducer 2. Therefore, the concentration of inducer 2 was not

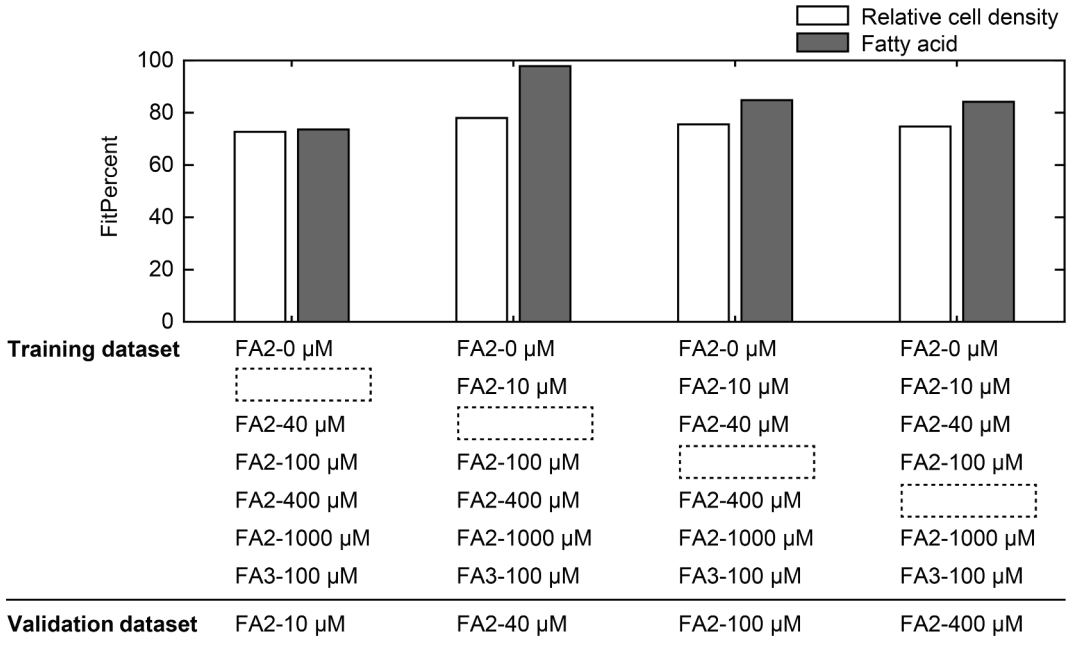

Fig. 4. Holdout validation of FA2 and FA3 models. Model validation was performed separately for cell density and fatty acid concentration using FitPercent as a model fit indicator. In each validation round, data from one batch of FA2 was used as the validation set, while data from the remaining FA2 and FA3 batches served as the training dataset.

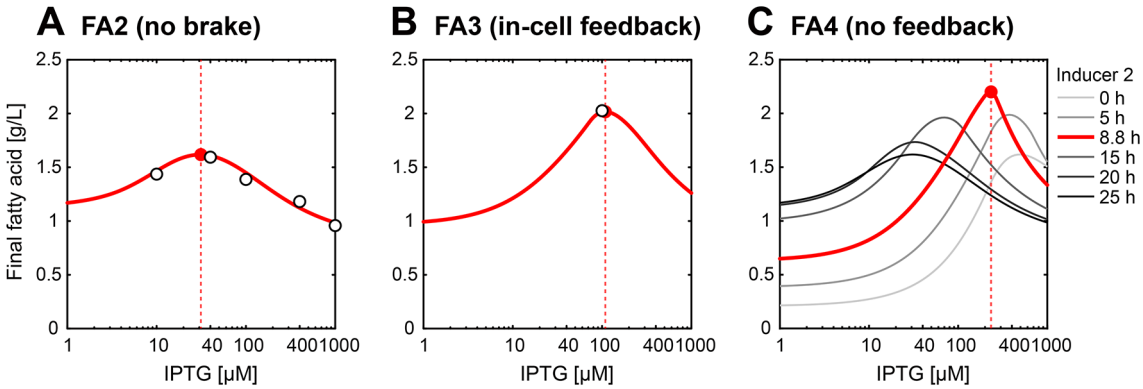

Fig. 5. Model-based optimization of IPTG input for maximizing fatty acid yields. White dots represent experimental data, and red lines represent simulations. Red dots and vertical dashed lines indicate optimized IPTG concentrations for *E. coli* strains FA2 (**A**), FA3 (**B**), and FA4 (**C**). In (**C**), gray and red lines show simulations with different addition times of inducer 2 and a simulation with the optimal addition time, respectively.

optimized. The feasible input range was defined as 0–1,000 μM for IPTG concentration and 0–25 h for the time of addition of inducer 2. In addition, comprehensive model simulations were performed to visualize the overall distribution of fatty acid yields over the feasible input range. The predicted fatty acid yields for all strains exhibited a convex-upward function with a distinct peak, and the input values optimized by the in silico controllers were consistent with this peak. This observation suggests that in silico feedforward controllers contribute to product yield improvement through input optimization. In addition, the optimal IPTG concentration for FA2 was the lowest among the three strains, consistent with the inability of FA2 to reduce ACC expression during the mid-culture period; FA2 initially requires the in silico controller to maintain a lower initial IPTG concentration to maintain low ACC expression levels, unlike the other two strains. This suggests that the models successfully reproduced the trade-off inherent in ACC overexpression for fatty acid production.

Controller performance against PMMs

The robustness of HISICC against PMMs was evaluated by calculating the fatty acid yields of the three strains through five multi-round simulations (Fig. 6A-E). Each multi-round simulation focused on one of the five

parameters, k_E, d_E, $k_{R,2}$, d_R, and k_M, related to the proteins that constitute the in-cell feedback controller, ACC, LacI, and FapR. For all simulations, the parameters in the controller models for FA2 and FA3 were estimated using all experimental datasets (Fig. 3, Table S1), which were also employed in the controller model for FA4. Thus, the inducer input values employed in these simulations were identical to the optimized values shown in Fig. 5 (dotted vertical lines). Each simulation round introduced a PMM of different magnitudes; for the simulation focused on k_E (Fig. 6A), for example, ACC expression in the strains was assumed to be faster or slower than the values predicted by the corresponding in silico feedforward controllers. Specifically, we defined various values for k_E, which represents the maximum rate of ACC expression, for each round to calculate the dynamics of the process, which are denoted as k_E^* to distinguish them from the corresponding value k_E in the controller model. Notably, k_E^* represents an intrinsic property of cells that is difficult to manipulate in cell culture-based experiments. The simulation round where $k_E^*/k_E = 1$ corresponds to the situation in which no PMM was introduced (Fig. 6A, dotted vertical line). That is, this round represents an ideal situation in which the in silico feedforward controllers perfectly predict the ACC dynamics. As with k_E, we introduced PMMs with the other four parameters, d_E, $k_{R,2}$, d_R, and k_M, by simulating the process with parameter values deliberately shifted from those in the controller model (denoted as d_E^*, $k_{R,2}^*$, d_R^*, and k_M^*,), as shown in Fig. 6B-E.

PMMs associated with ACC expression and degradation

First, we describe the results of the simulation focusing on the maximum ACC expression rate k_E (Fig. 6A). Drifts in k_E can occur in real bioprocesses, for example, due to fluctuations in plasmid copy number before and after mathematical modeling. The highest fatty acid yields were obtained in the no-PMM round for all three strains. Lower yields were obtained when ACC expression was slower than predicted ($k_E^*/k_E < 1$) because sufficient ACC accumulation was not attained during the 25-h culture period. Similarly, a faster-than-predicted expression of ACC ($k_E^*/k_E > 1$) also resulted in lower fatty acid yields because of the cytotoxicity associated with excessive ACC.

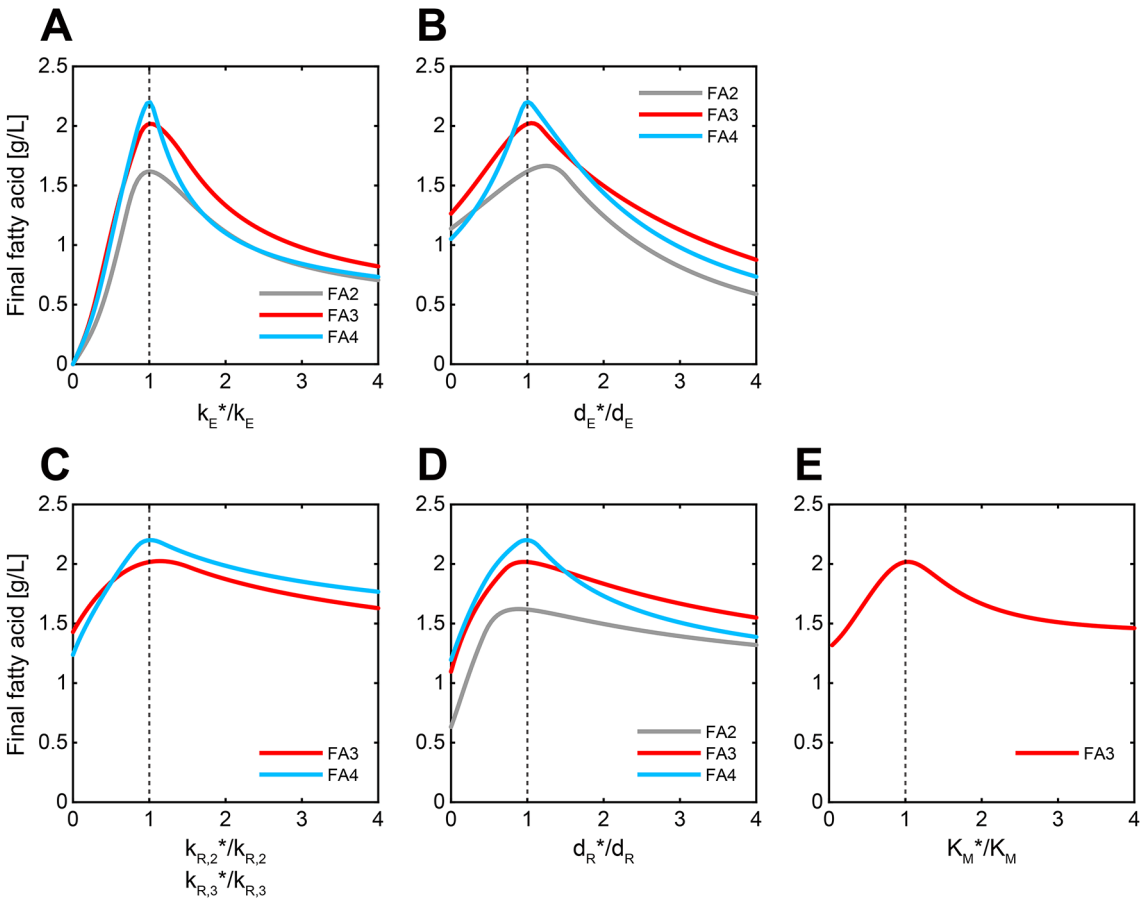

Fig. 6. Multi-round simulations examining fatty acid yields under different Process Model Mismatches (PMMs). The simulations explore the impact of PMMs in model parameters k_E (**A**), d_E (**B**), $k_{R,2}$ or $k_{R,3}$ (**C**), d_R (**D**), and K_M (**E**). In each simulation round, p^* represents the parameter value used for process dynamics calculations, while p refers to the estimated parameter value from Table S1 used in the controller model (k_E, d_E , $k_{R,2}$, $k_{R,3}$, d_R, or K_M). Inducer inputs (IPTG concentrations for FA2 and FA3, and both IPTG concentration and inducer 2 addition time for FA4) were set to the optimized values from Fig. 5. The vertical dashed line denotes a simulation round with no PMM, where the same parameter value was used in both the controller and process models.

The lowest fatty acid yield was obtained with FA2, regardless of the value of k_E^*. This is attributable to the inability of FA2 to decelerate ACC expression during the mid-culture period, unlike the other two strains. The consequent accumulation of excess ACC impairs fatty acid production in the later stages of culture in this strain.

Differences were observed between FA3 and FA4 with respect to the interference by PMMs, even though both strains were designed to decelerate ACC overexpression. Under slow ACC expression ($k_E^*/k_E < 1$), little difference was observed in the fatty acid yields between the two strains. However, when ACC expression was comparable to that predicted by the controller ($k_E^*/k_E \approx 1$), FA4 produced slightly higher quantities of fatty acids than FA3; this is because the in silico controller for FA4 could induce ACC overexpression with a higher concentration of IPTG and subsequently resort to abrupt braking enabled by the input of inducer 2 (Fig. 2I). Such abrupt braking is impossible for FA3 because the in-cell feedback controller embedded in this strain gradually responds to an increase in malonyl-CoA concentration (Fig. 2F). Notably, under rapid ACC expression ($k_E^*/k_E > 1$), FA4 produced lower yields of fatty acids than FA3. The reasons for this difference in the robustness of the PMM are as follows. First, the in silico feedforward controller for FA4 optimized the addition time of inducer 2 regardless of the actual ACC concentration, which increased faster than predicted by the controller. Consequently, inducer 2 was added later than the truly optimal timing, allowing excess ACC accumulation and reducing fatty acid yields. In contrast, FA3 contains an in-cell feedback controller that allows an autonomous change in the timing and intensity of brake application on ACC overexpression following the detected increase in malonyl-CoA concentration (reflecting ACC activity). Thus, a mitigated loss of fatty acid yields was obtained for FA3 compared to FA4. Additionally, under conditions of a very rapid rate of ACC expression ($k_E^*/k_E > 1.5$), the braking effect triggered by inducer 2 was almost nonexistent, resulting in low fatty acid yields nearly equivalent to that of FA2, which lacks a brake circuit. Similar results were obtained in the simulation that focused on the degradation rate of ACC, d_E (Fig. 6B). Fluctuations in d_E can occur, for example, when different proteases are activated during the transition from aerobic to anaerobic metabolism. When the degradation rate of ACC was either faster ($d_E^*/d_E > 1$) or slower ($d_E^*/d_E < 1$) than predicted by the controller, FA3 suppressed the decrease in fatty acid yields to a lesser extent than FA4. This is because, as in the case of k_E , the in-cell controller implemented in FA3 adjusts the strength and timing of the brake on ACC expression in real time depending on the actual level of ACC. These results indicate that the in-cell feedback controller within the HISICC can support the in silico feedforward controller to handle PMMs associated with the key enzyme.

PMMs associated with other proteins

The decrease in fatty acid yields due to PMMs associated with LacI and FapR was less severe than that due to PMMs associated with ACC (Fig. 6C-E). For LacI, there was almost no difference in the decrease in fatty acid yields between FA3 and FA4, when both simulations focused on the expression rate from plasmids, $k_{R,2}$ (Fig. 6C) or the degradation rate, d_R (Fig. 6D).

Discussion

This study extended the HISICC concept we previously developed to address the PMM problem[39] to a system that regulates the overexpression of a key enzyme based on feedback from an intracellular metabolite biosensor. The engineered *E. coli* strains FA2 and FA3, developed by Liu et al. for fatty acid production[41], were employed for a proof-of-concept study. Unlike FA2, FA3 has an in-cell feedback controller that indirectly detects an increase in the concentration of the rate-limiting enzyme, ACC, by sensing malonyl-CoA produced by its enzymatic activity and then applying a brake on the expression of the enzyme. Therefore, only FA3 forms an HISICC in combination with the in silico feedforward controller. We hypothesized that the in-cell feedback controller allows an HISICC to adjust the timing and strength of brake application according to the detected enzyme concentration, which, in turn, allows a mitigation of fatty acid yields loss attributable to the PMM associated with the rate of enzyme expression. As a first step in proving this hypothesis, we constructed mathematical models and designed in silico feedforward controllers for the strains. The models captured the dynamics of cell growth and fatty acid production in response to various inducer concentrations and exhibited good predictive performances against experimental data during holdout validation. Next, we evaluated the robustness of the HISICC against PMMs using a multi-round simulation by comparing FA2, FA3, and FA4. FA4 was conceptualized based on the circuit designs of FA2 and FA3 and contained a brake circuit that decelerated the enzyme overexpression in response to another inducer. Comparing the fatty acid yields from the three strains revealed that when the rate of enzyme overexpression was higher than that predicted by the in silico controllers, the minimum fatty acid yield loss was obtained with FA3, consistent with our hypothesis. This observation demonstrates the potential of the extended HISICC with intracellular biosensing as a solution to the PMM problem.

The robustness of the HISICC to PMMs was evaluated by simulations rather than by culture experiments as in previous modeling studies on in-cell feedback controllers[35,51], because the actual cell dynamics, which are only partially manipulable in real culture experiments, can be arbitrarily defined in simulations. As a result, the robustness of HISICCs was extensively and quantitatively tested in a controlled manner. Of course, additional experimental evaluations of robustness would further confirm the feasibility of our approach. Specifically, PMMs could be introduced into real culture experiments with some degree of control by deliberately shifting the parameter values in the controller model from those estimated in model training. In this case, the results of the exhaustive simulations presented in this study are helpful in determining the appropriate range of parameter shift for assessing controller robustness.

Our simulations do not demonstrate that the proposed approach can address all possible PMMs that may occur in bioprocesses. While the HISICC is more robust than no-feedback systems to PMMs related to expression or degradation rates of ACC, PMMs related to LacI and FapR, which constitute the in-cell feedback controller, may negate this advantage. However, according to simulations, the effect of PMMs for LacI and FapR is small compared to that for ACC (Fig. 6). In addition, there is little difference in robustness between the HISICC and

no-feedback counterparts with respect to the PMMs associated with the expression or degradation rates of LacI (Figs. 6C and D). Furthermore, the metabolic burden of LacI and FapR is lower, and their expression rates are less likely to drift due to plasmid instability than ACC, which is overexpressed to the point of cytotoxicity. In other words, there is a reasonable chance that the PMMs associated with ACC will be the main factor affecting product yields. Therefore, our simulations suggest that our approach could be a potential solution to the PMM problem in practical bioprocesses.

The in-cell controller used in this study is simple in structure and crude in operation; as described in the Results section, due to the gradual response of FapR to malonyl-CoA, it takes time for the HISICC to stabilize the ACC level at the desired level, resulting in a lower product yield than the no-feedback counterpart in the absence of PMMs. Therefore, it is expected that an advanced in-cell controller capable of differential or second-order differential control[52,53] can be used as the in-cell controller for HISICC to shorten the response time and improve performance.

In this study, in silico feedforward controllers were designed to optimize only one or two inputs: the IPTG concentration for FA2 and FA3, and the IPTG concentration and time of addition of inducer 2 for FA4. These inputs can also be optimized by an exhaustive experimental search, instead of in silico controllers. However, there are two directions for extending the application of our control framework where in silico controllers would play a more important role: multiple enzyme regulation and time-varying process inputs. First, if multiple rate-limiting enzymes need to be regulated, as in another *E. coli* strain engineered for fatty acid production by Xu et al.[47], the cost of the experimentation for the optimal input will likely be high. Second, IPTG has the problem of irreversibility; in batch processes, the added IPTG remains unconsumed and continues to activate IPTG-inducible promoters. In contrast, inducing gene expression using light[11,26,27], temperature[25], or consumable inducers such as glucose[54], lactose[55], and arabinose[56] is reversible and allows time-varying inputs to the in-cell feedback controller. Although these reversible input channels enable the effective and dynamic regulation of the expression of key enzymes, they require determining the optimal input values for each control interval, making the experimental approach very challenging. Thus, the HISICC concept extended in this study is expected to provide a practical solution to PMM-associated issues in more complicated bioprocesses involving the manipulation of multiple enzymes or time-varying process inputs, the demonstration of which is an important topic for future research.

Methods

Experimental data

Experimental data from the fatty acid production cultures used to train and validate the FA2 and FA3 models were obtained from Liu et al.[41] (Figures S2 and S4 in the Supplementary Information appended to the original article[41]). Regarding the culture conditions, seed cultures of both strains were grown overnight at 37 °C on a rotary shaker at 220 rpm in Luria–Bertani medium supplemented with the appropriate antibiotics (50 mg/L ampicillin, 50 mg/L kanamycin, and 30 mg/L chloramphenicol). The seed cultures were then transferred to a minimal medium (M9 minimal medium supplemented with 2% glucose, 75 mM MOPS, 2 mM $MgSO_4$, 1 mg/L thiamine, 10 μM $FeSO_4$, 0.1 mM $CaCl_2$, and the following micronutrients: 3 μM $[NH_4]_6Mo_7O_{24}$, 0.4 mM boric acid, 30 μM $CoCl_2$, 15 μM $CuSO_4$, 80 μM $MnCl_2$, and 10 μM $ZnSO_4$) at a concentration of 2% (v/v) and grown overnight for adaptation; the M9 media was supplemented with the same antibiotics as before. Fatty acid production cultures were initiated from the overnight culture for adaptation with an initial OD_{600} of 0.08 and grown in fresh minimal medium supplemented with 2% glucose and 0.01% arabinose. When an OD_{600} of 0.6 was attained, the strains were induced with 200 nM anhydrotetracycline and different concentrations of IPTG. The relative cell density (arbitrary units) was defined previously[41] using OD_{600} values recorded every 1,000 s. Samples for fatty acid quantification were collected 25 h after induction.

Parameter estimation

MATLAB/Simulink 2022a was used for model construction and simulation. Simulink Design Optimization was used to estimate the model parameters. To simulate the models, Simulink automatically selected an appropriate variable step solver. As described in the previous subsection, only the relative cell density and fatty acid concentration were used as training data. We downsampled the relative cell density into datasets at 4-h intervals. For fatty acid concentrations, only endpoint values (25 h) were available. The parameter values were chosen to minimize the sum of the squared errors between the model predictions and the training data, as shown in Eqs. 12 and 13. The errors were normalized to the maximum values of the measurements in the same culture. In Eqs. 12 and 13, V represents the objective function for optimization. The vectors $\boldsymbol{\theta}$ and $\widehat{\boldsymbol{\theta}}$ represent the model parameters and their estimated values, respectively. The parameters y and $\hat{y}$ represent the measured and predicted process outputs, respectively. I represents the IPTG input. The subscripts i, j, k, and l represent the process output ($i = 1$ for relative cell density and $i = 2$ for fatty acid concentration), strain ($j = 1$ for FA2 and $j = 2$ for FA3), IPTG input, and measurement time index, respectively. N_i and M_j represent the number of measurement time points for process output i and the IPTG input levels for strain j, respectively. The compositions of the datasets and the corresponding indices are summarized in Tables S2 and S3.

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left[V\left(\boldsymbol{\theta}\right) \right] \tag{12}$$

$$V\left(\boldsymbol{\theta}\right) = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{1}{M_j} \sum_{k=1}^{M_j} \frac{1}{N_i} \sum_{l=1}^{N_i} \left(\frac{y_{i,j,k,l} - \hat{y}_{i,j}\left(t_l; I_{j,k} | \boldsymbol{\theta}\right)}{\max\limits_{l'} \left[y_{i,j,k,l'} \right]} \right)^2 \tag{13}$$

The *lsqnonlin* command of Simulink Design Optimization was used for optimization. The trust region method was selected as the optimization algorithm for this command. Initial parameter estimates were approximated using general knowledge of the physical properties of biomolecules, as well as *E. coli* growth and gene expression (Table S1). A scaling factor equal to the initial estimate was specified for each parameter to prevent large absolute values from disproportionately affecting the overall parameter estimation.

Model validation

The holdout method was used to validate whether the models could predict the experimental results for various IPTG concentrations without overlearning. For each round of validation, one batch of FA2 was selected from six batches, excluding the batches induced at the highest (1,000 μM) or lowest (0 μM) IPTG concentrations; the data obtained from this batch were defined as the validation dataset (Fig. 4). The data obtained from the remaining five batches of FA2 and one batch of FA3 were collectively defined as the training datasets. The normalized root-mean-squared error, expressed as a percentage (FitPercent[50]), was employed to measure how well the model response fitted the validation dataset. For each validation round, the FitPercent was calculated for the cell density or fatty acid concentration as follows:

$$\mathrm{FitPercent}_{i,k} = 1 - \frac{\sqrt{\sum_{l=1}^{N_i}\left(y_{i,1,k,l} - \hat{y}_{i,1}(t_l; I_{1,k}|\hat{\boldsymbol{\theta}}_k)\right)^2}}{\sqrt{\sum_{l=1}^{N_i}\left(y_{i,1,k,l} - \frac{1}{N_i}\sum_{l'=1}^{N_i} y_{i,1,k,l'}\right)^2}}$$

$$\hat{\boldsymbol{\theta}}_k = \arg\min_{\boldsymbol{\theta}} \left[V_k(\boldsymbol{\theta})\right]$$

$$V_k(\boldsymbol{\theta}) = \sum_{i=1}^{2}\left[\frac{1}{M_1}\sum_{k'\in D_{\mathrm{train},k}}\frac{1}{N_i}\sum_{l=1}^{N_i}\left(\frac{y_{i,1,k',l} - \hat{y}_{i,1}(t_l; I_{1,k'}|\boldsymbol{\theta})}{\max_{l'}\left[y_{i,1,k',l'}\right]}\right)^2 + \frac{1}{M_2}\sum_{k'=1}^{M_2}\frac{1}{N_i}\sum_{l=1}^{N_i}\left(\frac{y_{i,2,k',l} - \hat{y}_{i,2}(t_l; I_{2,k'}|\boldsymbol{\theta})}{\max_{l'}\left[y_{i,2,k',l'}\right]}\right)^2\right]$$

Subscript k represents the batch of strain FA2 selected for the validation dataset. The subscript set $D_{\mathrm{train},k}$ represents the set of FA2 batches selected for the training dataset corresponding to k.

Model-based input optimization

Simulink Design Optimization was used to optimize the IPTG input. The optimal values I_{opt} for FA2 and FA3 were chosen to maximize the fatty acid yields obtained at the end of the 25-h culture period, as follows:

$$I_{\mathrm{opt}} = \arg\min_{I}\left[-\hat{y}_{2,j}(t_N; I)\right] \text{ for } j = 1, 2.$$

There is another decision variable for FA4: the time of addition of the inducer 2, denoted as t_u. The optimal values I_{opt} and $t_{u,\mathrm{opt}}$ for FA4 were selected as follows:

$$(I_{\mathrm{opt}}, t_{u,\mathrm{opt}}) = \arg\min_{(I,t_u)}\left[-\hat{y}_{2,j}(t_N; I, t_u)\right] \text{ for } j = 3.$$

The subscript j represents the strain index ($j = 1$ for FA2, $j = 2$ for FA3, and $j = 3$ for FA4). The culture duration t_N in the simulation was defined as 25 h, as in the experiment performed. As with the parameter estimation, the *lsqnonlin* command was used for optimization and the trust region method was selected as the algorithm.

Data availability

The datasets and computer code used in this study are available on GitHub (https://github.com/TomokiOhkubo/HISICC2).

Received: 21 June 2024; Accepted: 9 October 2024
Published online: 18 November 2024

References

1. Kallscheuer, N. Engineered microorganisms for the production of food additives approved by the European Union—a systematic analysis. *Front. Microbiol.***9**, 1746 (2018).
2. Dunlop, M. J. Engineering microbes for tolerance to next-generation biofuels. *Biotechnol. Biofuels***4**, 32 (2011).
3. Siu, Y., Fenno, J., Lindle, J. M. & Dunlop, M. J. Design and selection of a synthetic feedback loop for optimizing biofuel tolerance. *ACS Synth. Biol.***7**, 16–23 (2018).
4. Tang, T.-C. et al. Materials design by synthetic biology. *Nat. Rev. Mat.***6**, 332–350 (2020).
5. Yan, X., Liu, X., Zhao, C. & Chen, G.-Q. Applications of synthetic biology in medical and pharmaceutical fields. *Signal Transduct. Target Ther.***8**, 199 (2023).
6. Villadsen, J., Nielsen, J. & Lidén, G. *Bioreaction engineering principles* (Springer, 2011).
7. Krivoruchko, A., Siewers, V. & Nielsen, J. Opportunities for yeast metabolic engineering: Lessons from synthetic biology. *Biotechnol. J.***6**, 262–276 (2011).
8. Soma, Y., Tsuruno, K., Wada, M., Yokota, A. & Hanai, T. Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch. *Metab. Eng.***23**, 175–184 (2014).
9. Soma, Y. & Hanai, T. Self-induced metabolic state switching by a tunable cell density sensor for microbial isopropanol production. *Metab. Eng.***30**, 7–15 (2015).

10. Batianis, C. et al. A tunable metabolic valve for precise growth control and increased product formation in *Pseudomonas putida*. *Metab. Eng.***75**, 47–57 (2023).
11. Lalwani, M. A., Zhao, E. M. & Avalos, J. L. Current and future modalities of dynamic control in metabolic engineering. *Curr. Opin. Biotechnol.***52**, 56–65 (2018).
12. Stargardt, P., Feuchtenhofer, L., Cserjan-Puschmann, M., Striedner, G. & Mairhofer, J. Bacteriophage inspired growth-decoupled recombinant protein production in *Escherichia coli*. *ACS Synth. Biol.***9**, 1336–1348 (2020).
13. Zha, W., Rubin-Pitel, S. B., Shao, Z. & Zhao, H. Improving cellular malonyl-CoA level in *Escherichia coli* via metabolic engineering. *Metab. Eng.***11**, 192–198 (2009).
14. Magnuson, K., Jackowski, S., Rock, C. O. & Cronan, J. E. Jr. Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiol. Rev.***57**, 522–542 (1993).
15. Davis, M. S., Solbiati, J. & Cronan, J. E. Jr. Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in *Escherichia coli*. *J. Biol. Chem.***275**, 28593–28598 (2000).
16. Jeung, K., Kim, S., Yeon Lee, J. & Yeol Jung, G. Optimization of mevalonate production using acetate by precursor balancing and flux redistribution in *Escherichia coli*. *J. Ind. Eng. Chem.* **120**, 421–428 (2023).
17. Turner, W. J. & Dunlop, M. J. Trade-offs in improving biofuel tolerance using combinations of efflux pumps. *ACS Synth. Biol.***4**, 1056–1063 (2015).
18. Liu, D. & Zhang, F. Metabolic feedback circuits provide rapid control of metabolite dynamics. *ACS Synth. Biol.***7**, 347–356 (2018).
19. David, F., Nielsen, J. & Siewers, V. Flux control at the malonyl-CoA node through hierarchical dynamic pathway regulation in Saccharomyces cerevisiae. *ACS Synth. Biol.***5**, 224–233 (2016).
20. Wen, J. et al. A synthetic malonyl-CoA metabolic oscillator in *Komagataella phaffii*. *ACS Synth. Biol.***9**, 1059–1068 (2020).
21. Verma, B. K., Mannan, A. A., Zhang, F. & Oyarzún, D. A. Trade-offs in biosensor optimization for dynamic pathway engineering. *ACS Synth. Biol.***11**, 228–240 (2022).
22. Del Vecchio, D., Dy, A. J. & Qian, Y. Control theory meets synthetic biology. *J. R. Soc. Interf.***13**, (2016).
23. Hsiao, V., Swaminathan, A. & Murray, R. M. Control theory for synthetic biology: Recent advances in system characterization, control design, and controller implementation for synthetic biology. *IEEE Control Syst. Mag.***38**, 32–62 (2018).
24. Khammash, M. H. Cybergenetics: Theory and applications of genetic control systems. *Proc. IEEE***110**, 631–658 (2022).
25. Harder, B.-J., Bettenbrock, K. & Klamt, S. Temperature-dependent dynamic control of the TCA cycle increases volumetric productivity of itaconic acid production by *Escherichia coli*. *Biotechnol. Bioeng.***115**, 156–164 (2018).
26. Benisch, M., Aoki, S. K. & Khammash, M. Unlocking the potential of optogenetics in microbial applications. *Curr. Opin. Microbiol.***77**, 102404 (2023).
27. Carrasco-López, C., García-Echauri, S. A., Kichuk, T. & Avalos, J. L. Optogenetics and biosensors set the stage for metabolic cybergenetics. *Curr. Opin. Biotechnol.***65**, 296–309 (2020).
28. Ohkubo, T., Sakumura, Y. & Kunida, K. On-line reoptimization of mammalian fed-batch culture using a nonlinear model predictive controller. *New Gener. Comput.* **42**, 283–302 (2024).
29. Teixeira, A. P., Alves, C., Alves, P. M., Carrondo, M. J. T. & Oliveira, R. Hybrid elementary flux analysis/nonparametric modeling: application for bioprocess control. *BMC Bioinf.***8**, 30 (2007).
30. Mahadevan, R. & Doyle, F. J., III. On-line optimization of recombinant product in a fed-batch bioreactor. *Biotechnol. Prog.* **19**, 639–646 (2003).
31. Xiong, Z. & Zhang, J. Neural network model-based on-line re-optimisation control of fed-batch processes using a modified iterative dynamic programming algorithm. *Chem. Eng. Process Process Intensif.***44**, 477–484 (2005).
32. Espinel-Ríos, S. et al. Toward a modeling, optimization, and predictive control framework for fed-batch metabolic cybergenetics. *Biotechnol. Bioeng.***121**, 366–379 (2024).
33. Hsiao, V., de los Santos, E. L. C., Whitaker, W. R., Dueber, J. E. & Murray, R. M. Design and implementation of a biomolecular concentration tracker. *ACS Synth. Biol.* **4**, 150–161 (2015).
34. Zhang, F. & Keasling, J. Biosensors and their applications in microbial metabolic engineering. *Trends Microbiol.***19**, 323–329 (2011).
35. Dunlop, M. J., Keasling, J. D. & Mukhopadhyay, A. A model for improving microbial biofuel production using a synthetic feedback loop. *Syst. Synth. Biol.***4**, 95–104 (2010).
36. Xie, W., Ye, L., Lv, X., Xu, H. & Yu, H. Sequential control of biosynthetic pathways for balanced utilization of metabolic intermediates in *Saccharomyces cerevisiae*. *Metab. Eng.***28**, 8–18 (2015).
37. Honjo, H. et al. Synthetic microbial consortium with specific roles designated by genetic circuits for cooperative chemical production. *Metab. Eng.***55**, 268–275 (2019).
38. Soma, Y. *et al.* Design of synthetic quorum sensing achieving induction timing-independent signal stabilization for dynamic metabolic engineering of *E. coli*. *ACS Synth. Biol.* **10**, 1384–1393 (2021).
39. Ohkubo, T., Soma, Y., Sakumura, Y., Hanai, T. & Kunida, K. A hybrid *in silico*/in-cell controller for microbial bioprocesses with process-model mismatch. *Sci. Rep.***13**, 1–12 (2023).
40. Åström, K. J. & Murray, R. Feedback Systems: An Introduction for scientists and engineers, Second Edition. (Princeton University Press, 2021).
41. Liu, D., Xiao, Y., Evans, B. S. & Zhang, F. Negative feedback regulation of fatty acid production based on a malonyl-CoA sensor-actuator. *ACS Synth. Biol.***4**, 132–140 (2015).
42. Lu, X., Vora, H. & Khosla, C. Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. *Metab. Eng.* **10**, 333–339 (2008)
43. Monod, J. Recherches sur la croissance des cultures bactériennes. (Hermann, 1942).
44. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature***434**, 1130–1134 (2005).
45. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature***403**, 339–342 (2000).
46. Alon, U. An introduction to systems biology: Design principles of biological circuits. (CRC Press, 2019).
47. Xu, P., Li, L., Zhang, F., Stephanopoulos, G. & Koffas, M. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proc. Natl. Acad. Sci. USA***111**, 11299–11304 (2014).
48. Xu, P. Branch point control at malonyl-CoA node: A computational framework to uncover the design principles of an ideal genetic-metabolic switch. *Metab. Eng. Commun.***10**, e00127 (2020).
49. Kuldell, N., Bernstein, R., Ingram, K. & Hart, K. M. BioBuilder. (O'Reilly, 2015).
50. Ljung, L. System identification toolbox: User's guide. https://www.mathworks.com/help/pdf_doc/ident/ident_ug.pdf (2022).
51. Harrison, M. E. & Dunlop, M. J. Synthetic feedback loop model for increasing microbial biofuel production using a biosensor. *Front. Microbiol.***3**, 360 (2012).
52. Alexis, E., Espinel-Ríos, S., Kevrekidis, I. G. & Avalos, J. L. Biochemical implementation of acceleration sensing and PIDA control. *bioRxiv* 2024.07.02.601775 (2024).
53. Alexis, E., Schulte, C. C. M., Cardelli, L. & Papachristodoulou, A. Regulation strategies for two-output biomolecular networks. *J. R. Soc. Interface***20**, 20230174 (2023).
54. Bothfeld, W., Kapov, G. & Tyo, K. E. J. A Glucose-sensing toggle switch for Autonomous, high productivity genetic control. *ACS Synth. Biol.***6**, 1296–1304 (2017).

55. Gombert, A. K. & Kilikian, B. V. Recombinant gene expression in *Escherichia coli* cultivation using lactose as inducer. *J. Biotechnol.***60**, 47–54 (1998).
56. Chae, H. J. et al. Framework for online optimization of recombinant protein expression in high-cell-density *Escherichia coli* cultures using GFP-fusion monitoring. *Biotechnol. Bioeng.***69**, 275–285 (2000).

Acknowledgements
We thank ACS Authoring Services (https://authoringservices.acs.org/) and Editage (https://www.editage.jp) for English language editing. This study was supported by the Next Generation Interdisciplinary Research Project of the Nara Institute of Science and Technology (NAIST) and AMED under Grant Number JP23wm0425017 and JP23tm0524001.

Author contributions
T.O. designed the study, performed the modeling and simulation, and wrote the manuscript. F.Z. obtained the experimental data. All authors contributed to the study and approved the final manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-76029-1.

Correspondence and requests for materials should be addressed to T.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.