Dynamical Modeling of Optogenetic Circuits in Yeast for Metabolic Engineering Applications

Robert J. Lovelett, Evan M. Zhao, Makoto A. Lalwani, Jared E. Toettcher, Ioannis G. Kevrekidis, and José L. Avalos*
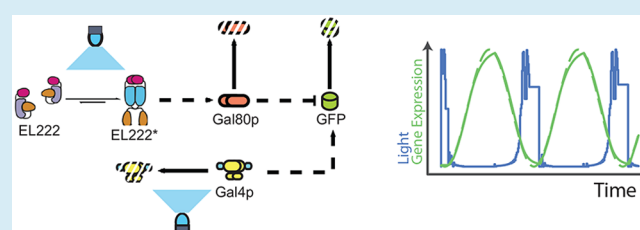
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Dynamic control of engineered microbes using light via optogenetics has been demonstrated as an effective strategy for improving the yield of biofuels, chemicals, and other products. An advantage of using light to manipulate microbial metabolism is the relative simplicity of interfacing biological and computer systems, thereby enabling *in silico* control of the microbe. Using this strategy for control and optimization of product yield requires an understanding of how the microbe responds in real-time to the light inputs. Toward this end, we present mechanistic models of a set of yeast optogenetic circuits. We show how these models can predict short- and long-time response to varying light inputs and how they are amenable to use with model predictive control (the industry standard among advanced control algorithms). These models reveal dynamics characterized by time-scale separation of different circuit components that affect the steady and transient levels of the protein under control of the circuit. Ultimately, this work will help enable real-time control and optimization tools for improving yield and consistency in the production of biofuels and chemicals using microbial fermentations.

INTRODUCTION

Effective dynamical models that describe how systems respond in time to manipulated inputs and perturbations are required for developing the optimization and control algorithms that are ubiquitous in modern engineered systems. In particular, sophisticated control algorithms have enabled the success of the modern petrochemical industry by delivering consistent product quality at high yield.[1] Partly inspired by this success, there has been increasing interest in developing and deploying such mathematical techniques to living systems for biotechnological applications, leading to the growth of a research area termed "cybergenetics."[2−6] A potential application of these ideas is to optimize production of biofuels, chemicals, or pharmaceuticals by engineered microbes.[7]

Deploying optimization and control techniques to a specific problem requires effective tools to manipulate the system in real-time. Toward this end, synthetic biology provides an extensive toolkit for dynamic control of cell cultures, and dynamic control strategies have already been demonstrated as effective for increasing product yield in microbial fermentations.[8,9] Common strategies to control cellular metabolism include induction of metabolic pathways using chemicals (e.g., IPTG[10] or doxycycline[11]), alternative carbon sources (e.g., galactose,[12] arabinose,[13] or ethanol[14]), or nutrient deprivation (e.g., methionine[15]) that can be added to the batch. Alternatively, our group has developed optogenetic circuits that induce or repress metabolic pathways using light.[7,16] Light has several advantages compared to chemical induction,

including that it can be added or removed instantly from a cell culture, it does not rely on changes in media composition, it is relatively inexpensive, it is easily tunable by either intensity or duty cycle, and different wavelengths could be used to orthogonally control different sets of genes. The use of light in controlling the production levels of useful products such as biofuels, chemicals, and recombinant proteins has already been demonstrated.[7,16,17] We envision that these and similar methods will find widespread use in biotechnology.

A key advantage of using light to manipulate metabolism is the ease with which it can interface with computer control systems. Several research groups have used optogenetics for real-time control of gene expression, using feedback control techniques such as model predictive control to track the set point of fluorescent reporter proteins or cell growth rates.[18−20] Additional challenges, however, are introduced when the optogenetic circuit does not drive the expression of a fluorescent reporter that is directly measured, but rather controls enzyme expression, thereby acting as a metabolic valve.[21] Therefore, understanding the dynamics of optogenetic circuits is valuable for predicting how the circuits will perform
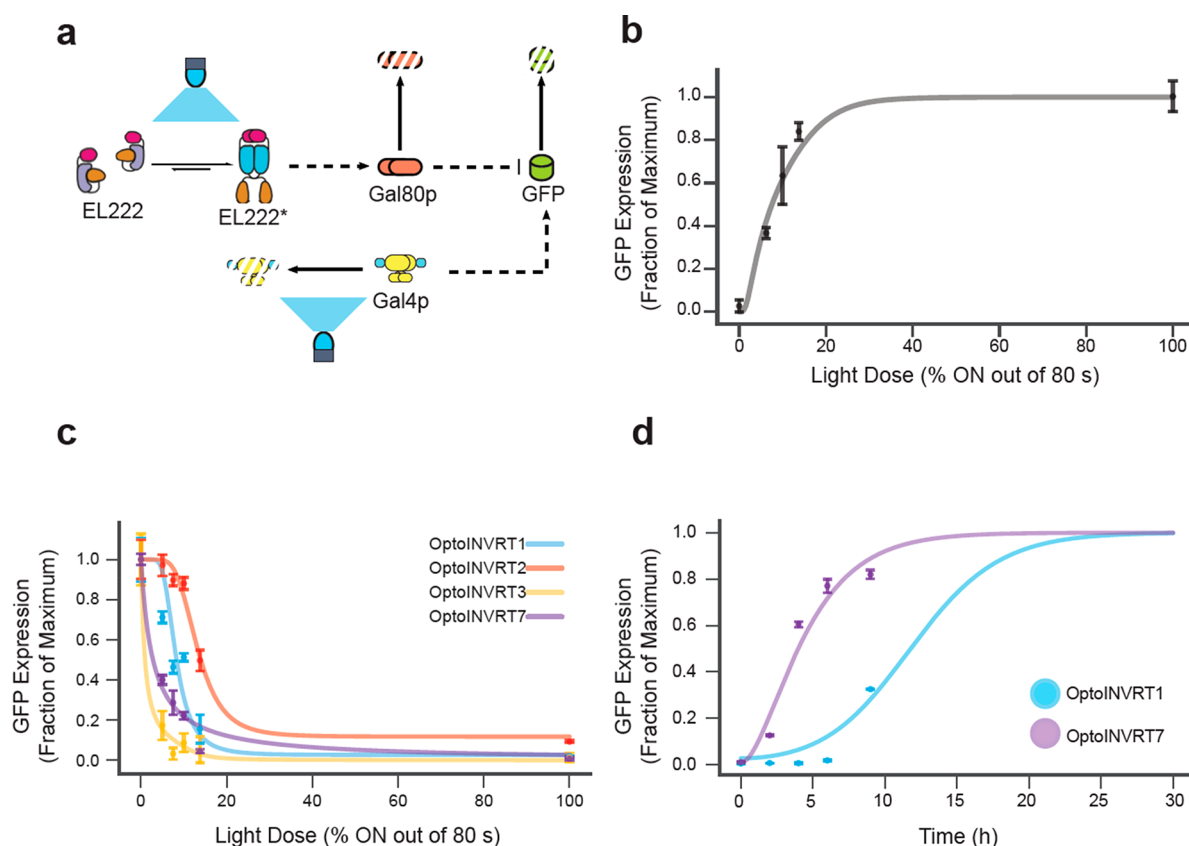
Figure 1. Model design and parameter fitting. (A) Schematic of the *OptoINVRT7* model; simpler circuits can be obtained by removing connections as appropriate. Solid lines indicate biochemical transformations (including degradations) and dashed lines indicate activation or repression. (B) Dose−response of OptoEXP circuit: model (curve) and experimental data (points) show GFP expression in yeast using an 80 s duty cycle with varying light dose. (C) Dose−response of OptoINVRT circuits: model (curve) and experimental data (points) show GFP expression in yeast using an 80 s duty cycle with a varying blue light dose. (D) Model (curve) and experimental data (points) showing change in GFP expression versus time when blue light is switched from on (minimal expression) to off (maximum expression) for OptoINVRT1 (blue) and OptoINVRT7 (purple) circuits. In all experiments, data are first normalized by average GFP expression under the *TEF1* promoter, and then normalized by the maximum observed expression level for that particular circuit. Data are from our other publications;[7,16] error bars represent the s.d. of four biologically independent 1 mL sample replicates exposed to the same conditions. All experiments were repeated at least three times. The dose−response curves computed from the models show time-averaged GFP expression of 80 s limit cycles.

in different contexts and selecting a circuit for a specific application.

We recently presented a set of rapid optogenetic circuits for controlling yeast metabolism that were used to improve yields of biofuels and chemicals,[16] as well as a data-driven methodology for learning the nonlinear optogenetic circuit dynamics.[22] In this Letter, we develop mechanistic dynamical models of these circuits and show that they accurately predict transient and long-time gene expression under various light conditions. We then analyze the models to reveal how the circuits respond to complex light patterns and demonstrate their amenability to modern control algorithms. Our purpose is to use mechanistic understanding combined with mathematical models to reveal how choices made in circuit design result in both qualitative and quantitative differences intransient and steady levels of the protein under control of these circuits. The results we present inform circuit design, help with pairing existing circuits with particular tasks, and may be useful in developing real-time control of microbial chemical production using optogenetics.

■ RESULTS AND DISCUSSION

We developed mechanistic models of five optogenetic circuits for *Saccharomyces cerevisiae* described in our previous work:[7,16] OptoEXP, OptoINVRT1, OptoINVRT2, OptoINVRT3, and OptoINVRT7. OptoEXP is designed to induce transcription in blue light via VP16-EL222, an engineered blue light-activated transcription factor derived from *Erythrobacter litoralis*.[23] OptoINVRT circuits are designed to invert the transcriptional response to light by repurposing the galactose regulon native to *S. cerevisiae*. Specifically, the desired protein is expressed under a promoter activated by Gal4p, which itself is natively regulated by the repressor Gal80p. Gal80p is then controlled by the light-dependent activity of VP16-EL222. OptoINVRT1 and OptoINVRT2 express the transcriptional activator Gal4p using different constitutive promoters, ADH1 and PGK1, respectively. OptoINVRT3 and OptoINVRT7 are further engineered by adding a light-activated degron tag to the transcriptional activator Gal4p. However, Gal80p in Opto-INVRT1, OptoINVRT2, and OptoINVRT3 is highly persistent, which slows the circuit dynamics. This was remedied in OptoINVRT7 by adding a constitutive degron tag to Gal80p. Degron tags are small peptide sequences that target a protein for degradation.[24,25] Such tags can be appended to protein

sequences to artificially reduce their half-lives.[26,27] Protein degradation can be made inducible by placing constraints on proteolytic activation; for example, by only expressing the protease gene in the presence of galactose.[28] Light-activated degron tags work by fusing the core degradation tag sequence with a photosensitive domain that undergoes a conformational shift when exposed to certain wavelengths of light, thereby hiding or exposing the degradation tag for protease recognition based on light conditions.[29] Further details of the optogenetic circuits studied in this work are available in our previous publications.[7,16]

To study these circuits, we used ordinary differential equation models, which have proven valuable in analyzing gene networks of low to moderate complexity in several different systems,[30−33] including some optogenetic circuits.[2,34−36] A schematic of the OptoINVRT circuit design is shown in Figure 1A. The models exclude some mechanistic details (e.g., binding of Gal80p to Gal4p) to simplify the analysis and avoid overfitting and unidentifiable parameters (see Methods). The OptoINVRT models are composed of four ordinary differential equations, representing species balance equations for the fraction of light-activated VP16-EL222 (i.e., VP16-EL222 in the dimer state, bound to DNA, and activating transcription), denoted EL222* in the model, and the normalized protein concentrations of Gal80p (the repressor), Gal4p (the activator), and P_{target} (the protein being control by the circuit). These models describe the average protein levels over the entire population of cells. The model units are fraction of maximum value, which implies that all variables range from 0 to 1 and that P_{target} in each circuit is normalized by the maximum value of expression level for that specific circuit. Further details on model development and assumptions are available in the Methods. For the OptoINVRT circuits, the model equations are

$$\frac{d\,EL222^*}{dt} = \frac{u(t) - EL222^*}{\tau_{EL222^*}}$$

$$\frac{d\,Gal80p}{dt} = \frac{(H_a(EL222^*|K_{EL222^*}, n_{EL222^*}) - Gal80p)}{\tau_{Gal80p}} - \frac{Gal80p}{\tau_{DD}}$$

$$\frac{d\,Gal4p}{dt} = \frac{1 - Gal4p}{\tau_{Gal4p}} - \frac{u(t)Gal4p}{\tau_{LAD}}$$

$$\frac{dP_{target}}{dt} = \frac{(H_r(Gal80p|K_{Gal80p}, n_{Gal80p})Gal4p - P_{target})}{\tau_{target}}$$

Here, EL222*, Gal80p, Gal4p, and P_{target} are the state variables, t is time, and $u(t)$ is the binary input variable indicating that the blue light is on ($u = 1$) or off ($u = 0$). $H_a(C|K,n) = C^n/(C^n + K^n)$ and $H_r(C|K,n) = K^n/(C^n + K^n)$ are Hill functions for transcriptional activation and repression, respectively. The time constants, τ, are empirical and can be considered reciprocals of first order rate constants. OptoINVRT1 and OptoINVRT2 circuits are the simplest designs; they assume $\tau_{DD} = \infty$ and $\tau_{LAD} = \infty$, implying that $\frac{d\,Gal4p}{dt} = 0$. OptoINVRT1 and OptoINVRT2 differ from each other only in the use of different promoters for Gal4p, which is accounted for in the models by using differing Hill parameters in the expression for

$\frac{dP_{target}}{dt}$. OptoINVRT3 is similar to OptoINVRT2, except that Gal4p has been modified with a light activated degradation tag on Gal4p; therefore, it has nonzero $\frac{d\,Gal4p}{dt}$ due to the term containing finite τ_{LAD}. Finally, OptoINVRT7, with its modified Gal80p, includes the term containing finite τ_{DD} leading to lower steady state levels of Gal80p; we further allow a different value of τ_{LAD} in OptoINVRT7 vs OptoINVRT3, as an empiricism that may further correct for the large difference in protein levels across circuits. All other parameter definitions and their values are available in Supplementary Table 1. The blue light input, $u(t)$, can be periodic, so the system can be periodically forced, and one may expect long-term solutions to be oscillations characterized by the forcing frequency (what we will refer to as "periodically forced limit cycles"). The OptoEXP circuit model can be obtained by simplifying the OptoINVRT models, where equations for Gal80p and Gal4p are eliminated, and the equation for P_{target} is replaced by

$$\frac{dP_{target}}{dt} = \frac{(H_a(EL222^*|K_{EL222^*}, n_{EL222^*}) - P_{target})}{\tau_{target}}$$

The models predict both the long-time and transient responses of P_{target} expression under the OptoEXP and OptoINVRT circuits. Transient responses are calculated by integrating the state equations in time. The "long-time" P_{target} expression is found for a given periodic light schedule by calculating the average P_{target} expression over a stable periodically forced limit cycle solution[37] (see Methods for details). Because the dynamics of P_{target} are much slower than the light-active species in the model, the P_{target} expression is essentially constant during short and intermediate forcing period cycles (Supplementary Figure 1). Therefore, the periodically forced limit cycle calculations can be used to calculate a dose–response curve for each optogenetic circuit (Figure 1b,c), where the dose is the fraction of time that light is on, and the response is the normalized P_{target} expression. The dose response curve also depends on the *forcing period* (the duration of time until the cycle is repeated). If the forcing period is 80 s (Figure 1B,C, and Supplementary Figure 1A,B), then a 10% dose means that blue light is continuously ON for 8 s and then continuously OFF for 72 s.

Physiologically relevant parameters are necessary to ensure predictive accuracy of the model. We estimate model parameters using data from our recent publications[7,16] in which GFP is expressed under each of the circuits (i.e., $P_{target} = $ GFP). Except for the activation time of VP16-EL222 (which was previously reported[38]), model parameters were estimated from experimental data of GFP expression from OptoEXP and each OptoINVRT circuit after 8 h, under different light conditions (Figure 1B,C), and from circuit activation experiments (Figure 1D), and can be found in Supplementary Table 1. Our parameter estimation strategy (available in the Methods) allows our model to predict gene expression when light input variations occur in short time scales, such as with 80 s forcing periods, as well as in time scales of hours. The parameter values discovered from the dose–response curve (Figure 1B,C) and kinetics experiments (Figure 1D) are consistent with our expectations, observations, and known biology. For example, Gal80p has a very long estimated decay constant (16 h), which makes this protein linger in the system for a relatively long time, unless a degradation tag is added to the protein (as in OptoINVRT7, which adds a second time
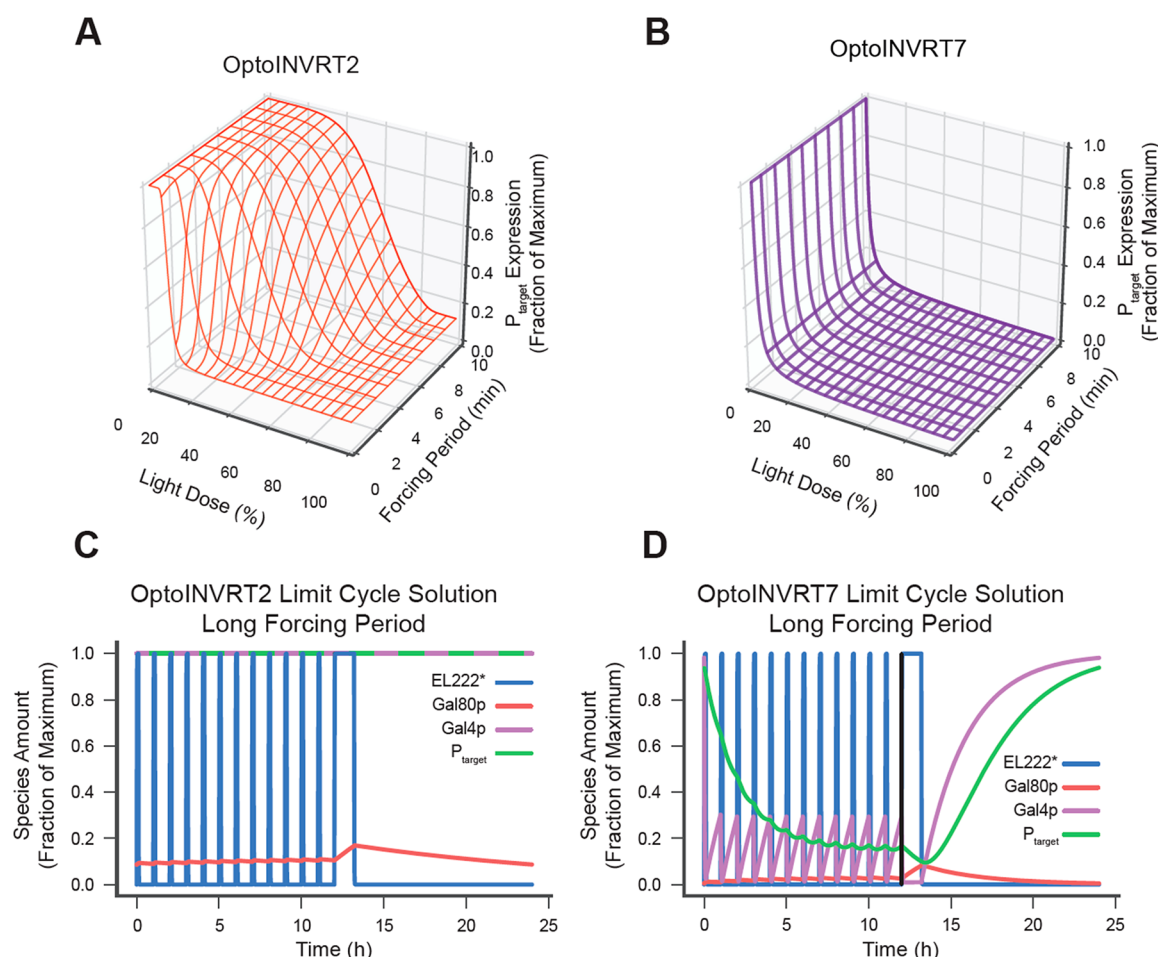
Figure 2. Optogenetic circuit model predictions: Expression surface for OptoINVRT2 (A), and OptoINVRT7 (B), showing the circuit expression response dependence on light dose and forcing period (see Supplementary Figure 2 for expression surfaces for OptoEXP, OptoINVRT1 and OptoINVRT3). Stable limit cycle solution for OptoINVRT2 (C) and OptoINVRT7 (D) circuits, where each forcing period is 1 day. For the first 12 h, blue light is ON for 0.1 out of 1 h, and for the second 12 h, blue light is ON for the initial 1.2 h, and off the remaining time (the black line demarcates the split between the regimes). The average light dose over the whole 24-h period is 10%. For OptoINVRT2 (C), neither forcing regime supplies light for sufficient time to deactivate the circuit. However, for OptoINVRT7 (D), significantly different behaviors in protein levels are observed: for the first 12 h where Gal4p does not have sufficient time to fully regenerate between light doses, P_{target} levels consistently drop; however, during the second 12 h, Gal4p is given enough time to regenerate in the dark, allowing the accumulation of P_{target} at higher levels. (Note that P_{target} is shown as a dashed line in panel C to prevent it from obscuring the line for Gal4p).

constant, $\tau_{DD} = 5$ h). On the basis of the standard errors of the parameter estimates (see Methods), the estimates are mostly robust and indicate the parameters are identifiable from the data, with the important exception of τ_{DD} which is not identifiable from the kinetics experiments. We note that due to time scale separation, the time constants ($\tau_{P_{target}}, \tau_{Gal80p}, \tau_{Gal4p}, \tau_{DD}$) have almost no effect on the shapes of the dose–response curves (Figure 1B,C), while the Hill parameters have almost no effect on the activation kinetics (Figure 1D).

The shapes of the dose–response curves (Figure 1C) and the activation kinetics (Figure 1D) for the OptoINVRT circuits are consequences of the circuit designs that are reflected in the mathematical models. We note that circuits containing light activated degradation of the activator, Gal4p, (OptoINVRT3 and OptoINVRT7) have "decay-like" dose response curves, whereas those that do not (OptoINVRT1 and OptoINVRT2) have "sigmoid-like" dose response curves. This distinction is a result of the fast response of Gal4p to light when it is engineered to include a light-activated degron domain.

The dynamic nature of optogenetic circuits and their distinct responses to different forcing periods may be exploited for improved process design. With this goal in mind, we used the mechanistic circuit models to explore how the shape of the dose–response curve depends on the forcing period. We calculated the dose–response curve of each optogenetic circuit model to various forcing periods from 10 s to 10 min—we call these plots "expression surfaces" (Figure 2A,B and Supplementary Figure 2). We limited the forcing periods to 10 min because this is approximately 10 times faster than the response time of GFP (Supplementary Table 1), which ensures that the target protein remains constant over a forcing period (Supplementary Figure 1C,D) assuming the target protein has a similar response time as GFP. For OptoINVRT1 (Supplementary Figure 2A), OptoINVRT2 (Figure 2A), and OptoEXP (Supplementary Figure 2C), the shapes of the dose–response curves change substantially with the forcing period, because the key regulator in the OptoINVRT systems (Gal80p) and the target protein itself in OptoEXP are tethered to the average activation of VP16-EL222, which due to its rapid activation depends on the forcing period (Supplementary
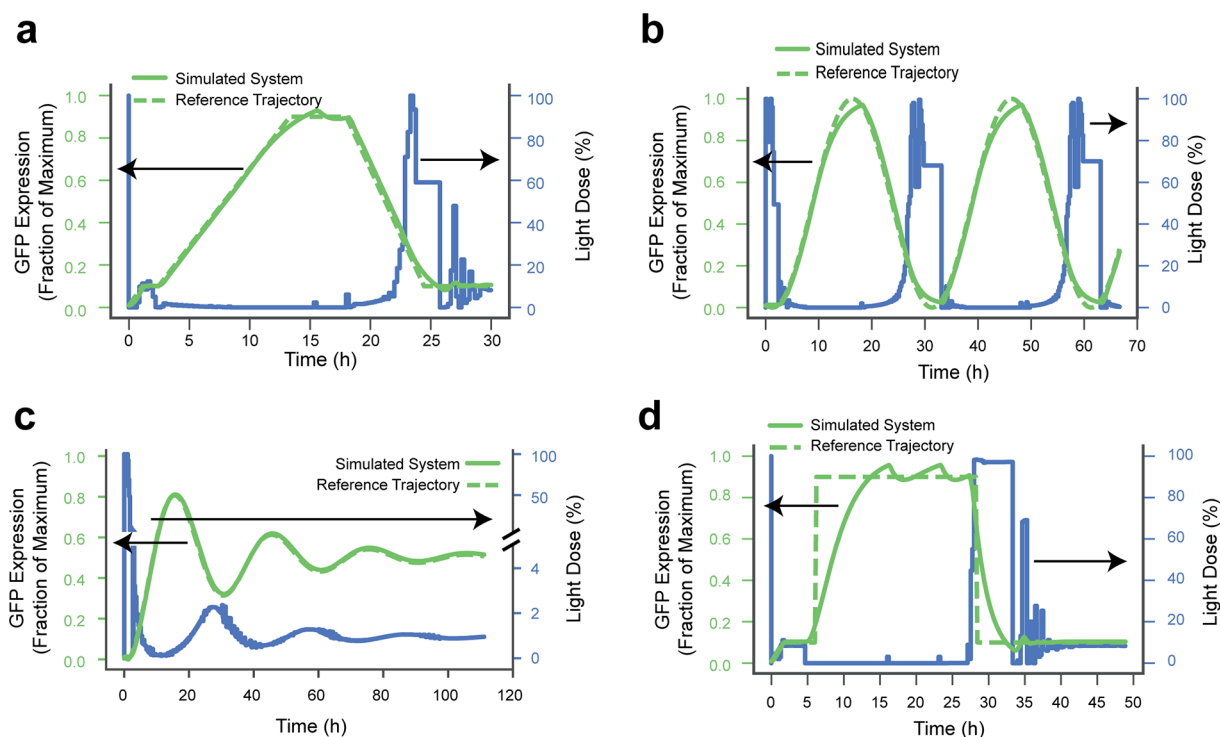
Figure 3. Simulation of model predictive control. *In silico* demonstration of model predictive control of GFP expression (green curves) using the OptoINVRT7 circuit and several reference trajectories (green dashed curves). The manipulated variable is the light dose (blue curves) which indicates the fraction of time that blue light is on out of an 80 s duty cycle.

Figure 1). Conversely, OptoINVRT3 (Supplementary Figure 2B) and OptoINVRT7 (Figure 2B) show a sharper decay-like behavior with light dose, which is a consequence of the light-activated degradation of Gal4p, rendering these circuits more light-sensitive and thus virtually unaffected by the forcing period at short time scales. The expression surfaces reflect the response dynamics of optogenetic circuits and are useful in determining the light dose and short time scale forcing periods required to obtain a desired constant gene expression outcome. Additionally, we speculate that this type of analysis could be used to develop single-input, multi-output controls systems. We imagine a single binary manipulated input variable (light ON or OFF), where there are two degrees of freedom for control (dose and forcing period). For example, if a single organism expressed two optogenetic circuits with orthogonal chemistries, then it may be possible to keep one circuit active while deactivating the other by changing only the forcing period, without everchanging the overall light dose.

The mechanistic models presented here reveal qualitatively different behaviors for the optogenetic circuits that are based on different architectures, which result from differences in the time scale of the response of the important proteins that regulate the expression (Gal80p and Gal4). When we apply longer forcing periods and more complex forcing functions, the target protein can no longer be assumed constant over a periodically forced limit cycle, but we can still calculate the full, time-varying solution for the target protein expression. For example, OptoINVRT2 (Figure 2C) and OptoINVRT7 (Figure 2D) respond quite differently to an irregular forcing function over a 24 h period, with the relatively fast response of OptoINVRT7 enabling comparatively rapid changes in the target protein expression. For OptoINVRT7, 6 min pulses each hour over the first 12 h are sufficient to partially degrade

Gal4p, resulting in a reduction of protein expression; subsequently, the absence of light during the final 10.8 h (after a 1.2-h incubation in light) leads to the regeneration of Gal4p, and in turn, gradual recovery of expression, aided also in part by the more rapid constitutive degradation of Gal80p in this circuit. In contrast, OptoINVRT2 does not respond fast enough to exhibit any changes in expression using this light schedule, keeping expression constantly high. This shows that when light inputs induce significant fluctuations in the target expression, for example during the limit cycle for Opto-INVRT7 (Figure 2D), where simple dose−response relationships no longer apply, our models can still make predictions about time varying expression under such complex forcing periods.

The models and the trajectories of expression levels that we compute reveal how the dynamics of the regulators in the optogenetic circuit (in our case, VP16-EL222, Gal4p, and Gal80) affect both the transient and steady level of the protein under the circuit's control and offer suggestions for process design. For example, if tunability of protein levels is more important than rapid response or maximum expression levels, slowly varying regulators (as in OptoINVRT1 and Opto-INVRT2) with longer forcing periods may be desired. Conversely, if rapid response and high expression levels are needed, then fast decaying regulators (such as OptoINVRT7) are preferable.

Mathematical modeling of optogenetic systems such as VP16-EL222 provides more accurate predictions of response times for light-controlled transcriptional activation and protein degradation. Such models may be the most effective way to optimize performance of the system in bioreactors, given the infinite number of possible time schedules that could be applied during a fermentation. Model predictive control

(MPC), which is widely used in industrial practice,[1,39−41] can utilize these models to enable robust closed-loop feedback control of complex biological systems when supplemented with real-time data, and has previously been used with some optogenetic circuit applications.[2,18,42] We anticipate that by combining system models with live fermentation output tracking, *in silico* controllers can be designed and used to generate well-timed light pulse inputs that optimize chemical production.

Toward this end, we developed an MPC system to track set point changes in GFP expression under the OptoINVRT7 circuit using our mechanistic models. We use a nonlinear MPC algorithm (see Methods) that relies on feedback from a fluorescent reporter protein to estimate the internal state of the optogenetic circuit (i.e., the current levels of Gal4p, Gal80p, and VP16-EL222 activation). The MPC system effectively tracks a variety of reference trajectories (i.e., desired time-varying GFP expression profiles), including ramps, holds, oscillations, and steps (Figure 3). For every case, the MPC controller effectively drives the simulated system along the reference trajectory, even if we consider the effects of imperfect modeling, limited only by the kinetics of the circuits. These results support the possibility of online, model-based control of gene expression for the optogenetic circuits developed in this work.

The models developed here predict the levels of protein controlled by the optogenetic circuit as functions of light input. Future development could extend these models to include important quantities such as the production of metabolites or products synthesized using the proteins under control of the circuit, as well as account for the effects of the circuit dynamics on cell growth/biomass production. One possible way to incorporate these effects would be to include phenomenological expressions, such as Michaelis Menten rate laws for metabolite production or Monod expressions for growth rate that are coupled to the predicted P_{target} levels. Alternatively, these dynamic models could be coupled with genome-scale flux balance analyses, especially more recent versions that explicitly include constraints from enzyme levels.[43] With such a model in place, it would be straightforward to design state estimators to provide real-time estimates of protein or metabolite levels; the state estimator could use real-time data from fluorescent biosensors[44] or infrequent measurements from off-line assays for feedback.

In conclusion, we built computer models to predict the behaviors of optogenetics circuits at different time scales and light doses. These biophysical models are consistent with expected biology and closely fit experimental data from short (Figure 1B,C) and long (Figure 1D) time scales. Because they have a biophysical foundation, we anticipate these models to generalize better than empiricisms and help choose the best optogenetic circuit for a particular task. They enable prediction of expression for any conceivable light schedule and are amenable to modern control systems such as MPC. Ultimately, we envision modeling approaches such as the ones described here will be important for deploying computer-assisted dynamic control and optimization of microbial chemical production.

■ METHODS

Model Development and Assumptions. The mathematical model is composed of dynamic species balance equations on the four species of interest: activated (i.e.,

dimerized) VP16-EL222, Gal80p, Gal4p, and P_{target} (the gene being expressed). We assume that steady-state expression of Gal80p and P_{target} are described by Hill equations, and that the relationship between P_{target} and Gal4p is in the linear regime of the Hill equation (we tested using the full Hill equation, with no improvement in model fidelity). We also assume the total quantity of VP16-EL222 is constant, and some fraction of VP16-EL222 is activated; VP16-EL222 activation kinetics are first order; Gal80p, Gal4p, and P_{target} are produced and degraded with first order kinetics. This implies that transcription or translation are not explicitly modeled but their rates are accounted for by the parameter estimates. We also assume the light activated degron domain on Gal4p induces a first order degradation mode when light is on; the degron domain on Gal80p induces a first order degradation mode at all times. Finally, we assume light penetration does not need to be explicitly modeled and that all cells receive effectively the same light dose (justified by our previous observations[7]).

We assume that the expression of VP16-EL222 is constant in time and constant across each of the different circuit designs. We also assume that the kinetics of the different protein species are constant across different circuits. However, we do not assume that maximum protein levels (other than VP16-EL222) are constant across different circuits. The expression of GFP in each circuit was normalized by its maximum expression specific to the circuit (to more easily compare qualitative differences in the circuit responses), and by its expression under a TEF (constitutive) promoter (to place measurements from the dose response curves and kinetics experiments on a common scale). We allowed variable Hill parameters (which are phenomenological rather than mechanistic) across some circuits, as well as included the term τ_{DD} for optoINVRT7, to account for different maximum expression levels across circuits. We also note that it may not be trivial to extend this model to situations when the light *intensity*, rather than the duty cycle, is variable. Under these circumstances, it is likely that the process gains are nonlinear, which would need to be modeled explicitly.

Limit Cycle Calculation. In our mathematical model, we obtain periodically forced stable limit cycles of our ordinary differential equation model by (1) finding a single point on the limit cycle solution, and (2) integrating the solution from time 0 to T (T = forcing period). Because the dynamics of P_{target} production and degradation are much slower than the forcing period, P_{target} is essentially constant during this limit cycle; therefore, we can average P_{target} expression in time and consider this value the "steady" P_{target} expression. To find such stable limit cycles (1:1 resonant oscillations), we examine the *discrete* dynamical system defined by the change in state variables after one forcing period, which is called the stroboscopic map. Fixed points on the stroboscopic map provide initial conditions for the limit cycle solution. To find such fixed points, we employ a Newton−Raphson algorithm based on a shooting formulation in which we integrate the state equations and variational equations for one period per iteration; integrating the variational equations supplies the Jacobian for Newton−Raphson iterations (*c.f.*, Kevrekidis et al.[37]). We use the VODE algorithm[45] through the SciPy interface[46] to integrate the equations. An illustration of VP16-EL222* changing limit cycles due to a change in forcing period is shown in Supplementary Figure 3.

Parameter Estimation. Parameters are estimated by minimizing an ordinary (unweighted) least-squares objective

function. For the dose–response curves (Figure 1B,C), P_{target} expression is calculated as described in the Limit Cycle Calculation; for the kinetics curves (Figures 1D), GFP (P_{target}) expression is calculated by integrating the state equations in time. In both cases, the objective function is minimized using the BFGS algorithm[47,48] (Jacobian from finite differences) implemented in SciPy. Supplementary Table 1 shows the parameter values and experiments used to obtain them.

We used the progression of the optogenetic circuits from simple (OptoEXP) to complex (OptoINVRT7) and the time-scale separation between the fast changing variables (EL222* and Gal4p) and slow changing variables (P_{target} and Gal80p) to fit parameters in a stepwise fashion. First, we inserted arbitrarily large time constants for P_{target}, Gal80p, and Gal4p to estimate the Hill parameters for EL222* and Gal80p from the dose response curves for OptoEXP and OptoINVRT1/2. Next, we used data from the OptoINVRT3 dose response curve to estimate the time constant for Gal4p light-activated degradation. We repeated this procedure with OptoINVRT7. Then, we estimated the protein time constants using the data from kinetics experiments. Finally, we confirmed our assumption that the time constants for the proteins do not affect the dose–response curves.

We estimated standard errors of the parameters by using a finite difference approximation of the Hessian of the objective function. The standard errors can then be estimated using

$$\widehat{SE(\beta)} = \hat{\sigma} \, \text{diag}(H^{-1}), \text{ with } \hat{\sigma} = \sqrt{\frac{f_{obj}}{N-M}},$$

where f_{obj} is the sum of squared errors, N is the number of data points, and M is the number of parameters. Supplementary Table 1 shows the value, standard error, and the procedure for estimating each of the model parameters for each of the optogenetic circuits. We note that one parameter, τ_{DD} is unidentifiable from our data (i.e, the objective function gradient is numerically zero over a large range of values, resulting in a singular Hessian matrix).

Model Predictive Control. The general idea in model predictive control (MPC) is to use a dynamical system model to predict the system output over a finite prediction horizon and optimize input variables over this prediction horizon to best match a desired output trajectory called the *reference trajectory*. Then, only the first step of the optimized input trajectory is applied, and the process is repeated at each control interval.

For this system, we developed a full nonlinear implementation of MPC where duty cycle is the manipulated variable. We use discrete-time (stroboscopic) MPC with a control interval (period) of 800 s. The forcing period (which, for simplicity, should be an integer divisor of the MPC period) was chosen to be 80 s to correspond with our experiments. The time horizon was chosen to be 6 control intervals (~1.3 h). Choosing the time horizon requires balancing controller performance (improved with longer time horizons) and computational intensity (improved with shorter time horizons). Unmeasured states (Gal80p, Gal4p, EL222*) were obtained from GFP measurements online using a deterministic observer with an innovation function gain of 0.1 chosen for stability[49] where estimates were constrained to lie between 0 and 1. The state estimator reliably converged in the first several iterations (Supplementary Figure 4). Our objective function is the ordinary least-squares difference between the predicted GFP trajectory and the reference GFP trajectory during the time horizon, and has a box constraint on the inputs (0 < light dose < 1). At each control interval, we use a sequential least-squares

quadratic programming (SLSQP) algorithm[50] to obtain the best light dose at each of the next 6 control intervals. The Jacobian for optimization was estimated through finite differences. The light dose from the optimized trajectory is implemented for the next control interval and the process is repeated. To test the control system, we used a "perturbed" model where a Gaussian random variable ($\mu = 0$, $\sigma = 0.05*\theta_i$) was added to each parameter to simulate plant–model mismatch. This model is called the "Simulated Trajectory" in Figure 3.

Numerical Continuation. To efficiently investigate the effect of varying the input waveform and forcing period, we used numerical continuation to calculate GFP expression with light dose ranging from 0% to 100% and forcing period ranging from 10 s to 10 min. The numerical computation package AUTO[51] was used for calculations. The "maps" function was used in AUTO to perform continuation of the stroboscopic map using the VODE integrator to numerically construct the map. The state Jacobian was obtained from the variational equations and the sensitivity matrix (or Jacobian with respect to dose and forcing period parameters) from finite differences.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00372.

Table of model parameter estimates; figure illustrating periodically forced limit cycles of optogenetic circuit models; expression surfaces from OptoINVRT1, Opto-INVRT3, and OptoEXP circuit models; figure illustrating transition between limit cycles; figure demonstrating state estimation procedure used in model predictive control (PDF)

■ AUTHOR INFORMATION

Corresponding Author

José L. Avalos − *Department of Chemical and Biological Engineering and The Andlinger Center for Energy and the Environment, Princeton University, Princeton, New Jersey 08544, United States;* ⓞ orcid.org/0000-0002-7209-4208; Email: javalos@princeton.edu

Authors

Robert J. Lovelett − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States; Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, Maryland 21218, United States;* ⓞ orcid.org/ 0000-0003-1428-0931

Evan M. Zhao − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States*

Makoto A. Lalwani − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States*

Jared E. Toettcher − *Department of Molecular Biology, Princeton, New Jersey 08544, United States;* ⓞ orcid.org/ 0000-0002-1546-4030

Ioannis G. Kevrekidis − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States; Department of Chemical and Biomolecular Engineering, Johns Hopkins University,*

Baltimore, Maryland 21218, United States; orcid.org/
0000-0003-2220-3522

Complete contact information is available at:
https://pubs.acs.org/10.1021/acssynbio.0c00372

■ REFERENCES

(1) Qin, S. J., and Badgwell, T. A. (2003) A survey of industrial
model predictive control technology. *Control Eng. Pract.* 11, 733−764.
(2) Milias-Argeitis, A., Summers, S., Stewart-Ornstein, J., Zuleta, I.,
Pincus, D., El-Samad, H., Khammash, M., and Lygeros, J. (2011) In
silico feedback for in vivo regulation of a gene expression circuit. *Nat.
Biotechnol.* 29, 1114−1116.
(3) Rullan, M., Benzinger, D., Schmidt, G. W., Milias-Argeitis, A.,
and Khammash, M. (2018) An Optogenetic Platform for Real-Time,
Single-Cell Interrogation of Stochastic Transcriptional Regulation.
Mol. Cell 70, 745.
(4) Fiore, G., Perrino, G., Di Bernardo, M., and Di Bernardo, D.
(2016) In Vivo Real-Time Control of Gene Expression: A
Comparative Analysis of Feedback Control Strategies in Yeast. *ACS
Synth. Biol.* 5, 154−162.
(5) Lugagne, J. B., Sosa Carrillo, S., Kirch, M., Köhler, A., Batt, G.,
and Hersen, P. (2017) Balancing a genetic toggle switch by real-time
feedback control and periodic forcing. *Nat. Commun.* 8, 1−7.
(6) Uhlendorf, J., Miermont, A., Delaveau, T., Charvin, G., Fages, F.,
Bottani, S., Batt, G., and Hersen, P. (2012) Long-term model
predictive control of gene expression at the population and single-cell
levels. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14271−6.
(7) Zhao, E. M., Zhang, Y., Mehl, J., Park, H., Lalwani, M. A.,
Toettcher, J. E., and Avalos, J. L. (2018) Optogenetic regulation of
engineered cellular metabolism for microbial chemical production.
Nature 555, 683.
(8) Tan, S. Z., and Prather, K. L. J. (2017) Dynamic pathway
regulation: recent advances and methods of construction. *Curr. Opin.
Chem. Biol.* 41, 28−35.
(9) Lalwani, M. A., Zhao, E. M., and Avalos, J. L. (2018) Current
and future modalities of dynamic control in metabolic engineering.
Curr. Opin. Biotechnol. 52, 56−65.
(10) Donovan, R. S., Robinson, C. W., and Glick, B. R. (1996)
Review: optimizing inducer and culture conditions for expression of
foreign proteins under the control of the lac promoter. *J. Ind.
Microbiol.* 16, 145−154.
(11) Tan, S. Z., Manchester, S., and Prather, K. L. J. (2016)
Controlling Central Carbon Metabolism for Improved Pathway Yields
in Saccharomyces cerevisiae. *ACS Synth. Biol.* 5, 116−124.
(12) Xie, W., Ye, L., Lv, X., Xu, H., and Yu, H. (2015) Sequential
control of biosynthetic pathways for balanced utilization of metabolic
intermediates in Saccharomyces cerevisiae. *Metab. Eng.* 28, 8−18.
(13) Rodrigues, A. L., Becker, J., de Souza Lima, A. O., Porto, L. M.,
and Wittmann, C. (2014) Systems metabolic engineering of
Escherichia coli for gram scale production of the antitumor drug
deoxyviolacein from glycerol. *Biotechnol. Bioeng.* 111, 2280−2289.
(14) Peng, B., Williams, T. C., Henry, M., Nielsen, L. K., and
Vickers, C. E. (2015) Controlling heterologous gene expression in
yeast cell factories on different carbon substrates and across the
diauxic shift: a comparison of yeast promoter activities. *Microb. Cell
Fact.* 14, 91.
(15) Paradise, E. M., Kirby, J., Chan, R., and Keasling, J. D. (2008)
Redirection of flux through the FPP branch-point in Saccharomyces
cerevisiae by down-regulating squalene synthase. *Biotechnol. Bioeng.*
100, 371−378.
(16) Zhao, E. M., Lalwani, M. A., Lovelett, R. J., García-Echaurí, S.
A., Hoffman, S. M., Gonzalez, C. G., Toettcher, J. E., Kevrekidis, I. G.,
and Avalos, J. L. (2020) Design and characterization of rapid
optogenetic circuits for dynamic control in yeast metabolic engineer-
ing. *ACS Synth. Biol.* 9, 3254−3266.
(17) Lalwani, M. A., Ip, S. S., Carrasco-López, C., Day, C., Zhao, E.
M., Kawabe, H., and Avalos, J. L. (2021) Optogenetic control of the
lac operon for bacterial chemical and protein production. *Nat. Chem.
Biol.* 17, 71.
(18) Milias-Argeitis, A., Rullan, M., Aoki, S. K., Buchmann, P., and
Khammash, M. (2016) Automated optogenetic feedback control for
precise and robust regulation of gene expression and cell growth. *Nat.
Commun.* 7, 12546.
(19) Melendez, J., Patel, M., Oakes, B. L., Xu, P., Morton, P., and
McClean, M. N. (2014) Real-time optogenetic control of intracellular
protein concentration in microbial cell cultures. *Integr. Biol.* 6, 366−
372.
(20) Harrigan, P., Madhani, H. D., and El-Samad, H. (2018) Real-
Time Genetic Compensation Defines the Dynamic Demands of
Feedback Control. *Cell* 175, 877−886.
(21) Moser, F., Espah Borujeni, A., Ghodasara, A. N., Cameron, E.,
Park, Y., and Voigt, C. A. (2018) Dynamic control of endogenous
metabolism with combinatorial logic circuits. *Mol. Syst. Biol.* 14, 1−18.
(22) Lovelett, R. J., Avalos, J. L., and Kevrekidis, I. G. (2020) Partial
Observations and Conservation Laws: Gray-Box Modeling in
Biotechnology and Optogenetics. *Ind. Eng. Chem. Res.* 59, 2611−2620.
(23) Motta-Mena, L. B., Reade, A., Mallory, M. J., Glantz, S., Weiner,
O. D., Lynch, K. W., and Gardner, K. H. (2014) An optogenetic gene
expression system with rapid activation and deactivation kinetics. *Nat.
Chem. Biol.* 10, 196−202.
(24) Gottesman, S., Roche, E., Zhou, Y., and Sauer, R. T. (1998)
The ClpXP and ClpAP proteases degrade proteins with carboxy-
terminal peptide tails added by the SsrA-tagging system. *Genes Dev.*
12, 1338−1347.
(25) Zhang, M., MacDonald, A. I., Hoyt, M. A., and Coffino, P.
(2004) Proteasomes Begin Ornithine Decarboxylase Digestion at the
C Terminus. *J. Biol. Chem.* 279, 20959−20965.
(26) Hoyt, M. A., Zhang, M., and Coffino, P. (2003) Ubiquitin-
independent mechanisms of mouse ornithine decarboxylase degrada-
tion are conserved between mammalian and fungal cells. *J. Biol. Chem.*
278, 12135−12143.
(27) Matsuzawa, S., Cuddy, M., Fukushima, T., and Reed, J. C.
(2005) Method for targeting protein destruction by using a ubiquitin-

independent, proteasome-mediated degradation pathway. *Proc. Natl. Acad. Sci. U. S. A. 102*, 14982−14987.

(28) Taxis, C., Stier, G., Spadaccini, R., and Knop, M. (2009) Efficient protein depletion by genetically controlled deprotection of a dormant N-degron. *Mol. Syst. Biol. 5*, 267.

(29) Renicke, C., Schuster, D., Usherenko, S., Essen, L.-O., and Taxis, C. (2013) A LOV2 Domain-Based Optogenetic Tool to Control Protein Degradation and Cellular Function. *Chem. Biol. 20*, 619−626.

(30) Chen, K. C., Calzone, L., Csikasz-Nagy, A., Cross, F. R., Novak, B., and Tyson, J. J. (2004) Integrative analysis of cell cycle control in budding yeast. *Mol. Biol. Cell 15*, 3841−3862.

(31) Brandman, O., Ferrel, J. E., Jr., Li, R., and Meyer, T. (2005) Interlinked Fast and Slow Positive Feedback Loops Drive Reliable Cell Decisions. *Science 310*, 496−498.

(32) Karlebach, G., and Shamir, R. (2008) Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol. 9*, 770−780.

(33) Venturelli, O. S., El-Samad, H., and Murray, R. M. (2012) Synergistic dual positive feedback loops established by molecular sequestration generate robust bimodal response. *Proc. Natl. Acad. Sci. U. S. A. 109*, E3324−E3333.

(34) Olson, E. J., Tzouanas, C. N., and Tabor, J. J. (2017) A photoconversion model for full spectral programming and multi-plexing of optogenetic systems. *Mol. Syst. Biol. 13*, 926.

(35) Olson, E. J., Hartsough, L. A., Landry, B. P., Shroff, R., and Tabor, J. J. (2014) Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nat. Methods 11*, 449−455.

(36) Toettcher, J. E., Voigt, C. A., Weiner, O. D., and Lim, W. A. (2011) The promise of optogenetics in cell biology: interrogating molecular circuits in space and time. *Nat. Methods 8*, 35−38.

(37) Kevrekidis, I. G., Schmidt, L. D., and Aris, R. (1986) Some common features of periodically forced reacting systems. *Chem. Eng. Sci. 41*, 1263−1276.

(38) Zoltowski, B. D., Motta-Mena, L. B., and Gardner, K. H. (2013) Blue light-induced dimerization of a bacterial LOV-HTH DNA-binding protein. *Biochemistry 52*, 6653−6661.

(39) García, C. E., Prett, D. M., and Morari, M. (1989) Model predictive control: Theory and practice-A survey. *Automatica 25*, 335−348.

(40) Meadows, E. S., and Rawlings, J. B. (1996) Model Predictive Control, in *Nonlinear Process Control* (Seborg, D. E., and Henson, M. A., Eds.), pp 233−310. Prentice Hall, Englewood Cliffs, NJ.

(41) Mayne, D. Q., Rawlings, J. B., Rao, C. V., and Po (2000) Constrained model predictive control: Stability and optimality. *Automatica- 36*, 789−814.

(42) Milias-Argeitis, A., and Khammash, M. (2015) Adaptive model predictive control of an optogenetic system. *Conf. Decis. Control*, 1265−1270.

(43) Sánchez, B. J., Zhang, C., Nilsson, A., Lahtvee, P.-J., Kerkhoven, E. J., and Nielsen, J. (2017) Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol. 13*, 935.

(44) Carrasco-López, C., García-Echauri, S. A., Kichuk, T., and Avalos, J. L. (2020) Optogenetics and biosensors set the stage for metabolic cybergenetics. *Curr. Opin. Biotechnol. 65*, 296−309.

(45) Brown, P. N., Byrne, G. D., and Hindmarsh, A. C. (1989) VODE: A Variable-Coefficient ODE Solver. *SIAM J. Sci. Stat. Comput. 10*, 1038−1051.

(46) Jones, E., Oliphant, T., and Peterson, P. *SciPy: Open source scientific tools for python*, version 1.4.1, SciPy Developers.

(47) Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997) Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-con-strained optimization. *ACM Trans. Math. Softw. 23*, 550−560.

(48) Nocedal, J., and Wright, S. J. (2006) *Numerical Optimization*, Springer, New York.

(49) Muske, K. R., and Edgar, T. F. (1996) Nonlinear State Estimation, in *Nonlinear Process Control* (Seborg, D. E., and Henson, M. A., Eds.) pp 311−370, Prentice Hall.

(50) Kraft, D. (1988) *A Software Package for Sequential Quadratic Programming*, Institut fur dynamik der Flugsysteme Oberpfaffenhofen, Cologne, Germany.

(51) Doedel, E. J. (1981) AUTO: A program for the automatic bifurcation analysis of autonomous systems. *Congr. Numer. 30*, 265.