ACS SyntheticBiology

pubs.acs.org/synthbio Letter

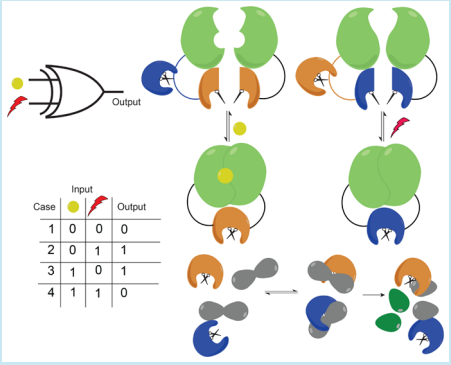
Mathematical Models of Protease-Based Enzymatic Biosensors

Deepak K. Agrawal, Elliott M. Dolan, Nancy E. Hernandez, Kristin M. Blacklock, Sagar D. Khare,* and Eduardo D. Sontag*

Cite This: *ACS Synth. Biol.* 2020, 9, 198−208 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: An important goal of synthetic biology is to build biosensors and circuits with well-defined input−output relationships that operate at speeds found in natural biological systems. However, for molecular computation, most commonly used genetic circuit elements typically involve several steps from input detection to output signal production: transcription, translation, and post-translational modifications. These multiple steps together require up to several hours to respond to a single stimulus, and this limits the overall speed and complexity of genetic circuits. To address this gap, molecular frameworks that rely exclusively on post-translational steps to realize reaction networks that can process inputs at a time scale of seconds to minutes have been proposed. Here, we build mathematical models of fast biosensors capable of producing Boolean logic functionality. We employ protease-based chemical and light-induced switches, investigate their operation, and provide selection guidelines for their use as on−off switches. As a proof of concept, we implement a rapamycin-induced switch *in vitro* and demonstrate that its response qualitatively agrees with the predictions from our models. We then use these switches as elementary blocks, developing models for biosensors that can perform OR and XOR Boolean logic computation while using reaction conditions as tuning parameters. We use sensitivity analysis to determine the time-dependent sensitivity of the output to proteolytic and protein−protein binding reaction parameters. These fast protease-based biosensors can be used to implement complex molecular circuits with a capability of processing multiple inputs controllably and algorithmically. Our framework for evaluating and optimizing circuit performance can be applied to other molecular logic circuits.



Cellular signaling networks perform efficient computation in a complex environment by sensing and processing a multitude of chemical and physical signals into various responses that play a central role in cell metabolism and function. This biological computation has inspired the creation of synthetic genetic networks that follow defined input−output characteristics to generate a diversity of response outputs in response to a set of inputs as a proof-of-principle for biological computation.[1−3] These genetic networks require transcription and translation reactions (after input sensing) to produce proteins. The period of performance of a single input−output layer is several hours. These output proteins (e.g., transcription factors) can be used in turn, to control the expression of other genes in cascade or multilayer network topologies. However, the long (∼hours) time-scales involved preclude efficient construction of multilayer networks. For biosensing applications where rapid sensing and response are required,[4,5] genetic networks are, thus, of limited value. Moreover, heterogeneity of the intracellular environment and crosstalk between synthetic and endogenous components limit not only the speed but also the overall robustness of circuits based on transcription.[3] An alternative biomolecular component library that can potentially operate at a much faster time scale compared to transcription-based genetic networks may help realize complex circuitry for situations, such as real-time detection, for which a fast and precise response is required.

Recently, an alternative paradigm has been developed to realize fast biomolecular computations. This approach uses recombinantly expressed synthetic proteins fused to enzyme fragments.[6−8] Unlike transcription-translation reactions, enzyme catalysis occurs at a time scale of seconds to minutes, and thus protein-based enzymatic biochemical circuits operate at much faster time scales than the transcription-based circuits[9] (Figure 1a,b). These features led to the development of several enzyme-based reaction networks to implement analog and digital logic functions;[7,8] yet, there are few generalizable approaches that can be robustly tailored to implement biosensors with a capability of processing multiple inputs at tunable speeds and controllable sensitivity.[10−12] Many enzymatic circuits are built ad-hoc, using specific substrates/products of particular enzymes as input/output signals.[7,8,10,13,14] However, recent advances in the construction

Received: July 4, 2019
Published: February 4, 2020

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

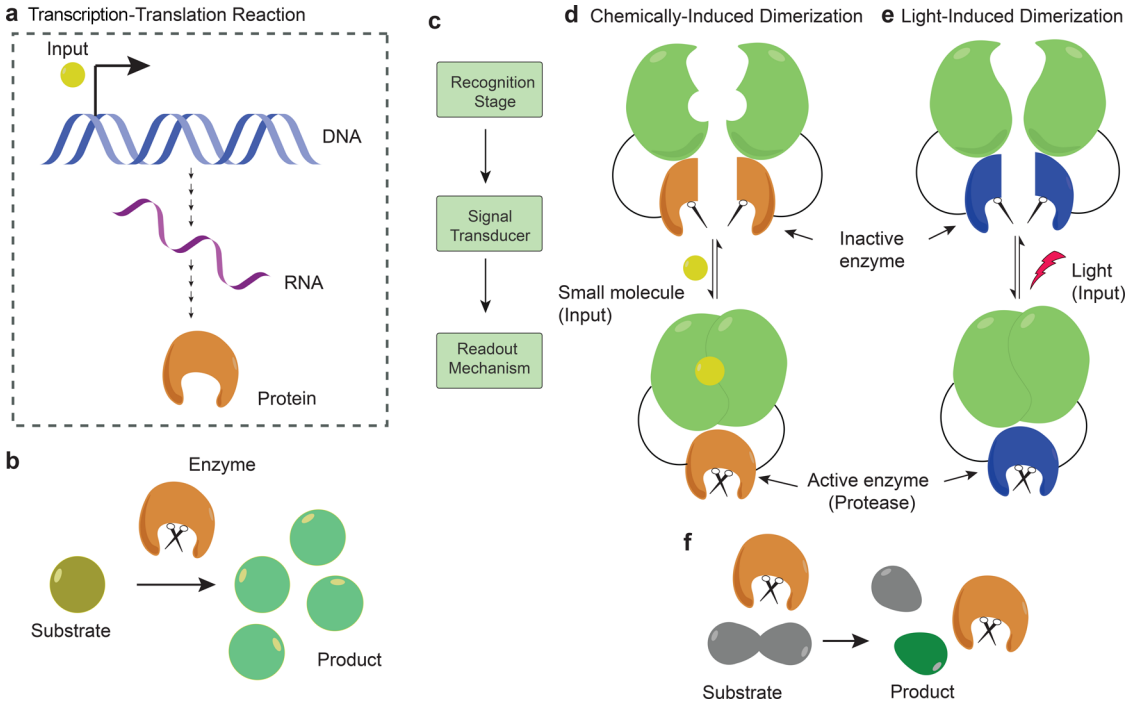


Figure 1. Enzyme-based reaction networks. (a) Genetic regulatory circuits typically use gene expression, which requires transcription (converting DNA to RNA), and translation (converting RNA to protein) reactions. The requirement for these steps leads to long time scales for circuit operation (hours-days) (b) In contrast, enzyme catalysis, in which a protein enzyme acts as a catalyst for converting a substrate molecule into a product, occurs on seconds to minutes time scales. (c) Block diagram representation of the biosensor. (d,e) Concept of induced molecular dimerization. In this scheme, two inactive fragments of a protease are attached to two different protein molecules, which cooperatively undergo dimerization in the presence of (d) chemical or (e) optical signals to induce folding of the attached enzyme fragments and restore its activity. (f) Quenched fluorescent protein system in which only an active protease can cleave the substrate, leading to an increased fluorescence signal.

of entirely bio-orthogonal, post-translationally responsive and controllable protease-based systems have enabled the development of modular components that can potentially be used as building blocks for generating fast and controllable biomolecular circuits both *in vitro* and *in vivo*.[3,7,8]

In this paper, we present a bottom-up design framework for the rapid detection and response to chosen chemical and optical inputs using protease-based logic circuits. We first design two elementary circuits (called switches) based respectively upon chemical and light-induced dimerization mechanisms and capable of producing the same type of output in response to different kinds of inputs. We evaluate their responses for a broad range of reaction conditions and provide screening criteria to optimize their performance. We tested the predictions of our model by experimentally characterizing a chemical (rapamycin) induced switch, and found good qualitative agreement with predictions from our model. Motivated by this, we then use these switches to design biosensors, which can process two different types of inputs simultaneously and produce an output that follows either OR or XOR Boolean logic functionality. We develop comprehensive ordinary differential equation (ODE) models for these biosensors and analyze their dynamic response through numerical analysis at a realistic set of reaction parameters. We further conduct sensitivity analysis to determine the influence of reaction parameters on the output dynamics over time. Our results indicate that a variety of digital signal processing functions can be implemented by using these switches as elementary blocks. Additionally, our approach illustrates several quantitative methods useful for assessing the performance of biomolecular logic gates.

RESULTS

Designing Biosensor Components. Our overall goal is to design biosensors that can process two different types of inputs and rapidly produce a well-defined output signal. The output should be tunable to meet performance specifications that govern the overall dynamics of the system. For example, the biosensor should be able to differentiate between cases where both inputs are absent and the cases where either or both inputs are present (an OR gate). The OR gate design should be extendable to achieve an XOR Boolean logic function, which allows differentiation between cases in which both inputs are either absent or present and cases for which only one input is present at a time.

As a starting point, we choose to divide the operation of the biosensor into three modules (Figure 1c); recognition stage, signal transducer, and read-out mechanism. A recognition stage detects the presence of a particular input, which in our design can be a chemical or physical signal. A signal transducer then converts the detected signal resulting from the interaction between the input and the recognition stage into a measurable molecular activity. In the final stage, the readout mechanism allows generating and reading an optical signal corresponding to the output.

To process chemical and physical input signals via the recognition stage, we used chemically and light induced dimerization mechanisms, respectively. In chemically induced dimerization (CID), a small molecule or a dimerizer acts to bring two proteins or protein fragments together, leading to an increase in the effective concentration of the dimerized complex.[15] CID mechanisms have been widely used to rapidly manipulate molecular activities in cells from a variety of

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

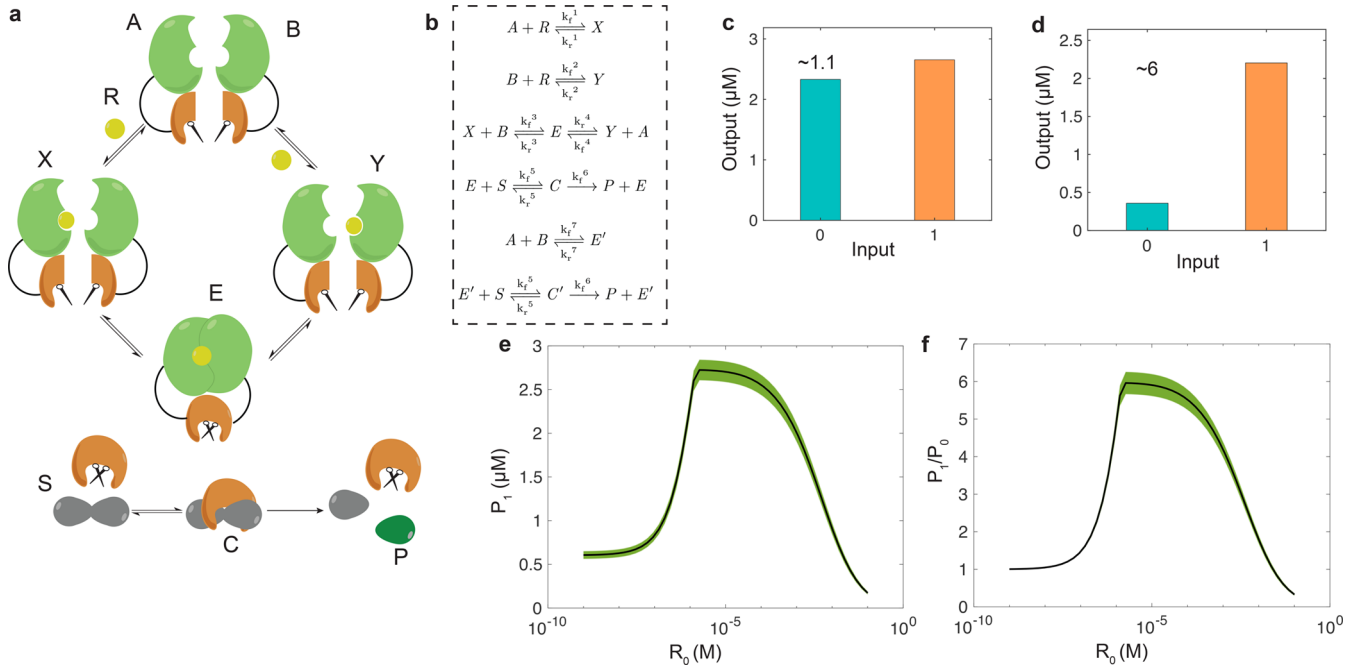


Figure 2. Small molecule induced switch. (a) Schematic diagram represents a switch that uses the chemically induced dimerization (CID) mechanism. In the presence of rapamycin molecule (R), FKBP (A) and FRB (B) proteins form a complex, which leads to restore the activity of split protease E. The protease E then cleaves a fluorophore substrate (S) to produce green fluorescent protein output (P). (b) The corresponding chemical reaction equations for the CID switch. We derived an ODE model (see Supplementary Note S1) from the chemical reactions and used the model to simulate the response of the CID switch with parameters shown in Table 1. (c,d) Simulated response of the CID switch at the initial concentrations of (c) $A_0 = B_0 = 5$ μM, and $R_0 = 0.5$ μM; (d) $A_0 = B_0 = 1.44$ μM, and $R_0 = 3.01$ μM. Here, 0 and 1 represent the absence ($R_0 = 0$ μM) and presence of the input signal (R), and S_0 was 5 μM, while the rest of the molecular species were initially set to 0 μM. (e,f) Performance evaluation of the CID switch at different values of R_0 while keeping A_0 and B_0 fixed at 1.44 μM each. For each concentration of R_0, 1000 simulations were conducted where we randomly sampled a set of parameter values from a uniform distribution (see Methods). Averaged metrics of P_1 and the P_1/P_0 are shown in panels e and f, respectively. All the values of P were determined at 30 min. The error bars are shown in the shaded region and were determined using the standard error of the mean.

species, including *E. coli*.[16] Similar to the technique of CID, light-induced dimerization (LID) exploits a pair of specialized protein domains that can be driven into a high-affinity binding state by illumination with a specific wavelength of light.[17] Light-induced dimers are especially useful because they can be turned on and off with high spatial and temporal resolution in living systems, allowing for control of protein localization and, in our application, enzyme activity. Several LIDs are currently available and have been used to control signaling pathways in living cells.[18−21]

For signal transduction, split proteases can be fused to the pair of proteins used by CID and LID reactions.[22,23] Each split protease is inactive on its own. In the presence of the input, the pair of proteins come together to form a dimer. While this happens, the split proteases come in close proximity, which allows them to reconstitute and thereby restore protease activity (Figure 1d,e). The signal transduction along with the recognition stage allows converting a chemical or physical input signal into a sufficient concentration of the active protease. Finally, for a read-out signal, we use a green fluorescent protein (GFP) which is fused to a quencher protein. The reconstituted protease can cleave the link between GFP and the quencher protein, which causes an increase in the measured fluorescence signal (Figure 1f). In the absence of the active protease, the fluorescent signal is negligible as the quencher is in close proximity to the fluorescent protein.[24] Signal amplification occurs when the transducer enzymatically converts the protease concentration into GFP output. Our approach is to connect the recognition stage, signal transducer, and the read-out mechanism to build mathematical models for the elementary blocks, which can produce a GFP output in response to either chemical or optical signal. These blocks can then be used to design biosensors capable of processing two inputs simultaneously.

Modeling CID Switch. To process a chemical signal, we start by developing a model for a CID switch. To be realistic in our approach, we considered a switch that uses FK506 binding protein (FKBP) and the FKBP rapamycin binding protein (FRB) as the two protein fragments. These two protein fragments can form a dimer in the presence of rapamycin.[25−27] By fusing a fragment of an inactive protease with each protein, that is otherwise unfolded but folds upon enhanced proximity with one another, rapamycin-dependent enzyme activity can be observed. The high affinity of the ternary complex means that small concentrations of rapamycin can be used to trigger enzyme folding, and the entire action can be induced on a time scale of seconds. A previous report of such a CID-split protease fusion shows that Tobacco Etch Virus (TEV) protease activity can be robustly reconstituted in response to rapamycin addition by fusing fragments of TEV protease to FKBP and FRB proteins.[27] For the read-out signal, we considered a recombinant intramolecular FRET construct consisting of GFP and resonance energy-accepting chromoprotein (REACh), fused by a recognition peptide sequence for TEV protease. TEV can cleave this sequence, causing the GFP output to

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

increase.[24] The concentration of the GFP output, therefore, depends on the concentration of the input (rapamycin).

In the presence of input, an interaction between FKPB (denoted as A) and FRB (denoted as B) proteins leads to the formation of TEV protease (denoted as E), which then cleaves the GFP-REACh substrate (denoted as S) to produce the GFP output, which is denoted as P (Figure 2a). The molecular interactions involved in the formation of the FKPB−rapamycin−FRB ternary complex (denoted as E) have been investigated earlier and demonstrated to have two different pathways through which the ternary complex can form (Figure 2a,b).[26] Moreover, split TEV protease fragments can interact with low affinity to generate (at low levels) the active enzyme even in the absence of rapamycin (R).[23] We refer to this rapamycin-independent association of the components as a "leak reaction". To understand the operation of the CID switch, we model its kinetics using an ODE model (see Supplementary Note S1) at a feasible set of reactions parameters (Table 1). For simplicity, hereon FKBP, FRB, rapamycin, TEV, GFP-REACh, and GFP are denoted as A, B, R, E, S, and P, respectively. The subscripts 0 correspond to the initial concentration of the respective species except for P. We determined the switch responses in the absence and presence of R, which are denoted as 0 and 1, respectively. We observed almost the same output response for input 0 as for input 1 (Figure 2c). The lack of difference is likely due to the leak reaction between A and B, which takes place even in the absence of R and leads to the production of E', which then reacts to S to produce P (Figure 2b).

Table 1. Model Parameters for Each Reaction Network[a]

parameters	values	units	refs
k_f^1	1×10^6	$M^{-1}\ s^{-1}$	28
k_r^1	2×10^{-4}	s^{-1}	26
k_r^2	26	s^{-1}	26
k_r^3	12	s^{-1}	26
k_r^4	2×10^{-3}	s^{-1}	26
k_r^5	74	s^{-1}	29
k_f^6	0.02	s^{-1}	29
k_r^7	20	s^{-1}	
k_r^8	0.8	s^{-1}	18
k_r^9	47	s^{-1}	18
k_r^{10}	0.01	s^{-1}	
k_f^{11}	1	s^{-1}	

[a]In this study, k_f^2, k_f^3, k_f^4, k_f^5, k_f^7, k_f^8, k_f^{10}, k_f^{12} were same as k_f^1, $k_r^{10} = k_r^{12}$, and $k_f^{11} = k_f^{13}$. A discussion on how the parameters values were derived can be found in the methods section.

Changing the Reaction Conditions Allows Tuning the Response of the CID Switch. From a switch, typically, a high output (P_1) is desired when the input is 1 (presence of R) compared to a low output (P_0) when the input is 0 (absence of R). This requires defining a specific output range for each input condition for efficiently characterizing the switch operation as off (P_0) or on (P_1). Therefore, we set the following prerequisite conditions; a high output P_1 should be more than 1 μM, and a low output P_0 should be less than 0.5 μM, so that $P_1/P_0 > 2$. To achieve a response from the CID switch that meets our specification, initial concentrations of A, B, and R were optimized (see Methods and Materials), and a desired response was achieved at reduced concentrations of A_0 and B_0, and an increased concentration of R_0 compared to the earlier case (Figure 2d) while using the same parameters values. At reduced initial concentrations of A and B, the total amount of E and E' also reduces, but an increase in R_0 increases only the amount of E (see Figure S1). Notably, the output of the transduction stage (E) takes only a few seconds to reach a steady-state (see Figure S2). It should be noted that P_1 and P_0 values can be different from the ones we selected here as long as they can be measured accurately, and it is easy to differentiate P_1 from P_0.

To understand the CID switch operation comprehensively, we evaluated its performance at a wide range of reaction parameter values, considering the absolute value of P_1 and the P_1/P_0 ratio as metrics. To ensure that our results were not specific to the particular parameter value, we performed 1000 simulations for which each parameter was randomly sampled from a uniform distribution from a bounded interval (see methods) at the optimized initial concentrations of A and B. For simplicity, only the average metrics of P_1 and the P_1/P_0 are shown in Figure 2 panels e and f, respectively. The result obtained at a different value of A_0 and B_0 is shown in Figure S3.

The response curves reveal that the switch performed best in terms of much higher values of P_1 and the P_1/P_0 at a specific concentration of R_0 (Figure 2e,f). This is because the recognition stage of the CID switch has two possible reaction pathways through which E can form. An increase in R_0 increases the amount of E as the interaction between A and R is much stronger than the interaction between B and R (see Table 1). However, after a critical point, the increase in R_0 led to a higher interaction between B and R as did the fact that less B is available to bind to X in order to form E (see Figure S4). Moreover, at the lower concentrations of R_0, P is much higher than at the higher concentrations of R_0 because, at lower values of R_0, A and B are freely available to produce E (product of the leak reaction) than at the higher values of R_0 where A and B are sequestered by R, and forms X and Y (see Figure S5). Following the prerequisite conditions ($P_1 > 1$ μM and the $P_1/P_0 > 2$), we found that the minimum detectable concentration of rapamycin is 0.2 μM and the range of detection is 15.3 mM. By changing this specification ($P_1 > 0.5$ μM and the $P_1/P_0 > 2$), it is possible to reduce the minimum detectable concentration to 2.1 nM, and simultaneously increase the range of detection to 32.4 mM (Figure S3).

Implementation and Characterization of the FKBP-FRB TEV Switch. A key prediction of the CID model developed above is emergent biphasic behavior of the system in response to an increase in the rapamycin concentration: as the concentration of rapamycin increases at a fixed concentration of the protease fragments, the fluorescence response first increases and then decreases (Figure 2e,f). This behavior results from the thermodynamics of rapamycin binding to the FKBP and FRB proteins.[30] To experimentally test the predictions of our model developed for the CID switch, we recombinantly expressed and purified FKBP-NTEV and FRB-CTEV proteins and measured protease activity as a function of rapamycin concentration using a fluorogenic substrate peptide (see Methods) (Figure 3a). We observed an initial increase in proteolytic activity with increased rapamycin concentration (from 0 to 14 μM), followed by a decrease at rapamycin concentrations >14 μM (Figure 3b,c). These trends are remarkably qualitatively similar to the predictions of our model (Figure 2e,f), which was built using "off-the-shelf" parameters obtained from the literature. Fitting the parameters of the model to our experimental data yields a

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

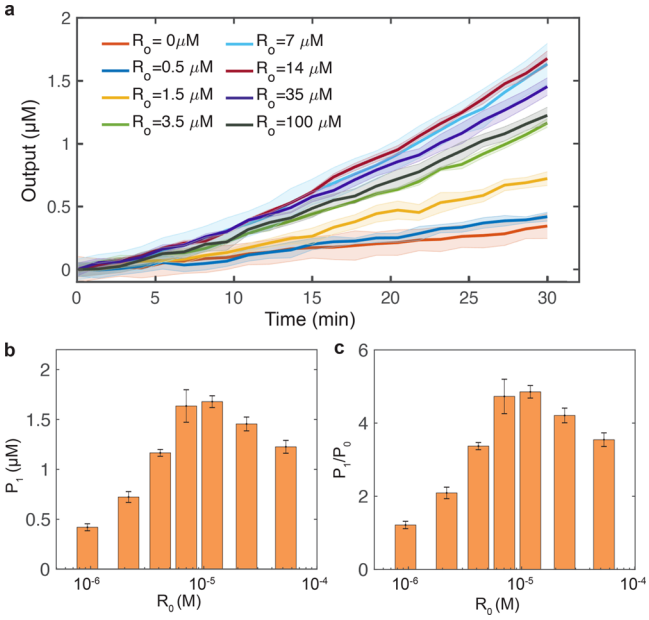


Figure 3. Response of the experimentally implemented FKBP-FRB TEV switch qualitatively agreed with the model, which we developed for the CID switch. (a) Measured response of the FKBP-FRB TEV switch at different concentrations of R_0 (rapamycin) while A_0 (FKBP-PTEV) and B_0 (FRB-NTEV) were fixed at 7 μM each, and S_0 (fluorogenic substrate peptide) was 50 μM. Each kinetic trace was normalized with respect to the first time point recorded. A calibration factor was used to convert the measured signal into the concentration. Experimentally observed variance from reactions run in triplicate (error bars) is shown in the shaded region. Summary of P_1 and the P_1/P_0, defined similarly as in Figure 2, at 30 min is shown in panels b and c, respectively. Error bars are from the SEM of at least three repeats.

parameter set (Figure S6 and Table S1) that can quantitatively reproduce the experimentally observed behavior. Thus, the biphasic behavior predicted from the model is qualitatively validated experimentally, and quantitative fits to the experimental data can be obtained by adjusting two rate constant parameters, thereby providing insight into the modifications made to the underlying molecular components of the switch.

Modeling and Improving the LID Switch Response. Our goal is to design biosensors that can process two different kinds of inputs simultaneously. We, therefore, sought to design a LID switch to detect a physical signal such as an optical signal. To build a realistic model of a LID switch, we considered iLID and SSPB proteins, which form a dimer in the presence of blue light.[18] This reaction was investigated earlier, and an approximately 50-fold increase in the concentration of the dimer was reported in the presence of blue light.[18] Similar to that of the CID switch, by fusing fragments of the inactive split protease with each protein, it should be possible to robustly recover the active protease in response to a physical light signal. For simplicity, we assume that this protease, which is denoted as F, is orthogonal to the one used by the CID switch and forms ~50× more in the presence of input compared to when it is absent. Similar to E, F cleaves the substrate S to produce a GFP output P (Figure 4a,b). Hereon, iLID and SSPB proteins are denoted as U and V, respectively.

We used an ODE model (see Supplementary Note S2) to determine the response of the LID switch at typical initial concentration values of U and V (denoted as U_0 and V_0), and found that P_0 was more than 1 μM (Figure 4c). As per our prerequisite conditions, a low P_0 value should be less than 0.5 μM (P_1 > 1 μM and P_0 < 0.5 μM). A high P_0 value was observed because of the undesired interaction between U and V in the absence of the input, that led to form substantial amounts of F (see Figure S7). We then found reduced U_0 and V_0 values through optimization at which the switch response agreed with our specification (Figure 4d).

To understand these results further, we determined the response curve for the LID switch as a function of U_0 and V_0 ($U_0 = V_0$) at different random sets of reaction parameters (see methods). The mean performances are shown in Figure 4e,f. Unlike the CID switch, the LID switch demonstrated completely different input−output characteristics (Figure 2 and Figure 4). For the LID switch, we observed that P_1 was reduced with a reduction in U_0 and V_0, but the P_1/P_0 ratio increased (within a bound). This means that lowering U_0 and V_0 led to a substantial reduction in P_0 compared to P_1. The binding affinity between U and V can be quantified using the dissociation constant (k_d). The reported k_d value of U (iLID) and V (SSBP) is 47 μM in the absence of light compared to 0.8 μM in the presence of light.[18] Typically, a smaller value of k_d suggests a high affinity between the two species, and so by definition, if ($U_0 = V_0$) < k_d, the reaction is unlikely to form any F. Therefore, at reduced values of U_0 and V_0, the leak reaction produced negligible amounts of F (see Figure S8).

Designing a Biosensor with Boolean OR Gate Functionality. One remarkable property of the CID and LID switches is the potential to network them together to make more complex circuits that can process versatile inputs. As a model system, we aim to design a protease-based biosensor that can mimic a Boolean OR gate functionality. A typical OR gate conventional symbol and a truth table are shown in Figure 5 panels a and b, respectively. It has two inputs and one output, and can produce a high output only when either or both inputs are high. Such a response can, therefore, be used to detect the presence of either or both the inputs simultaneously. To design an OR gate based biosensor, we combined CID and LID switches for which rapamycin and light inputs control the production rate of E and F proteases, respectively. These proteases can cleave the same substrate S to provide a common output P. The two inputs, which are rapamycin and light, therefore, control the production of output P (Figure 5c).

We next sought to model the kinetics of the chemical reaction network for this biosensor (Figure 5d) for four different cases (Figure 5b). In case 1, both the inputs are absent, while in cases 2 and 3 only the light signal or rapamycin is present respectively. In case 4, both the inputs are present. An OR gate operation requires a high output in cases 2, 3, and 4, and a low output in case 1 (Figure 5b). We follow the same specifications as used earlier to categorize the output as either high or low (P_1 > 1 μM and P_0 < 0.5 μM). Using an ODE model (see Supplementary Note S3) and the parameter values shown in Table 1, we simulated the OR gate based biosensor response at the reaction conditions that were used to achieve the desired response from the CID and LID switches separately. Even though high responses were observed for cases 2, 3, and 4, in case 1 the output was more than 0.5 μM, which contradicts our specification. We therefore sought to reduce the output activity when neither input is present while simultaneously maintaining a high output when either or both the inputs are present. To achieve this, we optimized the initial

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

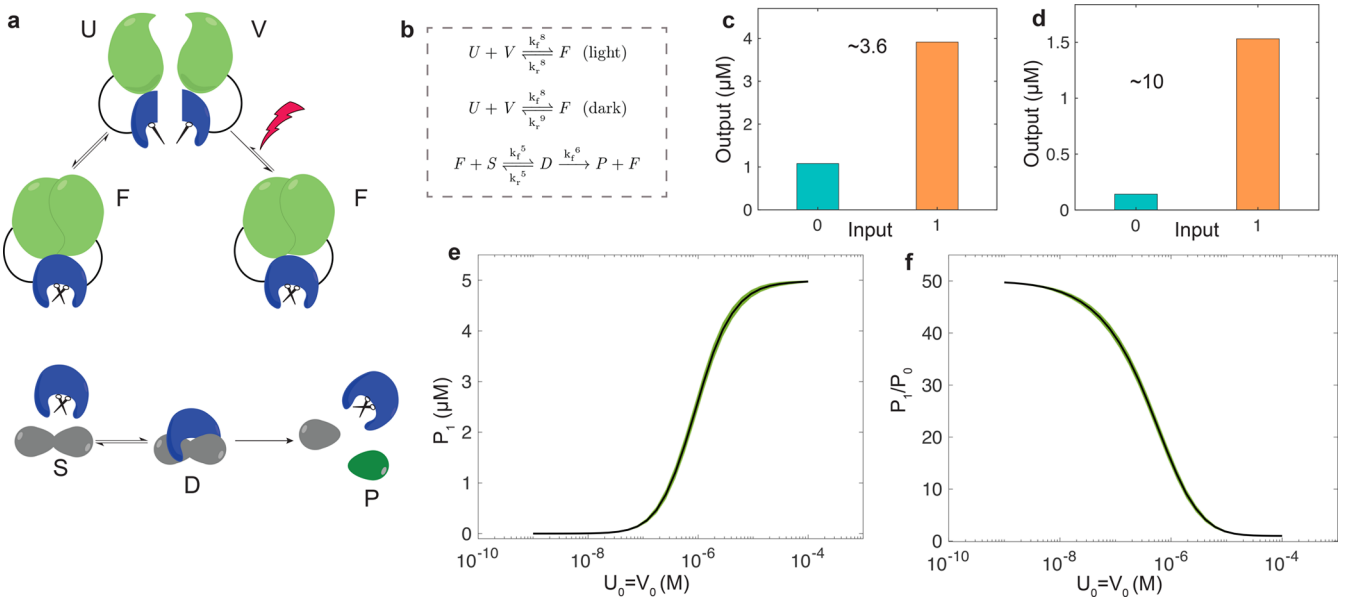


Figure 4. Light-induced switch. (a) Schematic diagram represents a switch that uses the light-induced dimerization (LID) mechanism. In the presence of an optical signal (blue light), iLID microproteins (U) and SSPB proteins (V) form a complex, which leads to the formation of a different version of TEV protease, denoted as F. (b) Corresponding chemical reaction equations for the LID switch. We derived an ODE model (see Supplementary Note S2) from the chemical reaction equations and used the model to simulate the response of the LID switch with parameters shown in Table 1. (c,d) Simulated response of the LID switch at the initial concentrations of (c) $U_0 = V_0 = 5$ μM; (d) $U_0 = V_0 = 1.59$ μM. Here, 0 and 1 represent the absence and presence of the input signal (R), and S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM. (e–f) Performance evaluation of the switch at different values of U_0 and V_0 ($U_0 = V_0$). For each initial concentration value, 1000 simulations were conducted in which we randomly sampled a set of parameter values from a uniform distribution (methods). Averaged metrics of P_1 and the P_1/P_0 are shown in panels e and f, respectively. All the values of P were determined at 30 min. The error bars are shown in the shaded region and were determined using the standard error of the mean.

concentrations of A, B, R, U, and V, and found an OR gate response that met our screening criteria without changing any reaction parameters (Figure 5f).

The operation of an interconnecting reaction network with several species and reaction parameters can be challenging to understand, especially, in the presence of undesired interactions such as the leak reactions. To carry-out our optimization, it is required to determine how the reaction parameters govern the output dynamics at different reaction conditions for the same input combinations. For this purpose, we used sensitivity analysis to get an insight into how each model parameter affects the dynamics of the system. We, therefore, calculated the time-dependent sensitivity coefficient matrix to measure how sensitive the output is with respect to each parameter over time (see methods).[31,32] The output sensitivity for the input combination in case 1 is shown in Figure 5g,h before and after optimization respectively. The coefficients with a high value indicate that variations in the associated parameter cause a significant change in the output dynamics. Note that the coefficients of the normalized sensitivity matrix depend on time.

In Figure 5 panels g and h, the high sensitivity of k_f^7 and k_f^8 suggests that the leak reactions contributed to form the proteases (E' and F) which then led to produce a high output value even in the absence of the inputs. At the optimized condition, a reduced sensitivity of k_f^7 and an increased sensitivity of k_r^8 suggest a reduction in the amount of E' and F, which resulted in lowering the output value (Figure 5f,d). The output sensitivity for the rest of the three cases is shown in Figure S9.

Extending the OR Gate Design to Achieve Boolean XOR Gate Functionality. A biosensor capable of processing a complex computation requires a complex circuit with a capability in which different logic gates can read the same combination of inputs to produce an entirely different logical functionality. Our approach is advantageous over others[7] in the sense that instead of designing a completely new reaction network for each logic gate, CID and LID switches can be considered as elementary blocks to design new logic gates. To demonstrate this capability of our approach, we aim to design a biosensor that can mimic an XOR gate functionality. The conventional symbol and a truth table of an XOR gate are shown in Figure 6a,b, respectively. Unlike the OR gate, which produces a high output when either or both the inputs are high (Figure 5b), an XOR gate provides a high output when only one input is high (Figure 6b).

Similar to the OR gate, the XOR gate has two inputs, rapamycin and light, that restore the activity of E and F proteases respectively (Figure 6c). However, to achieve an XOR functionality, we added additional reactions to the OR gate design in such a manner that these reactions can limit the output production only when both the inputs are present. In these reactions, protease E can degrade U, and protease F can degrade A and the intermediate complex X. Therefore, the two proteases can limit the production of each other, which limits the output production (Figure 6c). For example, in the presence of rapamycin, the active protease E cleaves the substrate S to produce the output P and also degrades U. As in either of these reactions, E is not consumed, a high output should be produced. Similarly, in the presence of the light signal, a high output should be observed as the degradation of A (and X) by F does not affect the output activity (Figure 6d). However, when both the inputs are present, E and F actively degrade U and A (and X), respectively, and this should limit

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198–208

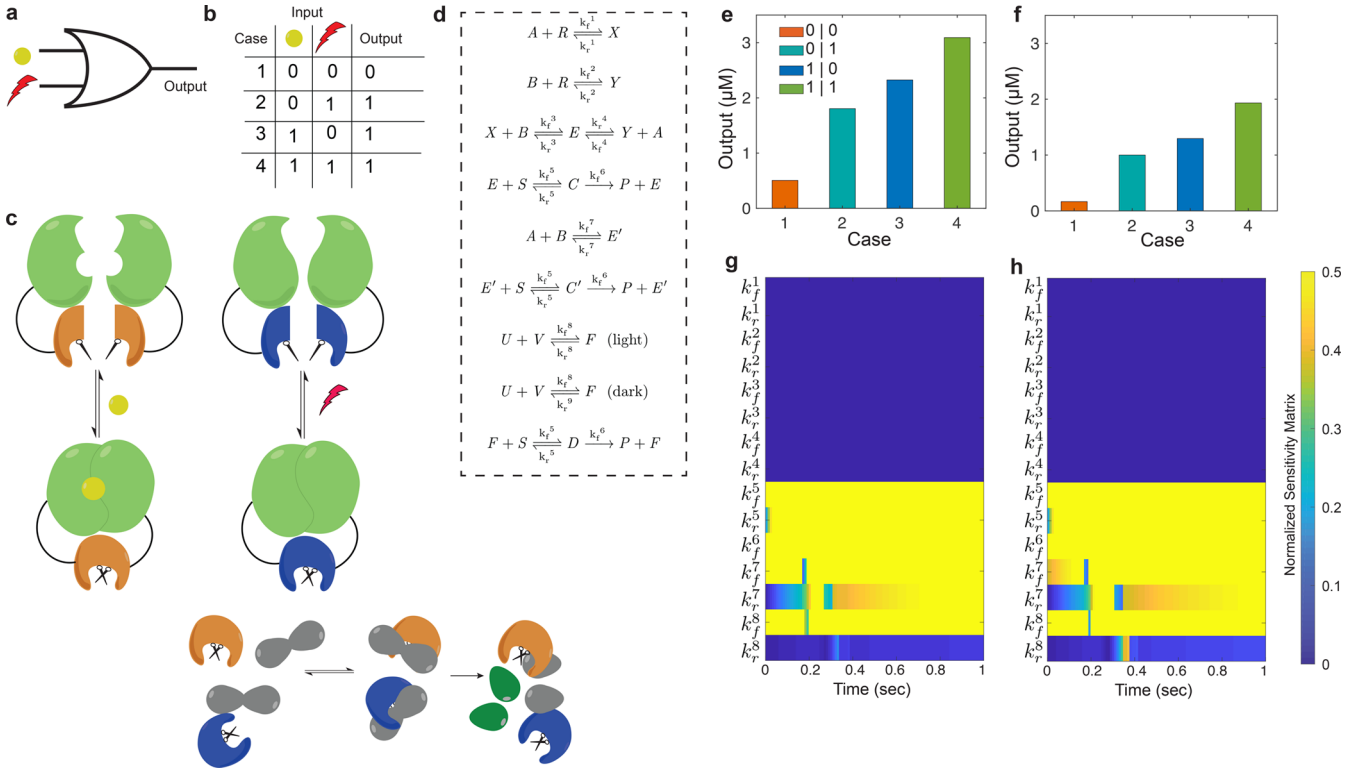


Figure 5. Protease-based Boolean OR gate. (a) Conventional symbol and (b) a truth table of a Boolean OR logic gate. (c) Design of an OR gate that uses CID and LID switches, and (b) the corresponding chemical reactions. The OR gate has two inputs, rapamycin and light, that restore the activity of two split proteases E and F, respectively. A high output in terms of the amount of GFP (P) results when either or both of the inputs are present demonstrating the operation of an OR gate. We model the kinetics of the OR gate using an ODE model (see Supplementary Note S3) with parameters shown in Table 1. (e,f) Simulated response of the OR gate at (e) unoptimized ($A_0 = B_0 = 1.44$ μM, $R_0 = 3.01$ μM, and $U_0 = V_0 = 1.59$ μM) and (f) optimized ($A_0 = B_0 = 0.72$ μM, $R_0 = 1.67$ μM, and $U_0 = V_0 = 1$ μM) conditions, and the corresponding results of the sensitivity analysis in panels g and h, respectively, for case 1 (both the inputs are absent). Normalized sensitivity matrix is shown with respect to the output (P). Here, yellow and blue correspond to the most sensitive and least sensitive values respectively. All the values of P were determined at 30 min. For cases 1 and 2, $R_0 = 0$ μM.

the production of P to a minimal level. Note that for the XOR gate to be functional, the rates at which E and F degrade A (and X) and U, respectively, should be much faster than the rate at which E interacts with S. This constrain enforces a low output when both the inputs are present.

Using an ODE model (see Supplementary Note S4), we determined the biosensor response for four different combinations of inputs (Figure 6b). The same specification was used to categorize the output as either low or high as was used for the OR gate ($P_1 > 1$ μM and $P_0 < 0.5$ μM). Instead of observing a low output in case 4 (both the inputs are present), we observed a high output, that has almost the same activity as in case 2 (Figure 6e). To achieve an XOR logical functionality for all input combinations, we then performed an optimization in A_0, B_0, R_0, U_0, and V_0 analogously to what was done in the OR case and found a desired response that met our specification (Figure 6f).

To understand these results further, we conducted a sensitivity analysis for all four combinations of inputs. The results are shown for case 4 in Figure 6g,h before and after optimizing the reaction conditions, respectively. At the optimized condition, a reduced sensitivity of k_f^1 and k_f^7 suggests reduction in the amount of E and E' while at the same time, increased sensitivity of k_f^8 suggests an increased production of F. This leads to a higher sensitivity to the k_f^{11} and k_f^{12} reaction parameters, which quantify how E and F degrade respectively U and A (and X), reducing the output in case 4. The results for the other cases are shown in Figure S10. This type of analysis provides a way to understand the operation of a complex reaction network where several parameters govern the output dynamics.

DISCUSSION

In the past decade, several biomolecular network topologies capable of performing computations similar to analog and digital circuits were proposed. A majority of these biomolecular circuits are based on gene expression, which requires multiple reactions to happen in a sequential order to produce an output, and because of that these circuits operate at a time scale of hours to days. This limits the computational complexity that can be achieved using molecular computation. Here, we presented a generalized framework to model an emerging class of protease-based Boolean logic gates for biosensing applications.[7] These biosensors use enzymatic post-translational protein modifications, thereby allowing fast circuit operation. We considered a realistic design framework to analyze the input−output characteristics of the CID and LID switches responsive to chemical and light input signal, respectively. We then tested the model of CID switch by experimentally constructing and characterizing the response of FKBP-FRB based rapamycin induced switch, and found good qualitative agreement between experiment and modeling.

We used these switches as elementary blocks to design biosensors to process two different inputs simultaneously, and

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

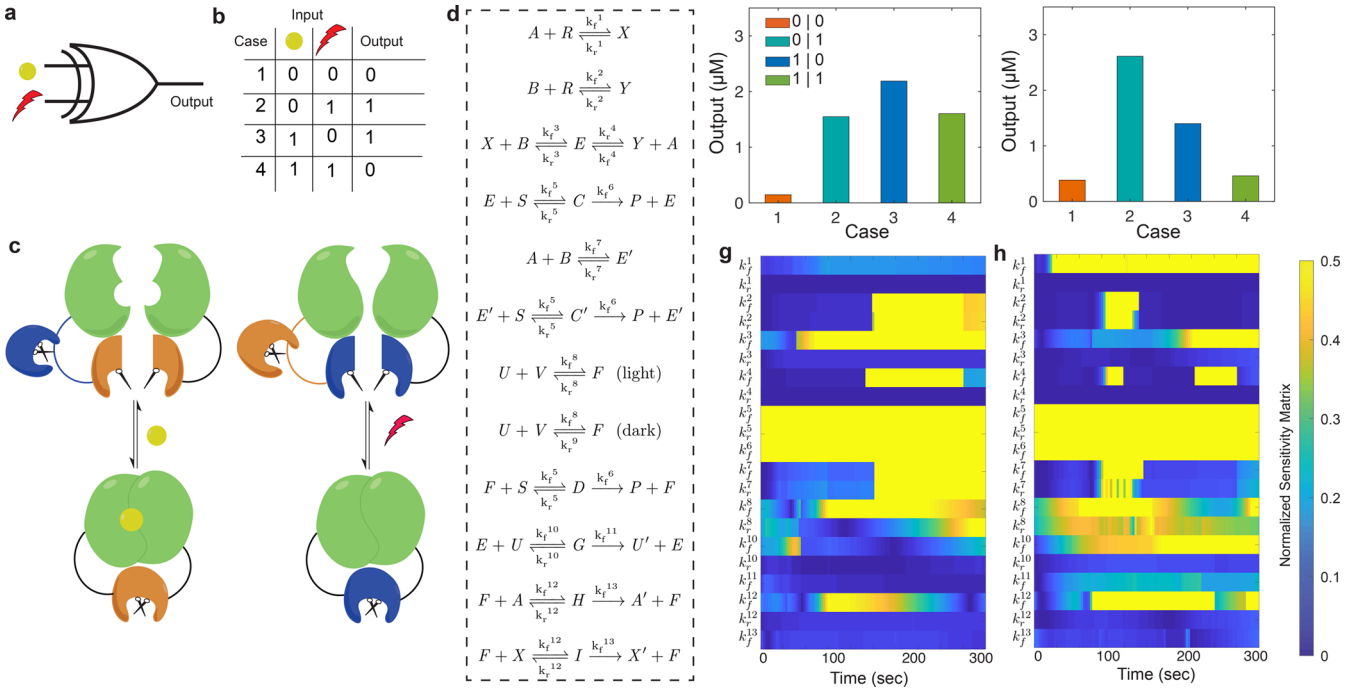


Figure 6. Protease based Boolean XOR gate. (a) Conventional symbol and (b) a truth table of a Boolean XOR logic gate. (c) Design of an XOR gate, which uses the CID and LID switches, and (d) the corresponding chemical reactions. Similar to the OR gate, the XOR gate has two inputs, rapamycin and light, that restore the activity of two split proteases E and F, respectively. However, each of these two proteases can also inactivate the other enzyme by degrading the protein fragment orthogonally. These reactions limit the production of the two proteases when both the inputs are present, resulting in a low output. We model the kinetics of the XOR gate using an ODE model (see Supplementary Note S4) with parameters shown in Table 1. (e–f) Simulated response of the XOR gate at (e) unoptimized ($A_0 = B_0 = 1.44$ μM, $R_0 = 3.01$ μM, and $U_0 = V_0 = 1.59$ μM) and (f) optimized ($A_0 = 1.22$ μM, $B_0 = 0.88$ μM, $R_0 = 0.94$ μM, and $U_0 = V_0 = 2.74$ μM) conditions and the corresponding results of the sensitivity analysis in panels g and h, respectively for case 1 (both the inputs are absent). Normalized sensitivity matrix is shown with respect to the output (P). Here, yellow and blue correspond to the most sensitive and least sensitive values, respectively. All the values of P were determined at 30 min. For cases 1 and 2, $R_0 = 0$ μM.

produce either standard OR or XOR Boolean logic functionality. We improved the expected response of these biosensors through rigorous optimization, and the improvements were explained using sensitivity analysis. The biosensor's capacity to meet the performance specifications considering biologically feasible reaction parameters suggests that this approach is viable for realistic chemical computing circuits. The recent literature on designing biomolecular logic gates lacks a generalized approach that can be used to investigate the operation of an elementary circuit or a complex network.[33,34] This is partially because analyzing protease-based circuits can be difficult due to the irreversibility of cleavage reactions and a lack of a steady-state response. Lack of control over recombinantly expressed protein levels can add to the complexity of accurately modeling circuit behavior inside cells. The mathematical optimization framework developed here with purified protein components can be in principle adapted to model and understand the operation of a complex molecular circuit inside cells by including the kinetics of additional processes for protein degradation, as well as with cell-free systems.

The protease enzyme used in this study, TEV protease, is a common biotechnological reagent and can be produced in high quantities in recombinant form. However, it is a relatively inefficient enzyme compared to many other proteases. For example, the catalytic efficiency of TEV protease is lower than that of SUMO protease.[35] While split SUMO proteases are yet to be developed, we investigated how k_{cat}/K_m variation affects the CID switch response (Figure S11). We found that that irrespective of the speed of protease, the emergent dynamic behavior of the switch remains qualitative similar. The quantitatively optimized response however will be different while implementing logic gates that utilize the modified switch. Note that the dynamic range of the switch is highly sensitive to the background self-association of the split enzyme fragments (rapamycin-independent association of the components): a 2-fold decrease in the background self-association can lead to up to ~30-fold increase in the switch dynamic range and up to ~4.5-fold improvement in the sensitivity (Figure S11). Therefore, our analysis suggests that significant performance improvements can be achieved by molecular design to decrease background self-association (without affecting reconstituted protease activity). Finally, the two-component systems are highly co-operative, and our modeling suggests that when the CID switch model is noncooperative, the dynamic response will be substantially different (Figure S12) than the cooperative CID switch model response (Figure 5). Thus, controlling the binding interactions in the CID system will be key for the use of these systems in biomolecular circuitry.

In this work, we relied on the values of the parameters that were reported in the literature. Our data fitting to the measured response of a CID switch suggested that the differences in the best-fit and "off-the-shelf" parameter set are limited to two parameters out of 13 parameters. These parameters are the dissociation constant of rapamycin from FKBP-NTEV (k_r^1) and the apparent Michaelis constant for the reconstituted enzyme (k_f^5). The former difference likely originates from a decrease in the affinity of rapamycin with

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198–208

FKBP-NTEV compared to FKBP by itself, and the latter reflects a possible activation step (arising from a conformational change) for the reconstituted protease. This additional reaction step in which the reconstituted protease (intermediate species) takes time to get fully activated before it can interact with the substrate can be phenomenologically considered as a reduction in the binding rate of E to the substrate S, which is denoted by k_f^5 in our model. However, such a delay can only result in lowering the output values and should not change the qualitative biphasic characteristics of the CID switch. This can be inferred from Figure S11, where we reduced k_f^6, which should have the same effect on the output P as a reduction in k_f^5.

Several designs and implementations of protein-based circuits have been proposed, that can mimic the operation of Boolean logic gates, but these designs have not yet been tailored to biosensing applications.[7] Moreover, most of the previously designed biosensing mechanisms can detect only one input at a time.[36−38] Our approach is unique as we designed these sensors to recognize two different kinds of inputs simultaneously to produce a programmable output. However, the current design is limited in the sense that the biosensors cannot distinguish different sets of input combinations that provide the same output (either high or low). For example, the output of the OR gate is high for three possible scenarios: either or both the inputs are high (Figure 5). Similarly, for the XOR gate, we cannot distinguish between cases 2 and 3 (Figure 6) for which one or the other input is present. To address this, a multiplexer-based approach might be used to design protease-based circuits where each input combination results in a unique output.[39]

Cellular mechanisms use molecular computation to detect multiple chemical and physical input signals to execute an output that aids cellular function. Our approach can be extended to design other Boolean logic gates and eventually new multilayered, multi-input circuits with complex network connectivity using enzymatic reactions that can be used to develop new types of biosensors with a capability of rapidly detecting and responding to multiple inputs.

METHODS AND MATERIALS

Mathematical Modeling and Sensitivity Analysis. The simulated response of each reaction network was determined by numerically integrating ODE models (shown in the Supporting Information) using the MATLAB ode23s solver unless otherwise specified. Initial conditions for each molecular species are described in figure captions, and the values of reaction parameters are shown in Table 1.

All the forward reaction rate constants (k_f^1, k_f^2, k_f^3, k_f^4, k_f^5, k_f^7, k_f^8, k_f^{10}, and k_f^{12}) were considered to be diffusion-limited,[28] except k_f^6, k_f^{11}, and k_f^{13}. The reverse rate constants k_r^1, k_r^2, k_r^3, and k_r^4 were calculated using the dissociation constant values reported in ref 26 for the rapamycin-induced FKBP-FRB interaction while keeping the respective k_f values fixed at 1×10^6 M^{-1} s^{-1}. The parameters k_r^5 and k_r^6 were inferred k_{cat} and K_m measurements in ref 29 and assuming $k_f^5 = 1 \times 10^6$ M^{-1} s^{-1}. The reverse rate constants k_r^8 and k_r^9 were calculated using the dissociation constant values reported in ref 18 for the iLID and SSPB proteins interactions while using $k_f^8 = 1 \times 10^6$ M^{-1} s^{-1}. The values of k_r^7, k_r^{10}, k_f^{11}, k_r^{12}, and k_f^{13} were picked so as to be consistent with preliminary experiments carried out in our lab. Note that a small variation in k_r^7 does not alter the qualitative behavior of the CID switch (see Figure S11). Similarly, the response of the XOR gate remains almost the same when k_r^{10} and k_r^{12} were increased by a factor of 10 (see Figure S13).

To optimize the responses, we used the MATLAB fmincon function. We used the prerequisite conditions of each reaction as constraints to meet the specific performance criteria. For plots shown in Figures 2e,f, and 4ef, to generate the 1000 combinations of kinetic parameters, each parameter was randomly sampled from a uniform distribution from an interval bounded by a lower bound of 0.1× the nominal value and an upper bound of 10× the nominal value given in Table 1. To determine the output sensitivity of each parameter, we calculated the sensitivity coefficient matrix over time ($s_{i,j}$), which is defined as[40]

$$s_{i,j}(t) = \left.\frac{\partial y_i}{\partial p_j}\right|_t \quad i = 1,2,3,\ldots,N \quad j = 1,2,3,\ldots,M \tag{1}$$

where y_i is a molecular species and p_j is a reaction parameter, the subscript i corresponds to a particular species, and the subscript j to a particular parameter in the system. In our study, y_i is P and p_j can be any of the parameters shown in Table 1. The ODE model equations can be written as

$$\frac{dy_i}{dt} f_i = f_i(\mathbf{y}, \mathbf{p}, t) \tag{2}$$

Here, y and p are the vectors of all the species and parameters, respectively. To calculate $s_{i,j}$, we use a sensitivity differential equation, which can be expressed as

$$\frac{ds_{i,j}}{dt} = \sum_{k=1}^{N} \frac{\partial f_i}{\partial y_k}(\mathbf{y}, \mathbf{p}, t) s_{i,j} + \frac{\partial f_i}{\partial p_j}(\mathbf{y}, \mathbf{p}, t) \tag{3}$$

Equation 3 is solved numerically to calculate $s_{i,j}(t)$ for each parameter and the normalized values of $s_{i,j}$ ($\bar{s}_{i,j}$) are reported in Figures 5 and 6 and Figures S9 and S10 using

$$\bar{s}_{i,j}(t) = \left.\frac{\partial y_i / y_i}{\partial p_j / p_j}\right|_t \tag{4}$$

Creating the FKBP-FRB TEV Switch Fusion Constructs. The FKBP-CTEV and FRB-NTEV constructs were created using Gibson cloning, where each two-piece assembly reaction included (1) amplified linear gene fragments (Thermo Fischer Scientific) containing either the FKBP and CTEV or the FRB and NTEV (sequences from Matthew Good, University of Pennsylvania) and (2) a linearized, modified pet15b vector containing an N-terminal 8-histadine tag followed by a maltose binding protein (MBP) tag, primers and linearized gene fragments found in Table S2. The sequences were confirmed with Sanger sequencing (Genewiz). Protein sequences are provided in Table S3.

Fusion Protein Expression and Purification. Plasmids containing the fusion constructs were transformed into *Escherichia coli* BL21(DE3). Both proteins were expressed by first preparing small cultures of 5 mL of LB with 100 μg/mL of ampicillin, inoculating the media with a single colony, and incubating the culture at 37 °C at 200 rpm overnight. The small cultures were used to inoculate a 500 mL culture of ZYP-5052 media with 100 μg/mL ampicillin, grown at 37 °C at 200 rpm to an OD of 1.5−2.0, and then reduced the temperature to 18 °C to grow overnight.[41] Cultures were spun down at 3000 rcf to harvest the cells and resuspended in lysis buffer (25

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

mM Tris, 150 mM NaCl, 5% glycerol, 30 mM imidazole, pH 7.4), supplemented with 1 mM PMSF, 20 μg/mL lysozyme, and 1 μg/mL DNase (GoldBio) and held on ice for 20 min. The cells were lysed with sonication and clarified with centrifugation (50 krcf for 1 h), and the supernatant was then filtered with 0.22 μm filters. The supernatant was run by gravity over 4 mL of Ni-NTA agarose columns, washed with 5 column volumes of lysis buffer, and eluted with 15 mL of elution buffer (25 mM Tris, 150 mM NaCl, 5% glycerol, 300 mM imidazole, pH 7.4). Both proteins were dialyzed against TNG buffer (25 mM Tris, 300 mM NaCl, 10% glycerol, pH 7.4). Proteins were stored by flash-freezing with liquid nitrogen and storing at −80 °C.

FRET-Quenched TEV Protease Switch Assay. Components for the switch assay are as follows: FRB-NTEV protein (thawed and kept on ice), FKBP-CTEV protein (thawed and kept on ice), rapamycin (dissolved in DMSO at 20× the concentrations below), DTT, the above TNG buffer, and a FRET-quenched fluorophore (reconstituted in DMSO). The FRET-quenched fluorophore peptide substrate was ordered from LifeTein, a methoxycoumarin acetic acid on the *n*-terminus of the TEV-cleavable peptide GENLYFQGSTK, with the c-terminal lysine functionalized with dinitrophenol—TEV substrate. The assay was performed in black half a rea 96-well plates by first preparing reactions of 7 μM FRB-NTEV, 7 μM FKBP-CTEV, 5 mM DTT, and rapamycin at 100, 50, 35, 21, 14, 7, 3.5, 1.5, 1, 0.5, or 0 μM in 100 μL TNG buffer. This was incubated at 34 °C for 10 min before adding 5 μL of a 1 M solution of the above TEV substrate, mixed by pipetting and spun down. Fluorescence (excitation 323 nm/emission 393 nm) was measured continuously for 3 h at 34 °C using the Tecan Infinite M200 Pro plate reader.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.9b00279.

Supplementary Notes S1−S4 contain mathematical models of the switches and the logic gates; Supplementary Figures S1−S13 contain additional simulation results; Supplementary Tables S1−S3 contain parameters values, primers, and proteins sequences respectively (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Sagar D. Khare − *Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854, United States*; Email: khare@chem.rutgers.edu

Eduardo D. Sontag − *Department of Bioengineering and Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts 02120, United States; Laboratory of Systems Pharmacology, Program in Therapeutic Science, Harvard Medical School, Boston, Massachusetts 02115, United States;* orcid.org/0000-0001-8020-5783; Email: e.sontag@northeastern.edu

Authors

Deepak K. Agrawal − *Department of Bioengineering and Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts 02120, United States*

Elliott M. Dolan − *Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854, United States*

Nancy E. Hernandez − *Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854, United States*

Kristin M. Blacklock − *Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.9b00279

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported in part by Grants DARPA FA8650-18-1-7800 and NSF 1817936.

■ REFERENCES

(1) Cameron, D. E., Bashor, C. J., and Collins, J. J. (2014) A brief history of synthetic biology. *Nat. Rev. Microbiol. 12*, 381−90.

(2) Elowitz, M. B., and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature 403*, 335−8.

(3) Nandagopal, N., and Elowitz, M. B. (2011) Synthetic biology: integrated gene circuits. *Science 333*, 1244−8.

(4) Pushkarsky, M. B., Dunayevskiy, I. G., Prasanna, M., Tsekoun, A. G., Go, R., and Patel, C. K. (2006) High-sensitivity detection of TNT. *Proc. Natl. Acad. Sci. U. S. A. 103*, 19630−4.

(5) Antunes, M. S., Morey, K. J., Smith, J. J., Albrecht, K. D., Bowen, T. A., Zdunek, J. K., Troupe, J. F., Cuneo, M. J., Webb, C. T., Hellinga, H. W., and Medford, J. I. (2011) Programmable ligand detection system in plants through a synthetic signal transduction pathway. *PLoS One 6*, No. e16292.

(6) Isaacs, F. J., Dwyer, D. J., and Collins, J. J. (2006) RNA synthetic biology. *Nat. Biotechnol. 24*, 545−54.

(7) Gao, X. J., Chong, L. S., Kim, M. S., and Elowitz, M. B. (2018) Programmable protein circuits in living cells. *Science 361*, 1252−1258.

(8) Fink, T., Lonzarić, J., Praznik, A., Plaper, T., Merljak, E., Leben, K., Jerala, N., Lebar, T., Strmšek, v., Lapenta, F., et al. (2019) Design of fast proteolysis-based signaling and logic circuits in mammalian cells. *Nat. Chem. Biol. 15*, 115.

(9) Miyamoto, T., Razavi, S., DeRose, R., and Inoue, T. (2013) Synthesizing biomolecule-based Boolean logic gates. *ACS Synth. Biol. 2*, 72−82.

(10) Katz, E., and Privman, V. (2010) Enzyme-based logic systems for information processing. *Chem. Soc. Rev. 39*, 1835−57.

(11) Zhou, J., Arugula, M. A., Halamek, J., Pita, M., and Katz, E. (2009) Enzyme-based NAND and NOR logic gates with modular design. *J. Phys. Chem. B 113*, 16065−70.

(12) Privman, M., Tam, T. K., Pita, M., and Katz, E. (2009) Switchable electrode controlled by enzyme logic network system: Approaching physiologically regulated bioelectronics. *J. Am. Chem. Soc. 131*, 1314−1321.

(13) Neurath, H., and Walsh, K. A. (1976) Role of proteolytic-enzymes in biological regulation. *Proc. Natl. Acad. Sci. U. S. A. 73*, 3825−3832.

(14) Amour, A., Bird, M., Chaudry, L., Deadman, J., Hayes, D., and Kay, C. (2004) General considerations for proteolytic cascades. *Biochem. Soc. Trans. 32*, 15−16.

(15) Fegan, A., White, B., Carlson, J. C. T., and Wagner, C. R. (2010) Chemically controlled protein assembly: Techniques and applications. *Chem. Rev. 110*, 3315−3336.

(16) Voss, S., Klewer, L., and Wu, Y. W. (2015) Chemically induced dimerization: reversible and spatiotemporal control of protein function in cells. *Curr. Opin. Chem. Biol. 28*, 194−201.

(17) Kim, B., and Lin, M. Z. (2013) Optobiology: optical control of biological processes via protein engineering. *Biochem. Soc. Trans. 41*, 1183−1188.
(18) Guntas, G., Hallett, R. A., Zimmerman, S. P., Williams, T., Yumerefendi, H., Bear, J. E., and Kuhlman, B. (2015) Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc. Natl. Acad. Sci. U. S. A. 112*, 112−117.
(19) Cheng, P., He, Q. Y., Yang, Y. H., Wang, L. X., and Liu, Y. (2003) Functional conservation of light, oxygen, or voltage domains in light sensing. *Proc. Natl. Acad. Sci. U. S. A. 100*, 5938−5943.
(20) Strickland, D., Lin, Y., Wagner, E., Hope, C. M., Zayner, J., Antoniou, C., Sosnick, T. R., Weiss, E. L., and Glotzer, M. (2012) TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods 9*, 379−U92.
(21) Malzahn, E., Ciprianidis, S., Kaldi, K., Schafmeier, T., and Brunner, M. (2010) Photoadaptation in Neurospora by competitive interaction of activating and inhibitory LOV domains. *Cell 142*, 762−772.
(22) Fang, L., Jia, K. Z., Tang, Y. L., Ma, D. Y., Yu, M., and Hua, Z. C. (2007) An improved strategy for high-level production of TEV protease in Escherichia coli and its purification and characterization. *Protein Expression Purif. 51*, 102−109.
(23) Wehr, M. C., Laage, R., Bolz, U., Fischer, T. M., Grunewald, S., Scheek, S., Bach, A., Nave, K. A., and Rossner, M. J. (2006) Monitoring regulated protein-protein interactions using split TEV. *Nat. Methods 3*, 985−993.
(24) Ganesan, S., Ameer-beg, S. M., Ng, T. T. C., Vojnovic, B., and Wouters, F. S. (2006) A dark yellow fluorescent protein (YFP)-based Resonance Energy-Accepting Chromoprotein (REACh) for forster resonance energy transfer with GFP. *Proc. Natl. Acad. Sci. U. S. A. 103*, 4089−4094.
(25) Spencer, D. M., Wandless, T. J., Schreiber, S. L., and Crabtree, G. R. (1993) Controlling signal-transduction with synthetic ligands. *Science 262*, 1019−1024.
(26) Banaszynski, L. A., Liu, C. W., and Wandless, T. J. (2005) Characterization of the FKBP center dot Rapamycin center dot FRB ternary complex. *J. Am. Chem. Soc. 127*, 4715−4721.
(27) Bierer, B. E., Mattila, P. S., Standaert, R. F., Herzenberg, L. A., Burakoff, S. J., Crabtree, G., and Schreiber, S. L. (1990) Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin. *Proc. Natl. Acad. Sci. U. S. A. 87*, 9231−5.
(28) Schlosshauer, M., and Baker, D. (2004) Realistic protein−protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. *Protein Sci. 13*, 1660−1669.
(29) Yi, L., Gebhard, M. C., Li, Q., Taft, J. M., Georgiou, G., and Iverson, B. L. (2013) Engineering of TEV protease variants by yeast ER sequestration screening (YESS) of combinatorial libraries. *Proc. Natl. Acad. Sci. U. S. A. 110*, 7229−7234.
(30) Douglass, E. F., Jr, Miller, C. J., Sparer, G., Shapiro, H., and Spiegel, D. A. (2013) A comprehensive mathematical model for three-body binding equilibria. *J. Am. Chem. Soc. 135*, 6092−6099.
(31) Agrawal, D. K., Tang, X., Westbrook, A., Marshall, R., Maxwell, C. S., Lucks, J., Noireaux, V., Beisel, C. L., Dunlop, M. J., and Franco, E. (2018) Mathematical modeling of RNA-based architectures for closed loop control of gene expression. *ACS Synth. Biol. 7*, 1219−1228.
(32) Yao, K. Z., Shaw, B. M., Kou, B., McAuley, K. B., and Bacon, D. W. (2003) Modeling ethylene/butene copolymerization with multi-site catalysts: Parameter estimability and experimental design. *Polym. React. Eng. 11*, 563−588.
(33) Privman, V., Arugula, M. A., Halamek, J., Pita, M., and Katz, E. (2009) Network Analysis of biochemical logic for noise reduction and stability: A system of three coupled enzymatic AND gates. *J. Phys. Chem. B 113*, 5301−5310.
(34) Fedichkin, L., Katz, E., and Privman, V. (2008) Error correction and digitalization concepts in biochemical computing. *J. Comput. Theor. Nanosci. 5*, 36−43.
(35) Hickey, C. M., Wilson, N. R., and Hochstrasser, M. (2012) Function and regulation of SUMO proteases. *Nat. Rev. Mol. Cell Biol. 13*, 755.
(36) De Silva, A. P., and Uchiyama, S. (2007) Molecular logic and computing. *Nat. Nanotechnol. 2*, 399−410.
(37) Szacilowski, K. (2008) Digital information processing in molecular systems. *Chem. Rev. 108*, 3481−3548.
(38) Credi, A. (2007) Molecules that make decisions. *Angew. Chem., Int. Ed. 46*, 5472−5475.
(39) Wakerly, J. F. (2005) in *Digital Design: Principles and Applications*, Prentice Hall.
(40) Beck, J. V., and Arnold, K. J. (1977) in *Parameter estimation in engineering and science*, James Beck.
(41) Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression Purif. 41*, 207−234.

https://dx.doi.org/10.1021/acssynbio.9b00279
ACS Synth. Biol. 2020, 9, 198−208

Supplemental Material

Mathematical models of protease-based enzymatic biosensors

Deepak K. Agrawal,[†,‡] Elliott M. Dolan,[¶] Nancy E. Hernandez,[¶] Kristin M. Blacklock,[¶] Sagar D. Khare,[*,¶] and Eduardo D Sontag[*,†,‡,§]

†*Department of Bioengineering, Northeastern University, Boston, MA, USA*

‡*Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA*

¶*Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ, USA*

§*Laboratory of Systems Pharmacology, Program in Therapeutic Science, Harvard Medical School, Boston, MA, USA*

E-mail: khare@chem.rutgers.edu; e.sontag@northeastern.edu

Supplementary Note S1

Using the mass-action law, the kinetics of the CID switch can be modeled as:

$$\dot{A} = k_r^1 X - k_f^1 AR + k_r^4 E - k_f^4 YA - k_f^7 AB + k_r^7 E', \tag{1}$$

$$\dot{R} = k_r^1 X - k_f^1 AR + k_r^2 Y - k_f^2 BR, \tag{2}$$

$$\dot{B} = k_r^2 Y - k_f^2 BR + k_r^3 E - k_f^3 XB - k_f^7 AB + k_r^7 E', \tag{3}$$

$$\dot{X} = k_f^1 AR - k_r^1 X - k_f^3 XB + k_r^3 E, \tag{4}$$

$$\dot{Y} = k_f^2 BR - k_r^2 Y - k_f^4 YA + k_r^4 E, \tag{5}$$

$$\dot{E} = k_f^4 YA + k_f^3 XB - k_r^3 E - k_r^4 E - k_f^5 ES + k_r^5 C + k_f^6 C, \tag{6}$$

$$\dot{S} = -k_f^5 ES + k_r^5 C - k_f^5 E'S + k_r^5 C', \tag{7}$$

$$\dot{C} = k_f^5 ES - k_r^5 C - k_f^6 C, \tag{8}$$

$$\dot{P} = k_f^6 C + k_f^6 C', \tag{9}$$

$$\dot{E'} = k_f^7 AB - k_r^7 E' - k_f^5 E'S + k_r^5 C' + k_f^6 C', \tag{10}$$

$$\dot{C'} = k_f^5 E'S - k_r^5 C' - k_f^6 C'. \tag{11}$$

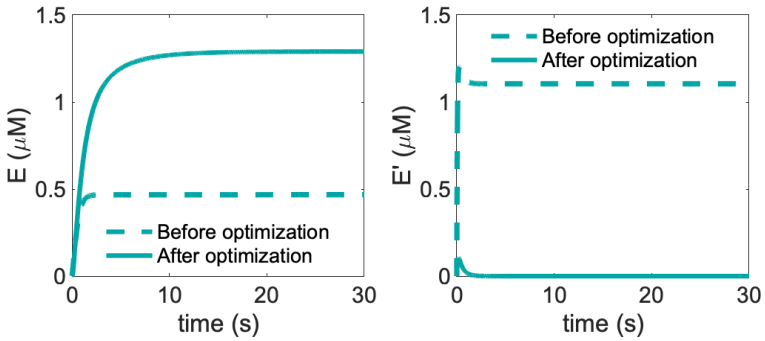


Figure S1: Timing response of the CID switch in the presence of input (R) before optimization where $A_0 = B_0 = 5$ μM, and $R_0 = 0.5$ μM (dashed-line), and after optimization where $A_0 = B_0 = 1.44$ μM and $R_0 = 3.01$ μM (solid-line). A reduction in the initial concentrations of A and B reduced the amount of E and E', but an increase in R_0 increased only the amount of E, resulting in an overall increase in the amount of E. The ODE model shown in Note S1 was used to simulate the response with parameters shown in Table 1. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

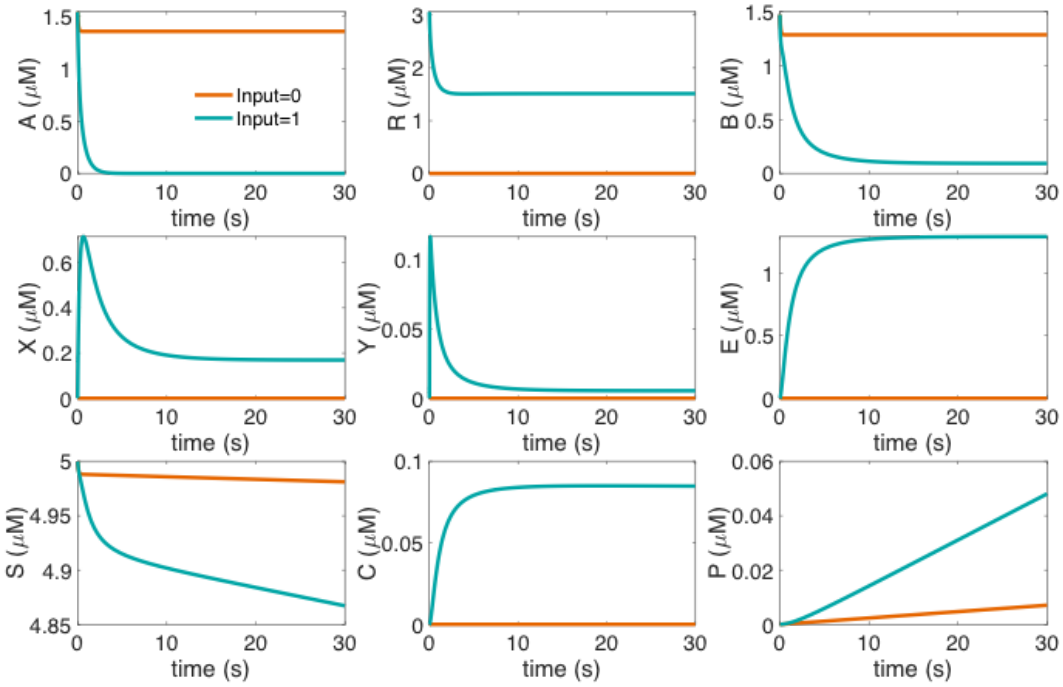


Figure S2: Simulated timing response of different species that make up the CID switch while keeping $A_0 = B_0 = 5$ μM, and $R_0 = 0.5$ μM. Species involved in the sensing element reached steady-state within a few seconds. The ODE model shown in Note S1 was used to simulate the response with parameters shown in Table 1. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

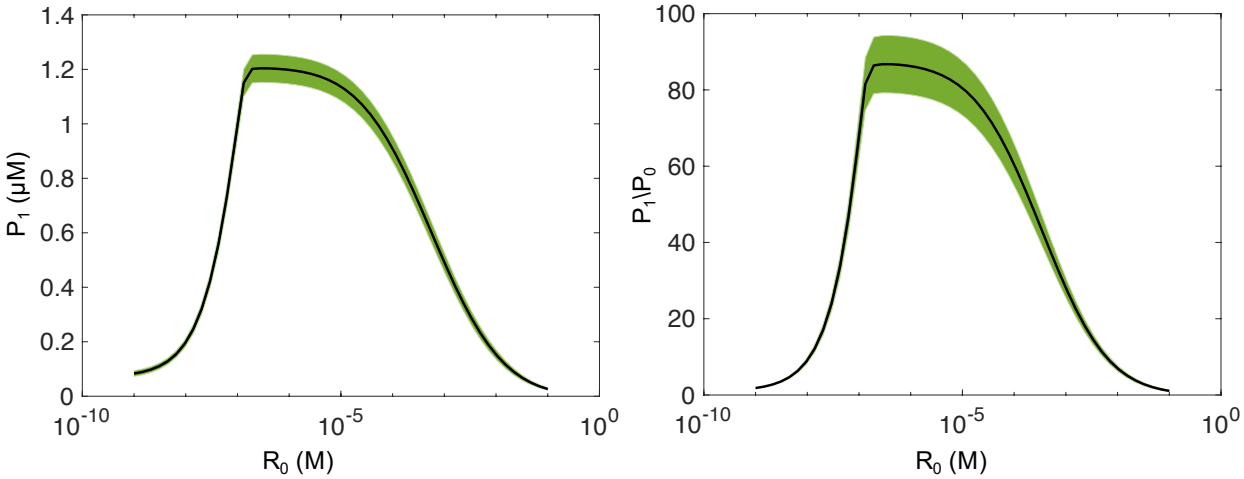


Figure S3: Performance evaluation of the CID switch as a function of R_0 while keeping A_0 and B_0 fixed at 0.144 μM each. For each concentration of R_0, 1,000 simulations were conducted. Parameter values were randomly sampled from a uniform distribution (see Methods). The ODE model shown in Note S1 was used to simulate the response with parameters shown in Table 1. Averaged metrics of P_1 and the P_1/P_0 are shown here. All the values of P were determined at 30 min. The error bars are shown in the shaded region and were determined using the standard error of the mean. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

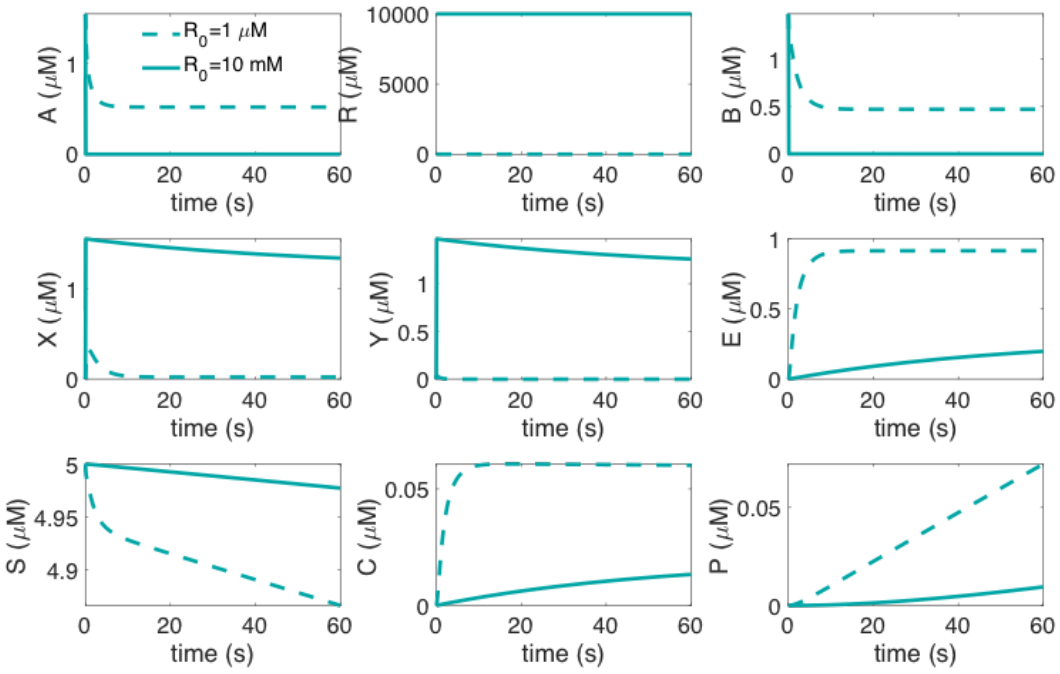


Figure S4: Simulated timing response of different species that make up the CID switch at different initial concentrations of R while keeping A_0 and B_0 fixed at 1.44 μM each. An increase in R_0, led to a higher interaction between B and R, and because of that, less B is available to bind to X in order to form E. The ODE model shown in Note S1 was used to simulate the response with parameters shown in Table 1. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

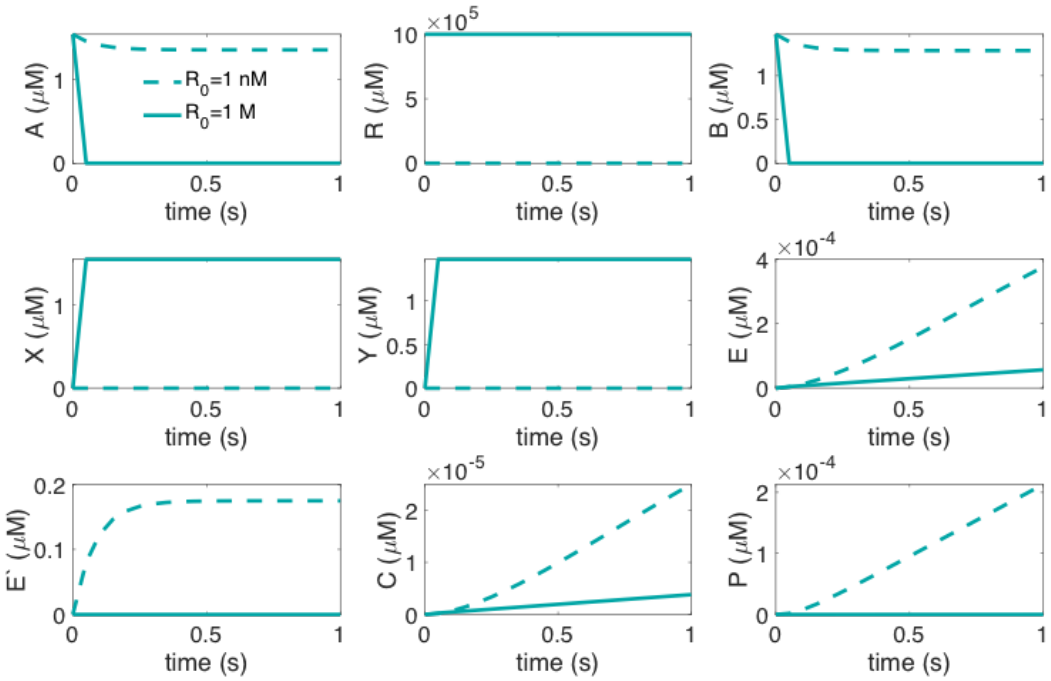


Figure S5: Simulated timing response of different species that make up the CID switch at different initial concentrations of R while keeping A_0 and B_0 fixed at 1.44 μM each. At the lower concentrations of R_0, A and B are freely available to produce E' than at the higher values of R_0 where A and B are sequestered by R, and forms X and Y. The ODE model shown in Note S1 was used to simulate the response with parameters shown in Table 1. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

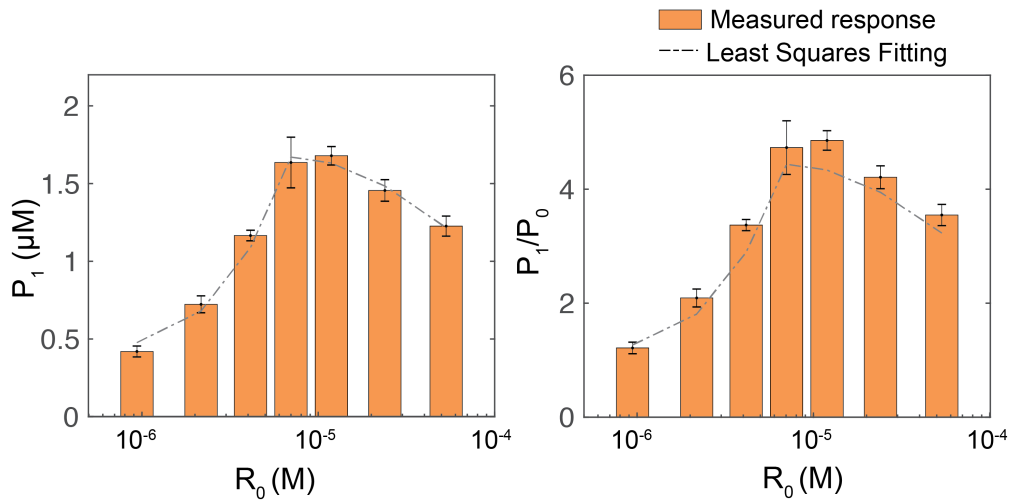


Figure S6: Comparing the measured response of the rapamycin induced FKBP-FRB TEV switch at different concentrations of R_o (rapamycin) while A_0 (FKBP-PTEV) and B_0 (FRB-NTEV) fixed at 7 μM each and S_0 fluorogenic substrate peptide) was 50 μM. Summary of P_1 and the P_1/P_0 are shown at 30 mins. Here, P_0 and P_1 correspond to the cases with R_o is zero and a nonzero value respectively. Each kinetic trace was normalized with respect to the first time point recorded. A calibration factor was used to convert the measured signal into the concentration. The ODE model shown in Note S1 was used to calculate the switch response with parameters shown in Table S1. Least squares fitting was used to generate the best-fit model response considering all the values of the output (P) at 30 min. Error bars are from the SEM of at least three repeats.

Table S1: Estimated model parameters obtained from the least squares fitting for the ODE model shown in Figure 2b. Here, k_f^2, k_f^3, k_f^4, k_f^7 were same as k_f^1.

Parameters	Values	Units
k_f^1	1×10^{6}	$M^{-1}s^{-1}$
k_r^1	5.1×10^{-3}	s^{-1}
k_r^2	16.71	s^{-1}
k_r^3	0.36	s^{-1}
k_r^4	$5.7\times\ 10^{-3}$	s^{-1}
k_f^5	9.99×10^{3}	$M^{-1}s^{-1}$
k_r^5	29.43	s^{-1}
k_f^6	0.0104	s^{-1}
k_r^7	27.85	s^{-1}

Supplementary Note S2

Using the mass-action law, the kinetics of the LID switch can be modeled as:

$$\dot{U} = -k_f^8 UV + k_r^8 F, (\text{dark}) \tag{12}$$

$$\dot{V} = -k_f^8 UV + k_r^8 F, (\text{dark}) \tag{13}$$

$$\dot{U} = -k_f^8 UV + k_r^9 F, (\text{light}) \tag{14}$$

$$\dot{V} = -k_f^8 UV + k_r^9 F, (\text{light}) \tag{15}$$

$$\dot{F} = k_r^5 D + k_f^6 D - k_f^5 FS + k_f^8 UV - k_r^8 F, \tag{16}$$

$$\dot{S} = -k_f^5 FS + k_r^5 D, \tag{17}$$

$$\dot{D} = k_f^5 FS - k_r^5 D - k_f^6 D, \tag{18}$$

$$\dot{P} = k_f^6 D. \tag{19}$$

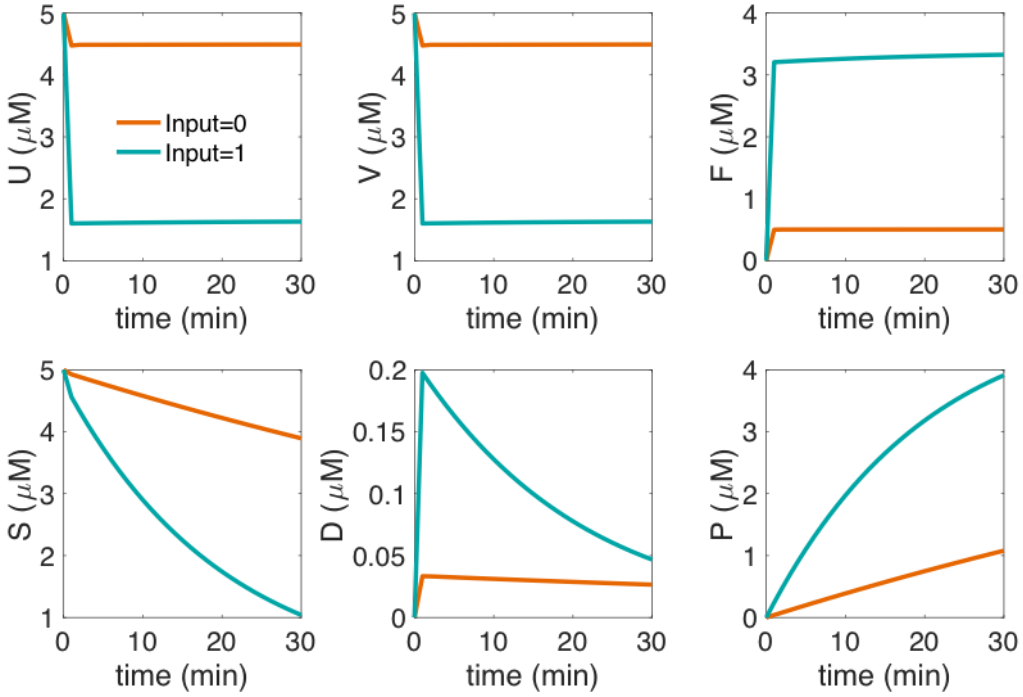


Figure S7: Simulated timing response of different species that make up the LID switch in the absence (denoted as 0) and presence of input (denoted as 1) while keeping U_0 and V_0 at 5 μM each. U and V dimerized even in the absence of input, which allowed to form F, resulted in a high amount of P. The ODE model shown in Note S2 was used to simulate the response with parameters shown in Table 1. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

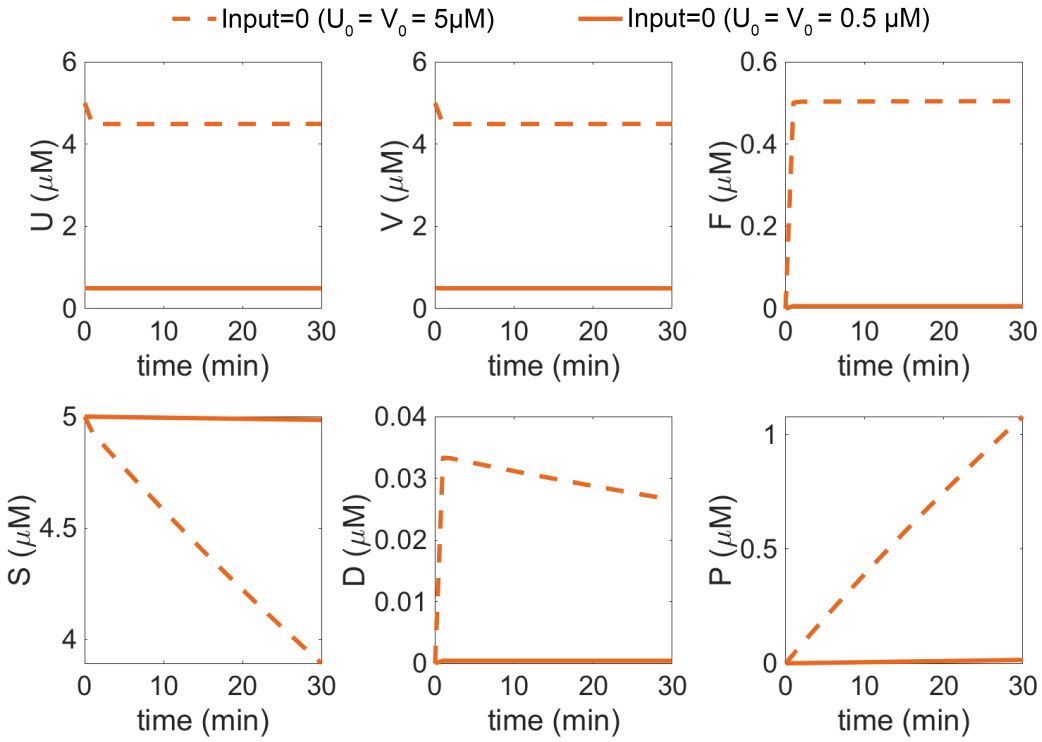


Figure S8: Simulated timing response of different species that make up the LID switch in the absence of input while keeping U_0 and V_0 fixed at 5 μM (dashed-line) or at 0.5 μM (solid-line) each. At a reduced initial concentrations of U and V, F is almost negligible. The ODE model shown in Note S2 was used to simulate the response with parameters shown in Table 1. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

Supplementary Note S3

Using the mass-action law, the kinetics of the OR gate can be modeled as:

$$\dot{A} = k_r^1 X - k_f^1 AR + k_r^4 E - k_f^4 YA - k_f^7 AB + k_r^7 E', \tag{20}$$

$$\dot{R} = k_r^1 X - k_f^1 AR + k_r^2 Y - k_f^2 BR, \tag{21}$$

$$\dot{B} = k_r^2 Y - k_f^2 BR + k_r^3 E - k_f^3 XB - k_f^7 AB + k_r^7 E', \tag{22}$$

$$\dot{X} = k_f^1 AR - k_r^1 X - k_f^3 XB + k_r^3 E - k_f^{12} FX + k_r^{12} I, \tag{23}$$

$$\dot{Y} = k_f^2 BR - k_r^2 Y - k_f^4 YA + k_r^4 E, \tag{24}$$

$$\dot{E} = k_f^4 YA + k_f^3 XB - k_r^3 E - k_r^4 E - k_f^5 ES + k_r^5 C + k_f^6 C, \tag{25}$$

$$\dot{S} = -k_f^5 ES + k_r^5 C - k_f^5 FS + k_r^5 D - k_f^5 E'S + k_r^5 C', \tag{26}$$

$$\dot{C} = k_f^5 ES - k_r^5 C - k_f^6 C, \tag{27}$$

$$\dot{U} = -k_f^8 UV + k_r^8 F, (\text{dark}) \tag{28}$$

$$\dot{V} = -k_f^8 UV + k_r^8 F, (\text{dark}) \tag{29}$$

$$\dot{U} = -k_f^8 UV + k_r^9 F, (\text{light}) \tag{30}$$

$$\dot{V} = -k_f^8 UV + k_r^9 F, (\text{light}) \tag{31}$$

$$\dot{F} = k_r^5 D + k_f^6 D - k_f^5 FS + k_f^8 UV - k_r^8 F \tag{32}$$

$$- k_f^{12} FX + k_r^{12} I + k_f^{13} I, \tag{33}$$

$$\dot{D} = k_f^5 FS - k_r^5 D - k_f^6 D, \tag{34}$$

$$\dot{P} = k_f^6 C + k_f^6 D + k_f^6 C', \tag{35}$$

$$\dot{E'} = k_f^7 AB - k_r^7 E' - k_f^5 E'S + k_r^5 C' + k_f^6 C', \tag{36}$$

$$\dot{C'} = k_f^5 E'S - k_r^5 C' - k_f^6 C'. \tag{37}$$

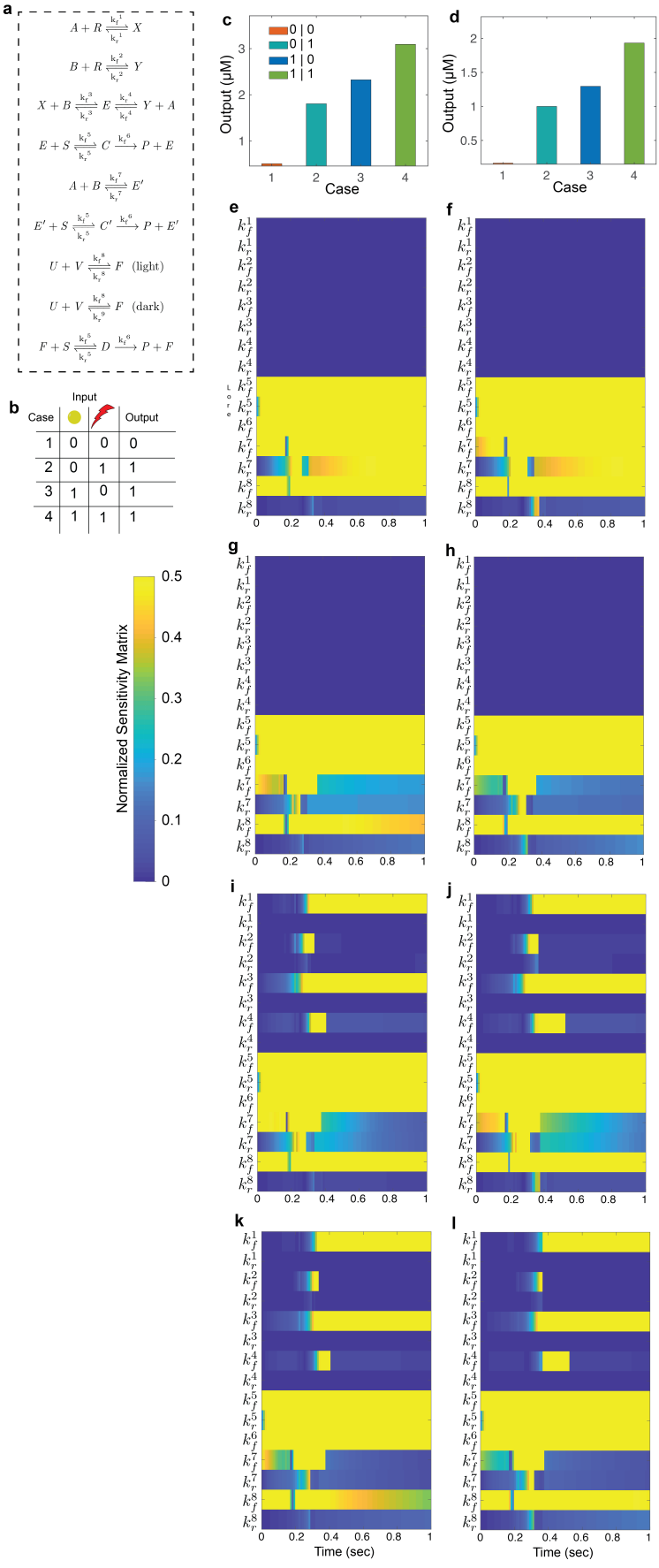


Figure S9: Protease based Boolean OR gate. (a) Chemical reactions of the OR gate and (b) the corresponding truth table. The ODE model shown in Note S3 was used to simulate the response with parameters shown in Table 1. (c-d) Simulated response of the OR gate at (c) unoptimized ($A_0 = B_0 = 1.44$ μM, $R_0 = 3.01$ μM, and $U_0 = V_0 = 1.59$ μM) and **(f)** optimized ($A_0 = B_0 = 0.72$ μM, $R_0 = 1.67$ μM, and $U_0 = V_0 = 1$ μM) conditions and the corresponding results of the sensitivity analysis are shown in (e) and (f) for case 1; in (g) and (h) for case 2; in (i) and (j) for case 3; in (k) and (l) for case 4, respectively. Normalized sensitivity matrix is shown with respect to the output (P) where yellow and blue correspond to the most sensitive and least sensitive values respectively. All the values of P were determined at 30 min. For cases 1 and 2, $R_0 = 0$ μM while in cases 2 and 4, k_r^8 was replaced with k_r^9. Here, S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

Supplementary Note S4

Using the mass-action law, the kinetics of the XOR gate can be modeled as:

$$\dot{A} = k_r^1 X - k_f^1 AR + k_r^4 E - k_f^4 YA - k_f^7 AB + k_r^7 E' - k_f^{12} FA + k_r^{12} H, \tag{38}$$

$$\dot{R} = k_r^1 X - k_f^1 AR + k_r^2 Y - k_f^2 BR, \tag{39}$$

$$\dot{B} = k_r^2 Y - k_f^2 BR + k_r^3 E - k_f^3 XB - k_f^7 AB + k_r^7 E', \tag{40}$$

$$\dot{X} = k_f^1 AR - k_r^1 X - k_f^3 XB + k_r^3 E - k_f^{12} FX + k_r^{12} I, \tag{41}$$

$$\dot{Y} = k_f^2 BR - k_r^2 Y - k_f^4 YA + k_r^4 E, \tag{42}$$

$$\dot{E} = k_f^4 YA + k_f^3 XB - k_r^3 E - k_r^4 E - k_f^5 ES + k_r^5 C + k_f^6 C \tag{43}$$

$$- k_f^{10} EU + k_r^{10} G + k_f^{11} G, \tag{44}$$

$$\dot{S} = -k_f^5 ES + k_r^5 C - k_f^5 FS + k_r^5 D - k_f^5 E'S + k_r^5 C', \tag{45}$$

$$\dot{C} = k_f^5 ES - k_r^5 C - k_f^6 C, \tag{46}$$

$$\dot{U} = -k_f^8 UV + k_r^8 F - k_f^{10} EU + k_r^{10} G, (\text{dark}) \tag{47}$$

$$\dot{V} = -k_f^8 UV + k_r^8 F, (\text{dark}) \tag{48}$$

$$\dot{U} = -k_f^8 UV + k_r^9 F - k_f^{10} EU + k_r^{10} G, (\text{light}) \tag{49}$$

$$\dot{V} = -k_f^8 UV + k_r^9 F, (\text{light}) \tag{50}$$

$$\dot{F} = k_r^5 D + k_f^6 D - k_f^5 FS + k_f^8 UV - k_r^8 F - k_f^{12} FA + k_r^{12} H + k_f^{13} H \tag{51}$$

$$- k_f^{12} FX + k_r^{12} I + k_f^{13} I, \tag{52}$$

$$\dot{D} = k_f^5 FS - k_r^5 D - k_f^6 D, \tag{53}$$

$$\dot{P} = k_f^6 C + k_f^6 D + k_f^6 C', \tag{54}$$

$$\dot{E'} = k_f^7 AB - k_r^7 E' - k_f^5 E'S + k_r^5 C' + k_f^6 C', \tag{55}$$

$$\dot{C'} = k_f^5 E'S - k_r^5 C' - k_f^6 C', \tag{56}$$

$$\dot{G} = k_f^{10} EU - k_r^{10} G - k_f^{11} G, \tag{57}$$

$$\dot{U}' = k_f^{11} G, \tag{58}$$

$$\dot{H} = k_f^{12} FA - k_r^{12} H - k_f^{13} H, \tag{59}$$

$$\dot{I} = k_f^{12} FX - k_r^{12} I - k_f^{13} I, \tag{60}$$

$$\dot{A}' = k_f^{13} H, \tag{61}$$

$$\dot{X}' = k_f^{13} I. \tag{62}$$

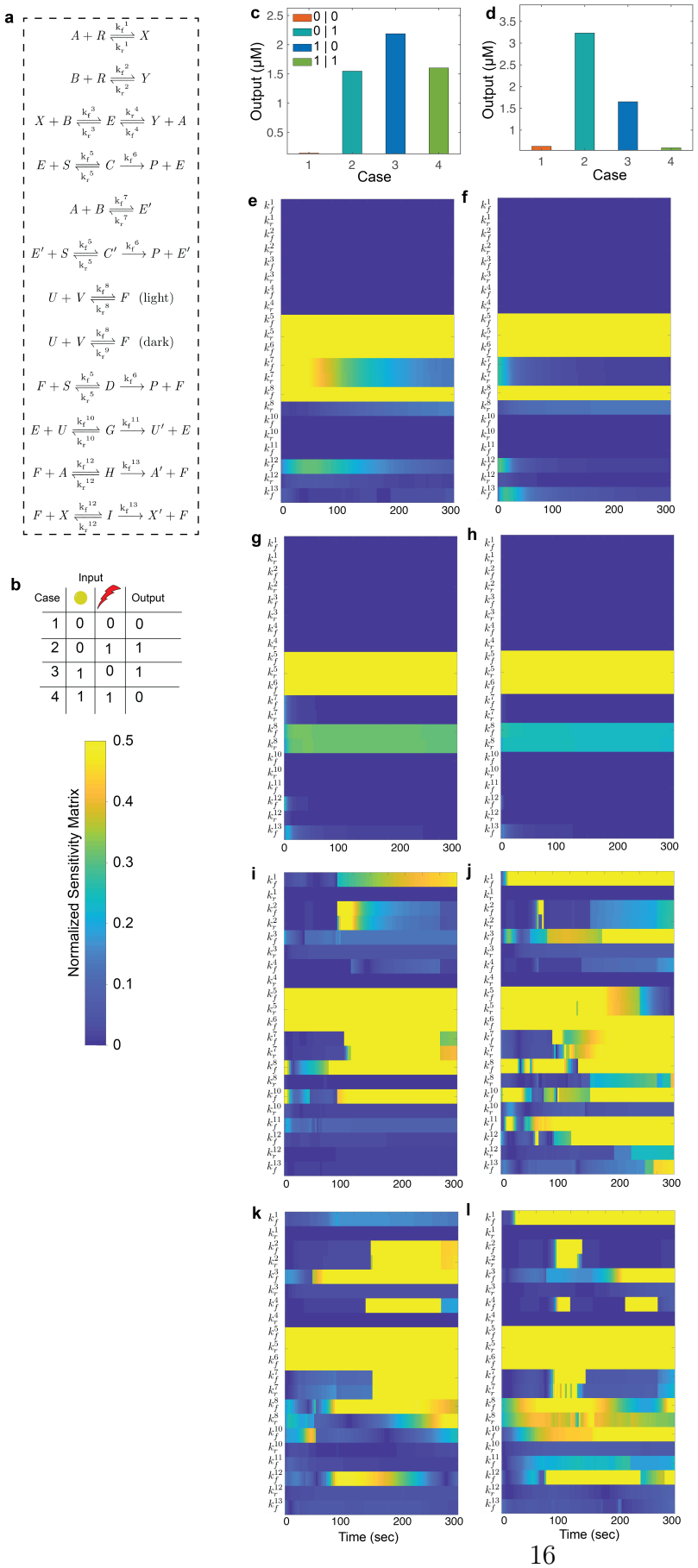
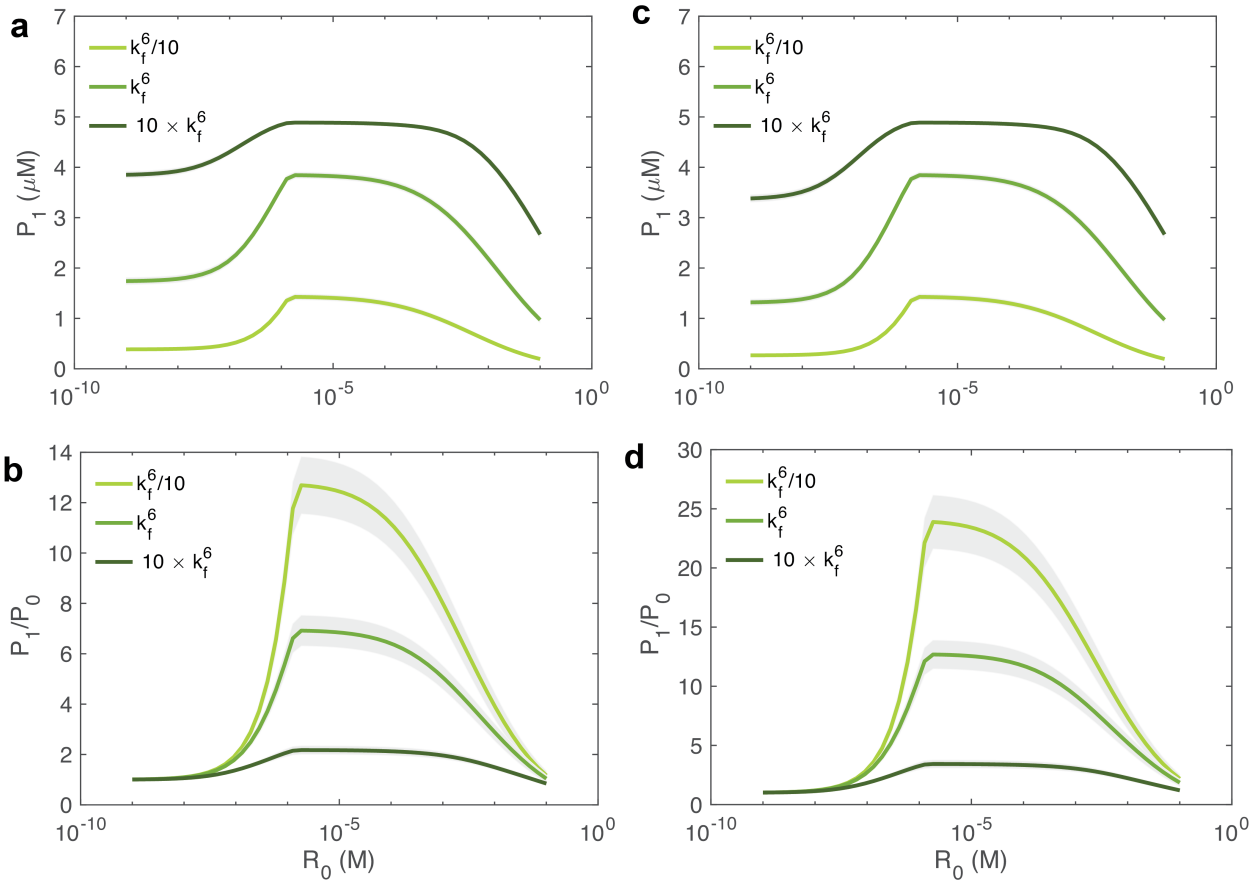


Figure S10: Protease based Boolean XOR gate. (a) Chemical reactions of the XOR gate and (b) the corresponding truth table. The ODE model shown in Note S4 was used to simulate the response with parameters shown in Table 1. (c-d) Simulated response of the OR gate at (c) unoptimized ($A_0 = B_0 = 1.44$ μM, $R_0 = 3.01$ μM, and $U_0 = V_0 = 1.59$ μM) and **(f)** optimized ($A_0 = 1.22$ μM, $B_0 = 0.88$ μM, $R_0 = 0.94$ μM, and $U_0 = V_0 = 2.74$ μM) conditions and the corresponding results of the sensitivity analysis are shown in (e) and (f) for case 1; in (g) and (h) for case 2; in (i) and (j) for case 3; in (k) and (l) for case 4, respectively. Normalized sensitivity matrix is shown with respect to the output (P) where yellow and blue correspond to the most sensitive and least sensitive values respectively. All the values of P were determined at 30 min. For cases 1 and 2, $R_0 = 0$ μM while in cases 2 and 4, k_r^8 was replaced with k_r^9. Here, S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.



e

For a-b	$k_f^6/10$	k_f^6	$10 \times k_f^6$
Range of detection (mM)	32.37	22.22	0.52
Minimum detectable amount (μM)	0.133	0.133	0.87

f

For c-d	$k_f^6/10$	k_f^6	$10 \times k_f^6$
Range of detection (mM)	99.99	68.66	15.26
Minimum detectable amount (μM)	0.063	0.063	0.19

Figure S11: Change in the catalytic proficiency of TEV protease and background binding equilibrium constant (leak reaction) result in dramatic change in the range of detection and minimum detectable concentration of rapamycin when (a-b) $k_r^7 = 20$ s^{-1} (c-d) $k_r^7 = 10$ s^{-1} and the corresponding summary are shown in (e) and (f) respectively. Here, k_f^6 is the reaction rate constant that governs the turnover rate of an enzyme-substrate complex to product and enzyme. An increase in k_f^6 increases the catalytic proficiency while reduction in k_r^7 reduces the formation of TEV protease, which happens via leak reaction. For each concentration of R_0, 1,000 simulations were conducted. Parameter values were randomly sampled from a uniform distribution (see Methods). The ODE model shown in Note S1 was used to simulate the response with parameters shown in Table 1. Averaged metrics of P_1 and the P_1/P_0 are shown here. All the values of P were determined at 30 min. The error bars are shown in the shaded region and were determined using the standard error of the mean. The prerequisite conditions were $P_1 > 1$ μM and the $P_1/P_0 > 2$. Here, A_0 and B_0 fixed at 1.44 μM each and S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

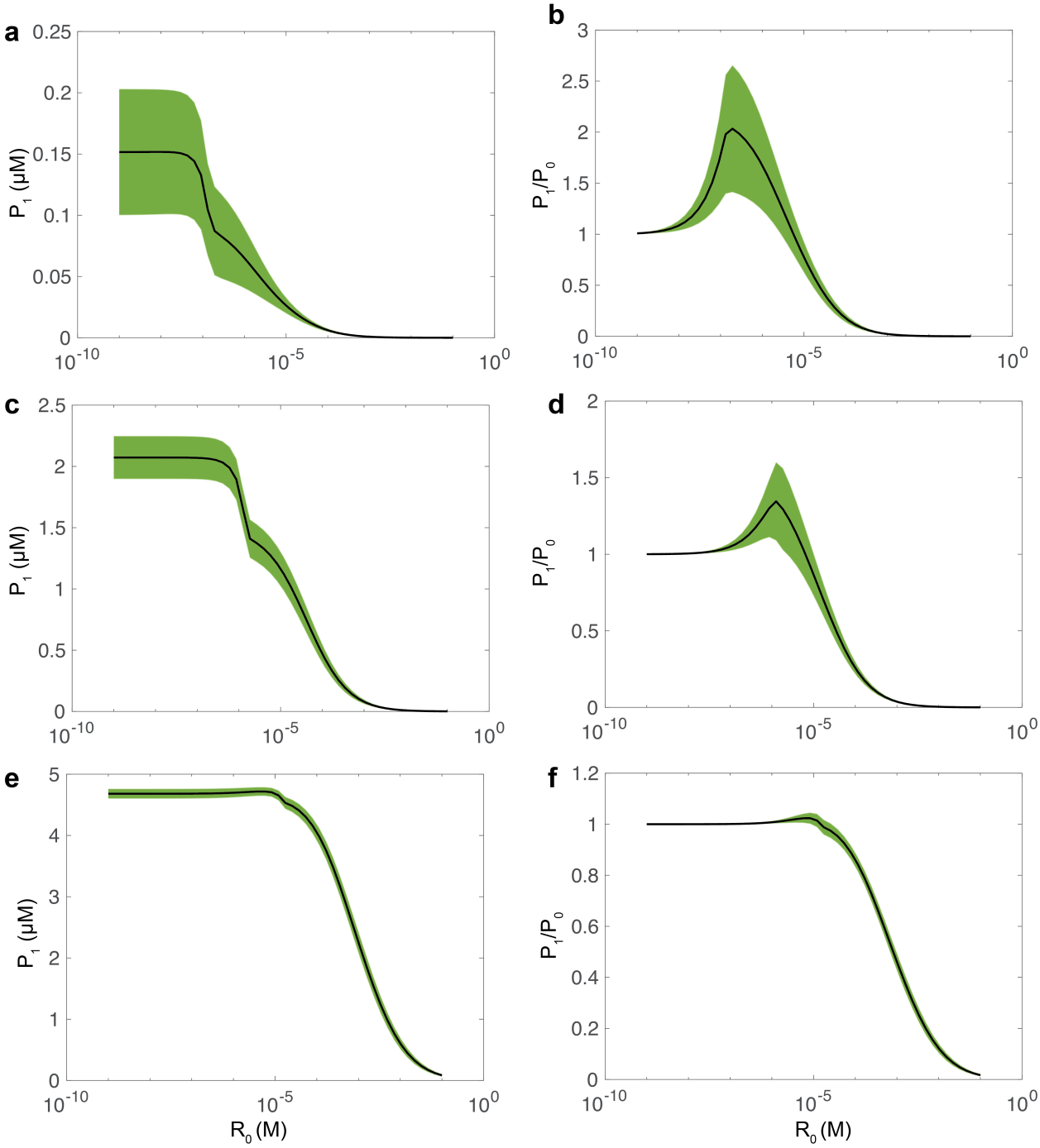


Figure S12: A non-cooperative behavior of the CID switch as a function of R_0 while (a-b) $A_0 = B_0$ =0.144 μM; (c-d) $A_0 = B_0 = 1.44$ μM; (e-f) $A_0 = B_0 = 14.4$ μM; For each concentration of R_0, 1,000 simulations were conducted. Parameter values were randomly sampled from a uniform distribution (see Methods). The ODE model shown in Note S1 was used to simulate the response with parameters shown in Table 1 except $k_f^4 = k_f^1$, $k_r^4 = k_r^1$, $k_f^3 = k_f^2$, and $k_r^3 = k_r^2$. Averaged metrics of P_1 and the P_1/P_0 are shown here. All the values of P were determined at 30 min. The error bars are shown in the shaded region and were determined using the standard error of the mean. Here S_0 was 5 μM while the rest of the molecular species were initially set to 0 μM.

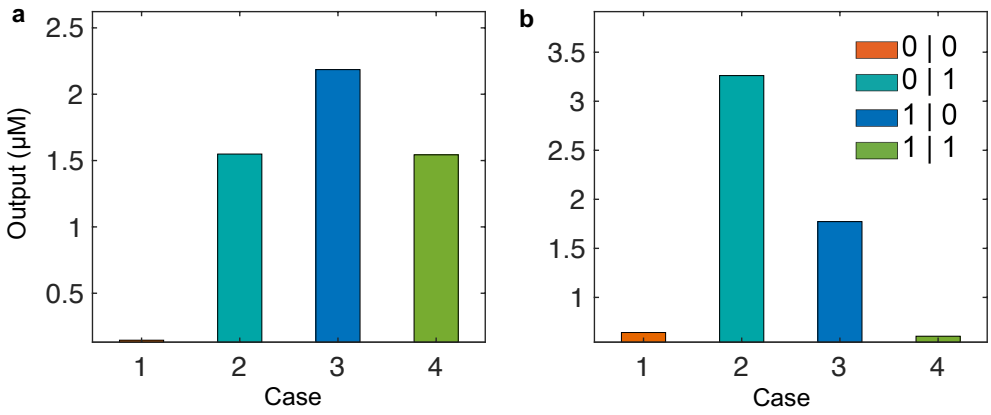


Figure S13: Simulated response of the XOR gate while considering $k_r^{10} = k_r^{12} = 0.1$ s^{-1} in (a) unoptimized ($A_0 = B_0 = 1.44\mu$M, $R_0 = 3.01$ μM, and $U_0 = V_0 = 1.59$ μM) and (b) optimized ($A_0 = 1.69$ μM, $B_0 = 1.13$ μM, $R_0 = 1.58$ μM, and $U_0 = V_0 = 3.67$ μM) conditions. All the values of P (output) were determined at 30 min. We model the kinetics of the XOR gate using an ODE model (see Note S4) with parameters shown in Table 1 except k_r^{10} and k_r^{12} values.

Table S2: The primers (F-LB-forXOR) were designed for linearization of the modified pet15b vector. The primers (F-Amp-XORInserts) were designed to amplify the FKBP/FRB split TEV insert dsDNA fragments (FRB-SbMVcut-TEVn-pet15 and FKBP-SbMVcut-TEVc-pet15) for gibson into the modified pet15b vector.

Oligo	Amplification on which component	Primer Sequence
F-LB-forXOR	vector backbone	TAG CTG GAA CAT CAT CAT CAT CAT CAT TGA
R-LB-forXOR	vector backbone	AGT CTG CGC GTC TTT CAG GG
F-Amp-XORInserts	FKBP/FRB-split TEV dsDNA fragments	CC CTG AAA GAC GCG CAG ACT
R_Amp_XORInserts	TEV dsDNA fragments	TCA ATG ATG ATG ATG ATG ATG TTC CAG CTA
FRB-SbMVcut-TEVn-pet15	dsDNA fragments	GTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAA CAACAACAGTGGACTGGTGCCGCGCGGCAGCTCTGGTGGTGGTGGCGGTTCCATGTGGCATGA AGGATTAGAAGAAGCGTCACGCCTGTATTTCGGAGAACGCAACGTAAAGGGAATGTTTGAAGTA CTGGAACCGCTCCATGCGATGATGGAACGCGGACCGCAAACCCTGAAGGAAACCAGCTTCAACC AAGCGTATGGCCGCGATCTCATGGAAGCGCAAGAATGGTGTCGCAAATATATGAAGAGCGGTAA CGTAAAGGATCTGACCCAAGCGTGGGATCTGTATTATCATGTATTTCGCCGCATTTCAAAGCAAG GTGGTGGCGGTTCCGAAAGCGTGTCGTTACAGTCGGGTGGTGGCGGTTCCGGAGAAAGCTTGT TTAAGGGGCCGCGTGATTACAACCCGATATCGAGCACCATTTGTCATTTGACGAATGAATCTGAT GGGCACACAACATCGTTGTATGGTATTGGATTTGGTCCCTTCATCATTACAAACAAGCACTTGTT TAGAAGAAATAATGGAACACTGTTGGTCCAATCACTACATGGTGTATTCAAGGTCAAGAACACCA CGACTTTGCAACAACACCTCATTGATGGGAGGGACATGATAATTATTCGCATGCCTAAGGATTTC CCACCATTTCCTCAAAAGCTGAAATTTAGAGAGCCACAAAGGGAAGAGCGCATATGTCTTGTGAC AACCAACTTCCAAACTTAGCTGGAACATCATCATCATCATCATTGA
FKBP-SbMVcut-TEVc-pet15	dsDNA fragments	AATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACAGTGGACTGGTGCCGCGCGGCAG CTCTGGTGGTGGTGGCGGTTCCATGAAAGGCGTGCAGGTGGAGACTATCTCCCCAGGAGACGG GCGCACCTTCCCCAAGCGCGGCCAGACCTGCGTGGTGCACTACACCGGGATGCTTGAAGATGG AAAGAAATTTGATTCCTCCCGGGACAGAAACAAGCCCTTTAAGTTTATGCTAGGCAAGCAGGAGG TGATCCGAGGCTGGGAAGAAGGGGTTGCCCAGATGAGTGTGGGTCAGAGAGCCAAACTGACTA TATCTCCAGATTATGCCTATGGTGCCACTGGGCACCCAGGCATCATCCCACCACATGCCACTCTC GTCTTCGATGTGGAGCTTCTAAAACTGGAAGGTGGTGGCGGTTCCGAAAGCGTGTCGTTACAGT CGGGTGGTGGCGGTTCCAAGAGCATGTCTAGCATGGTGTCAGACACTAGTTGCACATTCCCTTC ATCTGATGGCATATTCTGGAAGCATTGGATTCAAACCAAGGATGGGCAGTGTGGCAGTCCATTA GTATCAACTAGAGATGGGTTCATTGTTGGTATACACTCAGCATCGAATTTCACCAACACAAACAAT TATTTCACAAGCGTGCCGAAAAACTTCATGGAATTGTTGACAAATCAGGAGGCGCAGCAGTGGG TTAGTGGTTGGCGATTAAATGCTGACTCAGTATTGTGGGGGGGCCATAAAGTTTTCATGGTGAA ACCTGAAGAGCCTTTTCAGCCAGTTAAGGAAGCGACTCAACTCATGAATCGTCGTCGCCGTCGC TAGCTGGAACATCATCATCATCATCATTGA

Table S3: These are the proteins sequence were used in the FKBP/FRB-split TEV fluorescence assay. Each construct has a N-terminal 8-histadine tag, followed by a maltose binding protein, and finally a thrombin cutsite. After this thrombin cutsite there is the FKBP or FRB domain, followed by a linker and terminating with the appropriate split TEV fragment.

Protein Construct	1-Letter Amino Acid Sequence
MBP-FRB-TEVn	MGHHHHHHHHHASKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGD GPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNP PKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLT FLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGV LSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIM PNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNSGLVPRGSSGGGGGS MWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYM KSGNVKDLTQAWDLYYHVFRRISKQGGGGSESVSLQSGGGGSGESLFKGPRDYNPISSTICHLTNE SDGHTTSLYGIGFGPFIITNKHLFRRNNGTLLVQSLHGVFKVKNTTTLQQHLIDGRDMIIIRMPKDFPPF PQKLKFREPQREERICLVTTNFQT
MBP-FKBP-TEVc	MGHHHHHHHHHASKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGD GPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNP PKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLT FLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGV LSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIM PNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNSGLVPRGSSGGGGGS MKGVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKFDSSRDRNKPFKFMLGKQEVIRGWEEGV AQMSVGQRAKLTISPDYAYGATGHPGIIPPHATLVFDVELLKLEGGGGSESVSLQSGGGGSKSMSSM VSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGFIVGIHSASNFTNTNNYFTSVPKNFMELLTN QEAQQWVSGWRLNADSVLWGGHKVFMVKPEEPFQPVKEATQLMNRRRRR