

In the format provided by the authors and unedited.

Modular and tunable biological feedback control using a de novo protein switch

Andrew H. Ng^{1,2,3,4}, Taylor H. Nguyen¹, Mariana Gómez-Schiavon¹, Galen Dods¹, Robert A. Langan^{5,6,7}, Scott E. Boyken^{5,6}, Jennifer A. Samson², Lucas M. Waldburger², John E. Dueber², David Baker^{5,6,8} & Hana El-Samad^{1,4,9*}

¹Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA, USA. ²Department of Bioengineering, University of California, Berkeley, Berkeley, CA, USA. ³The UC Berkeley-UCSF Graduate Program in Bioengineering, UC Berkeley, Berkeley, CA, USA. ⁴Cell Design Initiative, University of California, San Francisco, CA, USA. ⁵Department of Biochemistry, University of Washington, Seattle, WA, USA. ⁶Institute for Protein Design, University of Washington, Seattle, WA, USA. ⁷Graduate Program in Biological Physics, Structure, and Design, University of Washington, Seattle, WA, USA. ⁸Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. ⁹Chan-Zuckerberg Biohub, San Francisco, CA, USA. *e-mail: hana.el-samad@ucsf.edu

Modular and tunable biological feedback control using a *de novo* protein switch SUPPLEMENTARY INFORMATION

1 Model description

We describe below the computational model used to generate all plots in the manuscript.

1.1 Species

- G : degSwitch-transcription factor (e.g. degSwitch-GEM)
- Z : Transcription factor (e.g. ZPM)
- K : Key
- C : Key-degSwitch-transcription factor complex
- Y_* : "Immature" output (e.g. immature YFP)
- Y : Output (e.g. mature YFP)

Note: For simplicity and readability, in the model analysis and figures we omit the square brackets when referring to the concentration of these species.

1.2 Parameters

- μ_G ([nM/min]) : Constitutive synthesis rate of G
- γ_G ([1/min]) : Degradation/loss rate of free G (i.e. leaky degradation)
- η_+ ([1/(nM min)]) : Binding rate of G and K
- η_0 ([1/min]) : Unbinding rate of G and K
- η_- ([1/min]) : Active degradation rate of G in the complex form (i.e. $C \rightarrow K$)
- $f_Z(G + C, E)$ ([nM/min]) : Synthesis rate of Z regulated by G and E (e.g. $\mu_Z \cdot (G + C) \cdot E$)
- E ([nM]) : Co-factor of G to regulate Z synthesis (e.g. oestradiol)
- γ_Z ([1/min]) : Degradation/loss rate of Z
- $f_K(X, P)$ ([nM/min]) : Synthesis rate of K regulated by Z and P (e.g. $\mu_K \cdot X \cdot P$)
- P ([nM]) : Co-factor of Z to regulate K synthesis (e.g. progesterone)
- γ_K ([1/min]) : Degradation/loss rate of K and C
- $f_Y(X, P)$ ([nM/min]) : Synthesis rate of Y regulated by Z and P (e.g. $\mu_Y \cdot X \cdot P$)
- γ_Y ([1/min]) : Degradation/loss rate of Y
- κ_+ ([1/min]) : Maturation rate of Y_*

1.3 ODE system

$$\frac{d}{dt}G = \mu_G - \gamma_G G - \eta_+ GK + \eta_0 C \quad (1)$$

$$\frac{d}{dt}Z = f_Z(G + C, E) - \gamma_Z Z \quad (2)$$

$$\frac{d}{dt}K = f_K(Z, P) - \gamma_K K - \eta_+ GK + (\eta_0 + \eta_-)C \quad (3)$$

$$\frac{d}{dt}C = -\gamma_C C + \eta_+ GK - (\eta_0 + \eta_-)C \quad (4)$$

$$\frac{d}{dt}Y_* = f_Y(Z, P) - \gamma_Y Y_* - \kappa_+ Y_* \quad (5)$$

$$\frac{d}{dt}Y = \kappa_+ Y_* - \gamma_Y Y \quad (6)$$

2 Qualitative properties

Notice that in this system (Eqs. 1-5), total K concentration (i.e. $K_T = K + C$) and output Y_T are given by:

$$\frac{d}{dt}K_T = \frac{d}{dt}(K + C) = f_K(Z, P) - \gamma_K(K + C) \quad (7)$$

$$\frac{d}{dt}Y_T = \frac{d}{dt}(Y_* + Y) = f_Y(Z, P) - \gamma_Y(Y_* + Y) \quad (8)$$

assuming $f_K(Z, P)$ and $f_Y(Z, P)$ have similar qualitative form (e.g. $f_K(Z, P) \propto f_Y(Z, P)$), then K_T and Y_T have analogous functional form.

At steady state :

$$\frac{d}{dt}G = 0 \Leftrightarrow G_{ss} = \frac{\mu_G + \eta_0 C_{ss}}{\gamma_G + \eta_+ K_{ss}} \quad (9)$$

$$\begin{aligned} \frac{d}{dt}C = 0 &\Leftrightarrow C_{ss} = \frac{\eta_+ G_{ss} K_{ss}}{\gamma_C + \eta_0 + \eta_-} \\ &\Rightarrow C_{ss} = \frac{\mu_G K_{ss}}{\gamma_G \left(\frac{\gamma_C + \eta_0 + \eta_-}{\eta_+} \right) + K_{ss}(\gamma_C + \eta_-)} \end{aligned} \quad (10)$$

Then, when $K_{ss} \gg \frac{\gamma_G(\gamma_C + \eta_0 + \eta_-)}{\eta_+(\gamma_C + \eta_-)}$, the $C_{ss} \approx \frac{\mu_G}{\gamma_C + \eta_-}$. This occurs regardless of the specific form of $\frac{d}{dt}K$, and in particular the presence or absence of feedback. This is actually an important limit in the control system, as the feedback action occurs through and only through C formation, and this is a required step for active degradation –and then the feedback– to occur. For example, consider that $f_K(z, P)$ is an increasing function of Z and P and consider a positive disturbance implemented through an increase in P : initially, as P increases, K synthesis increase, potentially increasing C and then the effective active degradation of G . Once G decreases, the synthesis of Z will decrease accordingly, decreasing then K synthesis, i.e. “compensating” for the increment on P . Nevertheless, if C has reached its maximum value ($\lim_{K \rightarrow \infty} C = \frac{\mu_G}{\gamma_C + \eta_-}$), increasing K does not have an effect on G degradation, and the feedback is effectively broken.

2.1 Basal activity & saturation on synthesis functions

We incorporate the complexity of the synthesis process by using a Hill-type function as the synthesis function for each of the regulated genes:

$$f_Z(G + C, E) = \mu_Z \left(\alpha_Z + (1 - \alpha_Z) \frac{(E \cdot (G + C))^{n_Z}}{(E \cdot (G + C))^{n_Z} + K_Z^{n_Z}} \right) \quad (11)$$

$$f_K(Z, P) = \mu_K \left(\alpha_K + (1 - \alpha_K) \frac{(P \cdot Z)^{n_K}}{(P \cdot Z)^{n_K} + K_K^{n_K}} \right) \quad (12)$$

$$f_Y(Z, P) = \mu_Y \left(\alpha_Y + (1 - \alpha_Y) \frac{(P \cdot Z)^{n_Y}}{(P \cdot Z)^{n_Y} + K_Y^{n_Y}} \right) \quad (13)$$

where $\{\mu_Z, \mu_K, \mu_Y\}$ represent the maximum synthesis rates, $\{\alpha_Z\mu_Z, \alpha_K\mu_K, \alpha_Y\mu_Y\}$ the basal synthesis rates, $\{n_Z, n_K, n_Y\}$ the Hill coefficients, and $\{K_Z, K_K, K_Y\}$ the activation thresholds relative to the active regulators (i.e. transcription factor and co-factor complex) for $\{Z, K, Y\}$ synthesis rate functions, respectively. We assume that the Hill coefficient and activation threshold are the same for K and Y synthesis function, as they are both regulated by Z and P ; the maximum and basal synthesis rate are expected to depend more strongly on the gene sequence, and then we allow them to differ. When the feedback is removed, $f_K = \mu_{K^*}$, where μ_{K^*} represents the constitutive synthesis rate of K .

The Hill function is often used as a phenomenological description of gene regulation. Nevertheless, regardless of the specific form of $\{f_Z, f_K, f_Y\}$ functions, a basal and a maximum synthesis rate are expected for every gene. The functions in Eqs. 11-13 allow us to explore the effect of these limits.

Extended Data Figure 3a shows an example of the steady state results for different values of P using this model (Eqs. 11-13). As expected, the feedback is active only when C is neither too small nor too high. We define the feedback as “active” whenever the relative change in total G ($G_T = G + C$) over the relative change of the P -disturbance is higher than an arbitrary threshold (e.g. $\epsilon = 0.15$):

$$\frac{(\Delta G_T)/G_T}{(\Delta P)/P} \geq \epsilon \quad (14)$$

Feedback directly changes the amount of G_T , thus this metric is useful because the amount of feedback is correlated to the sensitivity of G_T to the P -disturbance. The value of ΔG_T as well as the metric in Eq. 14 are equal to zero in a system without feedback.

Similar to perturbing the system by increasing the concentration of P , it is possible to test the feedback action by perturbing the degradation rate of Z (see *Figure 3b* in the main text). *Extended Data Figure 3b* shows that this “negative” perturbation has a similar response compared to the previous “positive” perturbation (*Extended Data Figure 3a*), just changing the direction of the effect of the perturbation (i.e. Z increases as γ_Z increases, to “compensate” for the faster degradation of X). Once again, the feedback control is active (substituting $\Delta(P)/P$ by $\Delta(\gamma_Z)/\gamma_Z$ in Eq. 14) when C has not reach saturation and its concentration is high enough to contribute significantly to the degradation of G_T .

2.2 Tuning the feedback efficiency

Extended Data Figure 7 shows a few examples of how steady state solution and feedback “activity” regime varies as either η_+ or μ_K change. In both cases, decreasing the parameter value initially shifts the feedback “activity” regime towards higher P -disturbance values (*Extended Data Figure 7c-d*), and eventually breaks the feedback reducing both the range and magnitude of the feedback activity (as defined by Eq. 14). Nevertheless, for high values of η_+ , the sensitivity of G_T increases dramatically, compared to high values of μ_K , where the feedback actually saturates and eventually breaks. Therefore, systems with higher η_+ can potentially result in a stronger and more efficient feedback. Nevertheless, *Extended Data Figure 7a* shows an example where the feedback activity occurs where the output Y expression is very close to its basal value, making the benefit of this increment of η_+ almost imperceptible (compare black and purple continuous lines in *Extended Data Figure 7a*). In this case, decreasing the basal expression of Y would be enough to reveal the feedback action on this system. This example highlights the advantage of defining this model to do a full tuning of the experimental system.

Table : **Simulation parameter values in Figures & Extended Data Figures**

	μ_G	γ_G	η_+	η_0	η_-	μ_Z	α_Z	n_Z	K_Z	E	γ_Z
Fig.	nM min ⁻¹	min ⁻¹	nM ⁻¹ min ⁻¹	min ⁻¹	min ⁻¹	nM min ⁻¹			nM	nM	min ⁻¹
3b	0.006	0.02	0.0375	1E-4	0.05	11	1E-6	2.2	36	7.5	0.01
4a	0.006	0.02	0.0375*	1E-4	0.05	11	1E-6	2.2	36	7.5	0.01
E3a	0.006	0.02	0.0375	1E-4	0.05	11	1E-6	2.2	36	7.5	0.01
E3b	0.006	0.02	0.0375	1E-4	0.05	11	1E-6	2.2	36	30	0.01
E7a,c	0.006	0.02	0.0375 [†]	1E-4	0.05	11	1E-6	2.2	36	7.5	0.01
E7b,d	0.006	0.02	0.0375	1E-4	0.05	11	1E-6	2.2	36	7.5	0.01
	μ_K	α_K	n_K	K_K	P	γ_K	μ_{K*}^{\ddagger}	μ_Y	α_Y	γ_Y	κ_+
Fig.	nM min ⁻¹			nM	nM	min ⁻¹	nM min ⁻¹	nM min ⁻¹		min ⁻¹	min ⁻¹
3b	2	1E-5	2.6	12	[0.8, 3] [†]	0.01	0.0028	0.75	0.03	0.05	0.01
4a	2*	1E-5	2.6	12	... [†]	0.01	5E-4	0.75	0.03	0.05	0.01
E3a	2	1E-5	2.6	12	... [†]	0.01	0.0028	0.75	0.03	0.05	0.01
E3b	2	1E-5	2.6	12	1.57	0.01	0.28	0.75	0.03	0.05	0.01
E7a,c	2	1E-5	2.6	12	... [†]	0.01	5E-4	0.75	0.03	0.05	0.01
E7b,d	2 [†]	1E-5	2.6	12	... [†]	0.01	5E-4	0.75	0.03	0.05	0.01

[‡] “No feedback” system.

[†] Unless directly perturbed.

* When η_+ is “tuned”: $\eta_+ = [0.01125, 0.001125] \text{ nM}^{-1} \text{ min}^{-1}$;

when μ_K is “tuned”: $\mu_K = [0.2, 0.02] \text{ nM min}^{-1}$.

NOTE: E# refers to Extended Data Figure #.