ChemPubSoc Europe

DOI: 10.1002/cbic.201800007

CHEMBIOCHEM Communications

SPECIAL ISSUE

A Miniaturized *Escherichia coli* Green Light Sensor with High Dynamic Range

Nicholas T. Ong[a] and Jeffrey J. Tabor*[a, b]

Genetically engineered photoreceptors enable unrivaled control over gene expression. Previously, we ported the *Synechocystis* PCC 6803 CcaSR two-component system, which is activated by green light and deactivated by red, into *Escherichia coli*, resulting in a sensor with a sixfold dynamic range. Later, we optimized pathway protein expression levels and the output promoter sequence to decrease transcriptional leakiness and to increase the dynamic range to approximately 120-fold. These CcaSR v1.0 and v2.0 systems have been used for precise quantitative, temporal, and spatial control of gene expression for a variety of applications. Recently, other workers deleted two PAS domains of unknown function from the CcaS sensor histidine kinase in a system similar to CcaSR v1.0. Here we apply these deletions to CcaSR v2.0, resulting in a v3.0 light sensor with an output four times less leaky and a dynamic range of nearly 600-fold. We demonstrate that the PAS domain deletions have no deleterious effect on CcaSR green light sensitivity or response dynamics. CcaSR v3.0 is the best-performing engineered bacterial green light sensor available, and should have broad applications in fundamental and synthetic biology studies.

CcaSR is a green-light-activated/red-light-repressed two-component signal transduction system (TCS) that controls expression of the phycobilisome linker protein CpcG2 in the *Synechocystis* PCC6803 complementary chromatic adaptation pathway.[1] CcaS is a cyanobacteriochrome sensor histidine kinase (SK) consisting of a putative N-terminal transmembrane region, a photosensing cGMP phosphodiesterase/adenyl cyclase/FhlA (GAF) domain that covalently binds the chromophore phycocyanobilin (PCB), two Per/Arnt/Sim (PAS) domains of unknown function, and a C-terminal histidine kinase (HisK) domain (Figure 1 A).[2] Holo-CcaS is produced in a green-absorbing ground state with low kinase activity. Under green light, CcaS switches to a red-absorbing conformation that phosphorylates the response regulator (RR) CcaR. Phosphorylated CcaR binds to the P_{cpcG2} output promoter, where it activates transcription.[2]

[a] *Dr. N. T. Ong, Dr. J. J. Tabor*
Department of Bioengineering, Rice University
6100 Main Street, Houston, TX 77005 (USA)
E-mail: jeff.tabor@rice.edu

[b] *Dr. J. J. Tabor*
Department of Biosciences, Rice University
6100 Main Street, Houston, TX 77005 (USA)

Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201800007.

This article is part of a Special Issue on the Optical Control of Biological Processes.

To construct CcaSR v1.0, we cloned the *Synechocystis* PCC 6803 *ccaS-ccaR-cpcG2* genomic region into an *Escherichia coli* plasmid and replaced *cpcG2* with a reporter gene.[3] We co-transformed this plasmid with a second encoding *Synechocystis* PCC6803 *ho1* and *pcyA*, which encode a heme oxygenase and phycocyanobilin/ferredoxin oxidoreductase that catalyze the conversion of heme into PCB.[4] The resulting system exhibits some leaky transcriptional output in red light, and sixfold activation by green.[5] Despite these relatively modest features, CcaSR v1.0 has been used to perform bacterial photography,[3] to program tailor-made gene expression signals,[6] to characterize the input/output dynamics[6] and adaptive noise-filtering properties[7] of synthetic gene circuits, to control cellular growth rate by modulating expression of a metabolic enzyme,[8] and to implement in silico feedback control over gene expression.[9]

To construct CcaSR v2.0, we utilized libraries of synthetic promoters and ribosome binding sites to optimize the expression levels of CcaS, CcaR, Ho1, and PcyA while removing a putative unregulated transcriptional start site within P_{cpcG2}.[5] CcaSR v2.0 exhibits 17 times less leakiness than v1.0, and 117-fold activation.[5] We recently combined CcaSR v2.0 with an engineered red light sensor to program two gene expression signals simultaneously and independently,[10] whereas Guet and co-workers used it to extend in silico control of gene expression to the single-cell level.[11]

Recently, several blue-light-activated bacterial photoreceptors have been reported to exhibit very low leakiness and between 300- and 10000-fold activation.[12–14] However, blue light has higher phototoxicity than green, and these photoreceptors suffer from wide absorbance spectra and slower response dynamics relative to CcaSR. We therefore sought to improve the dynamic range (ratio of output gene expression in green light to that in red light) of CcaSR v2.0.

Recently, Sode and co-workers constructed a series of miniaturized CcaS variants (hereafter mini-CcaSs) lacking the two PAS domains.[15] In particular, they deleted the amino acids between Q221, which follows the GAF domain, and E504–A514, a linker region upstream of the HK domain. In a plasmid setup similar to our CcaSR v1.0 system, two variants, named mini-CcaS #3 and #10 (Figure 1 A), yielded slightly lower gene expression output than full-length CcaS in green light, combined with lower leakiness under red.[15]

Inspired by these results, we hypothesized that substituting CcaS for mini-CcaS #3 or #10 might increase the dynamic range of CcaSR v2.0. To examine this possibility, we introduced the corresponding deletions into plasmid pSR43.6,[5] resulting in plasmids pNO286-1 and pNO286-3 (Figure 1 A and Figure S1 in the Supporting Information). Next, we co-transformed each

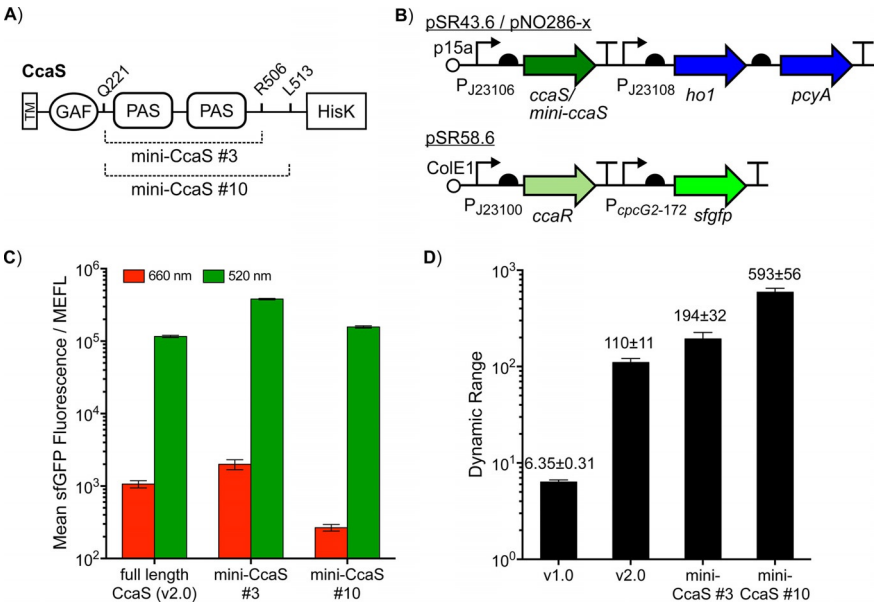

Figure 1. Engineering CcaSR v3.0. A) CcaS domain truncations. The two PAS domains in full-length CcaS were deleted, and the remaining N-terminal CcaS fragment ending at residue Q221 was directly fused to the remaining C-terminal CcaS fragments starting at R506 or L513 to build miniaturized CcaS constructs mini-CcaS #3 and mini-CcaS #10, respectively.[15] B) Device schematics for CcaSR systems tested in this paper. CcaSR v2.0[5] uses two plasmids—pSR43.6 and pSR58.6—for constitutive expression of *ccaS*, *ccaR*, and an optimized metabolic pathway for PCB production (*ho1* and *pcyA*), whereas the sensor uses $P_{cpcG2\text{-}172}$ as a transcriptional output for sfGFP reporter expression. Here, the same plasmid architecture for testing mini-CcaS constructs is retained. The domain truncations described in Figure 1 A are applied to *ccaS* on pSR43.6 to build two separate plasmids—pNO286-1 and pNO286-3 (collectively denoted as pNO286-*x*)—for expression of genes encoding mini-CcaS #3 and #10, respectively. C) sfGFP response of CcaS sensors to saturating green (520 nm) and red (660 nm) light. Bar heights and error bars represent the means and SDs, respectively, of $n=3$ experiments run on separate days. D) Dynamic range of CcaS sensor sfGFP responses to green/red light; the previously reported dynamic range of the v1.0 sensor[5] is shown for comparison.

of the three plasmids pairwise with pSR58.6, which encodes *ccaR* and a superfolder green fluorescent protein (*sfGFP*) reporter (Figures 1 B and S1).[5] We then utilized flow cytometry to measure sfGFP expression levels under green and red light, following our previous protocols (see the Experimental Section).[6,16] We observed a CcaSR v2.0 output of $(1.06\pm0.12)\times10^3$ molecules of equivalent fluorescein (MEFL) sfGFP under red light and $(1.163\pm0.042)\times10^5$ under green (Figure 1 C). These data constitute a (110 ± 11)-fold (Figure 1 D) dynamic range, consistent with our previous measurements.[5] Mini-CcaS #3 results in higher sfGFP expression levels under both red [$(2.00\pm0.31)\times10^3$ MEFL] and green [$(3.806\pm0.089)\times10^5$ MEFL] light (Figure 1 C), and an improved dynamic range of (194 ± 32)-fold (Figure 1 D). Mini-CcaS #10 simultaneously yields a (3.99 ± 0.26)-fold reduction in red light output [$(2.67\pm0.27)\times10^2$ MEFL] and higher green output [$(1.574\pm0.053)\times10^5$ MEFL] than full-length CcaS (Figure 1 C), resulting in a (593 ± 56)-fold dynamic range (Figure 1 D). Although the results we observe in the v2.0 context differ qualitatively from those reported by Sode and co-workers,[15] the truncations do decrease leakiness for mini-CcaS #10 and increase dynamic range in both cases. Because mini-CcaS #10 exhibits the lowest leakiness and highest dynamic range of all green sensors we have built (Figure S2), we renamed the corresponding construct "CcaSR v3.0".

To validate that the PAS deletions have no deleterious effect on sensor performance, we went on to characterize the input/output properties of CcaSR v3.0 in more detail. Firstly, we compared the steady-state transfer functions—or relationships between green light intensity and sfGFP expression level—of CcaSR v2.0 and v3.0. As expected, gene expression output increases sigmoidally with green light intensity for both systems (Figure 2). Interestingly, CcaSR v3.0 exhibits a slightly steeper transfer function, with a Hill coefficient of 2.737 ± 0.044, in comparison with 2.114 ± 0.052 for the v2.0 system (Table S1). Although outside the scope of this study, future biophysical experiments could reveal the origins of this increase. On the other hand, CcaSR v3.0 requires a photon flux of (1.075 ± 0.025) µmol m^{-2} s^{-1} 520 nm light for 50% maximal activation, which is similar to the value for CcaSR v2.0 [$(0.793\pm$

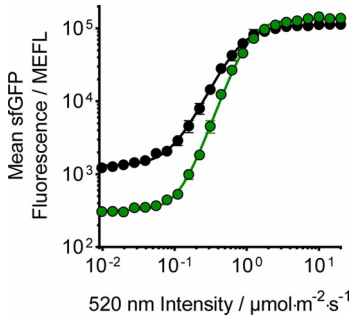

Figure 2. CcaSR v2.0 and v3.0 steady-state transfer functions. Markers (v2.0 black, v3.0 green) and error bars represent the means and SDs, respectively, of $n=3$ experiments run on separate days. Solid lines represent the Hill function fit. Fit parameters are shown in Table S1.

0.028) μmol m^{-2} s^{-1}, Table S1]. This result suggests that the sensitivity of CcaS to photons is unaltered by the PAS domain deletions.

Finally, we compared the response dynamics of CcaSR v3.0 to those of the v2.0 system. Specifically, we performed step activation and deactivation experiments in which bacteria were optically preconditioned to the high or low sfGFP expression state in saturating green or red light, the light was switched to the opposite condition at time zero, and sfGFP levels were measured over time (see the Experimental Section). We observed no substantial difference in the response dynamics of CcaSR v3.0 relative to its predecessor (Figure 3).

Whereas PAS domains have been shown to mediate light sensing in phototropins[17] and protein–protein interactions between plant phytochrome dimers,[18] the role of the CcaS PAS

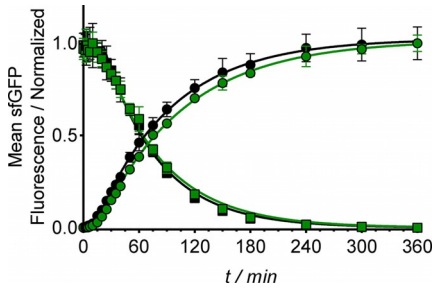

Figure 3. CcaSR v2.0 and v3.0 step response dynamics. Response to step changes in light from 660 to 520 nm (step ON, circles) and 520 to 660 nm (step OFF, squares). Markers (v2.0 black, v3.0 green) and error bars represent the means and SDs, respectively, of $n = 3$ experiments run on separate days. Lines represent model fits (see the Experimental Section).

domains is unknown. Firstly, the purified CcaS GAF domain binds PCB and undergoes the same photocycle as a nearly full-length (missing the 23-residue N-terminal transmembrane domain) PCB-bound version of CcaS in vitro.[2] Combined with the fact that the PAS domain deletions do not affect CcaSR green light sensitivity[15] (Figure 2) or response dynamics (Figure 3), these results suggest that the PAS domains are not involved in CcaS photoreception. Additionally, on the basis of available literature, the CcaS PAS domains are not known to mediate protein–protein interactions. It is possible that the CcaS PAS domains mediate undiscovered protein signaling interactions, or are even involved in ligand binding—other PAS domains are known to bind to light-absorbing ligands such as hemes, flavins, or cinnamic acid.[19] Future studies, such as genetic screens in cyanobacteria, will be enlightening.

Here we have utilized the CcaS PAS domain truncations reported by Sode and co-workers[15] to construct a third-generation CcaSR *E. coli* green light sensor with four times lower leakiness than our second-generation system and nearly 600-fold dynamic range. Both of these performance features are superior to those of any previously reported bacterial green light sensor. Furthermore, CcaSR v3.0 approaches the dynamic range of the best-performing bacterial (blue) light sensors in the literature. We have achieved these improvements without compromising the high photosensitivity, photoreversibility, or rapid response dynamics of CcaSR. These features make CcaSR v3.0 ideal for fine-tuning expression of toxic RNAs or proteins, or for characterizing the dynamics of synthetic or evolved gene regulatory circuits,[20] among other applications.

Experimental Section

Plasmids, strains, and media: pNO286-1 and pNO286-3 were constructed by using Golden Gate assembly[21] with pSR43.6[5] as the template vector. Assembled plasmids were transformed into *E. coli* NEB 10-β for validation (New England Biolabs, catalog no. C3019H). Lysogeny broth (LB) medium was used to culture transformed bacteria at 37 °C, with shaking at 250 rpm and appropriate antibiotics added [chloramphenicol (34 μg mL^{-1}) and/or spectinomycin (100 μg mL^{-1})]. Primers were ordered from Integrated DNA Technologies, Inc., IA, USA. DNA sequencing was performed by Genewiz (South Plainfield, NJ, USA). Sequence-validated plasmids were transformed into *E. coli* BW29655 [BW28357 Δ(envZ-ompR)520(::FRT)][22] for all experiments. All experiments were conducted in M9 medium [M9 salts (1×), glucose (0.4%, *w/v*), casamino acids (0.2%, *w/v*), $MgSO_4$ (2 mM), $CaCl_2$ (100 μM)] supplemented with appropriate antibiotics.

Plasmids will be available from Addgene.

Optical hardware: All experiments were conducted in four 24-well light plate apparatus (LPA) devices[23] mounted in a shaking incubator at 37 °C and 250 rpm. Green light (520 nm) and red light (660 nm) were generated from 520-2-KB (WP7083ZGD/G, Kingbright, CA, USA) and 660-LS (L2-0-R5TH50-1, LEDSupply, VT, USA) LEDs, respectively. Photon flux outputs were calibrated with a spectroradiometer (StellarNet UVN-SR-25 LT16) by our previous method.[10] Replicates of each set of experimental conditions were conducted in randomized well positions to minimize possible systematic errors.

Experimental protocols: All experiments were performed by using a previously described protocol.[24] In short, −80 °C aliquots of transformed bacteria were thawed at room temperature and inoculated to $OD_{600} = 5 \times 10^{-5}$ in fresh M9 medium. We determined that this inoculation density enables CcaSR gene expression responses to reach steady state and the culture to remain in exponential phase ($OD_{600} < 0.2$) for the duration of the experiment (8 h). Inoculated M9 cultures (500 μL) were then added into each well of a clear-bottomed black-walled 24-well plate (ArcticWhite AWLS-303008) that was then sealed with adhesive foil (VWR 60941-126). The culture plates were mounted into preconfigured LPA devices running the desired light program.[23] For dynamics experiments, cells were preconditioned for 2 h under green or red light (20 μmol m^{-2} s^{-1} 520 or 660 nm, respectively) and then switched to the final set of light conditions according to our previous staggered-start protocol[6] over the remaining 6 h. At the end of the experiment, the plates were removed from the incubator and placed in an ice/water slurry for 15 min to inhibit growth. Culture samples (100 μL) were pipetted from each well into prechilled flow cytometry tubes containing rifampicin dissolved in 1×PBS (500 μg mL^{-1}, 1 mL) and placed in a 37 °C water bath for 1 h to allow for maturation of translated sfGFP while inhibiting additional transcription. They were placed back in ice/water for 15 min, and then measured for sfGFP fluorescence by using flow cytometry.

Flow cytometry and data analysis: sfGFP fluorescence was measured with a BD FACScan flow cytometer, as previously described.[6] Calibration bead samples (Spherotech RCP-30-5A) diluted in PBS were measured each day an experiment was conducted. After data

acquisition, raw flow cytometry files were processed with FlowCal, as previously described.[16] Specifically, cell populations were density-gated (80%), and the raw fluorescence values [a.u.] were transformed into calibrated MEFL values by using a standard curve. Because of the wide range of fluorescence values, residual cells from previous samples measured with the flow cytometer might unduly affect arithmetic means of subsequent fluorescence distributions. We applied a simple Python script to identify and remove outlying bins of MEFL counts outside of the main cell population distribution to obtain a more robust estimation of the arithmetic mean fluorescence of the measured cell population. For each data sample, the script traversed outwards from the median of a smoothed histogram of MEFL counts with a preset bandwidth until the count values fell below a preset threshold value of 0.5% of the peak value. These bin points were then set as the upper and lower bounds of the population to be processed, and the outlying values were removed. The arithmetic means of each individually gated and trimmed population were calculated, and the arithmetic mean autofluorescence value of *E. coli* lacking sfGFP measured on the same day was subtracted from the arithmetic mean of each of these sfGFP-producing *E. coli* populations, resulting in the reported sfGFP fluorescence values.

Hill function model: We performed model fits in Figure 2 to the Hill equation [Eq. (1)]:

$$f(I) = b + \frac{aI^n}{K_{1/2}^n + I^n} \quad (1)$$

by using a nonlinear least-squares curve-fitting algorithm in Prism v7.0a for Mac OSX (GraphPad Software, Inc., CA, USA). Three replicates per data point were each weighted by their autofluorescence-corrected sfGFP fluorescence values, and the model function parameters were fit locally to each weighted replicate.[10]

Dynamics model [Eqs. (2), (3)]**:**

$$\frac{df}{dt} = a - k \cdot f(t) \quad (2)$$

$$f(t) = \begin{cases} b, & t < \tau \\ b + (a-b) \bullet (1 - e^{(-k(t-\tau))}), & t \geq \tau \end{cases} \quad (3)$$

The simplified ordinary differential equation (ODE) model describing sfGFP production $f(t)$ is shown, followed by the analytical solution to the ODE. The variables for the solution are defined as follows: a [MEFL] represents the fit parameter for the final sfGFP fluorescence reached, b [MEFL] represents the fit parameter for the initial sfGFP fluorescence measured in the system before any change in promoter activity occurs, k [min^{-1}] represents the exponential rate fit parameter, whereas τ [min] represents the fit parameter for the pure time delay before the change in promoter activity. Dynamics model fits to the analytical solution in Figure 3 were performed by use of the same software algorithm for Hill function fits. Each triplicate set of step response datasets was normalized to its minimum and maximum autofluorescence-corrected sfGFP fluorescence mean values. The model function parameters were then fit locally to each replicate weighted by their individual normalized values.

Acknowledgements

*This work was supported by the Office of Naval Research (MURI N000141310074) and a fellowship from the Agency of Science, Technology and Research (A*STAR), Singapore [National Science Scholarship (PhD), to N.T.O]. We thank Evan Olson for assistance with calibrating LPA hardware and programming scripts for flow cytometry data analysis, and the Joel Moake Lab for generously sharing their flow cytometer. We also thank Shelly Cheng, Evan Olson, and Sebastián Castillo-Hair for their helpful comments on the manuscript.*

Conflict of Interest

The authors declare no conflict of interest.

Keywords: gene expression · optogenetics · photoreceptors · synthetic biology

[1] Y. Hirose, R. Narikawa, M. Katayama, M. Ikeuchi, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8854–8859.
[2] Y. Hirose, T. Shimada, R. Narikawa, M. Katayama, M. Ikeuchi, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 9528–9533.
[3] J. J. Tabor, A. Levskaya, C. A. Voigt, *J. Mol. Biol.* **2011**, *405*, 315–324.
[4] G. A. Gambetta, J. C. Lagarias, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10566–10571.
[5] S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3*, 820–831.
[6] E. J. Olson, L. A. Hartsough, B. P. Landry, R. Shroff, J. J. Tabor, *Nat. Methods* **2014**, *11*, 449–455.
[7] C. Zechner, G. Seelig, M. Rullan, M. Khammash, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 4729–4734.
[8] E. A. Davidson, A. S. Basu, T. S. Bayer, *J. Mol. Biol.* **2013**, *425*, 4161–4166.
[9] A. Milias-Argeitis, M. Rullan, S. K. Aoki, P. Buchmann, M. Khammash, *Nat. Commun.* **2016**, *7*, 12546.
[10] E. J. Olson, C. N. Tzouanas, J. J. Tabor, *Mol. Syst. Biol.* **2017**, *13*, 926.
[11] R. Chait, J. Ruess, T. Bergmiller, G. Tkačik, C. C. Guet, *Nat. Commun.* **2017**, *8*, 1535.
[12] R. Ohlendorf, R. R. Vidavski, A. Eldar, K. Moffat, A. Möglich, *J. Mol. Biol.* **2012**, *416*, 534–542.
[13] X. Chen, R. Liu, Z. Ma, X. Xu, H. Zhang, J. Xu, Q. Ouyang, Y. Yang, *Cell Res.* **2016**, *26*, 854–857.
[14] A. Baumschlager, S. K. Aoki, M. Khammash, *ACS Synth. Biol.* **2017**, *6*, 2157–2167.
[15] M. Nakajima, S. Ferri, M. Rögner, K. Sode, *Sci. Rep.* **2016**, *6*, 37595.
[16] S. M. Castillo-Hair, J. T. Sexton, B. P. Landry, E. J. Olson, O. A. Igoshin, J. J. Tabor, *ACS Synth. Biol.* **2016**, *5*, 774–780.
[17] S. Crosson, S. Rajagopal, K. Moffat, *Biochemistry* **2003**, *42*, 2–10.
[18] M. D. Edgerton, A. M. Jones, *Biochemistry* **1993**, *32*, 8239–45.
[19] B. L. Taylor, I. B. Zhulin, *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 479–506.
[20] E. J. Olson, J. J. Tabor, *Nat. Chem. Biol.* **2014**, *10*, 502–511.
[21] C. Engler, R. Gruetzner, R. Kandzia, S. Marillonnet, *PLoS One* **2009**, *4*, e5553.
[22] L. Zhou, X. H. Lei, B. R. Bochner, B. L. Wanner, *J. Bacteriol.* **2003**, *185*, 4956–4972.
[23] K. P. Gerhardt, E. J. Olson, S. M. Castillo-Hair, L. A. Hartsough, B. P. Landry, F. Ekness, R. Yokoo, E. J. Gomez, P. Ramakrishnan, J. Suh, et al., *Sci. Rep.* **2016**, *6*, 35363.
[24] N. T. Ong, E. J. Olson, J. J. Tabor, *ACS Synth. Biol.* **2018**, *7*, 240–248.

Manuscript received: January 7, 2018
Accepted manuscript online: February 8, 2018
Version of record online: March 23, 2018