A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: A miniaturized E. coli green light sensor with high dynamic range

Authors: Nicholas Ting Xun Ong and Jeffrey J Tabor

To be cited as: *ChemBioChem* 10.1002/cbic.201800007

Link to VoR: http://dx.doi.org/10.1002/cbic.201800007

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

A miniaturized *E. coli* green light sensor with high dynamic range

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

[a] Dr. N.T. Ong, Dr. J.J. Tabor
Department of Bioengineering, Rice University
6100 Main St, Houston, TX 77005, USA
*Email: jeff.tabor@rice.edu

[b] Dr. J.J. Tabor
Department of Biosciences, Rice University
6100 Main St, Houston, TX 77005, USA

ABSTRACT

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 de-activated by red, into *E. coli*, resulting in a sensor with 6-fold dynamic range. Later, we optimized pathway protein expression levels and the output promoter sequence to decrease transcriptional leakiness and increase the dynamic range to approximately 120-fold. These CcaSR v1.0 and 2.0 systems have been used for precise quantitative, temporal, and spatial control of gene expression for a variety of applications. Recently, others have deleted two PAS domains of unknown function from the CcaS sensor histidine kinase in a CcaSR v1.0-like system. Here, we apply these deletions to CcaSR v2.0, resulting in a v3.0 light sensor with 4-fold lower leaky output and nearly 600-fold dynamic range. 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.

Keywords: synthetic biology, optogenetics

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) comprised 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 1A**).[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]

To construct CcaSR v1.0, we cloned the *Synechocystis* PCC 6803 *ccaS-ccaR-cpcG2* genomic region into an *E. 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 to PCB.[4] The resulting system exhibits some leaky transcriptional output in red light, and 6-fold activation by green.[5] Despite these relatively modest features, CcaSR v1.0 has been used to perform bacterial photography,[3] program tailor-made gene expression signals,[6] characterize the input/output dynamics [6] and adaptive noise filtering properties [7] of synthetic gene circuits, control cellular growth rate by modulating expression of a metabolic enzyme,[8] and 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-fold 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] while Guet and coworkers 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 10,000-fold activation.[12ï 14] However, blue light has higher phototoxicity than green, and these photoreceptors suffer from wide absorbance spectra and slow response dynamics relative to CcaSR. We therefore sought to improve the dynamic range (ratio of output gene expression in green to 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 ï A514, a linker region upstream of the HK domain. When compared to full-length CcaS in a plasmid setup similar to our CcaSR v1.0 system, two variants named mini-CcaS #3 and #10 (**Figure 1A**) yielded slightly lower gene expression output 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 may 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, pNO286-3 (**Figure 1A**; **S1**). Next, we co-transformed each of the three plasmids pairwise with pSR58.6, which encodes *ccaR* and a superfolder green fluorescent protein (*sfGFP*) reporter (**Figures 1B, S1**).[5] Then, we utilized flow cytometry to measure sfGFP expression levels under green and red light, following our previous protocols (**Experimental Section**).[6,16] We observed a CcaSR v2.0 output of 1.06 ± 0.12 x 10^3 molecules of equivalent fluorescein (MEFL) sfGFP under red light and 1.163 ± 0.042 x 10^5 under green (**Figure 1C**). These data constitute a 110 ± 11-fold (**Figure 1D**) dynamic range, consistent with our previous measurements.[5] Mini-CcaS #3 results in higher sfGFP expression levels under both red (3.806 ± 0.089 x 10^5 MEFL) and green (2.00 ± 0.31 x 10^3 MEFL) (**Figure 1C**), and an improved dynamic range of 194 ± 32-fold (**Figure 1D**). Mini-CcaS #10 simultaneously yields a 3.99 ± 0.26ï fold reduction in red light output (2.67 ± 0.27 x 10^2 MEFL) and higher green output (1.574 ± 0.053 x 10^5 MEFL) than full length CcaS (**Figure 1C**), resulting in 593 ± 56-fold dynamic range (**Figure 1D**). Though 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 ñCcaSR v3.0ò.

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. First, 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 compared to 2.114 ± 0.052 for the v2.0 system (**Table S1**). Though 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 Ñ 0.025 ε mol Ö m^{-2} Ö s^{-1} 520 nm light for 50 % maximal activation, which is similar to the value for CcaSR v2.0 (0.793 ± 0.028 ε mol Ö m^{-2} Ö 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 de-activation experiments where 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 (**Experimental Section**). We observe no substantial difference in the response dynamics of CcaSR v3.0 relative to its predecessor (**Figure 3**).

While 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 domains is unknown. First, the purified CcaS GAF domain binds PCB and undergoes the same photocycle as a nearly full-length (missing the 23 amino acid 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, based on 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 ï other PAS domains are known to bind to light-absorbing ligands like 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 4-fold lower leakiness than our second-generation system and nearly 600-fold dynamic range. Both of these performance features are superior to 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 achieve the aforementioned 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 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 using Golden Gate assembly [21] with pSR43.6 [5] as the template vector. Assembled plasmids were transformed into *E. coli* NEB 10-b for validation (New England Biolabs, catalog no. C3019H). LB media was used to culture transformed bacteria at 37 ºC, with shaking at 250 rpm, and appropriate antibiotics added (chloramphenicol (34 ε g/mL), and/or spectinomycin (100 ε g/mL)).

Primers were ordered from Integrated DNA Technologies, Inc., IA, USA. DNA sequencing was performed using Genewiz, NJ, USA. Sequence validated plasmids were transformed into *E. coli* BW29655 (BW28357 φ(envZ-ompR)520(::FRT)) [22] for all experiments. All experiments were conducted using M9 media (M9 salts (1×), glucose (0.4 % w/v), cas-amino acids (0.2 % w/v), $MgSO_4$ (2 mM), $CaCl_2$ (100 εM)) supplemented with appropriate antibiotics.

Optical Hardware

All experiments were conducted in four 24-well Light Plate Apparatus (LPA) devices [23] mounted in a shaking incubator at 37 ºC and 250 rpm. 520 nm green light and 660 nm red light were generated using 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 using a spectroradiometer (StellarNet UVN-SR-25 LT16) via our previous method.[10] Replicates of each experimental condition were conducted in randomized well positions to minimize possible systematic errors.

Experimental Protocols

All experiments were performed 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 x 10^{-5} in fresh M9 media. 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 pre-configured LPA devices running the desired light program.[23] For dynamics experiments, cells were pre-conditioned for 2 h under green or red light (20 εmolΘm^{-2}Θs^{-1} 520 nm or 660 nm respectively) and then switched to the final light condition using 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 pre-chilled flow cytometry tubes containing rifampicin dissolved in 1× PBS (500 µg/mL, 1 mL) and placed in a 37 ºC 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 using flow cytometry.

Flow Cytometry and Data Analysis

sfGFP fluorescence was measured using 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 using a standard curve. Because of the wide range of fluorescence values, residual cells from previous samples measured on the flow cytometer could 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 $f(I) = b + \frac{a \cdot I^n}{K_{1/2}^n + I^n}$ using a non-linear 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

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

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

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, while 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 using the same software algorithm for Hill function fits. Each triplicate set of step response datasets were normalized to their 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. Plasmids will be available from Addgene.

REFERENCES

[1] Y. Hirose, R. Narikawa, M. Katayama, M. Ikeuchi, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 8854ï 8859.

[2] Y. Hirose, T. Shimada, R. Narikawa, M. Katayama, M. Ikeuchi, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9528ï 9533.

[3] J. J. Tabor, A. Levskaya, C. A. Voigt, *J. Mol. Biol.* **2011**, *405*, 315ï 324.

[4] G. A. Gambetta, J. C. Lagarias, *Proc. Natl. Acad. Sci.* **2001**, *98*, 10566ï 10571.

[5] S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3*, 820ï 31.

[6] E. J. Olson, L. A. Hartsough, B. P. Landry, R. Shroff, J. J. Tabor, *Nat. Methods* **2014**, *11*, 449ï 455.

[7] C. Zechner, G. Seelig, M. Rullan, M. Khammash, *Proc. Natl. Acad. Sci.* **2016**, *113*, 4729ï 4734.

[8] E. A. Davidson, A. S. Basu, T. S. Bayer, *J. Mol. Biol.* **2013**, *425*, 4161ï 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. Tkaļik, C. C. Guet, *Nat. Commun.* **2017**, *8*, DOI 10.1038/s41467-017-01683-1.

[12] R. Ohlendorf, R. R. Vidavski, A. Eldar, K. Moffat, A. Möglich, *J. Mol. Biol.* **2012**, *416*, 534ï 542.

[13] X. Chen, R. Liu, Z. Ma, X. Xu, H. Zhang, J. Xu, Q. Ouyang, Y. Yang, *Cell Res.* **2016**, *26*, 854ï 857.

[14] A. Baumschlager, S. K. Aoki, M. Khammash, *ACS Synth. Biol.* **2017**, *6*, 2157ï 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*, 774ï 780.

[17] S. Crosson, S. Rajagopal, K. Moffat, *Biochemistry* **2003**, *42*, 2ï 10.

[18] M. D. Edgerton, A. M. Jones, *Biochemistry* **1993**, *32*, 8239ï 45.

[19] B. L. Taylor, I. B. Zhulin, *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 479ï 506.

[20] E. J. Olson, J. J. Tabor, *Nat. Chem. Biol.* **2014**, *10*, 502ï 11.

[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*, 4956ï 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*, 240ï 248.

FIGURE LEGENDS

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 to constitutively express *ccaS*, *ccaR* and an optimized metabolic pathway for PCB production (*ho1* and *pcyA*), while 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 1A 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 SD 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.

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 SD respectively of n = 3 experiments run on

separate days. Solid lines represent the Hill function fit. Fit parameters are shown in **Table S1**.

Figure 3

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

Table of Contents Text

We previously engineered v1.0 and v2.0 *E. coli* green light sensors from the *Synechocystis* PCC 6803 CcaSR two-component system, with 6- and 117-fold dynamic range respectively. Here, we introduce a double PAS domain deletion recently demonstrated to decrease leakiness in a v1.0-like system by Sode and coworkers into the v2.0 system, resulting in a v3.0 CcaSR system with 4-fold lower leak, almost 600-fold dynamic range, and otherwise identical optical response properties.

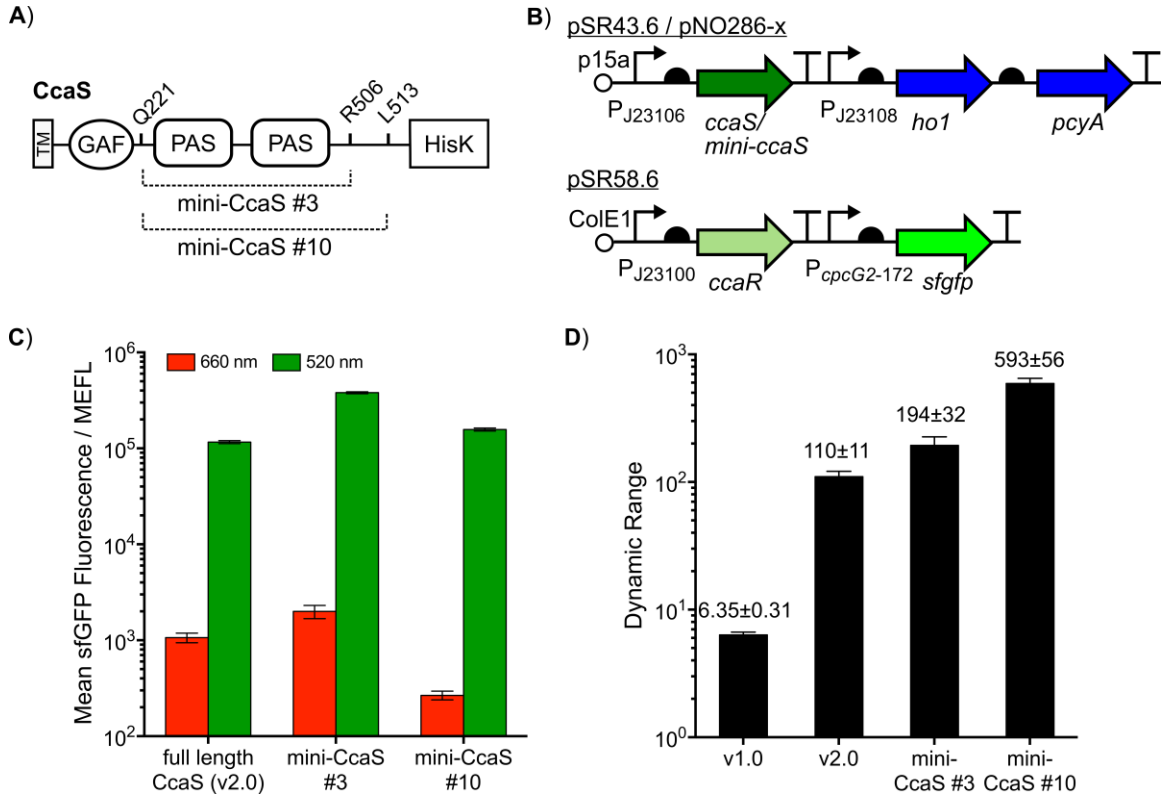
A)
CcaS
TM
GAF
Q221
PAS
PAS
R506
L513
HisK
mini-CcaS #3
mini-CcaS #10
B)
pSR43.6 / pNO286-x
p15a
P J23106
ccaS/
mini-ccaS
P J23108
ho1
pcyA
pSR58.6
ColE1
P J23100
ccaR
P cpcG2-172
sfgfp
C)
Mean sfGFP Fluorescence / MEFL
660 nm
520 nm
full length CcaS (v2.0)
mini-CcaS #3
mini-CcaS #10
D)
Dynamic Range
6.35±0.31
110±11
194±32
593±56
v1.0
v2.0
mini-CcaS #3
mini-CcaS #10

Figure 1

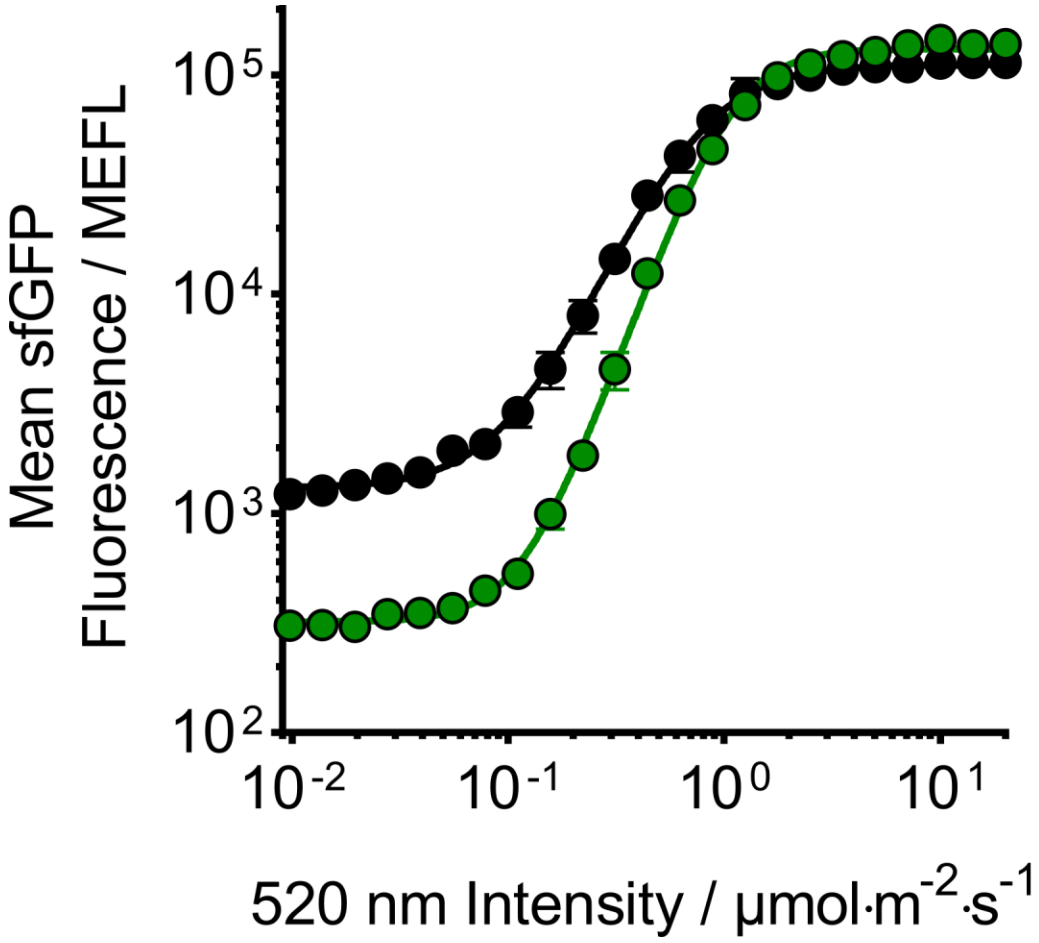
Mean sfGFP
Fluorescence / MEFL
10^5
10^4
10^3
10^2
10^-2
10^-1
10^0
10^1
520 nm Intensity / µmol·m^-2·s^-1

Figure 2

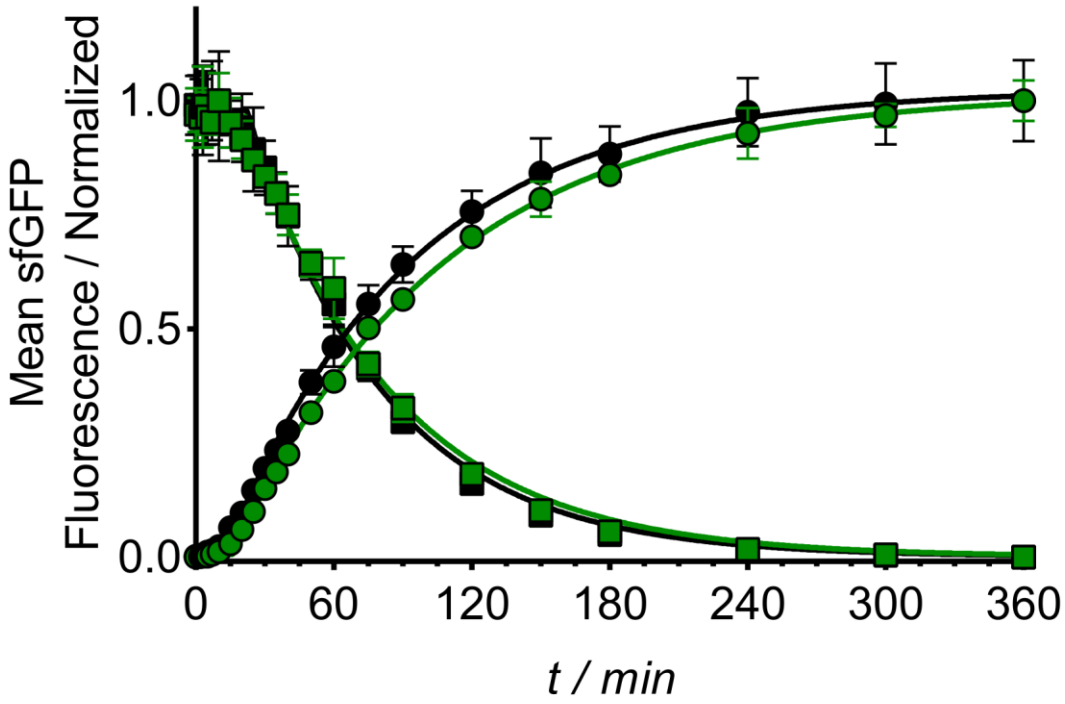
Mean sfGFP
Fluorescence / Normalized
1.0
0.5
0.0
0
60
120
180
240
300
360
t / min

Figure 3