ARL-TR-8872 ● Dec 2019

DEVCOM ARMY RESEARCH LABORATORY

Evaluation of Genetic Switches and Reporters as Synthetic Biology Tools for *Clostridium acetobutylicum*

by Alexander V Tobias

OPSEC Review

Approved

The OPSEC Review serves as an approval record for ARL generated information to be presented or disseminated external to ARL.

Documents

ARL-TR-8872 Evaluation of Genetic Switches and Reporters as Synthetic Biology Tools for Clostridium acetobutylicum

Type:	Report / Technical Report
Authors:	Alexander Tobias
Description:	Tech report is currently published with government-only restricted access. I would like the access changed to public disclosure and to have it made available to the public.
Contains Human Subjects study information?	No
Project number:	
Period covered:	Jul 1, 2017 - Jul 31, 2019
Sponsoring agency:	DoD ARAP (Synthetic Biology for Military Environments
File name:	ARL-TR-8872_For_Public_Release.docx

Package Details

Unclassified Title or Solicitation Number/Title:	ARL-TR-8872, Evaluation of Genetic Switches and Reporters as Synthetic Biology Tools for Clostridium acetobutylicum
Distribution Statement:	Approved for public release; distribution is unlimited.
Reasons:	0. Public release
Supplemental Information:	The purpose of this Form-1 package is to obtain authorization to convert this published ARL Technical Report from Government-only access to publicly accessible.
Security Classification:	Unclassified
Was the author extended a personal invitation by the sponsoring organization to give a presentation or submit a publication?	No
Is the manuscript subject to export control?	No
Will foreign nationals have access to this information?	Yes
Core competencies:	Materials and Manufacturing Science
Collaborative mechanisms:	Synthetic Biology for Military Environments ARAP
Created on:	Sep 3, 2020, 9:45 AM
Approved on:	Sep 17, 2020, 7:46 AM

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-8872 ● DEC 2019

Evaluation of Genetic Switches and Reporters as Synthetic Biology Tools for *Clostridium acetobutylicum*

Alexander V Tobias
Biotechnology Branch, Sensors and Electron Devices Directorate,
CCDC Army Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
December 2019	Technical Report	1 July 2017–31 July 2019

4. TITLE AND SUBTITLE
Evaluation of Genetic Switches and Reporters as Synthetic Biology Tools for *Clostridium acetobutylicum*

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Alexander V Tobias

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
CCDC Army Research Laboratory
ATTN: FCDD-RLS-CB
2800 Powder Mill Road
Adelphi, MD 20783-1138

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-8872

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Department of Defense (DOD) Applied Research for the Advancement of Science and Technology Priorities (ARAP) Program in Synthetic Biology for Military Environments (SBME)

10. SPONSOR/MONITOR'S ACRONYM(S)
DOD ARAP SBME

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
ORCID ID: Alexander V Tobias, 0000-0002-5866-5254

14. ABSTRACT
This report presents an evaluation of reporter proteins and inducible gene expression elements in *Clostridium (C.) acetobutylicum*, a bacterium of interest to the Army. Reporter proteins evaluated were LacZ (beta-galactosidase), PhiLOV, Evoglow C-Bs2, SNAP-Tag, FAST, and NanoLuc luciferase. Inducible systems evaluated included six riboswitches and seven chemically inducible promoter systems. Most of these genetic tools had not been previously tested in *C. acetobutylicum*. The theophylline-inducible riboswitch and the NanoLuc luminescent reporter protein proved useful, convenient, and modular. The FAST and SNAP-Tag fluorescent reporters also proved worthwhile for flow cytometry. A subset of these genetic tools was deployed as part of a study and evaluation of four different Gram-positive origins of replication in *C. acetobutylicum*: pBP1, pCB102, pCD6, and pIM13, with respect to their utility for elevated protein expression. The pIM13 origin yielded the highest level of protein expression, which was about 20-fold greater than that conferred by pBP1, the origin with the next-highest copy number. An outlook perspective is provided, based on this and related work, of genetic tools that could be tested or retested in *C. acetobutylicum*, with lessons from this work as a guide.

15. SUBJECT TERMS
synthetic biology, genetic engineering, Clostridium, riboswitch, reporter protein

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	42	Alexander V Tobias
Unclassified	Unclassified	Unclassified			19b. TELEPHONE NUMBER (Include area code) (301) 394-4162

Contents

List of Figures

List of Tables

Acknowledgments

I thank several members of the Biotechnology Branch for their assistance and teaching: Nathan Schwalm, Rebecca Renberg, Deborah Sarkes, and Christian Sund. This work was supported by the Department of Defense Applied Research for the Advancement of Science and Technology Priorities Program in Synthetic Biology for Military Environments.

Summary

Clostridium acetobutylicum (*C. acetobutylicum*) is a solventogenic bacterium of interest to industry and the military for its abilities to utilize and grow on an array of 5- and 6-carbon sugars, to form stable spores for simple transportation logistics, and for its potential to synthesize organic compounds of import via simple fermentation in quiescent anaerobic tanks. Development of new applications for *C. acetobutylicum* suffers, however, from a dearth of genetic and synthetic biology tools for constructing novel, controllable metabolic pathways and stimuli-responsive genetic switches and circuits. Herein, I present my evaluation of a set of reporter proteins and inducible gene expression elements in *C. acetobutylicum*. Reporter proteins evaluated were LacZ (beta-galactosidase), PhiLOV, Evoglow C-Bs2, SNAP-Tag, FAST, and NanoLuc luciferase. Inducible systems evaluated were six riboswitches and seven chemically inducible promoter systems. These efforts quickly uncovered that it was best to avoid reporters that fluoresce in the green color spectrum to avoid interference with the substantial cellular autofluorescence of this bacterium. While the reporters and inducible elements I evaluated were gathered from the literature, most had not been previously tested in *C. acetobutylicum*. The best success was with the theophylline-inducible riboswitch and the NanoLuc luminescent reporter protein system. The FAST and SNAP-Tag fluorescent reporters also proved useful for flow cytometry but suffered from some drawbacks. A subset of these genetic tools was deployed as part of a study and evaluation of four different Gram-positive plasmid origins of replication in *C. acetobutylicum:* pBP1 from a *Clostridium botulinum* indigenous plasmid, pCB102 from a *Clostridium butyricum* plasmid, pCD6 from a *Clostridium difficile* plasmid, and pIM13 from *Bacillus subtilis* plasmid. The results of these measurements correlated strongly with quantitative polymerase chain-reaction measurements of plasmid copy number and plasmid-based gene transcript levels, and helped rank these origins with respect to their utility for elevated protein expression. The pIM13 origin yielded the highest level of protein expression, which was about 20-fold greater than that conferred by pBP1, the origin with the second-highest copy number. These results also point to a noteworthy and potentially useful observation about the interaction between the theophylline riboswitch's maximal induction level and the plasmid copy number. An outlook perspective is provided, based on this and related work, of genetic tools that could be tested or retested in *C. acetobutylicum*, with lessons from this work serving as a guide.

1. Introduction

Solventogenic clostridia such as *Clostridium acetobutylicum (C. acetobutylicum)* have several desirable characteristics for serving as chassis organisms for industrial and military applications. They possess the natural ability to catabolize a wide range of 5- and 6-carbon sugars and show potential to be engineered to efficiently produce a number of specialty chemicals. Being strict anaerobes, they can be cultured in simple vessels without the need for aeration or agitation. The high tolerance of these bacteria to solvents and low pH, and their ability to form stable spores, represent additional attractive features.

However, although clostridial species have a long history of anchoring industrial fermentation processes for solvent and biofuel production, they have been much less extensively genetically modified and engineered compared with laboratory mainstay microorganisms such as *Escherichia coli* (*E. coli*) or *Saccharomyces cerevisiae (S. cerevisiae)*.

Clostridia are certainly more challenging to work with and genetically engineer than many other microbes. They are obligate anaerobes, have low-guanine-cytosine genomes that constrain locus targeting with several clusters of regularly interspaced short palindromic repeats (CRISPR)-based tools and encumber synthesis of many native or codon-adapted deoxyribonucleic acid (DNA) sequences, and typically have low transformation efficiency. These "high barriers to entry" have kept the community of solventogenic clostridia researchers small and limited in terms of focus areas and bandwidth. Consequently, there remains a dearth of genetic and synthetic biology tools to support and enable further discovery and engineering efforts with solventogenic clostridia.[1,2]

With funding from the Department of Defense's Applied Research for the Advancement of Science and Technology Priorities Program, *Synthetic Biology for Military Environments*, I was able to conduct the research reported herein on the testing, evaluation, and development of genetic tools to enable further genetic engineering of *C. acetobutylicum*. In this report, I present some additions to the clostridial synthetic biology toolbox: reporter proteins and inducible gene expression elements that I characterized and in some cases, recommend for adoption by those working to engineer *C. acetobutylicum* and other clostridia. My results with the genetic parts deemed less suitable for this organism (Evoglow, phiLOV2.1) are useful to the clostridial research community for the effort they can help conserve. The more suitable "newly validated" parts on which I focused this report—theophylline riboswitch, SNAP-Tag with SNAP-Cell-647-SiR, FAST protein with HBR-3,5-DOM fluorogen, NanoLuc luciferase with Nano-Glo Live

Cell Reagent—function well within the constraints of this chassis organism. They avoid interference by green cellular autofluorescence[3–6] and, except for NanoLuc, function under fully anoxic conditions. These parts provide greatly needed additional options to researchers for exerting control over, monitoring, and quantifying the expression of just about any native or heterologous gene(s).

2. Methods, Assumptions, and Procedures

2.1 Chemicals and Materials

When not stated otherwise, chemicals were obtained from Sigma Aldrich (St Louis, Missouri). The 4-hydroxy-3,5-dimethoxybenzylidene rhodanine (HBR-3,5-DOM) was purchased from Matrix Scientific (Columbia, South Carolina). SNAP-Cell-647-SiR was from New England Biolabs (Ipswitch, Massachusetts). FACSFlow flow cytometry sheath fluid was from BD Biosciences (San Jose, California). Anhydrotetracycline (aTc) hydrochloride and 2,4-diacetylphloroglucinol (DAPG) were from Cayman Chemical (Ann Arbor, Michigan). Nano-Glo Live Cell Assay System was from Promega (Madison, Wisconsin).

2.2 Bacterial Strains and Propagation

E. coli strains Zymo-5α (Zymo Research, Irvine, California) or NEB-5 α (New England Biolabs) were used for plasmid cloning and replication. *E. coli* cells harboring plasmids were grown in Luria-Bertani (LB) medium supplemented with 30 µg/mL of chloramphenicol and plated onto LB plates supplemented with the same concentration of the antibiotic. *C. acetobutylicum* ATCC 824 was grown in clostridial growth medium (CGM)[7] supplemented with 5 mg/mL of glucose and, for liquid cultures of strains harboring plasmids, 30 µg/mL of chloramphenicol or thiamphenicol. Agar plates for plating transformations of *C. acetobutylicum* and maintenance of said transformants by streaking were CGM, supplemented with 5 mg/mL of glucose and 15 µg/mL of chloramphenicol. Clostridia were handled in a Coy anaerobic chamber (Coy Laboratory Products, Grass Lake, Michigan) and cultured at 37 °C in an incubator within the chamber without agitation. The gas mixture in the chamber was approximately 93% nitrogen, 5% carbon dioxide, and 1–2% hydrogen. Chemical induction of *C. acetobutylicum* was performed on freshly 10-fold diluted subcultures with the specified concentration of theophylline or aTc dissolved in dimethyl sulfoxide (DMSO). DMSO-only controls were run with an equivalent dose of DMSO (generally 1% v/v).

2.3 Plasmids

Shuttle plasmids were based on the pMTL80000 modular plasmid system purchased from CHAIN Biotechnology (Marlow, United Kingdom)[8] specifically pMTL85141, which contains a pIM13 Gram-positive replicon, ColE1 Gram-negative replicon, *catP* gene for chloramphenicol or thiamphenicol resistance, and a multi-cloning site. Table A1 in the Appendix contains additional information about the plasmids and genetic parts used in this work. Plasmid derivatives of pMTL85141 were constructed using a variety of assembly techniques: standard restriction and ligation, NEBuilder assembly (New England Biolabs), or Golden Gate assembly (New England Biolabs). The DNA cassettes assembled with the pMTL85141 backbone were custom-synthesized by IDT (Coralville, Iowa), Genscript (Piscataway, New Jersey), or Twist Bioscience (South San Francisco, California). Plasmid pAVT020 was built using components of the anhydrotetracycline-inducible promoter system from plasmid pWKS1583[9] (based on pRPF185[10]), which was obtained from Addgene (catalog #70188, Watertown, Massachusetts). Plasmid miniprep and midiprep purification kits from Qiagen (Germantown, Maryland) and Zymo Research were used for purification of plasmids from *E. coli* cultures. Plasmids and DNA assembly mixtures were transformed into frozen chemically competent *E. coli* (previously listed strains) following the supplier's recommended protocol. Genome Compiler software version 2.2.88 was used for design, visualization, and archiving of plasmid sequences. Plasmids were sequenced by Illumina NextSeq, with Illumina Nextera library preparation.

2.4 Plasmid Transformation into *C. acetobutylicum*

All procedures in Section 2.4 were performed in an anaerobic chamber as previously described. Plasmids were methylated before transformation into Clostridium in vitro with CpG methyltransferase (New England Biolabs) or by transformation into an *E. coli* strain harboring a second plasmid encoding a phage methyltransferase,[11] followed by overnight culturing and plasmid purification. Competent *C. acetobutylicum* cells were grown in 50-mL CGM + 5-mg/mL glucose to an OD600 nm of approximately 1.3 as measured on an Ultrospec 10 Cell Density Meter (Amersham Biosciences, Amersham, United Kingdom). The culture was iced for 10 min, then centrifuged at 6000×g for 10 min. The supernatant was decanted and the cells were resuspended in 25 mL of ice-cold sodium phosphate buffer (5 mM) supplemented with 275 mM sucrose. The cells were then put on ice for 10 min, and the centrifugation–resuspension cycle was repeated two additional times, except the final resuspension was with 2.4 mL of the sodium phosphate–sucrose buffer. Six-tenths of a milliliter (0.6 mL) of these resuspended cells was

transferred to a 4-mm gap electroporation cuvette (Bio-Rad, Hercules, California). Added to the cells in the cuvette was 1–2 µg of plasmid DNA. The cuvette was tapped to mix, chilled on ice for 30 min, then placed in a Bio-Rad Gene Pulser Xcell electroporator and pulsed with settings of 2000 V, 25 µF, and infinite resistance. The contents of the cuvette were then added to 5 mL of CGM + 5-mg/mL glucose prewarmed to 37 °C. This culture of electroporated cells was incubated at 37 °C for 2–4 h, then concentrated to approximately 0.5 mL by centrifugation and resuspension in a reduced amount of supernatant, followed immediately by plating onto CGM–glucose agar supplemented with 15 µg/mL of chloramphenicol and incubation at 37 °C. Transformed colonies typically appeared within 24–48 h.

2.5 Measurements

Beta-galactosidase (LacZ) enzyme assays on clostridial culture lysates were conducted as follows: 2-mL cultures were grown overnight in 48-well blocks in CGM medium with 5-mg/mL glucose and 30-µg/mL chloramphenicol (for plasmid-harboring strains). Cell density was measured as optical density at 600 nm on a Synergy HT plate reader (BioTek, Winooski, Vermont) after diluting a 15-µL aliquot of each well 10-fold in fresh medium in a flat-bottom 96-well assay plate (Corning Costar, Corning, New York). The block was centrifuged at 4500×g in a 5804 R Centrifuge (Eppendorf, Hauppauge, New York) for 10 min, then the supernatants were gently removed by pipette. To each well was added 300 µL of lysis buffer consisting of 50-mM sodium phosphate buffer pH 7, 20-µL/mL lysozyme, 10-µL/mL protease inhibitor cocktail, and 0.2-µL/mL Benzonase (all from Sigma Aldrich). Approximately 300 µL of beads (0.1-mm zirconia/silica) (Biospec Products, Bartlesville, Oklahoma) were then dispensed to each well. The plate was fastened inside a 1600 MiniG Automated Tissue Homogenizer and Cell Lyser (SPEX, Metuchen, New Jersey) and agitated for two cycles of 45 s each at 1500 strokes/min, with 2 min of rest between cycles. The plate was centrifuged again with the same settings. For each enzyme assay, 20 µL of lysate was used, and was added to 128.5 µL of 50 mM sodium phosphate buffer pH 7 and 1.5 µL of 5-mg/mL fluorescein di-β-D-galactopyranoside (FDG) fluorescent substrate (ThermoFisher Scientific, Waltham, Massachusetts) prewarmed to 37 °C in a black-walled, clear-bottom 96-well assay plate (Corning Costar). After pipette mixing, the plate was read in the Synergy HT plate reader in kinetic mode for 1 h using the 485/20-nm excitation filter and the 528/20-nm emission filter. Relative LacZ expression levels were determined as the slope of the FDG hydrolysis signal between 8 and 20 min, and were normalized to culture OD600 nm.

Flow cytometry was conducted with a FACS Canto II instrument (BD Biosciences) with the threshold voltage set to 700 V. The instrument is equipped with 488- and

633-nm lasers. Cellular autofluorescence was detected in the fluorescein channel (488-nm laser; 530/30-nm emission filter). FAST protein complexed with the HBR-3,5-DOM fluorogen was detected in the phycoerythrin (PE) channel (488-nm laser; 585/40-nm emission filter). SNAP-Tag protein conjugated with SNAP-Cell-647-SiR was detected in the allophycocyanin (APC) channel (633-nm laser; 660/20-nm filter). For each sample, 10,000 events were recorded. Subcultures were prepared for flow cytometry by diluting overnight *C. acetobutylicum* cultures 10-fold into fresh medium containing inducers and growing to exponential phase (OD600 nm between ~0.2–0.4, generally 4–5 h). For treatment with substrate or fluorogen, culture aliquots between 200–800 µL in volume were taken. For measurement of SNAP-Tag expression, a stock solution of 30-µM SNAP-Cell-647-SiR substrate dissolved in DMSO was added to a *C. acetobutylicum* culture at a final concentration of 250 nM and the culture was left to incubate for an additional 30 min at 37 °C. Treated cultures were then washed three times to remove unbound substrate from the cells before flow cytometry. Each wash cycle consisted of centrifugation at 6000×g for 3 min, removal of supernatant, resuspension in 1 mL of FACSFlow sheath fluid (BD Biosciences), and static incubation at room temperature for 5 min to allow unbound substrate to diffuse out of cells. If cells were also to be treated with HBR-3,5-DOM for measurement of FAST expression, the fluorophore was added last from a 5-mM stock solution in DMSO to a final concentration of 25 µM. After incubation at room temperature for 15 min, flow cytometry was conducted on the treated samples.

Luciferase assays on NanoLuc-expressing cultures of *C. acetobutylicum* and related control cultures were performed according to the instructions provided in the Nano-Glo Live Cell Assay System kit from Promega. Aliquots of *C. acetobutylicum* cultures between 2 and 4 µL in volume were added to 106–108 µL of Nano-Glo Live Cell Substrate Dilution buffer without lysis. All-white assay plates were used for the luciferase assays, which were run in a Synergy Neo2 plate reader (BioTek). With the plate on the input tray of the reader, 10 µL of substrate solution containing 1 µL of Nano-Glo Live Cell Substrate was rapidly dispensed to each well (for a total of 120 µL per assay well). Luminescence readings were then taken by the plate reader every 30 s for 5 min. These readings tended to plateau for 1–3 min, then decrease as the substrate was consumed. To ensure that readings were taken with sufficient substrate to accurately measure luciferase expression, the maximum luminescence value read within that timeframe was taken as the luminescence value for each well. Relative luminescence units (RLU) were normalized to culture OD600 nm.

3. Results and Discussion

3.1 Development of a Ligand-inducible Riboswitch in *C. acetobutylicum*

A core goal of synthetic biology is to control the behavior and/or metabolism of an organism. As such, a bevy of sense-and-respond genetic switches have been developed and implemented in workhorse laboratory microbes such as *E. coli* and *S. cerevisiae*. These inducible switches can be activated (or deactivated) by the addition of chemical ligands,[12] light,[13,14] or other stimuli such as electrical current.[15] In turn, these switches are used as components for the construction of increasingly complex "genetic circuits" and "genetic logic gates" that enable their host cells to process information inputs from their surroundings in a manner analogous to electronic computers[16] or process control modules.[17]

To demonstrate the function of these switches and circuits in cells, researchers generally use them to control the expression of reporter proteins, which are fluorescent proteins (e.g., green fluorescent protein) or enzymes (e.g., LacZ, luciferase) whose presence is relatively easy to measure with straightforward assays or ubiquitous equipment.

Prior to this study, few inducible gene-expression systems had been described in *Clostridium*. The anhydrotetracycline-inducible promoter system found widespread application in microbial synthetic biology and related research, and its implementation in *C. acetobutylicum* was published in 2012.[18] For my first attempt at introducing a new genetic switch into *C. acetobutylicum*, I chose to try a riboswitch. This was an attractive option because it only required a small amount of DNA (~80 nucleotides) in addition to the reporter protein, unlike most inducible promoter systems, which require the additional expression of a transcription factor protein consisting of approximately 1000 or more nucleotides. More specifically, I chose theophylline riboswitch "E" (TheoRS^{E}) described by Topp et al.[19] Figure 1 illustrates the method of action of this riboswitch.

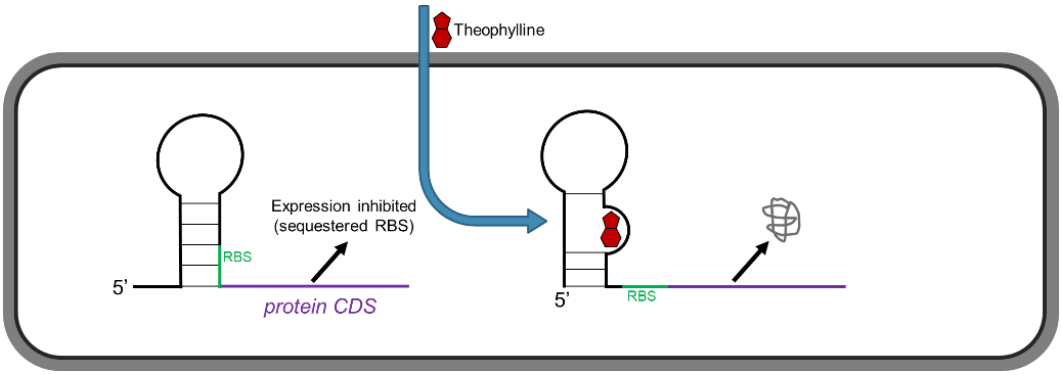

Fig. 1 Schematic of the theophylline-on riboswitch. In the absence of inducer, the riboswitch portion of the mRNA forms a stem–loop structure that occludes the ribosomal binding site (RBS) from the ribosome. When theophylline is added, it binds to the riboswitch and alters the stem–loop structure, exposing the RBS to the ribosome. Protein translation may then proceed.

3.2 Evaluation of Reporter Proteins in *C. acetobutylicum*

With TheoRS^{E} appended to the native constitutive thiolase promoter from *C. acetobutylicum*[8] I was able to control the expression of several reporter proteins by induction with added theophylline. Table 1 summarizes these results. LacZ was a safe reporter enzyme to try first, as it had previously been reported to function in *C. acetobutylicum*.[20] The data plotted in Fig. 2 show that LacZ enabled demonstration of the inducibility of the riboswitch with an induction ratio of approximately 10 (ratio of induced vs. uninduced expression level).

Table 1 Reporter proteins evaluated in *C. acetobutylicum*

Protein	Type	References	Notes	Results
LacZ	Beta-galactosidase enzyme	21	Many substrates available, including fluorescent ones	Works for lysis and enzyme assays. Green fluorescent substrate confounded with *C. acetobutylicum* autofluorescence. Poor for flow cytometry.
PhiLOV	Fluorescent protein with flavin chromophore	3	Does not require oxygen; 19 kDa	Weak green fluorescence in *E. coli*. Confounded with *C. acetobutylicum* autofluorescence. Undetectable signal in *C. acetobutylicum*.
Evoglow C-Bs2	Fluorescent protein with flavin chromophore	22	Does not require oxygen; 19 kDa	Bright green fluorescence in *E. coli*. Confounded with autofluorescence and low signal in *C. acetobutylicum*.
FAST	Variant of photoactive yellow protein. Noncovalently binds to fluorogenic dye.	23,24	Does not require oxygen; 14 kDa. Can be fused to other proteins. Three fluorogenic dyes available. Fluorogen can dissociate in ~10 s.[24]	Works well in *E. coli*. Functional but inconsistent with HBR-3,5-DOM in *C. acetobutylicum*. Some interference with *C. acetobutylicum* autofluorescence.
SNAP-Tag	Engineered alkylguanine-DNA alkyltransferase. Covalently binds to fluorescent benzylguanine derivatives.	5,25–28	Does not require oxygen; 19 kDa. Can be fused to other proteins. Seven cell-penetrating substrates available.	Very good for flow cytometry in *C. acetobutylicum* with SNAP-Cell 647 SiR substrate. No interference with autofluorescence. Laborious: requires multiple washes.
NanoLuc	Engineered luciferase from deep sea shrimp	29–31	Furimazine substrate. 19 kDa, 171 aa. Requires oxygen, but needs only a small aliquot of cells (~2 μL).	Very sensitive and useful reporter for *C. acetobutylicum*. Very low background signal (<0.1%) from nonexpressing cells. Glowing visible to naked eye. Low labor—no washes or lysis with NanoGlo Live Cell Kit (Promega). Not directly compatible with flow cytometry.

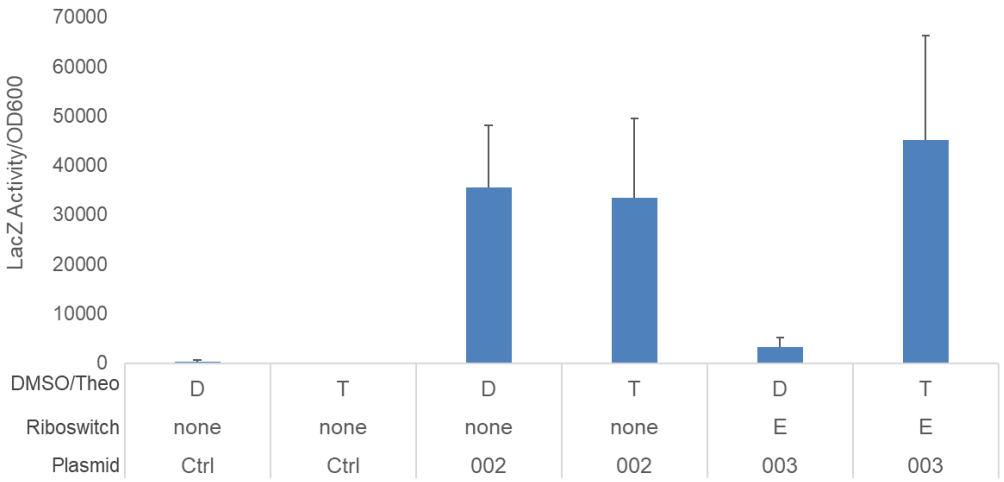

Fig. 2 Results of enzyme assays on *C. acetobutylicum* cultures harboring *lacZ* plasmids. Ctrl: *C. acetobutylicum* harboring plasmid pMTL85141 with no *lacZ* gene; 002: *C. acetobutylicum* harboring plasmid pAVT002 with *lacZ* expression driven by the constitutive thiolase promoter (P_{thl}-*lacZ*); 003: *C. acetobutylicum* harboring plasmid pAVT003 with *lacZ* expression driven by a fusion of the thiolase promoter and TheoRS^{E} (P_{thl}-TheoRS^{E}-*lacZ*); D: only DMSO provided; T: 2-mM theophylline provided. Error bars represent standard deviations: n=3.

Use of the fluorescent FDG substrate was an upgrade over the colorimetric X-Gal substrate (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside [results not shown]) for the prevention of false signals on the plate reader. However, the assay is quite laborious, requiring cell lysis and many other steps. An inherently fluorescent protein such as the green fluorescent protein (GFP), or a color-shifted variant such as red fluorescent protein, would be very advantageous to use as a reporter protein in *C. acetobutylicum*. These fluorescent proteins do not require cell lysis or use of chemical substrates and are compatible with a broad array of imaging and measurement techniques, including simple visualization under an appropriate light source, fluorescence-quantifying plate readers, and flow cytometry. However, this family of fluorescent proteins requires molecular oxygen for chromophore formation. It is possible to take an aliquot of cells from an anaerobic culture into the air, whereupon expressed GFP or similar protein can "recover" its fluorescence in approximately 2 h.[32] However, this time delay is also rather undesirable, and *C. acetobutylicum* cells exposed to ambient oxygen for that long will certainly all be dead, limiting some applications of this reporter protein.

I tested some alternative fluorescent reporter proteins that do not require oxygen for chromophore maturation. These proteins, PhiLOV[3] and Evoglow C-Bs2[22] have lower quantum yields than GFP and are thus not as bright. I found that PhiLOV did not provide sufficient green fluorescence signal even in *E. coli*, and also exhibited undetectably weak fluorescence signal in *C. acetobutylicum*. Others recently

obtained the same result.[4] Evoglow C-Bs2 did provide substantial visible green fluorescence when expressed in *E. coli*, but none was observed when it was expressed in *C. acetobutylicum* cells.

Like many other clostridia, *C. acetobutylicum* cells exhibit substantial green autofluorescence.[3–6] I observed this during flow cytometry of plasmid-free and uninduced *C. acetobutylicum* cultures on many occasions. It is thus desirable for a reporter protein expressed in *C. acetobutylicum* to fluoresce or otherwise emit its primary signal outside of the green wavelength range (~495–570 nm).

The next two reporter proteins I subcloned for testing in *C. acetobutylicum* were SNAP-Tag[5,25–28,33] and FAST.[23,24] SNAP-Tag is an engineered alkylguanine-DNA alkyltransferase enzyme that will form covalent bonds between itself and benzylguanine derivative substrates. Nothing about the protein or the reaction requires oxygen, so SNAP-Tag is fully compatible with anaerobic conditions. A variety of fluorescently labeled SNAP-Tag substrates are available on the market. I chose SNAP-Cell-647-SiR from New England Biolabs. This labeled SNAP-Tag substrate is compatible with our flow cytometer's 633-nm laser and APC (660/20 nm) emission filter. A nice feature of this fluorophore is that it fluoresces far outside the green wavelength range, so is not interfered with by *C. acetobutylicum* autofluorescence. I found that the combination of SNAP-Tag and SNAP-Cell-647-SiR worked fairly well in *C. acetobutylicum* but not in *E. coli*. This is likely due to the difference in membrane structure between the two microbes. Gram-negative *E. coli*, with its two phospholipid membranes, is known to be less permeable to many small molecules than Gram-positive bacteria such as *C. acetobutylicum.*[34,35]

Figure 3b shows flow cytometry histograms of uninduced versus induced *C. acetobutylicum* cells harboring plasmid pAVT017, which encodes the SNAP-Tag protein under control of P_{thl}-TheoRS^{E}. A similar induction ratio was observed for the riboswitch, about 10-fold, as that obtained for P_{thl}-TheoRS^{E}-*lacZ* (Fig. 2). The main downside to SNAP-Tag as a reporter protein is that multiple wash cycles are required to rid the cells of unbound fluorescent substrate before the measurement is taken. Because the substrate is fully fluorescent whether it is bound to SNAP-Tag protein or not, it will permeate equally into all cells, induced or not. For cells in which SNAP-Tag expression has not been induced, the low concentration of the enzyme will mean a low retention of substrate in the cell, but only if sufficient wash cycles are performed. These washes consist of repeated cycles of centrifugation, supernatant removal, resuspension in fresh buffer, and incubation time to allow for unbound substrate to diffuse out of cells. Because these washes are fairly laborious and time-consuming, especially for large numbers of samples, SNAP-Tag was an

effective but not very convenient or high-throughput anaerobic reporter protein for *C. acetobutylicum*.

FAST is an engineered variant of the photoactive yellow protein. The acronym stands for, "fluorescence-activating and absorption-shifting tag."[24] FAST binds to particular fluorogenic dyes in a noncovalent and reversible fashion, and the binding event itself converts the fluorogenic dye into a fluorescent protein-dye complex.[23,24] By varying the fluorogenic dye molecule, the fluorescence spectrum of the complex can be varied. As with SNAP-Tag, oxygen is not required for functional FAST protein to be produced, nor for the binding of fluorogen to it. I constructed plasmid pAVT012 featuring P_{thl}-TheoRS^{E}-FAST, and tested it in *C. acetobutylicum* with the fluorogen HBR-3,5-DOM. This fluorogen, when bound to the FAST protein, has a fluorescence excitation peak at 518 nm and an emission peak at 600 nm.[23] This combination is compatible with the PE channel of our flow cytometer (488-nm laser excitation; 585/40-nm emission filter).

Figure 3a shows flow cytometry results for uninduced and induced *C. acetobutylicum* cells harboring pAVT012. The observed induction ratio here was about 4. Although FAST with HBR-3,5-DOM sometimes functioned well, I found it to be inconsistent, and *C. acetobutylicum* autofluorescence would often "leak" into the PE channel of the flow cytometer. I therefore considered FAST with HBR-3,5-DOM (which is the furthest red-shifted FAST fluorogen currently described and commercially available) insufficiently reliable for use with *C. acetobutylicum*. Earlier this year, Streett et al. published their results with FAST and the HMBR fluorogen in *C. acetobutylicum*.[36] HMBR is less red-shifted than HBR-3,5-DOM. They conducted flow cytometry with a 488-nm excitation laser and a 530/30-nm filter appropriate for the HMBR fluorogen. The flow cytometry histograms in their report are very broad, and they prewashed their *C. acetobutylicum* cells several times before adding the fluorogen. Using the native thiolase promoter to drive FAST expression, they observed a small "tail" of fluorescent cells over time, but the peak of the fluorescence histogram did not move much, in contrast to Fig. 3a herein. Only when they switched to a "super" high-strength variant of the thiolase promoter did the histogram peaks shift to substantially higher fluorescence values, but this was accompanied by substantial broadening of the histograms. They did not couple FAST expression to an inducible promoter or riboswitch.

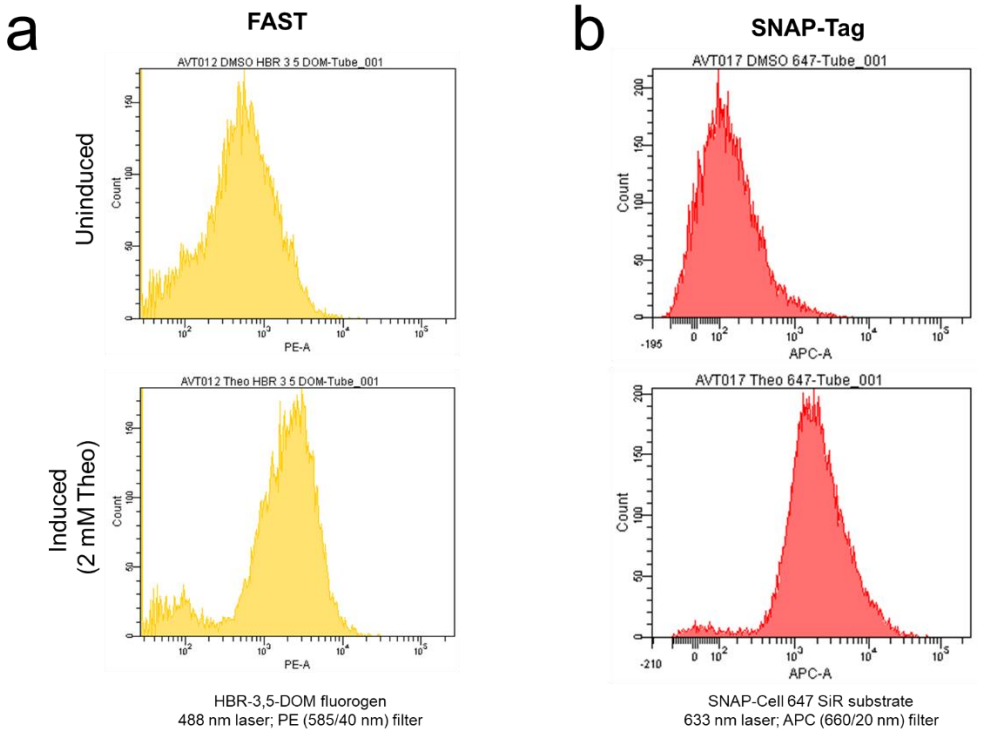

Fig. 3 Flow cytometry histograms of *C. acetobutylicum* cells harboring inducible genetic switches a) Cells transformed with plasmid pAVT012 with a theophylline riboswitch-inducible FAST protein; b) Cells transformed with plasmid pAVT017 with a riboswitch-inducible SNAP-Tag protein. Uninduced and induced labels apply to both a and b.

It seems fair to conclude that FAST can be a useful reporter protein in *C. acetobutylicum* in certain cases. Prewashing of cells, although laborious, may reduce the background green autofluorescence. The difference between induced and uninduced cells can be enhanced via use of a very strong promoter, when possible. By using a more modern flow cytometer, such a "spectral analyzer" with many more emission channels, the potential of FAST for use in *C. acetobutylicum* would probably be more fully realized. Finally, if a variant of FAST or a new fluorogen were developed that could be efficiently excited with a longer-wavelength laser, such as 561 nm, interference by green autofluorescence might be completely abolished.

Despite the shortcomings of both SNAP-Tag and FAST, I was able to use them together in the same *C. acetobutylicum* cells to demonstrate independent dual-switch, dual-reporter induction by two different ligands. This is significant because the ability to use two different inducer molecules would allow more complex genetic circuits to be constructed, such as "OR" or "AND" gates, which require two inputs. Figure 4 shows the flow cytometry results of *C. acetobutylicum* cells transformed with plasmid pAVT020. This plasmid features P_{thl}-TheoRS^{E}-SNAP-

Tag from pAVT017 and an aTc-inducible promoter system[10] that I subcloned to drive expression of FAST. The 2-D dot plots in Fig. 4 show effective independent (orthogonal) induction of each individual reporter by its cognate inducer molecule, as well as simultaneous induction of both reporter proteins from addition of both inducers together.

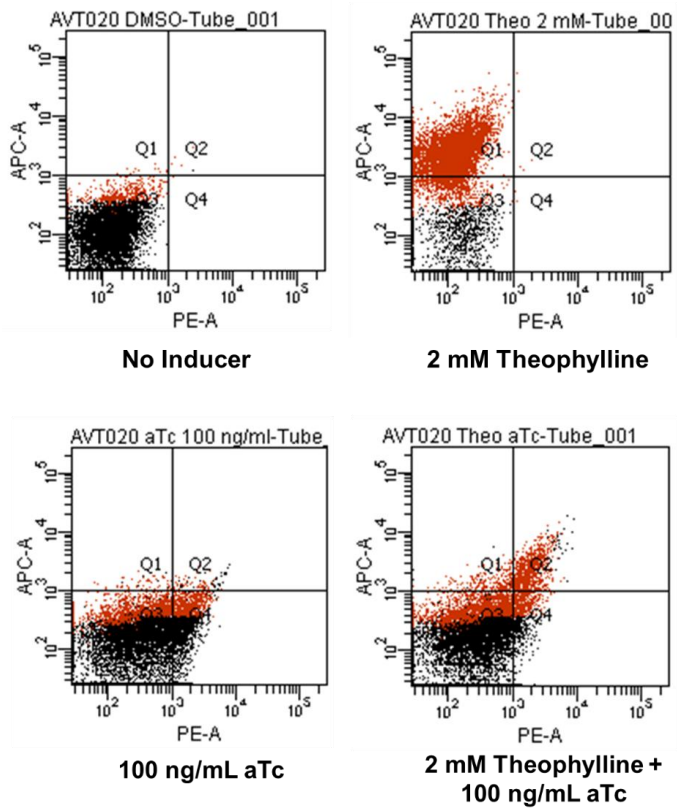

Fig. 4 Flow cytometry dot plots of *C. acetobutylicum* cells harboring dual-switch plasmid pAVT020. The plasmid consists of a theophylline riboswitch-inducible SNAP-Tag protein and an aTc promoter/repressor-inducible FAST protein. SNAP-Cell 647 SiR and HBR-3,5-DOM were added to all samples. The former fluoresces in the APC channel (vertical axis) while the latter fluoresces in the PE channel (horizontal axis). Therefore, the vertical axis represents SNAP-Tag induction by theophylline and the horizontal axis represents FAST induction by aTc.

The final reporter protein I tested in *C. acetobutylicum* was NanoLuc. This is an engineered variant of a luciferase enzyme from the deep-sea shrimp *Oplophorus gracilirostris*. Researchers at Promega developed the NanoLuc enzyme and its cognate substrate to be more than 100-fold brighter than other known luciferases.[31] A different luciferase was previously used as a reporter enzyme in *C. acetobutylicum*.[21] In that work, LucB, a luciferase from the firefly *Photinus pyralis* was evaluated. This luciferase requires adenosine triphosphate, acts on luciferin, a different substrate, and its reaction produces much less luminescence.

Luciferase-based luminescence is a visually striking phenomenon that has proven extremely valuable to science. For this application, because luminescence requires no incident excitation light source, it is immune to interference by the cellular autofluorescence that plagues *C. acetobutylicum* and many other microbes. It is simple to measure on most plate readers, and the luminescence from NanoLuc-furimazine reactions can easily be bright enough to see with the naked eye in a darkened laboratory, so instrumentation beyond a standard camera is not required.

I purchased synthetic DNA encoding the NanoLuc polypeptide, codon-adapted for *C. acetobutylicum*, and subcloned it after P_{thl}-TheoRS^{E} to generate plasmid pAVT074. I quickly found that NanoLuc was a very functional and convenient reporter enzyme for *C. acetobutylicum*. NanoLuc does require molecular oxygen for the luciferase reaction to proceed, but with the Nano-Glo Live Cell Assay System, the cells do not need to be lysed because the furimazine substrate rapidly penetrates into them to react with luciferase expressed therein. No centrifugation or washing steps are required either. Therefore, it is quite facile to perform rapid and fairly high-throughput measurements of NanoLuc expression on *C. acetobutylicum* cultures still growing in an anaerobic chamber by withdrawing a very small volume of cells for a plate-reader assay in the ambient laboratory air. The main disadvantages of NanoLuc assays are that they are not directly compatible with flow cytometry (unless they are fused to a fluorescent protein)[37] and because the reaction requires oxygen, it could be challenging (though not impossible) to maintain the viability of strictly anaerobic cells run through a luciferase assay.

With as little as a 2-µL aliquot of *C. acetobutylicum* cells, I was able to run NanoLuc assays that yielded blue luminescence visible to the naked eye for several minutes (Fig. 5). I performed some optimization of the assay to ensure that sufficient substrate was provided for accurate quantification of expressed luciferase (see Section 2.5).

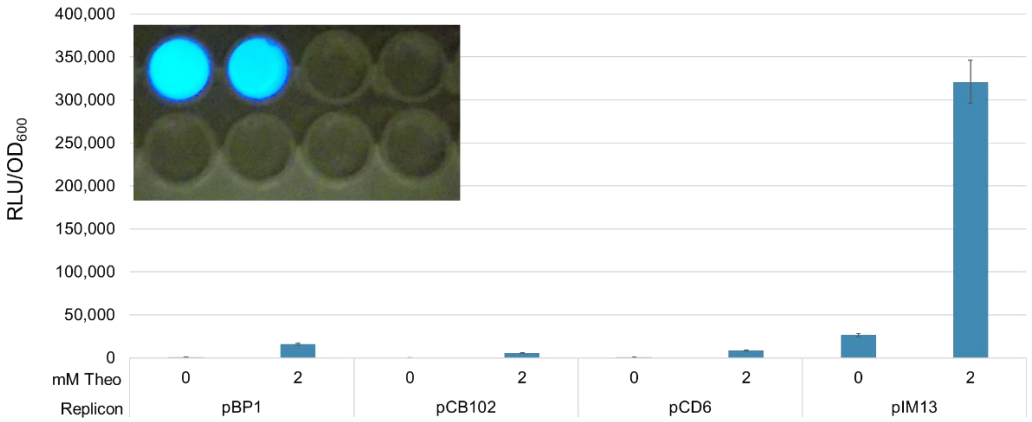

Fig. 5 Comparison of luciferase expression in *C. acetobutylicum* from plasmids containing four different Gram-positive origins of replication. The Gram-positive replicons included in the pMTL80000 set of modular plasmids[8] are: pBP1, pCB102, pCD6, and pIM13. The four plasmids compared all contained the same TheoRSE-NanoLuc cassette. Luminescence (RLU) was scaled by the OD600 nm of the cells in each well (proxy measurement of cell density). With this system, the values of RLU/OD600 at 2-mM theophylline (Theo) should be proportional to the average plasmid copy number per cell, which is a strong function of the origin of replication. Error bars represent standard deviations, n≥3. Inset: photograph of microplate wells in a NanoLuc assay showing two positive wells emitting bright blue light.

The riboswitch functioned well to drive NanoLuc expression upon induction with theophylline (Fig. 5). As part of a related study exploring the potential of *C. acetobutylicum* to biosynthesize the small molecule phloroglucinol[38], I generated variants of pAVT074 with the other three Gram-positive origins of replication included with the pMTL80000 modular plasmid system: pBP1, pCB102, and pCD6 (pAVT074, being based on pMTL-85141, contains the pIM13 replication origin). Different plasmid replication origins are characterized by their "copy number"—the average number of plasmid copies per cell, which is primarily determined by the identity of the origin. Copy number can vary widely between different origins. We measured relative copy number for these four plasmids as well as related plasmids expressing the phloroglucinol synthase instead of NanoLuc by quantitative polymerase chain reaction (qPCR). Not surprisingly, the qPCR results correlated just about exactly with the luminescence assay results shown in Fig. 5. The pIM13 origin that I had been using for nearly all the plasmids described in this report gave by far the highest copy number per cell, at about 24–32. Once again, an induction ratio of about 10 was observed for the theophylline riboswitch, regardless of plasmid origin.

Figure 6 is a plot of a related dataset that compares NanoLuc expression from high- and low-copy number plasmids as well as constitutive expression driven by P_{thl} versus inducible expression driven by P_{thl}-TheoRS^{E}. The results for pIM13 mirror

those shown in Fig. 2 with LacZ, which makes sense, as the measurements presented in Fig. 2 were of *C. acetobutylicum* cells harboring plasmids with the pIM13 origin. Specifically, reporter enzyme expression, whether LacZ or NanoLuc, upon induction of the riboswitch for the P_{thl}-TheoRS^{E} plasmids was equivalent to that observed with the constitutive P_{thl} plasmids. When this result was first observed with LacZ in Fig. 2, it led me to conclude that the riboswitch did not limit the expression level of the downstream protein, when the switch was fully induced.

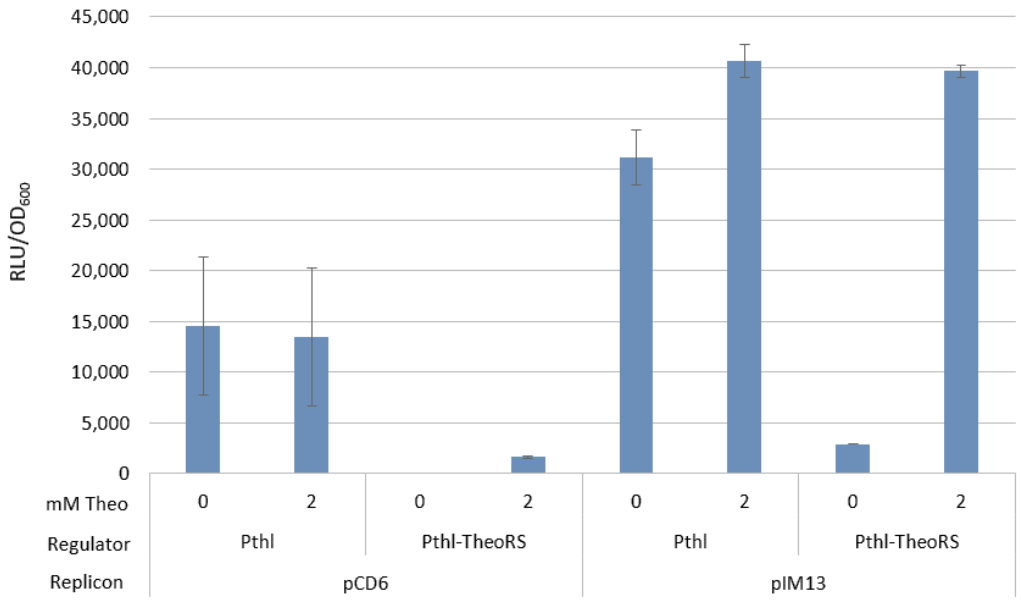

Fig. 6 Two-way comparison of replication origins and constitutive vs. riboswitch-induced NanoLuc expression. Low-copy origin pCD6 was compared with high-copy origin pIM13. For each origin, constitutive P_{thl} was compared with inducible P_{thl}-TheoRS. Note the large difference in maximal NanoLuc expression (RLU/OD600) between the constitutive and riboswitch expression systems for pCD6, compared with no difference in maximal expression for pIM13.

However, the results for low-copy origin pCD6 in Fig. 6 tell a different story. For these pCD6 plasmids, luciferase expression from the induced riboswitch plasmid was much lower than from the constitutive P_{thl} plasmid. This mirrors data published this year by Canadas et al. with the same riboswitch versus a constitutive promoter control on pBP1-based plasmids (low-copy) in *Clostridium sporogenes*.[39]

How should these divergent observations be reconciled? To me, the simplest possible explanation is that the riboswitch does limit protein expression compared to the same promoter without a riboswitch. The best plasmids with which to evaluate whether riboswitches limit downstream protein expression compared with constitutive promoters are low-copy plasmids, because they help ensure that all the other cellular components involved in protein synthesis are plentiful and not rate-

limiting. When plasmid copy number is increased 20-fold or more with a high-copy origin like pIM13, the cell's protein synthesis machinery can become taxed. This is a less ideal situation for assessing the effects of a riboswitch, and we would expect protein concentrations to be elevated but also "compressed" (have less variance), making it more difficult to isolate the effects of a riboswitch.

Does this mean that with pIM13 and P_{thl} we have reached the maximum possible rate of protein synthesis in *C. acetobutylicum*? Possibly, but not necessarily. Yang et al. developed mutants of P_{thl} such as 1200-9-9, which they measured to be about 12-fold stronger in terms of LacZ expression than P_{thl}.[40] This evaluation was done in *C. acetobutylicum* with plasmids that possess the low-copy pCB102 replicon, which is appropriate for evaluating the strength of strong promoters, as just discussed. Streett et al. from the Papoutsakis laboratory used promoter 1200-9-9 with a pIM13 origin in *C. acetobutylicum* to obtain about 2–3-fold higher expression levels of FAST compared with P_{thl}.[36] This implies that it is possible to obtain higher protein expression than with the pIM13/P_{thl} combination. However, there is a chance that the pIM13 I used and the pIM13 used by the Papoutsakis laboratory are not exactly the same. They measured a copy number of approximately 8 for their pIM13-based plasmids in *C. acetobutylicum*,[41] which is about a third to a quarter as high as the copy number we measured by qPCR.

I am comfortable concluding that the $TheoRS^E$ riboswitch does have a limiting effect on downstream protein synthesis compared with the same riboswitch-free promoter, as Canadas et al. and I both observed this with low-copy plasmids.[39] I am not yet prepared to comment on whether or not it is possible to exceed the level of protein synthesis in *C. acetobutylicum* we have obtained with the pIM13/P_{thl} combination. However, this would be fairly straightforward to test by comparing NanoLuc expression from pIM13/P_{thl} with that from pIM13/1200-9-9. Unfortunately, I was not able to run this experiment before this project concluded.

3.3 Evaluation of Other Riboswitches and Inducible Promoters in *C. acetobutylicum*

Four other riboswitches were tested in *C. acetobutylicum* after the theophylline riboswitch. I was not able to observe inducible gene expression from any of them in *C. acetobutylicum* using SNAP-Tag as the reporter protein. Table 2 summarizes these other tested riboswitches.

Table 2 Other riboswitch systems evaluated in *C. acetobutylicum*

Name	Inducer	Reference	Notes
add A	adenine	42	Use of a natural purine as inducer not ideal; could be metabolized
M6'	ammeline	43	Mutant of *add* A. Inducer poorly soluble.
pbuE(Δ1-11)	2-aminopurine	44	Mutant of *pbuE* adenine riboswitch from *Bacillus subtilis*
*pbuE/pbuE**(6U)	2-aminopurine	44	Mutant of *pbuE* adenine riboswitch from *Bacillus subtilis*
xpt(C74U)*/pbuE**(7U)	2-aminopurine	44	Chimera of mutants of *xpt* and *pbuE* guanine riboswitches

After the riboswitches, I tested seven inducible promoter systems that I adapted for *C. acetobutylicum*, which I "sourced" from a recent publication about "Marionette," an engineered strain of *E. coli* that was endowed with 12 small-molecule sensors.[12] Since these sensors are all based on inducible promoters, they all consist of a regulator protein (either a repressor or an activator) that binds to the small molecule inducer and changes conformation to alter its DNA-binding affinity, and an associated DNA promoter sequence to which the regulator binds. The amount of synthetic DNA required for these inducible promoter systems is thus greater than for riboswitches, because the regulator protein sequences are several hundred to nearly 1000 base pairs long, versus about 80–150 nucleotides for a riboswitch. Nevertheless, being protein-based, these systems generally have greater affinity and selectivity toward small molecule ligands, so their performance could be superior to that of riboswitches.

I constructed plasmids encoding the seven regulator-promoter systems, and for each system, I made one plasmid with FAST as the reporter protein and one with SNAP-Tag as the reporter protein. I codon-adapted the regulator protein sequences for *C. acetobutylicum*, but left the promoter sequences unaltered. I first tested each inducible promoter system in *E. coli*, then tested them in *C. acetobutylicum*. Table 3 qualitatively summarizes these results. Unfortunately, despite some of the systems working well in *E. coli*, none of them functioned well in *C. acetobutylicum*. The aTc system is similar to the one I used in pAVT020 (Fig. 4), but has differences as well. I was not able to perform much troubleshooting in *C. acetobutylicum* to diagnose the source of the lack of function.

Table 3 Adapted "Marionette"[12] inducible promoter systems evaluated in *C. acetobutylicum*

Inducer	Type of regulator	Induction performance in *E. coli*	Induction performance in *C. acetobutylicum*
Diacetylphloroglucinol (DAPG)	Repressor	Moderate to good	No induction observed
Cuminic acid	Repressor	No induction observed	Moderate-to-high constitutive-on phenotype. Induction not observed.
$3OC_6$-HSL	Activator	No induction observed	Moderate-to-high constitutive-on phenotype. Induction not observed.
Vanillic acid	Repressor	Good	Moderate-to-high constitutive-on phenotype. Induction not observed.
aTc	Repressor	Moderate to good	Moderate-to-high constitutive-on phenotype. Small amount of induction possibly observed, but hard to tell due to high leakiness.
Salicylate	Activator	Excellent	Moderate-to-high constitutive-on phenotype. Induction not observed.
$3OHC_{14}$-HSL	Activator	Moderate	Moderate-to-high constitutive-on phenotype. Induction not observed.

If I had an opportunity to revisit this effort, I would use NanoLuc as the reporter and would focus first on repressor systems for which I could "port" the operator (short section of DNA to which the repressor protein binds). These repressor systems have a good chance of functioning with native *C. acetobutylicum* promoters or other promoters known to work well in *C. acetobutylicum*. However, *C. acetobutylicum* is quite different than *E. coli* in terms of its metabolism, and more specifically, its propensity to degrade small molecules. So some of these inducers may be metabolically incompatible with *C. acetobutylicum*, unless a strain can be developed that does not degrade or metabolize them (see Section 4).

4. Conclusions

This report details my efforts to "domesticate" *C. acetobutylicum* to make it more "engineerable" and useful to researchers interested in controlling its behavior or metabolism. As this report exemplifies, the number of genetic parts that must be tested to find suitably functional ones can be large and the success rate can be low. I was able to make some progress in developing the theophylline riboswitch and several reporter proteins to function well in this chassis microbe, but much work remains to render this bacterium even one-tenth as engineerable as *E. coli* or

S. cerevisiae in terms of the time and effort to complete a design–build–test cycle and the number of genetic tools available.

For example, plasmid transformation efficiency is low with the type strain *C. acetobutylicum* ATCC 824 because of known restriction enzymes it possesses that cut foreign DNA as a defense mechanism.[45,46] *C. acetobutylicum* strains in which the main restriction enzymes have been deleted have been reported, but are not easily obtained, for example, from public repositories. So it can be difficult and laborious to obtain a lot of colonies when *C. acetobutylicum* is transformed with plasmid DNA and this can make many types of experiments challenging, especially those where the researcher needs to search through "libraries" of random or partially random variants of protein or regulatory DNA sequences to find rare ones with desired function.

Three of the reporter proteins that function in *C. acetobutylicum*—FAST, SNAP-Tag, and NanoLuc—are small and able to be used as fusion tags to track other proteins of interest.[5,23–28,33,47,48] This bodes well for the possibility of using these reporters to help quantify or follow other proteins such as metabolic enzymes, heterologously expressed proteins, surface proteins, or others, in a straightforward and high-throughput fashion.

During the related project mentioned earlier to engineer *C. acetobutylicum* to biosynthesize the small molecule phloroglucinol, we discovered that *C. acetobutylicum* was degrading or metabolizing phloroglucinol via an unknown enzyme or mechanism. Depending on one's perspective, clostridia are famous or notorious for this type of phenomenon.[49] Therefore, an early part of any testing regimen when engineering clostridia to produce or respond to a small molecule should be an exogenous recovery study, whereupon the molecule is added to a culture and then the culture supernatant is recovered after some time and analyzed for the concentration of the added molecule.

As mentioned in Section 3.1, other inducible systems for gene expression have been developed that respond to nonchemical stimuli such as light. Such a system, if adapted for clostridia or other related microbes, would add significantly to the synthetic biology toolbox for these organisms and might find important practical usage.

The field of synthetic biology has embraced many novel and nonstandard microbial chassis organisms and risen to the challenge of engineering them for impressive feats of biosynthesis, biosensing, remediation, and information processing. As more tools are developed to endow these organisms with new functions and to enable more facile and rapid subsequent engineering of these organisms, we can

look forward to a future in which biological technologies will be increasingly sought and leveraged to tackle the world's most urgent challenges.

5. References

1. Gyulev IS, Willson BJ, Hennessy RC, Krabben P, Jenkinson ER, Thomas GH. Part by part: synthetic biology parts used in solventogenic clostridia. ACS Synth Biol. 2018;7(2):311–327.

2. Joseph RC, Kim NM, Sandoval NR. Recent developments of the synthetic biology toolkit for *Clostridium*. Front Microbiol. 2018;9:154.

3. Buckley AM, Jukes C, Candlish D, Irvine JJ, Spencer J, Fagan RP, Roe AJ, Christie JM, Fairweather NF, Douce GR. Lighting up *Clostridium difficile*: reporting gene expression using fluorescent LOV domains. Sci Rep. 2016;6:23463.

4. Mordaka PM, Heap JT. Stringency of synthetic promoter sequences in *Clostridium* revealed and circumvented by tuning promoter library mutation rates. ACS Synth Biol. 2018;7(2):672–681.

5. Pereira FC, Saujet L, Tome AR, Serrano M, Monot M, Couture-Tosi E, Martin-Verstraete I, Dupuy B, Henriques AO. The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. PLoS Genet. 2013;9(10):e1003782.

6. Ransom EM, Ellermeier CD, Weiss DS. Use of mCherryOpt fluorescent protein in *Clostridium difficile*. Methods Mol Biol. 2016;1476:53–67.

7. Servinsky MD, Liu S, Gerlach ES, Germane KL, Sund CJ. Fermentation of oxidized hexose derivatives by *Clostridium acetobutylicum*. Microb Cell Fact. 2014;13:139.

8. Heap JT, Pennington OJ, Cartman ST, Minton NP. A modular system for *Clostridium* shuttle plasmids. J Microbiol Methods. 2009;78(1):79–85.

9. Oliveira PAM, Friggen, AH, Hossein-Javaheri S, Smits WK. The signal sequence of the abundant extracellular metalloprotease PPEP-1 can be used to secrete synthetic reporter proteins in *Clostridium difficile*. ACS Synth Biol. 2016;5(12):1376–1382.

10. Fagan RP, Fairweather NF. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem. 2011;286(31):27483–93.

11. Mermelstein LD, Papoutsakis ET. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi-3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol. 1993;59(4):1077–81.

12. Meyer AJ, Segall-Shapiro TH, Glassey E, Zhang J, Voigt CA. *Escherichia coli* "Marionette" strains with 12 highly optimized small-molecule sensors. Nat Chem Biol. 2019;15(2):196–204.

13. Baumschlager A, Aoki SK, Khammash M. Dynamic blue light-inducible T7 RNA polymerases (Opto-T7RNAPs) for precise spatiotemporal gene expression control. ACS Synth Biol. 2017;6(11):2157–2167.

14. Ohlendorf R, Vidavski RR, Eldar A, Moffat K, Moglich A. From dusk till dawn: one-plasmid systems for light-regulated gene expression. J Mol Biol. 2012;416(4);534–42.

15. Tschirhart T, Kim E, McKay R, Ueda H, Wu HC, Pottash AE, Zargar A, Negrete A, Shiloach J, Payne GF, et al. Electronic control of gene expression and cell behaviour in *Escherichia coli* through redox signalling. Nat Commun. 2017;8:14030.

16. Andrews LB, Nielsen AAK, Voigt CA. Cellular checkpoint control using programmable sequential logic. Science. 2018;361:6408.

17. Moser F, Espah Borujeni A, Ghodasara AN, Cameron E, Park Y, Voigt CA. Dynamic control of endogenous metabolism with combinatorial logic circuits. Mol Syst Biol. 2018;14(11):e8605.

18. Dong H, Tao W, Zhang Y, Li Y. Development of an anhydrotetracycline-inducible gene expression system for solvent-producing *Clostridium acetobutylicum*: a useful tool for strain engineering. Metab Eng. 2012;14(1):59–67.

19. Topp S, Reynoso CM, Seeliger JC, Goldlust IS, Desai SK, Murat D, Shen A, Puri AW, Komeili A, Bertozzi CR, et al. Synthetic riboswitches that induce gene expression in diverse bacterial species. Appl Environ Microbiol. 2010;76(23):7881–4.

20. Tummala SB, Welker NE, Papoutsakis ET. Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol. 1999;65(9):3793–9.

21. Feustel L, Nakotte S, Durre P. Characterization and development of two reporter gene systems for Clostridium acetobutylicum. Appl Environ Microbiol. 2004;70(2):798–803.

22. Drepper T, Eggert T, Circolone F, Heck A, Krauss U, Guterl JK, Wendorff M, Losi A, Gartner W, Jaeger KE. Reporter proteins for in vivo fluorescence without oxygen. Nat Biotechnol. 2007;25(4):443–5.

23. Li C, Plamont MA, Sladitschek HL, Rodrigues V, Aujard I, Neveu P, Le Saux T, Jullien L, Gautier A. Dynamic multicolor protein labeling in living cells. Chem Sci. 2017;(8):5598–5605.

24. Plamont MA, Billon-Denis E, Maurin S, Gauron C, Pimenta FM, Specht CG, Shi J, Querard J, Pan B, Rossignol J, et al. Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. Proc Natl Acad Sci U S A. 2016;113(3):497–502.

25. Cassona CP, Pereira F, Serrano M, Henriques AO. A fluorescent reporter for single cell analysis of gene expression in *Clostridium difficile*. Methods Mol Biol. 2016;1476:69–90.

26. Cole NB. Site-specific protein labeling with SNAP-Tags. Curr Protoc Protein Sci. 2013;73:Unit 30 1.

27. Gautier A, Juillerat A, Heinis C, Correa IR Jr, Kindermann M, Beaufils F, Johnsson K. An engineered protein tag for multiprotein labeling in living cells. Chem Biol. 2008;15(2):128–36.

28. Pereira FC. Spore differentiation in relation to the infectious cycle of the enteric pathogen *Clostridium difficile* [dissertation]. [Oeiras, Lisbon (Portugal)]: Universidade Nova de Lisboa; 2014.

29. Chen S, Kaufman MG, Korir ML, Walker ED. Ingestibility, digestibility, and engineered biological control potential of *Flavobacterium hibernum*, isolated from larval mosquito habitats. Appl Environ Microbiol. 2014;80(3):1150–8.

30. Mimee M, Tucker AC, Voigt CA, Lu TK. Programming a human commensal bacterium, *Bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota. Cell Syst. 2015;1(1):62–71.

31. Hall MP, Unch J, Binkowski BF, Valley MP, Butler BL, Wood MG, Otto P, Zimmerman K, Vidugiris G, Machleidt T, et al. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS Chem Biol. 2012;7(11):1848–57.

32. Zhang C, Xing XH, Lou K. Rapid detection of a gfp-marked *Enterobacter aerogenes* under anaerobic conditions by aerobic fluorescence recovery. FEMS Microbiol Lett. 2005;249(2):211–8.

33. Gronemeyer T, Chidley C, Juillerat A, Heinis C, Johnsson K. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. Protein Eng Des Sel. 2006;19(7):309–16.

34. Henriques AO. New University of Lisbon, Portugal. Personal communication, 2018 Apr 17.

35. Scorciapino MA, Acosta-Gutierrez S, Benkerrou D, D'Agostino T, Malloci G, Samanta S, Bodrenko I, Ceccarelli M. Rationalizing the permeation of polar antibiotics into Gram-negative bacteria. J Phys Condens Matter. 2017;29(11):113001.

36. Streett HE, Kalis KM, Papoutsakis ET. A strongly fluorescing anaerobic reporter and protein-tagging system for Clostridium organisms based on the fluorescence-activating and absorption-shifting tag protein (FAST). Appl Environ Microbiol. 2019;85(14).

37. Schaub FX, Reza MS, Flaveny CA, Li W, Musicant AM, Hoxha S, Guo M, Cleveland JL, Amelio AL. Fluorophore-nanoLuc BRET reporters enable sensitive in vivo optical imaging and flow cytometry for monitoring tumorigenesis. Cancer Res. 2015;75(23):5023–33.

38. Meyer AJ, Saaem I, Silverman A, Varaljay V, Mickol R, Blum SM, Tobias AV, Schwalm ND, Mojadedi W, Onderko E, et al. Organism engineering for the bioproduction of the triaminotrinitrobenzene (TATB) precursor phloroglucinol (PG). ACS Synth Biol. 2019 Nov 21.

39. Canadas IC, Groothuis D, Zygouropoulou M, Rodrigues R, Minton NP. RiboCas: a universal CRISPR-based editing tool for *Clostridium*. ACS Synth Biol. 2019;8(6):1379–1390.

40. Yang G, Jia D, Jin L, Jiang Y, Wang Y, Jiang W, Gu Y. Rapid generation of universal synthetic promoters for controlled gene expression in both gas-fermenting and saccharolytic *Clostridium* species. ACS Synth Biol. 2017;6(9):1672–1678.

41. Lee SY, Mermelstein LD, Papoutsakis ET. Determination of plasmid copy number and stability in *Clostridium acetobutylicum* ATCC 824. FEMS Microbiol Lett. 1993;108(3):319–23.

42. Serganov A, Yuan YR, Pikovskaya O, Polonskaia A, Malinina L, Phan AT, Hobartner C, Micura R, Breaker RR, Patel DJ. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. Chem Biol. 2004;11(12):1729–41.

43. Dixon N, Duncan JN, Geerlings T, Dunstan MS, McCarthy JE, Leys D, Micklefield J. Reengineering orthogonally selective riboswitches. Proc Natl Acad Sci USA. 2010;107(7):2830–5.

44. Ceres P, Trausch JJ, Batey RT. Engineering modular ‘ON’ RNA switches using biological components. Nucleic Acids Res. 2013;41(22):10449–61.

45. Croux C, Nguyen NP, Lee J, Raynaud C, Saint-Prix F, Gonzalez-Pajuelo M, Meynial-Salles I, Soucaille P. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. Biotechnol Biofuels. 2016;9:23.

46. Dong H, Zhang Y, Dai Z, Li Y. Engineering *Clostridium* strain to accept unmethylated DNA. PLoS One. 2010;5(2):e9038.

47. Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K. A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat Biotechnol. 2003;21(1):86–9.

48. Sun X, Zhang A, Baker B, Sun L, Howard A, Buswell J, Maurel D, Masharina A, Johnsson K, Noren CJ, et al. Development of SNAP-Tag fluorogenic probes for wash-free fluorescence imaging. Chembiochem. 2011;12(14):2217–26.

49. Regan KM, Crawford RL. Characterization of Clostridium bifermentans and its biotransformation of 2,4,6-trinitrotoluene (TNT) and 1,3,5-triaza-1,3,5-trinitrocyclohexane (RDX). Biotechnol Lett. 1994;16(10):1081–1086.

Appendix. Table of Plasmids

Table A-1 Plasmids used in this study

Plasmid	Regulon	Reporter	Gram-positive replication origin	Intended usage
pMTL85141[1]	None	None	pIM13	Negative control (empty plasmid)
pAVT002	P_{thl}[1,2]	LacZ *(Tt)*	pIM13	Constitutive expression of LacZ
pAVT003	P_{thl}-TheoRS^{E} [(3)]	LacZ *(Tt)*	pIM13	Theophylline-induced expression of LacZ
pAVT011	P_{thl}-TheoRS^{E}	Evoglow C-Bs2[4,5]	pIM13	Theophylline-induced expression of Evoglow C-Bs2
pAVT012	P_{thl}-TheoRS^{E}	FAST[6,7]	pIM13	Theophylline-induced expression of FAST
pAVT013	P_{thl}-TheoRS^{E}	PhiLOV2.1[8]	pIM13	Theophylline-induced expression of PhiLOV2.1
pAVT017	P_{thl}-TheoRS^{E}	SNAP-Tag[9,10]	pIM13	Theophylline-induced expression of SNAP-Tag
pAVT020	P_{thl}-TheoRS^{E}	SNAP-Tag	pIM13	Theophylline-induced expression of SNAP-Tag and aTc-induced expression of FAST
	P_{tet}	FAST		
pAVT074	P_{thl}-TheoRS^{E}	NanoLuc[11]	pIM13	Theophylline-induced expression of NanoLuc
pAVT075	P_{thl}	NanoLuc	pIM13	Constitutive expression of NanoLuc
pAVT076	P_{thl}-TheoRS^{E}	NanoLuc	pBP1	Theophylline-induced expression of NanoLuc
pAVT077	P_{thl}-TheoRS^{E}	NanoLuc	pCB102	Theophylline-induced expression of NanoLuc
pAVT078	P_{thl}-TheoRS^{E}	NanoLuc	pCD6	Theophylline-induced expression of NanoLuc
pAVT086	P_{thl}	NanoLuc	pBP1	Constitutive expression of NanoLuc
pAVT087	P_{thl}	NanoLuc	pCB102	Constitutive expression of NanoLuc
pAVT088	P_{thl}	NanoLuc	pCD6	Constitutive expression of NanoLuc
pAVT023	P_{thl}(1200-6)-M6' [(2,12)]	SNAP-Tag	pIM13	Ammeline-induced expression of SNAP-Tag
pAVT026	P_{thl}(1200-6)-*add*A[13]	…	pIM13	Adenine-induced expression of SNAP-Tag
pAVT032	P_{fdx}[1]	FAST	pIM13	Constitutive expression of FAST
pAVT033	P_{fdx}-*pbuE*(Δ1-11) [14]	FAST	pIM13	2-aminopurine-induced expression of FAST
pAVT034	P_{fdx}-*pbuE/pbuE**(6U) [14]	FAST	pIM13	2-aminopurine-induced expression of FAST
pAVT035	P_{fdx}-*xpt*(C74U)/ *pbuE**(7U)[14]	FAST	pIM13	2-aminopurine-induced expression of FAST

Table A-1 Plasmids used in this study (continued)

Plasmid	Regulon	Reporter	Gram-positive replication origin	Intended usage
pAVT036	PhlF^{AM} (DAPG response regulator)[15]	FAST	pIM13	Diacetylphloroglucinol-induced expression of FAST
pAVT037	CymR^{AM} (cuminic acid response regulator)[15]	FAST	pIM13	Cuminic acid-induced expression of FAST
pAVT038	LuxR (3OC$_6$-HSL response regulator)[15]	FAST	pIM13	3OC$_6$-HSL-induced expression of FAST
pAVT039	VanR^{AM} (vanillic acid response regulator)[15]	FAST	pIM13	Vanillic acid-induced expression of FAST
pAVT040	TetR (aTc response regulator)[15]	FAST	pIM13	aTc-induced expression of FAST
pAVT041	NahR^{AM} (sodium salicylate response regulator)[15]	FAST	pIM13	Salicylate-induced expression of FAST
pAVT042	CinR^{AM} (3OHC$_{14}$-HSL response regulator)[15]	FAST	pIM13	3OHC$_{14}$-HSL-induced expression of FAST
pAVT047	PhlF^{AM} (DAPG response regulator)[15]	SNAP-Tag	pIM13	Diacetylphloroglucinol-induced expression of SNAP-Tag
pAVT048	CymR^{AM} (cuminic acid response regulator)[15]	SNAP-Tag	pIM13	Cuminic acid-induced expression of SNAP-Tag
pAVT049	LuxR (3OC$_6$-HSL response regulator)[15]	SNAP-Tag	pIM13	3OC$_6$-HSL-induced expression of SNAP-Tag
pAVT050	VanR^{AM} (vanillic acid response regulator)[15]	SNAP-Tag	pIM13	Vanillic acid-induced expression of SNAP-Tag
pAVT051	TetR (aTc response regulator)[15]	SNAP-Tag	pIM13	aTc-induced expression of SNAP-Tag
pAVT052	NahR^{AM} (sodium salicylate response regulator)[15]	SNAP-Tag	pIM13	Salicylate-induced expression of SNAP-Tag
pAVT053	CinR^{AM} (3OHC$_{14}$-HSL response regulator)[15]	SNAP-Tag	pIM13	3OHC$_{14}$-HSL-induced expression of SNAP-Tag

A.1 Appendix References

1. Heap JT, Pennington OJ, Cartman ST, Minton NP. A modular system for *Clostridium* shuttle plasmids. J Microbiol Methods. 2009;78(1):79–85.

2. Yang G, Jia D, Jin L, Jiang Y, Wang Y, Jiang W, Gu Y. Rapid generation of universal synthetic promoters for controlled gene expression in both gas-fermenting and saccharolytic *Clostridium* species. ACS Synth Biol. 2017;6(9):1672–1678.

3. Topp S, Reynoso CM, Seeliger JC, Goldlust IS, Desai SK, Murat D, Shen A, Puri AW, Komeili A, Bertozzi CR, et al. Synthetic riboswitches that induce gene expression in diverse bacterial species. Appl Environ Microbiol. 2010;76(23):7881–4.

4. Evoglow® *Clostridia* Basic Kit product manual (#FP-21060). Düsseldorf (Germany): Evocatal GmbH; 2009.

5. Drepper T, Eggert T, Circolone F, Heck A, Krauss U, Guterl, JK, Wendorff M, Losi A, Gartner W, Jaeger KE. Reporter proteins for in vivo fluorescence without oxygen. Nat Biotechnol. 2007;25(4):443–5.

6. Li C, Plamont MA, Sladitschek HL, Rodrigues V, Aujard I, Neveu P, Le Saux T, Jullien L, Gautier A. Dynamic multicolor protein labeling in living cells. Chem Sci. 2017;8(8):5598–5605.

7. Plamont MA, Billon-Denis E, Maurin S, Gauron C, Pimenta FM, Specht CG, Shi J, Querard J, Pan B, Rossignol J, et al. Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. Proc Natl Acad Sci U S A. 2016;113(3):497–502.

8. Buckley AM, Jukes C, Candlish D, Irvine JJ, Spencer J, Fagan RP, Roe AJ, Christie JM, Fairweather NF, Douce GR. Lighting up *Clostridium difficile*: reporting gene expression using fluorescent LOV domains. Sci Rep. 2016;6:23463.

9. Cassona CP, Pereira F, Serrano M, Henriques AO. A fluorescent reporter for single cell analysis of gene expression in *Clostridium difficile*. Methods Mol Biol. 2016;1476:69–90.

10. Pereira FC, Saujet L, Tome AR, Serrano M, Monot M, Couture-Tosi E, Martin-Verstraete I, Dupuy B, Henriques AO. The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. PLoS Genet. 2013;9(10):e1003782.

11. Hall MP, Unch J, Binkowski BF, Valley MP, Butler BL, Wood MG, Otto P, Zimmerman K, Vidugiris G, Machleidt T, et al. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS Chem Biol. 2012;7(11):1848–57.

12. Dixon N, Duncan JN, Geerlings T, Dunstan MS, McCarthy JE, Leys D, Micklefield J. Reengineering orthogonally selective riboswitches. Proc Natl Acad Sci U S A. 2010;107(7):2830–5.

13. Serganov A, Yuan YR, Pikovskaya O, Polonskaia A, Malinina L, Phan AT, Hobartner C, Micura R, Breaker RR, Patel DJ. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. Chem Biol. 2004;11(12):1729–41.

14. Ceres P, Trausch JJ, Batey RT. Engineering modular 'ON' RNA switches using biological components. Nucleic Acids Res. 2013;41(22):10449–61.

15. Meyer AJ, Segall-Shapiro TH, Glassey E, Zhang J, Voigt CA. *Escherichia coli* "Marionette" strains with 12 highly optimized small-molecule sensors. Nat Chem Biol. 2019;15(2):196–204.

List of Symbols, Abbreviations, and Acronyms

2-D	two-dimensional
APC	allophycocyanin
aTc	anhydrotetracycline
C. acetobutylicum	*Clostridium acetobutylicum*
CGM	clostridial growth medium
CRISPR	clusters of regularly interspaced short palindromic repeats
DAPG	2,4-diacetylphloroglucinol
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DOD	Department of Defense
E. coli	*Escherichia coli*
FAST	fluorescence-activating and absorption-shifting tag
FDG	fluorescein di-β-D-galactopyranoside
GFP	green fluorescent protein
HBR-3,5-DOM	4-hydroxy-3,5-dimethoxybenzylidene rhodanine
LacZ	beta-galactosidase
LB	Luria-Bertani
pBP1	replication origin from *Clostridium botulinum* plasmid
pCB102	replication origin from *Clostridium butyricum* plasmid
pCD6	replication origin from *Clostridium difficile* plasmid
PE	phycoerythrin
pH	potential hydrogen
pIM13	replication origin from *Bacillus subtilis* plasmid
qPCR	quantitative polymerase chain reaction
RBS	ribosomal binding site
RLU	relative luminescence units
S. cerevisiae	*Saccharomyces cerevisiae*
TheoRS^{E}	theophylline riboswitch “E”

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	CCDC ARL FCDD RLD DCI TECH LIB
5 (PDF)	CCDC ARL FCDD RLS CB A TOBIAS JJ SUMNER CJ SUND JD GOLLIHAR KL AKINGBADE
1 (PDF)	CCDC ARL FCDD RLS EC SHAFFER
2 (PDF)	USAFRL RR NAIK C SULLIVAN
1 (PDF)	USNRL G VORA