Engineered Regulon to Enable Autonomous Azide Ion Biosensing, Recombinant Protein Production, and *In vivo* Glycoengineering

Chandra Kanth Bandi,[a] Kyle S. Skalenko,[b] Ayushi Agrawal,[a] Neelan Sivaneri,[a] Margaux Thiry,[a] and Shishir P. S. Chundawat*[a]

[a] Department of Chemical & Biochemical Engineering, Rutgers, The State University of New Jersey, 98 Brett Road, Piscataway, New Jersey 08854, USA

[b] Department of Genetics and Waksman Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Rd, Piscataway, New Jersey 08854, USA

KEYWORDS: azide biosensor • promoter engineering • protein expression • glycoengineering • chemical-biology

ABSTRACT: Detection of azide-tagged biomolecules (e.g., azido sugars) inside living cells using 'click' chemistry has been revolutionary to the field of chemical-biology. However, we currently still lack suitable synthetic biology tools to autonomously and rapidly detect azide ions. Here, we have developed an engineered synthetic promoter system called *cyn* regulon, and complementary *Escherichia coli* engineered strains, to selectively detect azide ions and autonomously induce downstream expression of reporter genes. The engineered *cyn* azide operon allowed highly-tunable reporter green fluorescent protein (GFP) expression over three orders of inducer azide ion concentrations (0.01-5 mM) and rapidly induce GFP expression by over 600-fold compared to uninduced control. Next, we showcase the superior performance of this engineered *cyn*-operon over the classical *lac*-operon for recombinant protein production. Finally, we highlight how this synthetic biology toolkit can enable glycoengineering-based applications by facilitating *in-vivo* activity screening of mutant carbohydrate-active enzymes (CAZymes), called glycosynthases, using azido sugars as donor substrates.

Introduction

Azide ion is an excellent nucleophile that readily participates in nucleophilic substitution reactions.[1] This has facilitated the discovery and synthesis of organic azides that have been ubiquitously used as building blocks in organic chemistry for over 150 years.[2] Azides have found widespread commercial applications as automobile airbag propellants, biocides, and as functional groups in fine chemicals and pharmaceutical drugs synthesis.[3,4] In particular, organic azides have been clinically approved as potent antibiotics and pharmaceutical drugs (e.g., azidamfenicol, azidocillin, and zidovudine).[5–8] With the invention of biorthogonal click chemistry[9] that utilizes alkyne-azide chemistry for diverse applications in chemical-biology, drug development, and bioorganic chemistry, there has been a reemergence of interest towards azides within the scientific community. Azide functionalized molecular probes are being developed as imaging reagents and biomarkers for diseased cell/tissue detection,[10] pharmaceutical drugs delivery,[11] as well as glycomics, metabolomics,[12] and chemoproteomics related applications.[13] However, *in vivo* use of azido-based drugs or molecular imaging probes necessitates monitoring organic azide stability towards degradative release of azide ions.[14] To date, spectrophotometry,[15,16] spectrometry,[17,18] fluorescence,[19–22] and redox sensing[23] have been used for *in vitro* azide ion detection. However, these methods require extensive sample derivatization and are not suited for real-time *in vivo* detection.

Glycosyl azides have also facilitated *in vivo* imaging of carbohydrates to improve our understanding of the function, localization, and metabolic role of glycans.[24] Glycosyl azides are also used as donor substrates for engineered glycosidases called glycosynthases to synthesize complex glycans.[25,26] Compared to other donor substrates (e.g., glycosyl fluorides and p-nitrophenyl glycosides), glycosyl azides carry a small azide leaving group with strong nucleophilic character making them ideal substrates for glycoengineering applications. However, there are limited *in vivo* options available for rapid azide detection unlike other leaving groups (e.g., p-nitrophenyl or pNP).[27] This has limited the widespread use of glycosyl azides as standard reagents in glycosciences unlike pNP-glycosides. Therefore, availability of a synthetic biology toolkit for rapid *in vivo* azide ion detection can broadly enable the fields of chemical-biology and glycoengineering.

Here, we have developed a biosensing toolkit to selectively and autonomously detect azide ions inside living cells. Briefly, an *Escherichia coli* operon was engineered to generate a synthetic promoter plasmid that is selectively inducible by azide anions. The tunable expression of a model green fluorescence protein (GFP) was demonstrated using this novel azide based promoter system and compared to the standard *E. coli* lactose (*lac)* operon. Finally, the biosensing potential of this promoter system to detect azide anions, versus organic glycosyl azides, was showcased. In addition, a proof-of-concept glycoengineering application is showcased where we selectively monitored the *in vivo* glycosynthase enzyme activity for a mutant fucosidase that can synthesize fucosylated disaccharides using fucosyl azide as the donor sugar and pNP-xylose as the acceptor sugar.

Results and Discussion

The *E. coli* genome consists of several hundred operons (~700) performing specific functions essential for bacterial metabolism and survival in harsh environments.[28] In particular, the cyanate or *cyn* operon enables *E. coli* cells to survive in cyanate-rich environments.[29] This gene had evolved in archaea and cyanobacteria for energy production and nitrogen assimilation during the early history of single-celled life in extreme marine environments.[30] The *cyn* operon, analogous to the *lac* operon, is comprised of three structural genes; *cynT*, *cynS*, and *cynX* (**Figure 1A**) that encode carbonic anhydrase, cyanate hydratase, and cyanate transporter proteins, respectively. These genes code for enzymes that catalyze the bicarbonate-dependent decomposition of cyanate ions into carbon dioxide and ammonia. Along with *cynTSX* genes, *cynR* repressor gene encoding cynR protein is present for regulatory purposes. The operon is under tight negative regulation of the cynR repressor protein which upon binding to the operator region results in unfavorable DNA bending to prevent transcription of the *cynTSX* genes. The *cyn* operon is only activated when a cyanate molecule binds to the allosteric repressor protein also bound to the promoter/operator region. Exogenous cyanate can bind to the cynR inducer binding domain and cause conformational changes in the binding domain thereby reducing the DNA bend in the operator region to facilitate transcription.[29,31,32]

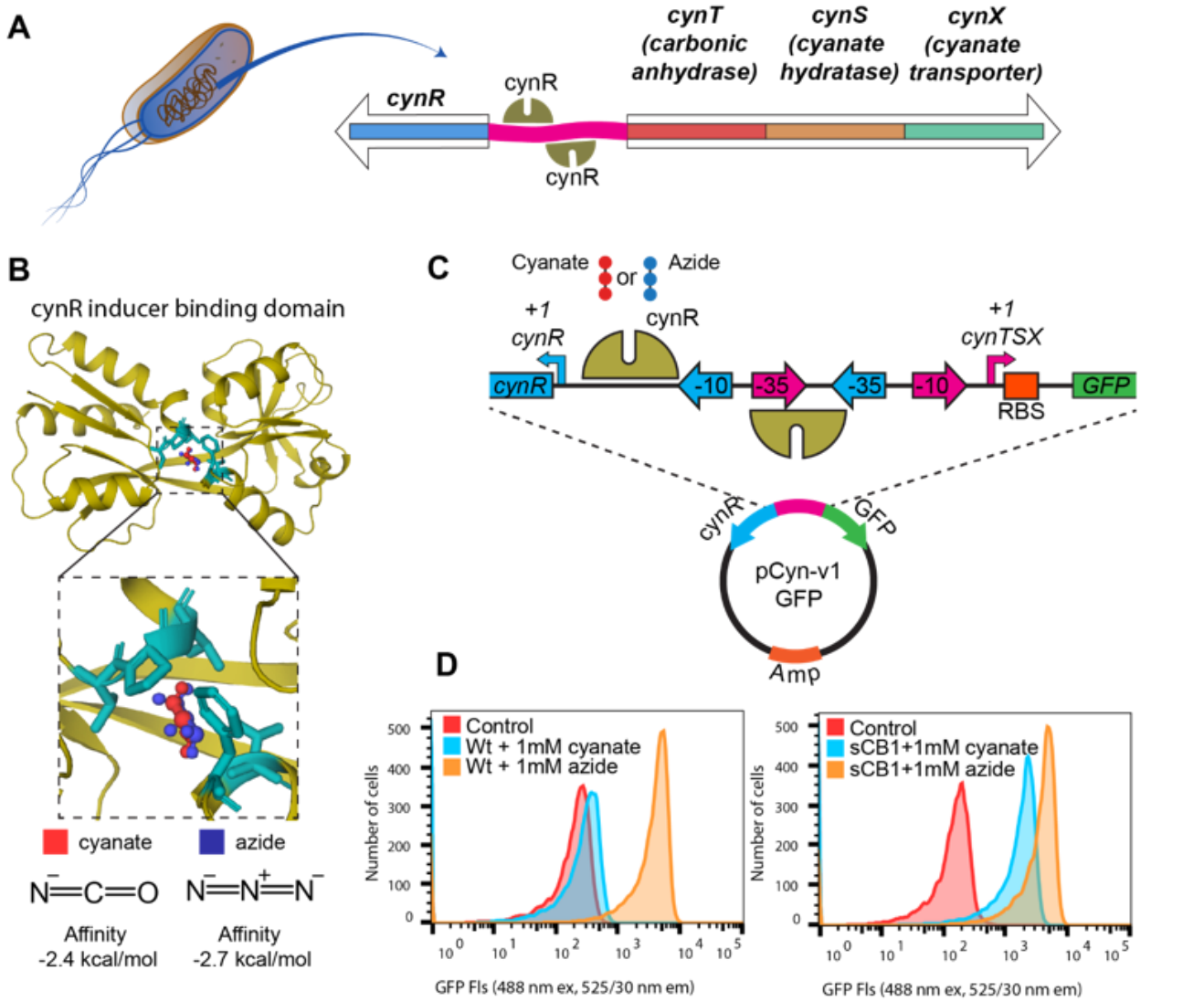


Figure 1. Promoter engineering to enable reporter *gfp* gene expression in *E. coli* using azide and cyanate as inducers. A. Native cynTSX operon facilitates *E. coli* cell growth in a cyanate-rich environment. Cyanate molecule binds to the cynR repressor protein to regulate downstream protein expression of three essential native genes (i.e., cynT, cynS, cynX). **B**. Molecular docked structures of cyanate versus azide ions in the cynR binding pocket is shown here. Docking simulations confirmed that azide also binds tightly in the cynR binding pocket with a similar binding affinity like cyanate, a close structural homolog of azide. **C**. Engineered plasmid map design of pCyn-v1-GFP containing the native cynR and *cyn* operator region cloned upstream of a GFP reporter gene. As proof of concept, the version 1 (v1) design was used first to monitor GFP expression using 1 mM azide or cyanate as inducers. **D**. Flow cytometry based individual cells analysis of GFP fluorescence signal confirms cyanate versus azide inducible heterologous *in vivo* protein expression using both native (BW25113-wt) and engineered *ΔcynS* knockout (BW25113-sCB1) *E. coli* strains.

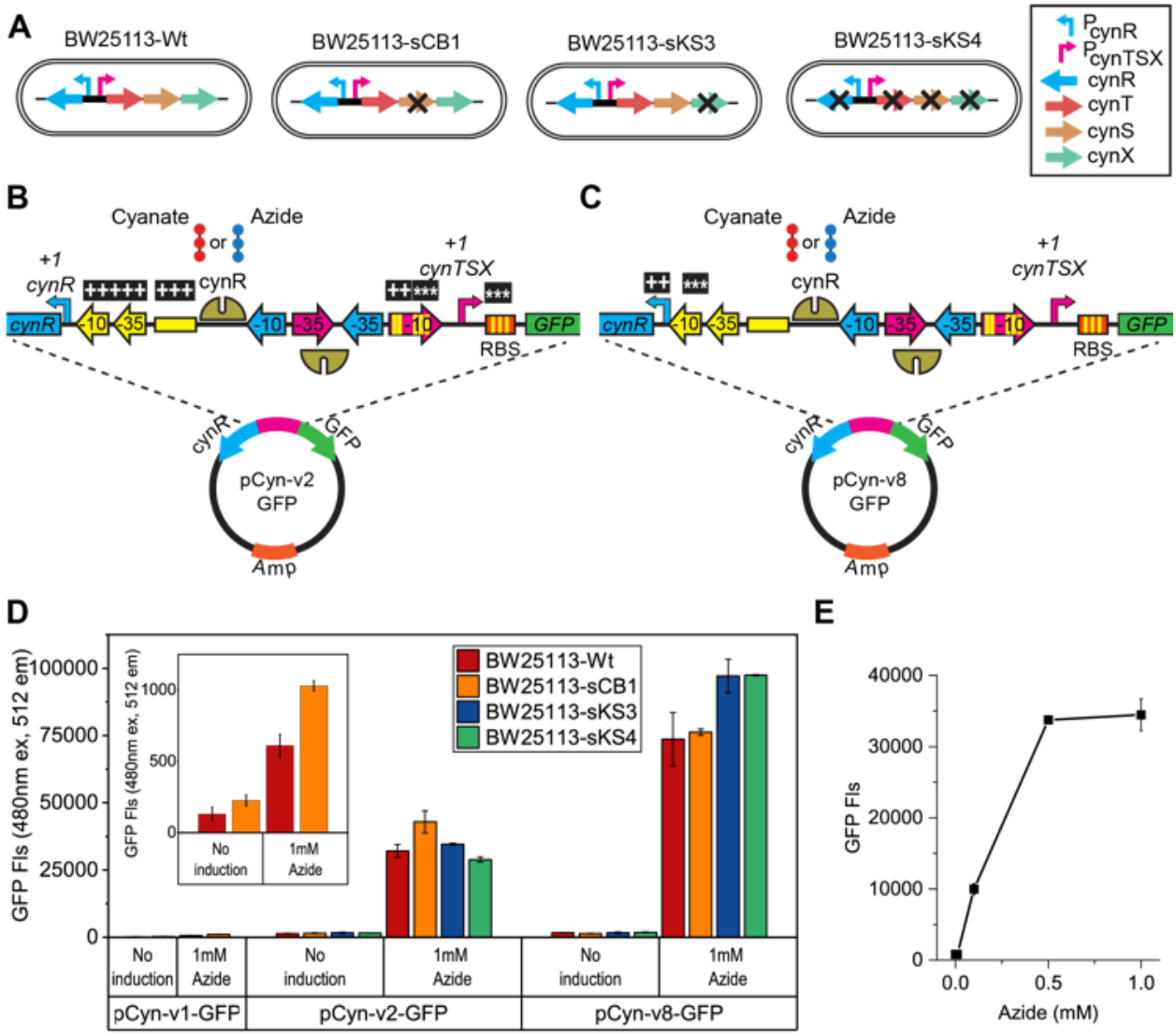


Figure 2. **Strain engineering and promoter design co-optimization to increase reporter GFP protein expression upon induction by azide**. **A**. *E. coli* BW25113 wild type and various knockout strains are represented here in the cartoon schematics. P_{cynR} and P_{cynTSX} are promoters for cynR and cynTSX genes, respectively. Individual genes are color coded and genes knocked-out are illustrated with "X" mark. **B-C**. Plasmid map design of pCyn-v2-GFP and pCyn-v8-GFP containing the engineered regulatory region between *cynR* and *gfp* genes are highlighted here. The mutational replacements in the promoter region are represented by asterisk (*) while additions are denoted by plus sign (+). **D**. Bar graphs representing GFP fluorescence intensities from cell lysates of engineered strains containing three distinct plasmid variants (v1, v2, v8) are shown. The magnified inset shows GFP fluorescence for native promoter construct (pCyn-v1-GFP). **E**. Fluorescence plot for cell lysate of BW25113-wt strain containing pCyn-v2-GFP and induced with increasing concentrations of sodium azide. Error bars indicate one standard deviation from reported mean values from three biological replicates.

Since cyanate is a linear molecular ion that is structurally homologous to azide, we hypothesized that azide could also bind to cynR and trigger the *cyn* operon. To confirm if azide can bind with similar binding affinity to cynR, Autodock vina[33] was used to perform docking simulations for both cyanate and azide in the binding pocket of cynR. The binding orientation of the docked ligands revealed a good overlap in the binding site with similarly predicted binding affinities (**Figure 1B**). This suggested that azide could indeed function as a gratuitous inducer for the native *cyn* operon.

Next, we subcloned the regulatory segment consisting of *cynR* gene and the *cyn* operator region into a plasmid vector with an ampicillin resistance marker and green fluorescent protein (*gfp*) as the reporter gene. The regulatory segment was placed upstream of the *gfp* reporter gene to facilitate GFP expression using either cyanate or azide ions as inducers (**Figure 1C**). The resultant plasmid (called pCyn-v1-GFP) was transformed into *E. coli* BW25113 wildtype (Wt) strain and induced with 1 mM of sodium cyanate or sodium azide. The supernatant of the lysed cells after induction was analyzed for fluorescence using a spectrophotometer (see supporting information or **SI Figure S1**). In addition, the fluorescence of individual intact cells was analyzed through a flow cytometer and **Figure 1D** presents the resultant histogram plots of the induced cells fluorescence after 19 hours. Compared to the control (uninduced cells), the azide induced cell lysate had a

nearly 20-fold increase in GFP fluorescence while cyanate did not result in any significant GFP expression. Specifically, there was a significant difference (two-tailed t-test $P<.001$ for 2 biological replicates per test group) in the fluorescence observed for pCyn-v1-GFP transformed cell lysate after 19 hours induction time in the presence of 1 mM sodium azide (3840+135; mean±s.d.) versus uninduced cells (169±38). The absence of GFP expression in cyanate induced cells was likely due to the breakdown of cyanate by the endogenous cyanate hydratase enzyme encoded by the native *E. coli* BW25113-Wt genome. Hence, we next tested the induction capacity of the synthetic promoter using a Δ*cynS* knockout strain (BW25113-sCB1) to clearly show GFP expression upon induction by cyanate (**SI Figure S1**). However, the amount of protein expressed based on total cell lysate GFP fluorescence was very low suggesting that the native *cyn* promoter strength was quite poor. This is not surprising since the native *cyn* promoter regulates associated CynTSX proteins expression required to overcome cyanate toxicity and that seldom requires high protein expression yields. The native operator region also contains suboptimal -10 sequence and Shine-Delgarno (SD) sequences (or ribosomal binding site; RBS) which could play a role in poor expression strength.

To improve the native promoter strength for enabling higher inducible protein expression levels, the native *cyn* promoter was next engineered to contain consensus -10 and SD sequences. The *cynR* gene, that is negatively regulated by *cyn* promoter, was placed under control of an independent constitutive promoter[34] (see SI Text for sequence information). Additionally, to making these modifications, the promoters were separated by inserting a random DNA spacer sequence of 100 bp (pCyn-v2-GFP) and 1000 bp (pCyn-v4-GFP) to avoid any interference and steric hinderances between the two promoters (**Figure 2B** and **SI Figure S2**). A control construct that contained no spacer sequence (pCyn-v3-GFP) was also generated. The modified constructs were individually transformed into *E. coli* cells and tested for GFP expression using sodium azide as inducer. All three engineered constructs showed a greater than 50-fold increase in the fluorescence signal in the cell lysate, indicating a significant increase in GFP expression as compared to the proof-of-concept pCyn-v1-GFP version 1 promoter (**Figure 2D** and **SI Figure S2B**). Specifically, there was a significant difference (two-tailed t-test $P=.002$ for 3 biological replicates per test group) in the fluorescence observed for pCyn-v2-GFP transformed cell lysate after 2 hours induction time in the presence of 1 mM sodium azide (32068±2170; mean±s.d.) versus pCyn-v1-GFP transformed cell lysate (610±80). The pCyn-v4-GFP construct with longest spacer region showed around 37% increased fluorescence when compared to the no spacer pCyn-v3-GFP control. Specifically, there was a significant difference (two-tailed t-test $P=.001$ for 3 biological replicates per test group) in the fluorescence observed for pCyn-v4-GFP transformed cell lysate after 2 hours induction time in the presence of 1 mM sodium azide (65468±932; mean±s.d.) versus pCyn-v3-GFP transformed cells (47765±28). However, significant leaky GFP expression was seen for the no induction controls of v4 construct (**SI Figure S2C**). The extra-long spacer sequences could have potentially allowed undesirable interactions in the plasmid causing a change in the DNA bending properties for CynR protein.

Decreasing the length between the promoters largely reduced the leaky expression under no induction while maintaining the largely improved GFP yield upon azide induction (**Figure S2**). For all subsequent work, pCyn-v2-GFP design was chosen to keep the promoters at optimal distance from each other. The nearly 626-fold increase in GFP fluorescence observed was now promising to be able to utilize the pCyn-v2-GFP engineered design for autonomously sensing azide ions. Specifically, there was a significant difference (two-tailed t-test $P<.001$ for 3 biological replicates per test group) in the fluorescence observed for pCyn-v2-GFP transformed cell lysate after 4 hours induction time in the presence of 1 mM sodium azide (105179±3802; mean±s.d.) versus uninduced cells (168±17). To determine the minimal azide amount required to induce GFP protein expression, we induced cells transformed with the pCyn-v2-GFP plasmid with varying concentrations of azide ions. The maximum amount of azide used for induction was limited to 5 mM since higher amounts had a negative influence on bacterial cell growth (**Figure S3**). GFP fluorescence of cell lysate showed a strong correlation in protein expression as a function of inducer dosage. GFP expression increased until 1 mM azide concentration induction after which there was reduction in the GFP signal for 5 mM inducer concentration likely due to cellular toxicity of azide (**Figure 2E**). The lowest tested azide concentration (10 μM) showed marginal GFP expression indicating that the detection limit for this synthetic promoter would have to be greater than 10 μM. The bacterial growth phase during the time of induction also played an important role and we observed that induction at early exponential phase yielded higher expressed protein amounts (**Figure S4**).

CynR protein acts as a repressor for the promoter by causing a bend at the -35 site hindering the binding of RNA polymerase. The cynR protein consists of two domains: an inducer binding domain and a DNA binding domain. The inducer binding domain binds to the inducer (cyanate or azide) which decreases the bend at the promoter site to facilitate transcription. For regulating protein expression levels, along with the inducer amounts the basal level of cynR protein present is also critical. We therefore further engineered the cynR constitutive promoter to adjust the background level of repressor protein expressed. Four different constructs (v5, v6, v7 and v8) were created by introducing mutations at the promoter -10 site, regions between -10 and -35 site, and between RBS and translation start site (**Figure 2C** and **SI Figure S5A**). The expression strength of these modified promoters was tested in the wild type (BW2113) and several additional knockout strains (sCB1, sKS3, sKS4) by monitoring GFP fluorescence in cell lysates. The knockout strains sKS3 and sKS4 were generated to remove the native *cynX* gene and *cyn* operon (**SI Table S1**), respectively, to minimize the export of azide ions using native transporter proteins and reduce interference from the endogenous *cyn* operon. The pCyn-v8-GFP construct, which had reduced efficiency at -10 site but optimal length between RBS and translation start site, gave about 120-160 fold higher reporter GFP fluorescence as compared to the native promoter, with the maximum fold increase observed in *ΔcynX* (sKS3), *ΔcynR*, and *ΔcynTSX* (sKS4) knockout strains (**Figure 2D**). Specifically, there was a significant difference (two-tailed t-test $P=.006$ for 3 biological replicates per test group) in the fluorescence observed for pCyn-v8-GFP transformed cell lysate after 2 hours induction time in the presence of 1 mM sodium azide (73686±9930; mean±s.d.) versus pCyn-v1-GFP transformed cell lysate (610±80). On the other hand, while other designs (v5, v6, v7) showed comparable fluorescence with respect to the v2

design, there was undesirable leaky background GFP expression seen even without inducer addition (**SI Figure S5**). The optimized plasmid designs (pCyn-v2-GFP and pCyn-v8-GFP) could now function as highly tunable synthetic biosensors for rapid *in vivo* detection of azide ions. The amount of GFP expressed within 2 hours of 1 mM azide induction in BW25113-Wt strain was estimated to be 0.4 mg and 0.9 mg from 1 ml culture for pCyn-v2-GFP and pCyn-v8-GFP, respectively, based on the protein concentration and GFP fluorescence calibration curve (**Figure S6**).

Next, the heterologous protein expression efficiency of the engineered promoter (pCyn-v2-GFP) was compared against the *E. coli lac* operon system.[35,36] The synthetic *cyn* promoter (P_{cyn}) in BW25113-Wt (**Figure 3A**) and *lac* promoter (P_{lac}) in BL21 cells (**Figure 3B**) were induced using sodium azide and isopropyl-β-D-1-thiogalactopyranoside (or IPTG), respectively, at various concentrations ranging between 0.01 mM to 1 mM. The amount of protein expressed in the P_{lac} system was higher during the early time points after induction and at the lowest inducer concentrations. Even at lower IPTG concentrations (0.01 mM and 0.1 mM), GFP expression increased with induction time and reached a maximum after 24 hours of induction. Whereas we see a relatively lower GFP expression at similar low azide concentrations (**Figure 3A** and **3B**). However, we see a highly tunable reporter gene expression during azide induction as the amount of GFP expressed was closely proportional to inducer concentration, unlike IPTG. Also, the maximum GFP expression achieved after 24 hours was at least two-fold higher for the *cyn* promoter than *lac* promoter at inducer concentrations higher than 0.5 mM (**Figure 3A** and **3B**). Specifically, there was a significant difference (two-tailed t-test P=.005 for 3 biological replicates per test group) in the fluorescence observed for pCyn-v2-GFP transformed cell lysate in the presence of 1 mM sodium azide (114238±3721; mean±s.d.) versus pEC-GFP transformed cell lysate in the presence of 1 mM IPTG (67336±8892) after 23 hours induction time. These results showcase the utility of the novel engineered plasmid harboring the P_{cyn} promoter as an alternative system for heterologous protein expression and other biotechnology applications.

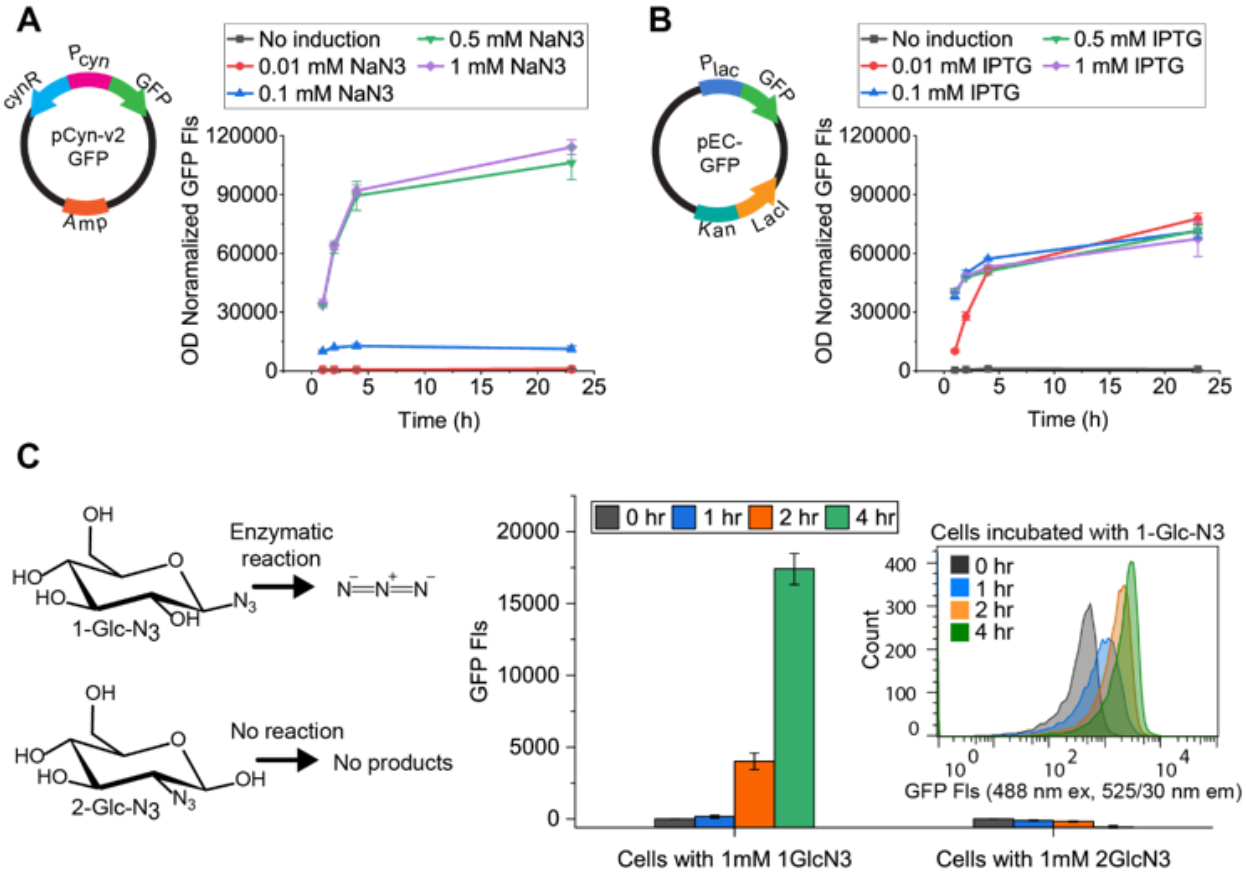


Figure 3. **Utility of novel synthetic azide promoter system towards heterologous protein expression and biorthogonal chemical-biology applications**. **A**. *E. coli* BW25113-Wt containing pCyn-v2-GFP plasmid was induced with varying azide concentrations for 24 hours and GFP fluorescence of cell lysate was measured at various time points. **B**. *E. coli* BL21 containing pEC-GFP plasmid was induced with different IPTG concentrations for 24 hours and GFP fluorescence of the cell lysate was measured at different time points. **C**. *E. coli* BW25113-Wt cells with pCyn-v2-GFP plasmid was incubated with 1-azido-β-D-glucopyranosyl azide (1-Glc-N_3) and 2-Deoxy-2-azido-β-D-glucopyranosyl azide (2-Glc-N_3) and GFP fluorescence of respective cell lysates are shown as bar graphs. The inset depicts the flow cytometry analysis data for cells incubated with 1-Glc-N_3 for 4 hours. Error bar indicates one standard deviation from reported mean values from three biological replicates.

In addition to application as a heterologous protein production system, the developed azide promoter can be useful for chemical biologists especially in glycobiology and glycosynthase engineering related applications. The proposed

system is a powerful synthetic biology toolkit that will allow selective identification of inorganic azides amongst other chemically linked organic azides. Enzymatic reactions which result in the release of azide ion (as leaving group) can be potentially identified using the proposed *cyn* promoter system. Carbohydrate-Active enZymes (CAZymes) such as glycosyl hydrolases and transglycosidases have been shown to release azide from azido-based sugars during glycosidic bond hydrolysis[37] and synthesis,[26,37–40] respectively. Multiple glycosyl azides such as galactosyl-, glucosyl-, and mannosyl-azides have been reported to be hydrolyzed by galactosidases, glucosidases, and mannosidases, respectively.[37] Similarly, engineered glycosidases (i.e., transglycosidases and glycosynthases) have been shown to use azido-hexoses and other N-acetyl derivatives as donor sugars for oligosaccharides synthesis.[38,40] β-glucosidases from *Aspergillus* sp. and *Agrobacterium* sp. belonging to glycosyl hydrolase (GH) families GH3 and GH1 were reported to actively release azide ion from glucosyl azides.[39] Similarly, homologous β-glucosidases belonging to the GH1 and GH3 families present in native *E. coli* can potentially also show similar substrate specificity to release azide ions when incubated with glucosyl azides.

Here, we incubated *E. coli* BW25113-Wt cells containing the engineered pCyn-v2-GFP plasmid with 1-azido-β-D-glucopyranosyl azide (1-Glc-N_3) for 4 hours. As illustrated in **Figure 3C**, a small change in GFP fluorescence was seen in the first hour after which a rapid increase in fluorescence was observed within 1-4 hours owing to the release of azide when using 1-Glc-N_3. Specifically, there was a significant difference (two-tailed t-test P=.035 for 3 biological replicates per test group) in the fluorescence observed for pCyn-v2-GFP transformed cell lysate in the presence of 1 mM 1-Glc-N_3 after 1 hour induction time (170±93; mean±s.d.) versus 0-hour induction time (0±15). To verify if the detected GFP fluorescence is due to release of azide ions and not due to presence of azido-glucose, we also incubated cells with 2-Deoxy-2-azido-β-D-glucopyranosyl azide (2-Glc-N_3) analogous to the 2-deoxy-2-fluoro glycoside based β-glucosidase inhibitors.[41] No change in fluorescence was detected in the control samples since 2-Glc-N_3 is not hydrolyzed by β-glucosidases and hence cannot release free azide ions for induction of the engineered promoter. The preferential sensitivity of the designed promoter towards azide ions can therefore be used for *in vivo* glycoengineering to identify efficient CAZymes such as glycosyl hydrolases (GH) for glycosidic bonds synthesis and/or hydrolysis.[27]

Glycosynthases (GS) are engineered/mutant glycosyl hydrolases or glycosidases that can synthesize oligosaccharides by forming glycosidic bonds between suitable donor sugars and acceptor glycone or aglycone groups. The donor sugars are often activated sugars containing fluoride or azide moieties as leaving groups on the anomeric carbon that can be readily released in the enzyme active site in the presence of incoming acceptor substrates to form a glycosidic bond. The fluoride or azide ions are hence released as by-products of the glycosynthase reaction. Higher concentrations of by-product release are expected to be strongly correlated with higher glycosynthase activity. Our engineered azide specific promoter system can therefore be used for screening enzyme activity and identifying improved GS enzyme mutants. As proof of concept, a previously engineered fucosynthase mutant (TmAfc-D224G), isolated from GH29 family fucosidase enzyme (TmAfc-Wt),[38] was selected to determine the applicability of the azide biosensor for glycosynthase enzyme engineering and screening activity in live cells using flow cytometry. *E. coli* cells containing two plasmids (i.e., pEC-TmAfc and pCyn-v2-GFP) were induced for TmAfc enzyme expression and incubated with fucosyl-azide (donor) and pNP-xylose (acceptor) for conducting the fucosynthase reaction along with several controls. The resultant *E. coli* cells were analyzed for GFP expression using flow cytometry (**Figure 4**). The cells containing wild-type enzyme (TmAfc-Wt) incubated with donor sugar alone was used as a baseline control. This baseline control represents the response expected for inactive or non-glycosynthase based samples. Whereas the glycosynthase active variant (TmAfc-D224G) was also incubated with sodium azide alone as another control to depict the theoretical maximum glycosynthase activity that would be possible if 100% of the added fucosyl azide was converted by the enzymes into products. TmAfc-D224G showed very minor hydrolysis when cells expressing this D224G enzyme mutant were incubated with only donor sugars (red). However, the active GS mutant (TmAfc-D224G) showed a significant increase in GFP fluorescence when both the donor and acceptor sugars were added to the cell milieu (blue). Specifically, there was a significant difference (two-tailed t-test P<.001) in the fluorescence observed for transformed cells incubated with donor+acceptor substrates (230±148; mean±s.d.; n=71191; raw data shown in blue histogram in Fig. 4) versus cells incubated with only donor substrate (68±37; n=84611; raw data shown in red histogram in Fig. 4). Our recent work has shown that although the D224G single-point mutation results in an active fucosynthase unlike the wild-type enzyme, the total reaction yield is still well under 10% based on the donor sugar as limiting reagent.[27] This explains why TmAfc-D224G expressing cells gave an increase in GFP fluorescence that did not reach the maximum expected theoretical concentration (based on observed GFP fluorescence shown in orange histogram in Fig. 4), if 100% of the fucosyl azide was consumed in the GS reaction. Regardless, these proof-of-concept results clearly show that dual plasmids carrying *E. coli* cells (i.e., pEC-TmAfc and pCyn-v2-GFP) can be theoretically used for screening and isolating active glycosynthase mutants that release azide ions as the GS reaction by-product.

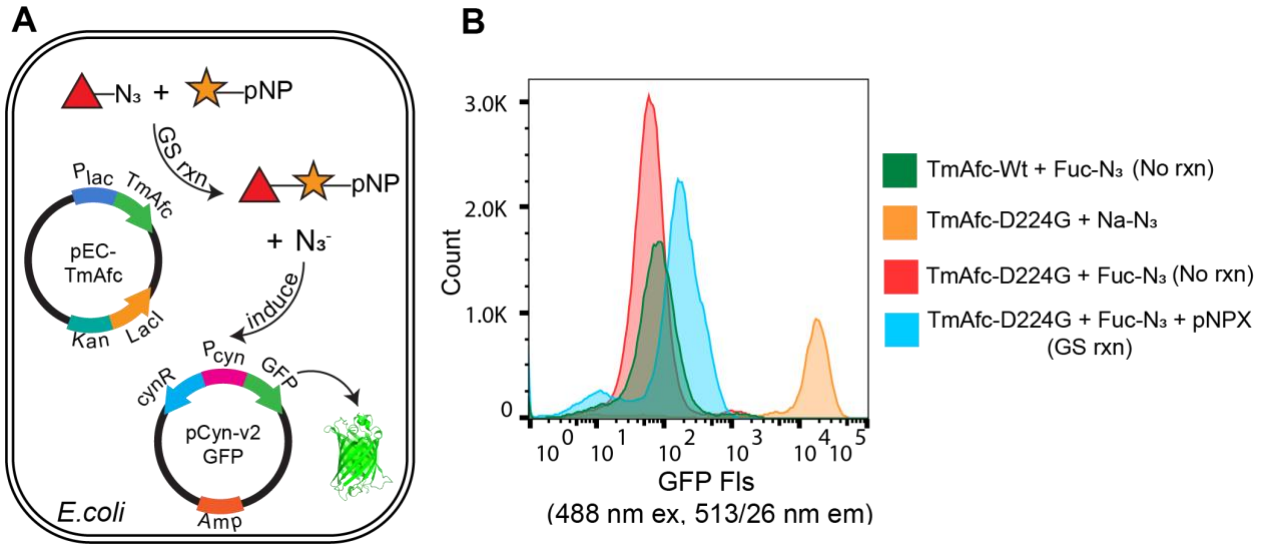


Figure 4. **Glycosynthase enzyme engineering and screening using flow cytometry is enabled using azide promoter system**. **A.** *E. coli* containing two plasmids (pEC-TmAfc and pCyn-v2-GFP) was incubated with glycosynthase reaction substrates; fucosyl-azide (red triangle) and pNP-xylose (orange star). TmAfc-D224G mutant alone catalyzes the glycosynthase (GS) reaction releasing azide ion as by-product which then induces the azide promoter on the secondary plasmid (pCyn-v2-GFP) resulting in GFP expression. **B**. Flow cytometry analysis of *E.coli* cells containing dual plasmids were incubated with glycosynthase substrates and various controls to demonstrate the utility of the azide promoter system. Only cells expressing the TmAfc-D224G mutant and in the presence of both GS reaction donor/acceptor sugar substrates gave a clear shift in cell fluorescence (blue) due to the release of azide ion during the GS reaction.

In summary, we have demonstrated the design and engineering of a novel azide (and cyanate) inducible promoter system for *E. coli*. This azide promoter system allows tunable expression by varying the inducer (i.e., azide ion) concentrations that also outperformed the conventional lactose/IPTG based system for heterologous reporter GFP expression. Additionally, the developed toolkit functions as a biosensor for detecting the presence of azide ions using living cells. This allowed us to employ this toolkit for engineering CAZymes such as glycosyl hydrolases (and/or glycosynthases) which use azido-sugars as substrates and autonomously monitor released azide ions upon substrate hydrolysis or during transglycosylation reactions. Furthermore, this biosensing platform can be evolved in the future for use with other prokaryotic or eukaryotic cells. Adaptation of an azide specific promoter system for diverse cell types would be also beneficial for various chemical-biology based research applications like azido-based drug development and molecular probes for real-time multicolor fluorescence imaging.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental Methods, DNA Sequences, Supporting Information (SI) Figures S1-S6 and Tables S1-S2 compiled in a single PDF.

AUTHOR INFORMATION

Corresponding Author

* Shishir P. S. Chundawat (98 Brett Road, Room C150A, Rutgers The State University of New Jersey, Piscataway, NJ 08854)

Author Contributions

CKB and SPSC conceived the project. CKB, KSS, AA, NS, MT contributed towards experimental work. CKB and SPSC performed data analysis and contributed towards manuscript writing. CKB generated all figures and TOC graphics using Adobe Illustrator. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

SPSC acknowledges support from the US National Science Foundation (CBET Award No. 1704679 and Chemistry Award No. 1904890) and Rutgers School of Engineering.

CONFLICT OF INTEREST

CKB and SPSC have a US provisional patent application filed by Rutgers University on June 30th, 2020. (Rutgers Provisional Application No. RU 2019-159).

ABBREVIATIONS

CAZymes, carbohydrate-active enzymes; GFP, green fluorescent protein; GH, glycosyl hydrolase; IPTG, isopropyl β-D-1-thiogalactopyranoside; pNP, para- nitrophenyl.

REFERENCES

(1) Soli, E. D.; Manoso, A. S.; Patterson, M. C.; DeShong, P.; Favor, D. A.; Hirschmann, R.; Smith, A. B. Azide and Cyanide Displacements via Hypervalent Silicate Intermediates. *J. Org. Chem.* **1999**, *64* (9), 3171–3177. https://doi.org/10.1021/jo982302d.

(2) Brse, S.; Banert, K. *Organic Azides: Syntheses and Applications*; Brse, S., Banert, K., Eds.; John Wiley & Sons, Ltd: Chichester, UK, 2009. https://doi.org/10.1002/9780470682517.

(3) Madlung, A. The Chemistry behind the Air Bag: High Tech in First-Year Chemistry. *J. Chem. Educ.* **1996**, *73* (4), 347. https://doi.org/10.1021/ed073p347.

(4) Dyukarev, E. A.; Knyazeva, A. G. Model of Detonation of Lead Azide (Pb(N3)2) with Regard to Fracture. *Int. J. Fract.* **1999**, *100*

(2), 197–205. https://doi.org/10.1023/A:1018788304044.

(5) Wereide, K. Sensitivity to Azidamphenicol. *Contact Dermatitis* **1975**, *1* (4), 271–272. https://doi.org/10.1111/j.1600-0536.1975.tb05433.x.

(6) Bengtsson, E.; Strandell, T.; Svanbom, M.; Tunevall, G. Azidocillin Treatment of Enterococcal Septicemia. *Scand. J. Infect. Dis.* **1972**, *4* (2), 143–148. https://doi.org/10.3109/inf.1972.4.issue-2.15.

(7) Blum, M. R.; Liao, S. H.; Good, S. S.; de Miranda, P. Pharmacokinetics and Bioavailability of Zidovudine in Humans. *Am. J. Med.* **1988**, *85* (2A), 189–194.

(8) Glassford, I.; Teijaro, C. N.; Daher, S. S.; Weil, A.; Small, M. C.; Redhu, S. K.; Colussi, D. J.; Jacobson, M. A.; Childers, W. E.; Buttaro, B.; Nicholson, A. W.; MacKerell, A. D.; Cooperman, B. S.; Andrade, R. B. Ribosome-Templated Azide-Alkyne Cycloadditions: Synthesis of Potent Macrolide Antibiotics by in Situ Click Chemistry. *J. Am. Chem. Soc.* **2016**, *138* (9), 3136–3144. https://doi.org/10.1021/jacs.5b13008.

(9) Devaraj, N. K. The Future of Bioorthogonal Chemistry. *ACS Cent. Sci.* **2018**, *4* (8), 952–959. https://doi.org/10.1021/acscentsci.8b00251.

(10) Devaraj, N. K.; Thurber, G. M.; Keliher, E. J.; Marinelli, B.; Weissleder, R. Reactive Polymer Enables Efficient in Vivo Bioorthogonal Chemistry. *Proc. Natl. Acad. Sci.* **2012**, *109* (13), 4762–4767. https://doi.org/10.1073/PNAS.1113466109.

(11) Au, K. M.; Wang, A. Z.; Park, S. I. Pretargeted Delivery of PI3K/MTOR Small-Molecule Inhibitor–Loaded Nanoparticles for Treatment of Non-Hodgkin's Lymphoma. *Sci. Adv.* **2020**, *6* (14), eaaz9798. https://doi.org/10.1126/sciadv.aaz9798.

(12) Laughlin, S. T.; Bertozzi, C. R. Imaging the Glycome. *Proc. Natl. Acad. Sci.* **2009**, *106* (1), 12 LP – 17. https://doi.org/10.1073/pnas.0811481106.

(13) Medina-Cleghorn, D.; Nomura, D. K. Exploring Metabolic Pathways and Regulation through Functional Chemoproteomic and Metabolomic Platforms. *Chem. Biol.* **2014**, *21* (9), 1171–1184. https://doi.org/10.1016/J.CHEMBIOL.2014.07.007.

(14) Dunge, A.; Chakraborti, A. K.; Singh, S. Mechanistic Explanation to the Variable Degradation Behaviour of Stavudine and Zidovudine under Hydrolytic, Oxidative and Photolytic Conditions. *J. Pharm. Biomed. Anal.* **2004**, *35* (4), 965–970. https://doi.org/10.1016/j.jpba.2004.03.007.

(15) Adhikari, S.; Guria, S.; Ghosh, A.; Pal, A.; Das, D. A Curcumin Derived Probe for Colorimetric Detection of Azide Ions in Water. *New J. Chem.* **2017**, *41* (24), 15368–15372. https://doi.org/10.1039/C7NJ03266H.

(16) K, T.; M, K.; Y, S. Rapid Determination of Cyanide and Azide in Beverages by Microdiffusion Spectrophotometric Method. *J. Anal. Toxicol.* **2001**, *25* (4). https://doi.org/10.1093/JAT/25.4.228.

(17) Wang, L.; Dai, C.; Chen, W.; Liu Wang, S.; Wang, B. Facile Derivatization of Azide Ions Using Click Chemistry for Their Sensitive Detection with LC-MSwz. *Chem. Commun. Chem. Commun* **2011**, *47* (47), 10377–10379. https://doi.org/10.1039/c1cc11199j.

(18) Dillen, L.; Sips, L.; De Vries, R.; Borgmans, C.; Vliegen, M.; Leclercq, L.; Cuyckens, F. Quantitative LC-MS/MS Analysis of Azide and Azidoalanine in in Vitro Samples Following Derivatisation with Dansyl Chloride. *Anal. Methods* **2013**, *5* (12), 3136–3141. https://doi.org/10.1039/c3ay00011g.

(19) Sahana, A.; Banerjee, A.; Guha, S.; Lohar, S.; Chattopadhyay, A.; Mukhopadhyay, S. K.; Das, D. Highly Selective Organic Fluorescent Probe for Azide Ion: Formation of a "Molecular Ring." *Analyst* **2012**, *137* (7), 1544. https://doi.org/10.1039/c2an16180j.

(20) Dhara, K.; Saha, U. C.; Dan, A.; Sarkar, S.; Manassero, M.; Chattopadhyay, P. A New Water-Soluble Copper(II) Complex as a Selective Fluorescent Sensor for Azide Ion. *Chem. Commun. (Camb).* **2010**, *46* (10), 1754–1756. https://doi.org/10.1039/b919937c.

(21) Wang, K.; Friscourt, F.; Dai, C.; Wang, L.; Zheng, Y.; Boons, G.-J.; Wang, S.; Wang, B. A Metal-Free Turn-on Fluorescent Probe for the Fast and Sensitive Detection of Inorganic Azides. *Bioorg. Med. Chem. Lett.* **2016**, *26* (7), 1651–1654. https://doi.org/10.1016/j.bmcl.2016.02.069.

(22) Puthiyedath, T.; Bahulayan, D. A Click-Generated Triazole Tethered Oxazolone-Pyrimidinone Dyad: A Highly Selective Colorimetric and Ratiometric FRET Based Fluorescent Probe for Sensing Azide Ions. *Sensors Actuators B Chem.* **2017**, *239*, 1076–1086. https://doi.org/10.1016/J.SNB.2016.08.044.

(23) Lim, J. Y. C.; Beer, P. D. A Halogen Bonding 1,3-Disubstituted Ferrocene Receptor for Recognition and Redox Sensing of Azide. *Eur. J. Inorg. Chem.* **2017**, *2017* (2), 220–224. https://doi.org/10.1002/ejic.201600805.

(24) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. Copper-Free Click Chemistry for Dynamic in Vivo Imaging. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (43), 16793–16797. https://doi.org/10.1073/pnas.0707090104.

(25) Bojarová, P.; Kren, V. Azido Leaving Group in Enzymatic Synthesis-Small and Efficient; 2010; pp 168–175. https://doi.org/10.1039/9781849730891-00168.

(26) Fialová, P.; Carmona, A. T.; Robina, I.; Ettrich, R.; Sedmera, P.; Přikrylová, V.; Petrásková-Hušáková, L.; Křen, V. Glycosyl Azide—a Novel Substrate for Enzymatic Transglycosylations. *Tetrahedron Lett.* **2005**, *46* (50), 8715–8718. https://doi.org/10.1016/j.tetlet.2005.10.040.

(27) Agrawal, A.; Bandi, C. K.; Burgin, T.; Woo, Y.; Mayes, H. B.; Chundawat, S. P. S. Click-Chemistry Enabled Directed Evolution of Glycosynthases for Bespoke Glycans Synthesis. *bioRxiv* **2020**, 2020.03.23.001982. https://doi.org/10.1101/2020.03.23.001982.

(28) Salgado, H.; Moreno-Hagelsieb, G.; Smith, T. F.; Collado-Vides, J. Operons in Escherichia Coli: Genomic Analyses and Predictions. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (12), 6652–6657. https://doi.org/10.1073/pnas.110147297.

(29) Sung, Y.-C.; Fuchs, J. A. Characterization of the Cyn Operon in Escherichia Coli K12* Young-Chul Sung And. *J. Biol. Chem.* **1988**, *263* (29), 14769–14775.

(30) Palatinszky, M.; Herbold, C.; Jehmlich, N.; Pogoda, M.; Han, P.; von Bergen, M.; Lagkouvardos, I.; Karst, S. M.; Galushko, A.; Koch, H.; Berry, D.; Daims, H.; Wagner, M. Cyanate as an Energy Source for Nitrifiers. *Nature* **2015**, *524* (7563), 105–108. https://doi.org/10.1038/nature14856.

(31) Sung, Y.-C.; Fuchs, J. A. The Escherichia Coli K-12 Cyn Operon Is Positively Regulated by a Member of the LysR Family. **1992**, *174* (11).

(32) Lamblin, A.-F. J.; Fuchs, J. A. Functional Analysis of the Escherichia Coli K-12 Cyn Operon Transcriptional Regulation. *J. BACrERIOLOGY* **1994**, *176* (21), 6613–6622.

(33) Trott,O., Olson, A. J. Autodock Vina: Improving the Speed and Accuracy of Docking. *J. Comput. Chem.* **2019**, *31* (2), 455–461. https://doi.org/10.1002/jcc.21334.AutoDock.

(34) Stringer, A. M.; Singh, N.; Yermakova, A.; Petrone, B. L.; Amarasinghe, J. J.; Reyes-Diaz, L.; Mantis, N. J.; Wade, J. T. FRUIT, a Scar-Free System for Targeted Chromosomal Mutagenesis, Epitope Tagging, and Promoter Replacement in Escherichia Coli and Salmonella Enterica. *PLoS One* **2012**, *7* (9). https://doi.org/10.1371/journal.pone.0044841.

(35) Studier, F. W.; Moffatt, B. A. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. *J. Mol. Biol.* **1986**, *189* (1), 113–130. https://doi.org/10.1016/0022-2836(86)90385-2.

(36) Eames, M.; Kortemme, T. Cost-Benefit Tradeoffs in Engineered Lac Operons. *Science (80-.).* **2012**, *336* (6083), 911–915. https://doi.org/10.1126/science.1219083.

(37) Bojarová, P.; Petrásková, L.; Ferrandi, E. E.; Monti, D.; Pelantová, H.; Kuzma, M.; Simerská, P.; Křen, V. Glycosyl Azides - An Alternative Way to Disaccharides. *Adv. Synth. Catal.* **2007**, *349* (8–9), 1514–1520. https://doi.org/10.1002/adsc.200700028.

(38) Cobucci-Ponzano, B.; Conte, F.; Bedini, E.; Corsaro, M. M.; Parrilli, M.; Sulzenbacher, G.; Lipski, A.; Dal Piaz, F.; Lepore, L.; Rossi, M.; Moracci, M. Beta-Glycosyl Azides as Substrates for Alpha-Glycosynthases: Preparation of Efficient Alpha-L-Fucosynthases. *Chem. Biol.* **2009**, *16* (10), 1097–1108. https://doi.org/10.1016/j.chembiol.2009.09.013.

(39) Müllegger, J.; Jahn, M.; Chen, H.-M.; Warren, R. A. J.; Withers, S. G. Engineering of a Thioglycoligase: Randomized Mutagenesis of the Acid-Base Residue Leads to the Identification of Improved Catalysts. *Protein Eng. Des. Sel.* **2005**, *18* (1), 33–40. https://doi.org/10.1093/protein/gzi003.

(40) Cobucci-Ponzano, B.; Zorzetti, C.; Strazzulli, A.; Carillo, S.; Bedini, E.; Corsaro, M. M.; Comfort, D. A.; Kelly, R. M.; Rossi, M.; Moracci, M. A Novel -d-Galactosynthase from Thermotoga Maritima Converts -d-Galactopyranosyl Azide to -Galacto-Oligosaccharides. *Glycobiology* **2011**, *21* (4), 448–456. https://doi.org/10.1093/glycob/cwq177.

(41) Rempel, B. P.; Withers, S. G. Covalent Inhibitors of Glycosidases and Their Applications in Biochemistry and Biology. *Glycobiology* **2008**, *18* (8), 570–586. https://doi.org/10.1093/glycob/cwn041.

Engineered promoters and microbial strains were developed to enable selective azide ion biosensing via inducible expression of reporter green fluorescent protein (GFP). Implementation of this synthetic biology toolkit for diverse biotechnology applications such as *in vivo* azide ion detection, heterologous protein production, and glycosynthase enzyme activity screening are highlighted here.

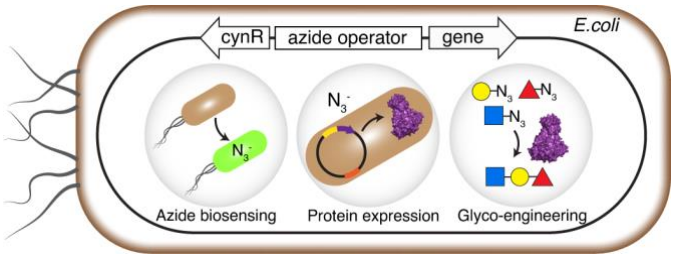


Table of Contents Caption and Artwork

Supporting Information (SI) Document

Engineered Regulon to Enable Autonomous Azide Ion Biosensing, Recombinant Protein Production, and *In vivo* Glycoengineering

Chandra Kanth Bandi,[a] Kyle S. Skalenko,[b] Ayushi Agrawal,[a] Neelan Sivaneri,[a] Margaux Thiry,[a] and Shishir P. S. Chundawat*[a]

[a] Department of Chemical & Biochemical Engineering, Rutgers, The State University of New Jersey, 98 Brett Road, Piscataway, New Jersey 08854, USA

[b] Department of Genetics and Waksman Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Rd, Piscataway, New Jersey 08854, USA

*Corresponding Author: shishir.chundawat@rutgers.edu

SI Table of Contents

SI Methods and Experimental Procedures

Bacterial strain engineering:

All the chemicals, reagents, and solvents were purchased from Fisher Scientific and Sigma-Aldrich and used without purification. *E. coli* strain used for cloning was E. cloni® 10G (Lucigen, WI, USA) and for protein expression using lac promoter was BL21-CodonPlus-RIPL [λDE3] (Stratagene, Santa Clara, CA, USA). Phusion 2X high fidelity PCR master mix (0.04 U/μL Phusion DNA polymerase, 400 μM dNTPs, Phusion 2XHF buffer, 3 mM MgCl2) was purchased from Thermo Fisher Scientific (USA) and restriction enzymes were procured from New England BioLabs Inc. (USA). The primers used for site directed mutagenesis (SDM), sequence and ligation independent cloning (SLIC), and sequencing reactions were obtained from Integrated DNA Technologies, Inc (USA). Successfully cloned plasmids were isolated from E. cloni® 10g cells using IBI Scientific (USA) plasmid extraction kit and sequences were confirmed through Sanger sequencing performed by Genscript Inc. (NJ, USA). The carbohydrate substrates used in reported assays were purchased from Synthose Inc, Canada.

The *E.coli* strains used for protein expression using *cyn* promoter were constructed using the protocol and reagents outlined in Datsenko 2000 [1]. Briefly, strains sCB1 (BW25113 *cynS*::FRT) and sKS3 (BW25113 *cynX*::FRT) were constructed by first streaking out the strains JW0331 (BW25113 *cynS*::Kan) and JW0332 (BW25113 *cynX*::Kan)[2] respectively from the Keio collection onto LB-agar kanamycin (50 μg/ml) plate. The kanamycin marker was then removed by transforming the individual strains with pCP20, following the protocol outlined in Datsenko 2000, and curing the strain of the plasmid. The final strain was diagnosed by PCR and for loss of kanamycin resistance.

Strain sKS4, BW25113 *cynR,cynTSX*::FRT, was constructed by first transforming the Keio parent strain, BW25113, with pKD46 and plated onto carbenicillin (100 μg/mL) plates. 5 mL overnights of BW25113 + pKD46 in LB + carbenicillin (100 μg/mL) were grown at 30°C and then back-diluted 1:100 into 50 mL LB + carbenicillin (100 μg/mL) + 0.2% L-arabinose. The 50 mL culture was grown to an OD600 of 0.8, at 30°C, and cells were washed 4 times with 50 mL of ice-cold water. The final cell pellet was resuspended 1:250th the starting culture volume with fresh water and sat on ice until ready to electroporate. The linear DNA fragment, that was used to knockout the *cynR* gene and *cynTSX* operon, was synthesized by amplifying the kanamycin gene from pKD4 using primers k1 and k2. The linear DNA fragment was checked by gel electrophoresis for purity and was then cleaned-up and concentrated using Qiagen's PCR clean-up kit, the product was eluted in water. The intermediate strain sKS1 (BW25113 *cynR*, *cynTSX*::Kan) was made by mixing 100 ng of linear DNA with 50 μL electrocompetent BW25113 + pKD46 cells, shocking immediately with 1.8 kV (with a pulse constant of 5.2ms), and recovered in 900 μL of SOC at 37°C for 3 hours. After recovery, cells were spun down at 10,000xg for 2 minutes at room temperature, resuspended to 100 μL, plating solution onto kanamycin (50 μg/mL) plates, and grown at 37°C overnight. sKS1 was transformed with pCP20 to remove the kanamycin selection marker, following the protocol in Datsenko 2000, making the final strain sKS4 (BW25113 *cynR,cynSTX*::FRT). sKS4 was diagnosed by loss of kanamycin resistance and PCR.

Design and construction of pCyn vectors:

The cloning of the plasmid constructs used in this study were performed using Sequence and Ligation-Independent Cloning (SLIC) protocol as outlined in Stevenson et.al [3]. Briefly, to create the pCyn-v1-GFP, the gene fragment consisting of native cynR and cyn promoter/operator region (gblock1) was custom synthesized from Genscript Inc, USA. The GFP gene fragment was taken from pEC-GFP plasmid available at Chundawat lab. Both the gene fragments were cloned into the parent plasmid, ptrc99a while getting rid of the intrinsic lac promoter using the primers p1-p4 and following the SLIC protocol (**SI Table S2A**). To generate pCyn-v2-GFP, an optimized promoter region was designed in-silico and the corresponding DNA fragment (gblock2) was custom synthesized and pCynv-1-GFP was used as starting DNA with primers used being p5-p8. The spacer region of 100 bp from the pCyn-v2-GFP was removed using primers p9-p10 to get pCyn-v3-GFP while an additional random sequence of 900 bp (gblock3) was added using p11-p14 to generate pCyn-v4-GFP. Further, pCyn-v5-GFP was constructed from pCyn-v2-GFP by site directed mutagenesis of the cynR constitutive promoter using primers p15-p16. In the end, the pCyn-v5-GFP was the parent DNA used to create pCyn-v6-GFP, pCyn-v7-GFP and pCyn-v8-GFP using the primers p17-p18, p19-p20 and p21-p22, respectively. All the constructs were diagnosed using Sanger sequencing and the sequence verified plasmids were preserved at -80°C and their corresponding transformed E. cloni cells were stored in 15% glycerol stocks at -80°C.

Induction of pCyn-v1/v2-GFP expression:

The pCyn-GFP plasmid constructs were transformed into BW25113 strains (wt, sCB1) and individual colonies were obtained on LB agar plates supplemented with 100 μg/ml Carbenicillin antibiotic. The transformants were inoculated into 10 ml of LB media with carbenicillin and grown at 37°C for 16 hrs. The 10 ml of overnight grown culture was transferred to 200 ml of fresh LB media with carbenicillin and incubated at 37°C until mid-exponential phase (OD 0.4-0.6) was reached. At this point, the 200 ml culture was split into six 25 ml falcon tubes. Two of the tubes were induced with 1 mM sodium azide and two tubes with 1 mM cyanate while the remaining two tubes were used as no induction control. The cultures were placed in the 37°C shaking incubator and 2 ml of sample for bulk fluorescence measurement and 200 ul of sample for flow cytometer runs were collected at every time point.

Induction of pCyn-v2/v3/v4-GFP expression:

The pCyn-GFP plasmid constructs were transformed into BW25113 strains (wt, sCB1, sKS3, sKS4) and individual colonies were obtained on LB agar plates supplemented with 100 µg/ml Carbenicillin antibiotic. The transformants were inoculated into 10 ml of LB media with carbenicillin and grown at 37°C for 16 hrs. The 10 ml of overnight grown culture was transferred to 200 ml of fresh LB media with carbenicillin and incubated at 37°C until mid-exponential phase (OD 0.4-0.6) was reached. At this point, the 200 ml culture was split into six 25 ml falcon tubes. Three of the tubes were induced with 1 mM sodium azide and remaining three tubes were used as no induction control. The cultures were placed in the 37°C shaking incubator and 2 ml of sample was collected at every time point for bulk fluorescence measurement and 200 ul was sampled to measure OD600.

Induction of pCyn-v2/v5/v6/v7/v8-GFP expression:

The pCyn-GFP plasmid constructs were transformed into BW25113 strains (wt, sCB1, sKS3, sKS4) and individual colonies were obtained on LB agar plates supplemented with 100 µg/ml Carbenicillin antibiotic. The transformants were inoculated into 10 ml of LB media with carbenicillin and grown at 37°C for 16 hrs. Then, 400 µl of overnight grown culture was transferred to 20 ml of fresh LB media with carbenicillin and incubated at 37°C until an OD600 of 0.3-0.4 was reached. At this point, 1 ml of culture was added to 12 wells in a 96 deep well plate. Three wells each were induced with 10 µM, 100 µM 1000 µM of sodium azide and three wells were left uninduced. The 96 well plate was placed in the 37°C shaking incubator for 2 hours. 200 ul of the sample was used for OD600 measurements and the residual sample was used for bulk GFP fluorescence measurement.

Bulk GFP fluorescence measurement:

The samples collected from each experiment were first centrifuged to remove the LB media. The pelleted cells were resuspended in 250 µl of BPER reagent (Bacterial Protein Extraction Reagent) and incubated at room temperature for 10 min to lyse the cells. The resultant samples were centrifuged and 200 ul of the supernatant was used to measure GFP fluorescence in a black opaque bottom 96 well plate using a spectrophotometer.

Flow cytometer data acquisition and analysis:

For measuring the GFP expression in individual cells, flow cytometer analysis was performed as reported previously by the Chundawat lab [4]. Briefly, at each timepoint for a given induction experiment, 200 µl of cells were collected. The cells were centrifuged to remove the LB media components, resuspended in 1x PBS buffer (phosphate buffered saline) and run through the Guava® easyCyte™ flow cytometer to measure the fluorescence distribution at 488 nm excitation and 525 nm emission. Guavasoft 3.3 software was used for gating live cells based on forward scatter (FSC) and side scatter (SSC) and the fluorescence associated with each cells was collected for 10,000 cells per sample and analyzed using FlowJo software.

Glycosynthase enzyme activity screening using pCyn-v2-GFP:

To test the utility of the engineered azide promoter for *in-vivo* glycosynthase enzyme activity screening, a glycosyl hydrolase family 29 fucosynthase from *Thermotoga maritima* (TmAfc-D224G)[5] was cloned into a lactose inducible vector with kanamycin resistance gene to generate pEC-TmAfc-D224G. Details about the pEC plasmid design, molecular cloning strategy, and protein expression/purification methods are provided in a recent report by our group.[4] Briefly, E. cloni® 10g competent cells (Lucigen, Middleton, WI) were first transformed with pCyn-v2-GFP plasmid and bacterial colonies were obtained on carbenicillin supplemented LB agar plates. The transformed colonies were picked and then transformed with pEC-TmAfc-D224G and plated on Kanamycin and Carbenicillin (K&C) supplemented LB-agar plates. The dual plasmid *E. coli* colonies obtained on K&C LB agar plates were grown overnight in LB media supplemented with K&C antibiotics at 37°C. Next, 750 µl of overnight grown cell culture volume was transferred to 25 ml of fresh LB K&C media and incubated at 37°C until an OD_{600} of 0.3-0.4 was reached. The cultures were then induced with 1 mM IPTG for facilitating protein expression for 1 hr at 37°C. At this point, 650 µl of culture was aliquoted into three different 3 ml culture tubes. A final concentration of 1 mM sodium azide was added to the first tube, 10 mM fucosyl-azide added to the second tube, and 10 mM fucosyl azide and 25 mM pNP-xylose added to the third tube. The culture tubes were then incubated overnight at 25 °C for 16 hrs to enable the glycosynthase reaction (along with GFP expression if azide was released as by-product from fucosyl-azide hydrolysis and/or transglycosylation with pNP-xylose). Next day, the cultures were centrifuged to recover cell pellet, cell pellet was washed and cells were resuspended in 1X PBS pH 7.4 buffer to a concentration of 10^6-10^7 cells/ml. The samples were next run on the Beckman Coulter Astrios EQ High Speed Cell Sorter instrument for single cell analysis. The GFP fluorescence distribution of the cells were measured and data collected at 488 nm excitation and 513 nm emission. The raw data was analyzed and processed using FlowJo software.

Results and Discussion

Supplementary Text (DNA Sequence Information):

>> Promoter region for pCyn-v1-GFP

TCGCAACCTATAAGTAAATCCAATGGAACTCATCATAAATGAGACTTTTACCTTATGACAATCGGCGAGTAGTCTGCCTCTCATTC
CAGAGACAGACAGAGGTTAACGATG

>> Promoter region for pCyn-v2-GFP

GGTTCCTCATTACCGTTATCATATGAACACACCATAACAAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGGCCGGGA
GCATGCGACGTCGGGCCCAATTCGCCCGATCTTAATGAATGGCCGGAAGAGGTACGGACGCGATATGCGGGGGTGAGAGGGC
AAATAGGCAGGTTCGCCTTCGTCACGCTAGGAGGCAATTCTATAAGGATCCTCGCAACCTATAAGTAAATCCAATGGAACTCGTC
AGAAATGAGACTTTTACCTTATGACAATCGGCTGGTATAATGCCTCTACTTCCAGAGACAGACATAAGGAGATTACGCATG

>> Promoter region for pCyn-v3-GFP

GGTTCCTCATTACCGTTATCATATGAACACACCATAACAAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGGCCGGGA
GCATGCGACGTCGGGCCCAATTCGGGATCCTCGCAACCTATAAGTAAATCCAATGGAACTCGTCAGAAATGAGACTTTTACCTTA
TGACAATCGGCTGGTATAATGCCTCTACTTCCAGAGACAGACATAAGGAGATTACGCATG

>> Promoter region for pCyn-v4-GFP

GGTTCCTCATTACCGTTATCATATGAACACACCATAACAAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGGCCGGGA
GCATGCGACGTCGGGCCCAATTCGCCCGATCTTAATGAATGGCCGGAAGAGGTACGGACGCGATATGCGGGGGTGAGAGGGC
AAATAGGCAGGTTCGCCTTCGTCACGCTAGGAGGCAATTCTATAAGATTGGACCGTACGCATGTCAAACTGCTGGCGAACCGCG
ATTCCACGACCGGTGCACGATTTAACTACGCCGACGTGACGACATTCCTGCTAATGCCTCGCCCGCCGGACCGCCCTCGTGATG
GGGTAGCTGGGCATGACCTTGTGACATATAACGAGAGTCTACTTGTTTAATCATCTCACGGCGAAAGTCGGGGGGACAGCAGCC
GCTGCAGACATTATACCGCAACTACACCAAGCTGAGATAACTCCGTAGTTGACTACGCATCCCTCTAGGCCTTACTTAACCGGAT
ACAGTGACTTTGACAGGTTTGTGGGCTACAGCAATCACTTGCATAGCTGCGTATGGAGGAAGCAACTCTTGGGTGTTAGTATGTT
GACCCCTGTATTAGGGATGCGGGTAGTAGATGTGGGCAGAGACACCCAGGTCAAGTACACGACCCTCTCGTAGGAGGTGTTCC
AGATCACCATACCACCATACCATTCGAGCATGGCACTATGTACGCTGTCCCCATTCTGGTAGTCATCATCCCTATCACGGTTTCG
AGTGACTGGTGACGGATATCCCCCACGAATGGAGATCTTATTCACAGTCGGTCACATTGGAGTGCTCCTTGACTAATCAGCTTGG
CCAGGTCTGTTGGGCCTCCGTGCCCCGAGTTTCGGCGCTGTGCTGCCGAGAGTCGGCCATTGTCATTGGGGCCTCACTTGTGG
ATACCCCGACCTATTTTGACGGGACCACTCGCGGTAGTCGTTGGGCTTATGCACCGTGAAGTCCTCCGCCGGCCTCCCCCCTAC
AAAAGATGATAAGCTCCGGCAAGCAATATTGAACAACGCAAGGATCGGCGATATAAACAGAGAAACGGCTGATTACTCTTGTTGG
TGTGGTATCGCTAAACTGGGATCCTCGCAACCTATAAGTAAATCCAATGGAACTCGTCAGAAATGAGACTTTTACCTTATGACAAT
CGGCTGGTATAATGCCTCTACTTCCAGAGACAGACATAAGGAGATTACGCATG

>> Promoter region for pCyn-v5-GFP

GGTTCCTCATTACCGTTATCATATGAACACACCATAAGGAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGGCCGGGA
GCATGCGACGTCGGGCCCAATTCGCCCGATCTTAATGAATGGCCGGAAGAGGTACGGACGCGATATGCGGGGGTGAGAGGGC
AAATAGGCAGGTTCGCCTTCGTCACGCTAGGAGGCAATTCTATAAGGATCCTCGCAACCTATAAGTAAATCCAATGGAACTCGTC
AGAAATGAGACTTTTACCTTATGACAATCGGCTGGTATAATGCCTCTACTTCCAGAGACAGACATAAGGAGATTACGCATG

>> Promoter region for pCyn-v6-GFP

GGTTCCTCATTACCGTTATCATATGAACACGCCATAAGGAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGGCCGGGA
GCATGCGACGTCGGGCCCAATTCGCCCGATCTTAATGAATGGCCGGAAGAGGTACGGACGCGATATGCGGGGGTGAGAGGGC
AAATAGGCAGGTTCGCCTTCGTCACGCTAGGAGGCAATTCTATAAGGATCCTCGCAACCTATAAGTAAATCCAATGGAACTCGTC
AGAAATGAGACTTTTACCTTATGACAATCGGCTGGTATAATGCCTCTACTTCCAGAGACAGACATAAGGAGATTACGCATG

>> Promoter region for pCyn-v7-GFP

GTTCCTCATTACCGTTATCATATGAACACACCATAAGGAAAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGGCCGGG
AGCATGCGACGTCGGGCCCAATTCGCCCGATCTTAATGAATGGCCGGAAGAGGTACGGACGCGATATGCGGGGGTGAGAGGG
CAAATAGGCAGGTTCGCCTTCGTCACGCTAGGAGGCAATTCTATAAGGATCCTCGCAACCTATAAGTAAATCCAATGGAACTCGT
CAGAAATGAGACTTTTACCTTATGACAATCGGCTGGTATAATGCCTCTACTTCCAGAGACAGACATAAGGAGATTACGCATG

>> Promoter region for pCyn-v8-GFP

GGAGAGTTCCTCATTACCGTTATCATATGAACACACCATAAGGAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGGCC
GGGAGCATGCGACGTCGGGCCCAATTCGCCCGATCTTAATGAATGGCCGGAAGAGGTACGGACGCGATATGCGGGGGTGAGA
GGGCAAATAGGCAGGTTCGCCTTCGTCACGCTAGGAGGCAATTCTATAAGGATCCTCGCAACCTATAAGTAAATCCAATGGAACT
CGTCAGAAATGAGACTTTTACCTTATGACAATCGGCTGGTATAATGCCTCTACTTCCAGAGACAGACATAAGGAGATTACGCATG

Figure S1. GFP fluorescence of cell lysate measured using spectrophotometer after inducing BW25113-Wt and BW25113-sCB1 strains containing pCyn-v1-GFP with 1 mM cyanate and azide. Total induction time was 4 hours. Error bars indicate one standard deviation from reported mean values from two biological replicates.

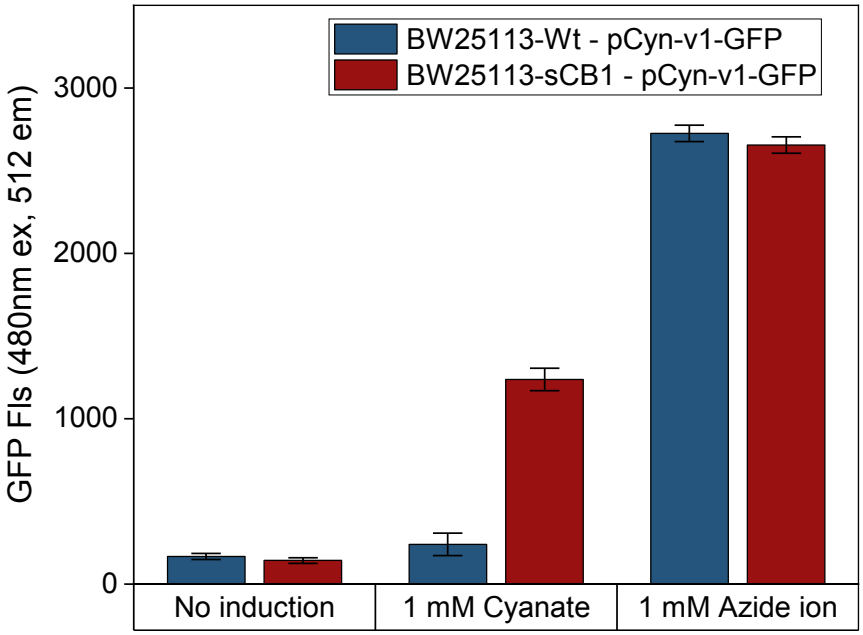


Figure S2. A. Illustration of the promoter regions of pCyn-v1-GFP, pCyn-v2-GFP, pCyn-v3-GFP, pCyn-v4-GFP. Fluorescence measurements of lysates from BW25113-Wt cells containing plasmids pCynv2-GFP (blue square), pCyn-v3-GFP (red triangle), and pCyn-v4-GFP (green circle) and induced with **B.** 1 mM sodium azide, or **C.** uninduced cells for 4 hours at 37°C. Error bars indicate one standard deviation from reported mean values from three biological replicates.

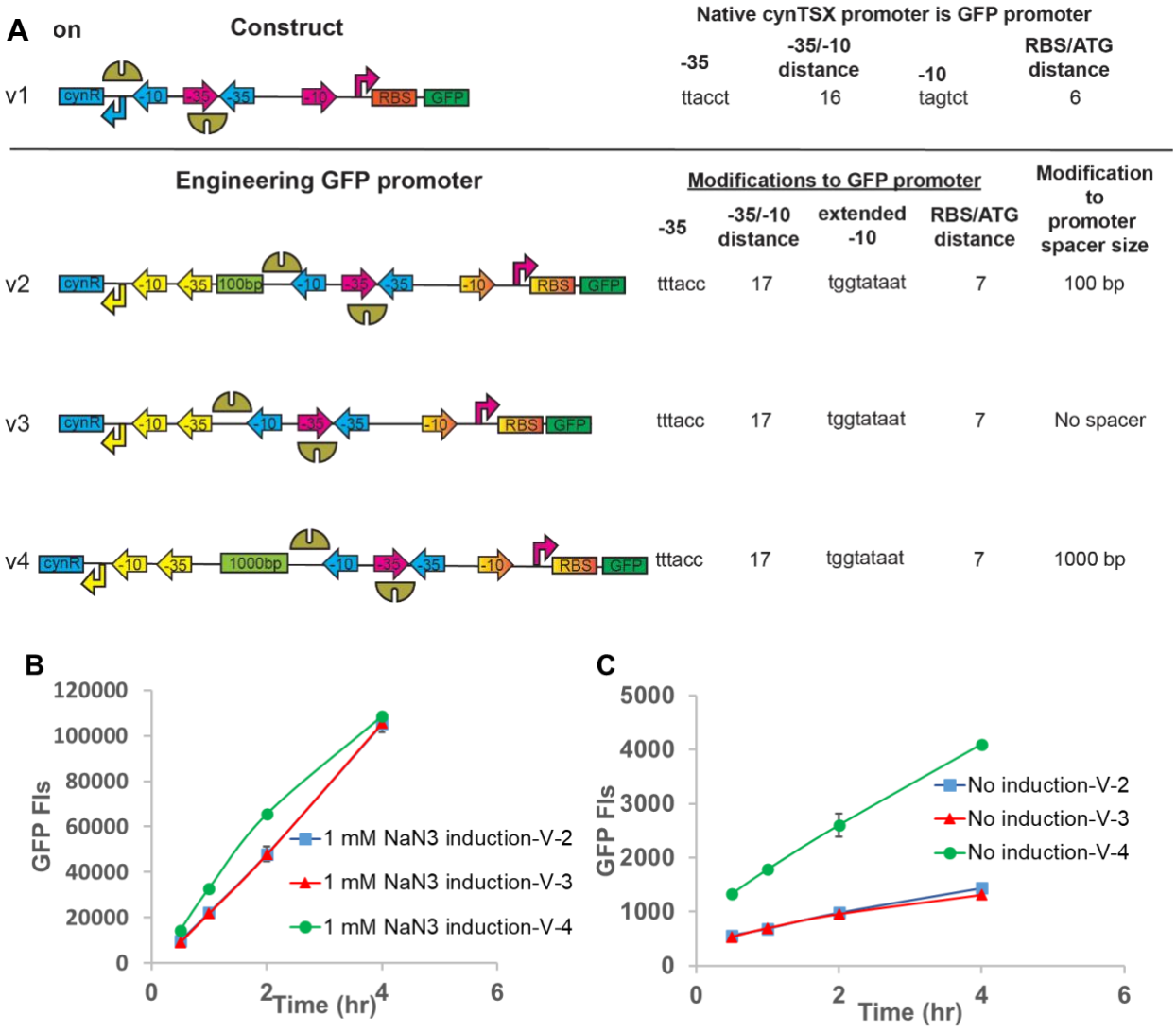


Figure S3. *E. coli* cells incubated with varying concentrations of sodium azide to study the inhibitory effect of azide on cell growth. Cells were grown in LB media at 37°C for 18 hours and the OD of the cells were measured at various intermittent time points after introduction of sodium aide at time t=0. Error bars indicate one standard deviation from reported mean values from three biological replicates.

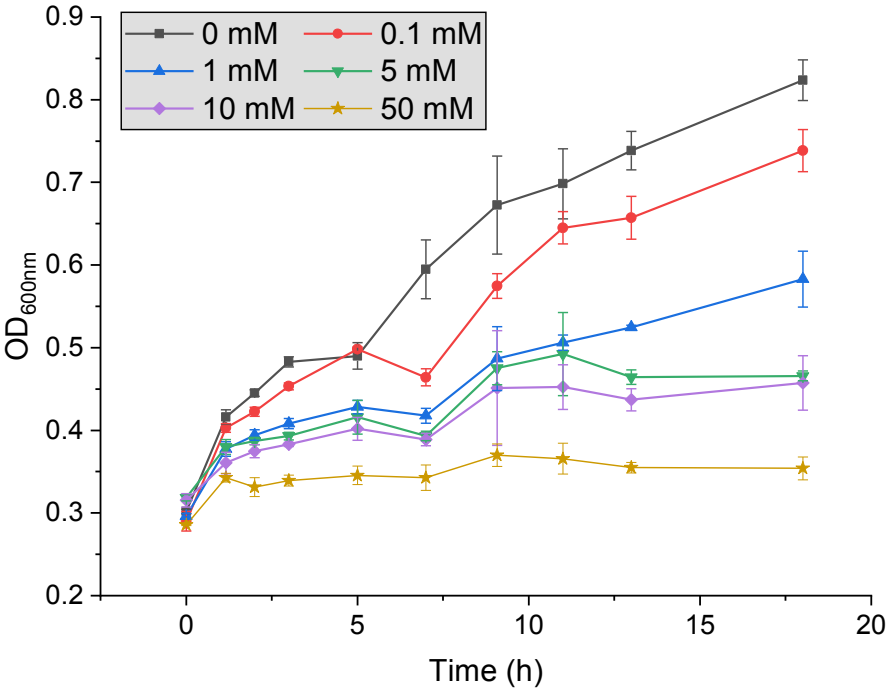


Figure S4. Cell lysate GFP fluorescence from *E. coli* BW25113 cells induced using 1 mM sodium azide but at different starting cell optical density (OD_{600}). Cells were grown until the desired OD in LB media and induced with 1 mM sodium azide and grown then for 4 hours with measurements taken at 1 hour (grey square, left axis) and 4 hours (red square, right axis), respectively. The GFP fluorescence on the y-axis is normalized with the measured OD after induction for specified culture times. Error bars indicate one standard deviation from reported mean values from three biological replicates.

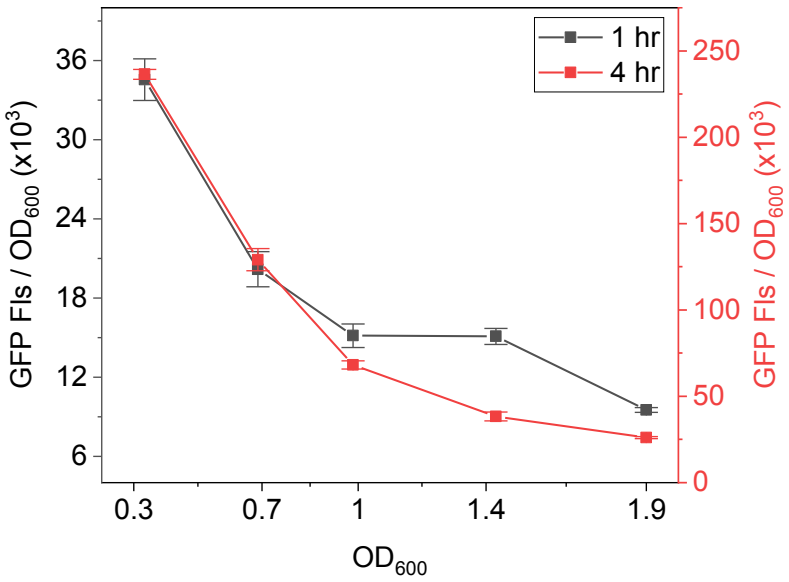
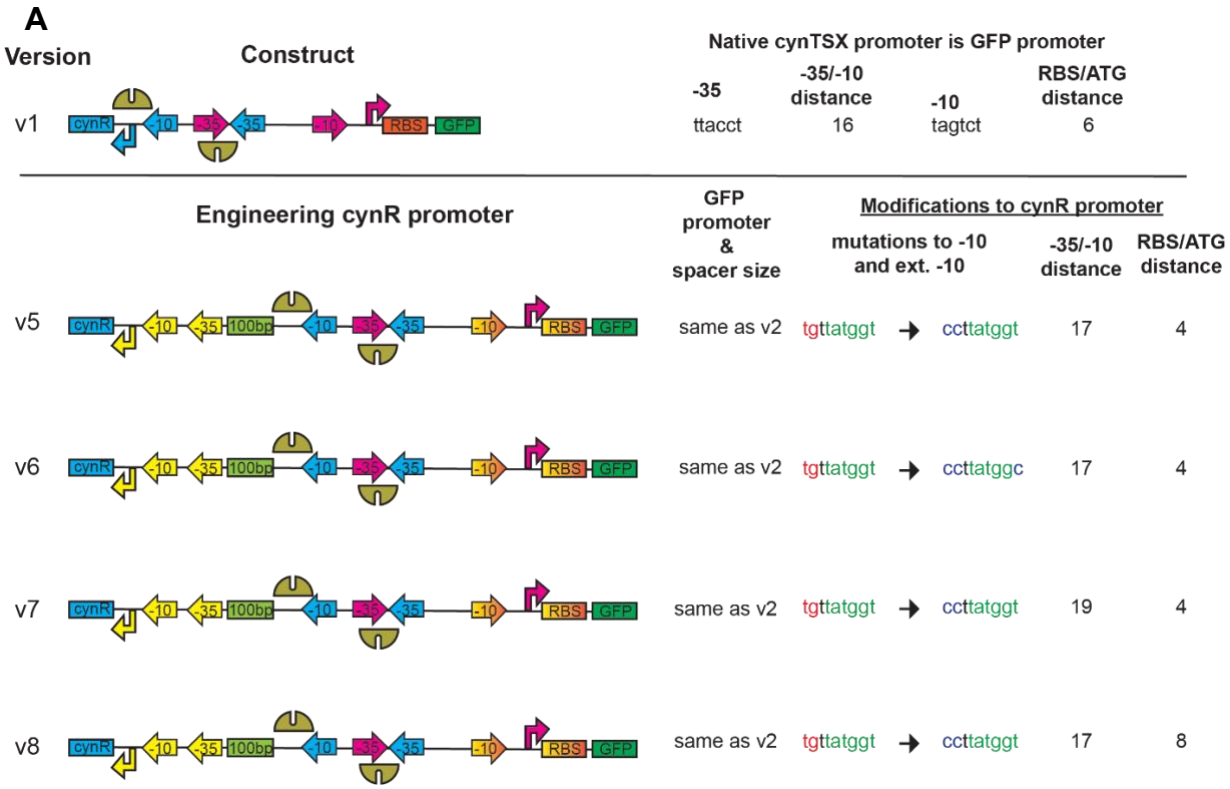


Figure S5. A. Illustration of the promoter regions of pCyn-v5-GFP, pCyn-v6-GFP, pCyn-v7-GFP, pCyn-v8-GFP. **B.** Fluorescence measurements of BW25113-Wt cells with plasmids pCyn-v2-GFP, pCyn-v5-GFP, pCyn-v6-GFP, pCyn-v7-GFP incubated with 1 mM sodium azide for 2 hours. The fluorescence and OD measurements were taken after 2 hours of induction. The GFP fluorescence was normalized with the OD and plotted as bar graphs. Error bars indicate one standard deviation from reported mean values from three biological replicates.



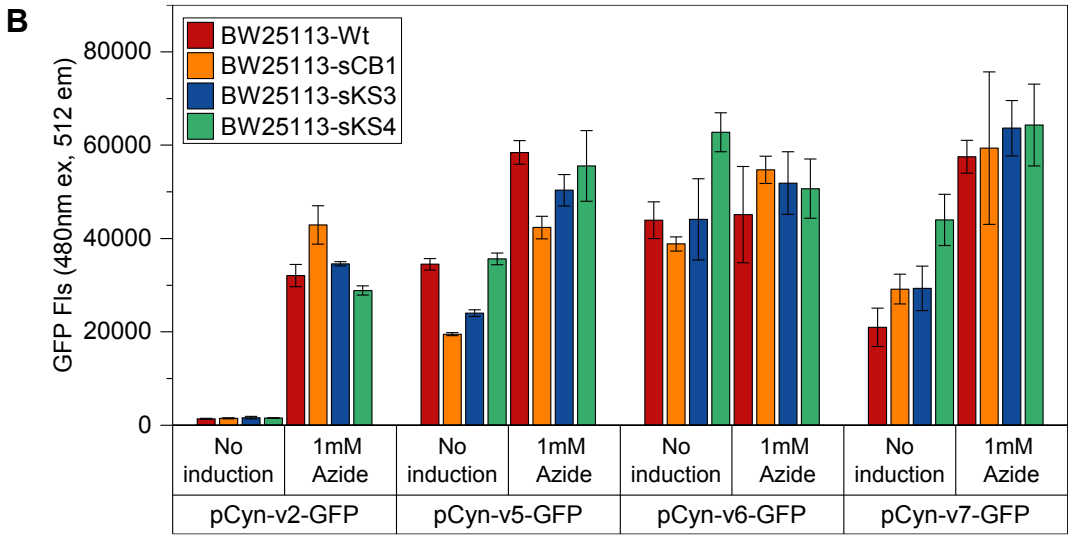


Figure S6. Calibration curve for GFP protein concentration (mg/ml) and measured cell lysate fluorescence measurements.

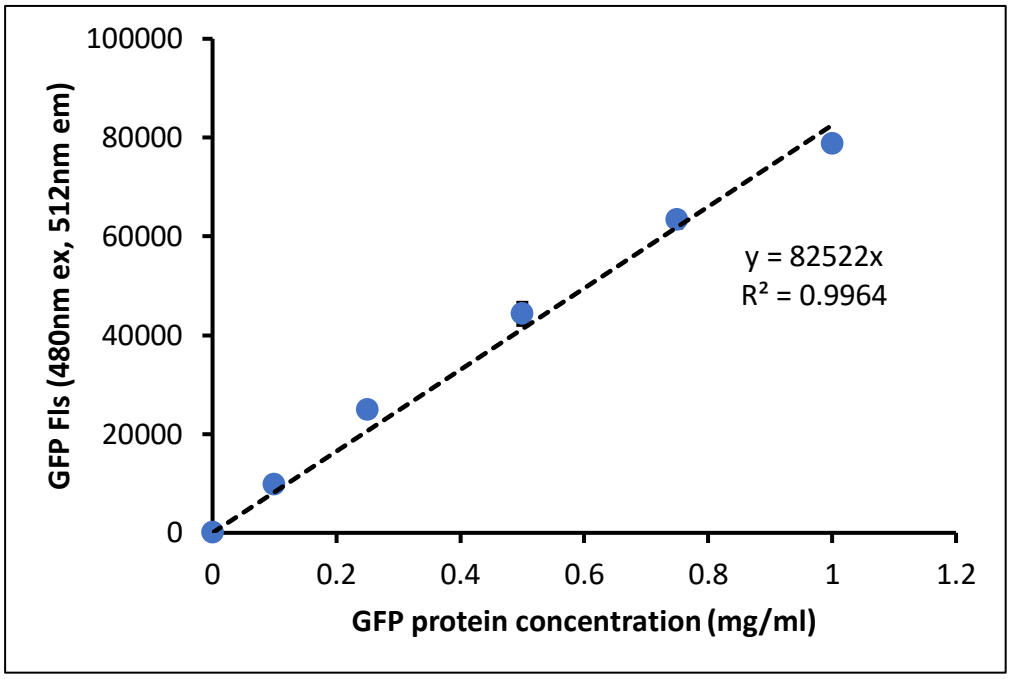


Table S1. Genotype of the bacterial knockout strains generated in this study.

Strain name	Genotype
BW25113-wt	Wild type
BW25113-sCB1	BW25113 cynS::FRT
BW25113-sKS3	BW25113 cynX::FRT
BW25113-sKS4	BW25113 cynR,cynSTX::FRT

Table S2. (A) Primers used for performing Sequence and ligation independent cloning (SLIC) and site directed mutagenesis (SDM).

Primer name	Sequence
k1	TGATAAGCGTAGCGCATCAGGCAATTCCAGCCGCAGACCTGTGTCAGCGGGTGTAGGCTGGAGCTGCTTCGAAG
k2	TCTACATTAGCCGCATCCGGCATGAACAAAGCGCAGGAACAAGCGTCGCACATATGAATATCCTCCTTAGTTCC
p1	AACCAGGATAATGATACAGATTAAATCAGAACGCAGAAG
p2	TTGTAATCGATAACGTAAATGCATGCCGCTTC
p3	GCATTTACGTTATCGATTACAAACGTTGAACGAC
p4	AATCTGTATCATTATCCTGGTTCTTTCCCCCA
p5	CGGTGGCAGCTCCAAAGGTGAAGAACTG
p6	AATGAGGAACCATGCTCTCTCGACATATC
p7	CGAGAGAGCATGGTTCCTCATTACCGTTATCATATG
p8	CCTTTGGAGCTGCCACCGCCATG
p9	CGGGCCCAATTCGGGATCCTCGCAACCTATAAGTAAATCC
p10	CGAGGATCCCGAATTGGGCCCGACGTC
p11	CGCTAAACTGGGATCCTCGCAACCTATAAGT
p12	GTACGGTCCAATCTTATAGAATTGCCTCCTAGCGTG
p13	CAATTCTATAAGATTGGACCGTACGCATGTC
p14	GCGAGGATCCCAGTTTAGCGATACCACACCA
p15	ATGAACACACCATAAGGAAGATGCATGCAGCT
p16	AGCTGCATGCATCTTCCTTATGGTGTGTTCAT
p17	GTTATCATATGAACACGCCATAAGGAAGATGCA
p18	TGCATCTTCCTTATGGCGTGTTCATATGATAAC
p19	CATATGAACACACCATAAGGAAAAGATGCATGCAGCTGTC
p20	GACAGCTGCATGCATCTTTTCCTTATGGTGTGTTCATATG
p21	AGAGTTCCTCATTACCGTTATCATATG
p22	CTCTCCATGCTCTCTCGACATATC

Table S2. (B) Gene block fragments used for cloning and generating plasmid constructs.

DNA fragment name	DNA Sequence
gblock1	ATCGATTACAAACGTTGAACGACTGGGTTACAGCGAGCTTAGTTTATGCCGGATGCGGCGTGAACGCCTTATCCGGC CTACGTAGAGCACTGAACTCGTAGGCCTGATAAGCGTAGCGCATCAGGCAATTCCAGCCGCTGATCTGTGTCAGCGG CTACCGTGATTCATTCCCGCCAACAACCGCGCATTCCTCCAACGCCATGTGCAAAAATGCCTTCGCAGCGGCTGTCT GCCAGCTGTAGTTTATGCCGGATGCGGCGTGAACGCCTTATCCGGCCTACGTAGAGCACTGAACTCGTAGGCCTGAT AAGCGTAGCGCATCAGGCAATTCCAGCCGCAGACCTGTGTCAGCGGCTACCGTGATTCATTTCCGCCAACAACCGCG CATTTATCCAACGCCATGTGCAAAAATGCCTTCGCGGCGGCTGTCTGCCAGCTATTTTTCCGCCGCAACAAAACCGCC GTTCTCTCCAGTAGTGGCGGGGCAAGAGAAATAGCTTTAAGCCCGTCATGTTGTGTGGCAATCGCTGCTGGTAACAA TGTGGAAAGGGAAGTGCGGCGAATCAGCTCCAGAACCGCGCTAATTGAGTTCGCCTCAATGACCACCTGTGGATGTA GCCCCGCTTTCTCGCAGTAGTGGTCAATTTGCTCTCTGGTGGCAAATTCCGCGCTGAGCAGGACCAGTTTTTCATCAT GCAAGCGACTCAACGCCACCTGTTCATGGACGGCCAGCGGATGATGTTGCGCCACGACTAACGCTAAACTTTCTGTC AGTAAAGGAATTGCCTCCAGCTCCGGCGAATGCACAGGCGCGAAGGCAATCCCAACGTCCAACTCGTCGCGGCAAA GCATATCCTCGATTTTCTCCTGCGACATTTCCTGTAGCTGGAGCGTGATGCTGGGATAGCGCGCATAGAAATCCGCC ATTAAGGGGCCGATAAAGTAGCTCGTAAAGGTGGGGGTGACGGCGATACGCAGCGATCCTCGCGTCAGATCGGCAA CATCATGAATCGCCCGTTTACCCGCCCCCAGTTCCTGTAACGCCCGGCTGGCGTACTGTCGCCAGACTTCTCCTGCA TCAGTGAGACGAATCGTTCGCCCGCTACGGTCAAACAGCGGCACGCCTAAACTCTCCTCTAACTGGCGAATCTGCTG GGAAAGCGCAGGTTGGGAGACGTGCAACGCACTGGCGGCACGGGTGAAGCTGCCATGTTCAGCCACGGCAAGAAA ATAATTGATATGTCGAGAGAGCATTCGCAACCTATAAGTAAATCCAATGGAACTCATCATAAATGAGACTTTTACCTTAT GACAATCGGCGAGTAGTCTGCCTCTCATTCCAGAGACAGACAGAGGTTAACGATG
gblock2	GGTTCCTCATTACCGTTATCATATGAACACACCATAACAAAGATGCATGCAGCTGTCTAAATCCCGCGGCCATGGCGG CCGGGAGCATGCGACGTCGGGCCCAATTCGCCCGATCTTAATGAATGGCCGGAAGAGGTACGGACGCGATATGCGG GGGTGAGAGGGCAAATAGGCAGGTTCGCCTTCGTCACGCTAGGAGGCAATTCTATAAGGATCCTCGCAACCTATAAG TAAATCCAATGGAACTCGTCAGAAATGAGACTTTTACCTTATGACAATCGGCTGGTATAATGCCTCTACTTCCAGAGAC AGACATAAGGAGATTACGCATGCATCACCATCATCACCATCACCATGGCGGTGGCAGC
gblock3	GATTGGACCGTACGCATGTCAAACTGCTGGCGAACCGCGATTCCACGACCGGTGCACGATTTAACTACGCCGACGTG ACGACATTCCTGCTAATGCCTCGCCCGCCGGACCGCCCTCGTGATGGGGTAGCTGGGCATGACCTTGTGACATATAA CGAGAGTCTACTTGTTTAATCATCTCACGGCGAAAGTCGGGGGGACAGCAGCCGCTGCAGACATTATACCGCAACTA CACCAAGCTGAGATAACTCCGTAGTTGACTACGCATCCCTCTAGGCCTTACTTAACCGGATACAGTGACTTTGACAGG TTTGTGGGCTACAGCAATCACTTGCATAGCTGCGTATGGAGGAAGCAACTCTTGGGTGTTAGTATGTTGACCCCTGTA TTAGGGATGCGGGTAGTAGATGTGGGCAGAGACACCCAGGTCAAGTACACGACCCTCTCGTAGGAGGTGTTCCAGA TCACCATACCACCATACCATTCGAGCATGGCACTATGTACGCTGTCCCCATTCTGGTAGTCATCATCCCTATCACGGT TTCGAGTGACTGGTGACGGATATCCCCCACGAATGGAGATCTTATTCACAGTCGGTCACATTGGAGTGCTCCTTGACT AATCAGCTTGGCCAGGTCTGTTGGGCCTCCGTGCCCCGAGTTTCGGCGCTGTGCTGCCGAGAGTCGGCCATTGTCA TTGGGGCCTCACTTGTGGATACCCCGACCTATTTTGACGGGACCACTCGCGGTAGTCGTTGGGCTTATGCACCGTGA AGTCCTCCGCCGGCCTCCCCCCTACAAAAGATGATAAGCTCCGGCAAGCAATATTGAACAACGCAAGGATCGGCGAT ATAAACAGAGAAACGGCTGATTACTCTTGTTGGTGTGGTATCGCTAAACTG

SI References

1. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640–6645 (2000).
2. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, (2006).
3. Stevenson, J., Krycer, J. R. ., Phan, L. & Brown, A. J. A Practical Comparison of Ligation-Independent Cloning Techniques. *PLoS One* **8**, e83888 (2013).
4. Agrawal, A. *et al.* Click-chemistry enabled directed evolution of glycosynthases for bespoke glycans synthesis. *bioRxiv* 2020.03.23.001982 (2020) doi:10.1101/2020.03.23.001982.
5. Cobucci-Ponzano, B. *et al.* beta-Glycosyl azides as substrates for alpha-glycosynthases: preparation of efficient alpha-L-fucosynthases. *Chem. Biol.* **16**, 1097–108 (2009).