bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Designing efficient genetic code expansion in Bacillus subtilis to gain biological insights

Devon A. Stork[1,2], Georgia R. Squyres[2], Erkin Kuru[1,3], Katarzyna A. Gromek[4], Jonathan Rittichier[1,3], Aditya Jog[1], Briana M. Burton[4], George M. Church[1,3]*, Ethan C. Garner[2]*, and Aditya M. Kunjapur[1,5]*

[1] Department of Genetics, Harvard Medical School, Boston, MA 02157

[2] Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

[3]Wyss Institute for Biologically Inspired Engineering, Boston, MA 02115

[4]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706

[5]Present Address: Department of Chemical and Biological Engineering, University of Delaware, Newark, DE 19716

* Corresponding authors.

Aditya M. Kunjapur Email: kunjapur@udel.edu

Ethan Garner Email: egarner@g.harvard.edu

George M. Church Email: gchurch@genetics.med.harvard.edu

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Bacillus subtilis is a model Gram-positive bacterium, commonly used to explore questions across bacterial cell biology and for industrial uses. To enable greater understanding and control of proteins in *B. subtilis*, we demonstrate broad and efficient genetic code expansion in *B. subtilis* by incorporating 20 distinct non-standard amino acids within proteins using 3 different families of genetic code expansion systems and two choices of codons. We use these systems to achieve click-labelling, photo-crosslinking, and translational titration. These tools allow us to demonstrate differences between *E. coli* and *B. subtilis* stop codon suppression, validate a predicted protein-protein binding interface, and begin to interrogate properties underlying bacterial cytokinesis by precisely modulating cell division dynamics *in vivo*. We expect that the establishment of this simple and easily accessible chemical biology system in *B. subtilis* will help uncover an abundance of biological insights and aid genetic code expansion in other organisms.

Introduction

Bacillus subtilis is a gram-positive soil bacterium naturally found in the plant rhizosphere[1,2] and employed in basic and applied research[3]. Genetic tractability and ease of culture have made *B. subtilis* the most commonly used gram-positive model-organism. In fundamental research, *B. subtilis* is a model organism for the study of endospore formation[4], asymmetric cell division[5], biofilm formation[6], and multicellular behavior[7]. In applied research, *B. subtilis* is a probiotic for plants and animals including humans, spanning agricultural, nutritional, and medical applications[8–10] and is used for industrial protein production[11]. A broad range of available genetic tools including inducible promoters and protein tags[5,12] have made *B. subtilis* an attractive model for both fundamental research and industrial biotech, but limitations in the chemistry and

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

flexibility of these tools prevent full utilization of the organism's potential. The recent use of CRISPRi[13] and optogenetic transcriptional control[14] transfer tools from *E. coli* to *B. subtilis*, though each is limited in their expression range and titratability. New types of peptidoglycan tags[15] allow chemical modification and imaging of the development of the cell wall, but not proteins. To examine broader mechanisms of growth and division in *B. subtilis*, finer chemical functionalization and titration tools are needed. We postulate that genetic code expansion through the incorporation of nonstandard amino acids (nsAAs) in proteins could achieve that in *B. subtilis* as has been demonstrated in *E. coli*.

The technique of site-specific nsAA incorporation widens the chemistry accessible to biological systems by allowing the use of diverse chemical functional groups in protein design and experimentation[16,17]. More than 200 different nsAAs derived from tyrosine, pyrrolysine, serine, leucine, and tryptophan have been incorporated into proteins, primarily in *E. coli* and mammalian systems[18]. Engineered variants of the corresponding orthogonal aminoacyl-tRNA synthetase (AARS) and tRNA pairs are used to attach nsAAs to tRNA that direct them for ribosomal incorporation at specific codons. An engineered AARS/tRNA system must be orthogonal to the native translational machinery of the host organism; thus, separately engineered systems must often be used in different organisms[16]. The requirements of orthogonality and balanced expression levels of AARS and tRNA have limited the utility of genetic code expansion in bacterial systems outside of *E. coli*[19,20].

Here we provide a broad investigation of the portability of many different sets of genetic code expansion technologies from *E. coli* to *B. subtilis*, and we use several of these systems to gain insights about protein translation and cell division in *B. subtilis*. This work builds substantially upon recent demonstrations of genetic code expansion in *B. subtilis*, where single

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

pyrrolysine[21] and tyrosine[22] nsAAs were incorporated. We expand the number of nsAAs in *B. subtilis* to 20, using 3 families of stable, genomically integrated AARS constructs and achieve higher incorporation efficiency than in previous studies. The functions of these nsAAs range from biorthogonal tagging to photocrosslinking and to fluorescence, with broad experimental utility. To facilitate the further application of this technology and explain complications noted in previous work[21,22], we firstly show that unlike in non-recoded *E. coli*[23–25], nsAAs incorporate efficiently at amber stop codons in native *B. subtilis* genes. Secondly, the incorporation of photocrosslinking nsAAs allows demonstration of binding interactions of secreted proteins homologous to virulence factors. While previous work has shown nsAAs can modulate translational rates[22], we demonstrate a tighter system with a larger dynamic range and make a detailed comparison of nsAA-titration to many different *B. subtilis* promoters. Finally, we use this system to facilitate novel biological discovery. We are able to precisely modulate the dynamics of the division protein FtsZ, an essential filament responsible for coordinating cell division[26]. Our data support the theory that FtsZ filaments must be above a minimal length to accomplish cell division[27]. These results and our deposited strains (at the Bacillus Genomic Stock Center) will facilitate the use of nsAAs for general use in *B. subtilis* across research and industrial applications.

Results

Activity of diverse Orthogonal Amino-Acyl tRNA Synthetases in *B. subtilis*

To begin testing genetic code expansion in *B. subtilis*, we genomically integrated a codon-optimized *Methanococcus jannaschii*-tyrosyl-tRNA synthetase (MjTyrRS) variant called bipARS capable of incorporating the nsAA **1** at the *lacA* locus. This synthetase was accompanied by the corresponding tRNA and a panel of constitutive promoters driving AARS and tRNA

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

expression. An IPTG-inducible reporter cassette was integrated at the *amyE* locus for expression of an mNeongreen fluorescent protein with a UAG stop codon at position 2 (Figure 1A). After an initial promoter screen with **1**, we determined that a pVeg/pSer AARS/tRNA promoter combination yielded the best incorporation with low background (Supp. text & Supp. Fig. 1A-C), and **1** incorporation within mNeongreen was confirmed by mass spectrometry (Supp. Fig. 2E).

To explore the potential of enabling more diverse chemical functionality in *B. subtilis,* we sought to incorporate additional nsAAs that contain sidechains known to function as fluorescent probes, handles for click chemistry, and photo-crosslinkers. To this end, we built multiple additional AARS cassettes. First, we created more MjTyrRSs variant cassettes to diversify the tyrosine-based nsAAs available. Second, to investigate the use of nsAAs that are based on non-phenyl sidechains, we used the *Saccharomyces cerevisiae* tryptophan synthetase (ScWRS)[28], and both *Methanosarcina barkeri* pyrrolysine synthetase (MbPylRS)[29] and *Methanomethylophilus alvus* pyrrolysine synthetase (MaPylRs)[30]. All AARS/tRNA pairs demonstrated activity in *B. subtilis* except for MbPylRS, which is known to be poorly soluble in bacteria due to the presence of an N-terminal domain that the homologous MaPylRS lacks[31,32] (Supp. Figure 1D). Encouragingly, the incorporation activity of these diverse synthetases on their corresponding nsAAs (**1-6**) reflected published activity in *E. coli*[18,33] (Fig 1B-D), which suggests that AARS engineering performed in *E. coli* would allow predictable activity and specificity in *B. subtilis*. We utilized the substrate promiscuity of the MjTyrRSs to incorporate many more nsAAs, bringing the total incorporated in *B. subtilis* up to 20 (Supp. Figure 1E-F), including nsAAs capable of fluorescence, photocrosslinking, click chemistry, metal chelation and more.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

These nsAAs cover most of the applications of genetic code expansion.

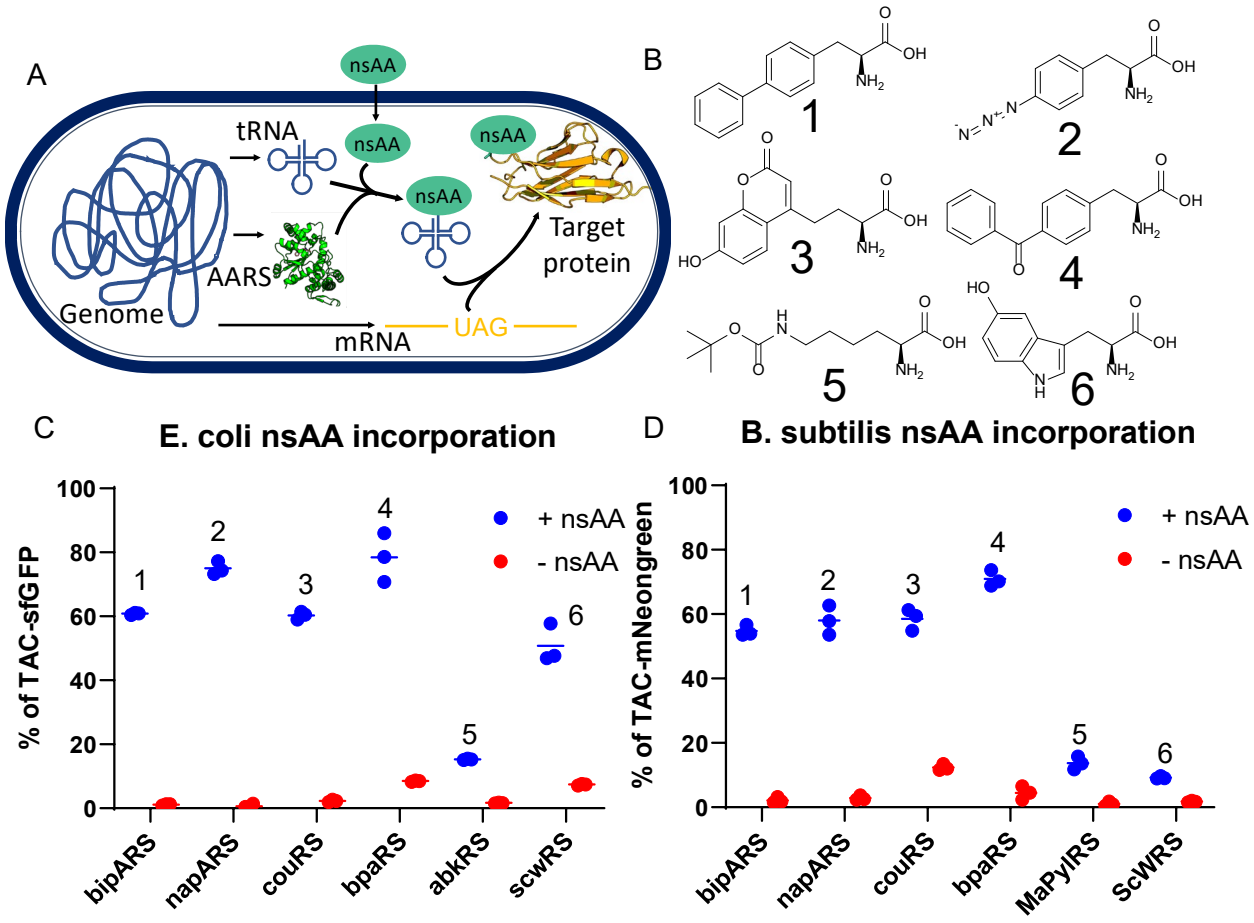

Figure 1: Genetic code expansion via nsAA incorporation in *B. subtilis*. A) Scheme of nsAA incorporation, with genomically integrated AARS and tRNA constructs incorporating externally provided nsAA into a genomically expressed gene containing an in-frame TAG amber stop codon. **B)** Chemical structure of 6 nsAAs of primary interest to this study. **1**:biphenylalanine (bipA), **2**:p-azidophenylalanine (pAzF) , **3**:Coumarin-nsAA (CouAA), **4**:Benzoylphenylalanine (BpA), **5**:boc-Lysine (boc-K), **6**:5-Hydroxytryptophan (5OHW) **C&D**) nsAA incorporation in **C)** *E. coli* and **D)** *B, subtilis*. nsAA used is indicated above datapoints, synthetase variant below each x-axis. In all cases, signal is normalized to an identical reporter containing a TAC Tyr codon in place of a TAG amber codon. Individual data points for biological triplicates shown, with a horizontal bar at mean. In **C)**, c321 recoded *E. coli* from plasmid based AARS & tRNA and a genomic sfGFP reporter containing a single UAG codon in an N-terminal linker. **D)** nsAA incorporation in *B. subtilis* from a genomic AARS & tRNA and a genomic M10SmNeongreen reporter containing a single UAG codon at position 2, immediately following the start codon.

To confirm nsAA incorporation, we expressed FLAG-tagged mNeongreen with nsAAs incorporated into an elastin-like peptide optimized for mass-spectrometry detection of nsAAs[34]. Purification and analysis of peptides demonstrated incorporation of all the tyrosine-based nsAAs shown in figure 1B (Supp. Figure 2). Instead of detecting **5** at the indicated position, lysine was detected instead. This is likely due to the known deprotection of boc during the chromatographic

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

step of peptide identification[35]. Additionally, **6** was not detected likely due to the inability to purify sufficient protein due to low expression levels (Supp. Figure 1G).

Proteome-wide incorporation of nsAAs at UAG sites

One potential application of nsAA incorporation is fluorescent tagging for subcellular microscopy[36,37]. We attempted to use nsAA **3** incorporation to specifically localize small, dynamic cell-division components that have not previously been tagged. However, high background fluorescence hindered this goal (Supp. Fig. 3). The fluorescent coumarin nsAA was washed out of cells effectively in the absence of synthetase; however, synthetase expression alone resulted in cellular retention of fluorescent nsAA (Supp. Fig. 4A). Whole protein lysate analysis revealed that the fluorescent nsAA was incorporated off-target into many proteins across the proteome (Fig. 2A). This is a surprising phenomenon as it is not observed in wild-type *E. coli*[23–25]. To enrich for the native proteins that are the most likely destinations for off-target nsAA incorporation, the nsAA **2** was incorporated for click-chemistry based enrichment. Mass-spectrometry analysis of enriched proteins revealed that **2** was highly enriched in proteins ending with a UAG stop codon as compared to UAA or UGA (Figure 2B). Analysis of the identity of the highly enriched proteins revealed that 23 of the 569 UAG-containing proteins made up 71% of the observed proteomic incorporation events (Supp. table 1). Despite the apparent suppression of genomic UAG stop codons, no significant decrease in doubling times was observed (Supp. Fig. 4B). These results explain speculation in previous works that nsAAs incorporate into native genes in *B. subtilis*[21].

Since general incorporation into the *B. subtilis* proteome will interfere with specific labelling approaches, we sought to reduce the level of this misincorporation. One approach to reduce the level of background nsAA incorporation and to further extend nsAAs in *B. subtilis*

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

would be to use a codon rarer than the UAG codon in *B. subtilis*, such as the quadruplet UAGA codon. To do this, we cloned quadruplet versions of tyrosine synthetase cassettes. These contained a tRNA with a quadruplet anticodon and synthetases modified with the F261S and D286E mutations, which have been shown to encourage UCUA-tRNA aminoacylation for p-aceytlphenylalanine[38]. This quadruplet system was able to successfully incorporate at UAGA codons with low efficiency but also incorporated into the UAG codon with similar efficiency (Supp. Fig. 4C). This is not the only report of lack of specificity of the UAGA-tRNAs[39] and suggests future issues for attempts to use quadruplet codons to increase the available codon space. In addition to the lack of specificity for the UAGA, 217 of the 596 *B. subtilis* UAG codons are UAGA, including 11/21 of those highly represented in the mass-spec results (Supp. Table 1). Therefore, it is not surprising we did not see any increase in specificity for a target incorporation site in whole-cell lysate (Supp. Fig. 4D). There are other potential avenues to prevent proteomic misincorporation, which include making translational termination stricter by upregulating RF1 or limited recoding of the genes which we found to misincorporate highly. These results indicate potentially unknown features of *B. subtilis* protein biology, such as the high rate of native UAG stop codon suppression of native genes compared to *E. coli*[23–25].

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

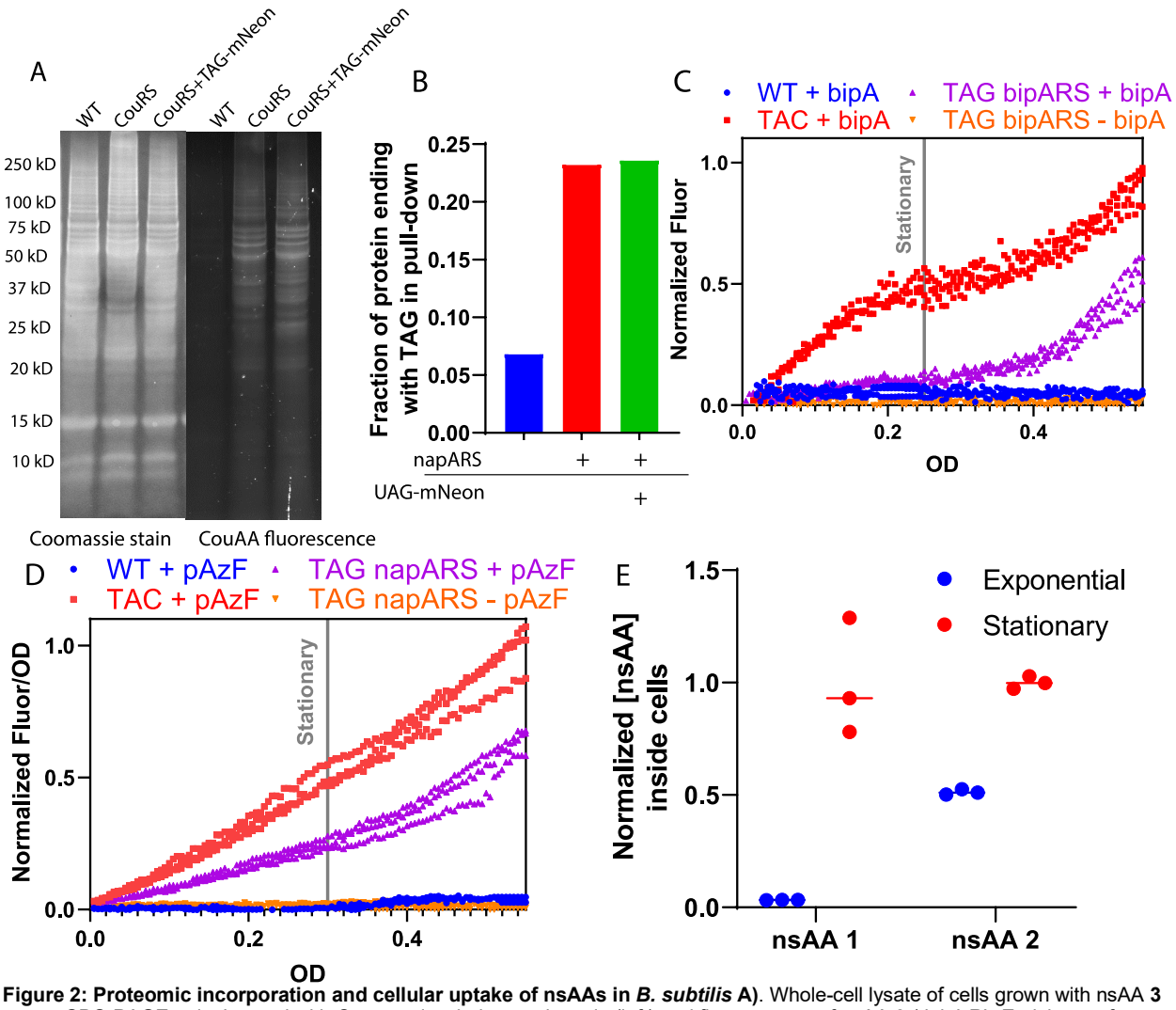

Figure 2: Proteomic incorporation and cellular uptake of nsAAs in *B. subtilis* A). Whole-cell lysate of cells grown with nsAA **3** run on SDS-PAGE gels. Imaged with Coomassie whole protein stain (left) and fluorescence of nsAA **3** (right**) B).** Enrichment for nsAA incorporation was followed by analysis for the fraction of peptides originating from proteins ending in the TAG codon. The azide-containing **2** was incorporated, and a click pulldown performed, after which peptides were detected by mass spectrometry **C&D)** Timecourse of mNeongreen fluorescence vs. OD normalized to maximal UAC-mNeongreen signal of **C) 1** and **D) 2** incorporation in the presence and absence of corresponding synthetases. 3 biological replicates are shown. Grey vertical lines indicate approximate start of stationary phase. (See Supp Fig 5 for corresponding growth curves). **E)** Relative concentration of nsAAs inside cells grown to indicated growth phase before being washed and lysed, then internal concentration of nsAAs **1** & **2** measured by LCMS. Values are relative, normalized to OD and to concentration in cells during stationary phase. 3 biological replicates are shown, with a horizontal bar at the median.

Cellular Uptake of nsAAs

As we were conducting nsAA incorporation experiments in *B. subtilis*, we discovered that a variety of factors seemed to influence nsAA transport that may not play a role in *E. coli*, including cell state, media richness, and the specific nsAA used. These findings constrain what nsAAs can be used to study certain phenomena. In rich media, nsAA incorporation is

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

significantly delayed for both **1** and **2** (Supp Fig 5 G-H, 6 G-H). Furthermore, in standard S750 minimal media containing glucose and glutamate, **2** incorporation was delayed until the onset of stationary phase (Fig. 2D & Supp Fig. 5A, 6A). For the smaller, more hydrophilic nsAA **1**, equal incorporation was observed in exponential and stationary phases (Fig. 2E & Supp. Fig. 5B, 6B). We also demonstrated that that **1** could be incorporated during sporulation (Supp. Fig. 8). nsAA **4**, another bulky hydrophobic amino acid, also showed limited incorporation in exponential phase, but the pyrrolysine analog **5** was able to incorporate in exponential phase (Supp. Fig 5 C-D, 6 C-D). We hypothesized that nsAA uptake into the cell was limited and inhibited by high concentrations of standard amino acids. Pluronic F-68, a surfactant shown to be non-toxic and to help bulky molecules cross the cell membrane[40] increased **1** incorporation, as did removing all amino acids from the growth media (Supp Fig. 5 E-F, 6 E-F).

To verify that import into the cell was limiting, we performed LC-MS experiments measuring internal nsAA concentrations under different conditions. We noted that **1** was only present in very low relative concentrations in cells during exponential phase while **2** was present at higher concentrations inside cells during exponential phase (Figure 2E). These findings confirm that the available pool of nsAAs inside the cell is limited for incorporation of **1** during exponential phase, suggesting problems for other large, hydrophobic nsAAs. As a result, in the near-term small and polar nsAAs can be used for a broader range of studies in *B. subtilis*.

Photocrosslinking

A primary application of nsAAs is the incorporation of UV-photocrosslinkers to probe protein structure and assembly *in vivo*. Previous work has shown that the YukE protein, a homolog of the mycobacterial virulence factor EsxA, requires homodimerization for efficient translocation by the Early secretory antigen (Esx) pathway[41]. We used the photocrosslinking

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

capabilities of **2** to demonstrate short-range specific crosslinking between YukE monomers in *B. subtilis* cells. Consistent with its use in *E. coli*, **2**-based photocrosslinking is capable of distinguishing short-range interactions, as placing the photocrosslinking nsAA on the interface of the homodimer (W44UAG) resulted in high-efficiency crosslinking, while placing it on the external face (V21UAG) did not yield any detectable crosslinking (Figure 3). The crosslinking efficiency was noted to be extremely high, with nearly 100% crosslinking achieved with only 15 minutes of UV exposure. The remaining monomer may have resulted from translational readthrough, judging from the **2**$^{-}$ control samples. The exceptional crosslinking efficiency suggests a very tight interaction between YukE monomers inside cells and supports previous work that suggests *in vivo* dimerization is required for export of the YukE dimer[41].

Translational titration with nsAAs

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Based on our work with **2** for studying cellular uptake and photo-crosslinking, we recognized that this nsAA may allow fine-tuned control of protein expression at the translational level by requiring incorporation for full length protein formation. Titration of **2** across two orders of magnitude controlled the translation of mNeongreen reporter over a wide linear range and a

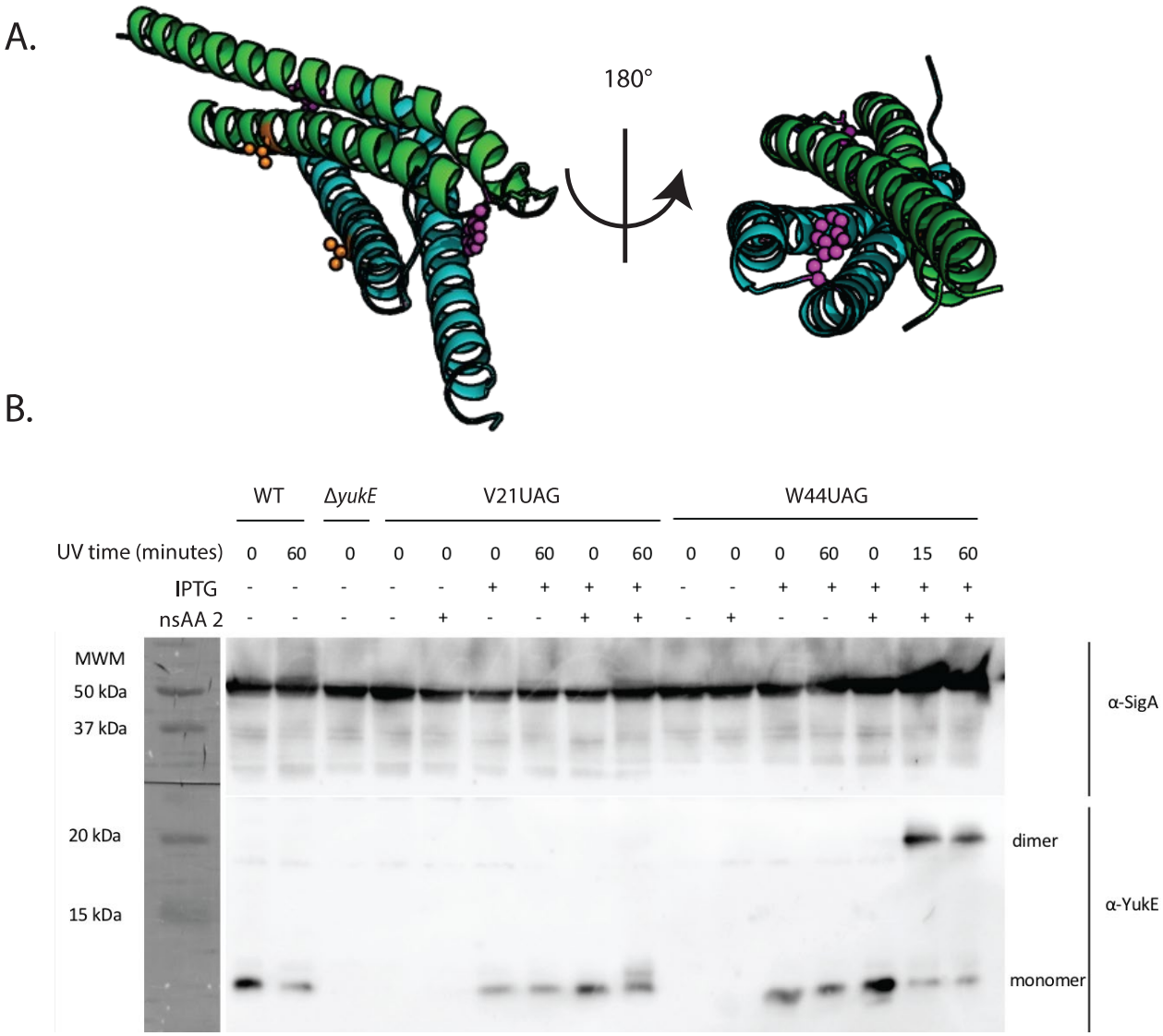

Figure 3: ***In vivo photocrosslinking*** **A).** Positions of V21 (orange space fill) and W44 (purple space fill) indicated. **B)** Immunoblotting results from in vivo crosslinking experiment. Cells producing YukE with **2** incorporated at either position 21 or 44 were treated with UV for indicated times. Top panel anti-SigA loading control immunoblot. Bottom panel anti-YukE immunoblot with monomeric and dimeric YukE positions noted.

dynamic range of approximately 50-fold (Fig. 4A and Supp Fig. 7A), roughly twice that previously observed for nsAAs in *B. subtilis*[22]. Notably, while most transcriptionally inducible

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

promoters show cooperative-like behavior and have steep induction curves with high Hill coefficients, nsAA-based translational control shows little to no cooperativity, with a Hill coefficient of similar value to 1 (Table 1).

Promoter	Inducer	Dynamic Range	Leakiness (absolute)	Hill coefficent	95% CI on Hill coefficent
pHyperspank	IPTG	13.7	7.1	1.22	1.01 to 1.48
pSpank	IPTG	4.1	3.6	2.28	1.67 to 3.30
pSpac	IPTG	4.3	4.5	2.33	1.15 to 6.45
pHyperspank*	IPTG	5.4	1.7	3.27	2.09 to 7.12
pXyl	Xylose	~10	11	Could not saturate	Could not saturate
UAG	2	49.9	1.3	1.08	0.97 to 1.2

Table 1. Quantitative characteristics of different promoters in *B. sub.* Characteristics were plotted by fitting sigmoidal curves to a titration, shown in figure 4A. Leakiness is measured as % of pHyperspank max expression. pXylose fit shown in figure 7A, and could not be saturated, even with 3 M xylose.

Because the transcriptionally inducible promoters act by controlling mRNA levels, and the UAG titration controls translation, it is possible to utilize both concurrently for '2-dimensional titration' (Fig. 4B and Supp Fig. 7B) allowing 570-fold induction when using **2** and 1300-fold protein induction when using **5** (Supp Fig 1G). Thus, these strategies are well-suited for experiments where leaky expression or limited dynamic range are barriers to obtaining insights.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

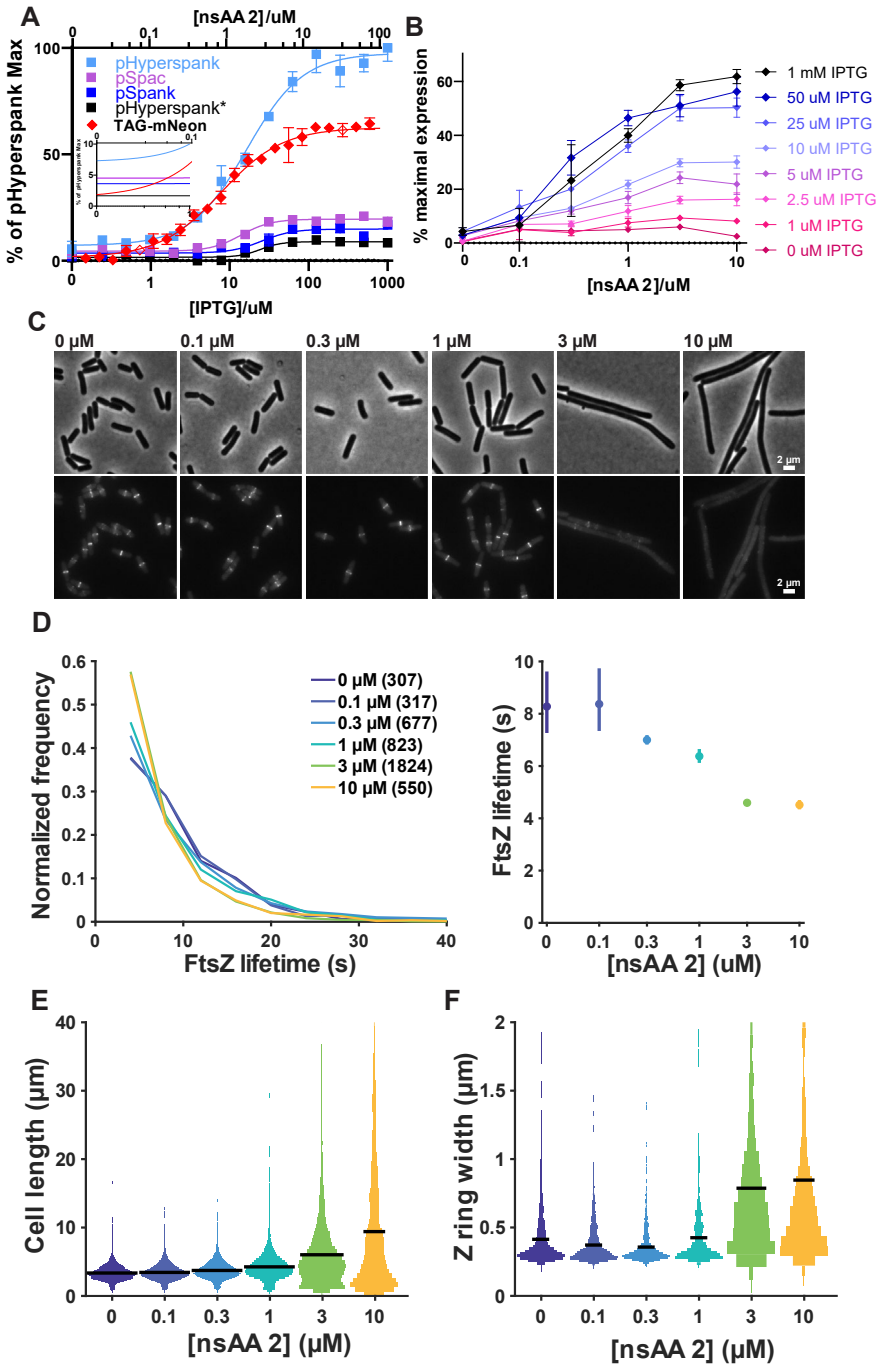

Figure 4: Protein titration with nsAAs A) Comparison between titration curves of standard IPTG promoters and nsAA titration with **2**. Hill coefficients and dynamic ranges of sigmoidal curves found in Table 1. Inset shows sigmoidal curve fits at low expression levelsAll values are averages of 3 replicates with standard deviation error bars. **B)** 2-dimensional titration of UAG-mNeongreen with both IPTG and **2**. All values are average of 3 replicates, errors are standard deviation. **C)** Titration of MciZ expression and its effects on cell division in vivo. Cell morphology (phase contrast, top) and Z ring morphology (epifluorescence images of mNeonGreen-FtsZ, bottom) in cells expressing MciZ under **2**-inducible control. The concentration of **2** added is indicated above each pair of panels. In each case, the UAG-MciZ construct was induced with 100 µM IPTG. **D)** Titration of FtsZ's filament dynamics by MciZ expression. Left: FtsZ single molecule lifetime distributions at various MciZ expression levels. FtsZ's lifetime reflects its dynamics, with a shorter lifetime corresponding to faster dynamics. Legend: **2** concentration (number of particles analyzed) Right: MciZ expression level vs mean FtsZ lifetime, computed from single exponential fits to the data at left. Error bars: 95% confidence interval. **E)** Increasing MciZ expression impairs Z ring formation. Frequency of Z rings along the length of cell was measured for N>400 cells in each condition. A control strain lacking the inducible MciZ system is included for comparison. **F)** Increasing MciZ expression causes aberrant Z ring assembly. Z ring width along the length of the cell was measured for N>400 Z rings in each condition as a proxy for successful Z ring condensation.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

One important phenomenon where improved ability to titrate protein concentration would facilitate biological insight is bacterial cell division. This process is orchestrated by the tubulin homolog FtsZ, which forms filaments that treadmill around the cell and assemble into the cytokinetic "Z ring"[26]. Recent work has demonstrated the importance of FtsZ dynamics in bacterial cell division[12,42]; modulating these dynamics is thus an attractive experimental target. A good candidate for this is expression of MciZ, a small protein that inhibits FtsZ polymerization, increases its treadmilling dynamics[27], and decreases its filament length[12]. Although the effects of varying levels of MciZ on FtsZ have been characterized *in vitro* and simulated *in silico*[43], *in vivo* experiments have been complicated by FtsZ's high sensitivity to low levels of MciZ[27], requiring a finely-controlled expression system. Experiments with standard promoters have been unable to examine this low-expression regime, as even low-level induction prevents Z ring assembly[27].

Using a pHyperspank-UAG-MciZ construct, we were able to finely control MciZ expression to modulate cell division in live *B. subtilis* cells. We observed that FtsZ filament dynamics & Z-ring assembly were altered with increasing MciZ expression (Fig. 4C, Supplementary Video 1). To quantitatively characterize the response of FtsZ filaments to MciZ titration, we assayed the single-molecule lifetimes of individual FtsZ filament subunits. The subunit lifetime reflects both the filament length and the rate of filament treadmilling[44]. Based on previous *in vitro* studies[27], we expected MciZ expression to increase FtsZ dynamics and decrease filament length, both of which should lead to a measurable decrease in FtsZ lifetime. Indeed, increasing MciZ expression caused FtsZ lifetimes to decrease in a dose-dependent manner. This demonstrated that we can precisely modulate FtsZ lifetime by controlling MciZ expression with nsAA translational titration (Fig. 4D).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Next, we investigated the effects of different MciZ levels on the process of cell division. As expected[27] , high levels of MciZ expression inhibited Z ring formation and cell division, as evidenced by increasing cell length (Fig. 4C-D). Notably, we were also able to find low-level induction conditions in which cells were still able to divide, conditions not previously accessible *in vivo* due to the leakiness and high cooperativity of standard inducible promoters[27]. We found that, although FtsZ subunit lifetimes are reduced even at low levels of MciZ induction (Fig. 4D), Z ring formation is only disrupted at higher MciZ inductions (Fig. 4E-F). The sharp switch in Z ring assembly seen at 3 μM **2** indicates that there is a critical point beyond which FtsZ filaments become too short and/or too dynamic to properly assemble a Z-ring, corresponding to a mean FtsZ lifetime between 4 and 6 s. These observations are consistent with a previous proposal that FtsZ filaments might need to be a minimal length to form into a Z ring[27]. These experiments use nsAA titration of MciZ expression to precisely modulate FtsZ filament dynamics *in vivo*, thereby interrogating properties underlying bacterial cytokinesis.

Discussion:

Here we demonstrate general nsAA incorporation tools in the gram-positive model organism *B. subtilis* and apply them to perform photocrosslinking and tight translational titration. We developed a general and extensible framework for tyrosine, pyrrolysine, and tryptophan-derived nsAAs in this gram-positive organism. These three families constitute 85% of demonstrated nsAAs incorporated in any organism[18], and our characterization of the limitations of nsAAs in this organism demonstrated extensive untargeted incorporation into native *B. subtilis* proteins, a phenomenon not previously observed in *E. coli*[23–25] or mammalian cells[45]. Additionally, we found that nsAA import can be limiting in rich media or with large,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

hydrophobic nsAAs, a phenomenon not previously observed in *B. subtilis* genetic code expansion.

However, despite the proteome-wide misincorporation that occurs in *B. subtilis*, we demonstrate both *in vivo* photocrosslinking and translational titration of protein expression with favorable induction characteristics. Photocrosslinking using nsAAs was extremely high-efficiency, nearing 100% crosslinking using the short-ranged crosslinker **2**. We used this technique to confirm a dimerization interface of the secreted virulence factor YukE, supporting previous studies that suggest dimerization happens inside of cells and is required for export and functionality[41]. Our use of translational titration adds nsAA-dependent protein expression to the toolkit of expression control systems available in *B. subtilis*. The nsAA-driven titration has very low cooperativity and leakiness, and the overall gain can be varied by titrating the associated rate of transcription & translation through inducible promoters or varying the ribosomal binding site.

Combining transcriptional and translational titration allowed us to modulate FtsZ's dynamics and therefore *B. subtilis* cell division. Specifically, we were able to access a previously unseen regime of FtsZ dynamics *in vivo*, and identify a critical transition in FtsZ function as we perturb these dynamics. Our data suggest that there is a minimum lifetime of FtsZ in the filament of 5 ± 1 seconds, below which cell division is impossible because FtsZ filaments are unable to form a functional Z ring. These results demonstrate that enabling genetic code expansion tools in *B. subtilis* will not only allow the use of nsAA-based techniques developed in *E. coli* but also serve as the foundation for developing new technologies taking advantage of *B. subtilis's* status as a gram-positive model, industrial and GRAS organism.

We discovered that AARSs that work in *E. coli* seem to usually function in *B. subtilis* with similar activity and range of substrates. Because of this portability of *E. coli* genetic code

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

expansion techniques to *B. subtilis*, future developments to *E. coli* nsAA tools will likely directly transfer to *B. subtilis*, enabling new applications and uses. However, future work in *B. subtilis* must be cautious of the innate differences between the two systems. First, there appears to be innate differences in the translational termination machinery that allows efficient nsAA incorporation via amber suppression with minimal engineering of the translational machinery, but also allows incorporation of the nsAA throughout the proteome. Though *B. subtilis* is known to relax translational termination during stress[46], further investigation of this phenomenon may uncover deeper differences between gram-positive and gram-negative protein biology. Second, the amino acid uptake machinery is quite different between these two organisms and will be subject to different limitations. Encouraging uptake of large, hydrophobic nsAAs could be accomplished by heterologous expression of a panel of appropriate membrane transporters[47], and minimal protein engineering may be necessary to maximize uptake[48]. Third, *B. subtilis* genomic insertion and modification is much simpler than in many bacterial species and will lend itself well to interrogating native genes with nsAAs. These findings will aid the expansion of the genetic code in organisms aside from *B. subtilis*.

Genetic code expansion in *B.* subtilis provides multiple potential applications in both experimental and applied spaces. Highly efficient and controlled nsAA incorporation will enable new research in *B. subtilis* allowing fluorescence, titration, crosslinking, and photocontrol[49] of different proteins. Furthermore, genetic code expansion will allow for biocontainment via synthetic auxotrophy[22,50], which will enhance use of *B. subitlis* as an engineered probiotic or therapeutic in humans, livestock, and plants[8,51]. Advances in metabolic engineering enabling *in-vivo* synthesis of nsAAs[52] will improve biocatalysis processes[53], enhancing the already capable *B. subtilis* chassis for further metabolic engineering[54] and protein production[11].

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Methods

Reagents:

Antibiotics and nsAAs were purchased from Sigma. nsAA stock solutions were prepared in water with minimal base, e.g., 0.3 M KOH to prepare 0.2 M Bpa stock solution, except for Cou stock, which was prepared in 50% DMSO with minimal base. All stock concentrations of nsAA were between 100–200mM.

B. subtilis media

$S7_{50}$ media was made in 500 mL, by mixing 50 mL 10X $S7_{50}$ salts (recipe below), 5 mL 100x $S7_{50}$ metals (recipe below), 5 mL 1 Molar glutamate and 10 mL 50% (w/w) glucose together. Made up to 500 mL with ddH_2O. Filter sterilized (not autoclaved) and distributed into 50 mL aliquots. Glycerol, sorbitol, or fumarate was substituted 1:1 for glucose for different carbon sources. Other amino acids can be substituted 1:1 for glutamate or use 20% (w/w) ammonium sulfate (final concentration 0.2% + additional 0.1% from 10x salts) for media lacking amino acids.

10x $S7_{50}$ salts were made in 1-liter aliquots. 104.7 g MOPS (free acid), 13.2 g ammonium sulfate $(NH_4)_2SO_4$, 6.8 grams potassium phosphate monobasic KH_2PO_4 were added and buffered to pH 7 with 50% KOH, then made up to 1 L with ddH_2O. Media was filter-sterilized, covered with foil, and store at 4 degrees C. If yellowed, the solution is no longer usable.

100x $S7_{50}$ metals have final concentrations of: 0.2 M $MgCl_2$, 70 mM $CaCl_2$, 5 mM $MnCl_2$, 0.1 mM $ZnCl_2$, 100 microgram/mL thiamine-HCl, 2 mM HCL, 0.5 mM $FeCl_3$. Added iron last to prevent precipitation. Media was sterilize and store foil-wrapped at 4 degrees C.

MC media for transformations was made at a 10x stock. The 10x stock has final concentrations of 1 M potassium phosphate pH 7, 30 mM sodium citrate, 20% (w/w) glucose, 220 mg/mL

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Ferric ammonium citrate, 1% casein hydrolysate, 2% potassium glutamate. Store in aliquots at -20 degrees, and make 1x media during the competence protocol, supplementing the 1x MC with 3-30 mM $MgSO_4$.

Strain construction

All *B. subtilis* strains were derived from the prototrophic strain PY79[55]. Strains for each figure, strain crosses, strain construction, and primers can be found in Supplemental excel file 1. All cloning was done via a combination of Gibson isothermal assembly and overlap-extension PCR. Primers were ordered from IDT with Q5 Tms (calculated via the NEB Tm calculator) of 70-72 degrees. Overlaps between parts were 20-35 bases, attempting to maintain overlap Tm of 55-65 degrees as calculated by the Geneious Tm calculator. PCRs were carried out in 25 uL, using 2x Q5 master mix with the maximum allowed temperature at 25 cycles and with 1 min/kB extension time. PCR fragments were gel-verified and either gel-extracted or PCR-purified using appropriate Qiagen kits. Isothermal reactions were done in 20 uL final volume homemade Gibson mix made according to the original 2011 Gibson paper[56], at 50 degrees C for 30-45 minutes. The entire volume of the Gibson was transformed directly into *B. subtilis* by the *B. subtilis* transformation protocol described below. In the case of transformation failure, overlap assembly PCR was carried out in a pairwise manner to assemble difficult-to-Gibson regions of DNA. 5-10 uL of each purified PCR product was added to overlap extension PCRs, which were run for 5 cycles, followed by the addition of primers to amplify the full-length piece and 25 more cycles. All constructs were confirmed by sequencing.

***B. subtilis* transformation protocol**

For transformation of either genomic or linearized DNA, LB plates were streaked with parent strains the day before. Freshly grown (not more than a day old) colonies were inoculated into 1

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

mL MC in a large glass test tube, with 3-30 uL of 1 M $MgSO_4$ supplemented. Strains were grown in MC at 37 degrees C in a roller drum for 4 hours, at which point culture should be visibly turbid. Then, 200 uL of culture was added to a standard 13 mL culture tube with transformation DNA. An entire 20 uL Gibson reaction or 2 uL of *B. subtilis* genomic DNA was added to each tube. Cultures were returned to the 37-degree roller for 2 more hours, and then the entire volume was plated on selective LB plates. Single colonies were picked and verified by sequencing of unpurified colony PCR. To do colony PCR with *B. subtilis* colonies, a colony was suspended in 50 uL TE + 10% CHELEX beads, then vortexed for 10 minutes, boiled for 30 minutes, vortexed for 10 minutes, and spun down. 1 uL was used as PCR template without getting any CHELEX beads into the PCR reaction.

***B. subtilis* genomic DNA preparation**

For Bacillus genomic DNA prep, used as a PCR template or to transform into other strains, 1-2 mL LB was inoculated from a fresh colony plated the night before. Cells were grown until dense, preferably not overgrown. Cells were pelleted at max speed and the supernatant aspirated. Cells were resuspended in 500 uL lysis buffer, final concentration 20 mM Tris-HCl pH 7.5, 50 mM EDTA, 10 mM NaCl, and 50 uL of freshly made 20 mg/mL egg white lysozyme in lysis buffer was added. The mixture was incubated at 37 degrees for 15-45 minutes, using the longer time if cells were overgrown. Then 60 uL 10% (w/w) sarkosyl (N-lauroylsarkosine) in ddH_2O was added and the mixture vortexed. The entire solution was transferred to a phase lock tube, and 600 uL phenol-chloroform was added. The tube was vortexed vigorously for 15-20 seconds until frothy, then spun at max speed for 5 minutes. The aqueous phase was removed to a fresh 1.5 mL tube and 1/10 volume (60 uL) 3 M sodium acetate added, then vortexed. 2 volumes (or top of tube) of 100% Ethanol was added and the tube inverted until DNA had visibly precipitated. Then

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the tube was spun at max speed for 1 min. The supernatant was aspirated, 150 uL 70% ethanol 30% ddH2O was added, and the tube quickly vortexed. The tube was then spun for 1 minuite at max speed. The supernatant was removed, and tubes were left open to dry on a bench for 5-15 minutes. DNA was resuspended in 350 uL ddH_2O and stored at -20 ° C.

nsAA incorporation

For nsAA incorporation, fresh colonies were picked from LB plates to make a starter in the same media as the experimental media, most often S750 minimal media. These cultures were grown to exponential phase, preferably between OD 0.1-0.4, though ODs as high as 0.7 were used without issue. Experimental cultures already containing IPTG and nsAA were seeded from starter cultures at an OD of 0.002. Experimental cultures were grown overnight either in a plate reader or culture tubes at 37 degrees, either in a shaker at 250 rpm or roller drum. For endpoint experiments, cultures were diluted 1:1 with PBS and then read in a Biotek spectrophotometric plate reader, unless the fluorescent couAA was being incorporated, in which case cells were pelleted at 5K rcf and washed 3x with PBS before being read in the plate reader. Fluorescence background signals were corrected for autofluorescence by subtracting the average fluorescence of PY79 cells grown in parallel in media with no additives.

Click-enrichment of proteins containing nsAAs

For protein expression of nsAA-containing proteins, 10mL of S750 culture containing 1 mM IPTG & 100uM pAzF was seeded at OD 0.002 with appropriate strain from an exponential-phase s750 culture (OD 0.1-0.5). Cultures were grown in a shaking incubator overnight at 37° C. Cells were pelleted at 5K RCF for 30 min and frozen at -80. Pellets were lysed with 40x 3-second sonication pulses in ice using a a QSonica Q125 sonicator using the urea lysis buffer provided by the ThermoFisher Click-iT™ Protein Enrichment Kit (Cat # C10416). Follow-up

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

purification was performed according to manufacturer's instructions. The click-enriched fraction was trypsin digested off of the beads with 2 ug/mL trypsin in 100 mM Tris, 2 mM CaCl2 buffer & 10% acetonitrile. Dialysis buffer exchange was used to convert the buffer to HPLC solvent A, after which the mass-spec protocol detailed below was followed.

Protein purification and mass spec

For protein expression of nsAA-containing proteins, 25mL of S750 culture containing 1 mM IPTG & nsAA (1 mM for boc-K & 5OHW, 100 uM for bipA, bpA and pAzF) was seeded at OD 0.002 with appropriate strain from an exponential-phase s750 culture (OD 0.1-0.5). Cultures were grown in a shaking incubator overnight at 37° C. Cells were pelleted at 5K RCF for 30 min and frozen at -80. Cells were lysed using the EMD Millipore Lysonase BugBuster kit (Cat. # 71370-3), following the manufacturer's instructions for gram-positive bacteria. For His-tag purification, Thermo Scientific HisPur cobalt resin (Cat. # 89964) was used. After washing with equilibration buffer, the lysate was bound to resin with 45-minute binding at room temperature followed by one wash step and 3 elution steps, using the following buffers: Equilibration buffer: 20 mM Tris-HCl, Ph 8.3, 0.5 M NaCl, 5 mM imidazole. Wash buffer: 20 mM Tris-HCl, Ph 8.3, 0.5 M NaCl, 20 mM imidazole. Elution buffer: 20 mM Tris-HCl, Ph 8.3, 0.5 M NaCl, 200 mM imidazole.

Elutions were combined, and 18 uL was run out on an Invitrogen 4-12% Bis-Tris NuPAGE gel (cat # NP0322PK2) following manufacturer's instructions. Purified tagged mNeongreen was seen at the correct weight and excised in approximately 5x20 mm pieces.

Excised gel bands were cut into approximately 1 mm^3 pieces. Gel pieces were then subjected to a modified in-gel trypsin digestion procedure. Gel pieces were washed and dehydrated with acetonitrile for 10 min, followed by removal of acetonitrile. Pieces were then completely dried in

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

a speed-vac. Rehydration of the gel pieces was with 50 mM ammonium bicarbonate solution containing 12.5 ng/µl modified sequencing-grade trypsin (Promega, 182 Madison, WI) at 4ºC. After 45 min., the excess trypsin solution was removed and replaced with 50 mM ammonium bicarbonate solution to just cover the gel pieces. Samples were then placed in a 37ºC room overnight. Peptides were later extracted by removing the ammonium bicarbonate solution, followed by one wash with a solution containing 50% acetonitrile and 1% formic acid. The extracts were then dried in a speed-vac (~1 hr). The samples were then stored at 4ºC until analysis. On the day of analysis, the samples were reconstituted in 5 - 10 µl of HPLC solvent A (2.5% acetonitrile, 0.1% formic acid). A nano-scale reverse-phase HPLC capillary column was created by packing 2.6 µm C18 spherical silica beads into a fused silica capillary (100 µm inner 191 diameter x ~30 cm length) with a flame-drawn tip. After equilibrating the column, each sample was loaded via a Famos auto sampler (LC Packings, San Francisco CA) onto the column. A gradient was formed, and peptides were eluted with increasing concentrations of solvent B (97.5% acetonitrile, 0.1% formic acid). As peptides eluted, they were subjected to electrospray 195 ionization and then entered into an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (Thermo Fisher Scientific, Waltham, MA). Peptides were detected, isolated, and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide. Peptide sequences (and hence protein identity) were determined by matching protein databases with the acquired fragmentation pattern by the software program, Sequest (Thermo Fisher Scientific, Waltham, MA). All databases include a reversed version of all the sequences, and the data were filtered to between a one and two percent peptide false discovery rate

LCMS nsAA import quantification

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Protocol adapted from[57]. To assay import of nsAAs into cells, serial dilutions of WT Py79 *B. subtilis* were seeded in 1 mL S750 media tubes and incubated at 37 ° C in a tube roller overnight to obtain turbid but not stationary cultures, OD 0.4-0.8. In the morning, 200 uL of this culture was added to 60 mL experimental cultures containing nsAA. After growth in a 37° shaker until OD ~0.2, exact OD's were recorded and 50 mL of culture taken, the rest left in the 37 ° shaker overnight. 50 mL of culture was pelleted at 4 ° C for 10 min at 5.25K RCF, and the supernatant was discarded. The pellet was washed 4 times with 1 mL ice-cold s750, with rapid but thorough resuspension and 2.5-minute spins at 14 k RCF. Cell pellets were then frozen. After the overnight, the saturated culture was diluted 10:1 for OD measurement, and then 5 mLs of saturated culture was pelleted at 4 ° C for 10 min at 5.25K, followed by 4 washes with 1 mL ice-cold s750, with rapid but thorough resuspension and 2.5-minute spins at 14 k RCF and brief freezing. To thoroughly lyse cells, pellets were resuspended in 400 uL 40:60 sterile methanol:water and transferred to screw-top tubes. 300 mg of 200 µm acid-washed glass beads were added to each tube, and the tubes were beaten in a bead beater for 10 minutes at 4 ° C, in 1-minute increments with a 5-minute gap between each bead beating step. The tubes were inverted and a small hole poked in the bottom with a syringe. Then the screw-top tubes were placed into 7 mL tubes, and the lysate was collected by a 5-minute spin at 4000 rcf. An additional 400 uL 40:60 methanol was added to the tubes, and the spin was repeated to ensure all lysate was collected. The collected lysate and wash were spun at 16 k rcf at 4 ° C for 30 hours, and 750 uL was transferred to fresh tubes, which were then centrifuged at 4 ° C for 2 hours, and 600 uL supernatant extracted to be used in LCMS experiments.

nsAA titration microscopy

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Two days prior to imaging, cells were struck from a glycerol stock at -80˚C onto an LB agar plate, and grown overnight at 37˚C. The following day, single colonies were inoculated from the LB plate into 1 mL $S7_{50}$ media containing 100 μM IPTG and the appropriate concentration of pAzF. A 1:10 dilution series was prepared out to 1:1,000, and all of these cultures were grown overnight at room temperature on a rolling drum. The next day, the culture in the dilution series that was nearest to mid-exponential phase growth was diluted 1:10 and grown in 1 mL $S7_{50}$+IPTG+pAzF at 37˚C on a rolling drum. For single molecule lifetime imaging, 20 pM JF549-HaloTag Ligand was additionally added to the media. Once the density of this culture reached mid-exponential phase, cells were ready to image. Cultures were pelleted by centrifugation for 2 minutes at 7,000 RPM. Approximately 900 μL of the supernatant media was removed; this volume was adjusted depending on the exact OD of cells in the culture. Cells were resuspended in the remaining media.

For imaging, 2 μL of these concentrated cells were pipetted on to a glass coverslip, and immobilized underneath a small pad of $S7_{50}$ agarose. Agarose pads were molded using 1.5 cm x 1.5 cm x 1 mm plastic frames, which were placed on a pane of glass that had been cleaned with detergent and 70% ethanol. Molten $S7_{50}$ + 2% agarose was poured into the frames, and a second glass plane was placed on top to complete the mold. Pads solidified at room temperature for at least 15 minutes, and excess agarose was cut away from the outside of the frame prior to use. Microscopy was performed on a Nikon Ti-E microscope equipped with TIRF optics, a perfect focus system (Nikon), and a MLC4008 laser launch (Agilent). The objective used was a Nikon CFI Plan Apochromat DM Lambda 100X oil immersion objective with 1.45 NA and a Ph3 phase contrast ring. The microscope was enclosed in a chamber heated to 37˚C. Images were collected using NIS-Elements software (Nikon).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Imaging of FtsZ filaments and Z rings and quantification of Z ring morphology was performed using strain bGS543, which contains the napARS system under the constitutive pVeg promoter and UAG-MciZ under the IPTG-inducible pHyperspank promoter. To visualize FtsZ, the strain contains a duplicate copy of the FtsAZ operon, including its promoter, in which FtsZ has been tagged with mNeonGreen on its N terminus. A 488 nm laser was used for excitation and a C-NSTORM QUAD filter (Nikon) and a ET525/50m filter (Chroma) were used for emission. Images were collected on an ORCA-Flash4.0 V2 sCMOS camera (Hamamatsu). TIRF time lapses were taken with 1 s exposures for 4 minutes total; after each time lapse, an epifluorescence image was taken to visualize Z ring morphology and a phase-contrast image was taken to visualize cells. Z ring morphology was analyzed as described previously[44].

Single molecule lifetime imaging and cell length measurements were performed using strain bGS545, which contains the napARS system under the constitutive pVeg promoter and UAG-MciZ under the IPTG-inducible pHyperspank promoter. To visualize FtsZ, the strain contains a duplicate copy of the FtsAZ operon, including its promoter, in which FtsZ has been tagged with HaloTag on its N terminus. The HaloTag-FtsZ construct is labeled with JF549-HaloTag Ligand for fluorescence microscopy. TIRF time lapses were taken using 500 ms exposures for 4 minutes total; after each time lapse, a phase-contrast image was taken to visualize cells. A 561 nm laser was used for excitation and a C-NSTORM QUAD filter (Nikon) and a ET600/50m filter (Chroma) were used for emission. Images were collected on an iXon Ultra 897 EMCCD camera (Andor). Single molecule lifetimes were measured as described previously[44]. To measure cell lengths, cell images were segmented using DeepCell and their dimensions were measured using Morphometrics. Cells that extended out of the field of view were excluded from the analysis because their lengths could not be accurately measured;

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

however, at the highest levels of MciZ expression, some cells were too long to be captured in a single field of view. Thus, these cell length measurements underestimate the true cell length for samples induced with 3 and 10 μM pAzF.

Incorporation of unnatural amino acids during sporulation

Sporulation was induced by resuspension as described[58] with the following modifications. Where appropriate, at the time of resuspension, the inducer, isopropyl β- d-1-thiogalactopyranoside (IPTG), and the unnatural amino acid p-azido-phenylalanine (pAzF) were added to the test cultures at 1 mM and 100 μM final concentration, respectively.

Samples for microscopy were collected and prepared as previously described[59]. Images were collected using an Axio M1 Imager (Zeiss). For strains producing GFP, the GFP channel excitation was 30 milliseconds (Ex band pass 470±40 and Em - 525±50). Images were analyzed and prepared using Zen 2.0 software (Zeiss) and ImageJ.

Detection of protein-protein interaction by crosslinking in vivo

Samples were prepared as described in[41] with the following modifications. Strains were grown in 30 mL LB medium at 37° C with shaking at 210 rpm to an OD 600 nm of 1.0-1.5. The cultures were split into separate flasks into 7 ml aliquots. As appropriate, cultures were either supplemented with inducer (IPTG) or with unnatural amino acid, pAzF, to final 1 mM and 100 μM concentration, respectively, or both or neither.

Growth at 37° C with shaking was continued for 13 hours. At this time, OD measurement of cultures was taken again (generally around OD600 3.4 – 4.9), and 1 mL aliquots were pelleted by centrifugation and resuspended in 1 mL of sterile 1x PBS buffer. Cell resuspensions were transferred to a 12-well polycarbonate plate (Corning, catalog # 353043) and exposed to UV light at 365 nm in Spectrolinker XL-1000 UV Crosslinker (Spectronics Corporation). The

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

distance between the bottom of the well with cell resuspension and the UV source was ~4 cm. Cell suspensions were exposed to the UV source for either 15 minutes or a total of 60 minutes, which was divided in two 30-minute intervals. 12-well plate was on ice, and thermal probe readings confirmed that the cell suspension temperature never exceeded 35.1° C.

After UV treatments, cell suspensions equivalent to 1.5 OD units were pelleted by centrifugation, Cell pellets were resuspended in 100 µL of Tris-HCl pH 6.8 lysis buffer, containing 10 mM EDTA, 100 µg/mL chicken egg white lysozyme, 10 µg/mL DNaseI, 0.25 mM phenylmethyl sulfonyl fluoride (PMSF) and 1 µM protease cocktail inhibitor E-64. Lysis was continued for 20 minutes at 37° C, and then 50 µL SDS-PAGE loading buffer (without reductant) was added. The lysed samples were mixed vigorously and then heated for 10 minutes at 65° C.

For crosslinking detection, 7.5 µL aliquots of all control and test samples were separated by 17% SDS-PAGE. Then proteins were transferred from the gel to polyvinylidene fluoride (PVDF) membrane for 52 minutes at 15 V in Trans Blot Semi Dry Transfer Cell (BioRad). The membrane was cut horizontally along the 25 kDa marker band to allow for probing with two different antibodies. After blocking, the lower part of the membrane was incubated with custom rabbit polyclonal anti-YukE antibodies used in a dilution of 1:5,000. The upper membrane portion (used as loading control) was probed with anti-SigmaA antibodies diluted 1:10,000[60]. The primary antibodies were tagged successively with goat anti-rabbit antibodies conjugated with horseradish peroxidase (HRP) (Abcam, catalog # ab6721). The bound antibodies were detected using luminol containing reagent Clarity Western ECL Substrate (BioRad) and visualized using ChemiDoc imager (BioRad).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgements

We thank Dr. Valerie Pivorunas for construction of strain bBB954. We thank Michaela Jones, Kamesh Narasimhan, Gabriel Filsinger for comments on this manuscript. We also thank S. Wilson, M. Holmes, W. Mallard, S. Hurlimann, M. Dion, M. Baas-Thomas, G. Filsinger, A. Rudolph, P. Smith, and Drs. G. Chao, A. Debnath, M. Shubert, A. Chatterjee, A. Bisson, A. Florez, J. Marchand, A. Mijalis, K. Narasimhan, A. Nyerges, N. Ostrov, D. Thompson, T. Wannier for helpful discussions and advice. The Taplin Mass Spectrometry Facility and the Analytical Chemistry Core at Harvard Medical School were also essential to this work. This work was supported by US Department of Energy Grant DE-FG02-02ER63445 (to GMC), NSF grant MCB-2027074 (To EG & AMK), NIH R01-GM121865 (to BMB), the Landry Cancer Biology Research Fellowship (to DS) and NSF GRFP DGE1144152 (to GRS).

Author Contributions

D.S, G.R.S, A.K, and A.J were involved in strain construction. D.S performed non-standard amino acid incorporation and growth experiments with help from E.K, A.K, and A.J. D.S and E. K. purified proteins and performed biochemical experiments. D.S and J.R performed cell import assays. G.R.S carried out microscopy experiments and analyzed data. K.G and B.B performed photocrosslinking assays and analysis. B.B, A.K, E.G & G.C. supervised the research. D.S, A.K, E.K, and E.G wrote the manuscript with feedback from all other authors.

Conflict of Interest Disclosure

G.M.C. has related financial interests in ReadCoor, EnEvolv, 64-X, and GRO Biosciences. For a complete list of G.M.C.'s financial interests, please visit arep.med.harvard.edu/gmc/tech.html. No other authors have any conflict of interests.

Data Availability

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Mahaffee, W. F. & Kloepper, J. W. Temporal changes in the bacterial communities of soil, rhizosphere, and endorhiza associated with field-grown cucumber (Cucumis sativus L.). *Microbial Ecology* **34**, 210–223 (1997).
2. Mahaffee, W. F. & Kloepper, J. W. Bacterial communities of the rhizosphere and endorhiza associated with field-grown cucumber plants inoculated with a plant growth-promoting rhizobacterium or its genetically modified derivative. *Canadian Journal of Microbiology* **43**, 344–353 (1997).
3. Earl, A. M., Losick, R. & Kolter, R. Ecology and genomics of Bacillus subtilis. *Trends in Microbiology* vol. 16 269–275 (2008).
4. Losick, R., Youngman, P. & Piggot, P. J. Genetics of Endospore Formation in Bacillus Subtilis. *Annual Review of Genetics* **20**, 625–669 (1986).
5. Mckenney, P. T., Driks, A. & Eichenberger, P. The Bacillus subtilis endospore: Assembly and functions of the multilayered coat. *Nature Reviews Microbiology* vol. 11 33–44 (2013).
6. Kearns, D. B., Chu, F., Branda, S. S., Kolter, R. & Losick, R. A master regulator for biofilm formation by Bacillus subtilis. *Molecular Microbiology* **55**, 739–749 (2005).
7. Branda, S. S., González-Pastor, J. E., Ben-Yehuda, S., Losick, R. & Kolter, R. Fruiting body formation by Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11621–11626 (2001).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

8. Bai, Y., D'Aoust, F., Smith, D. L. & Driscoll, B. T. Isolation of plant-growth-promoting Bacillus strains from soybean root nodules. *Canadian Journal of Microbiology* **48**, 230–238 (2002).

9. Sivasakthi, S., Usharani, G. & Saranraj, P. Biocontrol potentiality of plant growth promoting bacteria (PGPR) - Pseudomonas fluorescens and Bacillus subtilis: a review. *African Journal of Agricultural Research* **9**, 1265–1277 (2014).

10. Tam, N. K. M. *et al.* The intestinal life cycle of Bacillus subtilis and close relatives. *J. Bacteriol.* **188**, 2692–2700 (2006).

11. Westers, L., Westers, H. & Quax, W. J. Bacillus subtilis as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1694**, 299–310 (2004).

12. Bisson-Filho, A. W. *et al.* Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. *Science* **355**, 739–743 (2017).

13. Peters, J. M. *et al.* A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **165**, 1493 (2016).

14. Castillo-Hair, S. M., Baerman, E. A., Fujita, M., Igoshin, O. A. & Tabor, J. J. Optogenetic control of Bacillus subtilis gene expression. *Nature Communications* **10**, 3099 (2019).

15. Hsu, Y.-P. *et al.* Fluorogenic D-amino acids enable real-time monitoring of peptidoglycan biosynthesis and high-throughput transpeptidation assays. *Nat Chem* **11**, 335–341 (2019).

16. Chin, J. W. Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annual Review of Biochemistry* **83**, 379–408 (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

17. Ledbetter, M. P. & Romesberg, F. E. Editorial overview: Expanding the genetic alphabet and code. *Current Opinion in Chemical Biology* **46**, A1–A2 (2018).

18. Dumas, A., Lercher, L., Spicer, C. D. & Davis, B. G. Designing logical codon reassignment – Expanding the chemistry in biology †Electronic supplementary information (ESI) available: A comprehensive table of the UAAs incorporated to date (also summarized in), their reported/potential uses, and the required mutations in the aaRS to allow their uses. See DOI: 10.1039/c4sc01534g Click here for additional data file. *Chem Sci* **6**, 50–69 (2015).

19. Luo, X. *et al.* Recombinant thiopeptides containing noncanonical amino acids. *Proc Natl Acad Sci U S A* **113**, 3615–3620 (2016).

20. Bartholomae, M. *et al.* Expanding the Genetic Code of Lactococcus lactis and Escherichia coli to Incorporate Non-canonical Amino Acids for Production of Modified Lantibiotics. *Front. Microbiol.* **9**, (2018).

21. Scheidler, C. M., Vrabel, M. & Schneider, S. Genetic Code Expansion, Protein Expression, and Protein Functionalization in Bacillus subtilis. *ACS Synth. Biol.* (2020) doi:10.1021/acssynbio.9b00458.

22. Tian, R. *et al.* Titrating bacterial growth and chemical biosynthesis for efficient N -acetylglucosamine and N -acetylneuraminic acid bioproduction. *Nature Communications* **11**, 1–13 (2020).

23. Chemla, Y., Ozer, E., Algov, I. & Alfonta, L. Context effects of genetic code expansion by stop codon suppression. *Current Opinion in Chemical Biology* **46**, 146–155 (2018).

24. Kipper, K. *et al.* Application of Noncanonical Amino Acids for Protein Labeling in a Genomically Recoded Escherichia coli. *ACS Synth. Biol.* **6**, 233–255 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

25. Lajoie, M. J. *et al.* Genomically Recoded Organisms Expand Biological Functions. *Science* **342**, 357–360 (2013).

26. Du, S. & Lutkenhaus, J. At the Heart of Bacterial Cytokinesis: The Z Ring. *Trends in Microbiology* **27**, 781–791 (2019).

27. Bisson-Filho, A. W. *et al.* FtsZ filament capping by MciZ, a developmental regulator of bacterial division. *PNAS* **112**, E2130–E2138 (2015).

28. Hughes, R. A. & Ellington, A. D. Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA. *Nucleic Acids Res* **38**, 6813–6830 (2010).

29. Ai, H., Shen, W., Sagi, A., Chen, P. R. & Schultz, P. G. Probing Protein–Protein Interactions with a Genetically Encoded Photo-crosslinking Amino Acid. *ChemBioChem* **12**, 1854–1857 (2011).

30. Beránek, V., Willis, J. C. W. & Chin, J. W. An Evolved Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetase/tRNA Pair Is Highly Active and Orthogonal in Mammalian Cells. *Biochemistry* **58**, 387–390 (2019).

31. Jiang, R. & Krzycki, J. A. PylSn and the Homologous N-terminal Domain of Pyrrolysyl-tRNA Synthetase Bind the tRNA That Is Essential for the Genetic Encoding of Pyrrolysine. *J Biol Chem* **287**, 32738–32746 (2012).

32. Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H. & Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia coli. *Biochemistry* **52**, 1828–1837 (2013).

33. Kunjapur, A. M. *et al.* Engineering posttranslational proofreading to discriminate nonstandard amino acids. *PNAS* **115**, 619–624 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

34. Mohler, K. *et al.* MS-READ: Quantitative Measurement of Amino Acid Incorporation. *Biochim Biophys Acta* **1861**, 3081–3088 (2017).

35. Wolf, C. *et al.* Elucidation of the presence and location of t-Boc protecting groups in amines and dipeptides using on-column H/D exchange HPLC/ESI/MS. *Journal of the American Society for Mass Spectrometry* **16**, 553–564 (2005).

36. Lee, K. J., Kang, D. & Park, H.-S. Site-Specific Labeling of Proteins Using Unnatural Amino Acids. *Mol Cells* **42**, 386–396 (2019).

37. Charbon, G. *et al.* Subcellular Protein Localization by Using a Genetically Encoded Fluorescent Amino Acid. *Chembiochem* **12**, 1818–1821 (2011).

38. Hankore, E. D. *et al.* Genetic Incorporation of Noncanonical Amino Acids Using Two Mutually Orthogonal Quadruplet Codons. *ACS Synth. Biol.* **8**, 1168–1174 (2019).

39. DeBenedictis, E. In review at. *Nature Communications* (2021).

40. Berlatzky, I. A., Rouvinski, A. & Ben-Yehuda, S. Spatial organization of a replicating bacterial chromosome. *Proc Natl Acad Sci U S A* **105**, 14136–14140 (2008).

41. Sysoeva, T. A., Zepeda-Rivera, M. A., Huppert, L. A. & Burton, B. M. Dimer recognition and secretion by the ESX secretion system in Bacillus subtilis. *PNAS* **111**, 7653–7658 (2014).

42. Yang, X. *et al.* GTPase activity–coupled treadmilling of the bacterial tubulin FtsZ organizes septal cell wall synthesis. *Science* **355**, 744–747 (2017).

43. Corbin, L. & Erickson, H. P. Modelling FtsZ nucleation, hydrolysis and treadmilling activity with Monte Carlo methods. *bioRxiv* 2020.02.25.965095 (2020) doi:10.1101/2020.02.25.965095.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

44. Squyres, G. R. *et al.* Dynamics of bacterial cell division: Z ring condensation is essential for cytokinesis. *bioRxiv* 2020.06.30.180737 (2020) doi:10.1101/2020.06.30.180737.

45. Uttamapinant, C. *et al.* Genetic Code Expansion Enables Live-Cell and Super-Resolution Imaging of Site-Specifically Labeled Cellular Proteins. *J. Am. Chem. Soc.* **137**, 4602–4605 (2015).

46. Meyerovich, M., Mamou, G. & Ben-Yehuda, S. Visualizing high error levels during gene expression in living bacterial cells. *PNAS* **107**, 11543–11548 (2010).

47. Zhu, B. & Stülke, J. SubtiWiki in 2018: from genes and proteins to functional network annotation of the model organism Bacillus subtilis. *Nucleic Acids Res* **46**, D743–D748 (2018).

48. Ko, W., Kumar, R., Kim, S. & Lee, H. S. Construction of Bacterial Cells with an Active Transport System for Unnatural Amino Acids. *ACS Synth. Biol.* **8**, 1195–1203 (2019).

49. Courtney, T. & Deiters, A. Recent advances in the optical control of protein function through genetic code expansion. *Current Opinion in Chemical Biology* **46**, 99–107 (2018).

50. Mandell, D. J. *et al.* Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55–60 (2015).

51. Cutting, S. M. Bacillus probiotics. *Food Microbiology* vol. 28 214–220 (2011).

52. Völler, J.-S. & Budisa, N. Coupling genetic code expansion and metabolic engineering for synthetic cells. *Current Opinion in Biotechnology* **48**, 1–7 (2017).

53. Zhao, J., Burke, A. J. & Green, A. P. Enzymes with noncanonical amino acids. *Current Opinion in Chemical Biology* **55**, 136–144 (2020).

54. Calero, P. & Nikel, P. I. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microbial Biotechnology* **12**, 98–124 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.19.432053; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

55. Youngman, P. J., Perkins, J. B. & Losick, R. Genetic transposition and insertional mutagenesis in Bacillus subtilis with Streptococcus faecalis transposon Tn917. *Proc Natl Acad Sci U S A* **80**, 2305–2309 (1983).

56. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).

57. Zhang, M. S. *et al.* Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nature Methods* **14**, 729–736 (2017).

58. Harwood, C. R. & Cutting, S. M. *Molecular biological methods for Bacillus*. (Wiley, 1990).

59. Burton, B. M., Marquis, K. A., Sullivan, N. L., Rapoport, T. A. & Rudner, D. Z. The ATPase SpoIIIE Transports DNA across Fused Septal Membranes during Sporulation in Bacillus subtilis. *Cell* **131**, 1301–1312 (2007).

60. Huppert, L. A. *et al.* The ESX System in Bacillus subtilis Mediates Protein Secretion. *PLOS ONE* **9**, e96267 (2014).