Biochemistry

Communication

pubs.acs.org/biochemistry

# Expanding the Genetic Code of a Photoautotrophic Organism

Yonatan Chemla,[†] Mor Friedman,[†] Mathias Heltberg,[‡] Anna Bakhrat,[†] Elad Nagar,[§] Rakefet Schwarz,[§] Mogens Høgh Jensen,[‡] and Lital Alfonta*[,†]

[†]Department of Life Sciences and Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

[‡]Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark

[§]The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan 5290002, Israel

Supporting Information

**ABSTRACT:** The photoautotrophic freshwater cyanobacterium *Synechococcus elongatus* is widely used as a chassis for biotechnological applications as well as a photosynthetic bacterial model. In this study, a method for expanding the genetic code of this cyanobacterium has been established, thereby allowing the incorporation of unnatural amino acids into proteins. This was achieved through UAG stop codon suppression, using an archaeal pyrrolysyl orthogonal translation system. We demonstrate incorporation of unnatural amino acids into green fluorescent protein with 20 ± 3.5% suppression efficiency. The introduced components were shown to be orthogonal to the host translational machinery. In addition, we observed that no significant growth impairment resulted from the integration of the system. To interpret the observations, we modeled and investigated the competition over the UAG codon between release factor 1 and pyl-tRNA$_{CUA}$. On the basis of the model results, and the fact that 39.6% of the stop codons in the *S. elongatus* genome are UAG stop codons, the suppression efficiency in *S. elongatus* is unexpectedly high. The reason for this unexpected suppression efficiency has yet to be determined.

Cyanobacteria play a significant role on earth as aquatic primary producers and oxygen generators and are crucial in $CO_2$ fixation. Cyanobacteria are also prolific producers of natural products and unique enzymes[1] that are being systematically improved for biotechnological purposes.[2] *Synechococcus elongatus* sp. PCC7942 (*S. elongatus*) is a free-living, freshwater cyanobacterium strain. This strain is relatively simple to culture, is characterized by high biomass production, and can be easily genetically manipulated. By these virtues, *S. elongatus* has been utilized for a myriad of applications. Among these are biofuel production,[3] renewable energy,[4] and $CO_2$ reduction.[5] *S. elongatus* is also used as a model organism for studies of circadian rhythm,[6] biofilm formation,[7] carbon fixation,[8] and photosynthesis.[9] In recent years, multiple genetic tools and methods have been developed to facilitate *S. elongatus* genetic manipulation. However, while in the past 15 years, genetic code expansion tools were introduced and improved in *Escherichia coli*,[10] Actinobacteria,[11] *Caenorhabditis elegans*,[12] *Drosophila*,[13] and mammalian cells,[14] they were not adapted for cyanobacteria.

The most notable advantage of genetic code expansion is the ability to incorporate, site-specifically, a synthetic amino acid of choice aimed to introduce new chemical or physical properties into a desired protein. The common strategy utilized for this is stop codon suppression, using an archaeal orthogonal translation system (OTS) that does not cross-react with the host native tRNAs and tRNA synthetases. The archaea *Methanosarcina mazei* (*Mm*) were found to genetically encode pyrrolysine with the utilization of the UAG stop codon as a sense codon. This process is enabled by a unique *Mm*-pyrrolysyl tRNA synthetase (PylRS) and *Mm*-tRNA$_{CUA}^{pyl}$ (pyl-tRNA)[15] pair. Once cloned, this pair was used to incorporate more than 100 synthetic unnatural amino acids (UAAs) into proteins in the three different kingdoms of life[16] and was demonstrated to be incorporated in vivo and in vitro.[17] For these reasons, the pyl-orthogonal pair was chosen as a candidate for attempting to expand the genetic code of *S. elongatus*.

Herein, we aim to introduce and develop a general method for incorporation of UAAs into proteins in the cyanobacterium *S. elongatus*. To achieve this, the pyrrolysyl orthogonal translation system (pyl-OTS) was cloned to a pCB4 shuttle vector (Table S1), and the resulting construct was named pCOTS-pyl. In addition, a model protein for genetic code expansion, GFP Y35TAG, was cloned, resulting in pCOTS-pyl-35TAG-GFP, whereas the UAA was incorporated using the UAG stop codon as a sense codon, replacing tyrosine. Furthermore, a second genetic construct was designed for genomic recombination of the pyl-OTS into the *S. elongatus* chromosome (Figure S1). In this strain, the target gene, GFP, was expressed from the pES94-35TAG-GFP plasmid (Table S1). This genomically modified strain was named 7942-PO. In both cases, GFP was expressed under the P$_{cpcB1}$ promoter[18] and T$_{cpcA}$ terminator (Figure 1a) and PylRS was expressed under the P$_{rbcL}$ promoter and terminator. The pyl-tRNA was expressed under the *S. elongatus* Leu-tRNA$_{CAG}$ promoter and terminator, which were extracted from the *S. elongatus* genome (Figure 1b) (sequences are available in the Supporting Information). The two *S. elongatus* variants, pCOTS-pyl-

**Received:** February 15, 2017
**Revised:** April 9, 2017
**Published:** April 10, 2017

Downloaded via BAR-ILAN UNIV on February 9, 2025 at 14:09:26 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications © 2017 American Chemical Society



DOI: 10.1021/acs.biochem.7b00131
*Biochemistry* 2017, 56, 2161−2165

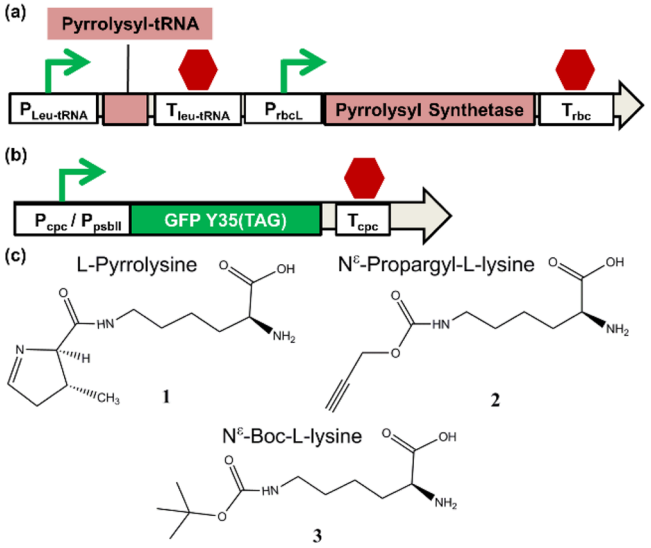



**Figure 1.** Genetic code expansion constructs that allow incorporation of UAAs into proteins in *S. elongatus*. (a) Genetic map of the OTS. (b) Genetic map for protein expression. (c) UAAs mentioned in this study.

35TAG-GFP and 7942-PO, were selected by antibiotic markers and validated by sequencing.

PylRS is a promiscuous enzyme as it was shown to recognize several UAAs in addition to its original substrate shown in Figure 1c, 1-pyrrolysine (**1**). Two of these UAAs are shown in Figure 1c: $N^{\varepsilon}$-propargyl-L-lysine (**2**) that allows a "click" reaction[19] and $N^{\varepsilon}$-boc-L-lysine (**3**) that is being used as a protected amino acid in bioorthogonal chemical conjugation reactions.[20] UAAs **2** and **3** were chosen to be incorporated as a general proof of concept for the genetically expanded *S. elongatus*. Either **2** or **3** was added to the growing bacterial cultures, containing either one of the genetic constructs, for incorporation into GFP(Y35TAG). To test the OTS in *S. elongatus*, we have grown the transformed cells in the absence of a UAA as a negative control. As a positive control, we have transformed the pCOTS-pyl-GFP-WT vector encoding the wild-type GFP gene. The incorporation of both **2** and **3** into GFP was successful, while no synthesis of protein was observed in the absence of a UAA (Figure 2a). To ensure the orthogonality of the pyl-tRNA in the *S. elongatus* cell, a genetic construct with only pyl-tRNA and the reporter Y35TAG GFP gene in the absence of the PylRS gene (pCOTS-ΔPylRS-GFP) was transformed and cultured in the presence of **2**. No expression of the reporter protein was observed (Figure S2), which led us to conclude that no endogenous tRNA synthetase can aminoacylate pyl-tRNA.

Aiming to optimize the growth conditions of *S. elongatus*, we examined variable UAA concentrations (Figure S3) and UAA addition at different $OD_{750}$ values (Figure S4). The optimum was found to be 1 mM UAA, and the optimal $OD_{750}$ for UAA addition was 0.01. Next, the optimal $OD_{750}$ for bacterial harvesting was found to be at an $OD_{750}$ of 1.1 (Figure S5). After optimization, the efficiency of the systems was evaluated. We have measured the UAA-incorporated GFP concentrations compared to the concentration of WT GFP (Figure 2b). This comparison is termed suppression efficiency. It is an indirect measurement of the ability of the OTS to reassign, at a specific site, the original "meaning" of the UAG codon. The pCOTS-pyl system exhibited up to 20 ± 3.5% suppression efficiency with yields of up to 15 μg/L (of original culture), while the 7942-PO system exhibited only 2% suppression efficiency with yields of up to 1.5 μg/L. Both systems resulted in relatively low yields; nonetheless, the vector-based pCOTS-pyl system showed results significantly better than those of the genome recombination system, and thus, further experiments were performed primarily with this system. Yields were further increased by changing the GFP expression promoter, $p_{cpcB1}$, to the *S. elongatus* $P_{psbII}$ promoter. This change improved the yields by a factor of 5 and resulted in yields of 375 and 75 μg/L [for WT and Y35(**2**)-GFP, respectively], with 22 ± 7.5% suppression efficiency (Figure 2c and Figure S6). In the future, protein yields can be increased by improving the combination of promoter strength, ribosome binding site strength, and adaptation of coding sequence to that of *S. elongatus* as was recently shown by us in *E. coli*.[21]

To validate the selectivity of UAA incorporation, the molecular masses of produced proteins were measured using electrospray ionization mass spectrometry (ESI-MS) (Figure 2d). The experimentally observed masses of the Y35(**2**) mutant and WT GFP were within 1 Da of the calculated masses.

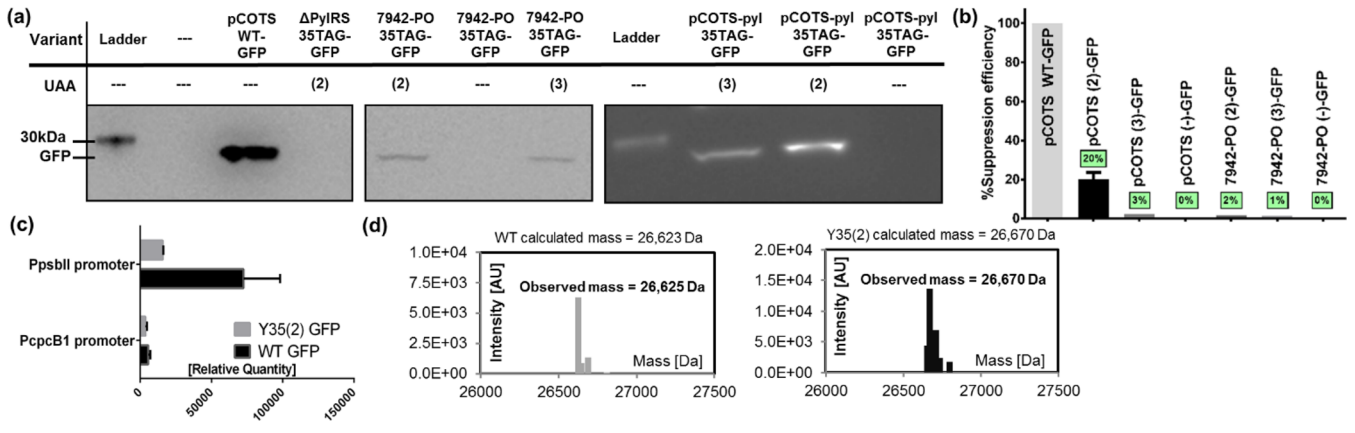



**Figure 2.** Incorporation of UAAs into GFP. (a) Representative Western blot analysis of GFP expression in *S. elongatus* variants, using different vectors, in the presence or absence of either **2** or **3**. (b) Suppression efficiencies of the different GFP variants relative to WT GFP expression. (c) Densitometry analysis comparing WT and Y35(**2**)-GFP expression levels between two promoters. (d) Mass spectrometry (electrospray ionization mass spectrometry with liquid chromatography) results for WT-GFP (gray) and Y35(**2**)-GFP (black).



DOI: 10.1021/acs.biochem.7b00131
*Biochemistry* 2017, 56, 2161−2165

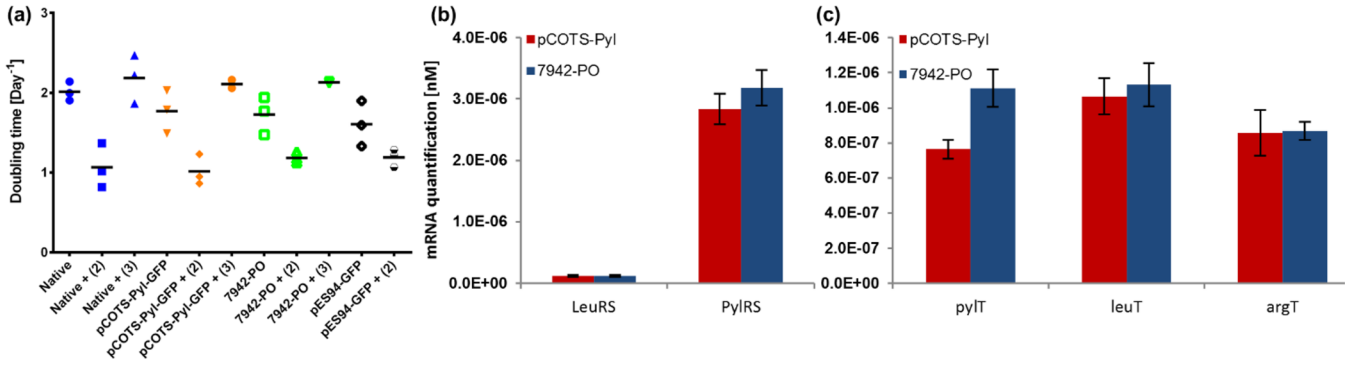



**Figure 3.** (a) Doublings per day of the different *S. elongatus* variants in the presence or absence of **2** or **3**. (b) Quantitative polymerase chain reaction (qPCR)-determined quantities of PylRS and LeuRS mRNA in *S. elongatus*. (c) qPCR quantities of pyl-tRNA, leu-tRNA$_{CAG}$, and arg-tRNA$_{CCT}$.

Observing these masses, we concluded that the UAA is selectively incorporated into GFP (Figure S7).

As pyl-tRNA suppresses the UAG stop codon, it should compete with release factor 1 (RF1) for the translation or termination of transcribed genes of *S. elongatus*. In fact, it was found that *S. elongatus* utilizes TAG termination in 39.6% of its open reading frames (ORFs), compared to only 7.4% in *E. coli*.[22] Hence, the important question of whether translation termination is being suppressed in TAG terminating genes arises, and if so, what are its adverse effects on the host organism if any? It is reasonable to assume that if the signal for the correct termination of almost 40% of the genes in *S. elongatus* is compromised, even to the smallest extent, *S. elongatus* growth should be impaired. This argument is emphasized when taking into account the fact that many of the photosystem genes terminate with TAG. To test this, the log-phase doubling times of the different strains in the presence or absence of different UAAs were followed (Figure 3a). It was found that pCOTS-pyl and 7942-PO variants, supplied with 1 mM **3**, did not show any significant reduction in growth rates (measured as the number of doublings of the population density per day) compared to that of native *S. elongatus* (1.9 ± 0.3 and 1.9 ± 0.4 day$^{-1}$ for pCOTS-pyl and 7942-PO variants, respectively, compared to 2.0 ± 0.1 day$^{-1}$ for the native form). However, the same variants, supplied with **2**, suffered a significant reduction in the number of doublings of the population density of 40−50% per day (1.0 ± 0.2 and 1.2 ± 0.1 day$^{-1}$, respectively).

The same reduction is observed when supplying native *S. elongatus* with **2**, 1.0 ± 0.2 day$^{-1}$ compared to the value of 2.0 ± 0.1 day$^{-1}$ of native *S. elongatus* in the absence of **2**. These results lead us to conclude that very low, if any, toxicity stems from the introduction of the OTS into this strain of cyanobacteria, while significant levels of toxicity stem from the introduction of **2** into *S. elongatus*. To further investigate the source of the toxicity of **2**, native and pCOTS-pyl strains were supplied with both **2** and L-lysine in equimolar quantities; in these cases, no toxicity was observed (data not shown). Therefore, we suspected that the cause of the toxicity is interference with the lysine synthesis metabolic pathway.

To better characterize the orthogonal translation system, the mRNA expression level of PylRS was measured using RT-PCR and was compared to that of the native leucine tRNA synthetase, which is a highly abundant synthetase (Figure 3b). The expression level of pyl-tRNA was also measured and compared to that of the most used leu-tRNA$_{CAG}$ and the least used arg-tRNA$_{CCT}$ (Figure 3c). The results suggest that the level of expression of the pyl-tRNA is relatively low and should be increased to improve the efficiency of the system. Counterintuitively, while the genomically recombined system shows an increased level of expression of the OTS, its efficiency is lower, and at present, we have no explanation for this intriguing observation. Finally, we tested the incorporation of **2** into different sites in GFP (Figure 4a) and simultaneously into multiple sites; however, multiple incorporation was undetectable (Figure S9).

The suppression efficiency achieved in this study was ∼20%, despite the low cellular concentration of pyl-tRNA and high

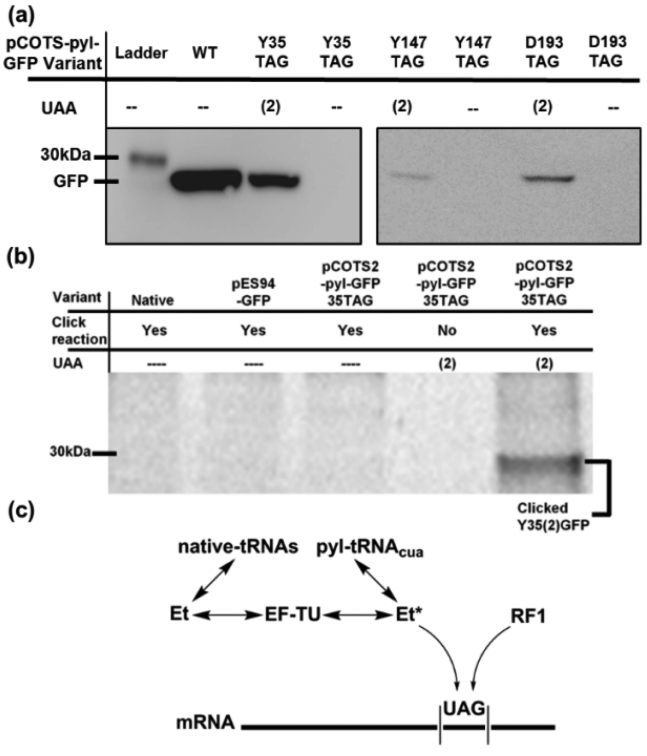



**Figure 4.** (a) Representative Western blot of *S. elongatus* lysates containing GFP variants with TAG mutations at different sites. (b) Representative fluorescent (excitation at 532 nm and emission at 575 nm) sodium dodecyl sulfate−polyacrylamide gel electrophoresis gel loaded with *S. elongatus* lysates that underwent click reaction to TAMRA-azide. (c) Scheme of the RF1 and pyl-tRNA competition model in which Et and Et* stand for the complex of the elongation factor (EF-Tu) bound to endogenous aminoacylated tRNAs and pyl-tRNA, respectively.



DOI: 10.1021/acs.biochem.7b00131
*Biochemistry* 2017, 56, 2161−2165

levels of RF1 needed for termination of ∼40% of ORFs in the *S. elongatus* genome. As a comparison, in *E. coli*, where the extent of genomic TAG termination is ∼7%, the initial suppression efficiency was <8%[23] and has gradually improved to 60%.[24] Given the suppression efficiencies, in *E. coli* and in *S. elongatus*, we expected that the off-target native chromosomal genes terminating with TAG would be suppressed with similar efficiencies. The system developed in this study allowed the exploration of this prediction in a bacterial strain with significant UAG termination usage. As described above, *S. elongatus* does not exhibit any significant difference in growth rates with and without an aminoacylated pyl-tRNA; this implies that there is no significant UAG suppression in chromosomal genes.

To test if this lack in toxicity is unique to *S. elongatus*, we examined two *E. coli* strains (DH5α and BL21) transformed with pEVOL-pyl-OTS and supplied with **2** or **3** (Figure S10); no reduction in growth rates was observed in either form of *E. coli*. The toxicity could possibly be alleviated by the use of multiple stop signals downstream from the suppressed UAG stop codon, but if this is the case, it should be expected that the UAA will be incorporated into those suppressed UAG sites. To test this claim directly and to evaluate one possible downstream application of the presented system, both *E. coli* and *S. elongatus* were cultured with vectors harboring pyl-OTS and **2**. The cultures were lysed and underwent a "click" reaction between proteins incorporated with **2** and a TAMRA-azide fluorescent dye. The fluorescent gels of the reaction products were analyzed (Figure 4b and Figures S11 and S12). The incorporation of **2** could be detected only in Y35TAG GFP, while no significant incorporation was observed in off-target proteins that terminate with TAG. Granted, it could be that the efficiency of the click reaction as well as the low suppression efficiency of the Pyl-tRNA will not afford a signal that is strong enough to be detected; however, in *S. elongatus*, strongly expressed photosystem components terminate with UAG. Taken together, these results may imply the existence of a mechanism by which the organism avoids unwarranted read-through events; this observation is in agreement with previous studies in *E. coli*.[25,26] The intriguing bias in suppression has also been recently observed in *E. coli* and HEK293T cells and was attributed to context effects, an increased level of degradation due to incorrect termination, and low expression levels of the off-target genes.[27,28] One of the causes of context dependence was shown to be that the identity of the fourth nucleotide in the 3′ end of the UAG codon affects the affinity of the suppressor or the terminator for the UAG codon.[29] Another hypothesis suggested that the position of the UAG codon in the gene may also affect local ribosomal density that, in turn, will influence protein expression.[21] Furthermore, and in agreement with these findings, Ozer et al. showed that RF1 has no observable influence on near-cognate pyl-tRNA mis-suppression of the UAG codon.[30] However, while these factors may play a part in this bias, we find it unlikely that they account for the apparently complete absence of pyl-tRNA suppression in the chromosomal genes of *S. elongatus*.

To further characterize this phenomenon, we modeled the cellular competition between RF1 and pyl-tRNA over the UAG codon in *E. coli* (Figure 4c). The competition between native tRNAs and pyl-tRNA over the elongation factor (EF-Tu) was calculated. It was done by using the reported affinity between EF-Tu and both pyl-tRNA[31] and native tRNAs,[32] it also took into account the approximate cellular concentrations of these molecules. This revealed the steady-state concentration of EF-Tu molecules bound to aminoacylated pyl-tRNA (Et*). Next, the competition over the UAG codon between Et* and RF1 could be assessed using their measured affinities (detailed in the Supporting Information). Our results show that RF1 should outcompete the pyl-$tRNA_{CUA}$ by more than 3 orders of magnitude (RF1 is ∼1365 times more likely to react with a UAG codon, according to the model). This explains the fact that no reduction in generation time was observed in *E. coli* or *S. elongatus*.

However, the model predicts that the suppression efficiencies should be negligible (<0.1%). This calculation for *E. coli* is in direct contradiction with the observed data by 2.5 orders of magnitude. Considering the significant abundance of UAG termination in *S. elongatus* compared to *E. coli*, and the correlation between stop codon usage and RF1 abundance,[22] the model results are further contradictory to our finding that the suppression efficiency is 20% in *S. elongatus*. These results stipulate the existence of a mechanism that reduces the level of premature stop codon termination. The existence of such a mechanism suggests an evolutionarily selective advantage by reducing the level of premature termination by nonsense mutations or translation errors, which requires further investigation.

In conclusion, this study reports a facile method of genetic code expansion in the cyanobacterium *S. elongatus*. The incorporation of UAAs was demonstrated by incorporating **2** or **3** into three different sites of GFP as a proof of concept. This protein augmentation ability combined with the unique attributes of cyanobacteria should give rise to many new applications, such as single-molecule studies in cyanobacteria with site-specifically labeled proteins, coupling of the photosynthetic system to modified proteins of interest, and large-scale production of proteins with unique new properties utilizing the solar energy harvested by these bacteria.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b00131.

Biochemical model, methods, additional figures, and gene sequences (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*E-mail: alfontal@bgu.ac.il.

### ORCID

Lital Alfonta: 0000-0002-3805-8625



### Funding

This research was supported in part by ERC Grant 260647.


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Eden Ozer for assistance with "click" experiments and Michael Shaferman, Orr Schlesinger, Rami Parnassa, and Ashok Bhagat for providing assistance and advice.



DOI: 10.1021/acs.biochem.7b00131
*Biochemistry* 2017, 56, 2161−2165

## REFERENCES


(1) Kehr, J. C., Gatte Picchi, D., and Dittmann, E. (2011) *Beilstein J. Org. Chem. 7*, 1622−1635.
(2) Bernstein, H., et al. (2016) *mBio 7*, e00949-16.
(3) Dexter, J., and Fu, P. (2009) *Energy Environ. Sci. 2*, 857−864.
(4) Quintana, N., Van Der Kooy, F., Van De Rhee, M. D., Voshol, G. P., and Verpoorte, R. (2011) *Appl. Microbiol. Biotechnol. 91*, 471−490.
(5) Atsumi, S., Higashide, W., and Liao, J. C. (2009) *Nat. Biotechnol. 27*, 1177−1180.
(6) Cohen, S. E., and Golden, S. S. (2015) *Microbiol. Mol. Biol. Rev. 79*, 373−385.
(7) Parnasa, R., Nagar, E., Sendersky, E., Reich, Z., Simkovsky, R., Golden, S., and Schwarz, R. (2016) *Sci. Rep. 6*, 32209.
(8) Savage, D. F., Afonso, B., Chen, A. H., and Silver, P. A. (2010) *Science 327*, 1258−1261.
(9) Nelson, N., and Junge, W. (2015) *Annu. Rev. Biochem. 84*, 659−683.
(10) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) *Science 292*, 498−500.
(11) He, J., Van Treeck, B., Nguyen, H. B., and Melançon, C. E. (2016) *ACS Synth. Biol. 5*, 125−132.
(12) Greiss, S., and Chin, J. W. (2011) *J. Am. Chem. Soc. 133*, 14196−14199.
(13) Bianco, A., Townsley, F. M., Greiss, S., Lang, K., and Chin, J. W. (2012) *Nat. Chem. Biol. 8*, 748−750.
(14) Liu, W., Brock, A., Chen, S., Chen, S., and Schultz, P. G. (2007) *Nat. Methods 4*, 239−244.
(15) Srinivasan, G., James, C. M., and Krzycki, J. A. (2002) *Science 296*, 1459−1462.
(16) Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015) *Chem. Sci. 6*, 50−69.
(17) Chemla, Y., Ozer, E., Schlesinger, O., Noireaux, V., and Alfonta, L. (2015) *Biotechnol. Bioeng. 112*, 1663−1672.
(18) Sawaki, H., Sugiyama, T., and Omata, T. (1998) *Plant Cell Physiol. 39*, 756−761.
(19) Thirumurugan, P., Matosiuk, D., and Jozwiak, K. (2013) *Chem. Rev. 113*, 4905−4979.
(20) Nguyen, D. P., Alai, M. M. G., Virdee, S., and Chin, J. W. (2010) *Chem. Biol. 17*, 1072−1076.
(21) Schlesinger, O., Chemla, Y., Heltberg, M., Ozer, E., Marshall, R., Noireaux, V., Jensen, M. H., and Alfonta, L. (2017) *ACS Synth. Biol.*, DOI: 10.1021/acssynbio.7b00019.
(22) Korkmaz, G., Holm, M., Wiens, T., and Sanyal, S. (2014) *J. Biol. Chem. 289*, 30334−30342.
(23) Ryu, Y., and Schultz, P. G. (2006) *Nat. Methods 3*, 263−266.
(24) Chatterjee, A., Xiao, H., and Schultz, P. G. (2012) *Proc. Natl. Acad. Sci. U. S. A. 109*, 14841−14846.
(25) Eggertsson, G., and Söll, D. (1988) *Microbiol. Rev. 52*, 354−374.
(26) Johnson, D. B. F., et al. (2011) *Nat. Chem. Biol. 7*, 779−786.
(27) Uttamapinant, C., Howe, J. D., Lang, K., Beránek, V., Davis, L., Mahesh, M., Barry, N. P., and Chin, J. W. (2015) *J. Am. Chem. Soc. 137*, 4602−4605.
(28) Kipper, K., Lundius, E. G., Curic, V., Nikic, I., Lemke, E. A., Wiessler, M., and Elf, J. (2017) *ACS Synth. Biol. 6*, 233−255.
(29) Smolskaya, S., Zhang, Z. J., and Alfonta, L. (2013) *PLoS One 8*, e68363.
(30) Ozer, E., Chemla, Y., Schlesinger, O., Aviram, H., Riven, I., Haran, G., and Alfonta, L. (2017) *Biotechnol. Bioeng. 114*, 1065−1073.
(31) Wang, J., Kwiatkowski, M., and Forster, A. C. (2016) *Biotechnol. Bioeng. 113*, 1552−1559.
(32) Ieong, K., Pavlov, M. Y., Kwiatkowski, M., Ehrenberg, M., and Forster, A. C. (2014) *RNA 20*, 632−643.




DOI: 10.1021/acs.biochem.7b00131
*Biochemistry* 2017, 56, 2161−2165