Published in final edited form as:

Chem. 2020 October 8; 6(10): 2717–2727. doi:10.1016/j.chempr.2020.07.013.

Creation of Bacterial cells with 5-Hydroxytryptophan as a 21st Amino Acid Building Block

Yuda Chen¹, Juan Tang¹, Lushun Wang¹, Zeru Tian¹, Adam Cardenas¹, Xinlei Fang¹,
Abhishek Chatterjee², Han Xiao^{1,3,4,5,*}

¹Department of Chemistry, Rice University, 6100 Main Street, Houston, Texas, 77005. U.S.A.

²Department of Chemistry, Boston College, 2609 Beacon Street, 246B Merkert Chemistry Center, Chestnut Hill, MA, 02467, U.S.A.

³Department of Biosciences, Rice University, 6100 Main Street, Houston, Texas, 77005, U.S.A.

⁴Department of Bioengineering, Rice University, 6100 Main Street, Houston, Texas, 77005, U.S.A.

⁵Lead Contact

SUMMARY

While most organisms utilize 20 canonical amino acid building blocks for protein synthesis, adding additional candidates to the amino acid repertoire can greatly facilitate the investigation and manipulation of protein structures and functions. In this study, we report the generation of completely autonomous organisms with a 21st ncAA, 5-hydroxytryptophan (5HTP). Like 20 canonical amino acids, 5-hydroxytryptophan can be biosynthesized *in vivo* from simple carbon sources and is subsequently incorporated into proteins in response to the amber stop codon. Using this unnatural organism, we have prepared a single-chain immunoglobulin variable fragment conjugated with a fluorophore and demonstrated the utility of these autonomous cells to monitor oxidative stress. Creation of this and other cells containing the 21st amino acid will provide an opportunity to generate proteins and organisms with novel activities, as well as to determine the evolutionary consequences of using additional amino acid buildings.

Graphical Abstract

The ability to introduce noncanonical amino acid building blocks into proteins in living cells provides a powerful tool for investigating and manipulating the structure and function of proteins and of the organisms themselves. In this study, we have created a completely autonomous organism that uses 5-hydroxytryptophan as a noncanonical 21st amino acid. We demonstrate the

*Correspondence: han.xiao@rice.edu.

AUTHOR CONTRIBUTIONS

Y.C. and H.X. designed all the experiments. Y.C., J.T., L.W., Z.T., C.A., and X.F. performed research; Y.C., A.C., and H.X. contributed new reagents/analytic tools; Y.C., J.T., and H.X. analyzed data; and Y.C., J.T., A.C., and H.X. wrote the manuscript.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DECLARATION OF INTERESTS

The authors declare no competing interests.

utility of this unnatural organism by using it to prepare site-specific protein conjugates and to monitor low levels of oxidative stress.

INTRODUCTION

With the rare exception of pyrrolysine and selenocysteine, 20 amino acid building blocks, with a limited number of functional groups, are used by almost all organisms for the biosynthesis of proteins. However, the literature overwhelmingly shows that biological systems require a still higher level of chemical complexity in proteins for carrying out many biological functions. This is evidenced by the discovery of a variety of function-altering post-translational protein modifications, the use of cofactors to enhance enzyme function, and nonproteinogenic amino acids within nonribosomal peptides.¹⁻⁵ These observations suggest that expansion of the genetic code could provide an additional means of generating novel protein structures and functions as well as an evolutionary impact for different species. To create a new organism with a 21st amino acid building block, one requires the use of bioorthogonal translational machinery, a “blank” (normally non-coding) codon, and the capacity for autonomous biosynthesis of the noncanonical amino acid (ncAA) of interest.³⁻⁸ Current efforts to expand the genetic code have focused on the addition of new components to the translational machinery of various organisms to achieve the selective and faithful genetic incorporation of ncAAs into proteins.^{4-6,9-11} Bioorthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA pairs have been developed that allow utilization of nonsense or frameshift codons for the incorporation of more than 200 ncAAs with diverse biological,

chemical, and physical properties in both prokaryotic and eukaryotic cells.^{4-6,10,12-19} To further expand the number of codons encoding amino acids, quadruplet codons and recoded organisms have been developed for providing more blank codons to encode ncAAs.²⁰⁻²⁵ Besides the canonical base pair, unnatural base pairs created via the use of hydrophobic and packing forces have recently been added to the genetic code to direct the site-specific incorporation of ncAAs into proteins in *E. coli*.¹⁶ Despite these areas of progress in genetic code expansion technology, current methodologies have largely relied on the exogenous feeding of chemically-synthesized ncAAs and successful uptake of these ncAAs by cells. In contrast, there have been very few reports of autonomous organisms capable of biosynthesizing ncAAs *in vivo*, followed by incorporation of the ncAAs into proteins.

The dearth of attempts at creating this type of completely autonomous organism is explained by the scarcity of verified biosynthetic pathways for producing ncAAs and by the frequent incompatibility of these verified pathways with genetic code expansion technology. Bioorthogonal translational machinery used by genetic code expansion systems is known to utilize the standalone ncAAs. But most enzymes modifying amino acids can only recognize the amino acids on the polypeptide or polyketide synthases, and are not able to generate standalone ncAAs.²⁷⁻²⁹ Furthermore, a significant intracellular level (ca. millimolar) of the ncAA must be produced via the biosynthetic pathway to allow efficient utilization of the “blank” codon to specify incorporation of the ncAA. The first example of completely autonomous organisms is an *E. coli* strain with the ability to genetically encode the 21st amino acid *p*-amino-phenylalanine (*pAF*).³⁰ Harboring a *pAF* synthetic pathway borrowed from *Streptomyces venezuelae* and an engineered *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS)/tRNA pair specific for *pAF*, this bacterium has the ability to biosynthesize the *pAF* from basic carbon sources as well as the ability to genetically incorporate *pAF* with high fidelity into proteins in response to the presence of amber codons.^{30,31} Using a similar strategy, phosphothreonine (pThr) can be biosynthesized in *E. coli* in the presence of *Salmonella enterica* kinase, followed by site-specific incorporation into proteins using an unusual phosphoseryl-tRNA synthetase (SepRS)/tRNA pair from methanogenic archaea.³² Surprisingly, compared to exogenous feeding of 1 mM pThr, a 40-fold increase in intracellular pThr concentration was achieved via this biosynthetic strategy, greatly facilitating the genetic incorporation of pThr.³² Recently, a unique pathway to biosynthesize terminal alkyne-containing amino acid from l-lysine has been discovered in *Streptomyces cattleya* and used to produce alkyne-labelled protein via residue-specific replacement of Met residues.^{33,34} However, the lack of aaRS/tRNA pair for this alkyne amino acid limits its potential applications in the genetic code expansion. In another example, the use of tyrosine phenol-lyase created an *E. coli* strain able to biosynthesize and incorporate dihydroxyphenylalanine (DOPA).³⁵ But the biosynthesis of DOPA in this strain requires the exogenous feeding of catechol, pyruvate, and ammonia.³⁵ Thus, further evolution of the genetic code expansion technology will be required for constructing improved versions of unnatural organisms with the ability to autonomously biosynthesize an ncAA and use it for protein synthesis.

Here, we report the generation of a completely autonomous *E. coli* strain that can biosynthesize 5-hydroxyl-tryptophan (5HTP) and genetically incorporate it into proteins (Fig. 1). We further demonstrate the utility of this autonomous strain by preparing a site-

specific antibody-fluorophore conjugate using a rapid bioorthogonal azo-coupling reaction recently developed and monitoring oxidative stress of bacterial with fluorescent proteins.

RESULTS

Screening of 5HTP Biosynthetic Pathways

5HTP is a naturally occurring ncAA primarily known as a precursor of the neurotransmitter serotonin in humans.³⁶ 5HTP is produced in eukaryotic cells through the hydroxylation of tryptophan by tryptophan 5-hydroxylase (T5H) in the presence of the co-factor tetrahydrobiopterin (BH4).^{37,38} Besides using eukaryotic hydroxylases to produce 5HTP, efforts have also been made to redirect the substrate specificity of prokaryotic phenylalanine 4-hydroxylase (P4H) from phenylalanine to tryptophan.³⁹⁻⁴¹ Instead of using BH4 as a cofactor, bacterial P4Hs are known to utilize the co-factor tetrahydromonapterin (MH4), a major form of pterin in *E. coli*.³⁹ To find the best hydroxylase for producing 5HTP in *E. coli*, we first tested four hydroxylases from different organisms, including hydroxylases from *Oryctolagus cuniculus* (OcT5H), *Schistosoma mansoni* (SmT5H), *Mesorhizobium ciceri biserrulae* (McP4H), and *Xanthomonas campestris* (XcP4H). All these hydroxylases are confirmed to utilize MH4 as a co-factor.⁴² Next, we introduced an artificial MH4 recycling pathway into *E. coli* as a means of enhancing MH4 levels as well as cell viability. This recycling pathway contains two genes: the *folM* gene, encoding *E. coli* dihydromonapterin reductase (DHMR) responsible for the conversion of dihydroxymonapterin (MH2) to MH4; and the *phhB* gene, encoding *Pseudomonas aeruginosa* pterin-4 α -carbinolamine dehydratase (PCD) to regenerate MH2 (Fig. 1).⁴³ The hydroxylase gene together with this MH4 recycling pathway were then subcloned into a reporter plasmid harboring a green fluorescence protein expression cassette with one amber codon at Tyr151.^{44,45} The resulting plasmid is designated pET22b-T5-sfGFP*-P4H/T5H-MH4R. To generate the bioorthogonal machinery for incorporating 5HTP, we used a modified *Saccharomyces cerevisiae* tryptophanyl-tRNA synthetase (*Sc*TrpRS)/tRNA^{Trp} pair.⁴⁶ We first subcloned the *Sc*TrpRS/tRNA pair into a recently reported pUltra vector.⁴⁷ The resulting pUltra-5HTP encodes lacI-driven *Sc*TrpRS and *proK*-driven *Sc*-tRNA^{Trp} cassettes. This new suppressor plasmid is very robust, exhibiting efficient suppression of the amber codon at exogenous levels of 5HTP as low as 0.1 mM (Fig. S1). To evaluate the relative abilities of different hydroxylases to produce biosynthetic 5HTP for genetic incorporation, pUltra-5HTP was transformed into BL21(DE3) cells containing pET22b-T5-sfGFP*-P4H/T5H-MH4R plasmids expressing different hydroxylase genes. Tests of protein expression were carried out in M9-glucose minimal medium (M9G) for 16 h in parallel with control cells harboring pET22b-T5-sfGFP* and pUltra-5HTP in the presence and absence of exogenously fed 1 mM 5HTP. As expected, sfGFP was only expressed in the presence of 1 mM 5HTP in the case of cells lacking a hydroxylase gene. Among the cells harboring a hydroxylase gene, we found that cells expressing XcP4H exhibited the largest increase in fluorescence, suggesting a higher expression level of full-length 5HTP-containing sfGFP (Fig. 2A).

Characterizations of XcP4H-mediated 5HTP Biosynthetic Pathway and its Incorporation into Protein

To confirm that P4H from *Xanthomonas campestris*, in conjunction with the MH4 recycling pathway and the *ScTrpRS/tRNA^{Trp}* pair, lead to the successful incorporation of biosynthesized 5HTP into sfGFP in response to the amber codon, tests of sfGFP expression were carried out in the absence of each of these gene cassettes. We found that removal of either the *XcP4H* or MH4 recycle genes dramatically decreases the expression level of sfGFP, suggesting that both the hydroxylase and the pterin regeneration pathway are required for the efficient biosynthesis of 5HTP (Fig. 2B). Furthermore, no background incorporation of biosynthesized 5HTP into sfGFP was observed in the absence of the *ScTrpRS/tRNA^{Trp}* pair (Fig. 2B). Next, we optimized the conditions for expressing proteins containing biosynthesized 5HTP. As shown in Figs. 2B and S2-5, protein expression time, medium, and carbon source do not have significant effects on the expression of proteins containing biosynthesized 5HTP. However, the M9-glucose minimal medium containing 0.5 g/L tryptophan exhibited a significantly enhanced expression level of sfGFP (Fig. 2C). Besides using normalized sfGFP fluorescence as a readout of 5HTP biosynthesis, we also measured the intracellular 5HTP concentration. Cells harboring both *XcP4H* and MH4 recycle or either gene cassette was expressed in M9-glucose minimal medium in the presence or absence 0.5 g/L tryptophan for 16 h (Table 1 and Fig. S6). At the same time, cells with or without exogenously fed 1 mM 5HTP were used as positive and negative controls, respectively. We found that high level of 5HTP could be detected in cells with both *XcP4H* and MH4 recycle pathway. Consistent with sfGFP fluorescence data, removal of either *XcP4H* or MH4 recycle will result in decreased intracellular 5HTP concentration in *E. coli*. Considering the complicated regulation of intracellular tryptophan concentration, the additional tryptophan consumption by 5HTP biosynthesis is not likely to disrupt its homeostasis and *E. coli* fitness, which is evidenced by similar growth curves and doubling times in all tested conditions with and without a 5HTP biosynthetic pathway (Fig. S7).^{48,49}

To further investigate the efficiency and specificity of the incorporation of biosynthesized 5HTP, sfGFPs containing exogenously fed or biosynthesized 5HTP were purified by Ni²⁺-NTA affinity chromatography and characterized by SDS-PAGE and ESI-MS. Full-length sfGFP was only expressed in the presence of exogenously fed 5HTP or when biosynthesis of 5HTP was induced (Fig. 2D). 10.5 mg L⁻¹ sfGFP containing biosynthesized 5HTP was produced, compared with the yield of 90 mg L⁻¹ for wildtype sfGFP and 12.5 mg L⁻¹ in the presence of 1 mM exogenously fed 5HTP. The mass of sfGFP-5HTP from the optimized biosynthetic condition was 27634 Da, which is in good agreement with the mass of sfGFP-5HTP expressed by feeding 5HTP (27635 Da, Fig. 2E and S8). We observed tryptophan incorporation for the sfGFP purified from control cells without 5HTP feeding (Fig. S9). Using 2x YT medium, we observed similar sfGFP-5HTP yields of 22.5 mg L⁻¹ and 20.1 mg L⁻¹ when 5HTP was exogenously fed or biosynthesized, respectively. To demonstrate that our 5HTP autonomous cells provide a general platform to generate 5HTP-containing proteins, we also expressed myoglobin protein containing 5HTP at position 99 from autonomous cells with good yield (Fig. S10 and 11).

Production of Antibody Fragment-fluorophore Conjugate from Autonomous Cells

Previously, we have developed a rapid chemoselective azo-coupling reaction (CRACR) for site-selective labeling of proteins containing 5HTP (Fig. 3A).⁵⁰⁻⁵¹ CRACR enables a rapid labeling reaction between 5HTP and various aromatic diazonium ions ($>100 \text{ M}^{-1} \text{ s}^{-1}$ for 4-carboxydiazonium).⁵⁰ To evaluate whether 5HTP-containing proteins isolated from the completely autonomous *E. coli* strain can be site-specifically modified using the CRACR reaction, the single-chain variable fragment (scFv) of anti-human epidermal growth factor 2 (HER2) antibody was used (anti-HER2-scFv).⁵²⁻⁵⁵ As a highly expressed membrane protein that regulates cell proliferation, invasion, and angiogenesis in breast and ovarian cancer cells, HER2 has been widely used as a therapeutic target.⁵⁶ To produce 5HTP-containing anti-HER2-scFv in *E. coli*, we first generated the pET22b-T5-scFv*-XcP4H-MH₄R plasmid encoding XcP4H, the BH₄ recycle pathway, and anti-HER2-scFv with an amber codon at position 113. *E. coli* BL21(DE3) cells transformed with pET22b-T5-scFv*-XcP4H-MH₄R and pUltra-5HTP were grown in M9G for 16 h in 30°C shaker to allow protein expression. To demonstrate the efficiency of the biosynthetic system, *E. coli* BL21(DE3) cells with pET22b-T5-scFv* and pUltra-5HTP, fed with and without 1 mM 5HTP were used as controls. 10.6 mg/L anti-HER2-5HTP was produced from the completely autonomous *E. coli* cells, compared with yield of 70 mg L⁻¹ for wildtype anti-HER2 and 13.8 mg L⁻¹ for that purified from control cells in the presence of 1 mM exogenously fed 5HTP (Fig. 2D). ESI-MS analysis of anti-HER2-5HTP produced by 5HTP biosynthesis revealed a mass of 27601 Da, in agreement with a calculated mass of 27601 Da, thus confirming the successful incorporation of biosynthesized 5HTP into anti-HER2-scFv (Fig. 3B, S12, and 13). In the absence of exogenous 5HTP, misincorporation of Trp was observed (Fig. S9). Next, we used CRACR to label the anti-HER2-5HTP with 5 coumarin-diazonium (Cou-Dz), which exhibits maximal absorption at 297 nm and emission at 386 nm (Fig. S14). The previously reported preparation method for the diazonium salts was further optimized for this labeling reaction. We found that by dissolving Cou-Dz in DMSO, followed by the addition of sodium nitrite and *p*-toluenesulfonic acid, we could generate diazonium salts at a higher yield, which enables us to achieve the site-specific labeling of anti-HER2-5HTP with a smaller quantity of Cou-Dz solution (Fig. S15). To our delight, the resulting anti-HER2-Cou conjugate exhibited strong fluorescence, with more than 90 % labeling of the 5HTP site, as determined by whole-protein ESI-MS analysis (Fig. 3B and S12). With the anti-HER2-Cou conjugate in hand, we then tested if this conjugate retained its function. HER2-positive SK-BR-3 cells and HER2-negative MDA-MB-468 cells were incubated with 100 nM anti-HER2-Cou conjugate and prepared for imaging. Confocal imaging revealed that this conjugate retains its bioactivity, as evidenced by its ability to immunolabel HER2-positive SK-BR-3 cells, but not HER2-negative MDA-MB-468 cells (Fig. 3C).

Detection of Oxidative Stress with Autonomous Cells

Taking advantage of 5HTP biosynthesis pathway, the autonomous bacterium could be redirected to monitor oxidative stress. Previous studies illustrate that the cellular level of reduced cofactors is associated with the oxidative stress.⁵⁷ Thus, our hypothesis is that consumption of reduced cofactor triggered by oxidative stress would make it relatively unavailable in the biosynthesis of 5HTP, which leads to a decrease in production of 5HTP-containing sfGFP (Fig. 4A). As a proof of concept study, we chose hydrogen peroxide as an

inducer for oxidative stress.^{58,59} To test this hypothesis, we first measured intracellular 5HTP concentration after overnight growth of autonomous bacterium containing pET22b-T5-sfGFP*-XcP4H-MH₄R and pUltra-5HTP plasmids in the presence of hydrogen peroxide with different concentrations. Identical control experiments were set up with control cells harboring pET22b-T5-sfGFP*, pUltra-5HTP plasmids, and 1 mM exogenously fed 5HTP. To our delight, we observed a gradually decreasing intracellular concentration of 5HTP in response to the increasing concentration of H₂O₂ (Fig. S16). In contrast, 5HTP cellular concentration maintains stable in the control cells fed with 1 mM 5HTP after H₂O₂ treatment (Fig. S16). As the precursor for 5HTP biosynthesis, tryptophan concentration maintains stable in the presence of H₂O₂ (Fig. S17). After confirming the decreased intracellular level of 5HTP associated with the H₂O₂ treatment, we monitored the fluorescence of cells to explore the usage of autonomous cells as a reactive oxygen species (ROS) indicator. Addition of 0.5 mM H₂O₂ to these autonomous cells triggers a significant decrease of green fluorescence (Fig. 4B, C, and S18), while control cells did not show a significant fluorescence response to H₂O₂ under similar conditions. In general, higher sensitivity of sfGFP fluorescence was observed with autonomous cells than that of control cells in response to surrounding hydrogen peroxide from 0.2 mM to 0.5 mM (Fig. 4B and S18). Thiourea is known as efficient hydroxyl radical scavengers. We found that the addition of hydroxyl radical quencher thiourea will significantly reduce the sensitivity of both autonomous cells and control cells to H₂O₂ (Fig. 4B and S18). To demonstrate that detectable concentrations of H₂O₂ will not influence the growth of autonomous cells, we measured the growth curves of autonomous cells and control cells in the presence of H₂O₂ (Fig. S19). There is no significant difference between the doubling times of control cells and autonomous cells after H₂O₂ treatments. These data demonstrate that living autonomous cells with a 21st amino acid can be used for the detection of oxidative stress in bacteria.

DISCUSSION

In conclusion, we have generated a completely autonomous bacterium that utilizes 5HTP as a 21st amino acid for protein synthesis. The resulting strain has the endogenous ability to biosynthesize 5HTP at a high level and genetically incorporate it into proteins with high efficiency and fidelity. Furthermore, we have demonstrated that 5HTP-containing proteins isolated from this engineered strain can be site-specifically labelled with aromatic diazonium ion compounds using the CRACR reaction. More importantly, the autonomous bacterium could also be used as a living oxidative indicator. Compared to cells exogenously fed with 5HTP, these autonomous cells with the endogenous ability to biosynthesize and incorporate 5HTP are more sensitive to oxidative stress.

So far, more than 300 ncAAs has been genetically added to proteins, which provides a powerful tool to probe the structure and function of proteins.^{4-6,14,15} Despite the broad application of this technology, the genetic incorporation of these ncAAs requires the exogenous feeding of chemically-synthesized ncAAs and successful uptake of these ncAAs by cells, which significantly restricts the utility of this approach in protein evolution studies and the preparation of novel protein/cell-based drugs. For example, studies indicate that the production yield of ncAA-containing protein by exogenous feeding is much lower than the yield when the ncAAs were biosynthesized.³² Furthermore, genetic ncAA incorporation has

been expanded into eukaryotic animals, including worms, fruit flies, zebrafish, and mice.⁶⁰⁻⁶⁹ But, most ncAAs have a poor pharmacokinetics *in vivo* and were cleared up within several hours. Thus, creation of cells with the endogenous ability to biosynthesize an ncAA and to use it for protein synthesis will enhance the efficiency of generation of ncAA-containing proteins as well as facilitate the applications of genetic code expansion at a more global whole-organism level. By combining bioorthogonal translational components, expanded genetic alphabets, and ncAA biosynthetic machinery, we hope to generate more autonomous cells and animals with expanded genetic codes. This will provide powerful tools for the evolution of novel proteins containing ncAAs, as well as for the preparation of new therapeutic proteins capable of revolutionizing modern medicine.

EXPERIMENTAL PROCEDURES

Full experimental procedures are provided in the Supplemental Information.

Resource Availability

Lead Contact—Further information and requests of resources should be directed to and will be fulfilled by the Lead Contact, Prof. Han Xiao (han.xiao@rice.edu).

Materials Availability—The plasmids generated in this study will be made available on request, but we may require payments and/or completed material transfer form if there are potentials for commercial applications.

Data and Code Availability—This study did not generate datasets or code.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the Cancer Prevention Research Institute of Texas (CPRIT RR170014), NIH (R35-GM133706), the Robert A. Welch Foundation (C-ig7o), the John S. Dunn Foundation Collaborative Research Award, and the Hamill Innovation Award. H. X. is a Cancer Prevention & Research Institute of Texas (CPRIT) Scholar in Cancer Research.

REFERENCES

1. Macek B, Forchhammer K, Hardouin J, Weber-Ban E, Grangeasse C, and Mijakovic I (2019). Protein posttranslational modifications in bacteria. *Nat. Rev. Microbiol* 17, 651–664. [PubMed: 31485032]
2. Frost JR, Smith JM, and Fasan R (2013). Design, synthesis, and diversification of ribosomally derived peptide macrocycles. *Curr. Opin. Struct. Biol* 23, 571–580. [PubMed: 23856642]
3. Walsh C (2006). *Posttranslational Modification of Proteins: Expanding Nature's Inventory* (Roberts and Company Publishers).
4. Wang L, Xie J, and Schultz PG (2006). Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct* 35, 225–249. [PubMed: 16689635]
5. Ambrogelly A, Palioura S, and Söll D (2007). Natural expansion of the genetic code. *Nat. Chem. Biol* 3, 29–35. [PubMed: 17173027]

6. Liu CC, and Schultz PG (2010). Adding new chemistries to the genetic code. *Annu. Rev. Biochem* 79, 413–444. [PubMed: 20307192]
7. Chin JW (2017). Expanding and reprogramming the genetic code. *Nature* 550, 53–60. [PubMed: 28980641]
8. Dien VT, Morris SE, Karadeema RJ, and Romesberg FE (2018). Expansion of the genetic code via expansion of the genetic alphabet. *Curr. Opin. Chem. Biol* 46, 196–202. [PubMed: 30205312]
9. Wang L (2017). Genetically encoding new bioreactivity. *New Biotechnol.* 38, 16–25.
10. Lang K, and Chin JW (2014). Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev* 114, 4764–4806. [PubMed: 24655057]
11. Wang L, Brock A, Herberich B, and Schultz PG (2001). Expanding the Genetic Code of *Escherichia coli*. *Science* 292, 498–500. [PubMed: 11313494]
12. Johnson JA, Lu YY, Van Deventer JA, and Tirrell DA (2010). Residue-specific incorporation of noncanonical amino acids into proteins: recent developments and applications. *Curr. Opin. Chem. Biol* 14, 774–780. [PubMed: 21071259]
13. Wan W, Tharp JM, and Liu WR (2014). Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta* 1844, 1059–1070. [PubMed: 24631543]
14. Xiao H, and Schultz PG (2016). At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code. *Cold Spring Harb. Perspect. Biol.* a023945. [PubMed: 27413101]
15. Young DD, and Schultz PG (2018). Playing with the Molecules of Life. *ACS Chem. Biol* 13, 854–870. [PubMed: 29345901]
16. Niu W, and Guo J (2013). Expanding the chemistry of fluorescent protein biosensors through genetic incorporation of unnatural amino acids. *Mol. Biosyst* 9, 2961–2970. [PubMed: 24080788]
17. Hao Z, Hong S, Chen X, and Chen PR (2011). Introducing Bioorthogonal Functionalities into Proteins in Living Cells. *Acc. Chem. Res* 44, 742–751. [PubMed: 21634380]
18. Li J, Wang J, and Chen P (2012). Unnatural amino acid mediated protein bioorthogonal labeling. *Acta Chim. Sin* 70, 1439–1445.
19. Chatterjee A, Xiao H, and Schultz PG (2012). Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci* 109, 14841–14846. [PubMed: 22927411]
20. Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, and Schultz PG (2004). An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci* 101, 7566–7571. [PubMed: 15138302]
21. Neumann H, Wang K, Davis L, Garcia-Alai M, and Chin JW (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441–444. [PubMed: 20154731]
22. Niu W, Schultz PG, and Guo J (2013). An Expanded Genetic Code in Mammalian Cells with a Functional Quadruplet Codon. *ACS Chem. Biol* 8, 1640–1645. [PubMed: 23662731]
23. Fredens J, Wang K, de la Torre D, Funke LFH, Robertson WE, Christova Y, Chia T, Schmied WH, Dunkelmann DL, Beránek V, et al. (2019). Total synthesis of *Escherichia coli* with a recoded genome. *Nature* 569, 514–518. [PubMed: 31092918]
24. Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, Lee D, Huang CLV, Chandrasegaran S, Cai Y, et al. (2017). Design of a synthetic yeast genome. *Science* 355, 1040–1044. [PubMed: 28280199]
25. Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, Dymond JS, Kuang Z, Scheifele LZ, Cooper EM, et al. (2014). Total synthesis of a functional designer eukaryotic chromosome. *Science* 344, 55–58. [PubMed: 24674868]
26. Zhang Y, Ptacin JL, Fischer EC, Aerni HR, Caffaro CE, San Jose K, Feldman AW, Turner CR, and Romesberg FE (2017). A semisynthetic organism that stores and retrieves increased genetic information. *Nature* 551, 644–647. [PubMed: 29189780]
27. Walsh CT, O'Brien RV, and Khosla C (2013). Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds. *Angew. Chem. Int. Ed Engl* 52, 7098–7124. [PubMed: 23729217]

28. Kudo F, Miyanaga A, and Eguchi T (2019). Structural basis of the nonribosomal codes for nonproteinogenic amino acid selective adenylation enzymes in the biosynthesis of natural products. *J. Ind. Microbiol. Biotechnol* 46, 515–536. [PubMed: 30291534]
29. Zou H, Li L, Zhang T, Shi M, Zhang N, Huang J, and Xian M (2018). Biosynthesis and biotechnological application of noncanonical amino acids: Complex and unclear. *Biotechnol. Adv* 36, 1917–1927. [PubMed: 30063950]
30. Mehl RA, Anderson JC, Santoro SW, Wang L, Martin AB, King DS, Horn DM, and Schultz PG (2003). Generation of a Bacterium with a 21 Amino Acid Genetic Code. *J. Am. Chem. Soc* 125, 935–939. [PubMed: 12537491]
31. Chen Y, Loredó A, Gordon A, Tang J, Yu C, Ordoñez J, and Xiao H (2018). A noncanonical amino acid-based relay system for site-specific protein labeling. *Chem. Commun* 54, 7187–7190.
32. Zhang MS, Brunner SF, Huguenin-Dezot N, Liang AD, Schmied WH, Rogerson DT, and Chin JW (2017). Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat. Methods* 14, 729–736. [PubMed: 28553966]
33. Marchand JA, Neugebauer ME, Ing MC, Lin C-I, Pelton JG, and Chang MCY (2019). Discovery of a pathway for terminal-alkyne amino acid biosynthesis. *Nature* 567, 420–424. [PubMed: 30867596]
34. Neugebauer ME, Sumida KH, Pelton JG, McMurphy JL, Marchand JA, and Chang MCY (2019). A family of radical halogenases for the engineering of amino acid-based products. *Nat. Chem. Biol* 15, 1009–1016. [PubMed: 31548692]
35. Kim S, Sung BH, Kim SC, and Lee HS (2018). Genetic incorporation of L-dihydroxyphenylalanine (DOPA) biosynthesized by a tyrosine phenol-lyase. *Chem. Commun* 54, 3002–3005.
36. Birdsall TC (1998). 5-Hydroxytryptophan: a clinically-effective serotonin precursor. *Altern. Med. Rev. J. Clin. Ther* 3, 271–280.
37. Nakamura K, and Hasegawa H (2009). Production and Peripheral Roles of 5-HTP, a Precursor of Serotonin. *Int. J. Tryptophan Res. IJTR* 2, 37–43. [PubMed: 22084581]
38. Hara R, and Kino K (2013). Enhanced synthesis of 5-hydroxyl-tryptophan through tetrahydropterin regeneration. *AMB Express* 3, 70. [PubMed: 24321061]
39. Lin Y, Sun X, Yuan Q, and Yan Y (2014). Engineering Bacterial Phenylalanine 4-Hydroxylase for Microbial Synthesis of Human Neurotransmitter Precursor 5-Hydroxytryptophan. *ACS Synth. Biol* 3, 497–505. [PubMed: 24936877]
40. Mora-Villalobos J-A, and Zeng A-P (2017). Protein and pathway engineering for the biosynthesis of 5-hydroxytryptophan in *Escherichia coli*. *Eng. Life Sci* 17, 892–899. [PubMed: 32624837]
41. Wang J, Shen X, Rey J, Yuan Q, and Yan Y (2018). Recent advances in microbial production of aromatic natural products and their derivatives. *Appl. Microbiol. Biotechnol* 102, 47–61. [PubMed: 29127467]
42. Zhang J, Wu C, Sheng J, and Feng X (2016). Molecular basis of 5-hydroxytryptophan synthesis in *Saccharomyces cerevisiae*. *Mol. Biosyst* 12, 1432–1435. [PubMed: 27008988]
43. Satoh Y, Tajima K, Muneoka M, Keasling JD, and Lee TS (2012). Engineering of L-tyrosine oxidation in *Escherichia coli* and microbial production of hydroxytyrosol. *Metab. Eng* 14, 603–610. [PubMed: 22948011]
44. Xiao H, Peters FB, Yang P-Y, Reed S, Chittuluru JR, and Schultz PG (2014). Genetic Incorporation of Histidine Derivatives Using an Engineered Pyrrolysyl-tRNA Synthetase. *ACS Chem. Biol* 9, 1092–1096. [PubMed: 24506189]
45. Chen Y, Wu K-L, Tang J, Loredó A, Clements J, Pei J, Peng Z, Gupta R, Fang X, and Xiao H (2019). Addition of Isocyanide-Containing Amino Acids to the Genetic Code for Protein Labeling and Activation. *ACS Chem. Biol* 14, 2793–2799. [PubMed: 31682403]
46. Ellefson JW, Meyer AJ, Hughes RA, Cannon JR, Brodbelt JS, and Ellington AD (2014). Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat. Biotechnol* 32, 97–101. [PubMed: 24185096]
47. Chatterjee A, Sun SB, Furman JL, Xiao H, and Schultz PG (2013). A Versatile Platform for Single and Multiple-Unnatural Amino Acid Mutagenesis in *Escherichia coli*. *Biochemistry* 52, 1828–1837. [PubMed: 23379331]

48. Crawford IP, and Stauffer GV (1980). Regulation of Tryptophan Biosynthesis. *Annu. Rev. Biochem* 49, 163–195. [PubMed: 6996563]
49. Niu H, Li R, Liang Q, Qi Q, Li Q, and Gu P (2019). Metabolic engineering for improving ltryptophan Production in *Escherichia coli*. *J. Ind. Microbiol. Biotechnol* 46, 55–65. [PubMed: 30426284]
50. Addy PS, Erickson SB, Italia JS, and Chatterjee A (2017). A Chemoselective Rapid Azo-Coupling Reaction (CRACR) for Unclickable Bioconjugation. *J. Am. Chem. Soc* 139, 11670–11673. [PubMed: 28787141]
51. Addy PS, Zheng Y, Italia JS, and Chatterjee A (2019). A “Quencherogenic” Chemoselective Protein Labeling Strategy. *Chembiochem Eur. J. Chem. Biol* 20, 1659–1663.
52. Ahmad ZA, Yeap SK, Ali AM, Ho WY, Alitheen NBM, and Hamid M (2012). scFv antibody: principles and clinical application. *Clin. Dev. Immunol* 2012, 980250. [PubMed: 22474489]
53. Du J, Cao Y, Liu Y, Wang Y, Zhang Y, Fu G, Zhang Y, Lu L, Luo X, Kim CH, et al. (2017). Engineering Bifunctional Antibodies with Constant Region Fusion Architectures. *J. Am. Chem. Soc* 139, 18607–18615. [PubMed: 29186655]
54. Zhao Y, Wang QJ, Yang S, Kochenderfer JN, Zheng Z, Zhong X, Sadelain M, Eshhar Z, Rosenberg SA, and Morgan RA (2009). A Herceptin-Based Chimeric Antigen Receptor with Modified Signaling Domains Leads to Enhanced Survival of Transduced T Lymphocytes and Antitumor Activity. *J. Immunol. Baltim. Md* 1950 183, 5563–5574.
55. Carter P, Presta L, Gorman CM, Ridgway JB, Henner D, Wong WL, Rowland AM, Kotts C, Carver ME, and Shepard HM (1992). Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc. Natl. Acad. Sci. U. S. A* 89, 4285–4289. [PubMed: 1350088]
56. Arteaga CL, Sliwkowski MX, Osborne CK, Perez EA, Puglisi F, and Gianni L (2012). Treatment of HER2-positive breast cancer: current status and future perspectives. *Nat. Rev. Clin. Oncol* 9, 16–32.
57. Kohanski MA, Dwyer DJ, Hayete B, Lawrence CA, and Collins JJ (2007). A Common Mechanism of Cellular Death Induced by Bactericidal Antibiotics. *Cell* 130, 797–810. [PubMed: 17803904]
58. Halliwell B, and Gutteridge JM (1984). Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochem. J* 219, 1–14. [PubMed: 6326753]
59. Imlay JA (2015). Diagnosing oxidative stress in bacteria: not as easy as you might think. *Curr. Opin. Microbiol* 24, 124–131. [PubMed: 25666086]
60. Young TS, Ahmad I, Brock A, and Schultz PG (2009). Expanding the genetic repertoire of the methylotrophic yeast *Pichia pastoris*. *Biochemistry* 48, 2643–2653. [PubMed: 19265424]
61. Greiss S, and Chin JW (2011). Expanding the Genetic Code of an Animal. *J. Am. Chem. Soc* 133, 14196–14199. [PubMed: 21819153]
62. Bianco A, Townsley FM, Greiss S, Lang K, and Chin JW (2012). Expanding the genetic code of *Drosophila melanogaster*. *Nat. Chem. Biol* 8, 748–750. [PubMed: 22864544]
63. Li F, Zhang H, Sun Y, Pan Y, Zhou J, and Wang J (2013). Expanding the Genetic Code for Photoclick Chemistry in *E. coli*, Mammalian Cells, and *A. thaliana*. *Angew. Chem. Int. Ed* 52, 9700–9704.
64. Parrish AR, She X, Xiang Z, Coin I, Shen Z, Briggs SP, Dillin A, and Wang L (2012). Expanding the genetic code of *Caenorhabditis elegans* using bacterial aminoacyl-tRNA synthetase/tRNA pairs. *ACS Chem. Biol.* 7, 1292–1302. [PubMed: 22554080]
65. Ernst RJ, Krogager TP, Maywood ES, Zanchi R, Beránek V, Elliott TS, Barry NP, Hastings MH, and Chin JW (2016). Genetic code expansion in the mouse brain. *Nat. Chem. Biol* 12, 776–778. [PubMed: 27571478]
66. Liu J, Hemphill J, Samanta S, Tsang M, and Deiters A (2017). Genetic Code Expansion in Zebrafish Embryos and Its Application to Optical Control of Cell Signaling. *J. Am. Chem. Soc* 139, 9100–9103. [PubMed: 28657738]
67. Brown W, Liu J, and Deiters A (2018). Genetic Code Expansion in Animals. *ACS Chem. Biol* 13, 2375–2386. [PubMed: 30125487]
68. Wang L (2017). Engineering the Genetic Code in Cells and Animals: Biological Considerations and Impacts. *Acc. Chem. Res* 50, 2767–2775. [PubMed: 28984438]

69. Chen Y, Ma J, Lu W, Tian M, Thauvin M, Yuan C, Volovitch M, Wang Q, Holst J, Liu M, et al. (2017). Heritable expansion of the genetic code in mouse and zebrafish. *Cell Res.* 27, 294–297 [PubMed: 27934867]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Construction of autonomous organism with 5-hydroxytryptophan as its 21st amino acid
- Use of autonomous organism for preparing proteins with site-specific modifications
- Creation of a Living oxidative stress indicator

The Bigger Picture

Most living organisms use 20 canonical amino acids to encode proteins. Despite this universality, nature appears to require an additional level of complexity in protein structure and function. Thus, there is tremendous excitement surrounding the addition of noncanonical amino acids to the genetic code of organisms. In this study, we create the first autonomous organism with 5-hydroxytryptophan as its 21st amino acid. The resulting cells harbor the endogenous ability to biosynthesize 5-hydroxytryptophan and site-specifically incorporate it into proteins. The utility of this unnatural organism is further demonstrated by our preparation of site-specific protein conjugates and our creation of a living indicator for oxidative stress. Creation of unnatural organisms with expanded genetic codes will allow us to examine the evolutionary consequences of adding new building blocks. Moreover, these new technologies will allow generation of proteins and organisms with novel and useful activities.

Figure 1. Generation of a completely autonomous *E. coli* strain with the 21st amino acid, 5HTP. 5HTP-containing proteins prepared using this strain can be site-specifically modified with aromatic diazonium ions by a recently reported azo-coupling reaction. These completely autonomous cells with the endogenous ability to biosynthesize and incorporate 5HTP into proteins were further used as a living indicator for the oxidative stress. DHMR: dihydromonapterin reductase; PCD: pterin-4 α -carbinolamine dehydratase; 4a-Hthb: 4a-Hydroxytetrahydrobiopterin; scFv: single-chain variable fragment.

Figure 2. Generation of completely autonomous *E. coli* with 5HTP.

(A) Comparison of the efficiency of different hydroxylases to produce 5HTP-containing sfGFPs. Cells fed with (+) or without (-) 1 mM 5HTP are used as controls. (B) The role of XcP4H, MH4 recycling pathway, and *Sc*TrpRS/tRNA in producing 5HTP-containing sfGFP. (C) Optimization of 5HTP biosynthesis by increasing tryptophan concentration in the expression medium. Cells with (+) and without (-) 1 mM exogenously fed 5HTP are used as controls. (D) SDS-PAGE analysis of sfGFP and anti-HER2-scFv expressed in M9G in the presence (+) or absence (-) of 5HTP, or when inducing the 5HTP biosynthetic plasmid (bio) without adding Trp. (E) Mass spectra of sfGFP-151-5HTP expressed from *E. coli* with exogenously fed 5HTP or biosynthesized 5HTP, see also Figure S8. Error bars represent the standard deviation.

Figure 3. Site-specific modification of 5HTP-containing proteins isolated from completely autonomous *E. coli* cells.

(A) CRACR reaction of 5HTP-containing proteins with aromatic diazonium ions. (B) Mass spectra of anti-HER2-5HTP before (left) and after labeling (right). (C) Binding of anti-HER2-Cou in SK-BR-3 and MD-MBA-468 cells visualized by confocal microscopy. Cells were incubated with 100 nM anti-HER2-Cou (blue) before imaging. Scale bar = 50 μm .

Figure 4. Indication of oxidative stress with autonomous cells.

(A) The increase of reactive oxygen species (ROS) would lead to a significant decrease in GFP fluorescence via a hypothesized decrease in reduced cofactors (NAD(P)H) and subsequent decrease in biosynthesized 5HTP. (B) Sensitivity-dose response curve of autonomous cells and control cells after the addition of various concentrations of H₂O₂ in the presence or absence of 30 mM hydroxyl radical quencher thiourea. (C) Confocal microscopic imaging of autonomous cells and normal cells after induction of protein expression in the presence or absence of 0.5 mM H₂O₂. ** P 0.01; *** P 0.001; **** P 0.0001; n.s. P > 0.05; scale bar: 10 μm.

Table 1.Cellular concentration (μM) of 5HTP in *E. coli*.

Entry	MH4 recycle ^a	XcP4H ^b	5HTP ^c	Trp	Conc. μM ^d
1	–	–	–	–	0.0
2	–	–	+	–	907.5
3	+	–	–	–	0.9
4	–	+	–	–	8.4
5	+	+	–	–	59.89
6	+	+	–	+	122.61

^aCells with the plasmid encoding dihydromonapterin reductase (DHMR) and pterin-4 α -carbinolamine dehydratase (PCD) for tetrahydromonapterin (MH4) recycle.

^bCells with plasmid encoding *Xanthomonas campestris* phenylalanine 4-hydroxylase (XcP4H).

^cCells exogenously fed with 1 mM 5-hydroxytryptophan (5HTP).

^dCells were grown in the M9-glucose minimal medium.