Non-Standard Amino Acid Incorporation in the New Era of Recoded Genomes

Aditya M. Kunjapur[1*] (aditya_kunjapur@hms.harvard.edu), Erkin Kuru,[1] Devon Stork,[1] Kamesh Narasimhan,[1] Po-yi Huang,[1] Daniel J. Mandell,[1] John Aach,[1], and **George M. Church[1]**

[1]Harvard Medical School, Department of Genetics, Boston, Massachusetts, USA

http://arep.med.harvard.edu

Project Goals: The advent of scalable genome engineering and assembly technologies has inspired efforts aimed at removing all instances of particular codons by synonymous substitution throughout a genome. In the resulting recoded organisms, codons are free for reassignment to amino acids beyond the canonical twenty without competition from endogenous tRNAs or release factors. This increases the ease of producing proteins containing non-standard amino acids (nsAAs), which have previously been shown to have numerous desirable properties. Here, we review our recent efforts to use nsAAs for new functions, such as for biocontainment. Moving forward, we are interested in improving the activity and selectivity of amino-acyl tRNA synthetases (AARSs) and other components of translational machinery for incorporation of D-amino acids and for novel nsAAs.

During the last decade, the Church Lab has been at the forefront of developing genome engineering technologies for recoding. The development of multiplex automatable genome engineering (MAGE) provided a foundational approach to recoding at the scale of hundreds of codons (*1*). Conjugative assembly genome engineering (CAGE) enabled the parallelization of recoding efforts within portions of an individual genome that could later be hierarchically reassembled (*2*). A major product of these efforts was the creation of an *Escherichia coli* strain devoid of amber (UAG) codons and release factor 1, known as C321.ΔA (*3*). Since then, MAGE and CAGE have also been used to construct an *E. coli* strain without the rare arginine codons AGA and AGG in any essential genes (*4*). Moving forward, genome synthesis and assembly are being used to construct a 57-codon strain, which is nicely reviewed in another poster (*5*).

Non-standard amino acids (nsAAs) can site-specifically be incorporated into proteins *in vivo* across prokaryotes and eukaryotes using orthogonal pairs of amino acyl tRNA synthetases (AARSs) and orthogonal tRNA (*6–9*). Some of the many functions that nsAAs have been shown to provide include photocrosslinking (*10*, *11*), functionalization (*12*), structure determination (*13*), fluorescence (*14*, *15*), and metal binding (*16*). Despite numerous demonstrations of nsAA potential, nsAAs use has been limited by the absence of a dedicated codon that is free of competition from endogenous translation machinery. In addition, we recently used the engineered C321.ΔA strain to demonstrate a new use for nsAAs: biocontainment dependent on the nsAA biphenylalanine (bipA) (*17*). Our ability to biocontain is especially important given that recoding can bestow an organism with increased virus resistance (*3*).

As we work toward a 57-codon genome that will include a dedicated codon for biocontainment,

we intend to first improve the activity and selectivity of BipARS, which is the AARS for bipA. Selectivity at the AARS level will be vital to ensuring the fidelity of translation during the envisioned simultaneous use of multiple types of nsAAs, and we are interested in ways to distinguish nsAAs at the molecular level *in vivo*. In addition, we are looking towards other nsAAs with interests in fluorogenic and D-amino acids (D-AAs). D-AA-containing proteins have potential to greatly expand protein conformation space (*18*) and thereby enable novel protein functions, and can also enhance protein thermostability (*19*). D-AAs are discriminated from L-AAs at 3 steps in core translation machinery: aminoacylation of tRNAs by aminoacyl-tRNA synthetases (aaRSes), formation of ternary complexes with EF-Tu-GTP and their delivery to the ribosome, and the ribosome's own catalysis of peptide bond formation.

References

1. H. H. Wang *et al.*, Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. **460**, 894–898 (2009).
2. F. J. Isaacs *et al.*, Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science (80-.).* **333**, 348–353 (2011).
3. M. J. Lajoie *et al.*, Genomically Recoded Organisms Expand Biological Functions. *Science (80-.).* **342**, 357–360 (2013).
4. M. G. Napolitano *et al.*, Emergent rules for codon choice elucidated by editing rare arginine codons in Escherichia coli. *Proc. Natl. Acad. Sci.* **113**, E5588–E5597 (2016).
5. N. Ostrov *et al.*, Design, synthesis, and testing toward a 57-codon genome. *Science (80-.).* **353**, 819–822 (2016).
6. L. Wang, A. Brock, B. Herberich, P. G. Schultz, Expanding the genetic code of Escherichia coli. *Science (80-.).* **292**, 498–500 (2001).
7. J. W. Chin *et al.*, An expanded eukaryotic genetic code. *Science (80-.).* **301**, 964–967 (2003).
8. L. Wang, J. Xie, P. G. Schultz, EXPANDING THE GENETIC CODE. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 225–249 (2006).
9. J. W. Chin, Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annu. Rev. Biochem*. **83**, 379–408 (2014).
10. J. W. Chin, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci.* **99**, 11020–11024 (2002).
11. J. W. Chin *et al.*, Addition of p-Azido-l-phenylalanine to the Genetic Code of Escherichia coli. *J. Am. Chem. Soc.* **124**, 9026–9027 (2002).
12. L. Wang, Z. Zhang, A. Brock, P. G. Schultz, Addition of the keto functional group to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 56–61 (2003).
13. J. Xie *et al.*, The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination. *Nat Biotechnol*. **22**, 1297–1301 (2004).
14. D. Summerer *et al.*, A genetically encoded fluorescent amino acid. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 9785–9 (2006).
15. J. Wang, J. Xie, P. G. Schultz, A Genetically Encoded Fluorescent Amino Acid. *J. Am. Chem. Soc.* **128**, 8738–8739 (2006).
16. J. Xie, W. Liu, P. G. Schultz, A Genetically Encoded Bidentate, Metal-Binding Amino Acid. *Angew. Chemie*. **119**, 9399–9402 (2007).
17. D. J. Mandell *et al.*, Biocontainment of genetically modified organisms by synthetic protein design. *Nature*. **518**, 55–60 (2015).
18. S. Rana *et al.*, A double catgrip mixed L and D mini protein only 20 residues long. *Bioorg. Med. Chem.*. **15**, 3874-82 (2007).
19. A. Rodriguez-Granillo *et al.*, Computational design of thermostabilizing D-amino acid substitutions. *J. Am. Chem. Soc.* **133**, 18750-9 (2011).

This project has been graciously funded by DOE grant DE-FG02-02ER63445.