Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Next-generation genetic code expansion

Pol Arranz-Gibert[1,2,3], Koen Vanderschuren[1,2,3] and Farren J Isaacs[1,2]

Engineering of the translation apparatus has permitted the site-specific incorporation of nonstandard amino acids (nsAAs) into proteins, thereby expanding the genetic code of organisms. Conventional approaches have focused on porting tRNAs and aminoacyl-tRNA synthetases (aaRS) from archaea into bacterial and eukaryotic systems where they have been engineered to site-specifically encode nsAAs. More recent work in genome engineering has opened up the possibilities of whole genome recoding, in which organisms with alternative genetic codes have been constructed whereby codons removed from the genetic code can be repurposed as new sense codons dedicated for incorporation of nsAAs. These advances, together with the advent of engineered ribosomes and new molecular evolution methods, enable multisite incorporation of nsAAs and nonstandard monomers (nsM) paving the way for the template-directed production of functionalized proteins, new classes of polymers, and genetically encoded materials.

Addresses
[1] Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520, USA
[2] Systems Biology Institute, Yale University, West Haven, CT 06516, USA
[3] Equal contribution

Corresponding author: Isaacs, Farren J. (farren.isaacs@yale.edu)

Current Opinion in Chemical Biology 2018, **46**:xx–yy

This review comes from a themed issue on **Synthetic biology**

Edited by **Michael Ledbetter** and **Floyd E. Romesberg**

https://doi.org/10.1016/j.cbpa.2018.07.020

Introduction

Proteins have evolved to assume diverse cellular functions, such as structural, catalytic or regulatory roles essential for cells. Enzymatic-driven post-translational modifications increase the chemical diversity of proteins well beyond the twenty standard amino acids (AA), thereby considerably enhancing their functions. Research in the field of genetic code expansion – that is, the addition of nonstandard amino acids (nsAAs) to the genetic code – has the potential to greatly expand the chemistries available for protein and peptide synthesis. More than 150 different nsAAs have been incorporated to date for a variety of scientific and biotechnological applications, such as handles for bioconjugation, photo-cross-linking, biophysical probes and genetically encoded "post-translational" modifications [1,2]. Additionally, there is increasing interest in genetically encoded monomers with backbones and chemistries that go beyond the L-α-AA configuration (i.e. non-standard monomers (nsM)), setting the stage for synthesis of entirely new classes of biopolymers and materials.

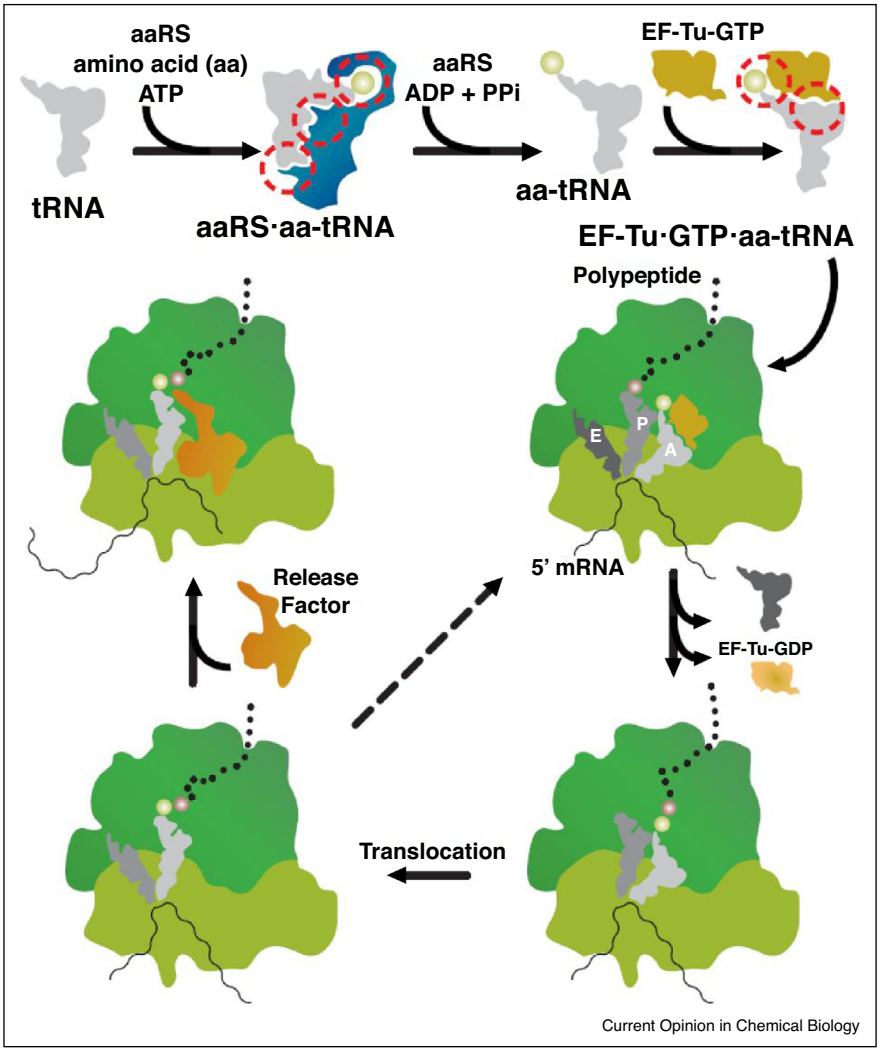
Expansion of the genetic code with nsAAs requires introduction of orthogonal translation systems to enhance the chemical repertoire of proteins while also preserving and not cross-reacting with native protein synthesis. Analogous to standard AAs, the first step is aminoacylation of the tRNA with its cognate AA by an aminoacyl-tRNA synthetases (aaRSs) (Figure 1). Next, the charged tRNA must form a ternary complex with EF-Tu·GTP, where both the tRNA body and the AA contribute to binding [3]. These ternary complexes sample the open codon by partially occupying the ribosomal A site (aminoacyl-tRNA binding site). If the mRNA codon and tRNA anticodon match, the tRNA is released from the ternary complex and the acceptor stem enters the ribosomal peptidyl transferase center (PTC). Here, the polypeptide is transferred from the P site (peptidyl-tRNA site) to the A site tRNA. Next, translocation of the tRNAs releases the empty tRNA from the E site (exit site), and the next codon enters the A site. When a stop codon is presented in the A site, translation is terminated and the polypeptide is released [4]. In prokaryotes, the stop codons UAG and UAA are recognized by the Release Factor 1 (RF1), and UAA and UGA by RF2.

Engineering of the translation apparatus

There are two main strategies that have been employed to genetically encode nsAAs. The first re-purposes an existing aaRS for the incorporation of an nsAA. This strategy involves the expression of the protein of interest (POI) in an auxotrophic expression host, that is, a strain unable to biosynthesize the AA being replaced by the target nsAA. Substitution of the standard AA with a close nsAA analog in the growth media will cause global suppression whereby the nsAA is incorporated in all positions throughout the proteome and POI instead of the standard AA [5,6]. In this strategy, all components are already in place; however, it does not allow the addition of an AA to the genetic code but rather the replacement of a standard one with a nsAA. Furthermore, the substitution of the

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

Figure 1



Translation system and sites of engineering for nsAA incorporation: tRNA, charged by its cognate aaRS with the corresponding AA, is recognized by EF-Tu·GTP; after the initiation of translation, the ternary complex aa-tRNA·EF-Tu·GTP binds the A site; later, the EF-Tu hydrolyzes GTP, changes its conformation and leaves the ribosome; tRNAs at A and P sites are translocated to P and E sites, respectively; the empty tRNA located at the E site is released; at the end of translation, a stop codon is recognized by a release factor that also interacts with acceptor stem of the P site tRNA and triggers hydrolysis of peptidyl-tRNA at the PTC. Red dotted circles show the identity elements for EF-Tu and aaRSs to recognize their cognate AAs and tRNAs. For aaRS, they often localize in the anticodon loop and acceptor stem; however, the variable arm (VA) and D-loop can also be recognized. EF-Tu virtually binds all aa-tRNAs through direct interaction with both AA and tRNA.

standard AA with the nsAA is oftentimes cytotoxic, preventing stable expression with the nsAA.

The second approach permits the site-specific incorporation of an nsAA alongside the full complement of the 20 standard AAs. This strategy involves the development of orthogonal translation systems (OTS) whereby engineered tRNA-aaRS pairs function alongside the native translation machinery. In this scheme, the orthogonal tRNA is not aminoacylated by any of the native aaRSs, and similarly, none of the native tRNAs are charged by the orthogonal aaRS. Two of the most widely used OTSs in bacteria are the natural amber suppressor pyrrolysyl-RS (PylRS)·$tRNA^{Pyl}$ pair derived from *Methanosarcina barkeri* [7], and the tyrosyl-RS (TyrRS)·$tRNA^{Tyr}$ from *Methanococcus jannaschii*. In eukaryotes, both the PylOTS and *E. coli* TyrRS·$tRNA^{Tyr}$ can be used to genetically encode nsAAs [7–9]. Although OTSs are generally employed to encode nsAAs, they may also be used to encode standard AA and thus functionally replace a native tRNA/aaRS pair. In turn, this native tRNA/aaRS pair can be repurposed for genetic code expansion [10•,11]. For example, the endogenous tryptophanyl-RS and tRNA pair from *E. coli* was functionally replaced with its *Saccharomyces cerevisiae* counterpart. Subsequently, the *E. coli* pair was reintroduced as either a UAG or UGA suppressor to

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

expand the genetic code with several tryptophan analogs, such as 5-hydroxytryptophan [10•].

As goals for genetic code expansion aim to go beyond the incorporation of a single nsAA into proteins, the enhancement of activity and orthogonality of OTSs that can operate simultaneously has emerged as a key challenge. Many of the current OTSs are polyspecific and permit charging of multiple nsAAs by the same orthogonal aaRS. To enable site-specific incorporation of multiple genetically encoded nsAAs, it will be imperative to ensure the compatibility and selectivity of the OTSs.

To address these challenges, powerful engineering strategies are needed to optimize properties of OTSs. Advances in molecular evolution and genome engineering have presented new methods to evolve the performance of OTSs in cells. For example, in a recent study by Bryson et al. phage-assisted continuous evolution (PACE) was used to generate 45-fold more active PylRS variants and MjRS variants with improved specificity [12•]. Another study by Amiram et al. combined the use of multiplex automated genome engineering (MAGE) in a genomically recoded organism (GRO; see next section) to produce 25-fold more active MjRS enzymes capable of introducing 30 instances of an nsAA in a single protein [13•]. The MAGE-OTS platform in this study also isolated aaRS enzymes with tunable specificities to >200 nsAAs, including one exclusive for *para*-azido-L-phenylalanine (pAzF).

In general, engineering the aaRS binding pocket to facilitate the accommodation of the desired nsAA may be sufficient; however, generating well-performing OTSs may require evolution of multiple components in parallel. For example, the performance of several aaRS/tRNA pairs was greatly enhanced by engineering the AA binding pocket of the aaRS in conjunction with the binding interface between the tRNA anticodon loop and the aaRS [13•]. In other cases, it became apparent that poor binding characteristics between the $tRNA^{nsAA}$ and EF-Tu greatly limited the incorporation of the nsAA *O*-phosphoserine [14]. Indeed, it was necessary to modify the EF-Tu [14] or the tRNA [15] to tune their binding kinetics for improved incorporation. A similar strategy improved the incorporation efficiency of selenocysteine and bulky residues (e.g. DL-2-anthraquinonylalanine) [16,17]. We anticipate future genetic code expansion efforts will also need to conduct EF-Tu engineering.

Attempts to incorporate bulky or charged nsAAs or even altered backbone monomers (nsM) will likely confront challenges dispersed across multiple steps in the translation process. Such efforts will likely require engineering of the ribosome. Despite the fact that the ribosome limits the incorporation of exotic backbones, *in vitro* work has shown that N^{α}-methylated [18,19], D-α- [20–22] and β-amino acids [23•] and *N*-substituted glycines [24] can be successfully incorporated in peptides. Notably, for several of these monomers, the efficiency was improved by mutating the ribosome in the PTC [18–23•,24]. Recently, the incorporation of β-amino acids has even been reported *in vivo* using an engineered ribosome [25].

A large challenge to ribosome engineering is that the fidelity and activity of the ribosome is inherently tied to protein synthesis and as a result to cell viability. To relieve this interdependency, it has been shown that in *E. coli* an additional, orthogonal ribosome (oRibosome) can be introduced, which is uncoupled from cell viability. The oRibosome has an altered anti Shine-Dalgarno (SD) sequence, which reduces its ability to initiate translation of native mRNA transcripts. Next, orthogonal genes can be introduced with a SD sequence that matches the anti-SD of the oRibosome (Figure 2) [26,27]. Despite these efforts, full orthogonality has not been achieved and complete isolation of oRibosome translation remains a challenge worthy of further pursuit.

A second challenge is that the anti-SD is located in the 16S rRNA, and the 23S rRNA can freely exchange between the native and orthogonal 16S rRNA populations. Recent work has shown that physical tethering of the 16S and 23S rRNAs establishes an oRibosome that will not cross-react with native ribosomes and therefore becomes an isolated ribosomal system [28••,29••]. As a result, this tethered oRibosome can be evolved beyond what native ribosomes can tolerate without impeding native translation and becoming lethal. This establishes a path toward fully isolated ribosomes whose PTC can be mutagenized to accommodate modified backbones and drive new catalysis beyond amide bond formation.

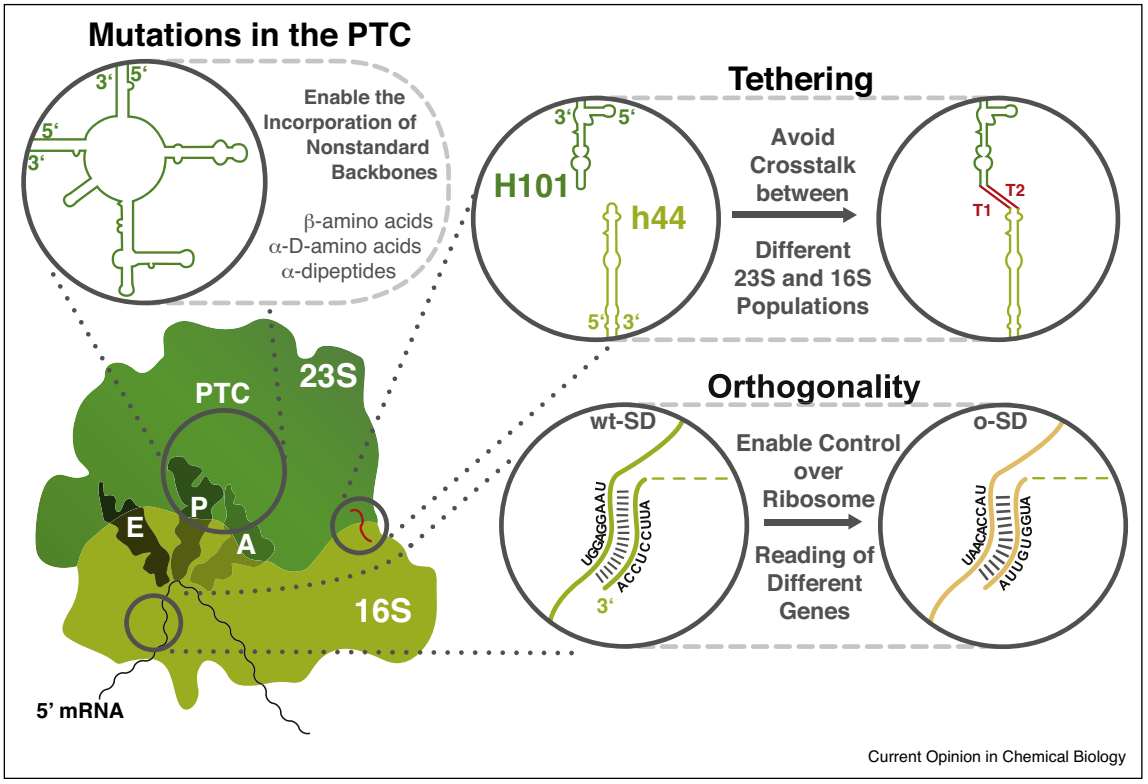
Lastly, cell-free protein synthesis (CFPS) is a unique opportunity to study and manipulate translational systems to produce proteins with nsAAs without the constraints of cell viability [30]. For example, amino acids that cannot pass cell membranes can be readily added in these systems [31]. Even more, if tRNA aminoacylation by a cognate aaRS cannot be achieved, chemical charging or flexizymes can be used to overcome this step [32,33]. Additionally, CFPS systems with defined tRNA composition, either reconstituted from purified components [34] or cell extracts lacking native tRNAs [35], facilitate genetic code reassignment [36].

Genomically recoded organisms (GROs): organisms with alternative genetic codes

An important challenge in genetic code expansion is that native systems typically do not possess open codons that can be dedicated to a nsAA. In other words, the OTS competes with a native tRNA or release factor for the assigned codon. For example, if the tRNA from the OTS decodes the UAG amber stop codon, then it competes

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

Figure 2



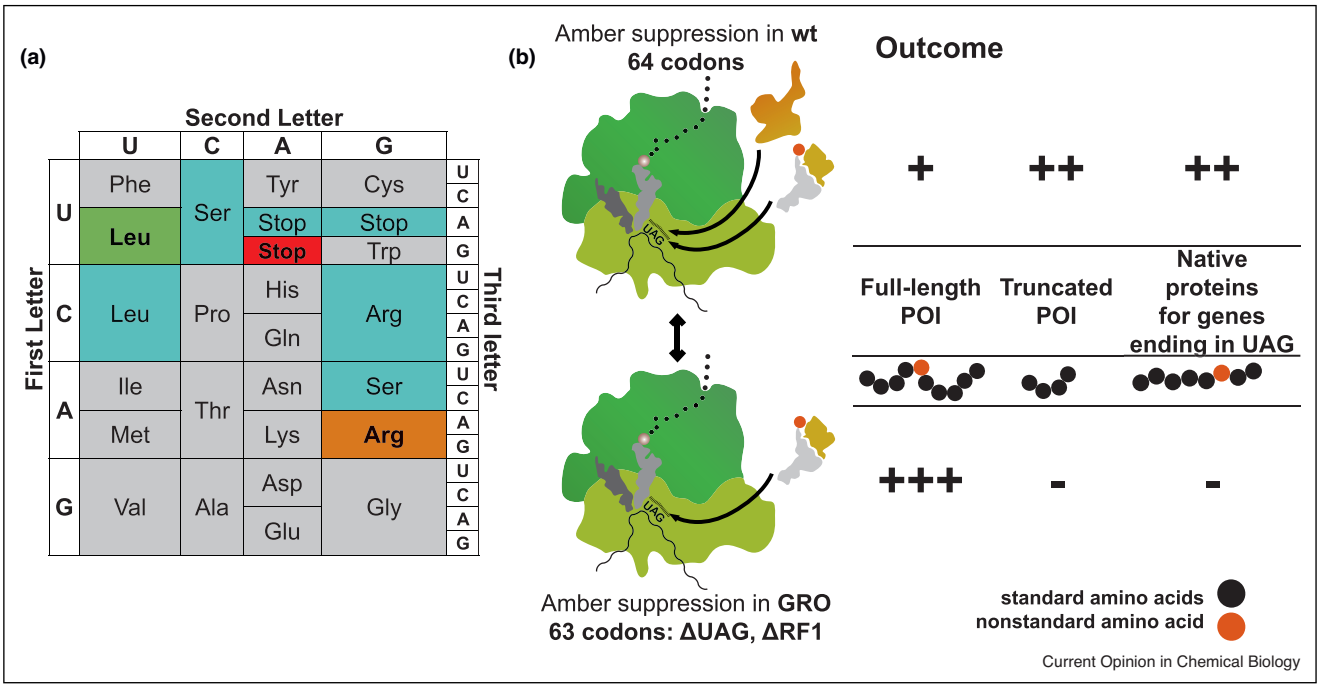
Ribosome engineering. The ribosome has been modified at the peptidyl transferase center (PTC) to enable the incorporation of nonstandard backbones; orthogonal ribosomes, that read different mRNAs, are obtained through changing the wt anti SD-sequence; tethering of 16S and 23S rRNAs avoids crosstalk between different rRNA subunits.

with RF1. In this scenario, a fraction of the POI will be terminated prematurely at the UAG codon, resulting in truncated POI and reduced yields of full-length protein. This issue is magnified for efforts aimed at multisite incorporation of nsAAs. Additionally, the OTS may also suppress native stop codons, leading to misincorporation of the nsAA at more than 300 UAG codons found in native proteins (Figure 3). This misincorporation may result in non-functional proteins and cytotoxicity, further lowering yields of the POI [37•].

Benefiting from recent advances in genome engineering, these issues can be overcome through whole-genome recoding, in which all instances of a codon can be eliminated from the genome. To achieve compatibility of recoding at the translational level, these efforts need to be coordinated with elimination or modification of native translation machinery that decodes the target codon at the ribosome. This was first demonstrated in *E. coli* MG1655 by converting all 321 UAG (amber) codons to UAA (ochre), constructing the first GRO (Figure 3) [38]. For this work, the UAG to UAA mutations were introduced in 32 strains in parallel through MAGE [39] and recoded genomic fragments were then combined through five hierarchical steps with conjugative assembly genome engineering (CAGE) [40]. In addition, RF1 (*prfA* gene) was deleted to eliminate native UAG function. This GRO, named C321.ΔA, exhibited enhanced functions [37], which include blocking of horizontally transferred genetic elements, resistance to multiple viruses [41], biocontainment [42,43], and improved properties for nsAA incorporation [13••]. Using a similar strategy, all 123 essential genes containing AGA and AGG codons (arginine) in *E. coli* were reassigned to CGN (Figure 3) [44]. Also, more aggressive attempts to recode the genome by removing 13 rare codons from 42 highly expressed essential genes have also been pursued using similar methodology [45]. This work has inspired several other ambitious recoding strategies that include sense recoding. One large effort aims to generate a GRO with a 57-codon genome by introducing synthetic DNA fragments (recoding of serine: AGC, AGU; leucine: UUA, UUG; arginine: AGG, AGA; stop: UAG; Figure 3) [46•]. Another project has developed SIRCAS (stepwise integration of rolling circle amplified segments) to construct a *Salmonella typhimurium* (*S. typhimurium*) strain where 1557 leucine codons were replaced with synonymous codons across 176 genes (Figure 3) [47•]. Lastly, a project dedicated to the design of a synthetic *Saccharomyces cerevisiae* genome (Sc 2.0) is

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

Figure 3



Genome Recoding. **(a)** The genetic code has been modified, in several species and diverse codons, to repurpose codon usage: In *E. coli* **all 321 UAG** codons/**all** genes were converted to UAA + **RF1** deleted (red); In *S. cerevisiae* **all UAG** codons/ genes are being converted to UAA (red); In *E. coli* **123 AGR** codons/**all** essential genes were converted to CGN (orange); In *S. typhimurium,* **1557 UUR** codons/**176** genes were converted to CTA/CTG (green); Ongoing project to obtain a 57-codon *E. coli.* Codons left for each amino acid: Leu (CUA, CUC), Ser (UCA, UCC), Arg (CGC, CGA), Stop (UAA) (all colours). **(b)** UAG suppression using an OTS in C321.ΔA does not compete with native translation, whereas standard amber suppression produces truncated products of the protein of interest (POI) due to transcription termination when RF1 binds UAG codons and misincorporation of nsAA in 321 native proteins terminating in UAG.

re-writing synthetic yeast genomes with numerous planned alterations, which include no introns, no IS elements, an inducible evolution system, and UAG to UAA codon reassignment across the genome to harness unique properties of GROs (Figure 3) [48••].

Expanding the genetic code: 4/5-base codon and nonstandard nucleobases (nsNB)

Other exciting strategies aim to expand the genetic code by stepping outside the standard 64 codons. The first of these aims to introduce tRNA species that read quadruplet codons, rather than the canonical triplets. Almost fifty years ago, quadruplet suppression was observed in strains of *S. typhimurium* carrying the mutation *sufD* (a CCCC quadruplet anticodon), thereby producing a $tRNA^{Gly}$ containing the frameshift suppressor mutation, a CCCC anticodon [49]. Recently, quadruplet decoding, and thus +1 frameshifting, has been used for nsAA incorporation [50,51]. Much work has focused on UAGN codons, which could theoretically expand the genetic code with four new open codons.

The second strategy is based on the introduction of nonstandard nucleobases (nsNB). Benner and co-workers designed an nsNB pair (P–Z) with a different hydrogen bonding donor/acceptor pattern compared to that found in standard nucleobases [52,53], whereas Hirao and Romesberg's groups developed hydrophobic nsNB pairs, Ds-Px [54–57] and NaM-TPT3 [58–60], respectively. Impressively, the Romesberg lab recently developed a strain of *E. coli* that imports [61,62] and uses NaM-TPT3 as the third pair of nucleotides: they can be maintained in the genome [63] and transcribed [64••]. Additionally, protein is translated from these mRNAs, enabling the incorporation of nsAA [64••]. Theoretically, the addition of this orthogonal nsNB pair to the genetic code would expand it from 64 to 216 possible codons; however, it is difficult to assess the physical and biological properties of this new nsNB pair, and it is unlikely that all new codons may allow genetic code expansion.

Conclusions

Three major efforts in the field of genetic code expansion are rapidly advancing the possibilities of nsAA incorporation in proteins: (1) advanced molecular evolution strategies to engineer OTSs, (2) efforts to free codons that can be reassigned for dedicated incorporation of nsAAs, and

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

(3) ribosome engineering to expand nsAA incorporation beyond L-α-AAs.

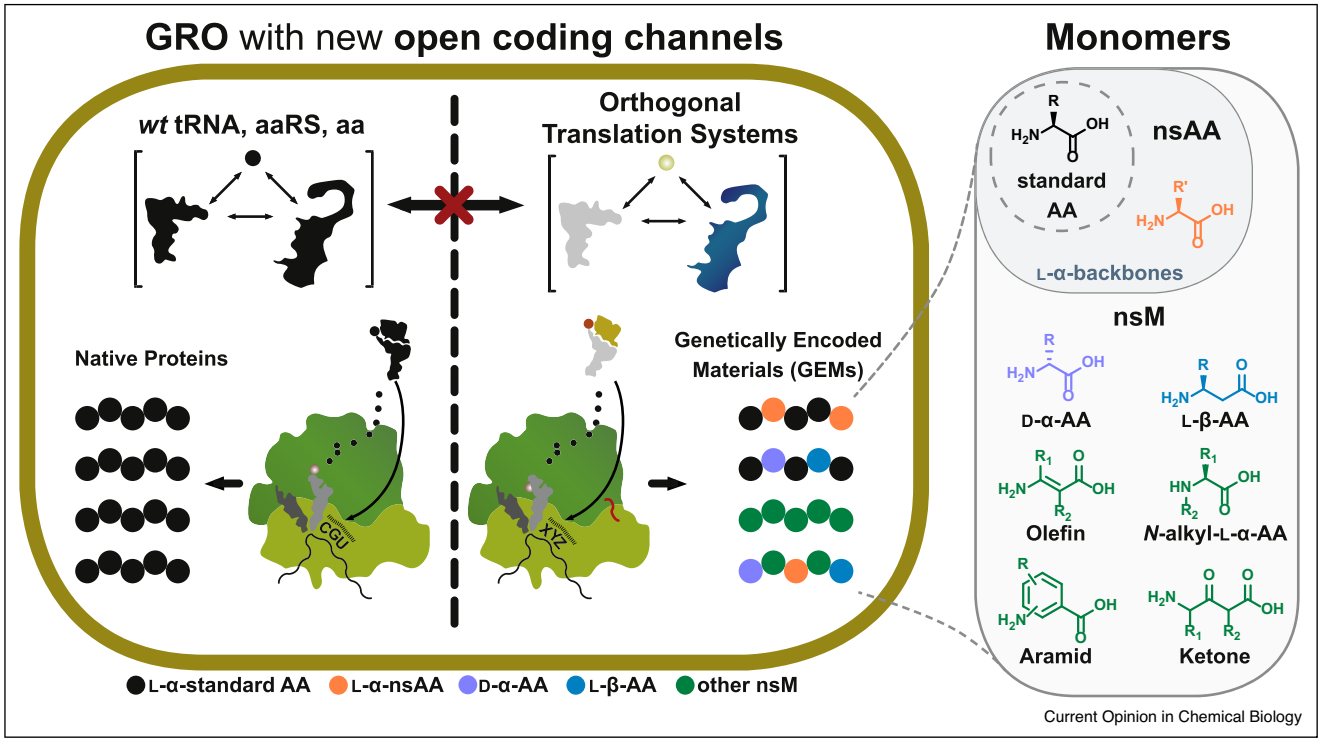
Engineering components of translation requires the use of *in vitro* or *in vivo* approaches that diversify and select for evolved components (e.g. aaRS, tRNA, and EF-Tu). However, the combinatorial possibilities to evolve more than one component at a time are vast and empirically unattainable, yet evolution of single components alone has limitations. Integration of combined efforts to evolve all of these individual components to work in a selective and efficient manner with the native translation apparatus will allow faithful expansion of the genetic code. In this regard, two main features need to be addressed using advanced molecular evolution strategies to enable efficient incorporation of multiple nsAAs, namely OTS efficiency (aaRS activity is typically 10^2–10^3-fold below their native counterparts) [65,66] and specificity, as many orthogonal aaRSs exhibit polyspecificity towards several nsAAs [67].

Historically, OTS performance limited the nsAA incorporation to only one or a few instances per protein, but this limitation is also due to the lack of open codons dedicated to the incorporation of an nsAA. The construction of a GRO has enabled the incorporation of multiple instances of an nsAA into a single protein. Further recoding efforts, advances in quadruplet suppression and introduction of nsNBs are likely to expand the number of codons available for genetic code expansion in the near future. This will pave the way for efficient genetic code expansion with multiple nsAAs.

Finally, the work on orthogonal and tethered ribosomes provides the cell with a complementary translation apparatus that can be fully dedicated to the synthesis of biotechnological products, breaking the link between cell survival and translational synthesis.

We envision GROs with *multiple* open codons combined with *multiple* OTSs and orthogonal, tethered ribosomes integrated in the genome, providing a biological framework capable of establishing a stable and expanded genetic code (Figure 4). Accurate, specific, and multi-site incorporation of diverse nsAAs is a major goal to be met soon and enable the production of functionalized proteins

Figure 4



Expansion of the genetic code in GROs containing multiple open coding channels. Native translation (left) is preserved and isolated from orthogonal translation systems (OTS, right) containing o-tRNA, o-aaRS, oRiboT. The OTS is designed to minimize cross-reacting or interfering with the native translational machinery. The open coding channels in the GRO permit efficient, multi-site incorporation of nonstandard amino acids (nsAAs) into proteins. Future efforts aimed at engineering all components of OTSs may enable the use of nonstandard monomers (nsM) possessing chemistries and modified backbones distinct from standard AA or nsAAs. Such capabilities could permit multi-site incorporation of multiple combinations of AA, nsAAs or nsMs for templated-directed production of genetically encoded materials (GEMs).

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

and polymers containing diverse nsAAs and nsM, leading to the production of entirely new forms of matter, genetically encoded materials (GEMs).

Acknowledgements

FJI gratefully acknowledges the National Science Foundation (CHE-1740549, MCB-1714860) and National Institutes of Health (1R01GM125951-01, 1R01GM117230-01) for funding research related to topics covered in this review.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as

- • of special interest
- •• of outstanding interest

1. Liu CC, Schultz PG: **Adding new chemistries to the genetic code**. *Annu Rev Biochem* 2010, **79**:413-444.

2. Davis L, Chin JW: **Designer proteins: applications of genetic code expansion in cell biology**. *Nat Rev Mol Cell Biol* 2012, **13**:168.

3. LaRiviere FJ, Wolfson AD, Uhlenbeck OC: **Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation**. *Science* 2001, **294**:165.

4. Petry S, Brodersen DE, Murphy FV, Dunham CM, Selmer M, Tarry MJ, Kelley AC, Ramakrishnan V: **Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon**. *Cell* 2005, **123**:1255-1266.

5. van Hest JCM, Kiick KL, Tirrell DA: **Efficient incorporation of unsaturated methionine analogues into proteins in vivo**. *J Am Chem Soc* 2000, **122**:1282-1288.

6. Johnson JA, Lu YY, Van Deventer JA, Tirrell DA: **Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications**. *Curr Opin Chem Biol* 2010, **14**:774-780.

7. Nozawa K, O'Donoghue P, Gundllapalli S, Araiso Y, Ishitani R, Umehara T, Söll D, Nureki O: **Pyrrolysyl-tRNA synthetase–tRNAPyl structure reveals the molecular basis of orthogonality**. *Nature* 2008, **457**:1163.

8. Kobayashi T, Nureki O, Ishitani R, Yaremchuk A, Tukalo M, Cusack S, Sakamoto K, Yokoyama S: **Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion**. *Nat Struct Biol* 2003, **10**:425.

9. Steer BA, Schimmel P: **Major anticodon-binding region missing from an archaebacterial tRNA synthetase**. *J Biol Chem* 1999, **274**:35601-35606.

10. • Italia JS, Addy PS, Wrobel CJJ, Crawford LA, Lajoie MJ, Zheng Y, Chatterjee A: **An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes**. *Nat Chem Biol* 2017, **13**:446.

This work shows the functional replacement of*E. coli* TrpRS-tRNATrp with its engineered *S. cerevisiae* counterpart. In turn, this enabled repurposing of the *E. coli* TrpRS-tRNATrp for genetic code expansion with several Trp-analogues.

11. Iraha F, Oki K, Kobayashi T, Ohno S, Yokogawa T, Nishikawa K, Yokoyama S, Sakamoto K: **Functional replacement of the endogenous tyrosyl-tRNA synthetase–tRNATyr pair by the archaeal tyrosine pair in *Escherichia coli* for genetic code expansion**. *Nucleic Acids Res* 2010, **38**:3682-3691.

12. • Bryson DI, Fan C, Guo L-T, Miller C, Söll D, Liu DR: **Continuous directed evolution of aminoacyl-tRNA synthetases**. *Nat Chem Biol* 2017, **13**:1253.

In this work, the selectivity and activity of aaRSs were improved through PACE. Using this advanced molecular evolution strategy, the authors reported up to a 45-fold increase in activity for PylRS.

13. • Amiram M, Haimovich AD, Fan C, Wang Y-S, Aerni H-R, Ntai I, Moonan DW, Ma NJ, Rovner AJ, Hong SH *et al.*: **Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids**. *Nat Biotechnol* 2015, **33**:1272.

Building on advances in genome engineering methods and GROs, these authors developed a platform for OTS optimization, which resulted in up to 25-fold improvement in aaRS activity. Additionally, the work demonstrates the tunability of substrate specificity for aaRSs.

14. Park H-S, Hohn MJ, Umehara T, Guo L-T, Osborne EM, Benner J, Noren CJ, Rinehart J, Söll D: **Expanding the genetic code of *Escherichia coli* with phosphoserine**. *Science* 2011, **333**:1151.

15. Rogerson DT, Sachdeva A, Wang K, Haq T, Kazlauskaite A, Hancock SM, Huguenin-Dezot N, Muqit MMK, Fry AM, Bayliss R *et al.*: **Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog**. *Nat Chem Biol* 2015, **11**:496.

16. K-I Haruna, Alkazemi MH, Liu Y, Söll D, Englert M: **Engineering the elongation factor Tu for efficient selenoprotein synthesis**. *Nucleic Acids Res* 2014, **42**:9976-9983.

17. Doi Y, Ohtsuki T, Shimizu Y, Ueda T, Sisido M: **Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system**. *J Am Chem Soc* 2007, **129**:14458-14462.

18. Kawakami T, Murakami H, Suga H: **Exploration of incorporation of Nα-methylated amino acids into peptides by sense-suppression method**. *Nucleic Acids Symp Ser* 2007, **51**:361-362.

19. Subtelny AO, Hartman MCT, Szostak JW: **Ribosomal synthesis of N-methyl peptides**. *J Am Chem Soc* 2008, **130**:6131-6136.

20. Goto Y, Murakami H, Suga H: **Initiating translation with D-amino acids**. *RNA* 2008, **14**:1390-1398.

21. Fujino T, Goto Y, Suga H, Murakami H: **Reevaluation of the D-amino acid compatibility with the elongation event in translation**. *J Am Chem Soc* 2013, **135**:1830-1837.

22. Achenbach J, Jahnz M, Bethge L, Paal K, Jung M, Schuster M, Albrecht R, Jarosch F, Nierhaus KH, Klussmann S: **Outwitting EF-Tu and the ribosome: translation with D-amino acids**. *Nucleic Acids Res* 2015, **43**:5687-5698.

23. • Fujino T, Goto Y, Suga H, Murakami H: **Ribosomal synthesis of peptides with multiple β-amino acids**. *J Am Chem Soc* 2016, **138**:1962-1969.

The authors tested, using a cell-free translation system, the incorporation during translation of several β-amino acids. Multiple instances were introduced using those that showed higher incorporation efficiency.

24. Kawakami T, Murakami H, Suga H: **Ribosomal synthesis of polypeptoids and peptoid–peptide hybrids**. *J Am Chem Soc* 2008, **130**:16861-16863.

25. Melo Czekster C, Robertson WE, Walker AS, Söll D, Schepartz A: **In vivo biosynthesis of a β-amino acid-containing protein**. *J Am Chem Soc* 2016, **138**:5194-5197.

This work reports the incorporation*in vivo* of a β-amino acid into a protein using wild-type EF-Tu, PheRS and mutant ribosomes.

26. Hui A, de Boer HA: **Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli***. *Proc Natl Acad Sci U S A* 1987, **84**:4762.

27. Rackham O, Chin JW: **A network of orthogonal ribosomemRNA pairs**. *Nat Chem Biol* 2005, **1**:159.

28. •• Orelle C, Carlson ED, Szal T, Florin T, Jewett MC, Mankin AS: **Protein synthesis by ribosomes with tethered subunits**. *Nature* 2015, **524**:119.

Successful tethering of the 23S rRNA to an orthogonal 16S rRNA establishes a parallel ribosome, which can be readily engineered without affecting cell viability. This work is paving the way for new opportunities for ribosome engineering, and ribosomal synthesis of novel biopolymers.

29. •• Fried SD, Schmied WH, Uttamapinant C, Chin JW: **Ribosome subunit stapling for orthogonal translation in *E. coli***. *Angew Chem Int Ed* 2015, **54**:12791-12794.

The authors constructed a tethered or "stapled" orthogonal ribosome to enable mutations in the 23S rRNA without affecting cell viability.

30. Kwon Y-C, Jewett MC: **High-throughput preparation methods of crude extract for robust cell-free protein synthesis**. *Sci Rep* 2015, **5**:8663.

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

31. Bundy BC, Swartz JR: **Site-specific incorporation of p-propargyloxyphenylalanine in a cell-free environment for direct protein–protein click conjugation**. *Bioconjug Chem* 2010, **21**:255-263.

32. Xiao H, Murakami H, Suga H, Ferré-D'Amaré AR: **Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme**. *Nature* 2008, **454**:358.

33. Ramaswamy K, Saito H, Murakami H, Shiba K, Suga H: **Designer ribozymes: programming the tRNA specificity into flexizyme**. *J Am Chem Soc* 2004, **126**:11454-11455.

34. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T: **Cell-free translation reconstituted with purified components**. *Nat Biotechnol* 2001, **19**:751.

35. Ahn J-H, Hwang M-Y, Oh I-S, Park K-M, Hahn G-H, Choi C-Y, D-M Kim: **Preparation method for *Escherichia coli* S30 extracts completely dependent upon tRNA addition to catalyze cell-free protein synthesis**. *Biotechnol Bioprocess Eng* 2006, **11**:420.

36. Cui Z, Stein V, Tnimov Z, Mureev S, Alexandrov K: **Semisynthetic tRNA complement mediates in vitro protein synthesis**. *J Am Chem Soc* 2015, **137**:4404-4413.

37. • Lajoie MJ, Rovner AJ, Goodman DB, Aerni H-R, Haimovich AD, Kuznetsov G, Mercer JA, Wang HH, Carr PA, Mosberg JA *et al.*: **Genomically recoded organisms expand biological functions**. *Science* 2013, **342**:357.

This work marks the completion of UAG to UAA whole-genome recoding in E. coli. The resulting GRO exhibited improved nsAA incorporation and increased viral resistance.

38. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman DB, Reppas NB *et al.*: **Precise manipulation of chromosomes in vivo enables genome-wide codon replacement**. *Science* 2011, **333**:348.

39. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM: **Programming cells by multiplex genome engineering and accelerated evolution**. *Nature* 2009, **460**:894.

40. Ma NJ, Moonan DW, Isaacs FJ: **Precise manipulation of bacterial chromosomes by conjugative assembly genome engineering**. *Nat Protoc* 2014, **9**:2285.

41. Ma NJ, Isaacs FJ: **Genomic recoding broadly obstructs the propagation of horizontally transferred genetic elements**. *Cell Syst* 2016, **3**:199-207.

42. Rovner AJ, Haimovich AD, Katz SR, Li Z, Grome MW, Gassaway BM, Amiram M, Patel JR, Gallagher RR, Rinehart J *et al.*: **Recoded organisms engineered to depend on synthetic amino acids**. *Nature* 2015, **518**:89.

43. Mandell DJ, Lajoie MJ, Mee MT, Takeuchi R, Kuznetsov G, Norville JE, Gregg CJ, Stoddard BL, Church GM: **Biocontainment of genetically modified organisms by synthetic protein design**. *Nature* 2015, **518**:55.

44. Napolitano MG, Landon M, Gregg CJ, Lajoie MJ, Govindarajan L, Mosberg JA, Kuznetsov G, Goodman DB, Vargas-Rodriguez O, Isaacs FJ *et al.*: **Emergent rules for codon choice elucidated by editing rare arginine codons in *Escherichia coli***. *Proc Natl Acad Sci U S A* 2016, **113**:E5588.

45. Lajoie MJ, Kosuri S, Mosberg JA, Gregg CJ, Zhang D, Church GM: **Probing the limits of genetic recoding in essential genes**. *Science* 2013, **342**:361.

46. • Ostrov N, Landon M, Guell M, Kuznetsov G, Teramoto J, Cervantes N, Zhou M, Singh K, Napolitano MG, Moosburner M *et al.*: **Design, synthesis, and testing toward a 57-codon genome**. *Science* 2016, **353**:819.

The authors design and work toward the construction of an*E. coli* with 57-codon genome.

47. • Lau YH, Stirling F, Kuo J, Karrenbelt Michiel AP, Chan YA, Riesselman A, Horton CA, Schäfer E, Lips D, Weinstock MT *et al.*: **Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA**. *Nucleic Acids Res* 2017, **45**:6971-6980.

The authors recoded a strain of Salmonella using a novel methodology called SIRCAS, which involves the direct iterative recombineering of 10–25 kb synthetic DNA constructs. They replaced 1557 synonymous leucine codons across 176 genes.

48. •• Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, Lee D, Huang CLV, Chandrasegaran S, Cai Y *et al.*: **Design of a synthetic yeast genome**. *Science* 2017, **355**:1040.

This paper describes the redesign and synthesis of a yeast genome with an 8% reduction in size, removing introns, IS elements, reassigning UAG codons to UAA, and introducing an inducible "SCRaMbLE" evolution system. This work opens up the possibility of customizable genome synthesis.

49. Riddle DL, Carbon J: **Frameshift suppression: a nucleotide addition in the anticodon of a glycine transfer RNA**. *Nat New Biol* 1973, **242**:230.

50. Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, Schultz PG: **An expanded genetic code with a functional quadruplet codon**. *Proc Natl Acad Sci U S A* 2004, **101**:7566.

51. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW: **Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome**. *Nature* 2010, **464**:441.

52. Yang Z, Chen F, Chamberlin SG, Benner SA: **Expanded genetic alphabets in the polymerase chain reaction**. *Angew Chem Int Ed* 2010, **49**:177-180.

53. Yang Z, Chen F, Alvarado JB, Benner SA: **Amplification, mutation, and sequencing of a six-letter synthetic genetic system**. *J Am Chem Soc* 2011, **133**:15105-15112.

54. Kimoto M, Kawai R, Mitsui T, Yokoyama S, Hirao I: **An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules**. *Nucleic Acids Res* 2009, **37**:e14.

55. Yamashige R, Kimoto M, Takezawa Y, Sato A, Mitsui T, Yokoyama S, Hirao I: **Highly specific unnatural base pair systems as a third base pair for PCR amplification**. *Nucleic Acids Res* 2012, **40**:2793-2806.

56. Okamoto I, Miyatake Y, Kimoto M, Hirao I: **High fidelity, efficiency and functionalization of Ds–Px unnatural base pairs in PCR amplification for a genetic alphabet expansion system**. *ACS Synth Biol* 2016, **5**:1220-1230.

57. Hirao I, Kimoto M, Mitsui T, Fujiwara T, Kawai R, Sato A, Harada Y, Yokoyama S: **An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA**. *Nat Methods* 2006, **3**:729.

58. Li L, Degardin M, Lavergne T, Malyshev DA, Dhami K, Ordoukhanian P, Romesberg FE: **Natural-like replication of an unnatural base pair for the expansion of the genetic alphabet and biotechnology applications**. *J Am Chem Soc* 2014, **136**:826-829.

59. Dhami K, Malyshev DA, Ordoukhanian P, Kubelka T, Hocek M, Romesberg FE: **Systematic exploration of a class of hydrophobic unnatural base pairs yields multiple new candidates for the expansion of the genetic alphabet**. *Nucleic Acids Res* 2014, **42**:10235-10244.

60. Feldman AW, Dien VT, Romesberg FE: **Chemical stabilization of unnatural nucleotide triphosphates for the in vivo expansion of the genetic alphabet**. *J Am Chem Soc* 2017, **139**:2464-2467.

61. Feldman AW, Fischer EC, Ledbetter MP, Liao J-Y, Chaput JC, Romesberg FE: **A tool for the import of natural and unnatural nucleoside triphosphates into bacteria**. *J Am Chem Soc* 2018, **140**:1447-1454.

62. Malyshev DA, Dhami K, Lavergne T, Chen T, Dai N, Foster JM, Corrêa IR, Romesberg FE: **A semi-synthetic organism with an expanded genetic alphabet**. *Nature* 2014, **509**:385.

63. Zhang Y, Lamb BM, Feldman AW, Zhou AX, Lavergne T, Li L, Romesberg FE: **A semisynthetic organism engineered for the stable expansion of the genetic alphabet**. *Proc Natl Acad Sci U S A* 2017, **114**:1317.

64. •• Zhang Y, Ptacin JL, Fischer EC, Aerni HR, Caffaro CE, San Jose K, Feldman AW, Turner CR, Romesberg FE: **A semi-synthetic**

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020

organism that stores and retrieves increased genetic information. *Nature* 2017, **551**:644.
The authors present a strain of*E. coli* in which nsNBs can be maintained in genomic DNA and used in transcription and in translation. The added nsNB pair has the potential to greatly expand the number of open codons that could be used to encode nsAA.

65. Umehara T, Kim J, Lee S, Guo L-T, Söll D, Park H-S: **N-acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess N-acetyl lysine specificity in vitro and in vivo**. *FEBS Lett* 2012, **586**:729-733.

66. Tanrikulu IC, Schmitt E, Mechulam Y, Goddard WA, Tirrell DA: **Discovery of *Escherichia coli* methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo**. *Proc Natl Acad Sci U S A* 2009, **106**:15285.

67. O'Donoghue P, Ling J, Wang Y-S, Söll D: **Upgrading protein synthesis for synthetic biology**. *Nat Chem Biol* 2013, **9**:594.

Please cite this article in press as: Arranz-Gibert P, et al.: Next-generation genetic code expansion, Curr Opin Chem Biol (2018), https://doi.org/10.1016/j.cbpa.2018.07.020