Reaching new heights in genetic code manipulation with high throughput screening

Briana R. Lino[a§], Sean J. Williams[a§], Michelle E. Castor[a§], and James A. Van Deventer[a,b*]

[a]*Chemical and Biological Engineering Department, Tufts University, Medford, Massachusetts 02155, United States*

[b]*Biomedical Engineering Department, Tufts University, Medford, Massachusetts 02155, United States*

[*]Email: James.Van_Deventer@tufts.edu

[§]B.R.L., S.J.W., and M.E.C. contributed equally to this paper.

Abstract

The chemical and physical properties of proteins are limited by the twenty canonical amino acids, but genetic code manipulation allows for the incorporation of noncanonical amino acids (ncAAs), enhancing protein functionality. This review explores three main strategies for introducing ncAAs into biosynthesized proteins, focusing on high-throughput screening's role in these advancements. The first section discusses engineering aminoacyl-tRNA synthetases (aaRSs) and tRNAs, emphasizing how novel selection methods improve selectivity and organismal fitness. The second section examines high-throughput techniques for optimizing protein translation machinery, expanding its capacity for alternative genetic codes. This includes opportunities to enhance ncAA incorporation through engineering cellular components unrelated to translation. The final section highlights various discovery platforms for high-throughput screening of ncAA-containing proteins, showcasing innovative binding ligands and enzymes that are challenging to create with only canonical amino acids. These advances have led to promising drug leads and biocatalysts. Overall, the ability to discover unexpected functionalities through high-throughput methods significantly influences ncAA incorporation and its applications. Future innovations in experimental techniques, along with advancements in computational protein design and machine learning, are poised to further elevate this field.

Table of Contents

1. Introduction

Proteins carry out an astonishing range of essential biological processes. However, their structures and functions are constrained by the properties of their constituent canonical amino acids. While both nature and humans have evolved, designed, and augmented proteins using post-translational modifications, biosynthetic introduction of noncanonical amino acids into proteins greatly expands the potential range of engineered protein functions (noncanonical amino acids (ncAAs) are also referred to as unnatural amino acids (UAAs), nonstandard amino acids (nsAAs), nonnatural amino acids (nnAAs) or nonproteinogenic amino acids (npAAs))[1]. NcAA side chains can confer proteins with a wide-ranging set of expanded features, including conjugation handles for probes or warheads, crosslinkable groups for spontaneous or light-induced covalent target engagement, and post-translational modifications at precisely defined sites[1-4]. Use of monomers that go beyond α-L-amino acids offer access to additional properties including unique conformational spaces as well as reduced recognition by the immune system and protein degradation pathways[5].

There are three primary strategies for adding ncAAs into proteins biosynthetically: 1) residue-specific incorporation to replace a canonical amino acid with a ncAA; 2) site-specific incorporation to introduce a ncAA into a protein without replacing a canonical amino acid; and 3) *in vitro* genetic code reprogramming to add one or several ncAAs into biosynthesized polypeptides (Figure 1A-E)[6-8]. The ribosomal synthesis of proteins containing ncAAs using any of these methods is technically demanding[9]. Although deep knowledge of biochemistry and structural biology has provided an excellent foundation for implementing ncAA incorporation strategies, combining this knowledge with approaches from biomolecular engineering and directed evolution has allowed this area to thrive[10-14].

This review focuses on the ways that high throughput screening and selection technologies (Table 1) have enabled researchers to reach new heights in genetic code manipulation and its applications. A common theme of this review is that higher throughput experimentation has resulted in the ability to "discover the unexpected": engineered translation systems with improved ncAA incorporation efficiency and fidelity, cells that better accommodate alternative genetic codes, therapeutic leads with improved potency or "drug-like" physicochemical properties, and more. Before focusing on uses of high throughput approaches in ncAA incorporation and its applications, we briefly review the three major strategies for inserting ncAAs into ribosomally synthesized proteins. Readers are referred elsewhere for more extensive discussions of fundamentals of genetic code manipulation[15-18].

The first major portion of this review focuses on advances in engineering aminoacyl-tRNA synthetases (aaRSs), tRNAs, and orthogonal aaRS-tRNA pairs (also called orthogonal translation systems (OTSs)). These biomolecules are pivotal in genetic code manipulation: their crucial role in controlling the interpretation of the genetic code requires that they be tailored to support the insertion of desired ncAAs into proteins with high efficiency and fidelity. Advanced screening systems have pushed ncAA incorporation efficiency and the diversity of biosynthetically accessible ncAA chemistries to impressive levels (Figure 1F-H).

The second major portion of this review highlights ambitious efforts to engineer cells to better accommodate alternative genetic codes. Primarily focusing on genetic code expansion, this section covers advances in manipulating components of the protein translation apparatus, reworking whole genomes, and conducting adaptive laboratory evolution and genome-wide screens. All parts of a cell must work in synergy to support efficient production of proteins containing ncAAs, and emerging studies powerfully demonstrate how engineered elongation factors, release factors, ribosomes, genomes, and more can all lead to enhanced genetic code manipulation systems.

The final major portion of this review describes applications in which ncAA-containing polypeptides are screened in high throughput to identify unique “chemically expanded” biomolecules including enzymes, cyclic and stapled peptides, and binding proteins. There is now a diverse array of discovery platforms (Figure 1F-L) that have enabled the discovery of high-performing biomolecules with properties that would be extremely challenging to engineer using only canonical amino acids. Emerging therapeutic leads and biocatalysts highlight the promising potential of integrating carefully chosen chemistries into high throughput biomolecule discovery and engineering.

Overall, this review provides a comprehensive exploration of the ways that high throughput experimentation has impacted this area. With a broad and rapidly growing set of screening tools now available, we will continue to see genetic code manipulation and its applications reach new heights in the years to come.

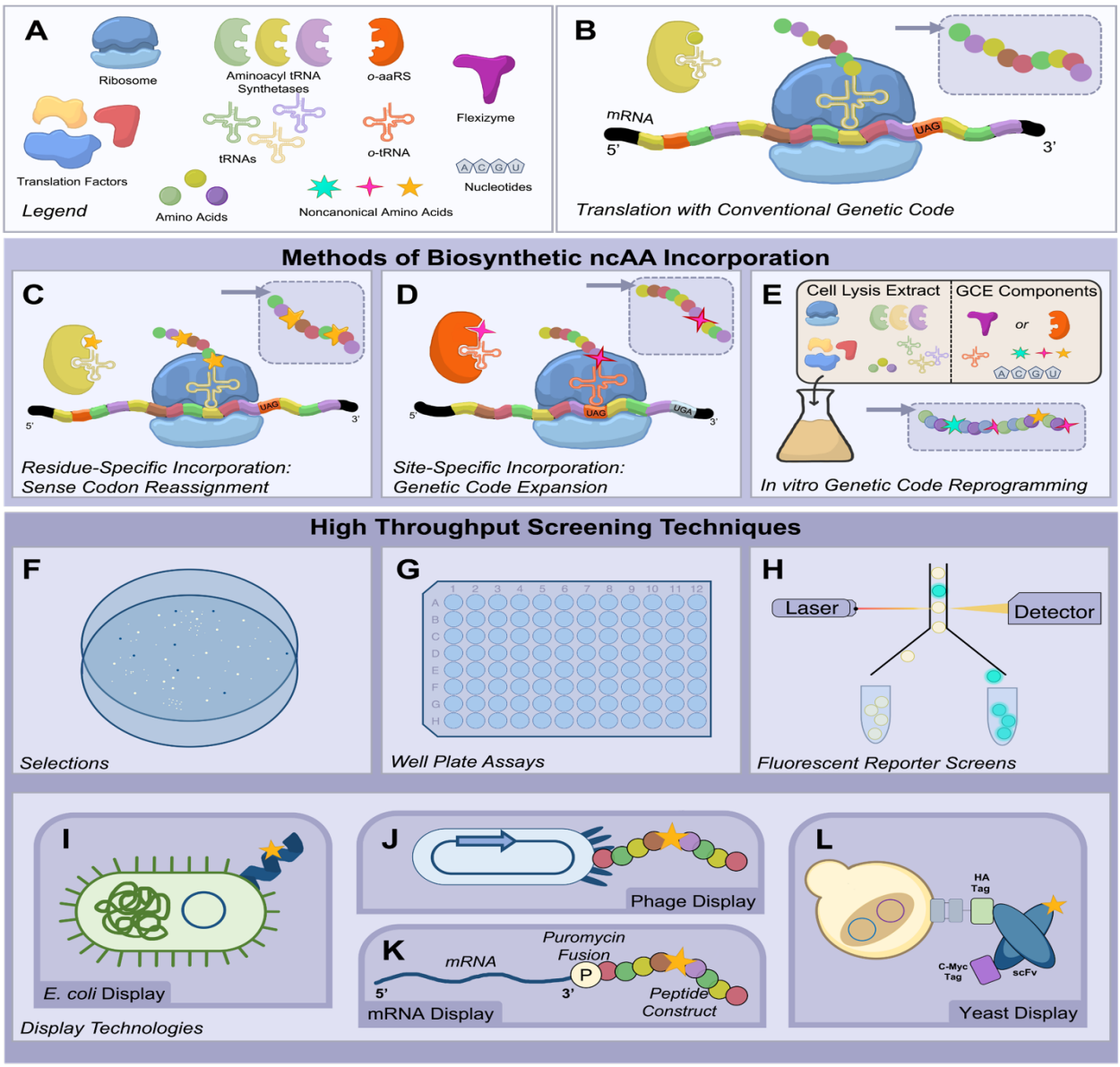

Figure 1. Overview of methods and high throughput screening techniques used for genetic code manipulation. A) Translation elements required for conventional translation and noncanonical amino acid (ncAA) incorporation. B) Graphical representation of translation using the conventional genetic code. An aminoacyl-tRNA synthetase first charges a tRNA with its cognate amino acid. Within the ribosome, the aminoacylated tRNA recognizes a specific codon(s) on mRNA and is positioned to catalyze the transfer of its bound amino acid to a growing polypeptide chain. When a stop codon is identified, translation terminates, resulting in release of the polypeptide. C-E) Schematics of biosynthetic ncAA incorporation methods. C) Residue-specific ncAA incorporation via sense codon reassignment. An aminoacyl-tRNA synthetase charges an ncAA onto a tRNA that recognizes a sense codon, resulting in ncAA incorporation at the sense codon. D) Site-specific ncAA incorporation via genetic code expansion. An orthogonal aminoacyl-tRNA synthetase charges a ncAA onto an orthogonal tRNA that recognizes one of the three stop codons, resulting in addition of a 21st amino acid to the genetic code in response to a stop codon (other "blank" codons can also be recoded with this approach). E) *In vitro* genetic code reprogramming via cell-free protein synthesis. Elements of the protein translation apparatus are obtained from cell extracts and combined with genetic code expansion components *in vitro*, which can facilitate extensively reworked genetic codes for biosynthesis of polypeptides containing multiple, distinct ncAAs. F-L) High throughput screening methods used in genetic code manipulation and its applications. F) Live/dead selections to identify clones by linking desired phenotype(s) to cell viability. G) Well plate assays to identify clones by linking desired phenotype(s) to fluorescent or colorimetric readouts in solution. H) Fluorescence-activated cell sorting to identify clones by linking desired phenotype(s) to cells expressing fluorescent reporters. I-L) Display technologies to identify clones by physically linking polypeptides with desired phenotype(s) to genetic material encoding polypeptides of interest. I) *E. coli* display. Common phenotypes assayed during enrichments: protein expression levels, antigen binding. J) Phage display. Common phenotype(s) assayed during enrichments: antigen binding. K) mRNA display. Common phenotypes assayed during enrichments: antigen binding. L) Yeast display. Common phenotypes assayed: protein expression levels, antigen binding, stability, specificity.

Table 1. Selected high throughput screening approaches.

HTS Method	Common Engineering Targets	Phenotype	Host System	Library Diversity
Live/Dead Selections	aaRS/tRNA	Growth	*E. coli;* *S. cerevisiae*	10^6- 10^9
Fluorescent Reporters	aaRS/tRNA	Fluorescence	*E. coli;* *S. cerevisiae*	10^6-10^8
Continuous Evolution	aaRS/tRNA	Phage propagation; Luminescence	Phage, *E. coli*	Experiment-dependent
Compartmentalized Partnered Replication (CPR)	aaRS/tRNA	DNA amplification	*E. coli*	10^8-10^{10}
Virus-Assisted Directed Evolution (VADER)	tRNA	Viral propagation	AAV, HEK293T	10^7
Yeast Display	Antibodies, enzymes, peptides, aaRS	Fluorescence	*S. cerevisiae*	10^8-10^9
E. coli Display	Peptides, select protein scaffolds	Fluorescence	*E. coli*	10^{10}-10^{11}
Phage Display	Peptides	Phage propagation	*E. coli*	10^{10}-10^{11}
mRNA Display	Peptides	DNA amplification	*In vitro*	10^{13}-10^{14}

2. Methods of ncAA Incorporation

Here, we introduce the three main approaches to incorporate ncAAs into biosynthesized peptides and proteins. It is important to recognize the commonalities and differences between these approaches, as their complementary strengths and limitations are ultimately what facilitates (or hinders) applications of ncAAs. For more detailed treatments of underlying methods, readers are referred to prior excellent reviews[19-24]. We use the term "genetic code manipulation" to encompass all methods for biosynthetically introducing ncAAs into proteins.

One method for ncAA incorporation is residue-specific ncAA incorporation, where one of the canonical amino acids is replaced by a ncAA[15]. This is a reinterpretation of an entire group of sense codons, which changes the composition of the genetic code without expanding it. Near-complete ncAA incorporation can be achieved using an auxotrophic host unable to synthesize the canonical amino acid being replaced, along with a close analog of the canonical amino acid. A primary advantage of residue-specific replacement is that preparing cells for ncAA incorporation does not require the addition of new components to the translation apparatus for close analogs. Straightforward media manipulations are sufficient to enable a native aaRS to aminoacylate its cognate tRNA with the ncAA of interest. The residue-specific approach facilitates insertion of ncAAs at several sites within the same protein, which can be advantageous for some protein engineering applications. Global insertion of ncAAs throughout the proteome is useful for sensitive identification of newly synthesized proteins[25-31] and related proteomics applications. A detailed treatment of proteomics approaches is beyond the scope of this review; the reader is referred to other excellent treatments of this area[29, 32-35]. Extensions of residue-specific replacement typically involve engineering a native aaRS to accept a broader range of ncAA substrates for aminoacylation[36]. Screening strategies for aaRS engineering will be discussed in Section 3.

The most widely practiced method for ncAA incorporation is site-specific incorporation in cells, where a "blank" codon is repurposed to add an ncAA to the genetic code alongside the 20 canonical amino acids[37]. This approach is also referred to as, "genetic code expansion." In its most common implementation, the amber stop codon (UAG) is utilized for ncAA incorporation[38]. Site-specific incorporation requires an additional aminoacyl-tRNA synthetase/tRNA pair to be added to the protein translation apparatus[39]; this pair is also referred to as an orthogonal translation system (OTS). The newly introduced OTS should operate orthogonally to the native translation machinery: the aaRS should recognize the ncAA and orthogonal tRNA but not recognize other amino acids or tRNAs present, while the orthogonal tRNA should not be a substrate of any aaRSs of the native translation apparatus. A major strength of this approach is that site-specific ncAA incorporation enables ncAA "point mutations" within proteins that minimally disrupt overall protein structure. However, engineering high-performing OTSs to accomplish such manipulations can be labor-intensive; advances in the discovery and engineering of these systems will be discussed at length in Section 3. Beyond repurposing the amber codon, researchers have also used the opal (UGA) and ochre (UAA) stop codons[40], frame-shift suppression with quadruplet and quintuplet codon-anticodon pairs[41-44], unnatural DNA base pairs[45], and genomically recoded organisms to create "blank" codons for expansion of the genetic code in cells[46]. Many of these efforts have used high throughput approaches to push the limits of genetic code expansion; they will be discussed in Section 4.

While both residue-specific and site-specific ncAA incorporation strategies are now well-established, these methods are constrained by the fact that they are performed within cells. Elimination of cells entirely with *in vitro* genetic code reprogramming frees researchers from the need to maintain cell viability during protein biosynthesis[47, 48]. This greatly increases the potential range of strategies for introducing ncAAs into proteins. Two common strategies for *in vitro* reprogramming are used: one where the translation apparatus is prepared from cell lysates, and another where it is reconstituted from purified components. The latter system, typically using components from *E. coli*, is called "protein synthesis using recombinant elements" (PURE)[49]. This system is usually combined with the use of ribozymes called flexizymes; these nucleic acid catalysts charge tRNAs with ncAAs[50-52] to enable ribosomal synthesis of chemically expanded (poly)peptides *in vitro*[53, 54]. With both cell extracts and purified components, protein synthesis reactions are supplemented with additional translation components that facilitate ncAA incorporation in response to codons of interest. As with cell-based genetic code manipulation, OTSs can be used to facilitate ncAA incorporation cotranslationally. In addition, a full set of acylated tRNAs can be prepared prior to protein synthesis to enable radical reprogramming of the genetic code using either OTSs or flexizymes. This can include removal of one or more cAAs as well as addition of multiple ncAAs. Selected methodological advances of *in vitro* genetic code manipulation systems related to engineering the translational apparatus will be highlighted in Section 4, but we also refer the reader elsewhere for excellent reviews that describe these systems more extensively[48, 55, 56].

When used well, all three genetic code manipulation strategies facilitate powerful modulations of protein structure and function. Perhaps unsurprisingly, the impressive reprogramming capabilities of *in vitro* systems have led to some of the most advanced peptide screening platforms reported to date. However, careful deployment of screening technologies using *E. coli* and yeast genetic code manipulation systems are also beginning to show promise for high throughput discovery of "chemically expanded" proteins. These applications of ncAAs with high throughput technologies will be discussed in Section 5.

3. High-Throughput Engineering of Aminoacyl-tRNA Synthetases and Orthogonal Translation Systems

The challenge of efficiently aminoacylating tRNAs with ncAAs of interest is common to all major types of genetic code manipulation[57, 58]. This section will focus on advances in engineering aaRSs and tRNAs for genetic code manipulation in cells. Powerful, advanced screening strategies for identifying high-performing aaRSs are conceptually similar for ncAA incorporation in response to sense or nonsense codons (residue-specific or site-specific ncAA incorporation, respectively). As such, these approaches will be discussed side-by-side. Exciting recent advances in engineering high-performing orthogonal tRNAs (site-specific ncAA incorporation) will also be discussed below.

There are numerous OTS engineering strategies, including both rational and high throughput approaches. The extensive body of literature on aaRS/tRNA biochemistry and structural biology can inform rational designs that can be quickly tested in the laboratory. In particular, a number of orthogonal tRNAs possess sequences that are unchanged from the sequences of naturally occurring tRNAs outside of the identity of the anticodon[59-63]. Some rationally designed aaRS variants have also been used for protein biosynthesis with ncAAs, but commonly exhibit low to moderate efficiencies[64-66]. A substantial confounding factor for rational design is that there are subtle (or even not so subtle)[67] sequence and structural differences between aaRS homologs. These differences make it challenging to transplant mutations from one homolog to another without substantial loss of ncAA incorporation efficiency[64]. More broadly, aaRSs must be able to use desired ncAA substrates efficiently while finely discriminating against cAAs. In some cases, it is also important to prevent undesired ncAA incorporation (i.e. off-target ncAAs), such as when two or more distinct ncAAs will be incorporated into the same protein[68, 69]. Despite the existence of substantial data, we lack the sophisticated understanding required to consistently use rational design to identify high-performing aaRSs. For the case of orthogonal tRNAs, our understanding of identity elements continues to improve[63, 70]. At the same time, recent screening efforts have yielded tRNAs with dramatically improved ncAA incorporation efficiencies but not yet mechanistic understanding explaining the basis for the improvements[12, 71]. Thus, high throughput screens and selections are critical tools for engineering efficient genetic code manipulation systems.

Historically, the most commonly used strategy for high throughput discovery of aaRSs or tRNAs employs a double sieve live/dead selection system. This approach has been used especially widely for genetic code expansion in *E. coli*[72, 73], and a durable version of this system is also available in yeast[74-76]. In these systems, the positive selection couples cell survival with stop codon suppression using a selectable marker, conferring a growth phenotype in the presence of the ncAA for a translationally active aaRS (for example, chloramphenicol resistance via chloramphenicol acetyltransferase (CAT) in *E. coli*). The negative selection typically couples cell death to stop codon readthrough in the absence of ncAAs using a second selectable marker (for example, the toxic barnase protein in *E. coli*)[39]. Over dozens of studies, these selections have been shown to be effective, cheap, and relatively straightforward to implement[72, 73, 77-79]. At the same time, these systems have limitations. Known issues with growth-based selections include poor discrimination between variants that exhibit different incorporation efficiencies and false positives[80]. Advanced selection and screening strategies that address these challenges are now available for engineering aaRS and tRNA variants. Sections 3.1 and 3.2 highlight a range of these emerging high-throughput methodologies. Section 3.3 describes initial data-driven and computational approaches for identifying promising aaRS and tRNA variants.

3.1 High-throughput Aminoacyl-tRNA Synthetase Engineering

Numerous strategies beyond live/dead selections enable efficient aaRS variant discovery, ranging from fluorescent reporters to continuous evolution systems. One simple but powerful approach is to implement fluorescent reporters at the conclusion of live/dead selections. Unlike selection systems, these reporters enable quantitative readouts; they also enable identification of false positives arising during selections[64, 80, 81]. Combining a reporter encoding a stop codon-containing green fluorescent protein (GFP) with conventional selections, Cooley et al. were able to rapidly compare the performance of first- and second-generation aaRS variants selected for the incorporation of 3-nitro-tyrosine (nitroTyr) in response to the amber stop codon in *E. coli*[80]. Following a modified selection procedure, assays with fluorescent reporters revealed that several second-generation clones greatly outperformed previously reported clones. This highlights both the utility of a fluorescent reporter as well as opportunities to tune selection conditions to facilitate isolation of enhanced aaRS variants[80]. As a further extension of this strategy, Sungwienwong et al. demonstrated that standard selections can be enhanced by adding a structurally similar “off-target” ncAA to negative selections to identify aaRSs that exhibit high selectivity. Assays with fluorescent reporters following selections revealed an aaRS variant that exhibited a 50-fold increase in selectivity for the fluorescent ncAA acridon-2-alanine (Acd) over the related nonfluorescent amino acid

N-phenyl-amino-phenylalanine (Npf); this amino acid is a byproduct of Acd synthesis that often contaminates the final product[82]. Homology modeling indicated that two separate mutations (G32A and E65D) play important roles in mediating selective recognition of Acd. These findings highlight the utility of even simple fluorescence reporters to evaluate the performance of individual aaRS candidates[82].

One challenge when combining fluorescent reporters with selections is that the genes encoding these distinct systems are usually located on separate plasmids. This requires extra steps to isolate aaRS variants of interest from cells used for selections and transform them into cells containing a fluorescent reporter. A solution is to create genes encoding multi-domain fusion proteins that encode for both a selectable phenotype and a fluorescent reporter. Cervettini et al described a triple reporter system for amber codon suppression that supports selections for ncAA incorporation with CAT, a GFP reporter to indicate misincorporation of tyrosine, and an E2Crimson reporter to indicate ncAA incorporation (Figure 2A)[63]. This reporter was used to identify a series of variants of *Archaeoglobus fulgidus* tyrosyl-tRNA synthetase *(*AfTyrRS)-G5 that each enabled incorporation of *p*-iodo-L-phenylalanine (pIF) and *p*-azido-L-phenylalanine (AzF) at higher rates than commonly used *Methanocaldococcus jannaschii* (Mj) aaRS derivatives[63]. Another sophisticated multi-domain reporter was implemented by Thyer et al. to support discovery of aaRS variants for the incorporation of 3, 4-dihydroxy-L-phenylanaline (L-DOPA) into proteins. A dual CAT-GFP reporter was established in which the spectral properties of the GFP reporter variant were altered upon L-DOPA incorporation (via crosslinking between oxidized L-DOPA and a cysteine introduced at position 205 of GFP)[83]. Compartmentalized Partnered Replication (CPR), an engineering approach that will be described further in Section 3.2, was used as a screening method in combination with the fluorescent reporter. Anticodon and binding pocket libraries yielded variants with minor improvements in activity and selectivity. Interestingly, the authors determined that a Y119F mutation, located outside the primary amino acid binding pocket of the aaRS, enhanced specificity for L-DOPA. Combining this mutation with other mutations that improved L-DOPA incorporation efficiency led to additive improvements in aaRS performance. It should be noted that the reporters are both somewhat "bespoke:" the dual CAT-GFP reporter exhibits spectral shifts only when L-DOPA is incorporated, and the triple CAT-GFP-E2Crimson reporter's cAA discrimination module identifies only tyrosine misincorporation events[83]. Further reporter development would be required to support readouts of additional types of activities, which is not necessarily trivial. Nonetheless, fusion reporters that integrate selections and fluorescent readouts are powerful tools to support high throughput discovery of aaRSs with desired properties.

As noted above, conventional selection systems can struggle to discriminate between aaRS variants that support even widely varying efficiencies of ncAA incorporation. Judicious changes to these systems can result in vastly improved selections. For example, Grasso et al. reported that a traditional selection system in *E. coli* based on CAT was unable to discriminate between aaRS variants known to exhibit varying levels of ncAA incorporation in response to the amber stop codon[84]. Replacement of the gene encoding CAT with a gene encoding a beta-lactamase led to the establishment of a selection system with a wider dynamic range (WiDr)[84]. Initial validations demonstrated discrimination between previously characterized variants with different activities that the prior selection system failed to differentiate. Furthermore, library selections with WiDr identified aaRS variants with 2-fold improved incorporation of benzoylphenyalanine (Bpa) into proteins compared to standard selections, and polyspecific aaRSs that support the incorporation of a chemically diverse range of aromatic ncAAs[84]. These results indicate the nuanced importance of selection system characteristics in supporting discovery of enhanced aaRSs for ncAA incorporation.

One of the fundamental challenges of selection systems is that coupling cell survival to a desired activity (or lack of activity in negative selections) is only an indirect indicator of the desired phenotype. Cells grow after a coupled activity reaches a specific threshold (or fail to grow if activity remains under a specific threshold); cellular survival mechanisms can be triggered and result in a high frequency of false positives during selections. Reporters that decouple aaRS activity from cell growth circumvent this issue. Owens et al. established a colorimetric LacZ reporter gene containing two TAG codons to facilitate evaluation of full-length reporter gene expression with blue/white colony screening in the presence of X-

gal[85]. This enabled qualitative identification of candidate aaRSs supporting ncAA incorporation on agar plates, followed by picking of individual colonies and quantitative measurement of galactosidase activity via para-nitrophenol-galactose in a multiwell plate format. Additional validation with a fluorescent reporter system provided a complementary assay to further evaluate aaRS variant performance. Application of this screening system with two rounds of random mutagenesis generated the variant CrtK-RS(4.1), which exhibited 250-370% improved incorporation of N^{ε}-crotonyl-lysine (CrtK) compared to prior variants. This reporter also enabled identification of mutations in the non-catalytic N-terminal domain of *Methanosarcina barkeri (Mb)* pyrrolysyl-tRNA synthetase (PylRS) that improved activity[85]. These findings demonstrate the feasibility and utility of decoupling aaRS activity from cell survival.

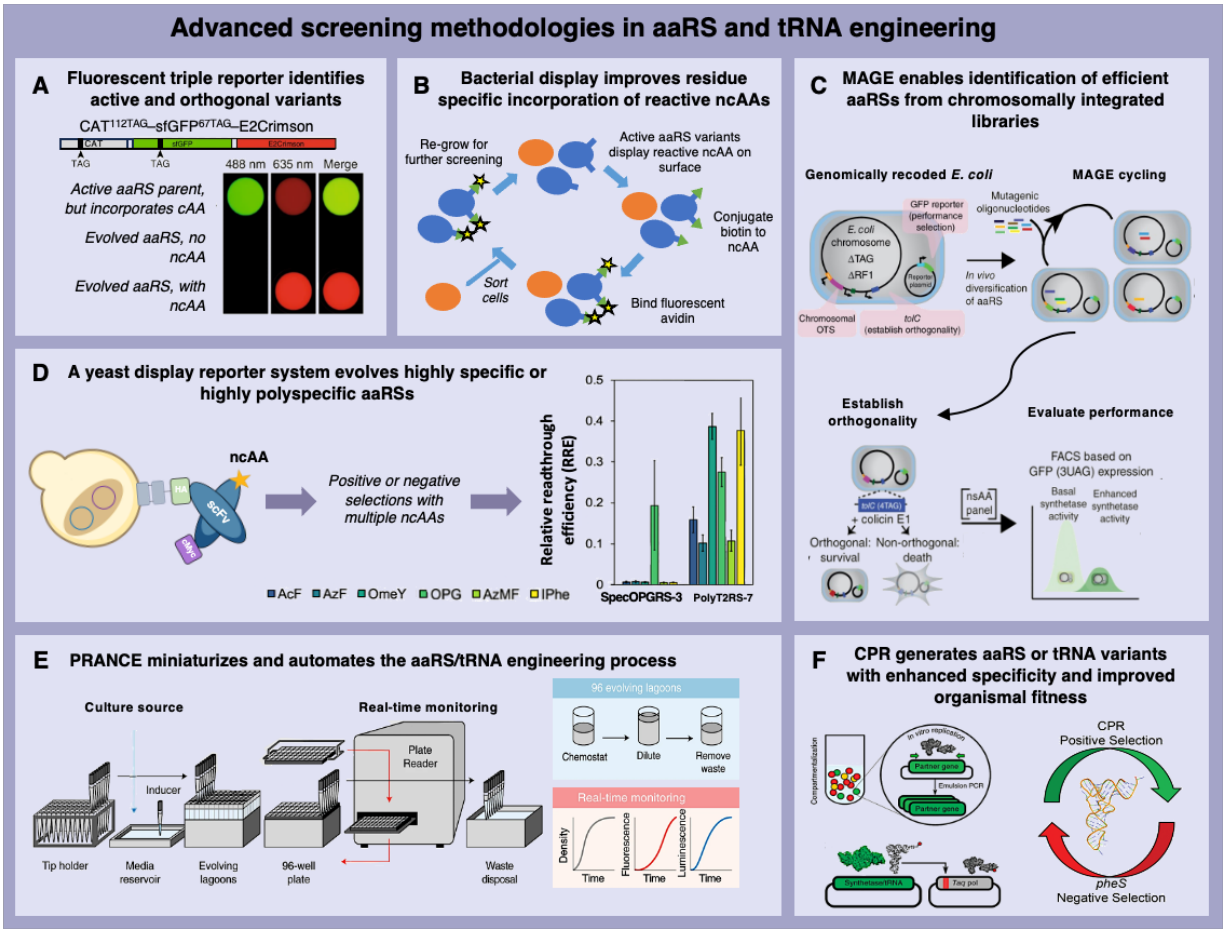

Figure 2. Examples of advanced screening methodologies in aaRS and tRNA engineering. A) A fluorescent triple reporter supports selections and fluorescence-based evaluations of ncAA incorporation and canonical amino acid misincorporation. Reprinted with permission from Ref [63]. *Copyright © 2020, Springer Nature America, Inc.* B) *E. coli* display of reactive amino acids and bioorthogonal conjugation enables improved sense codon reassignment of methionine for the long-chain azide azidonorleucine (Anl) by engineering more efficient MetRS variants. Reproduced with permission from Ref [92]. *Copyright © 2009 National Academy of Sciences*. C) Multiplex automated genome engineering (MAGE) enables screening of chromosomally integrated aaRS libraries. Resulting variants support production of proteins containing thirty ncAAs when used in combination with genomically recoded *E. coli.* Adapted with permission from Ref [11]. *Copyright © 2015, Springer Nature America, Inc*. D) A yeast display reporter[320] enables identification of highly ncAA-specific and broadly polyspecific aaRSs by positive or negative sorting of ncAAs *p*-acetyl-L-phenylalanine (AcF), *p*-azido-L-phenylalanine (AzF), *O*-methyl-L-tyrosine (OmeY), *p*-propargyloxy-L-phenylalanine (OPG), 4-azidomethyl-lLphenylalanine (AzMF), and 4-iodo-L-phenylalanine (IPhe). Adapted with permission from Ref [107] *Copyright © 2022, American Chemical Society.* E) Phage and robotics assisted near continuous evolution (PRANCE) enables 96 simultaneous evolution experiments of translation machinery for genetic code expansion. Reprinted with permission from Ref [13]. *Copyright © 2021, Springer Nature America, Inc.* F) Compartmentalized partnered replication (CPR) uses emulsion PCR to recover functional variants with improved organismal fitness. Application of this platform to aaRS and tRNA engineering yields efficient ncAA incorporation systems and reduced effects on cell viability. Adapted with Permission from Ref [124]. *Copyright © 2017, American Chemical Society.*

Like colorimetric reporters, fluorescent reporters decouple cell growth from ncAA incorporation efficiency. Combined with fluorescence-activated cell sorting (FACS), fluorescent reporters have emerged as powerful tools for implementing strategies for engineering aaRSs (and far beyond—see also the

following sections). Depending on available cell sorting instrumentation, FACS-based screens enable real-time sorting of 20-100 million single cell events per hour based on quantitative readouts of fluorescence and scattering parameters. While this throughput is lower than can be achieved with selection schemes, quantitative measurements and single-cell isolations based on detailed phenotypic criteria make FACS a uniquely powerful tool for enhancing genetic code manipulation systems. Numerous schemes involving fluorescent readouts of ncAA incorporation efficiency and fidelity are described in the literature, starting as early as 2002[77]. Powerful screens enabled by FACS are detailed in the following paragraphs[77, 86, 8788, 89].

While details of fluorescent reporter implementation are distinct between residue- and site-specific genetic code manipulation, the power of these reporters is equally evident in both methods. An early series of studies by Tirrell and coworkers demonstrated multiple strategies for identifying aaRS variants for efficient, residue-specific ncAA incorporation based on fluorescent readouts. One scheme relied on ncAA incorporation and bioorthogonal modification of ncAAs with an *E. coli* cell surface display system (Figure 2B). Azide-containing analogs of methionine were incorporated in response to ATG (methionine) codons within the membrane protein OmpC, after which cells were biotinylated via copper-catalyzed azide-alkyne cycloaddition (CuAAC), stained with a fluorescent biotin detection reagent, and evaluated for fluorescence levels via FACS[90]. Library screens of MetRS variants with this approach and with a smaller library after several adjustments to the scheme led to a range of variants that support efficient replacement of Met with azidonorleucine (Anl)[91, 92]. Detailed kinetic evaluation of MetRS variants revealed several mutants that exhibit Anl activation kinetics that are comparable to the activation kinetics of wild-type MetRS for Met; this is a rare achievement in aaRS engineering. Interestingly, structural characterization of the "MetRS-SLL" variant indicated a switch from an induced fit mechanism to a lock-and-key mechanism following Anl binding, explaining the faster activation[93]. A recent study described a conceptually similar surface display and click chemistry strategy adapted for enhancing aaRS specificity in *E. coli* stop codon suppression[94, 95]. Overall, these bioorthogonal chemistry-based strategies have led to the identification of high-performing, selective aaRS variants.

Although powerful, surface display methods that rely on bioorthogonal conjugations are limited to applications involving ncAAs with suitably reactive side chains. Fluorescent protein reporters and cell surface display reporters that do not depend on chemical reactivity are much more generalizable for aaRS engineering purposes. Yoo et al. engineered a form of GFP that tolerates global replacement of Met with ncAAs and used it to identify MetRS variants that support expanded ranges of ncAA incorporation. This approach was first used to identify variants that support global replacement of Met by 6,6,6-trifluoronorleucine. In this case, saturation mutagenesis within the MetRS active site was sufficient to identify efficient variants[96]. Truong et al. extended this approach to identify MetRS variants that support efficient incorporation of the short-chain alkyne propargylglycine (Pra). Initial screens using saturation mutagenesis in the active site yielded a variant with moderate Pra incorporation that retained some incorporation of Met[97]. Random mutagenesis of this variant followed by FACS-based screening led to identification of a variant exhibiting improved Pra specificity that was attributable to 1) mutations in the KMSKS motif (this motif plays a role in stabilizing the aminoacyl adenylate intermediate in stabilizing tRNA charging[98]); and 2) truncation of MetRS at E548[99]. These studies highlight the benefits of introducing mutations beyond the catalytic active site in evolving high-performing aaRS variants.

Almost all efforts to engineer aaRSs are performed using plasmid-based copies of aaRS genes that result in abnormally high intracellular expression levels of aaRSs. This enables protein biosynthesis with aaRSs that exhibit only modest catalytic properties[100-104]. Multiplex automated genome engineering (MAGE) has been used to evolve chromosomally integrated libraries of aaRSs in *E. coli* for stop codon suppression (Figure 2C). Placing aaRS variants within chromosomal DNA leads to reduced protein expression levels and thus a more stringent screening platform. After aaRS gene diversification and chromosomal insertion, cells are subjected to a tolC negative selection to remove variants that insert cAAs in response to the amber stop codon[11]. The selected library is then screened using a GFP reporter of ncAA incorporation through multiple rounds of FACS. When combined with genomically recoded *E. coli* strains, this powerful workflow led to a variant that supported efficient biosynthesis of a polypeptide

containing 30 copies of the same ncAA in response to the amber stop codon with >95% fidelity[11], which is an impressive feat.

Much of the aaRS engineering work described above was performed in *E. coli*, united by the idea that engineered aaRSs with enhanced ncAA incorporation properties would be challenging to identify with standard selection systems. Recent work in our lab has demonstrated that extensions of these concepts to the yeast *Saccharomyces cerevisiae* leads to powerful opportunities for identifying high-performing aaRS variants. The OTSs suitable for genetic code expansion in yeast are typically derived from *E. coli* aaRS-tRNA pairs, which prevents engineering these systems in most *E. coli* strains (although a handful of engineered *E. coli* strains partially address this issue)[105, 106]. We have used a yeast display approach to engineer numerous aaRS variants. In this approach, a reporter gene encodes a TAG-containing single chain variable fragment (scFv) tethered to the cell by the Aga1p-Aga2p display anchor and flanked by N-terminal and C-terminal epitope tags. An OTS is co-expressed alongside the reporter. If the OTS supports stop codon suppression, the TAG codon is read through, and the C terminal tag can be detected using flow cytometry following labeling of the epitope tags. We used this system to conduct positive and negative screens on libraries of aaRSs to discover variants that support the incorporation of new-to-yeast ncAAs, improve the performance of existing aaRSs, and discover highly specific or highly polyspecific aaRS variants. Notably, this included identification of aaRS variants that support protein biosynthesis with L-DOPA in yeast for the first time. When initially identified "DOPARS" variants from an aaRS active site library exhibited low but detectable incorporation activity, one variant was subjected to error prone PCR and stringent positive and negative enrichments. This led to variants that exhibited efficient ncAA incorporation even at reduced concentrations of ncAA during induction[107]. Additional library screening strategies with carefully controlled induction conditions enabled discovery of highly selective or highly polyspecific aaRS variants (Figure 2D). Resulting clones from selectivity screens support incorporation of *p*-propargyloxy-L-phenylalanine (OPG) into proteins but not any of five structurally related ncAAs; conventional screens did not yield any clones with comparable selectivity. Enrichments to identify polyspecific aaRSs led to variants that incorporate four to six structurally related aromatic ncAAs with high efficiency[107]. Overall, this work highlights opportunities to conduct aaRS engineering in organisms other than *E. coli*.

A limitation of both selections and FACS-based screens is that they typically require multiple rounds of enrichments (positive and negative), and sometimes even multiple rounds of library construction and screening. This can require weeks-long campaigns to identify candidates that support desired levels of ncAA incorporation efficiency and fidelity. Continuous or near-continuous evolution strategies address this challenge by enabling many rounds of selection and diversification to be conducted with minimal intervention from researchers. Phage-assisted continuous evolution (PACE), a technique that links phage propagation to phenotypes of interest, has been applied to aaRS engineering for selective ncAA incorporation in response to stop codons[108, 109]. With this system, Bryson et al. identified *Methanosarcina* spp. PylRS variants that exhibited Kcat/Km values 45-fold greater than the parent enzyme when incorporating N^{ε}-Boc-L-lysine (BocK). In addition, simultaneous positive and negative selections with the structurally similar ncAAs *p*-iodo-L-phenylalanine (pIF) and *p*-nitro-L-phenylalanine (pNF) resulted in a selectivity improvement of 23-fold for pIF compared to off-target ncAAs[109]. Phage assisted non-continuous evolution (PANCE) is a related serial transfer method that does not require the continuous flow instrumentation used for PACE[110]. Implementation of this method by Suzuki et al. led to diverse variants of PylRS that exhibited vastly improved protein biosynthesis with BocK[110].

Notably, a high-throughput version of PACE, phage and robotics assisted near continuous evolution (PRANCE), has recently been reported and applied to OTS engineering (Figure 2E). This method uses liquid handlers and Python scripts to enable scalability, miniaturization, parallelization of up to 96 simultaneous evolution experiments, and real-time activity monitoring to control selection stringency[13]. DeBenedictis and Chory et al. demonstrated PylRS evolutions for enhanced incorporation of BocK using varied levels of stringency, confirming the potential utility of this approach for aaRS engineering. The PRANCE platform has also shown utility in evolving tRNAs with quadruplet anticodons for improved decoding efficiency[13]. The implementation and underlying framework of PRANCE is

expected to be transferable to eukaryotic continuous evolution systems available in yeast or mammalian cells and associated viruses[111], raising the possibility of continuous OTS engineering in these organisms.

Despite impressive progress, most high throughput aaRS selections and screens are suitable only for enhancing incorporation of α-L-amino acids and closely related hydroxy acids[112]. Engineering aaRSs to charge tRNAs with monomers that are poor ribosomal substrates is a largely unexplored area. Recently, Dunkelmann et al. bypassed the ribosome during aaRS engineering by developing an *in vitro* selection scheme that selectively amplifies only acylated tRNAs linked directly to aaRS variants that support acylation[113]. Using a split-tRNA approach, "tRNA display" uses parallel selections in the presence and absence of ncAAs or non-canonical monomers followed by deep sequencing to identify aaRS variants of interest. With this approach, Dunkelmann et al identified orthogonal aaRS-tRNA pairs that are selective for eight new monomers, including β-amino acids, α,α-disubstituted amino acids, and β-hydroxy acids. The fidelity of protein biosynthesis with (*S*)-3-amino-3-(3-bromophenyl)propanoic acid ((S)β^3mBrF) was validated using fluorescence assays and a crystal structure of GFP encoding (S)β^3mBrF at position 13[113].

Another recently reported, ribosome-independent strategy for aaRS engineering is "selects tRNA-acylation without ribosomal translation" (START)[114]. In this approach, Soni et al. used a tRNA barcoding strategy to genetically link aaRS variant sequences to acylation activities of interest; the barcode is inserted directly within the anticodon of the tRNA. Library screens with this approach yielded variants of *M. alvus* pyrrolysyl-tRNA synthetase that support efficient translation with two ncAAs, confirming the feasibility of this selection scheme. The authors further demonstrated that START can detect acylation products with α-hydroxy BocK (OH-BocK) and desamino BocK (H-BocK), suggesting that selections with noncanonical monomers are feasible. The initial reports of both tRNA display and START indicate the promise of engineering translation components that use noncanonical monomers in place of α-L-amino acids.

Overall, advanced selections and screens have yielded new aaRS variants with properties unattainable by traditional live/dead selection schemes. This includes enhanced ncAA incorporation, stringent discrimination against cAAs and structurally similar off-target ncAAs, and polyspecificity to facilitate efficient incorporation of one of several ncAAs. An important emerging theme is that mutations outside of the aminoacylation active site can play a substantial role in enhancing ncAA incorporation efficiency and fidelity. The breadth and depth of screening technologies now available in this area provide exciting opportunities to push the limits aaRS discovery and engineering.

3.2 High Throughput tRNA Engineering

In genetic code expansion, the properties of orthogonal tRNAs are just as important as the properties of orthogonal aaRSs. Since these tRNAs are typically transferred between evolutionarily distant organisms, their performance in their new host cell may be poor; after all, they did not evolve to function as a part of the new host's translation apparatus. There are major unanswered questions regarding the properties of orthogonal tRNAs. For example, it is known that native tRNAs are extensively modified post-transcriptionally[115, 116]. The extent to which these modifications are retained and what their effects on ncAA incorporation might be in a new organism are not well understood[12, 117]. High throughput orthogonal tRNA screening efforts can yield high-performing variants even while gaps in our fundamental understanding of orthogonal tRNAs remain.

Several studies have used conventional live/dead selections in *E. coli* to identify tRNAs that support enhanced ncAA incorporation[72, 73, 118-121]. In fact, some of the earliest reports of ncAA incorporation in response to stop codons described the discovery of orthogonal tRNA variants of *Methanocaldococcous janaschii* (*Mj*) tRNA^{Tyr}[72, 77, 122][118, 123]. While the improvements of isolated variants were generally modest, these early studies highlight the utility of evolving orthogonal tRNAs and have motivated the establishment of more advanced tRNA evolution strategies in recent years.

As discussed in Section 3.1, tuning the stringency of live/dead selections is challenging. Maranhao et al. applied compartmentalized partnered replication (CPR) to tRNA engineering, (Figure

2F)[71, 124]. Importantly, CPR supports both positive enrichments to select for efficient ncAA incorporation and negative enrichments to reduce tRNA mischarging with enhanced control over selection pressures and growth biases compared to prior methods[71]. Directed evolution of the *Mj* tRNA^{Tyr} by Maranhao et al. led to the identification of variants that supported improved ncAA incorporation and 3-fold reduction in promiscuous aminoacylation by endogenous aaRSs. This reduction in promiscuous aminoacylation led to improved *E. coli* doubling times in comparison to the doubling times of *E. coli* expressing earlier generations of Mj tRNA^{Tyr} variants. This suggests that *E. coli* growth steps performed during CPR introduce selection pressure for cell fitness that other high throughput techniques may not[124]. Both these findings and impressive aaRS engineering studies that use CPR demonstrate the power of this technique for enhancing OTSs[71, 83].

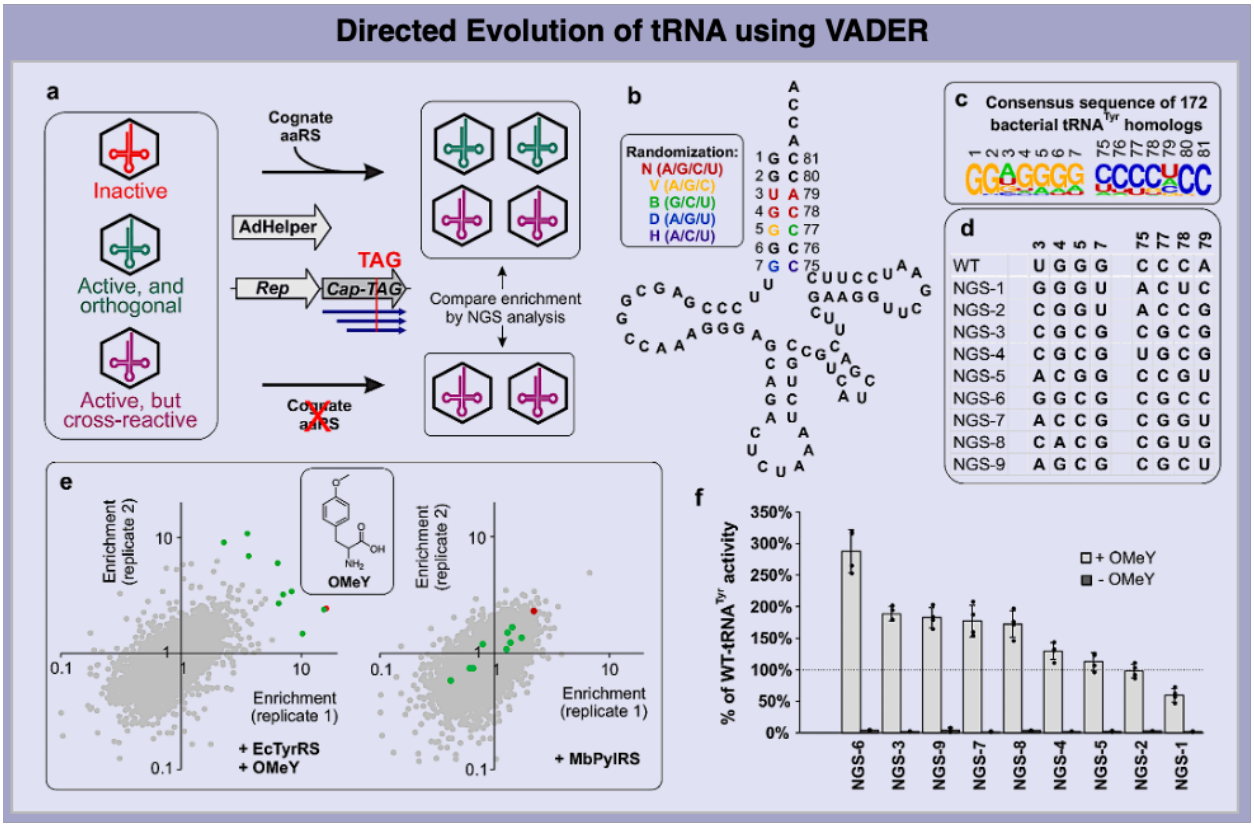

Figure 3. Directed evolution of *E. coli* tRNA^{Tyr} in mammalian cells using the virus assisted directed evolution of tRNA (VADER) system. A) Overview of selection and analysis scheme. Mammalian cells are infected with a collection of AAV2 encoding a tRNA library. Active mutants facilitate generation of AAV2 progeny, and sequences of tRNAs isolated from populations of mammalian cells expressing the orthogonal aaRS and control mammalian cells not expressing the orthogonal aaRS are analyzed by next-generation sequencing (NGS) to identify active and orthogonal tRNA variants. B) Library design for the *E. coli* tRNA$^{Tyr}_{CUA}$ mutant library. C) Consensus sequence of homologous tRNA^{Tyr} genes from 172 bacterial species that guided library design. D) Sequences of the most highly enriched tRNA^{Tyr} variants following selection with EcTyrRS. E) Enrichment after VADER scheme following selection with EcTyrRS charging *O*-methyl-L-tyrosine (OmeY) and MbPylRS that does not recognize tRNA^{Tyr}. Each selection was performed in duplicate; normalized enrichment factors are plotted against each other. Nine out of ten enriched sequences with EcTyrRS did not show strong enrichment with MbPylRS (green); a putatively cross-reactive mutant is depicted in red. F) Efficiency of ncAA incorporation of nine variants using the reporter protein EGFP-39TAG. Each tRNA$^{Tyr}_{CUA}$ variant was co-transfected into HEK293T cells with EcTyrRS and EGFP-39TAG in the presence or absence of 1 mM OmeY. Expression of EGFP was measured in cell extract and normalized to wildtype mCherry fluorescence. N = 3 independent experiments. *Reproduced with permission from ref [12]. Copyright © 2023 Springer Nature America, Inc.*

Due to species-to-species changes in the components of the protein translation apparatus, it is highly desirable to evolve orthogonal tRNAs directly in the final expression host. However, systems for evolving tRNAs (and other orthogonal components) in mammalian cells have been difficult to implement. In particular, challenges in robustly linking genotype and phenotype and generating suitably diversified tRNA libraries have all hindered screening and selection efforts[125]. To overcome these deficiencies, Jewel et al. established the groundbreaking platform "virus assisted directed evolution of tRNA" (VADER) (Figure 3). In this system, the activity of a suppressor tRNA is coupled to the replication of adeno-associated virus 2 (AAV2), a virus that can be used to deliver genetically encoded libraries to mammalian cells[126]. After delivery of a tRNA library, use of a stop codon inserted into the essential Cap gene of AAV2 facilitates selective amplification of tRNA variants that support stop codon suppression. Since this selective pressure does not discriminate between ncAA incorporation and mischarging of cAAs, an additional (but optional) step selects for orthogonality by using bioorthogonal chemistry to biotinylate AAVs containing reactive ncAAs[12]. With this approach, the researchers isolated numerous variants of the *E. coli* $tRNA^{Tyr}_{CUA}$. Characterizations of variant ncAA incorporation activities revealed substantially enhanced performance in mammalian cells. Notably, these improvements did not transfer over to *E. coli*: use of the top-performing A2.1 suppressor $tRNA^{Tyr}_{CUA}$ in an engineered *E. coli* strain led to ncAA incorporation activities that were unchanged from the activities with the parent tRNA[12]. This striking observation highlights the importance of evolving tRNAs in the desired protein biosynthesis host. Jewel et al. and Huang et al. recently reported advances that support screening of larger libraries and extension of the VADER concept to *E. coli* $tRNA^{Leu}$[127, 128]. Overall, tRNA engineering is a powerful but underutilized approach to enhancing the performance of genetic code manipulation systems.

3.3 Data-Driven Approaches to the Discovery of High-Performing aaRS and tRNA Variants

Despite the breadth of efforts focused on high throughput aaRS and tRNA engineering, only a small subset of these studies have leveraged data-driven approaches to identify high-performing variants. Rapid increases in the number of sequenced genomes[129], continued advances in deep sequencing[130], advances in machine learning models in biomolecular engineering[131-136], and impressive advances in computational protein design[137-139] all point to substantial opportunities to enhance the aaRSs and tRNAs used in genetic code manipulation systems. These opportunities are especially evident in genetic code expansion systems, where an extremely limited number of OTSs have been used in the vast majority of studies[17].

A few reports have leveraged genomic data to discover novel candidate OTSs. One such study by Qin et al mined sequencing data for thermostable OTSs by identifying genes encoding aaRS/tRNA pairs from organisms with high optimal growth temperatures. This approach resulted in identification of previously unknown PylRS/$tRNA^{Pyl}$ pairs from *Methanosarcina thermophila* and *Methanosarcina flavescens* that were determined to exhibit orthogonality to the native translation machineries of *E. coli* and mammalian cells[140]. Both pairs possess higher melting temperatures than their mesophilic counterparts[140, 141] that potentially make them more amenable to further engineering.

The Chin lab has established multiple strategies for mining genomic data and computationally prioritizing promising subsets of candidate sequences. These approaches identify candidate sequences that may be mutually orthogonal to existing OTSs as well as the native aaRS/tRNA pairs in *E. coli*[62, 142]. In one study, millions of tRNA genes from bacteria, archaea, chloroplasts, and bacteriophages were computationally filtered with a scoring system based on tRNA identity elements to narrow in on a limited but diverse set of candidate tRNAs that may exhibit orthogonality. These candidates were then evaluated using a tRNA extension (tREX) assay to investigate tRNA expression and aminoacylation with native polyacrylamide gel electrophoresis readouts[63]. Using this pipeline, 71 out of 243 candidate tRNAs evaluated met assay criteria for expression and orthogonality in *E. coli*. Further investigations of the activities of corresponding aaRSs led to the identification of five aaRS/tRNA pairs that exhibit

orthogonality with respect to the native E. coli machinery and with respect to the other newly identified pairs.

Another study combined genomic data with experimental information on PylRS/tRNA^{Pyl} cross-reactivity to identify mutually orthogonal PylRS-tRNA pairings[61]. Separate criteria for sequence identity thresholds of PylRS homologs and corresponding tRNA homologs were established to determine likelihood of crossreactivity. Agglomerative clustering of 351 PylRS sequences and further analysis of corresponding tRNA sequences identified candidate pairs likely to exhibit mutual orthogonality to one another. This led to the identification of two new classes of PylRSs, Class C and S. Numerous pairs tested in *E. coli* exhibited activity, orthogonality to the native *E. coli* translation apparatus, and orthogonality to previously reported classes of PylRS systems (Class N, A, and B)[61]. These informatics-driven analyses deftly exploit the wealth of genomic data to greatly increase the number of "starting points" for further engineering of OTSs.

For aaRSs and tRNAs, moving beyond data mining to computational design and machine learning remains a largely unexplored area. It is challenging to generate functional, diverse collections of aaRSs, as aminoacylation active sites can include more than 10 residues that contact the amino acid and numerous more distant residues that affect aaRS activity. Developing computational and machine learning pipelines could lead to vastly improved navigation of these complex sequence spaces[131-139]. A handful of early studies described attempts to computationally design aaRSs for enhanced activity and specificity[65, 143]. More recent efforts by Baumann et al. used Rosetta computational design in combination with live/dead selections to identify a variant of *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase that outperforms all other reported variants for the incorporation of *O*-nitrobenzyltyrosine (ONBY) in *E. coli*[144]. The rapidly expanding tools in biomolecular computational design, machine learning, and protein structure prediction offer powerful opportunities to push the performances of aaRSs and OTSs[133, 139, 145]. While we are not aware of their application to engineering OTSs, the undeniable success of machine learning tools – for instance, successful "grafting" of functional motifs onto scaffolds[139], antibody affinity maturation based only on sequence input[146], and prediction of enzymes with altered activity[147] – merits serious investigation for OTS engineering. These tools will be most effective when combined with the rapidly expanding collection of screening and selection platforms that have already started to push genetic code manipulation systems to new heights.

4. Engineered Translation Apparatuses for ncAA Incorporation

Native translation apparatuses did not evolve to efficiently incorporate ncAAs. While all forms of genetic code manipulation require careful implementation, the challenges posed by genetic code expansion in cells are particularly daunting. This section highlights progress in manipulating elements of the translation apparatus beyond OTSs to support ncAA incorporation, including elongation factors, release factors, ribosomes, and stop codon context. In many cases, precise manipulations provided initial evidence for the potential benefits of engineering translation apparatuses before the implementation of high throughput screens. An advantage of these screening approaches is that identification of high-performing variants does not require mechanistic understanding. In addition to enhancing discrete elements of protein translation apparatuses, genome engineering has emerged as a promising (but resource-intensive) approach to establishing prokaryotic and eukaryotic cells that support genetic code expansion efforts. While genome synthesis efforts are driven by rational and computational approaches, adaptive laboratory evolution of engineered *E. coli* strains has played an important role in improving the fitness of resulting strains[148]. In addition, a recent genome-wide screen for enhanced ncAA incorporation in yeast has indicated that deletion of genes with no known connection to the protein translation apparatus can enhance the efficiency of genetic code expansion.[149] These results, described further in the subsections below, indicate the power of high throughput screening to engineer cells to better accommodate alternative genetic codes.

4.1 Elongation Factor Engineering

Seminal work by Schrader et al. led to critical insights about elongation factor Tu (EF-Tu) and its interactions with aminoacylated tRNAs. Elongation factor Tu (EF-Tu) binds to and shuttles aminoacylated tRNAs to the ribosome; it also discriminates against misacylated tRNAs. Rigorous measurements of *E. coli* EF-Tu binding to aminoacylated tRNAs revealed that 1) AA-tRNA affinities all fall within a narrow range of affinities; and 2) increasing the affinity of AA-tRNA for EF-Tu can actually decrease the rate of peptide bond formation[150]. These insights have important implications for establishing translationally efficient OTSs because ncAA-tRNAs can interact with the host EF-Tu with altered affinities. An early rational study by Doi et al. confirmed the importance of EF-Tu in the incorporation of bulky ncAAs into proteins[151]; high throughput screening has been used to further engineer EF-Tu in *E. coli*.

O-phosphoserine (Sep) is a key posttranslationally modified form of serine in eukaryotes, but in certain species of methanogenic archaea Sep is inserted into proteins during translation with a dedicated aaRS (SepRS) and $tRNA^{Sep}$. However, the introduction of an OTS to add Sep to the genetic code of *E. coli* is not straightforward because native *E. coli* EF-Tu does not bind to Sep-tRNA with sufficiently high affinity to support translation. Park et al. completely randomized six residues within *E. coli* EF-Tu and selected for variants that support protein biosynthesis with Sep incorporated in response to the amber codon. This approach resulted in the discovery of a mutant form of EF-Tu, called EF-Sep, that enabled efficient incorporation of Sep during protein biosynthesis for the first time (Figure 4A; in combination with SepRS from *Methanococcus maripaludis* and $tRNA^{Cys}_{CUA}$ from *Methanocaldococcus jannaschii*)[79]. A study utilizing the bioinformatics approach known as reconstruction of evolutionary adaptive paths (REAP) identified further mutants of EF-Tu that support incorporation of Sep and other structurally related ncAAs[152].

EF-Tu engineering has also been used to enhance the performance of the commonly used *Mj* tyrosyl-tRNA synthetase/$tRNA^{Tyr}$ pair. Gan et al. used combinatorial alanine scanning to diversify the binding pocket of EF-Tu and screen for enhanced AzF incorporation with a GFP reporter. One variant was reported to improve incorporation by 4-fold over wildtype EF-Tu[104]. The identified mutations are hypothesized to alleviate possible steric hindrance of the tRNA acceptor stem in the EF-Tu binding pocket. These high throughput EF-Tu engineering approaches demonstrate the value of fine tuning the properties of elongation factors for enhanced ncAA incorporation systems.

4.2 Release Factor Engineering

NcAA incorporation via stop codon suppression is always in competition with release factor-mediated termination of translation. In *E. coli* and other bacteria, there are two release factors that independently recognize stop codons: RF1 recognizes UAA and UAG, whereas RF2 recognizes UAA and UGA. Johnson et al., Robertson et al., Kuru et al., Yi et. al, Zheng et al., and others have demonstrated that interfering with the expression or function of *E. coli* RF1 results in enhanced ncAA incorporation in response to the amber codon[153-157]. When combined with genome engineering strategies (described below), deletion of RF1 has enabled the construction of *E. coli* strains with greatly enhanced readthrough of amber codons. These findings demonstrate the feasibility of engineering the release factor activities of organisms to facilitate the use of expanded genetic codes.

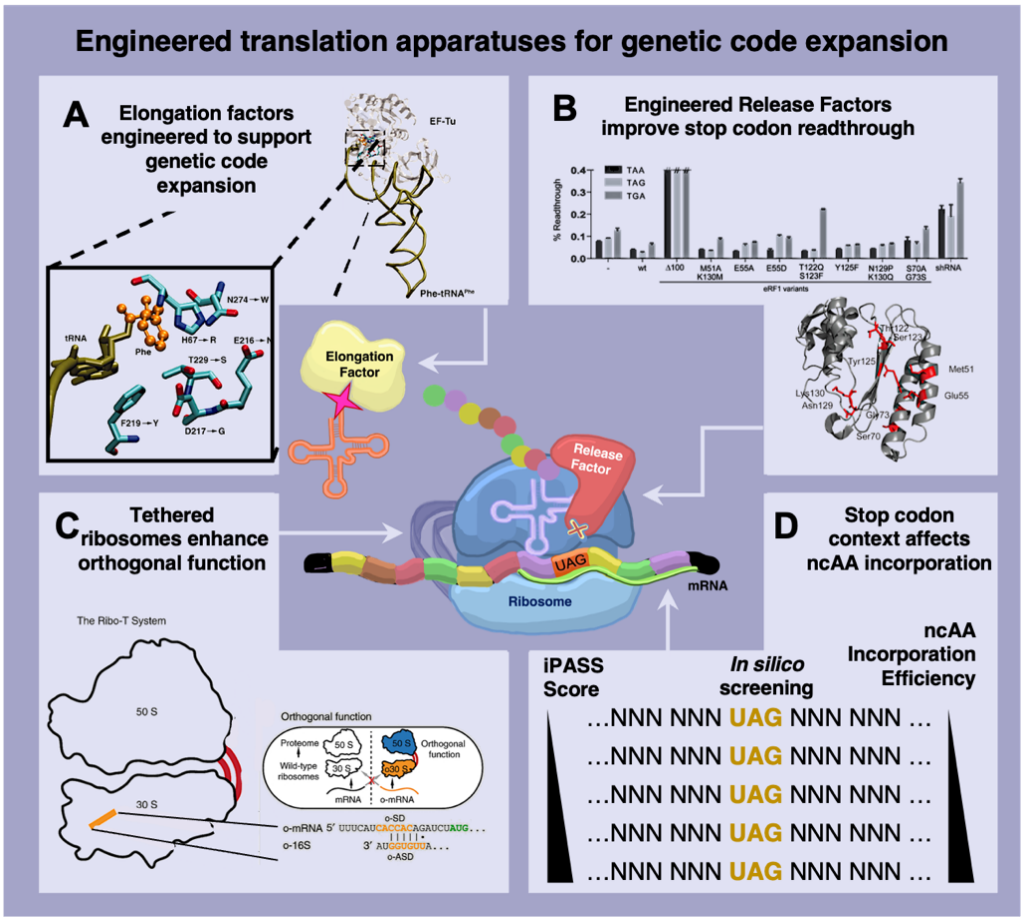

Figure 4. Engineered translation apparatuses for genetic code expansion. A) Ef-Tu mutations selected for improved phosphoserine (Sep)-tRNA^{Sep} binding. Adapted with permission from Ref [79]. *Copyright © 2011 American Association for the Advancement of Science.* B) Eukaryotic release factor 1 (eRF1) variants shown to exhibit improved stop codon readthrough and ncAA incorporation. Adapted with permission from Ref [172]. *Copyright © 2014 American Chemical Society.* C) Tethered ribosome system evolved for improved cellular fitness and orthogonality. Adapted with permission from Ref [198]. *Copyright © 2019 Springer Nature.* D) Screening data from engineered cell lines validated the Identification of Permissible Amber Sites for Suppression (iPASS) system's ability to predict and enhance ncAA incorporation efficiencies based on stop codon sequence context. Adapted with permission from Ref [204]. *Copyright © 2021 Oxford University Press.*

In contrast to bacteria, eukaryotic cells employ a single omnipotent release factor (eRF1) that recognizes all three stop codons. The lack of redundancy in the essential *SUP45* gene (which encodes eRF1) indicates opportunities to mutate, but not delete, this factor. There is a growing body of literature describing eRF1 variants with altered recognition capabilities in naturally occurring organisms[158-166]. In addition, multiple studies have used engineered release factors for ncAA incorporation. Merritt et al. evaluated a collection of eRF1 mutants that were previously identified to exhibit defective stop codon decoding *in vitro*[167-169], or identified through an *in vivo* suppressor screen[170]; they reported eRF1 mutants enhancing stop codon readthrough at all three stop codons[171]. Schmied et al. introduced mutations via site-directed mutagenesis into human eRF1 (Figure 4B)[172]. The resulting variants included E55D and Δ100 (broad deletions in the N-terminus of eRF1). When expressed in parallel with wild-type eRF1 in a

mammalian cell line, ncAA incorporation efficiencies increased 5-7 fold to 27% and 36%, respectively. The exogenously expressed E55D mutant has been successfully employed to improve mammalian ncAA incorporation in numerous additional studies[64, 173-182]. These findings suggest that high throughput screening approaches may lead to the identification of additional release factor variants that enhance ncAA incorporation.

4.3 Ribosome Engineering

Native ribosomes exhibit substantial versatility, supporting incorporation of diverse alpha amino acids into proteins beyond the 20 canonical amino acids. However, there are limits to the sizes of alpha amino acids that prokaryotic and eukaryotic ribosomes can tolerate, and most backbone structures beyond the alpha amino acid backbone are accepted either poorly or not at all by native ribosomes[183-185]. In addition, ribosomes primarily recognize 3-base codons, hindering the use of quadruplet codons in ncAA incorporation systems. There is demand for engineered ribosomes that are compatible with diverse ncAAs and coding schemes. Several groups have taken up the daunting challenge of ribosome engineering to better accommodate expanded genetic codes[186-190].

A major challenge in modifying the ribosome is that its native protein translation functions are essential for cell growth and survival. The creation of an orthogonal ribosome that functions in parallel with native ribosomes is a key enabling strategy for ribosome engineering. Pioneering work by Rackham and Chin led to the identification of orthogonal mRNA and 16s rRNA sequence pairs in *E. coli* via tunable positive and negative selections[191]. This was an essential step toward orthogonal ribosomes, as alterations led to an engineered small ribosomal subunit that predominantly recognized “orthogonal” messenger RNAs. Selection pressure was achieved by fusing a gene encoding chloramphenicol acetyltransferase (*cat*) and uracil phosphoribosyltransferase (*upp*) downstream of a ribosome binding site. Cells expressing the fusion construct (*cat-upp*) were able to survive the positive selection in the presence of chloramphenicol (due to CAT activity) enriching for ribosomes capable of translating mRNAs containing altered Shine-Dalgarno sequences. Conversely, negative selection with 5-Fluorouracil (5-FU) eliminated mRNA sequences that were substrates for endogenous ribosomes. Together, these selection systems enable the assessment of 10^9 more sequences than would be possible by rational design for the discovery of orthogonal ribosome-mRNA pairings.

Altered ribosome and mRNA elements enabled the Chin lab to further engineer the small subunit of the *E. coli* ribosome to support improved genetic code enhancement. In an effort by Wang et al., selections led to the identification of a 16s rRNA variant that improved ncAA incorporation in response to the amber codon from ~20% to ~60% efficiency with an established OTS[192]. This approach, which focused on reducing ribosomal interactions with *E. coli* RF1, was an important demonstration of ribosomal engineering, although later studies have shown than deleting RF1 from the genome leads to even larger benefits (see Section 4.2). In another noteworthy study, Neumann et al. identified ribosomal variants that support enhanced readthrough of both quadruplet codons and amber codons, enabling improved dual incorporation of azide- and alkyne-containing ncAAs during protein biosynthesis in *E. coli*[193]. Selections with orthogonal ribosomes identified mutants that can accommodate extended (4-base) anticodon-codon interactions; this was only possible with the availability of high throughput selection systems.

Engineered small ribosomal subunits and mRNAs are not sufficient to establish a fully orthogonal ribosome, as the bipartite nature of ribosomes enables exchanges between “orthogonal” and endogenous subunits. This leads to crossover between the translation of wild-type and orthogonal mRNAs, resulting in loss of orthogonality. Orelle et al. and Fried et al. have both reported engineered *E. coli* ribosomes in which the small and large ribosomal subunits are linked via RNA tethers[194, 195]. In the approach of Orelle et al., multiple selections were used to engineer tethered ribosomes. This included 1) a library of circularly permuted variants of the gene encoding the 23s rRNA to identify translationally competent, permuted variants; 2) a library of tethers varying in length to insert a circularly permuted 23s rRNA

within the 16s rRNA while retaining translational activity (called Ribo-T); and 3) adaptive laboratory evolution (ALE) to improve the doubling time of the *E. coli* strain SQ171 containing the monolithic tethered ribosome for protein translation (no wild-type ribosomal rRNA)[194]. Interestingly, following ALE, only genomic mutations were identified in strain SQ171fg, hinting that the genomic context plays an important role in determining efficiency of translation in cells containing Ribo-T (Figure 4C)[37, 194, 196]. Fried et al. used structural data and phylogenetic analysis to inform a strategy that also involved circularly permuting the 23s rRNA and inserting it into the 16s rRNA, but did not use library selections or adaptive laboratory evolution to generate monolithic ribosome variant O-ribo(h44H101)[195]. Subsequent reports have identified further improvements to both Ribo-T and O-ribo(h44H101), in which contacts between small and large ribosomal subunits are mutated to enhanced orthogonality of tethered ribosomes[186, 196-198]. Notably, the Jewett lab has demonstrated the feasibility of using improved versions of Ribo-T for ncAA incorporation[199]. This evidence combined with upgraded systems for designing and evolving O-ribosomes raises the possibility of engineering the *E. coli* ribosome to synthesize polymers containing noncanonical monomers that are poorly tolerated by native ribosomes.

4.4 Stop Codon Context

Recent work has shown that the termination of translation is dictated by numerous factors in addition to the literal presence of a stop codon. For example, in eukaryotes, the nucleotide following a stop codon is biased towards purines[200], and the proximity of the Poly(A)-Binding Protein (PABP) to the eukaryotic release factor complex stimulates translation termination[201]. This has important implications for genetic code expansion efforts, as it suggests that the context of an ncAA incorporation site can play an important role in dictating incorporation efficiency. In our group, Potts et al. and Lahiri et al. investigated a small set of stop codon contexts in yeast[69, 202]. Varying UAG or UGA stop codon position within a short linker led to substantial changes in ncAA incorporation efficiency. Higher throughput approaches in bacteria and eukaryotes have further elucidated features of codon context that influence ncAA incorporation efficiency and termination of translation.

Pott et al. used a directed evolution approach to investigate what sequence contexts resulted in the highest levels of amber suppression in *E. coli.* A random library of nucleotides consisting of two codons upstream and two codons downstream of an amber codon (NNK-NNK-TAG-NNK-NNK) was subjected to iterative rounds of positive selection until sequence convergence was observed. Two frequently isolated clones supported 70-110% increases in protein yields in comparison to controls[203], indicating the existence of favorable sequence contexts for genetic code expansion in bacteria.

Bartoschek et al. used high throughput experimentation in mammalian cells to establish a model for predicting stop codon readthrough efficiency based on codon context. Mouse embryonic stem cells (mESC) and human embryonic kidney HEK293T cells were genomically engineered to stably express an orthogonal PylRS/$tRNA^{Pyl}_{CUA}$ pair. These cell lines were then subjected to a modified version of stochastic orthogonal recoding of translation with enrichment (SORT-E) to characterize and enrich the amber suppressed proteome. The incorporation of ncAAs with bioorthogonal handles enabled selective conjugation of biotin tags to C-terminally extended proteins. Proteomics studies and additional characterizations enabled development and validation of Identification of Permissive Amber Sites for Suppression (iPASS) to predict ncAA incorporation efficiencies as a function of sequence context (Figure 4D)[204]. In a conceptually related approach, Ding et al. explored ncAA incorporation via rare codon recoding, noting that recoding rates of a tetrazine-containing-ncAA (TetBu) in response to TCG varied significantly as a function of flanking sequence context. Proteomics investigations enabled identification and estimation of the abundance of 1514 ncAA-containing peptides (and the identities of the surrounding sequence contexts). Randomization of the two codons on either side of a TCG codon in a reporter gene and measurements of recoding rates via LC-MS generated data used to refine a model predicting rare codon recoding rates. Overall, this study improved rare codon recoding from approximately 8% to over 95% in some instances[205]. These studies indicate the importance of sequence context for ncAA incorporation in response to a stop codon or a rare cognate codon. The impressive predictive capabilities

of the resulting models can be used to design favorable contexts to facilitate more efficient ncAA incorporation in mammalian cells.

Several additional studies of stop codon context highlight the nuanced importance of context, albeit without the use of ncAA incorporation. Anzalone et al. used a eukaryotic form of mRNA display to identify RNA elements downstream of stop codons that enhance stop codon readthrough. Interestingly, deep sequencing analysis revealed both primary sequence features and structural elements that were enriched. The sequencing dataset was used in combination with machine learning to predict human genes with 3' untranslated regions that would lead to elevated stop codon readthrough; several predictions were subsequently verified[206]. A number of ribosome profiling studies have led to complementary insights into features that influence stop codon readthrough, including the stop codon tetranucleotide (+1 to +4), the codon in the P site (-3 to -1), and the length of the 3' untranslated region[207-209]. These insights offer important opportunities to enhance genetic code expansion systems and further elucidate the rich biology underlying termination of translation.

4.5 Genome Engineering

4.5.1 Genome Engineering, Adaptive Evolution, and Genomic Diversification to Enhance Strains for ncAA Incorporation

Genome engineering efforts aimed at enhancing ncAA incorporation have progressed significantly over the past decade. Multiple groups have reported codon-compressed *E. coli* strains (63-codon and 61-codon genomes), and an international consortium has made substantial progress toward constructing a yeast strain with a 63-codon genome (Sc2.0)[7, 22, 210-216]. In contrast to high-throughput screening, even one reworked genome requires a large-scale effort, and resulting codon-compressed strains can exhibit slow doubling times compared to conventional strains. Given the low throughput of genomic manipulations, pinpointing specific mutations and synthetic lethal combinations that lead to slower doubling times is extremely challenging. Unbiased adaptive laboratory evolution (ALE) efforts have played an important role in improving the fitness of codon-compressed strains. In ALE, a strain is subjected to repeated passages over months (or years) in an attempt to improve its growth characteristics, with the assumption that beneficial mutations in the genome will spontaneously accumulate in faster-growing strain variants. Both Wannier et al. (63-codon strain) and Robertson, Funke, de la Torre, and Fredens et al. (61-codon strain) used ALE to improve growth characteristics of codon-compressed strains[148, 154]. These studies reported strains that contained multiple genomic changes that improved growth characteristics. While a detailed discussion of the specific mutations is beyond the scope of this review, both of these studies make it clear that the genomic modifications would have been difficult or impossible to predict rationally. The codon compression of chromosomes in the yeast genome is now well underway, along with substantial tool development to identify and address the genetic changes that lead to growth defects[217, 218]. A unique feature of the Sc2.0 effort is the introduction of recombination sites to facilitate rapid generation of genomic diversity throughout the genome, which could serve as a route to investigate how novel genome rearrangements lead to improved ncAA incorporation efficiency or other complex phenotypes of interest[212, 219]. These synthetic genome construction efforts clearly indicate that the use of synthetic genomes to push the limits of genetic code expansion will benefit greatly from high throughput screens and selections in years to come.

4.5.2 Genome-Wide Screens to Enhance ncAA Incorporation

A reworked genome is not essential for investigating genomic changes that enhance ncAA incorporation. In fact, there are incredible sets of genetic and genomic resources available in model organisms such as *E. coli* and *S. cerevisiae*. Based on this premise, we recently reported a genome-wide screen for enhanced ncAA incorporation in *S. cerevisiae*. In this study, Zackin et al. transformed a pooled, barcoded set of yeast genetic knockout strains with reporters of ncAA incorporation and OTSs, using FACS to identify clones exhibiting increased levels of the full-length reporter. This led to the identification of 55 candidate gene deletions that may enhance ncAA incorporation. Surprisingly, many of

the identified genes have no known role in protein synthesis. Two frequently identified deletions were of the gene YIL014C-A (unknown function) and the gene ALO1 (protects against oxidative stress). Studies of these deletions in standard genetic backgrounds confirmed that the knockouts support improved ncAA incorporation efficiency with several OTSs[149]. This study is significant because it strongly suggests that there are numerous genomic changes with the potential to enhance genetic code expansion beyond changes that can be predicted rationally. Discovery of unexpected factors that alter ncAA incorporation also provides opportunities to identify and investigate mechanisms of translation (and ncAA incorporation) more deeply. Overall, genome-wide screens are a powerful, complementary approach to genome engineering that can uncover new strategies for accommodating alternative genetic codes in cells[220].

4.5.3. Organisms with Naturally Evolved Genetic Codes

While laboratory efforts to engineer the genome are recent, selection pressures over millions of years have led to the natural evolution of various organisms and organelles that incorporate amino acids at one or more of the canonical stop codons, from ciliated protozoa to vertebrates and mitochondria[221, 222]. The genomic changes to accommodate these altered genetic codes may help to inform ongoing codon compression efforts in prokaryotes and eukaryotes. Furthermore, organisms (and organelles) with alternative genetic codes encode protein translation apparatuses that may be valuable for implementing genetic code manipulation strategies. For example, codon reassignment is frequently discovered alongside mutated forms of release factors in organisms with nonstandard genetic codes. In the recently characterized organism *Blastocrithidia nonstop*, release factor eRF1 contains several mutations that alter its recognition of stop codons. The S74G mutation (*B. nonstop* numbering) was introduced at equivalent positions within *S. cerevisiae* and human eRF1 homologues and examined for stop codon readthrough in yeast. Expression of the mutants significantly increased UGA readthrough without altering readthrough at UAG or UAA codons[223], suggesting that this variant may be a valuable starting point for additional translation apparatus engineering. This study and numerous others highlight the value of investigating organisms with divergent genetic codes to identify “nature-tested” strategies for genome engineering in combination with ncAA incorporation[224].

4.6 Combinatorial Experimentation for Genetic Code Manipulation with Cell-Free Protein Synthesis Platforms

In addition to naturally evolved and laboratory-engineered cellular platforms, cell-free protein synthesis platforms offer compelling strategies for expanding the capabilities of genetic code manipulation. These systems are freed of the requirement to support cellular replication and other cellular processes[187, 225], enabling investigation of a much broader range of conditions than can be surveyed within cells. One emerging area of opportunity is the ribosomal synthesis of polymers from monomers other than α-L-amino acids (noncanonical monomers)[9, 37, 183, 226-230]. While a complete discussion of advances in cell-free systems is beyond the scope of this review, we highlight the power of in vitro combinatorial studies to tailor translation apparatuses for syntheses that natural translation apparatuses cannot perform. Two important example are 1) adding a mutated form of EF-Tu to the PURE in vitro translation system to improve ribosomal protein synthesis with negatively charged d-amino acids[231]; and 2) evaluation of orthogonal tRNA compatibility with the *E. coli* translation apparatus to discover new tRNAs that support ncAA incorporation in response to UAG and UAA codons[232]. Systematic variations and experimental evaluations of the in vitro translation apparatuses were essential to identify the most performant genetic code manipulation approaches in each of these studies. Higher throughput explorations of these cell-free systems may provide powerful insights into protein biosynthesis with ncAAs and noncanonical monomers more broadly[233-235]. Overall, implementing high throughput screening strategies with protein translation apparatuses, genomic manipulations, and cell-free systems offer powerful routes to improve the performance of genetic code manipulation systems and elucidate molecular mechanisms of translation.

5. High Throughput Screening with Chemically Expanded Proteins and Peptides

The availability of high-performing genetic code manipulation systems has enabled the high throughput discovery of proteins and peptides with unique, ncAA-mediated properties. It is also important to recognize the power of accessing distinct chemical and structural features in proteins using rational or semi-rational insertion of ncAAs; the reader is referred elsewhere to learn about cases where pinpointed manipulations have yielded important biological insights[15, 21, 236-241]. However, not all rational ncAA substitutions result in the expected changes to protein properties. The interplay between a ncAA, its local context within a protein of interest, and its broader context amongst surrounding biomolecules and solvent may result in complex properties that only become evident during experimentation[242-244]. In these cases, the vast range of tools established for conventional protein engineering can be adapted for high throughput screening with ncAA-containing proteins. The discovery of the unexpected is increasingly routine with chemically expanded polypeptides. This includes proteins that tolerate numerous ncAA substitutions, proteins with enhanced stability, enzymes with new-to-nature catalytic mechanisms, and chemically expanded peptide and protein binding ligands.

5.1 Engineering Proteins that Tolerate Noncanonical Amino Acid Substitutions

There are many cases when ncAA substitutions are detrimental to protein stability or activity[15, 245-250]. Proteins in which a ncAA globally replaces a canonical amino acid via sense codon reassignment are especially prone to decreases in fitness. Montclare et al. observed this phenomenon when fluorinating the enzyme chloramphenicol acetyltransferase (CAT): replacement of the 13 leucine residues with the ncAA 5',5',5'-trifluoroleucine (TFL) resulted in a 20-fold reduction in CAT's half-life of thermal inactivation[245,251]. Two rounds of random mutagenesis and screening of fluorinated variants in 96-well plates led to the identification of a mutant exhibiting a 27-fold improved half-life of thermal inactivation. Inspection of the six mutations in the final variant revealed that no mutations were adjacent to TFL incorporation sites or within 15 angstroms of the CAT active site. These observations parallel insights gained from conventional protein engineering: mutations located far from sites of interest can lead to enhanced protein properties via mechanisms that are not readily apparent from structural data.

Yoo et al. extended this line of investigation to green fluorescent protein (GFP), where replacement of Leu with TFL in GFP was highly deleterious to fluorescence. Eleven rounds of random mutagenesis and screening via fluorescence-activated cell sorting (FACS) resulted in a highly fluorescent, fluorinated GFP; this variant contained 20 amino acid mutations and a 650-fold increase in fluorescence compared to the fluorinated parent. Analysis of protein folding kinetics revealed vastly enhanced folding of both the fluorinated and nonfluorinated forms of the evolved mutant, with the folding kinetics of the Leu-containing variant approaching the kinetics of "superfolder" GFP[252]. This is a striking result, as no explicit attempt to enhance folding kinetics was part of the evolution campaign. It appears that the structural perturbations caused by global fluorination of Leu were sufficient to promote the acquisition of mutations that enhance protein folding. While further investigations to evaluate potential generality are needed, the available evidence suggests that ncAA substitutions followed by directed evolution can identify protein variants that are more fit than parental sequences.

5.2 Enhancing Protein Stability with ncAAs and High Throughput Screening

In some cases, carefully placed ncAA substitution can lead to stabilized proteins. This includes several examples of computationally designed, "stapled" proteins stabilized via spontaneous crosslinking[253]. An alternative approach is to use high throughput screening to determine which ncAA(s) and substitution site(s) leads to enhanced protein stability. The primary advantage of this strategy is that it allows for unbiased approaches to identify candidate variants. Li et al. sought to increase the thermal stability of homoserine O-succinyltransferase (MetA), an enzyme that plays a key role in methionine biosynthesis and *E. coli* survival. A library of MetA variants encoding stop codons for site-specific ncAA incorporation was prepared and transformed into *E. coli* alongside one of several OTSs for ncAA incorporation. This enabled the library to be subjected to selections following incorporation of each of 12

distinct ncAAs. Growth-based selections at nonpermissive temperatures led to the identification of two different mutants exhibiting improved growth. One variant containing (p-benzoylphenyl)alanine (pBzF) resulted in a 21 °C increase in the melting temperature of MetA, while another variant containing *O*-tert-butyltyrosine (OtBuY) resulted in a 6 °C increase in the melting temperature[254]. Follow-up characterizations of the pBzF variant indicated a possible covalent interaction between a nearby cysteine residue and pBzF. A follow-on study used the ncAA *p*-isothiocyanate phenylalanine (pNCSF)[255] in screening MetA variants and identified additional mutants exhibiting enhanced stability due to putative crosslinking interactions. In conceptually related work with beta-lactamases, a library of variants containing ncAAs that facilitate extended disulfide bonds also led to a series of stabilized variants. In engineering both MetA and beta-lactamase, ncAAs enabled a level of stabilization that is difficult to achieve with a single round of mutagenesis and screening with canonical amino acids alone; such efforts would likely require multiple rounds of diversification. While protein stabilization with cAAs and conventional directed evolution remains a powerful approach[253, 256], experimentally driven ncAA-mediated stabilization is worthy of further investigation as a route to enzyme stabilization.

5.3 Accessing and Evolving New-to-Nature Biocatalysts with ncAAs

The introduction of ncAAs into enzyme active sites offers exciting opportunities to generate enzymes that catalyze new-to-nature reactions. This rapidly growing area has been the subject of multiple recent reviews; here we focus on advances in screening libraries of ncAA-containing enzymes and refer the reader to other sources for discussion of other aspects of this area[257-261]. Powerful engineering and design strategies have substantially extended the range of chemical reactions that can be catalyzed by enzymes using canonical amino acids[262-267]. However, numerous examples of ncAA-mediated biocatalysis exist[268-270] and have clearly established ncAAs as a means to access an impressive set of catalytic activities that are that are challenging or impossible to access with cAAs. Directed evolution has played an important role in enhancing activities of chemically expanded enzymes. One notable example is the establishment of ncAA-mediated photoenzymes. While some naturally occurring photoenzymes exist in photosynthetic pathways, triplet energy transfer catalysis is not known to occur in nature. Starting from a computationally designed enzyme scaffold, Trimble et al. introduced the photoactive ncAA 4-benzoylphenylalanine (BpA) at various locations in the enzyme active site (Figure 5)[271]. An initial variant exhibited the desired catalytic function but with poor activity and selectivity. Saturation mutagenesis and screening in 96-well plates yielded a mutant that supported over 300 turnovers and tolerated aerobic buffers, in contrast to small molecule photocatalysts. This establishment of an enzyme with bespoke photocatalytic properties is impressive, as triplet energy transfer catalysis is commonly used in organic synthesis. However, the authors also point out that the relatively small number of variants investigated (under 1000 total variants) highlights opportunities to further explore sequence space to identify additional ncAA-mediated photocatalysts.

In addition to photocatalysis, ncAAs have been incorporated into enzymes to generate artificial metalloenzymes[272-274]. Using myoglobin as a scaffold, Pott et al. site-specifically incorporated the metal-coordinating ncAA N^{ε}-methyl histidine (NMH) to generate a heme peroxidase. While ncAA substitution resulted in a modest 4-fold improvement in activity, several rounds of directed evolution resulted in an engineered metalloenzyme with 1140-fold increase in catalytic activity compared to the wild-type starting protein[275]. In addition to the subtle but important ncAA-mediated interactions that led to this increase in activity, this work demonstrates the ability to design and engineer designer metalloenzymes with activities comparable to wild-type metalloenzymes[275, 276]. Successes in identifying catalytically efficient, chemically expanded enzymes highlight the synergy between ncAA incorporation and subsequent evolution to yield biocatalysts with improved or new-to-nature functionalities.

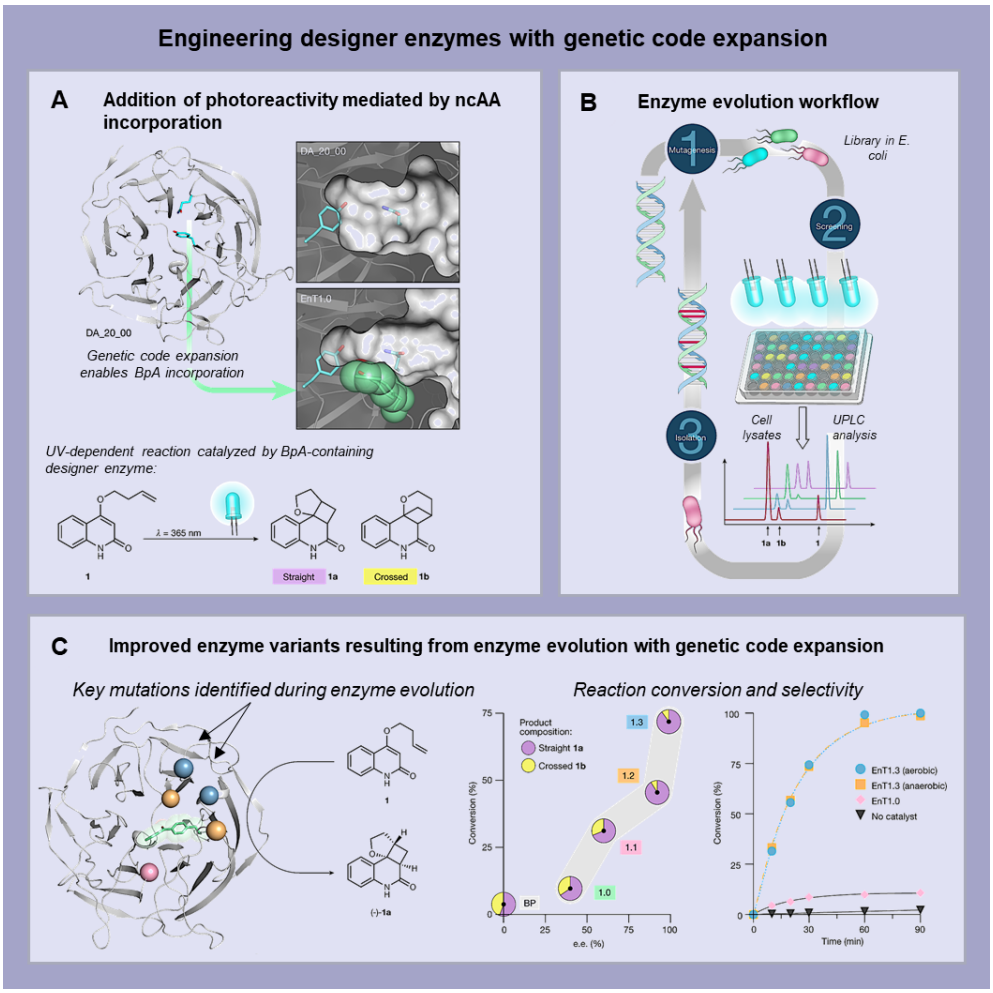

Figure 5. Using genetic code expansion to engineer enzymes with novel catalytic function. A) The photoreactive ncAA 4-benzoylphenylalanine (BpA) is site-specifically incorporated into the active site of a Diels–Alderase enzyme scaffold using genetic code expansion (Photoreactive enzyme dubbed EnT1.0). B) Enzyme evolution workflow to screen and isolate mutants with increased activity and selectivity. C) Highlighted mutations in the evolved clone, EnT1.3, along with data quantifying catalytic efficiency. Reactant conversion and reaction kinetics reported for the evolved clone (EnT1.3), the initial BpA-substituted construct (EnT1.0), and in the absence of the enzyme. (A-C) Adapted with permission from Ref [271]. *Copyright © 2022, Springer Nature America, Inc.*

5.4 High Throughput Discovery of Chemically Expanded Ligands

The maturation of display technologies has made the discovery of binding ligands against many targets routine (Here we define "binding ligands" as proteins or peptides that bind to a target of interest). However, not all molecular recognition challenges can be solved with genetically encoded binding ligands. Binders that modulate specific bioactivities (such as enzyme activity) or covalently bind to a target are difficult or impossible to identify without chemistries that go beyond what is found in the cAAs. NcAAs are one of several effective strategies for chemically diversifying binding ligands. Most other approaches use posttranslational diversification of peptides; several distinct phage display systems for cyclization or derivatization of peptide libraries are well-validated, and yeast display systems for peptide cyclization and disulfide-rich scaffold display are also available[277-288]. Beyond peptides and disulfide rich

scaffolds, Lewis et al. recently described a thiol-mediated strategy for preparing and screening libraries of fibronectin-small molecule "hybrids" in yeast display format[289]. The ongoing creativity in applying post-translational strategies to generate chemically diverse (poly)peptides offers complementary strengths and limitations in comparison to ncAA-mediated chemical expansion of ligands. Important advances in the use of ncAA-mediated chemical expansion of ligands have leveraged the unique strengths of each of the three major genetic code manipulation strategies.

5.4.1 In vitro Discovery of Chemically Expanded Peptides via High Throughput Screening

The most advanced chemically expanded discovery platform is mRNA display. This technique leverages the strengths of *in vitro* genetic code manipulation to implement numerous simultaneous changes to the genetic code during discovery. The versatility of mRNA display has been used to systematically change the chemical composition of genetically encoded peptide libraries. Modifications such as increasing the hydrophobicity of the set of encoded amino acids and introducing N-methyl groups within backbone amides improves the chances of identifying "druglike" hits during screening. While establishing an mRNA display platform requires considerable resources, the ability to screen trillions of candidates in parallel has led to an impressive range of discoveries in recent years. Here we highlight key advances in discovery and point to other recent reviews for further details[290-292].

While screening large libraries of cyclic peptides frequently yields potent binders, the use of hydrophilic cAAs can result in binding ligands that are unlikely to be amenable to therapeutic development. To address this challenge, Passioura et al. replaced the seven most hydrophilic canonical amino acids with hydrophobic ncAAs and expanded the total number of amino acids to 23 in an alternative genetic code within the Suga Lab's RAPID framework (tRNA acylation using flexizymes). Screens of cyclic peptides with this hydrophobic code against the cancer target interleukin-6 receptor yielded two nanomolar hits with affinities comparable to the hits emerging from a library containing only cAAs[293]. In an alternative *in vitro* genetic code reprogramming system, Guillen Schlippe et al. reassigned 12 canonical amino acids to ncAAs (tRNA acylation with aaRSs)[294]. A peptide library using the resulting code was screened against thrombin, a key protease in the anticoagulation cascade. Isolated clones exhibited apparent inhibitory constants in the single to double-digit nM range, and the importance of the ncAAs in facilitating inhibition was confirmed when replacement of the ncAAs by cAAs disrupted inhibition.

Specific chemical features can be installed that dramatically affect peptide discovery and function. NcAAs that facilitate peptide cyclization are now commonly used in many mRNA display libraries, as the benefits of peptide cyclization are well-known[294-297]. While important, these commonly used approaches will not be discussed further here. Following the reemergence of covalent warheads in small molecule discovery, recent mRNA display efforts have used crosslinkable ncAAs to facilitate the discovery of covalent peptides[298]; covalent target engagement can greatly enhance potency and duration of action[299]. Wu et al. demonstrated the feasibility of using photocrosslinkable ncAAs to discover cyclized peptides with the RaPID mRNA display platform. Using a UV-activated benzophenone ncAA, mRNA display campaigns were modified to include photocrosslinking and denaturation steps during screens against the second bromodomain within bromodomain-containing protein 3 (BRD3)[298]. Interestingly, the affinities of isolated hits were much lower than the affinities of hits identified in conventional mRNA display campaigns, suggesting that selection pressure for high affinity is decreased under the modified selection conditions. Work from Iskandar et al. described an mRNA display library in which the ncAA phenylselenocysteine was installed and then treated to reveal dehydroalanine, which can undergo proximity-induced crosslinking with thiols of cysteine side chain[299]. The inclusion of a denaturation step during selections promoted the discovery of spontaneously crosslinkable cyclic peptide hits against the targets calcium- and integrin-binding protein 1 (CIB1) and melanoma-associated antigen 4 (MAGE-A4). Multiple hits exhibited inhibitory activity against the targets with time-dependent IC_{50} values, providing strong evidence for covalent target engagement events that contribute to target inhibition. Overall, screening libraries encoding ncAA-mediated crosslinking is a promising, versatile route to covalent peptides.

In addition to an increasing set of academic investigations with mRNA display, a growing number of pharma and biotech companies have described advances in mRNA display in the scientific literature. Below we highlight several resource-intensive studies disclosing substantial technological advances and fundamental insights. A team from Merck has described their efforts to identify potent, orally bioavailable peptide inhibitors of proprotein convertase subtilisin-like/kexin type 9, PCSK9, an enzyme that regulates LDL cholesterol levels (Figure 6A)[300, 301]. An mRNA display campaign led to the identification of a lead with an inhibitory constant of approximately one micromolar. Extensive structure activity relationship (SAR) studies and structure-based design ultimately converted the monocyclic lead into a tricyclic peptide with a single-digit picomolar inhibitory constant. Importantly, one of the 5-fluorotryptophan residues present in the initial lead maintained its key role in determining potency throughout the optimization process; several additional features of the initial hit were also retained. This highlights the value of including chemically diverse elements within mRNA display libraries, while noting that intensive medicinal chemistry optimization was needed to identify a lead.

A team from Chugai Pharmaceuticals has established mRNA display libraries that enable rapid identification of orally bioavailable hits and leads directed against "tough intracellular targets"[302]. This ambitious effort first used limited diversities of cyclic peptides to quantitatively investigate "drug-likeness," including metabolic stability, cell permeability, and favorable pharmacokinetics. Insights from these studies enabled the design of an mRNA display library that encoded peptides exhibiting these features. However, the frequent use of *N*-alkylated amino acids in this design required further engineering of the underlying mRNA display platform to enable more efficient translation of two *N*-alkylated amino acids in a row. With an improved platform, the researchers were able to identify a series of cyclic peptides against the notoriously challenging target KRAS. Optimization of one of the hits resulted in a clinical candidate, LUNA18, that exhibited potent bioactivity in tissue culture and animal models of cancer; this candidate is currently in clinical development. This tour-de-force powerfully demonstrates the value of elucidating features that lead to druglikeness to advance the design of libraries of chemically expanded peptides.

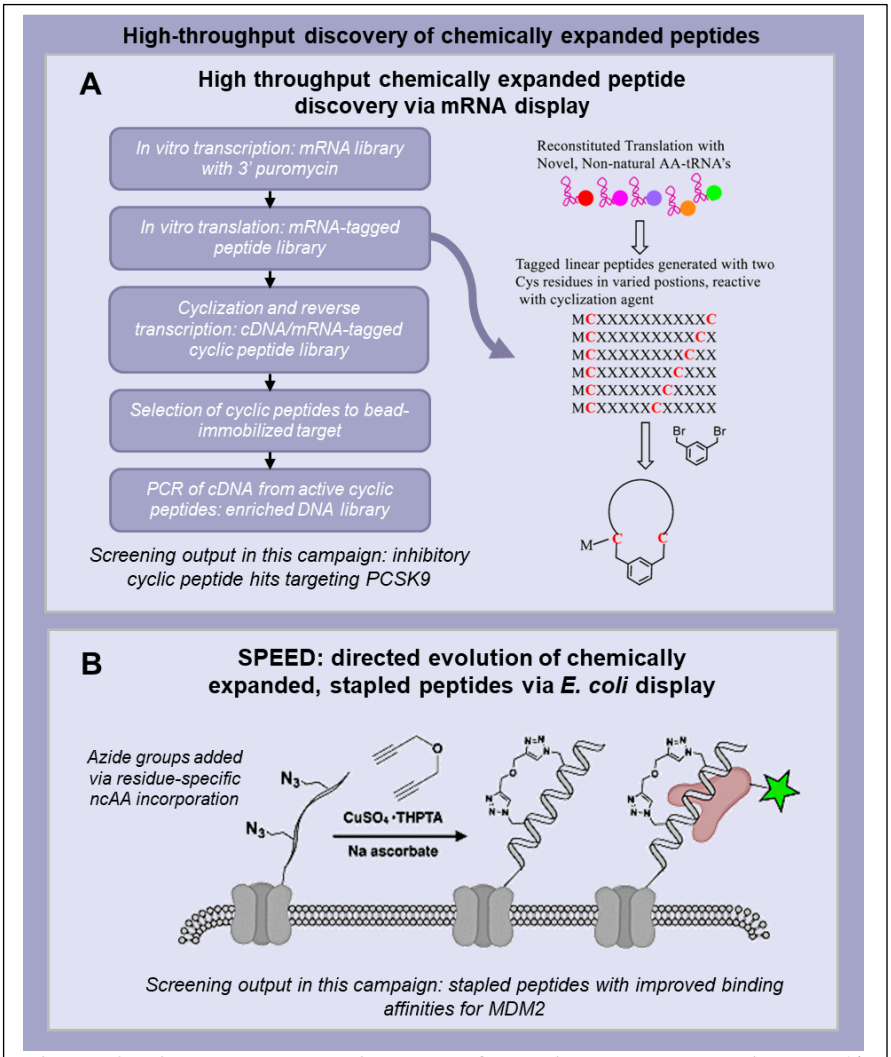

Figure 6. High throughput discovery of chemically expanded ligands. A) mRNA display and discovery workflow to generate and screen ncAA-containing cyclic peptide libraries against a bead-immobilized PCSK9 target. Reprinted with permission from Ref [300] *Copyright © 2020, American Chemical Society*. B) Stabilized Peptide Evolution by *E. coli* Display (SPEED) platform to site-specifically "staple" clickable ncAA-containing peptides using a bis-alkyne molecule in search of stapled peptides with improved binding affinities for MDM2. Adapted with permission from Ref [14] *Copyright © 2019. American Chemical Society.*

As mRNA display platforms continue to proliferate, screening campaigns with these platforms are expected to yield extensive deep sequencing datasets. This poses a challenge for analysis, as most sequence alignment programs are compatible with only canonical amino acids. A team at Genentech has described clustering algorithms that use chemical similarity to enable the identification of peptide families emerging during mRNA display screening campaigns[297]. This is an important development because it enables the rich datasets derived from deep sequencing to be used to identify a broader range of screening hits and directly inform SAR. Overall, mRNA display platforms continue to show tremendous promise for the discovery of chemically expanded peptides.

5.4.2. Phage and E. coli Display for Chemically Expanded Peptide Discovery

Built on the strong foundations of *E. coli*-based genetic code manipulation systems, there are now several phage display and *E. coli* display platforms for screening libraries of ncAA-containing peptides. While establishing these systems is less resource-intensive than establishing mRNA display systems, these display platforms usually only support the use of one to two ncAAs during library screening. Thus, to maximize their utility, it is critical to identify chemical or structural elements that will strongly enhance discovery against targets of interest. Many epigenetic readers and writers recognize acetylated lysine and other lysine derivatives; several of these modified lysines can be genetically encoded. Tharp et al. have used these biological insights and an advanced ncAA-containing phage display platform to conduct screens against lysine deacetylase sirtuin 2 (SIRT2). They discovered a series of acetyllysine-containing peptides that outperform previously reported small molecule inhibitors (IC_{50} values in the low double digit nanomolar range)[286]. The implementation of an amber-obligate phage replication scheme ensures that phage encode peptides that contain at least one ncAA. This eliminates replication biases that lead to preferential enrichment of phage encoding cAA-containing peptides. This unique system combined with well-chosen targets highlight fertile ground for future discovery efforts. Multiple groups have reported schemes for preparing cyclic peptides in phage display format using ncAAs[287, 303, 304]. The advantages and limitations of these systems are quite nuanced, especially when considered alongside related peptide phage display systems using only canonical amino acids. Hampton et al. recently reviewed chemically diversified phage libraries in a high level of detail; we refer readers there for deeper discussions[305].

E. coli peptide display systems are powerful alternatives to phage display systems, with flow cytometric approaches providing a key set of added capabilities[306]. Thurber and coworkers have established a peptide discovery platform called Stabilized Peptide Evolution by *E. coli* Display (SPEED) for stapled peptide construction and screening. This system uses the methionine analog azidohomoalanine (Aha) as the basis for installing staples (Figure 6B)[14]. Navaratna et al. used this platform to isolate high affinity ligands capable of disrupting the MDM2-p53 binding interaction[14]. Interestingly, the most potent evolved ligand included a disulfide loop in addition to the primary staple and exhibited enhanced resistance to proteolysis compared to parent sequences. A follow-up study by Case et. al. expanded upon this work, systematically investigating the properties of an extensive set of variants. Importantly, this work demonstrated that quantitative measurements of stapled peptides made on the *E. coli* cell surface are comparable to conventional biochemical measurements made in solution[307]. Very recently, this platform has been combined with machine learning techniques to map expanded stapled peptide sequence spaces. This demonstrates that the underlying bacterial display platform is robust enough to support powerful, data-driven approaches to stapled peptide design and engineering[308]. Further integration of chemically expanded peptide discovery platforms with data-driven screening and evolution has the potential to be highly impactful.

In addition to the isolation of individual constructs targeting a specific antigen, the wide availability of deep sequencing approaches has enabled broader investigations into the sequence-structure-activity relationships of many types of proteins; this area is growing rapidly and has been reviewed extensively elsewhere[136, 309, 310]. There are many opportunities to leverage ncAAs to better understand the binding preferences of naturally occurring proteins. Recently, Li et. al. employed bacterial display to screen peptides against tyrosine kinases and SH2 domains while using deep sequencing to map the role of amino acid sequence surrounding the key phosphotyrosine residue in binding and recognition[311]. A portion of these efforts used genetic code expansion to present phosphotyrosine analogs within peptide libraries to better mimic authentic substrates. In addition to enabling unique biological insights, this is an important demonstration of the integration of genetic code expansion strategies into *E. coli* peptide display. In an intracellular high-throughput screening approach, Barber et al. were able to evaluate phosphorylation-dependent binding events in high throughput with the use of genetic code expansion and a split fluorescent reporter[312]. Genetically encoding the ncAA phosphoserine in a genomically recoded organism, screens of phosphopeptides uncovered numerous biologically relevant interactions mediated by serine phosphorylation. Overall, the technological strengths of multiple *E. coli* genetic code manipulation systems have led to powerful chemically expanded peptide screening

approaches with *E. coli* and phage display (and beyond). These promising approaches will continue to play important roles in high throughput discovery of chemically expanded peptides in the future.

5.4.3 Engineering Binding Proteins using Genetic Code Manipulation and High Throughput Screening

Antibodies and other protein-based binding ligands offer distinct molecular recognition capabilities in comparison to peptides, with antibodies playing particularly expansive roles in basic research, diagnostics, and therapeutics[313, 314]. While chemically expanded antibody discovery remains in early phases, multiple discovery platforms are now technically validated and poised to facilitate high throughput discovery in protein-small molecule "hybrid" spaces. Van Deventer et al. established an *E. coli* display platform for engineering ncAA-substituted, clickable forms of antibody fragments[243]. Starting from an existing anti-digoxin antibody fragment, global replacement of methionine with structural analogs was found to be detrimental to antigen binding. Two rounds of random mutagenesis and screening with analog-substituted libraries led to clickable antibodies with enhanced display and binding properties. Notably, evolved, azide-containing variants exhibited improved affinity in comparison to Met-substituted forms of the same clones; this was an unexpected enhancement likely mediated by the unique properties of the azide side chain. Several azide-substituted variants tolerated strain-promoted azide-alkyne cycloadditions without loss of binding function. This work raises intriguing chemically expanded antibody engineering possibilities with clones containing multiple instances of a ncAA.

Liu et. al. reported a series of phage display studies using an expanded genetic code. A phage antibody library encoding ncAA substitution sites in antibody complementarity determining regions was constructed and paired with different OTSs for high-throughput discovery. Recognizing that some anti-HIV antibodies exhibit potent binding to gp120 with sulfated tyrosine residues, the researchers demonstrated that screens against gp120 resulted in the preferential isolation of sulfated antibodies. These screens yielded clones exhibiting affinities approaching those of known, naturally sulfated antibodies[315, 316]. This library was also used in combination with an OTS that supported the incorporation of the glycan-targeting ncAA *p*-borono-L-phenylalanine; screens led to preferential enrichment of ncAA-containing clones against a model sugar target. Overall, these studies demonstrate that ncAAs can substantially influence enrichment outcomes when presented within antibody libraries.

We have integrated genetic code expansion into yeast display for chemically expanded antibody discovery (and alternative binding ligand scaffolds)[69, 107, 149, 244, 284, 317-322]. To date, our published discovery efforts have described systematic evaluations of binding ligands. Alcala-Torano et al. established workflows for identifying covalent single-domain antibody variants on the yeast surface for both photocrosslinkable and spontaneously crosslinkable ncAAs (single-domain antibodies (sdAbs) are also referred to as nanobodies or VHHs)[244]. This work used an existing set of sdAbs targeting the light chain protease of botulinum neurotoxin serotype A (LC/A)[323]. Infection with botulinum neurotoxin enables LC/A to enter neurons and cause neuronal paralysis; an individual LC/A protein can retain its proteolytic activity inside neurons for months. As a proof of concept, Alcala-Torano et al. investigated over 40 chemically expanded sdAb variants by constructing and evaluating candidates on the yeast surface for covalent target engagement. Incorporation of the highly reactive, photocrosslinkable ncAA AzF into sdAbs led to numerous putatively crosslinkable variants (majority of variants tested), while incorporation of the weakly electrophilic ncAA *O*-bromoethyl-L-tyrosine (OBeY) into sdAbs led to identification of 3 putatively crosslinkable variants (out of 21 variants tested). Follow-up investigations of selected clones in solution confirmed crosslinking activities and indicated low or undetectable levels of off-target crosslinking in cell lysates. This represents a promising approach to covalent binding protein discovery that complements existing in-solution assays for covalent target engagement; it also has the potential to be applied with higher throughput screens in future work.

In a related approach, Huang et al. described the systematic evaluation of protein-small molecule hybrids on the yeast surface[318]. In a collaboration between the Van Deventer and Hackel groups, the labs investigated the roles of ncAA side chain structure and small molecule pharmacophore identity on the potency and specificity of hybrids targeting isoforms 2 and 9 of human carbonic anhydrase (hCA2, hCA9). Starting from previously reported hybrids based on reactive thiols and a potent acetazolamide

warhead[289], a series of clickable ncAAs and pharmacophores were used to construct dozens of chemical hybrid variants via CuAAC. Interestingly, numerous clones containing the acetazolamide warhead retained potent binding activity against hCA2 and hCA9, even though the chemical structures of the ncAA-acetazolimade linkers differed substantially from the parent linker structures. In addition, replacement of the potent warhead for a much weaker aryl sulfonamide warhead yielded several variants exhibiting activity. These observations suggest that chemical explorations of "hybrid space" have strong potential, especially when combined with 96 well-plate workflows established in the report. The identification of chemically expanded, crosslinkable antibodies and hybrids in these studies suggest that higher throughput screens on the yeast surface have the potential to yield a broader range of chemically expanded ligands. Overall, the availability of numerous technically validated, chemically expanded protein and peptide discovery platforms offers opportunities for discovery in unexplored hybrid binding ligand and drug discovery spaces.

6. Conclusions and Outlook

High throughput screening is strongly impacting both genetic code manipulation systems and their application to chemically expanded protein discovery. Expansion or alteration of the genetic code to accommodate ncAAs has benefitted greatly from efficient routes to the biocatalytic preparation of acylated tRNAs via engineered aminoacyl-tRNA synthetases (for cellular translation) and ribozymes (for *in vitro* translation). A growing range of components of the translation apparatus can now be screened or evolved in high throughput, enabling access to translational activities that cannot easily be identified through existing rational or computational approaches. By strategically engineering components involved in translation, researchers have demonstrated the ability to tune key interactions and mitigate competition with endogenous cellular processes. In addition to directly improving efficiencies of ncAA incorporation, this engineering paradigm has the potential to push large-scale biomanufacturing of ncAA-containing proteins and peptides beyond initial hard-won successes[324, 325] to the point of being routine. Beyond industrially relevant opportunities, many of the intricacies of protein translation apparatuses remain poorly understood. Unbiased screening approaches for ncAA incorporation present opportunities to uncover interaction networks and mechanisms that dictate protein translation efficiency and fidelity. This may even extend to components beyond the conventionally defined elements of the protein translation apparatus.

The availability of high-performing genetic code manipulation systems has fueled the establishment of high throughput discovery platforms for ncAA-containing proteins and peptides. In fact, there are compelling examples of screening platforms based on each of the three major types of genetic code manipulation. Exciting implementations of mRNA display are already being used as powerful tools in chemically expanded peptide drug discovery. In addition, the ongoing development of phage, yeast, and *E. coli* display platforms offers complementary strengths for chemically expanded binding ligand discovery efforts, including for antibodies and other protein-based binding ligands. Rapid advances in the discovery of enzymes that use ncAA-mediated, new-to-nature catalytic mechanisms are starting to push the boundaries of biocatalysis. While early screening campaigns have mostly used well plate formats (hundreds to thousands of variants), there are now numerous classes of enzymes amenable to directed evolution in display formats (millions of variants)[326-329]; advances in automation may also streamline discovery efforts in this area. For enzymes, binding ligands, and other polypeptides, a key challenge remains identifying the best "chemical starting points" that will lead to properties that cannot be accessed with canonical amino acids alone. In addition, the potential benefits of chemically expanded proteins must be weighed carefully against added costs or production complexities.

It is well-known that canonical amino acid sequence space is incomprehensibly large and must be navigated carefully; adding ncAAs to polypeptides results in further expanded sequence spaces. Computational approaches to navigating these spaces have the potential to be highly impactful. In the area of enhancing ncAA incorporation systems (with canonical amino acids only), existing frameworks for bioinformatics, biomolecular design, and machine learning can be applied to engineering protein translation apparatuses (and beyond)[131-139]. We expect this to be relatively straightforward because of the

availability of vast sequence and structural datasets on the proteins and nucleic acids that comprise the translation apparatus. In addition, fit-for-purpose datasets linking sequence and function can be generated using the powerful set of screening platforms described in Sections 3 and 4 of this review. Extending these tools to ncAA-containing proteins and peptides is likely to require substantial additional work. While some approaches, such as transfer learning[330], may inform models of ncAA-containing proteins, there is a dearth of experimental data on ncAA-containing proteins. There is no Protein Data Bank for structures of chemically expanded proteins, nor is there sequencing or evolutionary data on the DNA encoding such proteins. High throughput experimentation has an important role to play in establishing sufficient quantities of data in this area, and bioinformatics workflows will need to continue to be adapted to accommodate additional amino acids. Advances in laboratory automation, perhaps even including self-driving laboratories, will aid in the collection of this data[331-333]. Given the rapid rise of biomolecular machine learning, there are numerous opportunities to establish powerful design tools for ncAA-containing proteins.

From early on, researchers have used high throughput screening to "discover the unexpected" in genetic code manipulation and its applications. Fueled by advances in areas including directed evolution, DNA sequencing, biomolecular machine learning, and even laboratory automation, high throughput screening stands to play a critical role as genetic code manipulation continues to reach new heights.

Acknowledgments

Research on noncanonical amino acid incorporation and its applications in the Van Deventer Lab is supported by a grant from the National Institute of General Medical Sciences of the National Institutes of Health (R35GM133471), by a grant from the National Science Foundation (NSF2339201), and by the Bright Futures Assistant Professorship at Tufts University (all to J.A.V.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the National Science Foundation, or Tufts University.

Author Information

Briana R. Lino
Briana completed her undergraduate degree in Biochemistry from Wellesley College. Following graduation, she worked in the MIT Little Devices Lab developing affordable medical technologies. She then started a Biotechnology Engineering PhD at Tufts University working in the Van Deventer Lab. Her research focuses on engineering eukaryotic translation apparatuses to improve genetic code expansion in yeast.

Sean J. Williams
Sean Williams earned a bachelor's degree in Chemical Engineering at Northeastern University and conducted undergraduate research in the Auguste Lab for Bioresponsive Drug Delivery and Tissue Engineering. Following graduation, Sean joined the Chemical and Biological Engineering Department at Tufts University and is currently a Chemical Engineering PhD candidate in the Van Deventer Lab. His thesis work is focused on screening libraries of chemically expanded proteins to discover new classes of binding ligands.

Michelle E. Castor
Michelle Castor completed undergraduate studies at Northeastern University with dual degrees in Biology and Geology before completing a master's degree in Biotechnology at Brown University. Michelle works at Merck and Co. as a Senior Scientist. Michelle joined the Van Deventer Lab at Tufts as a part-time PhD student in Biotechnology Engineering, with her thesis work focusing on covalent antibody discovery.

James A. Van Deventer

Professor James Van Deventer completed undergraduate studies at Stanford University, PhD studies under the guidance of David Tirrell at the California Institute of Technology, and postdoctoral studies under the guidance of K. Dane Wittrup at the Massachusetts Institute of Technology. He then started his independent career in the Chemical and Biological Engineering Department at Tufts University. Dr. Van Deventer has won several awards for his research, including an R35 Outstanding Investigator Award for Early Career Researchers from the US National Institutes of Health and a CAREER Award from the US National Science Foundation. In addition, he has been recognized three times by the Tufts University Graduate Student Council for Outstanding Contributions to Graduate Education. The Van Deventer Laboratory focuses on 1) enhancing genetic code manipulation in yeast; and 2) chemically expanded antibody discovery.

References

(1) Rezhdo, A.; Islam, M.; Huang, M.; Van Deventer, J. A. Future prospects for noncanonical amino acids in biological therapeutics. *Curr Opin Biotechnol* **2019**, *60*, 168-178. DOI: 10.1016/j.copbio.2019.02.020

(2) Lugtenburg, T.; Gran-Scheuch, A.; Drienovska, I. Non-canonical amino acids as a tool for the thermal stabilization of enzymes. *Protein Eng Des Sel* **2023**, *36*. DOI: 10.1093/protein/gzad003

(3) Saleh, A. M.; Wilding, K. M.; Calve, S.; Bundy, B. C.; Kinzer-Ursem, T. L. Non-canonical amino acid labeling in proteomics and biotechnology. *J Biol Eng* **2019**, *13*, 43. DOI: 10.1186/s13036-019-0166-3

(4) Du, Y.; Li, L.; Zheng, Y.; Liu, J.; Gong, J.; Qiu, Z.; Li, Y.; Qiao, J.; Huo, Y. X. Incorporation of Non-Canonical Amino Acids into Antimicrobial Peptides: Advances, Challenges, and Perspectives. *Appl Environ Microbiol* **2022**, *88* (23), e0161722. DOI: 10.1128/aem.01617-22

(5) Castro, T. G.; Melle-Franco, M.; Sousa, C. E. A.; Cavaco-Paulo, A.; Marcos, J. C. Non-Canonical Amino Acids as Building Blocks for Peptidomimetics: Structure, Function, and Applications. *Biomolecules* **2023**, *13* (6). DOI: 10.3390/biom13060981

(6) Shandell, M. A.; Tan, Z.; Cornish, V. W. Genetic Code Expansion: A Brief History and Perspective. *Biochemistry* **2021**, *60* (46), 3455-3469. DOI: 10.1021/acs.biochem.1c00286

(7) Arranz-Gibert, P.; Vanderschuren, K.; Isaacs, F. J. Next-generation genetic code expansion. *Curr Opin Chem Biol* **2018**, *46*, 203-211. DOI: 10.1016/j.cbpa.2018.07.020

(8) Chung, C. Z.; Amikura, K.; Soll, D. Using Genetic Code Expansion for Protein Biochemical Studies. *Front Bioeng Biotechnol* **2020**, *8*, 598577. DOI: 10.3389/fbioe.2020.598577

(9) Liu, Y.; Kim, D. S.; Jewett, M. C. Repurposing ribosomes for synthetic biology. *Curr Opin Chem Biol* **2017**, *40*, 87-94. DOI: 10.1016/j.cbpa.2017.07.012

(10) Almhjell, P. J.; Boville, C. E.; Arnold, F. H. Engineering enzymes for noncanonical amino acid synthesis. *Chem Soc Rev* **2018**, *47* (24), 8980-8997. DOI: 10.1039/c8cs00665b

(11) Amiram, M.; Haimovich, A. D.; Fan, C.; Wang, Y. S.; Aerni, H. R.; Ntai, I.; Moonan, D. W.; Ma, N. J.; Rovner, A. J.; Hong, S. H.; et al. Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat Biotechnol* **2015**, *33* (12), 1272-1279. DOI: 10.1038/nbt.3372

(12) Jewel, D.; Kelemen, R. E.; Huang, R. L.; Zhu, Z.; Sundaresh, B.; Cao, X.; Malley, K.; Huang, Z.; Pasha, M.; Anthony, J.; et al. Virus-assisted directed evolution of enhanced suppressor tRNAs in mammalian cells. *Nat Methods* **2023**, *20* (1), 95-103. DOI: 10.1038/s41592-022-01706-w

(13) DeBenedictis, E. A.; Chory, E. J.; Gretton, D. W.; Wang, B.; Golas, S.; Esvelt, K. M. Systematic molecular evolution enables robust biomolecule discovery. *Nat Methods* **2022**, *19* (1), 55-64. DOI: 10.1038/s41592-021-01348-4

(14) Navaratna, T.; Atangcho, L.; Mahajan, M.; Subramanian, V.; Case, M.; Min, A.; Tresnak, D.; Thurber, G. M. Directed Evolution Using Stabilized Bacterial Peptide Display. *J Am Chem Soc* **2020**, *142* (4), 1882-1894. DOI: 10.1021/jacs.9b10716

(15) Johnson, J. A.; Lu, Y. Y.; Van Deventer, J. A.; Tirrell, D. A. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Curr Opin Chem Biol* **2010**, *14* (6), 774-780. DOI: 10.1016/j.cbpa.2010.09.013
(16) Singh-Blom, A.; Hughes, R. A.; Ellington, A. D. Residue-specific incorporation of unnatural amino acids into proteins in vitro and in vivo. *Methods Mol Biol* **2013**, *978*, 93-114. DOI: 10.1007/978-1-62703-293-3_7
(17) Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **2017**, *550* (7674), 53-60. DOI: 10.1038/nature24031
(18) Chin, J. W. Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Biochem* **2014**, *83*, 379-408. DOI: 10.1146/annurev-biochem-060713-035737
(19) Budisa, N. Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire. *Angew Chem Int Edit* **2004**, *43* (47), 6426-6463. DOI: 10.1002/anie.200300646
(20) Wu, Y.; Wang, Z.; Qiao, X.; Li, J.; Shu, X.; Qi, H. Emerging Methods for Efficient and Extensive Incorporation of Non-canonical Amino Acids Using Cell-Free Systems. *Front Bioeng Biotechnol* **2020**, *8*, 863. DOI: 10.3389/fbioe.2020.00863
(21) Leisle, L.; Valiyaveetil, F.; Mehl, R. A.; Ahern, C. A. Incorporation of Non-Canonical Amino Acids. *Adv Exp Med Biol* **2015**, *869*, 119-151. DOI: 10.1007/978-1-4939-2845-3_7
(22) de la Torre, D.; Chin, J. W. Reprogramming the genetic code. *Nat Rev Genet* **2021**, *22* (3), 169-184. DOI: 10.1038/s41576-020-00307-7
(23) Wang, L. Engineering the Genetic Code in Cells and Animals: Biological Considerations and Impacts. *Acc Chem Res* **2017**, *50* (11), 2767-2775. DOI: 10.1021/acs.accounts.7b00376
(24) Cornish, V. W.; Mendel, D.; Schultz, P. G. Probing Protein-Structure and Function with an Expanded Genetic-Code. *Angewandte Chemie-International Edition in English* **1995**, *34* (6), 621-633. DOI: DOI 10.1002/anie.199506211
(25) Tcherkezian, J.; Brittis, P. A.; Thomas, F.; Roux, P. P.; Flanagan, J. G. Transmembrane Receptor DCC Associates with Protein Synthesis Machinery and Regulates Translation. *Cell* **2010**, *141* (4), 632-644. DOI: 10.1016/j.cell.2010.04.008
(26) Zhang, M. M.; Tsou, L. K.; Charron, G.; Raghavan, A. S.; Hang, H. C. Tandem fluorescence imaging of dynamic S-acylation and protein turnover. *Proc Natl Acad Sci U S A* **2010**, *107* (19), 8627-8632. DOI: 10.1073/pnas.0912306107
(27) Gran-Scheuch, A.; Bonandi, E.; Drienovská, I. Expanding the Genetic Code: Incorporation of Functional Secondary Amines via Stop Codon Suppression. *Chemcatchem* **2023**. DOI: 10.1002/cctc.202301004
(28) Dieterich, D. C.; Hodas, J. J.; Gouzer, G.; Shadrin, I. Y.; Ngo, J. T.; Triller, A.; Tirrell, D. A.; Schuman, E. M. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat Neurosci* **2010**, *13* (7), 897-905. DOI: 10.1038/nn.2580
(29) Ngo, J. T.; Tirrell, D. A. Noncanonical amino acids in the interrogation of cellular protein synthesis. *Acc Chem Res* **2011**, *44* (9), 677-685. DOI: 10.1021/ar200144y
(30) Tom Dieck, S.; Muller, A.; Nehring, A.; Hinz, F. I.; Bartnik, I.; Schuman, E. M.; Dieterich, D. C. Metabolic labeling with noncanonical amino acids and visualization by chemoselective fluorescent tagging. *Curr Protoc Cell Biol* **2012**, *Chapter 7*, 7 11 11-17 11 29. DOI: 10.1002/0471143030.cb0711s56
(31) Dieterich, D. C.; Lee, J. J.; Link, A. J.; Graumann, J.; Tirrell, D. A.; Schuman, E. M. Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nat Protoc* **2007**, *2* (3), 532-540. DOI: 10.1038/nprot.2007.52
(32) Elliott, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsässer, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; et al. Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nature Biotechnology* **2014**, *32* (5), 465-U186. DOI: 10.1038/nbt.2860
(33) Elliott, T. S.; Bianco, A.; Townsley, F. M.; Fried, S. D.; Chin, J. W. Tagging and Enriching Proteins Enables Cell-Specific Proteomics. *Cell Chem Biol* **2016**, *23* (7), 805-815. DOI: 10.1016/j.chembiol.2016.05.018

(34) Hoffmann, J. E.; Dziuba, D.; Stein, F.; Schultz, C. A Bifunctional Noncanonical Amino Acid: Synthesis, Expression, and Residue-Specific Proteome-wide Incorporation. *Biochemistry* **2018**, *57* (31), 4747-4752. DOI: 10.1021/acs.biochem.8b00397
(35) Majekodunmi, T.; Britton, D.; Montclare, J. K. Engineered Proteins and Materials Utilizing Residue-Specific Noncanonical Amino Acid Incorporation. *Chem Rev* **2024**, *124* (15), 9113-9135. DOI: 10.1021/acs.chemrev.3c00855
(36) Wang, Y. S.; Fang, X. Q.; Wallace, A. L.; Wu, B.; Liu, W. S. R. A Rationally Designed Pyrrolysyl-tRNA Synthetase Mutant with a Broad Substrate Spectrum. *J Am Chem Soc* **2012**, *134* (6), 2950-2953. DOI: 10.1021/ja211972x
(37) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **1989**, *244* (4901), 182-188. DOI: 10.1126/science.2649980
(38) Ambrogelly, A.; Palioura, S.; Söll, D. Natural expansion of the genetic code. *Nature Chemical Biology* **2007**, *3* (1), 29-35. DOI: 10.1038/nchembio847
(39) Liu, C. C.; Schultz, P. G. Adding new chemistries to the genetic code. *Annu Rev Biochem* **2010**, *79*, 413-444. DOI: 10.1146/annurev.biochem.052308.105824
(40) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Wacker, D. A.; Lyttle, M. H.; Dix, T. A.; Chamberlin, A. R. Site-Specific Incorporation of Nonnatural Residues during Invitro Protein-Biosynthesis with Semisynthetic Aminoacyl-Transfer Rnas. *Biochemistry* **1991**, *30* (22), 5411-5421. DOI: DOI 10.1021/bi00236a013
(41) Hohsaka, T.; Ashizuka, Y.; Murakami, H.; Sisido, M. Incorporation of nonnatural amino acids into streptavidin through in vitro frame-shift suppression. *J Am Chem Soc* **1996**, *118* (40), 9778-9779. DOI: DOI 10.1021/ja9614225
(42) Hankore, E. D.; Zhang, L. Y.; Chen, Y.; Liu, K.; Niu, W.; Guo, J. T. Genetic Incorporation of Noncanonical Amino Acids Using Two Mutually Orthogonal Quadruplet Codons. *Acs Synthetic Biology* **2019**, *8* (5), 1168-1174. DOI: 10.1021/acssynbio.9b00051
(43) Hohsaka, T.; Ashizuka, Y.; Murakami, H.; Sisido, M. Five-base codons for incorporation of nonnatural amino acids into proteins. *Nucleic Acids Research* **2001**, *29* (17), 3646-3651. DOI: DOI 10.1093/nar/29.17.3646
(44) Chatterjee, A.; Lajoie, M. J.; Xiao, H.; Church, G. M.; Schultz, P. G. A Bacterial Strain with a Unique Quadruplet Codon Specifying Non-native Amino Acids. *Chembiochem* **2014**, *15* (12), 1782-1786. DOI: 10.1002/cbic.201402104
(45) Kimoto, M.; Hirao, I. Genetic Code Engineering by Natural and Unnatural Base Pair Systems for the Site-Specific Incorporation of Non-Standard Amino Acids Into Proteins. *Front Mol Biosci* **2022**, *9*, 851646. DOI: 10.3389/fmolb.2022.851646
(46) Wang, K.; Fredens, J.; Brunner, S. F.; Kim, S. H.; Chia, T.; Chin, J. W. Defining synonymous codon compression schemes by genome recoding. *Nature* **2016**, *539* (7627), 59-64. DOI: 10.1038/nature20124
(47) Passioura, T.; Suga, H. Reprogramming the genetic code in vitro. *Trends Biochem Sci* **2014**, *39* (9), 400-408. DOI: 10.1016/j.tibs.2014.07.005
(48) Katoh, T.; Suga, H. In Vitro Genetic Code Reprogramming for the Expansion of Usable Noncanonical Amino Acids. *Annu Rev Biochem* **2022**, *91*, 221-243. DOI: 10.1146/annurev-biochem-040320-103817
(49) 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* (8), 751-755. DOI: 10.1038/90802
(50) Murakami, H.; Kourouklis, D.; Suga, H. Using a solid-phase ribozyme aminoacylation system to reprogram the genetic code. *Chemistry & Biology* **2003**, *10* (11), 1077-1084. DOI: 10.1016/j.chembiol.2003.10.010
(51) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat Methods* **2006**, *3* (5), 357-359. DOI: 10.1038/nmeth877

(52) Ohuchi, M.; Murakami, H.; Suga, H. The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. *Curr Opin Chem Biol* **2007**, *11* (5), 537-542. DOI: 10.1016/j.cbpa.2007.08.011
(53) Kawakami, T.; Murakami, H.; Suga, H. Messenger RNA-programmed incorporation of multiple N-methyl-amino acids into linear and cyclic peptides. *Chem Biol* **2008**, *15* (1), 32-42. DOI: 10.1016/j.chembiol.2007.12.008
(54) Ohta, A.; Yamagishi, Y.; Suga, H. Synthesis of biopolymers using genetic code reprogramming. *Current Opinion in Chemical Biology* **2008**, *12* (2), 159-167. DOI: 10.1016/j.cbpa.2007.12.009
(55) Katoh, T.; Passioura, T.; Suga, H. Advances in in vitro genetic code reprogramming in 2014-2017. *Synth Biol (Oxf)* **2018**, *3* (1), ysy008. DOI: 10.1093/synbio/ysy008
(56) Richardson, S. L.; Dods, K. K.; Abrigo, N. A.; Iqbal, E. S.; Hartman, M. C. In vitro genetic code reprogramming and expansion to study protein function and discover macrocyclic peptide ligands. *Curr Opin Chem Biol* **2018**, *46*, 172-179. DOI: 10.1016/j.cbpa.2018.07.013
(57) Reynolds, N. M.; Vargas-Rodriguez, O.; Soll, D.; Crnkovic, A. The central role of tRNA in genetic code expansion. *Biochim Biophys Acta Gen Subj* **2017**, *1861* (11 Pt B), 3001-3008. DOI: 10.1016/j.bbagen.2017.03.012
(58) Melnikov, S. V.; Soll, D. Aminoacyl-tRNA Synthetases and tRNAs for an Expanded Genetic Code: What Makes them Orthogonal? *Int J Mol Sci* **2019**, *20* (8). DOI: 10.3390/ijms20081929
(59) Wan, W.; Tharp, J. M.; Liu, W. R. Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim Biophys Acta* **2014**, *1844* (6), 1059-1070. DOI: 10.1016/j.bbapap.2014.03.002
(60) Meineke, B.; Heimgartner, J.; Lafranchi, L.; Elsasser, S. J. Methanomethylophilus alvus Mx1201 Provides Basis for Mutual Orthogonal Pyrrolysyl tRNA/Aminoacyl-tRNA Synthetase Pairs in Mammalian Cells. *ACS Chem Biol* **2018**, *13* (11), 3087-3096. DOI: 10.1021/acschembio.8b00571
(61) Beattie, A. T.; Dunkelmann, D. L.; Chin, J. W. Quintuply orthogonal pyrrolysyl-tRNA synthetase/tRNA(Pyl) pairs. *Nat Chem* **2023**, *15* (7), 948-959. DOI: 10.1038/s41557-023-01232-y
(62) Dunkelmann, D. L.; Willis, J. C. W.; Beattie, A. T.; Chin, J. W. Engineered triply orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs enable the genetic encoding of three distinct non-canonical amino acids. *Nat Chem* **2020**, *12* (6), 535-544. DOI: 10.1038/s41557-020-0472-x
(63) Cervettini, D.; Tang, S.; Fried, S. D.; Willis, J. C. W.; Funke, L. F. H.; Colwell, L. J.; Chin, J. W. Rapid discovery and evolution of orthogonal aminoacyl-tRNA synthetase-tRNA pairs. *Nat Biotechnol* **2020**, *38* (8), 989-999. DOI: 10.1038/s41587-020-0479-2
(64) Avila-Crump, S.; Hemshorn, M. L.; Jones, C. M.; Mbengi, L.; Meyer, K.; Griffis, J. A.; Jana, S.; Petrina, G. E.; Pagar, V. V.; Karplus, P. A.; et al. Generating Efficient Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetases for Structurally Diverse Non-Canonical Amino Acids. *ACS Chem Biol* **2022**, *17* (12), 3458-3469. DOI: 10.1021/acschembio.2c00639
(65) Sun, R.; Zheng, H.; Fang, Z.; Yao, W. Rational design of aminoacyl-tRNA synthetase specific for p-acetyl-L-phenylalanine. *Biochem Biophys Res Commun* **2010**, *391* (1), 709-715. DOI: 10.1016/j.bbrc.2009.11.125
(66) Wang, Y. S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R. A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. *J Am Chem Soc* **2012**, *134* (6), 2950-2953. DOI: 10.1021/ja211972x
(67) Vincent, C.; Borel, F.; Willison, J. C.; Leberman, R.; Hartlein, M. Seryl-tRNA synthetase from Escherichia coli: functional evidence for cross-dimer tRNA binding during aminoacylation. *Nucleic Acids Res* **1995**, *23* (7), 1113-1118. DOI: 10.1093/nar/23.7.1113
(68) Osgood, A. O.; Zheng, Y.; Roy, S. J. S.; Biris, N.; Hussain, M.; Loynd, C.; Jewel, D.; Italia, J. S.; Chatterjee, A. An Efficient Opal-Suppressor Tryptophanyl Pair Creates New Routes for Simultaneously Incorporating up to Three Distinct Noncanonical Amino Acids into Proteins in Mammalian Cells. *Angew Chem Int Ed Engl* **2023**, *62* (19), e202219269. DOI: 10.1002/anie.202219269

(69) Lahiri, P.; Martin, M. S.; Lino, B. R.; Scheck, R. A.; Van Deventer, J. A. Dual Noncanonical Amino Acid Incorporation Enabling Chemoselective Protein Modification at Two Distinct Sites in Yeast. *Biochemistry* **2023**, *62* (14), 2098-2114. DOI: 10.1021/acs.biochem.2c00711
(70) Albers, S.; Beckert, B.; Matthies, M. C.; Mandava, C. S.; Schuster, R.; Seuring, C.; Riedner, M.; Sanyal, S.; Torda, A. E.; Wilson, D. N.; et al. Repurposing tRNAs for nonsense suppression. *Nat Commun* **2021**, *12* (1), 3850. DOI: 10.1038/s41467-021-24076-x
(71) Ellefson, J. W.; Meyer, A. J.; Hughes, R. A.; Cannon, J. R.; Brodbelt, J. S.; Ellington, A. D. Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat Biotechnol* **2014**, *32* (1), 97-101. DOI: 10.1038/nbt.2714
(72) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the genetic code of Escherichia coli. *Science* **2001**, *292* (5516), 498-500. DOI: 10.1126/science.1060077
(73) Wang, L.; Schultz, P. G. A general approach for the generation of orthogonal tRNAs. *Chem Biol* **2001**, *8* (9), 883-890. DOI: 10.1016/s1074-5521(01)00063-1
(74) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. An expanded eukaryotic genetic code. *Science* **2003**, *301* (5635), 964-967. DOI: 10.1126/science.1084772
(75) Cropp, T. A.; Anderson, J. C.; Chin, J. W. Reprogramming the amino-acid substrate specificity of orthogonal aminoacyl-tRNA synthetases to expand the genetic code of eukaryotic cells. *Nat Protoc* **2007**, *2* (10), 2590-2600. DOI: 10.1038/nprot.2007.378
(76) Wiltschi, B. Incorporation of non-canonical amino acids into proteins in yeast. *Fungal Genet Biol* **2016**, *89*, 137-156. DOI: 10.1016/j.fgb.2016.02.002
(77) Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat Biotechnol* **2002**, *20* (10), 1044-1048. DOI: 10.1038/nbt742
(78) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. *J Am Chem Soc* **2002**, *124* (31), 9026-9027. DOI: 10.1021/ja027007w
(79) Park, H. S.; Hohn, M. J.; Umehara, T.; Guo, L. T.; Osborne, E. M.; Benner, J.; Noren, C. J.; Rinehart, J.; Soll, D. Expanding the genetic code of Escherichia coli with phosphoserine. *Science* **2011**, *333* (6046), 1151-1154. DOI: 10.1126/science.1207203
(80) Cooley, R. B.; Feldman, J. L.; Driggers, C. M.; Bundy, T. A.; Stokes, A. L.; Karplus, P. A.; Mehl, R. A. Structural basis of improved second-generation 3-nitro-tyrosine tRNA synthetases. *Biochemistry* **2014**, *53* (12), 1916-1924. DOI: 10.1021/bi5001239
(81) Jiang, H. K.; Ambrose, N. L.; Chung, C. Z.; Wang, Y. S.; Soll, D.; Tharp, J. M. Split aminoacyl-tRNA synthetases for proximity-induced stop codon suppression. *Proc Natl Acad Sci U S A* **2023**, *120* (8), e2219758120. DOI: 10.1073/pnas.2219758120
(82) Sungwienwong, I.; Hostetler, Z. M.; Blizzard, R. J.; Porter, J. J.; Driggers, C. M.; Mbengi, L. Z.; Villegas, J. A.; Speight, L. C.; Saven, J. G.; Perona, J. J.; et al. Improving target amino acid selectivity in a permissive aminoacyl tRNA synthetase through counter-selection. *Org Biomol Chem* **2017**, *15* (17), 3603-3610. DOI: 10.1039/c7ob00582b
(83) Thyer, R.; d'Oelsnitz, S.; Blevins, M. S.; Klein, D. R.; Brodbelt, J. S.; Ellington, A. D. Directed Evolution of an Improved Aminoacyl-tRNA Synthetase for Incorporation of L-3,4-Dihydroxyphenylalanine (L-DOPA). *Angew Chem Int Ed Engl* **2021**, *60* (27), 14811-14816. DOI: 10.1002/anie.202100579
(84) Grasso, K. T.; Singha Roy, S. J.; Osgood, A. O.; Yeo, M. J. R.; Soni, C.; Hillenbrand, C. M.; Ficaretta, E. D.; Chatterjee, A. A Facile Platform to Engineer Escherichia coli Tyrosyl-tRNA Synthetase Adds New Chemistries to the Eukaryotic Genetic Code, Including a Phosphotyrosine Mimic. *ACS Cent Sci* **2022**, *8* (4), 483-492. DOI: 10.1021/acscentsci.1c01465
(85) Owens, A. E.; Grasso, K. T.; Ziegler, C. A.; Fasan, R. Two-Tier Screening Platform for Directed Evolution of Aminoacyl-tRNA Synthetases with Enhanced Stop Codon Suppression Efficiency. *Chembiochem* **2017**, *18* (12), 1109-1116. DOI: 10.1002/cbic.201700039

(86) Dongheon Lee, M.-K. K., and Jong-il Choi. Development of Orthogonal Aminoacyl tRNA Synthetase Mutant with Enhanced Incorporation Ability with Para-azido-L-phenylalanine. *Biotechnology and Bioprocess Engineering* **2022**. DOI: https://doi.org/10.1007/s12257-022-0252-0
(87) Kuhn, S. M.; Rubini, M.; Fuhrmann, M.; Theobald, I.; Skerra, A. Engineering of an orthogonal aminoacyl-tRNA synthetase for efficient incorporation of the non-natural amino acid O-methyl-L-tyrosine using fluorescence-based bacterial cell sorting. *J Mol Biol* **2010**, *404* (1), 70-87. DOI: 10.1016/j.jmb.2010.09.001
(88) Hohl, A.; Karan, R.; Akal, A.; Renn, D.; Liu, X.; Ghorpade, S.; Groll, M.; Rueping, M.; Eppinger, J. Engineering a Polyspecific Pyrrolysyl-tRNA Synthetase by a High Throughput FACS Screen. *Sci Rep* **2019**, *9* (1), 11971. DOI: 10.1038/s41598-019-48357-0
(89) Schwark, D. G.; Schmitt, M. A.; Fisk, J. D. Directed Evolution of the Methanosarcina barkeri Pyrrolysyl tRNA/aminoacyl tRNA Synthetase Pair for Rapid Evaluation of Sense Codon Reassignment Potential. *Int J Mol Sci* **2021**, *22* (2). DOI: 10.3390/ijms22020895
(90) Link, A. J.; Tirrell, D. A. Cell surface labeling of Escherichia coli via copper(I)-catalyzed [3+2] cycloaddition. *J Am Chem Soc* **2003**, *125* (37), 11164-11165. DOI: 10.1021/ja036765z
(91) Link, A. J.; Vink, M. K. S.; Agard, N. J.; Prescher, J. A.; Bertozzi, C. R.; Tirrell, D. A. Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proc Natl Acad Sci U S A* **2006**, *103* (27), 10180-10185. DOI: 10.1073/pnas.0601167103
(92) Tanrikulu, I. C.; Schmitt, E.; Mechulam, Y.; Goddard, W. A., 3rd; Tirrell, D. A. 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* (36), 15285-15290. DOI: 10.1073/pnas.0905735106
(93) Schmitt, E.; Tanrikulu, I. C.; Yoo, T. H.; Panvert, M.; Tirrell, D. A.; Mechulam, Y. Switching from an induced-fit to a lock-and-key mechanism in an aminoacyl-tRNA synthetase with modified specificity. *J Mol Biol* **2009**, *394* (5), 843-851. DOI: 10.1016/j.jmb.2009.10.016
(94) Kwok, H. S.; Vargas-Rodriguez, O.; Melnikov, S. V.; Soll, D. Engineered Aminoacyl-tRNA Synthetases with Improved Selectivity toward Noncanonical Amino Acids. *ACS Chem Biol* **2019**, *14* (4), 603-612. DOI: 10.1021/acschembio.9b00088
(95) Lacey, V. K.; Louie, G. V.; Noel, J. P.; Wang, L. Expanding the library and substrate diversity of the pyrrolysyl-tRNA synthetase to incorporate unnatural amino acids containing conjugated rings. *Chembiochem* **2013**, *14* (16), 2100-2105. DOI: 10.1002/cbic.201300400
(96) Yoo, T. H.; Tirrell, D. A. High-throughput screening for methionyl-tRNA synthetases that enable residue-specific incorporation of noncanonical amino acids into recombinant proteins in bacterial cells. *Angew Chem Int Ed Engl* **2007**, *46* (28), 5340-5343. DOI: 10.1002/anie.200700779
(97) Truong, F.; Yoo, T. H.; Lampo, T. J.; Tirrell, D. A. Two-strain, cell-selective protein labeling in mixed bacterial cultures. *J Am Chem Soc* **2012**, *134* (20), 8551-8556. DOI: 10.1021/ja3004667
(98) Schmitt, E.; Meinnel, T.; Blanquet, S.; Mechulam, Y. Methionyl-tRNA synthetase needs an intact and mobile 332KMSKS336 motif in catalysis of methionyl adenylate formation. *J Mol Biol* **1994**, *242* (4), 566-576. DOI: 10.1006/jmbi.1994.1601
(99) Crepin, T.; Schmitt, E.; Blanquet, S.; Mechulam, Y. Structure and function of the C-terminal domain of methionyl-tRNA synthetase. *Biochemistry* **2002**, *41* (43), 13003-13011. DOI: 10.1021/bi026343m
(100) O'Donoghue, P.; Ling, J.; Wang, Y. S.; Soll, D. Upgrading protein synthesis for synthetic biology. *Nat Chem Biol* **2013**, *9* (10), 594-598. DOI: 10.1038/nchembio.1339
(101) Guo, L. T.; Wang, Y. S.; Nakamura, A.; Eiler, D.; Kavran, J. M.; Wong, M.; Kiessling, L. L.; Steitz, T. A.; O'Donoghue, P.; Soll, D. Polyspecific pyrrolysyl-tRNA synthetases from directed evolution. *Proc Natl Acad Sci U S A* **2014**, *111* (47), 16724-16729. DOI: 10.1073/pnas.1419737111
(102) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J Am Chem Soc* **2000**, *122* (7), 1282-1288. DOI: DOI 10.1021/ja992749j
(103) Kiick, K. L.; Weberskirch, R.; Tirrell, D. A. Identification of an expanded set of translationally active methionine analogues in Escherichia coli. *FEBS Lett* **2001**, *502* (1-2), 25-30. DOI: 10.1016/s0014-5793(01)02657-6

(104) Gan, R.; Perez, J. G.; Carlson, E. D.; Ntai, I.; Isaacs, F. J.; Kelleher, N. L.; Jewett, M. C. Translation system engineering in Escherichia coli enhances non-canonical amino acid incorporation into proteins. *Biotechnol Bioeng* **2017**, *114* (5), 1074-1086. DOI: 10.1002/bit.26239
(105) Italia, J. S.; Latour, C.; Wrobel, C. J. J.; Chatterjee, A. Resurrecting the Bacterial Tyrosyl-tRNA Synthetase/tRNA Pair for Expanding the Genetic Code of Both E. coli and Eukaryotes. *Cell Chem Biol* **2018**, *25* (10), 1304-1312 e1305. DOI: 10.1016/j.chembiol.2018.07.002
(106) Ficaretta, E. D.; Wrobel, C. J. J.; Roy, S. J. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A. A Robust Platform for Unnatural Amino Acid Mutagenesis in E. coli Using the Bacterial Tryptophanyl-tRNA synthetase/tRNA pair. *J Mol Biol* **2022**, *434* (8), 167304. DOI: 10.1016/j.jmb.2021.167304
(107) Stieglitz, J. T.; Van Deventer, J. A. High-Throughput Aminoacyl-tRNA Synthetase Engineering for Genetic Code Expansion in Yeast. *ACS Synth Biol* **2022**, *11* (7), 2284-2299. DOI: 10.1021/acssynbio.1c00626
(108) Carlson, J. C.; Badran, A. H.; Guggiana-Nilo, D. A.; Liu, D. R. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat Chem Biol* **2014**, *10* (3), 216-222. DOI: 10.1038/nchembio.1453
(109) Bryson, D. I.; Fan, C.; Guo, L. T.; Miller, C.; Soll, D.; Liu, D. R. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Biol* **2017**, *13* (12), 1253-1260. DOI: 10.1038/nchembio.2474
(110) Suzuki, T.; Miller, C.; Guo, L. T.; Ho, J. M. L.; Bryson, D. I.; Wang, Y. S.; Liu, D. R.; Soll, D. Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat Chem Biol* **2017**, *13* (12), 1261-1266. DOI: 10.1038/nchembio.2497
(111) Molina, R. S.; Rix, G.; Mengiste, A. A.; Alvarez, B.; Seo, D.; Chen, H.; Hurtado, J.; Zhang, Q.; Donato Garcia-Garcia, J.; Heins, Z. J.; et al. In vivo hypermutation and continuous evolution. *Nat Rev Methods Primers* **2022**, *2*. DOI: 10.1038/s43586-022-00130-w
(112) Fricke, R.; Swenson, C. V.; Roe, L. T.; Hamlish, N. X.; Shah, B.; Zhang, Z.; Ficaretta, E.; Ad, O.; Smaga, S.; Gee, C. L.; et al. Expanding the substrate scope of pyrrolysyl-transfer RNA synthetase enzymes to include non-alpha-amino acids in vitro and in vivo. *Nat Chem* **2023**, *15* (7), 960-971. DOI: 10.1038/s41557-023-01224-y
(113) Dunkelmann, D. L.; Piedrafita, C.; Dickson, A.; Liu, K. C.; Elliott, T. S.; Fiedler, M.; Bellini, D.; Zhou, A.; Cervettini, D.; Chin, J. W. Adding alpha,alpha-disubstituted and beta-linked monomers to the genetic code of an organism. *Nature* **2024**, *625* (7995), 603-610. DOI: 10.1038/s41586-023-06897-6
(114) Soni, C.; Prywes, N.; Hall, M.; Nair, M. A.; Savage, D. F.; Schepartz, A.; Chatterjee, A. A Translation-Independent Directed Evolution Strategy to Engineer Aminoacyl-tRNA Synthetases. *ACS Cent Sci* **2024**, *10* (6), 1211-1220. DOI: 10.1021/acscentsci.3c01557
(115) Biela, A.; Hammermeister, A.; Kaczmarczyk, I.; Walczak, M.; Koziej, L.; Lin, T. Y.; Glatt, S. The diverse structural modes of tRNA binding and recognition. *J Biol Chem* **2023**, *299* (8), 104966. DOI: 10.1016/j.jbc.2023.104966
(116) Phizicky, E. M.; Hopper, A. K. The life and times of a tRNA. *RNA* **2023**, *29* (7), 898-957. DOI: 10.1261/rna.079620.123
(117) Baldridge, K. C.; Jora, M.; Maranhao, A. C.; Quick, M. M.; Addepalli, B.; Brodbelt, J. S.; Ellington, A. D.; Limbach, P. A.; Contreras, L. M. Directed Evolution of Heterologous tRNAs Leads to Reduced Dependence on Post-transcriptional Modifications. *ACS Synth Biol* **2018**, *7* (5), 1315-1327. DOI: 10.1021/acssynbio.7b00421
(118) Guo, J.; Melancon, C. E., 3rd; Lee, H. S.; Groff, D.; Schultz, P. G. Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. *Angew Chem Int Ed Engl* **2009**, *48* (48), 9148-9151. DOI: 10.1002/anie.200904035
(119) Niu, W.; Schultz, P. G.; Guo, J. An expanded genetic code in mammalian cells with a functional quadruplet codon. *ACS Chem Biol* **2013**, *8* (7), 1640-1645. DOI: 10.1021/cb4001662
(120) Wang, N.; Shang, X.; Cerny, R.; Niu, W.; Guo, J. Systematic Evolution and Study of UAGN Decoding tRNAs in a Genomically Recoded Bacteria. *Sci Rep* **2016**, *6*, 21898. DOI: 10.1038/srep21898

(121) Chatterjee, A.; Xiao, H.; Yang, P. Y.; Soundararajan, G.; Schultz, P. G. A tryptophanyl-tRNA synthetase/tRNA pair for unnatural amino acid mutagenesis in E. coli. *Angew Chem Int Ed Engl* **2013**, *52* (19), 5106-5109. DOI: 10.1002/anie.201301094
(122) Neumann, H.; Slusarczyk, A. L.; Chin, J. W. De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J Am Chem Soc* **2010**, *132* (7), 2142-2144. DOI: 10.1021/ja9068722
(123) Wang, N.; Ju, T.; Niu, W.; Guo, J. Fine-tuning interaction between aminoacyl-tRNA synthetase and tRNA for efficient synthesis of proteins containing unnatural amino acids. *ACS Synth Biol* **2015**, *4* (3), 207-212. DOI: 10.1021/sb500195w
(124) Maranhao, A. C.; Ellington, A. D. Evolving Orthogonal Suppressor tRNAs To Incorporate Modified Amino Acids. *ACS Synth Biol* **2017**, *6* (1), 108-119. DOI: 10.1021/acssynbio.6b00145
(125) Zheng, Y.; Lewis, T. L., Jr.; Igo, P.; Polleux, F.; Chatterjee, A. Virus-Enabled Optimization and Delivery of the Genetic Machinery for Efficient Unnatural Amino Acid Mutagenesis in Mammalian Cells and Tissues. *ACS Synth Biol* **2017**, *6* (1), 13-18. DOI: 10.1021/acssynbio.6b00092
(126) Kelemen, R. E.; Mukherjee, R.; Cao, X.; Erickson, S. B.; Zheng, Y.; Chatterjee, A. A Precise Chemical Strategy To Alter the Receptor Specificity of the Adeno-Associated Virus. *Angew Chem Int Ed Engl* **2016**, *55* (36), 10645-10649. DOI: 10.1002/anie.201604067
(127) Jewel, D.; Kelemen, R. E.; Huang, R. L.; Zhu, Z.; Sundaresh, B.; Malley, K.; Pham, Q.; Loynd, C.; Huang, Z.; van Opijnen, T.; et al. Enhanced Directed Evolution in Mammalian Cells Yields a Hyperefficient Pyrrolysyl tRNA for Noncanonical Amino Acid Mutagenesis. *Angew Chem Int Ed Engl* **2024**, *63* (9), e202316428. DOI: 10.1002/anie.202316428
(128) Huang, R. L.; Jewel, D.; Kelemen, R. E.; Pham, Q.; Yared, T. J.; Wang, S.; Singha Roy, S. J.; Huang, Z.; Levinson, S. D.; Sundaresh, B.; et al. Directed Evolution of a Bacterial Leucyl tRNA in Mammalian Cells for Enhanced Noncanonical Amino Acid Mutagenesis. *ACS Synth Biol* **2024**, *13* (7), 2141-2149. DOI: 10.1021/acssynbio.4c00196
(129) Giani, A. M.; Gallo, G. R.; Gianfranceschi, L.; Formenti, G. Long walk to genomics: History and current approaches to genome sequencing and assembly. *Comput Struct Biotechnol J* **2020**, *18*, 9-19. DOI: 10.1016/j.csbj.2019.11.002
(130) Hu, T.; Chitnis, N.; Monos, D.; Dinh, A. Next-generation sequencing technologies: An overview. *Hum Immunol* **2021**, *82* (11), 801-811. DOI: 10.1016/j.humimm.2021.02.012
(131) Krishna, R.; Wang, J.; Ahern, W.; Sturmfels, P.; Venkatesh, P.; Kalvet, I.; Lee, G. R.; Morey-Burrows, F. S.; Anishchenko, I.; Humphreys, I. R.; et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* **2024**, *384* (6693), eadl2528. DOI: 10.1126/science.adl2528
(132) Hie, B. L.; Yang, K. K. Adaptive machine learning for protein engineering. *Curr Opin Struct Biol* **2022**, *72*, 145-152. DOI: 10.1016/j.sbi.2021.11.002
(133) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583-589. DOI: 10.1038/s41586-021-03819-2
(134) Mirdita, M.; Schutze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: making protein folding accessible to all. *Nat Methods* **2022**, *19* (6), 679-682. DOI: 10.1038/s41592-022-01488-1
(135) Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **2023**, *379* (6637), 1123-1130. DOI: 10.1126/science.ade2574
(136) Notin, P.; Rollins, N.; Gal, Y.; Sander, C.; Marks, D. Machine learning for functional protein design. *Nat Biotechnol* **2024**, *42* (2), 216-228. DOI: 10.1038/s41587-024-02127-0
(137) Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; et al. Robust deep learning-based protein sequence design using ProteinMPNN. *Science* **2022**, *378* (6615), 49-56. DOI: 10.1126/science.add2187
(138) Sumida, K. H.; Nunez-Franco, R.; Kalvet, I.; Pellock, S. J.; Wicky, B. I. M.; Milles, L. F.; Dauparas, J.; Wang, J.; Kipnis, Y.; Jameson, N.; et al. Improving Protein Expression, Stability, and Function with ProteinMPNN. *J Am Chem Soc* **2024**, *146* (3), 2054-2061. DOI: 10.1021/jacs.3c10941

(139) Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; et al. De novo design of protein structure and function with RFdiffusion. *Nature* **2023**, *620* (7976), 1089-1100. DOI: 10.1038/s41586-023-06415-8
(140) Qin, X.; Tang, H.; Cao, W.; Dai, Z.; Hu, L.; Huang, Y.; Liu, T. An Orthogonal Tyrosyl-tRNA Synthetase/tRNA Pair from a Thermophilic Bacterium for an Expanded Eukaryotic Genetic Code. *Biochemistry* **2020**, *59* (1), 90-99. DOI: 10.1021/acs.biochem.9b00757
(141) Hu, L.; Qin, X.; Huang, Y.; Cao, W.; Wang, C.; Wang, Y.; Ling, X.; Chen, H.; Wu, D.; Lin, Y.; et al. Thermophilic Pyrrolysyl-tRNA Synthetase Mutants for Enhanced Mammalian Genetic Code Expansion. *ACS Synth Biol* **2020**, *9* (10), 2723-2736. DOI: 10.1021/acssynbio.0c00257
(142) Willis, J. C. W.; Chin, J. W. Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nat Chem* **2018**, *10* (8), 831-837. DOI: 10.1038/s41557-018-0052-5
(143) Zhang, D.; Vaidehi, N.; Goddard, W. A., 3rd; Danzer, J. F.; Debe, D. Structure-based design of mutant Methanococcus jannaschii tyrosyl-tRNA synthetase for incorporation of O-methyl-L-tyrosine. *Proc Natl Acad Sci U S A* **2002**, *99* (10), 6579-6584. DOI: 10.1073/pnas.052150499
(144) Baumann, T.; Hauf, M.; Richter, F.; Albers, S.; Moglich, A.; Ignatova, Z.; Budisa, N. Computational Aminoacyl-tRNA Synthetase Library Design for Photocaged Tyrosine. *Int J Mol Sci* **2019**, *20* (9). DOI: 10.3390/ijms20092343
(145) Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373* (6557), 871-876. DOI: 10.1126/science.abj8754
(146) Hie, B. L.; Shanker, V. R.; Xu, D.; Bruun, T. U. J.; Weidenbacher, P. A.; Tang, S.; Wu, W.; Pak, J. E.; Kim, P. S. Efficient evolution of human antibodies from general protein language models. *Nat Biotechnol* **2024**, *42* (2), 275-283. DOI: 10.1038/s41587-023-01763-2
(147) Schmitt, L. T.; Paszkowski-Rogacz, M.; Jug, F.; Buchholz, F. Prediction of designer-recombinases for DNA editing with generative deep learning. *Nat Commun* **2022**, *13* (1), 7966. DOI: 10.1038/s41467-022-35614-6
(148) Wannier, T. M.; Kunjapur, A. M.; Rice, D. P.; McDonald, M. J.; Desai, M. M.; Church, G. M. Adaptive evolution of genomically recoded Escherichia coli. *Proc Natl Acad Sci U S A* **2018**, *115* (12), 3090-3095. DOI: 10.1073/pnas.1715530115
(149) Zackin, M. T.; Stieglitz, J. T.; Van Deventer, J. A. Genome-Wide Screen for Enhanced Noncanonical Amino Acid Incorporation in Yeast. *ACS Synth Biol* **2022**, *11* (11), 3669-3680. DOI: 10.1021/acssynbio.2c00267
(150) Schrader, J. M.; Chapman, S. J.; Uhlenbeck, O. C. Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc Natl Acad Sci U S A* **2011**, *108* (13), 5215-5220. DOI: 10.1073/pnas.1102128108
(151) 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* (46), 14458-14462. DOI: 10.1021/ja075557u
(152) DeLey Cox, V. E.; Cole, M. F.; Gaucher, E. A. Incorporation of Modified Amino Acids by Engineered Elongation Factors with Expanded Substrate Capabilities. *ACS Synth Biol* **2019**, *8* (2), 287-296. DOI: 10.1021/acssynbio.8b00305
(153) Johnson, D. B.; Xu, J.; Shen, Z.; Takimoto, J. K.; Schultz, M. D.; Schmitz, R. J.; Xiang, Z.; Ecker, J. R.; Briggs, S. P.; Wang, L. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat Chem Biol* **2011**, *7* (11), 779-786. DOI: 10.1038/nchembio.657
(154) Robertson, W. E.; Funke, L. F. H.; de la Torre, D.; Fredens, J.; Elliott, T. S.; Spinck, M.; Christova, Y.; Cervettini, D.; Boge, F. L.; Liu, K. C.; et al. Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science* **2021**, *372* (6546), 1057-1062. DOI: 10.1126/science.abg3029
(155) Kuru, E.; Maattala, R. M.; Noguera, K.; Stork, D. A.; Narasimhan, K.; Rittichier, J.; Wiegand, D.; Church, G. M. Release Factor Inhibiting Antimicrobial Peptides Improve Nonstandard Amino Acid Incorporation in Wild-type Bacterial Cells. *ACS Chem Biol* **2020**, *15* (7), 1852-1861. DOI: 10.1021/acschembio.0c00055

(156) Yi, H.; Zhang, J.; Ke, F.; Guo, X.; Yang, J.; Xie, P.; Liu, L.; Wang, Q.; Gao, X. Comparative Analyses of the Transcriptome and Proteome of Escherichia coli C321. big up tri, openA and Further Improving Its Noncanonical Amino Acids Containing Protein Expression Ability by Integration of T7 RNA Polymerase. *Front Microbiol* **2021**, *12*, 744284. DOI: 10.3389/fmicb.2021.744284
(157) Zheng, Y.; Lajoie, M. J.; Italia, J. S.; Chin, M. A.; Church, G. M.; Chatterjee, A. Performance of optimized noncanonical amino acid mutagenesis systems in the absence of release factor 1. *Mol Biosyst* **2016**, *12* (6), 1746-1749. DOI: 10.1039/c6mb00070c
(158) Conard, S. E.; Buckley, J.; Dang, M.; Bedwell, G. J.; Carter, R. L.; Khass, M.; Bedwell, D. M. Identification of eRF1 residues that play critical and complementary roles in stop codon recognition. *RNA* **2012**, *18* (6), 1210-1221. DOI: 10.1261/rna.031997.111
(159) Lozupone, C. A.; Knight, R. D.; Landweber, L. F. The molecular basis of nuclear genetic code change in ciliates. *Curr Biol* **2001**, *11* (2), 65-74. DOI: 10.1016/s0960-9822(01)00028-8
(160) Kervestin, S.; Frolova, L.; Kisselev, L.; Jean-Jean, O. Stop codon recognition in ciliates: Euplotes release factor does not respond to reassigned UGA codon. *EMBO Rep* **2001**, *2* (8), 680-684. DOI: 10.1093/embo-reports/kve156
(161) Salas-Marco, J.; Fan-Minogue, H.; Kallmeyer, A. K.; Klobutcher, L. A.; Farabaugh, P. J.; Bedwell, D. M. Distinct paths to stop codon reassignment by the variant-code organisms Tetrahymena and Euplotes. *Mol Cell Biol* **2006**, *26* (2), 438-447. DOI: 10.1128/mcb.26.2.438-447.2006
(162) Fan-Minogue, H.; Du, M.; Pisarev, A. V.; Kallmeyer, A. K.; Salas-Marco, J.; Keeling, K. M.; Thompson, S. R.; Pestova, T. V.; Bedwell, D. M. Distinct eRF3 requirements suggest alternate eRF1 conformations mediate peptide release during eukaryotic translation termination. *Mol Cell* **2008**, *30* (5), 599-609. DOI: 10.1016/j.molcel.2008.03.020
(163) Lekomtsev, S.; Kolosov, P.; Bidou, L.; Frolova, L.; Rousset, J. P.; Kisselev, L. Different modes of stop codon restriction by the Stylonychia and Paramecium eRF1 translation termination factors. *Proc Natl Acad Sci U S A* **2007**, *104* (26), 10824-10829. DOI: 10.1073/pnas.0703887104
(164) Alkalaeva, E.; Eliseev, B.; Ambrogelly, A.; Vlasov, P.; Kondrashov, F. A.; Gundllapalli, S.; Frolova, L.; Soll, D.; Kisselev, L. Translation termination in pyrrolysine-utilizing archaea. *FEBS Lett* **2009**, *583* (21), 3455-3460. DOI: 10.1016/j.febslet.2009.09.044
(165) Eliseev, B.; Kryuchkova, P.; Alkalaeva, E.; Frolova, L. A single amino acid change of translation termination factor eRF1 switches between bipotent and omnipotent stop-codon specificity. *Nucleic Acids Res* **2011**, *39* (2), 599-608. DOI: 10.1093/nar/gkq759
(166) Kim, O. T.; Yura, K.; Go, N.; Harumoto, T. Newly sequenced eRF1s from ciliates: the diversity of stop codon usage and the molecular surfaces that are important for stop codon interactions. *Gene* **2005**, *346*, 277-286. DOI: 10.1016/j.gene.2004.11.046
(167) Frolova, L. Y.; Tsivkovskii, R. Y.; Sivolobova, G. F.; Oparina, N. Y.; Serpinsky, O. I.; Blinov, V. M.; Tatkov, S. I.; Kisselev, L. L. Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *RNA* **1999**, *5* (8), 1014-1020. DOI: 10.1017/s135583829999043x
(168) Frolova, L.; Seit-Nebi, A.; Kisselev, L. Highly conserved NIKS tetrapeptide is functionally essential in eukaryotic translation termination factor eRF1. *RNA* **2002**, *8* (2), 129-136. DOI: 10.1017/s1355838202013262
(169) Kolosov, P.; Frolova, L.; Seit-Nebi, A.; Dubovaya, V.; Kononenko, A.; Oparina, N.; Justesen, J.; Efimov, A.; Kisselev, L. Invariant amino acids essential for decoding function of polypeptide release factor eRF1. *Nucleic Acids Res* **2005**, *33* (19), 6418-6425. DOI: 10.1093/nar/gki927
(170) Bertram, G.; Bell, H. A.; Ritchie, D. W.; Fullerton, G.; Stansfield, I. Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. *RNA* **2000**, *6* (9), 1236-1247. DOI: 10.1017/s1355838200000777
(171) Merritt, G. H.; Naemi, W. R.; Mugnier, P.; Webb, H. M.; Tuite, M. F.; von der Haar, T. Decoding accuracy in eRF1 mutants and its correlation with pleiotropic quantitative traits in yeast. *Nucleic Acids Res* **2010**, *38* (16), 5479-5492. DOI: 10.1093/nar/gkq338

(172) Schmied, W. H.; Elsasser, S. J.; Uttamapinant, C.; Chin, J. W. Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J Am Chem Soc* **2014**, *136* (44), 15577-15583. DOI: 10.1021/ja5069728
(173) Sakin, V.; Hanne, J.; Dunder, J.; Anders-Osswein, M.; Laketa, V.; Nikic, I.; Krausslich, H. G.; Lemke, E. A.; Muller, B. A Versatile Tool for Live-Cell Imaging and Super-Resolution Nanoscopy Studies of HIV-1 Env Distribution and Mobility. *Cell Chem Biol* **2017**, *24* (5), 635-645 e635. DOI: 10.1016/j.chembiol.2017.04.007
(174) Jones, C. M.; Robkis, D. M.; Blizzard, R. J.; Munari, M.; Venkatesh, Y.; Mihaila, T. S.; Eddins, A. J.; Mehl, R. A.; Zagotta, W. N.; Gordon, S. E.; et al. Genetic encoding of a highly photostable, long lifetime fluorescent amino acid for imaging in mammalian cells. *Chem Sci* **2021**, *12* (36), 11955-11964. DOI: 10.1039/d1sc01914g
(175) Mondal, S.; Wang, S.; Zheng, Y.; Sen, S.; Chatterjee, A.; Thompson, P. R. Site-specific incorporation of citrulline into proteins in mammalian cells. *Nat Commun* **2021**, *12* (1), 45. DOI: 10.1038/s41467-020-20279-w
(176) Schifferdecker, S.; Zila, V.; Muller, T. G.; Sakin, V.; Anders-Osswein, M.; Laketa, V.; Krausslich, H. G.; Muller, B. Direct Capsid Labeling of Infectious HIV-1 by Genetic Code Expansion Allows Detection of Largely Complete Nuclear Capsids and Suggests Nuclear Entry of HIV-1 Complexes via Common Routes. *mBio* **2022**, *13* (5), e0195922. DOI: 10.1128/mbio.01959-22
(177) Zagotta, W. N.; Sim, B. S.; Nhim, A. K.; Raza, M. M.; Evans, E. G.; Venkatesh, Y.; Jones, C. M.; Mehl, R. A.; Petersson, E. J.; Gordon, S. E. An improved fluorescent noncanonical amino acid for measuring conformational distributions using time-resolved transition metal ion FRET. *Elife* **2021**, *10*. DOI: 10.7554/eLife.70236
(178) Shi, N.; Tong, L.; Lin, H.; Zheng, Z.; Zhang, H.; Dong, L.; Yang, Y.; Shen, Y.; Xia, Q. Optimizing eRF1 to Enable the Genetic Encoding of Three Distinct Noncanonical Amino Acids in Mammalian Cells. *Adv Biol (Weinh)* **2022**, *6* (11), e2200092. DOI: 10.1002/adbi.202200092
(179) Das, D. K.; Bulow, U.; Diehl, W. E.; Durham, N. D.; Senjobe, F.; Chandran, K.; Luban, J.; Munro, J. B. Conformational changes in the Ebola virus membrane fusion machine induced by pH, Ca2+, and receptor binding. *PLoS Biol* **2020**, *18* (2), e3000626. DOI: 10.1371/journal.pbio.3000626
(180) Beranek, V.; Reinkemeier, C. D.; Zhang, M. S.; Liang, A. D.; Kym, G.; Chin, J. W. Genetically Encoded Protein Phosphorylation in Mammalian Cells. *Cell Chem Biol* **2018**, *25* (9), 1067-1074 e1065. DOI: 10.1016/j.chembiol.2018.05.013
(181) Katrekar, D.; Moreno, A. M.; Chen, G.; Worlikar, A.; Mali, P. Oligonucleotide conjugated multi-functional adeno-associated viruses. *Sci Rep* **2018**, *8* (1), 3589. DOI: 10.1038/s41598-018-21742-x
(182) Chen, Z.; Zhang, S.; Li, X.; Ai, H. W. A high-performance genetically encoded fluorescent biosensor for imaging physiological peroxynitrite. *Cell Chem Biol* **2021**, *28* (11), 1542-1553 e1545. DOI: 10.1016/j.chembiol.2021.01.013
(183) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. Construction of modified ribosomes for incorporation of D-amino acids into proteins. *Biochemistry* **2006**, *45* (51), 15541-15551. DOI: 10.1021/bi060986a
(184) Englander, M. T.; Avins, J. L.; Fleisher, R. C.; Liu, B.; Effraim, P. R.; Wang, J.; Schulten, K.; Leyh, T. S.; Gonzalez, R. L., Jr.; Cornish, V. W. The ribosome can discriminate the chirality of amino acids within its peptidyl-transferase center. *Proc Natl Acad Sci U S A* **2015**, *112* (19), 6038-6043. DOI: 10.1073/pnas.1424712112
(185) Katoh, T.; Suga, H. Ribosomal Incorporation of Consecutive beta-Amino Acids. *J Am Chem Soc* **2018**, *140* (38), 12159-12167. DOI: 10.1021/jacs.8b07247
(186) Kofman, C.; Lee, J.; Jewett, M. C. Engineering molecular translation systems. *Cell Syst* **2021**, *12* (6), 593-607. DOI: 10.1016/j.cels.2021.04.001
(187) Hammerling, M. J.; Fritz, B. R.; Yoesep, D. J.; Kim, D. S.; Carlson, E. D.; Jewett, M. C. In vitro ribosome synthesis and evolution through ribosome display. *Nat Commun* **2020**, *11* (1), 1108. DOI: 10.1038/s41467-020-14705-2

(188) Hammerling, M. J.; Kruger, A.; Jewett, M. C. Strategies for in vitro engineering of the translation machinery. *Nucleic Acids Res* **2020**, *48* (3), 1068-1083. DOI: 10.1093/nar/gkz1011
(189) Wang, T.; Liang, C.; An, Y.; Xiao, S.; Xu, H.; Zheng, M.; Liu, L.; Wang, G.; Nie, L. Engineering the Translational Machinery for Biotechnology Applications. *Mol Biotechnol* **2020**, *62* (4), 219-227. DOI: 10.1007/s12033-020-00246-y
(190) Kim, S.; Yi, H.; Kim, Y. T.; Lee, H. S. Engineering Translation Components for Genetic Code Expansion. *J Mol Biol* **2022**, *434* (8), 167302. DOI: 10.1016/j.jmb.2021.167302
(191) Rackham, O.; Chin, J. W. A network of orthogonal ribosome x mRNA pairs. *Nat Chem Biol* **2005**, *1* (3), 159-166. DOI: 10.1038/nchembio719
(192) Wang, K.; Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat Biotechnol* **2007**, *25* (7), 770-777. DOI: 10.1038/nbt1314
(193) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **2010**, *464* (7287), 441-444. DOI: 10.1038/nature08817
(194) Orelle, C.; Carlson, E. D.; Szal, T.; Florin, T.; Jewett, M. C.; Mankin, A. S. Protein synthesis by ribosomes with tethered subunits. *Nature* **2015**, *524* (7563), 119-124. DOI: 10.1038/nature14862
(195) Fried, S. D.; Schmied, W. H.; Uttamapinant, C.; Chin, J. W. Ribosome Subunit Stapling for Orthogonal Translation in E. coli. *Angew Chem Int Ed Engl* **2015**, *54* (43), 12791-12794. DOI: 10.1002/anie.201506311
(196) Stan, A.; Mayer, C. Tethered Ribosomes: Toward the Synthesis of Nonproteinogenic Polymers in Bacteria. *Chembiochem* **2023**, *24* (12), e202200578. DOI: 10.1002/cbic.202200578
(197) Schmied, W. H.; Tnimov, Z.; Uttamapinant, C.; Rae, C. D.; Fried, S. D.; Chin, J. W. Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature* **2018**, *564* (7736), 444-448. DOI: 10.1038/s41586-018-0773-z
(198) Carlson, E. D.; d'Aquino, A. E.; Kim, D. S.; Fulk, E. M.; Hoang, K.; Szal, T.; Mankin, A. S.; Jewett, M. C. Engineered ribosomes with tethered subunits for expanding biological function. *Nat Commun* **2019**, *10* (1), 3920. DOI: 10.1038/s41467-019-11427-y
(199) Liu, Y.; Davis, R. G.; Thomas, P. M.; Kelleher, N. L.; Jewett, M. C. In vitro-Constructed Ribosomes Enable Multi-site Incorporation of Noncanonical Amino Acids into Proteins. *Biochemistry* **2021**, *60* (3), 161-169. DOI: 10.1021/acs.biochem.0c00829
(200) Matheisl, S.; Berninghausen, O.; Becker, T.; Beckmann, R. Structure of a human translation termination complex. *Nucleic Acids Res* **2015**, *43* (18), 8615-8626. DOI: 10.1093/nar/gkv909
(201) Ivanov, A.; Shuvalova, E.; Egorova, T.; Shuvalov, A.; Sokolova, E.; Bizyaev, N.; Shatsky, I.; Terenin, I.; Alkalaeva, E. Polyadenylate-binding protein-interacting proteins PAIP1 and PAIP2 affect translation termination. *J Biol Chem* **2019**, *294* (21), 8630-8639. DOI: 10.1074/jbc.RA118.006856
(202) Potts, K. A.; Stieglitz, J. T.; Lei, M.; Van Deventer, J. A. Reporter system architecture affects measurements of noncanonical amino acid incorporation efficiency and fidelity. *Mol Syst Des Eng* **2020**, *5* (2), 573-588. DOI: 10.1039/c9me00107g
(203) Pott, M.; Schmidt, M. J.; Summerer, D. Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids. *ACS Chem Biol* **2014**, *9* (12), 2815-2822. DOI: 10.1021/cb5006273
(204) Bartoschek, M. D.; Ugur, E.; Nguyen, T. A.; Rodschinka, G.; Wierer, M.; Lang, K.; Bultmann, S. Identification of permissive amber suppression sites for efficient non-canonical amino acid incorporation in mammalian cells. *Nucleic Acids Res* **2021**, *49* (11), e62. DOI: 10.1093/nar/gkab132
(205) Ding, W.; Yu, W.; Chen, Y.; Lao, L.; Fang, Y.; Fang, C.; Zhao, H.; Yang, B.; Lin, S. Rare codon recoding for efficient noncanonical amino acid incorporation in mammalian cells. *Science* **2024**, *384* (6700), 1134-1142. DOI: 10.1126/science.adm8143
(206) Anzalone, A. V.; Zairis, S.; Lin, A. J.; Rabadan, R.; Cornish, V. W. Interrogation of Eukaryotic Stop Codon Readthrough Signals by in Vitro RNA Selection. *Biochemistry* **2019**, *58* (8), 1167-1178. DOI: 10.1021/acs.biochem.8b01280

(207) Swart, E. C.; Serra, V.; Petroni, G.; Nowacki, M. Genetic Codes with No Dedicated Stop Codon: Context-Dependent Translation Termination. *Cell* **2016**, *166* (3), 691-702. DOI: 10.1016/j.cell.2016.06.020
(208) Mangkalaphiban, K.; He, F.; Ganesan, R.; Wu, C.; Baker, R.; Jacobson, A. Transcriptome-wide investigation of stop codon readthrough in Saccharomyces cerevisiae. *PLoS Genet* **2021**, *17* (4), e1009538. DOI: 10.1371/journal.pgen.1009538
(209) Wangen, J. R.; Green, R. Stop codon context influences genome-wide stimulation of termination codon readthrough by aminoglycosides. *Elife* **2020**, *9*. DOI: 10.7554/eLife.52611
(210) Lajoie, M. J.; Rovner, A. J.; Goodman, D. B.; Aerni, H. R.; Haimovich, A. D.; Kuznetsov, G.; Mercer, J. A.; Wang, H. H.; Carr, P. A.; Mosberg, J. A.; et al. Genomically recoded organisms expand biological functions. *Science* **2013**, *342* (6156), 357-360. DOI: 10.1126/science.1241459
(211) Dymond, J.; Boeke, J. The Saccharomyces cerevisiae SCRaMbLE system and genome minimization. *Bioeng Bugs* **2012**, *3* (3), 168-171. DOI: 10.4161/bbug.19543
(212) Richardson, S. M.; Mitchell, L. A.; Stracquadanio, G.; Yang, K.; Dymond, J. S.; DiCarlo, J. E.; Lee, D.; Huang, C. L.; Chandrasegaran, S.; Cai, Y.; et al. Design of a synthetic yeast genome. *Science* **2017**, *355* (6329), 1040-1044. DOI: 10.1126/science.aaf4557
(213) Isaacs, F. J.; Carr, P. A.; Wang, H. H.; Lajoie, M. J.; Sterling, B.; Kraal, L.; Tolonen, A. C.; Gianoulis, T. A.; Goodman, D. B.; Reppas, N. B.; et al. Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* **2011**, *333* (6040), 348-353. DOI: 10.1126/science.1205822
(214) Annaluru, N.; Muller, H.; Mitchell, L. A.; Ramalingam, S.; Stracquadanio, G.; Richardson, S. M.; Dymond, J. S.; Kuang, Z.; Scheifele, L. Z.; Cooper, E. M. Total synthesis of a functional designer eukaryotic chromosome. *science* **2014**, *344* (6179), 55-58.
(215) Rovner, A. J.; Haimovich, A. D.; Katz, S. R.; Li, Z.; Grome, M. W.; Gassaway, B. M.; Amiram, M.; Patel, J. R.; Gallagher, R. R.; Rinehart, J.; et al. Recoded organisms engineered to depend on synthetic amino acids. *Nature* **2015**, *518* (7537), 89-93. DOI: 10.1038/nature14095
(216) Wang, H. H.; Isaacs, F. J.; Carr, P. A.; Sun, Z. Z.; Xu, G.; Forest, C. R.; Church, G. M. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **2009**, *460* (7257), 894-898. DOI: 10.1038/nature08187
(217) Wu, Y.; Li, B. Z.; Zhao, M.; Mitchell, L. A.; Xie, Z. X.; Lin, Q. H.; Wang, X.; Xiao, W. H.; Wang, Y.; Zhou, X.; et al. Bug mapping and fitness testing of chemically synthesized chromosome X. *Science* **2017**, *355* (6329). DOI: 10.1126/science.aaf4706
(218) Zhao, Y.; Coelho, C.; Hughes, A. L.; Lazar-Stefanita, L.; Yang, S.; Brooks, A. N.; Walker, R. S. K.; Zhang, W.; Lauer, S.; Hernandez, C.; et al. Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions. *Cell* **2023**, *186* (24), 5220-5236 e5216. DOI: 10.1016/j.cell.2023.09.025
(219) Sanders, J.; Hoffmann, S. A.; Green, A. P.; Cai, Y. New opportunities for genetic code expansion in synthetic yeast. *Curr Opin Biotechnol* **2022**, *75*, 102691. DOI: 10.1016/j.copbio.2022.102691
(220) Barbieri, E. M.; Muir, P.; Akhuetie-Oni, B. O.; Yellman, C. M.; Isaacs, F. J. Precise Editing at DNA Replication Forks Enables Multiplex Genome Engineering in Eukaryotes. *Cell* **2017**, *171* (6), 1453-1467 e1413. DOI: 10.1016/j.cell.2017.10.034
(221) Osawa, S.; Ohama, T.; Jukes, T. H.; Watanabe, K. Evolution of the mitochondrial genetic code. I. Origin of AGR serine and stop codons in metazoan mitochondria. *J Mol Evol* **1989**, *29* (3), 202-207. DOI: 10.1007/bf02100203
(222) Osawa, S.; Jukes, T. H.; Watanabe, K.; Muto, A. Recent evidence for evolution of the genetic code. *Microbiol Rev* **1992**, *56* (1), 229-264. DOI: 10.1128/mr.56.1.229-264.1992
(223) Kachale, A.; Pavlikova, Z.; Nenarokova, A.; Roithova, A.; Durante, I. M.; Miletinova, P.; Zahonova, K.; Nenarokov, S.; Votypka, J.; Horakova, E.; et al. Short tRNA anticodon stem and mutant eRF1 allow stop codon reassignment. *Nature* **2023**, *613* (7945), 751-758. DOI: 10.1038/s41586-022-05584-2

(224) Bezerra, A. R.; Guimaraes, A. R.; Santos, M. A. Non-Standard Genetic Codes Define New Concepts for Protein Engineering. *Life (Basel)* **2015**, *5* (4), 1610-1628. DOI: 10.3390/life5041610
(225) Jewett, M. C.; Fritz, B. R.; Timmerman, L. E.; Church, G. M. In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Mol Syst Biol* **2013**, *9*, 678. DOI: 10.1038/msb.2013.31
(226) Kawakami, T.; Murakami, H.; Suga, H. Ribosomal synthesis of polypeptoids and peptoid-peptide hybrids. *J Am Chem Soc* **2008**, *130* (50), 16861-16863. DOI: 10.1021/ja806998v
(227) Tan, Z.; Forster, A. C.; Blacklow, S. C.; Cornish, V. W. Amino acid backbone specificity of the Escherichia coli translation machinery. *J Am Chem Soc* **2004**, *126* (40), 12752-12753. DOI: 10.1021/ja0472174
(228) Melo Czekster, C.; Robertson, W. E.; Walker, A. S.; Soll, D.; Schepartz, A. In Vivo Biosynthesis of a beta-Amino Acid-Containing Protein. *J Am Chem Soc* **2016**, *138* (16), 5194-5197. DOI: 10.1021/jacs.6b01023
(229) Daskalova, S. M.; Dedkova, L. M.; Maini, R.; Talukder, P.; Bai, X.; Chowdhury, S. R.; Zhang, C.; Nangreave, R. C.; Hecht, S. M. Elongation Factor P Modulates the Incorporation of Structurally Diverse Noncanonical Amino Acids into Escherichia coli Dihydrofolate Reductase. *J Am Chem Soc* **2023**, *145* (43), 23600-23608. DOI: 10.1021/jacs.3c07524
(230) Maini, R.; Nguyen, D. T.; Chen, S.; Dedkova, L. M.; Chowdhury, S. R.; Alcala-Torano, R.; Hecht, S. M. Incorporation of beta-amino acids into dihydrofolate reductase by ribosomes having modifications in the peptidyltransferase center. *Bioorg Med Chem* **2013**, *21* (5), 1088-1096. DOI: 10.1016/j.bmc.2013.01.002
(231) Katoh, T.; Suga, H. Ribosomal incorporation of negatively charged d-alpha- and N-methyl-l-alpha-amino acids enhanced by EF-Sep. *Philos Trans R Soc Lond B Biol Sci* **2023**, *378* (1871), 20220038. DOI: 10.1098/rstb.2022.0038
(232) Seki, K.; Galindo, J. L.; Karim, A. S.; Jewett, M. C. A Cell-Free Gene Expression Platform for Discovering and Characterizing Stop Codon Suppressing tRNAs. *ACS Chem Biol* **2023**, *18* (6), 1324-1334. DOI: 10.1021/acschembio.3c00051
(233) Martin, R. W.; Des Soye, B. J.; Kwon, Y. C.; Kay, J.; Davis, R. G.; Thomas, P. M.; Majewska, N. I.; Chen, C. X.; Marcum, R. D.; Weiss, M. G.; et al. Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat Commun* **2018**, *9* (1), 1203. DOI: 10.1038/s41467-018-03469-5
(234) Ranji Charna, A.; Des Soye, B. J.; Ntai, I.; Kelleher, N. L.; Jewett, M. C. An efficient cell-free protein synthesis platform for producing proteins with pyrrolysine-based noncanonical amino acids. *Biotechnol J* **2022**, *17* (9), e2200096. DOI: 10.1002/biot.202200096
(235) Cui, Z.; Johnston, W. A.; Alexandrov, K. Cell-Free Approach for Non-canonical Amino Acids Incorporation Into Polypeptides. *Front Bioeng Biotechnol* **2020**, *8*, 1031. DOI: 10.3389/fbioe.2020.01031
(236) Aydin, Y.; Coin, I. Genetically encoded crosslinkers to address protein-protein interactions. *Protein Sci* **2023**, *32* (5), e4637. DOI: 10.1002/pro.4637
(237) Tang, Y.; Tirrell, D. A. Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host. *J Am Chem Soc* **2001**, *123* (44), 11089-11090. DOI: 10.1021/ja016652k
(238) Bilgiçer, B.; Xing, X.; Kumar, K. Programmed self-sorting of coiled coils with leucine and hexafluoroleucine cores. *J Am Chem Soc* **2001**, *123* (47), 11815-11816. DOI: DOI 10.1021/ja016767o
(239) Budisa, N.; Rubini, M.; Bae, J. H.; Weyher, E.; Wenger, W.; Golbik, R.; Huber, R.; Moroder, L. Global replacement of tryptophan with aminotryptophans generates non-invasive protein-based optical pH sensors. *Angew Chem Int Edit* **2002**, *41* (21), 4066-4069. DOI: Doi 10.1002/1521-3773(20021104)41:21<4066::Aid-Anie4066>3.0.Co;2-6
(240) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-canonical amino acids in protein engineering. *Curr Opin Biotech* **2003**, *14* (6), 603-609. DOI: 10.1016/j.copbio.2003.10.011
(241) Chen, C.; Zhang, H.; Zhang, J.; Ai, H. W.; Fang, C. Structural origin and rational development of bright red noncanonical variants of green fluorescent protein. *Phys Chem Chem Phys* **2023**, *25* (23), 15624-15634. DOI: 10.1039/d3cp01315d

(242) Xiang, Z.; Ren, H.; Hu, Y. S.; Coin, I.; Wei, J.; Cang, H.; Wang, L. Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. *Nat Methods* **2013**, *10* (9), 885-888. DOI: 10.1038/nmeth.2595
(243) Van Deventer, J. A.; Yuet, K. P.; Yoo, T. H.; Tirrell, D. A. Cell surface display yields evolvable, clickable antibody fragments. *Chembiochem* **2014**, *15* (12), 1777-1781. DOI: 10.1002/cbic.201402184
(244) Alcala-Torano, R.; Islam, M.; Cika, J.; Lam, K. H.; Jin, R.; Ichtchenko, K.; Shoemaker, C. B.; Van Deventer, J. A. Yeast Display Enables Identification of Covalent Single-Domain Antibodies against Botulinum Neurotoxin Light Chain A. *ACS Chem Biol* **2022**, *17* (12), 3435-3449. DOI: 10.1021/acschembio.2c00574
(245) Montclare, J. K.; Tirrell, D. A. Evolving proteins of novel composition. *Angew Chem Int Ed Engl* **2006**, *45* (27), 4518-4521. DOI: 10.1002/anie.200600088
(246) Budisa, N.; Huber, R.; Golbik, R.; Minks, C.; Weyher, E.; Moroder, L. Atomic mutations in annexin V--thermodynamic studies of isomorphous protein variants. *Eur J Biochem* **1998**, *253* (1), 1-9. DOI: 10.1046/j.1432-1327.1998.2530001.x
(247) Wang, L.; Xie, J. M.; Deniz, A. A.; Schultz, P. G. Unnatural amino acid mutagenesis of green fluorescent protein. *J Org Chem* **2003**, *68* (1), 174-176. DOI: 10.1021/jo026570u
(248) Beiboer, S. H. W.; vandenBerg, B.; Dekker, N.; Cox, R. C.; Verheij, H. M. Incorporation of an unnatural amino acid in the active site of porcine pancreatic phospholipase A(2). Substitution of histidine by 1,2,4-triazole-3-alanine yields an enzyme with high activity at acidic pH. *Protein Eng* **1996**, *9* (4), 345-352. DOI: DOI 10.1093/protein/9.4.345
(249) Soumillion, P.; Fastrez, J. Incorporation of 1,2,4-triazole-3-alanine into a mutant of phage lambda lysozyme containing a single histidine. *Protein Eng* **1998**, *11* (3), 213-217. DOI: DOI 10.1093/protein/11.3.213
(250) Wolschner, C.; Giese, A.; Kretzschmar, H. A.; Huber, R.; Moroder, L.; Budisa, N. Design of anti- and pro-aggregation variants to assess the effects of methionine oxidation in human prion protein. *P Natl Acad Sci USA* **2009**, *106* (19), 7756-7761. DOI: 10.1073/pnas.0902688106
(251) Panchenko, T.; Zhu, W. W.; Montclare, J. K. Influence of global fluorination on chloramphenicol acetyltransferase activity and stability. *Biotechnol Bioeng* **2006**, *94* (5), 921-930. DOI: 10.1002/bit.20910
(252) Yoo, T. H.; Link, A. J.; Tirrell, D. A. Evolution of a fluorinated green fluorescent protein. *Proc Natl Acad Sci U S A* **2007**, *104* (35), 13887-13890. DOI: 10.1073/pnas.0701904104
(253) Moore, E. J.; Zorine, D.; Hansen, W. A.; Khare, S. D.; Fasan, R. Enzyme stabilization via computationally guided protein stapling. *P Natl Acad Sci USA* **2017**, *114* (47), 12472-12477. DOI: 10.1073/pnas.1708907114
(254) Li, J. C.; Liu, T.; Wang, Y.; Mehta, A. P.; Schultz, P. G. Enhancing Protein Stability with Genetically Encoded Noncanonical Amino Acids. *J Am Chem Soc* **2018**, *140* (47), 15997-16000. DOI: 10.1021/jacs.8b07157
(255) Li, J. C.; Nastertorabi, F.; Xuan, W. M.; Han, G. W.; Stevens, R. C.; Schultz, P. G. A Single Reactive Noncanonical Amino Acid Is Able to Dramatically Stabilize Protein Structure. *Acs Chem Biol* **2019**, *14* (6), 1150-1153. DOI: 10.1021/acschembio.9b00002
(256) Iannuzzelli, J. A.; Bacik, J. P.; Moore, E. J.; Shen, Z. F.; Irving, E. M.; Vargas, D. A.; Khare, S. D.; Ando, N.; Fasan, R. Tuning Enzyme Thermostability via Computationally Guided Covalent Stapling and Structural Basis of Enhanced Stabilization. *Biochemistry* **2022**, *61* (11), 1041-1054. DOI: 10.1021/acs.biochem.2c00033
(257) Hossack, E. J.; Hardy, F. J.; Green, A. P. Building Enzymes through Design and Evolution. *Acs Catal* **2023**, *13* (19), 12436-12444. DOI: 10.1021/acscatal.3c02746
(258) Birch-Price, Z.; Taylor, C. J.; Ortmayer, M.; Green, A. P. Engineering enzyme activity using an expanded amino acid alphabet. *Protein Eng Des Sel* **2023**, *36*. DOI: 10.1093/protein/gzac013
(259) Lovelock, S. L.; Crawshaw, R.; Basler, S.; Levy, C.; Baker, D.; Hilvert, D.; Green, A. P. The road to fully programmable protein catalysis. *Nature* **2022**, *606* (7912), 49-58. DOI: 10.1038/s41586-022-04456-z

(260) Zhao, J.; Burke, A. J.; Green, A. P. Enzymes with noncanonical amino acids. *Curr Opin Chem Biol* **2020**, *55*, 136-144. DOI: 10.1016/j.cbpa.2020.01.006
(261) Birch-Price, Z.; Hardy, F. J.; Lister, T. M.; Kohn, A. R.; Green, A. P. Noncanonical Amino Acids in Biocatalysis. *Chem Rev* **2024**, *124* (14), 8740-8786. DOI: 10.1021/acs.chemrev.4c00120
(262) Bornscheuer, U. T.; Kazlauskas, R. J. Catalytic promiscuity in biocatalysis: Using old enzymes to form new bonds and follow new pathways. *Angew Chem Int Edit* **2004**, *43* (45), 6032-6040. DOI: 10.1002/anie.200460416
(263) Busto, E.; Gotor-Fernández, V.; Gotor, V. Hydrolases: catalytically promiscuous enzymes for non-conventional reactions in organic synthesis. *Chem Soc Rev* **2010**, *39* (11), 4504-4523. DOI: 10.1039/c003811c
(264) Chatterjee, A.; Mallin, H.; Klehr, J.; Vallapurackal, J.; Finke, A. D.; Vera, L.; Marsh, M.; Ward, T. R. An enantioselective artificial Suzukiase based on the biotin-streptavidin technology. *Chemical Science* **2016**, *7* (1), 673-677. DOI: 10.1039/c5sc03116h
(265) Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. Artificial metalloenzymes for asymmetric allylic alkylation on the basis of the biotin-avidin technology. *Angew Chem Int Edit* **2008**, *47* (4), 701-705. DOI: 10.1002/anie.200703159
(266) Giver, L.; Gershenson, A.; Freskgard, P. O.; Arnold, F. H. Directed evolution of a thermostable esterase. *P Natl Acad Sci USA* **1998**, *95* (22), 12809-12813. DOI: DOI 10.1073/pnas.95.22.12809
(267) Kan, S. B. J.; Lewis, R. D.; Chen, K.; Arnold, F. H. Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. *Science* **2016**, *354* (6315), 1048-1051. DOI: 10.1126/science.aah6219
(268) Mayer, C.; Dulson, C.; Reddem, E.; Thunnissen, A. M. W. H.; Roelfes, G. Directed Evolution of a Designer Enzyme Featuring an Unnatural Catalytic Amino Acid. *Angew Chem Int Edit* **2019**, *58* (7), 2083-2087. DOI: 10.1002/anie.201813499
(269) Burke, A. J.; Lovelock, S. L.; Frese, A.; Crawshaw, R.; Ortmayer, M.; Dunstan, M.; Levy, C.; Green, A. P. Design and evolution of an enzyme with a non-canonical organocatalytic mechanism. *Nature* **2019**, *570* (7760), 219-223. DOI: 10.1038/s41586-019-1262-8
(270) Yu, M. Z.; Yuan, Y.; Li, Z. J.; Kunthic, T.; Wang, H. X.; Xu, C.; Xiang, Z. An Artificial Enzyme for Asymmetric Nitrocyclopropanation of alpha,beta-Unsaturated Aldehydes - Design and Evolution. *Angew Chem Int Ed Engl* **2024**, e202401635. DOI: 10.1002/anie.202401635
(271) Trimble, J. S.; Crawshaw, R.; Hardy, F. J.; Levy, C. W.; Brown, M. J. B.; Fuerst, D. E.; Heyes, D. J.; Obexer, R.; Green, A. P. A designed photoenzyme for enantioselective [2+2] cycloadditions. *Nature* **2022**, *611* (7937), 709-714. DOI: 10.1038/s41586-022-05335-3
(272) Moore, E. J.; Fasan, R. Effect of proximal ligand substitutions on the carbene and nitrene transferase activity of myoglobin. *Tetrahedron* **2019**, *75* (16), 2357-2363. DOI: 10.1016/j.tet.2019.03.009
(273) Almhjell, P. J.; Mills, J. H. Metal-chelating non-canonical amino acids in metalloprotein engineering and design. *Curr Opin Struc Biol* **2018**, *51*, 170-176. DOI: 10.1016/j.sbi.2018.06.001
(274) Kang, M. C.; Light, K.; Ai, H. W.; Shen, W. J.; Kim, C. H.; Chen, P. R.; Lee, H. S.; Solomon, E. I.; Schultz, P. G. Evolution of Iron(II)-Finger Peptides by Using a Bipyridyl Amino Acid. *Chembiochem* **2014**, *15* (6), 822-825. DOI: 10.1002/cbic.201300727
(275) Pott, M.; Hayashi, T.; Mori, T.; Mittl, P. R. E.; Green, A. P.; Hilvert, D. A Noncanonical Proximal Heme Ligand Affords an Efficient Peroxidase in a Globin Fold. *J Am Chem Soc* **2018**, *140* (4), 1535-1543. DOI: 10.1021/jacs.7b12621
(276) Drienovská, I.; Roelfes, G. Expanding the enzyme universe with genetically encoded unnatural amino acids. *Nat Catal* **2020**, *3* (3), 193-202. DOI: 10.1038/s41929-019-0410-8
(277) Wang, C. K.; Craik, D. J. Designing macrocyclic disulfide-rich peptides for biotechnological applications. *Nature Chemical Biology* **2018**, *14* (5), 417-427. DOI: 10.1038/s41589-018-0039-y
(278) Zha, J.; Li, J. J.; Fan, S. H.; Duan, Z. P.; Zhao, Y. B.; Wu, C. L. An evolution-inspired strategy to design disulfide-rich peptides tolerant to extensive sequence manipulation. *Chemical Science* **2021**, *12* (34), 11464-11472. DOI: 10.1039/d1sc02952e

(279) Li, S. W.; Roberts, R. W. A novel strategy for in vitro selection of peptide-drug conjugates. *Chemistry & Biology* **2003**, *10* (3), 233-239. DOI: 10.1016/S1074-5521(03)00047-4
(280) Bruce, A.; Adebomi, V.; Czabala, P.; Palmer, J.; McFadden, W. M.; Lorson, Z. C.; Slack, R. L.; Bhardwaj, G.; Sarafianos, S. G.; Raj, M. A Tag-Free Platform for Synthesis and Screening of Cyclic Peptide Libraries. *Angew Chem Int Edit* **2024**. DOI: 10.1002/anie.202320045
(281) Lin, Y. Y.; Penna, M.; Spicer, C. D.; Higgins, S. G.; Gelmi, A.; Kim, N.; Wang, S. T.; Wojciechowski, J. P.; Pashuck, E. T.; Yarovsky, I.; et al. High-Throughput Peptide Derivatization toward Supramolecular Diversification in Microtiter Plates. *Acs Nano* **2021**, *15* (3), 4034-4044. DOI: 10.1021/acsnano.0c05423
(282) Tian, F.; Tsao, M. L.; Schultz, P. G. A phage display system with unnatural amino acids. *J Am Chem Soc* **2004**, *126* (49), 15962-15963. DOI: 10.1021/ja045673m
(283) Liu, C. C.; Mack, A. V.; Tsao, M. L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V. Protein evolution with an expanded genetic code. *P Natl Acad Sci USA* **2008**, *105* (46), 17688-17693. DOI: 10.1073/pnas.0809543105
(284) Van Deventer, J. A.; Le, D. N.; Zhao, J.; Kehoe, H. P.; Kelly, R. L. A platform for constructing, evaluating, and screening bioconjugates on the yeast surface. *Protein Eng Des Sel* **2016**, *29* (11), 485-494. DOI: 10.1093/protein/gzw029
(285) Oller-Salvia, B.; Chin, J. W. Efficient Phage Display with Multiple Distinct Non-Canonical Amino Acids Using Orthogonal Ribosome-Mediated Genetic Code Expansion. *Angew Chem Int Ed Engl* **2019**, *58* (32), 10844-10848. DOI: 10.1002/anie.201902658
(286) Tharp, J. M.; Hampton, J. T.; Reed, C. A.; Ehnbom, A.; Chen, P. C.; Morse, J. S.; Kurra, Y.; Perez, L. M.; Xu, S.; Liu, W. R. An amber obligate active site-directed ligand evolution technique for phage display. *Nat Commun* **2020**, *11* (1), 1392. DOI: 10.1038/s41467-020-15057-7
(287) Owens, A. E.; Iannuzzelli, J. A.; Gu, Y.; Fasan, R. MOrPH-PhD: An Integrated Phage Display Platform for the Discovery of Functional Genetically Encoded Peptide Macrocycles. *ACS Cent Sci* **2020**, *6* (3), 368-381. DOI: 10.1021/acscentsci.9b00927
(288) Morse, J. S.; Sheng, Y. J.; Hampton, J. T.; Sylvain, L. D.; Das, S.; Alugubelli, Y. R.; Chen, P. C.; Yang, K. S.; Xu, S.; Fierke, C. A.; et al. Phage-assisted, active site-directed ligand evolution of a potent and selective histone deacetylase 8 inhibitor. *Protein Sci* **2022**, *31* (12), e4512. DOI: 10.1002/pro.4512
(289) Lewis, A. K.; Harthorn, A.; Johnson, S. M.; Lobb, R. R.; Hackel, B. J. Engineered protein-small molecule conjugates empower selective enzyme inhibition. *Cell Chem Biol* **2022**, *29* (2), 328-338 e324. DOI: 10.1016/j.chembiol.2021.07.013
(290) Iskandar, S. E.; Haberman, V. A.; Bowers, A. A. Expanding the Chemical Diversity of Genetically Encoded Libraries. *ACS Comb Sci* **2020**, *22* (12), 712-733. DOI: 10.1021/acscombsci.0c00179
(291) Passioura, T.; Katoh, T.; Goto, Y.; Suga, H. Selection-based discovery of druglike macrocyclic peptides. *Annu Rev Biochem* **2014**, *83*, 727-752. DOI: 10.1146/annurev-biochem-060713-035456
(292) Huang, Y.; Wiedmann, M. M.; Suga, H. RNA Display Methods for the Discovery of Bioactive Macrocycles. *Chem Rev* **2019**, *119* (17), 10360-10391. DOI: 10.1021/acs.chemrev.8b00430
(293) Passioura, T.; Liu, W.; Dunkelmann, D.; Higuchi, T.; Suga, H. Display Selection of Exotic Macrocyclic Peptides Expressed under a Radically Reprogrammed 23 Amino Acid Genetic Code. *J Am Chem Soc* **2018**, *140* (37), 11551-11555. DOI: 10.1021/jacs.8b03367
(294) Schlippe, Y. V.; Hartman, M. C.; Josephson, K.; Szostak, J. W. In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors. *J Am Chem Soc* **2012**, *134* (25), 10469-10477. DOI: 10.1021/ja301017y
(295) Millward, S. W.; Takahashi, T. T.; Roberts, R. W. A general route for post-translational cyclization of mRNA display libraries. *J Am Chem Soc* **2005**, *127* (41), 14142-14143. DOI: 10.1021/ja054373h
(296) White, E. R.; Sun, L.; Ma, Z.; Beckta, J. M.; Danzig, B. A.; Hacker, D. E.; Huie, M.; Williams, D. C.; Edwards, R. A.; Valerie, K.; et al. Peptide library approach to uncover phosphomimetic inhibitors of the BRCA1 C-terminal domain. *ACS Chem Biol* **2015**, *10* (5), 1198-1208. DOI: 10.1021/cb500757u
(297) Lee, M. L.; Farag, S.; Del Cid, J. S.; Bashore, C.; Hallenbeck, K. K.; Gobbi, A.; Cunningham, C. N. Identification of Macrocyclic Peptide Families from Combinatorial Libraries Containing Noncanonical

Amino Acids Using Cheminformatics and Bioinformatics Inspired Clustering. *ACS Chem Biol* **2023**, *18* (6), 1425-1434. DOI: 10.1021/acschembio.3c00159
(298) Wu, Y.; Bertran, M. T.; Joshi, D.; Maslen, S. L.; Hurd, C.; Walport, L. J. Identification of photocrosslinking peptide ligands by mRNA display. *Commun Chem* **2023**, *6* (1), 103. DOI: 10.1038/s42004-023-00898-2
(299) Iskandar, S. E.; Chiou, L. F.; Leisner, T. M.; Shell, D. J.; Norris-Drouin, J. L.; Vaziri, C.; Pearce, K. H.; Bowers, A. A. Identification of Covalent Cyclic Peptide Inhibitors in mRNA Display. *J Am Chem Soc* **2023**, *145* (28), 15065-15070. DOI: 10.1021/jacs.3c04833
(300) Alleyne, C.; Amin, R. P.; Bhatt, B.; Bianchi, E.; Blain, J. C.; Boyer, N.; Branca, D.; Embrey, M. W.; Ha, S. N.; Jette, K.; et al. Series of Novel and Highly Potent Cyclic Peptide PCSK9 Inhibitors Derived from an mRNA Display Screen and Optimized via Structure-Based Design. *J Med Chem* **2020**, *63* (22), 13796-13824. DOI: 10.1021/acs.jmedchem.0c01084
(301) Tucker, T. J.; Embrey, M. W.; Alleyne, C.; Amin, R. P.; Bass, A.; Bhatt, B.; Bianchi, E.; Branca, D.; Bueters, T.; Buist, N.; et al. A Series of Novel, Highly Potent, and Orally Bioavailable Next-Generation Tricyclic Peptide PCSK9 Inhibitors. *J Med Chem* **2021**, *64* (22), 16770-16800. DOI: 10.1021/acs.jmedchem.1c01599
(302) Ohta, A.; Tanada, M.; Shinohara, S.; Morita, Y.; Nakano, K.; Yamagishi, Y.; Takano, R.; Kariyuki, S.; Iida, T.; Matsuo, A.; et al. Validation of a New Methodology to Create Oral Drugs beyond the Rule of 5 for Intracellular Tough Targets. *J Am Chem Soc* **2023**, *145* (44), 24035-24051. DOI: 10.1021/jacs.3c07145
(303) Wang, X. S.; Chen, P. C.; Hampton, J. T.; Tharp, J. M.; Reed, C. A.; Das, S. K.; Wang, D. S.; Hayatshahi, H. S.; Shen, Y.; Liu, J.; et al. A Genetically Encoded, Phage-Displayed Cyclic-Peptide Library. *Angew Chem Int Ed Engl* **2019**, *58* (44), 15904-15909. DOI: 10.1002/anie.201908713
(304) Hampton, J. T.; Cho, C. D.; Coleman, D. D.; Geng, Z. Z.; Chen, P. C.; Dubey, G. K.; Sylvain, L. D.; Xu, S.; Liu, W. R. An amber-encoding helper phage for more efficient phage display of noncanonical amino acids. *Nucleic Acids Res* **2023**, *51* (13), 6566-6577. DOI: 10.1093/nar/gkad488
(305) Hampton, J. T.; Liu, W. R. Diversification of Phage-Displayed Peptide Libraries with Noncanonical Amino Acid Mutagenesis and Chemical Modification. *Chem Rev* **2024**, *124* (9), 6051-6077. DOI: 10.1021/acs.chemrev.4c00004
(306) Getz, J. A.; Schoep, T. D.; Daugherty, P. S. Peptide discovery using bacterial display and flow cytometry. *Methods Enzymol* **2012**, *503*, 75-97. DOI: 10.1016/B978-0-12-396962-0.00004-5
(307) Case, M.; Navaratna, T.; Vinh, J.; Thurber, G. Rapid Evaluation of Staple Placement in Stabilized alpha Helices Using Bacterial Surface Display. *ACS Chem Biol* **2023**, *18* (4), 905-914. DOI: 10.1021/acschembio.3c00048
(308) Case, M.; Smith, M.; Vinh, J.; Thurber, G. Machine learning to predict continuous protein properties from binary cell sorting data and map unseen sequence space. *Proc Natl Acad Sci U S A* **2024**, *121* (11), e2311726121. DOI: 10.1073/pnas.2311726121
(309) Obexer, R.; Walport, L. J.; Suga, H. Exploring sequence space: harnessing chemical and biological diversity towards new peptide leads. *Current Opinion in Chemical Biology* **2017**, *38*, 52-61. DOI: 10.1016/j.cbpa.2017.02.020
(310) Faber, M. S.; Whitehead, T. A. Data-driven engineering of protein therapeutics. *Curr Opin Biotechnol* **2019**, *60*, 104-110. DOI: 10.1016/j.copbio.2019.01.015
(311) Li, A.; Voleti, R.; Lee, M.; Gagoski, D.; Shah, N. H. High-throughput profiling of sequence recognition by tyrosine kinases and SH2 domains using bacterial peptide display. *Elife* **2023**, *12*. DOI: 10.7554/eLife.82345
(312) Barber, K. W.; Muir, P.; Szeligowski, R. V.; Rogulina, S.; Gerstein, M.; Sampson, J. R.; Isaacs, F. J.; Rinehart, J. Encoding human serine phosphopeptides in bacteria for proteome-wide identification of phosphorylation-dependent interactions. *Nat Biotechnol* **2018**, *36* (7), 638-644. DOI: 10.1038/nbt.4150
(313) Carter, P. J.; Rajpal, A. Designing antibodies as therapeutics. *Cell* **2022**, *185* (15), 2789-2805. DOI: 10.1016/j.cell.2022.05.029

(314) Wells, J. A.; Kumru, K. Extracellular targeted protein degradation: an emerging modality for drug discovery. *Nat Rev Drug Discov* **2024**, *23* (2), 126-140. DOI: 10.1038/s41573-023-00833-z
(315) Liu, C. C.; Choe, H.; Farzan, M.; Smider, V. V.; Schultz, P. G. Mutagenesis and evolution of sulfated antibodies using an expanded genetic code. *Biochemistry* **2009**, *48* (37), 8891-8898. DOI: 10.1021/bi9011429
(316) Liu, C. C.; Mack, A. V.; Tsao, M. L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V. Protein evolution with an expanded genetic code. *Proc Natl Acad Sci U S A* **2008**, *105* (46), 17688-17693. DOI: 10.1073/pnas.0809543105
(317) Rezhdo, A.; Lessard, C. T.; Islam, M.; Van Deventer, J. A. Strategies for enriching and characterizing proteins with inhibitory properties on the yeast surface. *Protein Eng Des Sel* **2023**, *36*. DOI: 10.1093/protein/gzac017
(318) Huang, M.; Rueda-Garcia, M.; Harthorn, A.; Hackel, B. J.; Van Deventer, J. A. Systematic Evaluation of Protein-Small Molecule Hybrids on the Yeast Surface. *ACS Chem Biol* **2024**, *19* (2), 325-335. DOI: 10.1021/acschembio.3c00524
(319) Stieglitz, J. T.; Van Deventer, J. A. Incorporating, Quantifying, and Leveraging Noncanonical Amino Acids in Yeast. *Methods Mol Biol* **2022**, *2394*, 377-432. DOI: 10.1007/978-1-0716-1811-0_21
(320) Stieglitz, J. T.; Kehoe, H. P.; Lei, M.; Van Deventer, J. A. A Robust and Quantitative Reporter System To Evaluate Noncanonical Amino Acid Incorporation in Yeast. *ACS Synth Biol* **2018**, *7* (9), 2256-2269. DOI: 10.1021/acssynbio.8b00260
(321) Islam, M.; Kehoe, H. P.; Lissoos, J. B.; Huang, M.; Ghadban, C. E.; Berumen Sanchez, G.; Lane, H. Z.; Van Deventer, J. A. Chemical Diversification of Simple Synthetic Antibodies. *ACS Chem Biol* **2021**, *16* (2), 344-359. DOI: 10.1021/acschembio.0c00865
(322) Hershman, R. L.; Rezhdo, A.; Stieglitz, J. T.; Van Deventer, J. A. Engineering Proteins Containing Noncanonical Amino Acids on the Yeast Surface. *Methods Mol Biol* **2022**, *2491*, 491-559. DOI: 10.1007/978-1-0716-2285-8_23
(323) Cenciarelli, O.; Riley, P. W.; Baka, A. Biosecurity Threat Posed by Botulinum Toxin. *Toxins (Basel)* **2019**, *11* (12). DOI: 10.3390/toxins11120681
(324) Roy, G.; Reier, J.; Garcia, A.; Martin, T.; Rice, M.; Wang, J.; Prophet, M.; Christie, R.; Dall'Acqua, W.; Ahuja, S.; et al. Development of a high yielding expression platform for the introduction of non-natural amino acids in protein sequences. *MAbs* **2020**, *12* (1), 1684749. DOI: 10.1080/19420862.2019.1684749
(325) Tian, F.; Lu, Y.; Manibusan, A.; Sellers, A.; Tran, H.; Sun, Y.; Phuong, T.; Barnett, R.; Hehli, B.; Song, F.; et al. A general approach to site-specific antibody drug conjugates. *Proc Natl Acad Sci U S A* **2014**, *111* (5), 1766-1771. DOI: 10.1073/pnas.1321237111
(326) Podracky, C. J.; An, C.; DeSousa, A.; Dorr, B. M.; Walsh, D. M.; Liu, D. R. Laboratory evolution of a sortase enzyme that modifies amyloid-beta protein. *Nat Chem Biol* **2021**, *17* (3), 317-325. DOI: 10.1038/s41589-020-00706-1
(327) Branon, T. C.; Bosch, J. A.; Sanchez, A. D.; Udeshi, N. D.; Svinkina, T.; Carr, S. A.; Feldman, J. L.; Perrimon, N.; Ting, A. Y. Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* **2018**, *36* (9), 880-887. DOI: 10.1038/nbt.4201
(328) Folger, I. B.; Frota, N. F.; Pistofidis, A.; Niquille, D. L.; Hansen, D. A.; Schmeing, T. M.; Hilvert, D. High-throughput reprogramming of an NRPS condensation domain. *Nat Chem Biol* **2024**. DOI: 10.1038/s41589-023-01532-x
(329) Cribari, M. A.; Unger, M. J.; Unarta, I. C.; Ogorek, A. N.; Huang, X.; Martell, J. D. Ultrahigh-Throughput Directed Evolution of Polymer-Degrading Enzymes Using Yeast Display. *J Am Chem Soc* **2023**, *145* (50), 27380-27389. DOI: 10.1021/jacs.3c08291
(330) Dieckhaus, H.; Brocidiacono, M.; Randolph, N. Z.; Kuhlman, B. Transfer learning to leverage larger datasets for improved prediction of protein stability changes. *Proc Natl Acad Sci U S A* **2024**, *121* (6), e2314853121. DOI: 10.1073/pnas.2314853121
(331) Carbonell, P.; Jervis, A. J.; Robinson, C. J.; Yan, C.; Dunstan, M.; Swainston, N.; Vinaixa, M.; Hollywood, K. A.; Currin, A.; Rattray, N. J. W.; et al. An automated Design-Build-Test-Learn pipeline for

enhanced microbial production of fine chemicals. *Commun Biol* **2018**, *1*, 66. DOI: 10.1038/s42003-018-0076-9
(332) HamediRad, M.; Chao, R.; Weisberg, S.; Lian, J.; Sinha, S.; Zhao, H. Towards a fully automated algorithm driven platform for biosystems design. *Nat Commun* **2019**, *10* (1), 5150. DOI: 10.1038/s41467-019-13189-z
(333) Rapp, J. T.; Bremer, B. J.; Romero, P. A. Self-driving laboratories to autonomously navigate the protein fitness landscape. *Nat Chem Eng* **2024**, *1* (1), 97-107. DOI: 10.1038/s44286-023-00002-4

For Table of Contents Use Only