ibe
Advancing Biology-Inspired Engineering
Official publication of the Institute of Biological Engineering

REVIEW   **Open Access**

Check for
updates

Non-canonical amino acid labeling in proteomics and biotechnology

Aya M. Saleh[1], Kristen M. Wilding[2], Sarah Calve[1], Bradley C. Bundy[2] and Tamara L. Kinzer-Ursem[1*]

Abstract

Metabolic labeling of proteins with non-canonical amino acids (ncAAs) provides unique bioorthogonal chemical groups during de novo synthesis by taking advantage of both endogenous and heterologous protein synthesis machineries. Labeled proteins can then be selectively conjugated to fluorophores, affinity reagents, peptides, polymers, nanoparticles or surfaces for a wide variety of downstream applications in proteomics and biotechnology. In this review, we focus on techniques in which proteins are residue- and site-specifically labeled with ncAAs containing bioorthogonal handles. These ncAA-labeled proteins are: readily enriched from cells and tissues for identification via mass spectrometry-based proteomic analysis; selectively purified for downstream biotechnology applications; or labeled with fluorophores for in situ analysis. To facilitate the wider use of these techniques, we provide decision trees to help guide the design of future experiments. It is expected that the use of ncAA labeling will continue to expand into new application areas where spatial and temporal analysis of proteome dynamics and engineering new chemistries and new function into proteins are desired.

Keywords: Metabolic labeling, Bioorthogonal chemistry, Residue-specific labeling, Site-specific labeling, Proteomics, Biotechnology

Overview of protein labeling with click chemistry functionality

Methods that allow for labeling of proteins co-translationally, i.e. as they are being synthesized, have wide ranging applications in engineering, biotechnology, and medicine. Incorporation of non-canonical amino acids (ncAAs) into proteins enables unique bioorthogonal chemistries, those that do not react with naturally occurring chemical functional groups, for conjugation. These conjugate substrates range from fluorophores, affinity reagents, and polymers to nanoparticle surfaces, enabling new advances in technology to study cellular systems and produce novel biocatalytic and therapeutic proteins. A key benefit of these techniques is the ability to enrich for labeled proteins of interest, whereas other labeling methods add or remove a mass (e.g. isotope labeling [1]) that can be difficult to identify when diluted within complex macromolecular mixtures. In this review, we focus specifically on techniques that incorporate click chemistry

functionality into proteins of interest and provide decision tree analyses to guide selection of optimal strategies for protein labeling methods.

Click chemistry functionality

First coined by Sharpless and colleagues in 2001, click chemistries are a set of chemical reactions that are readily catalyzed in aqueous solutions at atmospheric pressure and biologically-compatible temperatures, with few toxic intermediates, and relatively fast reaction kinetics [2]. The suite of specific click chemistry reactions that started with Staudinger ligation of azide and phosphine [3–5] and copper-catalyzed azide-alkyne cycloaddition [6, 7], has rapidly expanded to include more rapid and biologically friendly chemistries including strain promoted azide-alkyne cycloaddition [8, 9], oxime or hydrazine ligation [10, 11], strain-promoted alkyne nitrone cycloaddition [12, 13], tetrazine ligation [14, 15], and quadricyclane ligation [16, 17].

Here, we focus on azide-alkyne cycloaddition as it is one of the most widely used, with broad availability of commercial reagents, moderately fast kinetics, and well-established protocols. Copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC, Fig. 1a) has been implemented across disciplines,

* Correspondence: tursem@purdue.edu
[1]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA
Full list of author information is available at the end of the article

from biomaterials [18] and combinatorial chemistry [19] to polymer synthesis [20], protein activity tagging [21], and proteomics [22], some of which will be highlighted in later sections. One disadvantage of CuAAC is that there is significant cytotoxicity with using copper as the catalyst, hampering utilization in vivo [23]. To circumvent this limitation, Bertozzi and coworkers introduced a catalyst-free [3 + 2] cycloaddition reaction between azides and cyclooctyne derivatives, known as strain promoted azide-alkyne cycloaddition (SPAAC, Fig. 1b) [8, 23, 24]. The biocompatibility of this reaction was first demonstrated in Jurkat cells to label azide-tagged glycoproteins [8]. The strain-promoted azide-alkyne click reaction has since been applied in various in vivo settings with no apparent toxicity [24–27]. Importantly, CuAAC and SPAAC are bioorthogonal and will not interfere with natural biological chemistries.

Labeling of nascent proteins
Chemical biologists and bioengineers have found much utility in incorporating click chemistry functionality into nature's translational machinery. In these methods, known as genetic code expansion or ncAA labeling [28–31], a ncAA carrying a desired click chemistry functional group is introduced to the host expression system and is incorporated onto an aminoacyl tRNA synthetase (aaRS) that covalently attaches the ncAA to the corresponding tRNA (Fig. 2a). The ncAA-tRNA complex is brought into the ribosome where the tRNA recognizes the appropriate mRNA codon sequence and the ncAA is added to the growing polypeptide chain (Fig. 2b). ncAA labeling can be designed to occur either at specific amino acid residues of interest, for example using a Methionine (Met) analog that carries an azide or alkyne functionality to replace any Met in a newly synthesized protein [3], or at specific sites in a protein of interest [32].
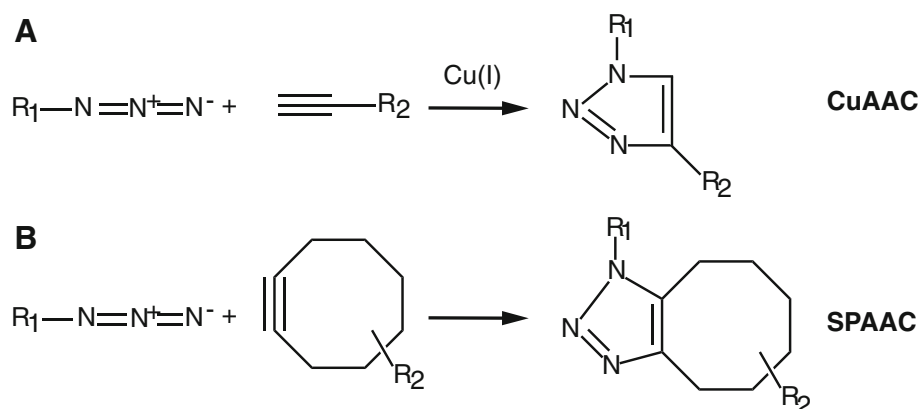
Though not the focus of this review, it is important to highlight other site-specific approaches for labeling proteins.

These include leveraging of enzymatic post-translational modification of proteins with click-chemistry functionalized non-canonical fatty acids, nucleic acids, and sugars. These methods utilize so called 'chemoenzymatic methods' to label proteins at specific residues via enzymatic recognition of specific peptide sequences. In this way, endogenous, engineered, and recombinantly expressed proteins can be efficiently labeled in situ. Some examples include glycosylation [33–35], sortagging [36, 37] and fatty acylation [38–41], including prenylation [10, 42], palmitoylation [43, 44], and myristoylation [45–49].
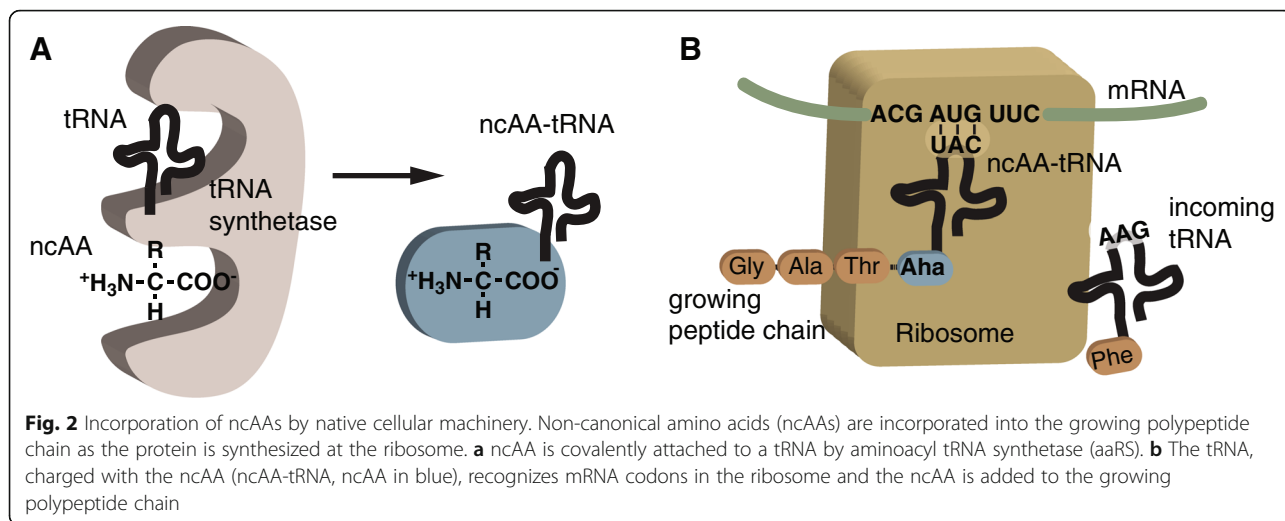
Residue-specific labeling of nascent proteins with non-canonical amino acids
First demonstrated by Tirrell and colleagues, native translational machinery in *E. coli* was found to readily incorporate noncanonical Met analogs into proteins in vivo [50–52]. In this way, alkene (homoallylglycine, Hag) and alkyne (homopropargylglycine, Hpg) side-chain functionalities were added at Met sites during protein biosynthesis (Fig. 3, and Table 1). Later, azide analogs of Met (e.g. Aha, Fig. 3) were also found to be readily incorporated in vivo [3].

These methods take advantage of the ability for some ncAAs to incorporate (or become charged) onto native aaRSs (Fig. 2a), covalently attach to the corresponding tRNA, and subsequently incorporate into growing polypeptide chains (Fig. 2b). The kinetics of Aha and Hpg binding to the methionyl tRNA synthetase (MetRS) are slower than that of Met (k_{cat}/K_m of 1.42×10^{-3} and $1.16 \times 10^{-3} \, s^{-1} \cdot \mu M^{-1}$ for Aha and Hpg, respectively vs $5.47 \times 10^{-1} \, s^{-1} \cdot \mu M^{-1}$ for Met) [3]. Nonetheless, this is a straightforward labeling method with no need for genetic engineering of the protein or organism under study (Fig. 4). For applications where 100% Met substitution is not necessary (e.g. enrichment for proteomics), adding the ncAA at concentrations where it can outcompete with Met provides sufficient functional incorporation. Alternatives

Fig. 1 Azide-alkyne cycloaddition reactions. **a** Copper(I)-catalyzed [3 + 2] azide-alkyne cycloaddition (CuAAC). **b** [3 + 2] cycloaddition of azides and strain-promoted alkynes (cyclooctynes) (SPAAC)

Fig. 2 Incorporation of ncAAs by native cellular machinery. Non-canonical amino acids (ncAAs) are incorporated into the growing polypeptide chain as the protein is synthesized at the ribosome. **a** ncAA is covalently attached to a tRNA by aminoacyl tRNA synthetase (aaRS). **b** The tRNA, charged with the ncAA (ncAA-tRNA, ncAA in blue), recognizes mRNA codons in the ribosome and the ncAA is added to the growing polypeptide chain
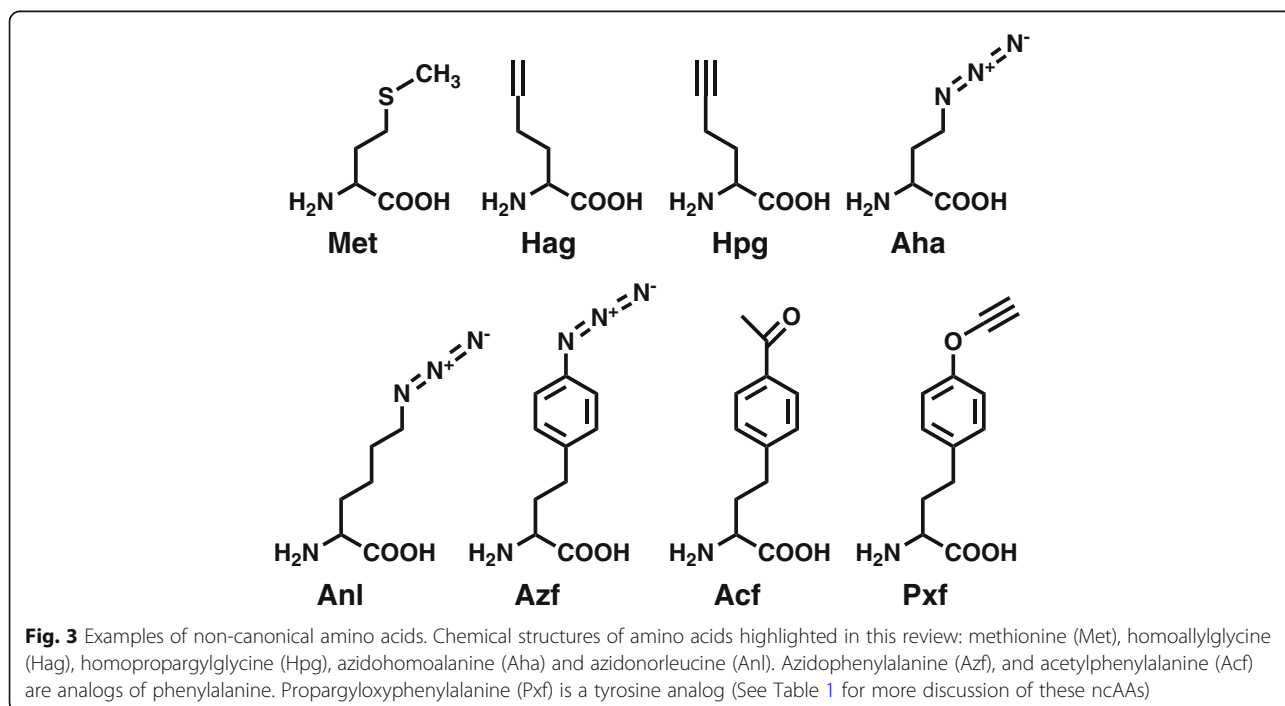
that increase ncAA incorporation include using Met auxotrophic strains of *E. coli* that cannot make their own Met [52], or Met-free media for mammalian cell culture. Orthogonal aaRSs have also been engineered to bind to ncAAs in cells expressing the mutant aaRS, allowing for protein labeling with ncAAs in specific cell types [53–57].

Site-specific labeling of proteins with non-canonical amino acids

An alternative to residue-specific ncAA incorporation is site-specific ncAA incorporation, in which a ncAA is incorporated exclusively at a pre-determined site.

Motivated by the implications for detailed studying of protein structure and function, Schultz and colleagues were one of the first to demonstrate the feasibility of site-specific incorporation of ncAAs into a full-length protein in 1989 [32]. To accomplish this, the anticodon of suppressor tRNA molecules was engineered to recognize the amber stop codon (UAG), chemically aminoacylated with the ncAA, and then added to an in vitro protein synthesis system. Later, Furter site-specifically incorporated ncAAs in vivo by using an engineered orthogonal tRNA/tRNA synthetase pair for amber suppression. As illustrated in Fig. 5, the tRNA/tRNA

Fig. 3 Examples of non-canonical amino acids. Chemical structures of amino acids highlighted in this review: methionine (Met), homoallylglycine (Hag), homopropargylglycine (Hpg), azidohomoalanine (Aha) and azidonorleucine (Anl). Azidophenylalanine (Azf), and acetylphenylalanine (Acf) are analogs of phenylalanine. Propargyloxyphenylalanine (Pxf) is a tyrosine analog (See Table 1 for more discussion of these ncAAs)

Table 1 List of ncAAs discussed in the review and their methods of incorporation

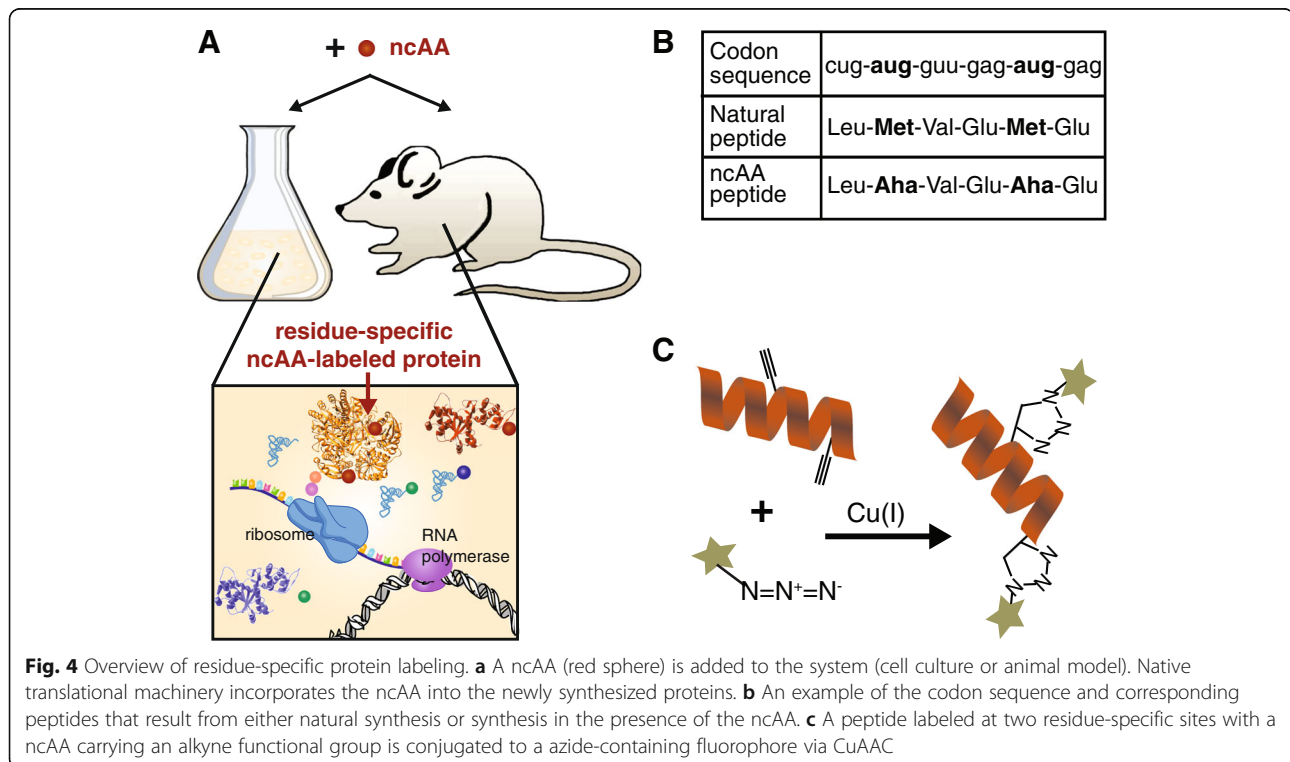| ncAA	Codon replacement	Natural amino acid analog	Genetic modification	Labeling method
Hag	AUG	Met	Not needed[a]	Residue-specific [50–52]
Hpg	AUG	Met	Not needed[a]	Residue-specific [22, 50, 51, 73]
Aha	AUG	Met	Not needed[a]	Residue-specific [3, 22, 75, 84–87]
Anl	AUG	Met	Transgenic lines that express a MetRS mutant capable of charging Anl[b]	Residue-specific [53, 54, 78, 88–90, 92–95]
Azf	UUC, UUU UAG	Phe	Transgenic lines that express a PheRS mutant capable of charging Azf[b] Transgenic lines that express an orthogonal aaRS/amber suppressor tRNA pair evolved for Azf specificity	Residue-specific [55] Site-specific [118]
Acf	UUC, UUU UAG	Phe	Transgenic lines that express a PheRS mutant capable of charging Acf[b] Transgenic lines that express an orthogonal aaRS/amber suppressor tRNA pair evolved for Acf specificity	Residue-specific [136] Site-specific [111]
Pxf	UAG	Tyr	Transgenic lines that express an orthogonal aaRS/amber suppressor tRNA pair evolved for Pxf specificity	Site-specific [124]

[a]The efficiency of the ncAA incorporation is greatly enhanced by using Met auxotrophic strains
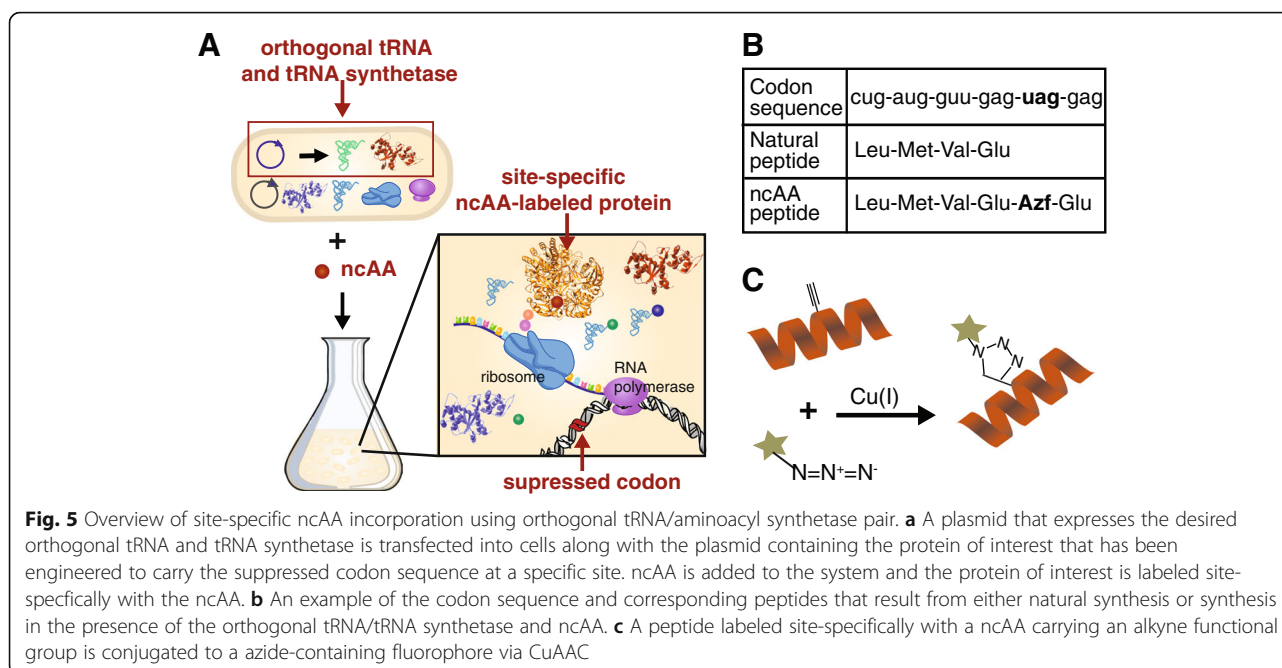[b]For cell selective labeling, the mutant aaRS is expressed under the control of cell-specific promoters

synthetase pair is exogenous and operates orthogonally and the tRNA is specific for UAG instead of AUG [58]. Since then, over 100 different ncAAs have been incorporated either in vivo or in vitro in a variety of systems including bacteria, yeast, plant, mammalian, and human cells [59, 60]. The methods for site-specific ncAA incorporation have also expanded beyond amber codon

suppression to include suppression of additional stop codons (nonsense suppression) [61, 62], recoding of sense codons [63], and recognition of 4-base codons (frameshift suppression) [62, 64, 65], though amber suppression is still the most widely used method.

As described above, initial ncAA incorporation was performed using chemically aminoacylated tRNA and an

Fig. 4 Overview of residue-specific protein labeling. **a** A ncAA (red sphere) is added to the system (cell culture or animal model). Native translational machinery incorporates the ncAA into the newly synthesized proteins. **b** An example of the codon sequence and corresponding peptides that result from either natural synthesis or synthesis in the presence of the ncAA. **c** A peptide labeled at two residue-specific sites with a ncAA carrying an alkyne functional group is conjugated to a azide-containing fluorophore via CuAAC

Fig. 5 Overview of site-specific ncAA incorporation using orthogonal tRNA/aminoacyl synthetase pair. **a** A plasmid that expresses the desired orthogonal tRNA and tRNA synthetase is transfected into cells along with the plasmid containing the protein of interest that has been engineered to carry the suppressed codon sequence at a specific site. ncAA is added to the system and the protein of interest is labeled site-specfically with the ncAA. **b** An example of the codon sequence and corresponding peptides that result from either natural synthesis or synthesis in the presence of the orthogonal tRNA/tRNA synthetase and ncAA. **c** A peptide labeled site-specifically with a ncAA carrying an alkyne functional group is conjugated to a azide-containing fluorophore via CuAAC

in vitro protein synthesis system [32, 65]. This method circumvents the need for evolving aaRSs to charge suppressor tRNA, and enables incorporation of virtually any ncAAs, including very large ncAAs such as those pre-conjugated to polyethylene glycol [64, 66]. Although chemically aminoacylated tRNA is still used for small-scale applications, it is not economically scalable for large-scale biotechnology applications, which must instead rely on enzymatic aminoacylation.

For large-scale applications, an orthogonal tRNA is engineered to recognize the specific codon sequence, and an orthogonal aaRS charges the engineered tRNA with the desired ncAA to enable continuous tRNA aminoacylation throughout protein expression (Fig. 5) [67]. The amber stop codon, UAG, is used less frequently by organisms than the other stop codons and is commonly targeted as the repurposed codon [68], though the other stop codons have also been successfully utilized [61, 62]. Frameshift suppression is executed similarly, by targeting a quadruplet codon [65]; however, suppression efficiencies are reportedly lower than nonsense suppression [62, 69]. By employing a combination of suppression techniques, multiple ncAAs can be site-specifically incorporated simultaneously [61, 62, 64, 69, 70]. In these cases, the suppression machinery must be mutually orthogonal in order to maintain site-specificity.

Overall, the site-specific approach provides significantly more control over the exact pre-defined location of ncAA incorporation into the protein as compared to other methods [71]. It also facilitates very high ncAA incorporation efficiencies [67]. As such, it is a powerful tool for biotechnology applications and will be detailed later in the paper. Potential uses of this technique for proteomics applications are still being developed and are briefly highlighted at the end of the following section.

Applications of ncAA labeling
Proteomics
Residue-specific labeling for proteomic applications
Residue-specific methods have since been applied to identify de novo protein synthesis in a variety of contexts. Dieterich et al. introduced the bioorthogonal non-canonical amino acid tagging (BONCAT) strategy for selective analysis of de novo protein synthesis with high temporal resolution [22, 72]. In this method, cells are cultured in media supplemented with Met analogs like Hpg or Aha, that are tagged with alkyne or azide functional groups, respectively (Fig. 4). Since azides and alkynes are bioorthogonal moieties, Hpg- and Aha-labeled proteins can be selectively conjugated to affinity tags even within complex cellular or tissue lysates to enrich the newly synthesized proteins from the pool of pre-existing unlabeled proteins. Additionally, labeled proteins can be ligated to fluorescent dyes for protein visualization using a sister technique referred to as fluorescent non-canonical amino acid tagging (FUNCAT) [25, 73].

Over the last decade, BONCAT has gained wide recognition because of its capability of tracking continuous changes in protein expression. It has been applied in mammalian cell cultures to study protein acylation [74], lysosomal protein degradation [75], and inflammation [76]. The method has also been used in various bacterial systems in order to explore quorum sensing [77], identify

virulence factors [78], and monitor bacterial degradation in phagocytes [79]. Moreover, BONCAT has proven effective in more complicated biological systems such as zebrafish [80], *Caenorhabditis elegans* [55, 81], and Xenopus [82].

Until recently, it was presumed that BONCAT cannot be applied to in vivo labeling of the rodent proteome because mammalian cells would favor incorporating endogenous Met, rather than an analog, into newly expressed proteins [83]. However, Schiapparelli et al. successfully labeled newly synthesized proteins in the retina of adult rats by intraocular injection of Aha [84]. Further, McClatchy et al. showed that in vivo labeling of the entire murine proteome is feasible by feeding animals an Aha-enriched diet for 4 to 6 days [85, 86]. More recently, Calve and Kinzer-Ursem demonstrated that two days of intraperitoneal injection of Aha and Hpg is sufficient for systemic incorporation of the Met analogs into the proteome of both juvenile mice and developing embryos [87]. In this study, neither perturbation of physiological function of the injected mice nor atypical embryonic development was observed. In addition, both Aha and Hpg were successfully incorporated into different murine tissues in a concentration-dependent manner [87]. Notably, labeling with Hpg was less efficient than Aha, which is in agreement with findings by Kiick et al. that the activation rate of Hpg by MetRS is slower than Aha [3]. In light of these results, the successful incorporation of Met analogs into the entire murine proteome through intraperitoneal injection will pave the road for using animal models to temporally map protein expression. This method provides several advantages over introducing ncAAs via diet because intraperitoneal injection is relatively easy to perform, global proteome labeling is achieved in a shorter time period, and injection warrants more accurate dose-effect calculations.

With the aim of probing proteomic changes in specific cell types, engineered aaRS technology was adopted to allow cell-selective labeling with ncAAs. This technique, pioneered by the Tirrell group, was first made possible by identifying *E. coli* MetRS mutants that can charge the Met analog azidonorleucine (Anl) to Met sites [88]. Anl is not a substrate to endogenous aaRSs, hence only cells bearing the mutant MetRS are labeled. Since its discovery, the mutant MetRS Anl labeling technique has been applied to label nascent proteomes of *E. coli* [51, 57, 89, 90], *Salmonella typhimurium* [91], *Yersinia enterocolitica* and *Yersinia pestis* strains [78], and *Toxoplasma gondii* [92] in infected host cells. The exclusive expression of the mutant MetRS in these pathogens allowed selective detection of pathogen proteins among the more abundant host proteins.

To further demonstrate the utility of this approach, other variants of aaRS have been evolved to enable cell-selective incorporation of ncAAs in mammalian cells and animals. Using *Caenorhabditis elegans* as a model organism, Yuet et al. employed a phenylalanyltRNA synthetase mutant capable of incorporating the ncAA azidophenylalanine (Azf, Fig. 3) into worm proteins [55]. In their study, cell-type-specific resolution was achieved by expressing the mutant synthetase in targeted cells under the control of cell-specific promoters. Similarly, Erdmann et al. demonstrated that cell selectivity in *Drosophila melanogaster* can be achieved by using murine (mMetRS) and drosophila MetRS (dMetRS) mutants that can activate Anl [93]. The dMetRS variant was further utilized by Niehues et al. to study protein synthesis rates in a drosophila model of Charcot–Marie–Tooth neuropathy [94], while the mMetRS variant was applied to selectively tag astrocyte proteins in a mixed culture system [95] and to label the nascent proteome of several mammalian cells [54].

More recently, Schuman and coworkers advanced the MetRS mutant technology to enable cell-selective labeling in live mammals for the first time [53]. In this seminal work, selective labeling and identification of nascent proteomes in hippocampal excitatory and cerebellar inhibitory neurons was achieved using a transgenic mouse line wherein a MetRS mutant is expressed under the control of Cre recombinase. The Conboy group expanded this technique to identify the "young" proteins that were transferred to old mice in a model of heterochronic parabiosis [96]. This application was further leveraged by the Conboy and Aran groups by designing a graphene-based biosensor capable of selective capturing and quantifying of azide-labeled blood proteins that traveled from young to old parabiotic pairings [97], signifying the potential utility of cell-selective technology in the field of diagnosis and biomarker discovery.

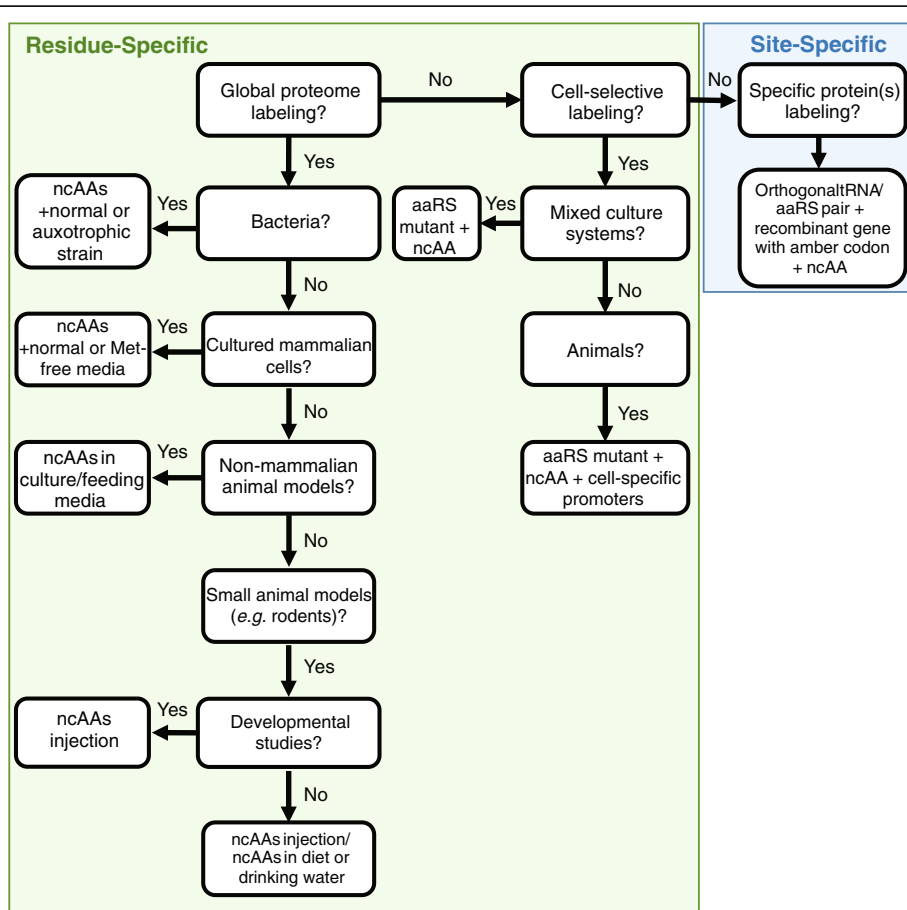### Site-specific labeling for proteomic applications

While residue-specific ncAA labeling has been primarily used for proteomic applications due to the ease of use and incorporation throughout the proteome, site-specific labeling has the potential to also assist in this area [53, 98]. For example, ncAAs could be used to label and trace a specific protein as it is expressed, migrates, and travels within a cell or tissue. In addition, ncAAs could be combined with proteomics to track a specific protein that is available at low levels. A factor that has limited site-specific ncAA use in proteomics is that this area of research has been focused on single-celled organisms, whereas proteomic studies are commonly performed in multicellular organisms. Recently, site-specific ncAA has been expanded to the multicellular organisms *Caenorhabditis elegans* and *Drosophila melanogaster* [99, 100], holding promise for the implementation into additional multicellular organisms. In the meantime, residue-specific labeling will continue to be the predominant approach when using ncAAs for most proteomics applications. With the increasing variety of approaches to ncAA

incorporation, it is important to identify which approaches are best suited to a given application. To help guide researchers find the optimal strategy to label proteins for a given proteomics application, a decision tree diagram is provided in Fig. 6.

Biotechnology applications

Traditional approaches to bioconjugation in biotechnology often target reactive side chains of natural amino acids such as lysines, though this results in a complex mixture of products modified at different locations and to different extents, complicating protein separation and

often reducing protein activity. For some applications, sufficient control is afforded by altering the pH of a conjugation reaction to enhance the reactivity of the N-terminal amino group [101, 102]. An advantage of this method is circumvention of protein mutation, however, restricting bioconjugation to the N-terminus limits the potential for conjugation site optimization and can be deleterious to the structure and function, as was found to be the case with parathyroid hormone [101]. Surface-exposed cysteines, either native or substituted into the proteins, may also be targeted for modification, as they are more limited than other reactive amino acids

Fig. 6 Decision tree for ncAA labeling in proteomics applications. If global proteome labeling is desired, consider residue-specific labeling. Residue-specific ncAA labeling is designed to replace a specific natural amino acid of interest in the entire proteome. Several natural amino acids analogs have been utilized (See Fig. 3 and Table 1). No genetic modification is needed for global proteome labeling with ncAA. Nevertheless, the labeling efficiency in bacterial cells is greatly enhanced if auxotrophic mutants are used. Similarly, labeling of cultured mammalian cells and non-mammalian animal models (e.g. nematodes) can be achieved by adding the ncAA directly to the culture/feeding media. However, if higher degree of labeling is required, consider using culture media that lacks the natural amino acid to be replaced. For in vivo labeling of small animal models (e.g. rodents), the ncAA can be injected or added to animal diet and/or drinking water. If embryonic labeling is desired, consider ncAA injection since it has been demonstrated that ncAAs are effectively incorporated into embryos when injected into pregnant animals without disturbing normal development [87]. If labeling of specific cell types in a mixed culture system is desired, consider using transgenic lines that express a mutant aaRS designed to charge the ncAA of interest. Since the ncAA is not a substrate of endogenous aaRSs, only cells expressing the mutant aaRS in the mixed culture system are labeled. Similarly, if cell-selective labeling of animals is required, consider use transgenic animals that express the mutant aaRS under cell-specific promoters. If specific protein labeling rather than global proteome labeling is needed, ncAA can be incorporated site-specifically in the polypeptide chain in response to an amber stop codon. This requires introducing the amber codon into the gene of interest and using an orthogonal aaRS/amber suppressor tRNA pair evolved for charging the desired ncAA

such as lysine [103]. However, successful application of these methods is limited by the inherent properties of the target protein – in some proteins, the N-terminus may be inaccessible or involved in protein function, and engineering cysteine sites into proteins with natural cysteines may interfere with native disulfide bond formation.

As uniquely reactive chemical moieties, ncAAs provide a tool for enhancing commercial and therapeutic applications of proteins in a bioorthogonal manner. ncAAs have been used to study protein stability and also generate proteins with improved stability [104]. Characterization of protein structure and conformation, properties essential for effective rational enzyme and drug design, can also be improved through FRET analysis following conjugation of fluorophores to incorporated ncAAs [63, 105, 106]. Click chemistry compatible ncAAs are also attractive methods for covalent protein bioconjugation, which has bearings on biocatalysis [104, 107], biochemical synthesis [108–110], therapeutic optimization [111, 112], and vaccine design [113, 114]. For example, enzyme immobilization is an established method for stabilization of proteins that allows recoverability of enzymes in biocatalysis [115–117], and has been demonstrated to improve the efficiency of enzymatic cascades by improving pathway flux [108–110]. Immobilization of such enzymes using ncAA can provide greater control over orientation, which is important for maintaining the activity of many enzymes. Similarly, polymer-protein conjugation is a well-established method for stabilizing therapeutic proteins against thermal or pH stress, proteolytic attack, and improving pharmacokinetic profiles [118–120], but is often accompanied by marked decreases in specific activity associated with imprecise control over location and extent of modification. These conjugates can be enhanced by the greater control and specificity of conjugation afforded by ncAA incorporation and targeting [111, 112, 118]. Finally, virus-like particles (VLPs) have emerged as promising candidates for safe, effective vaccines as well as functionalizable nanoparticles for drug delivery [121, 122]. The surface of these proteinaceous nanoparticles can be "decorated" with a variety of antigens or polymers to improve the generation of adequate immune response to presented antigens or mask immunogenicity of the VLP particle [71, 121]. ncAAs provide bioorthogonal conjugation targets to maintain the integrity of both the VLP and displayed antigens [114, 121].

Residue-specific labeling for biotechnology applications
In some cases, residue-specific labeling provides adequate control of conjugation site to maintain sufficient protein activity. For example, Met replacement was used to functionalize a VLP which contained only one Met in each capsid monomer [114]. For cases such as these, in which there are small numbers of accessible residues of a certain type, residue-specific labeling may be sufficient. For proteins in which the N-terminal methionine (fMet) is accessible, a mixture of products may still result due to ncAA incorporation at fMet. Additionally, for applications in which a mixture of conjugation sites in the product is acceptable, residue-specific ncAA incorporation offers a simplistic approach circumventing identification of necessary tRNA synthetases. A disadvantage of this approach, however, is that when multiple instances of a replaced residue are surface-accessible, targeting of the ncAA can still result in a mixture of products modified at different locations and to different extents, similar to that seen with targeting of natural amino acids such as lysine [101]. This limitation is particularly important in development of conjugated proteins for medicinal applications, where consistency in product specifications and performance is key.

Site-specific labeling for biotechnology applications
In many applications, including both the study of protein function and the design of enhanced proteins, it is desirable to incorporate the ncAA precisely at a pre-determined site. For example, conjugation site has been shown to have a significant effect on the stability and activity of antibody-drug conjugates [123], polymer-protein conjugates [111, 112, 118], and immobilized proteins [124]. Site-specific ncAA incorporation enables precise control of conjugation site to allow optimization as well as production of homogenous protein conjugates. This homogeneity is especially important for therapeutic applications such as antibody-drug conjugates and polymer-conjugated therapeutics where precise characterization is necessary [70, 111, 112, 123, 125, 126]. Therefore, protein conjugation for biotechnology applications must often be done in a site-specific manner to optimize conjugate homogeneity, activity, and protein stability. For example, using the ncAA acetylphenylalanine (Acf, Fig. 3), polyethylene glycol conjugation (PEGylation) of human growth hormone (hGH) was optimized for conjugation site, enabling mono-PEGylation and development of an active PEG-hGH with increased serum half-life [111]. Notably, Cho and coworkers reported as much as a 3.8-fold increase in the C_{max} of the optimally PEGylated hGH compared to hGH PEGylated at other sites, demonstrating the importance of site optimization and precise conjugation site targeting for pharmacokinetic properties [111].

In biocatalysis and enzyme production, site-specific incorporation of ncAA can be instrumental in preparation of robust, reusable proteins to improve industrial applicability. Deepankumar and coworkers immobilized transaminase to a chitosan substrate site-specifically to produce an immobilized enzyme which facilitated simple purification and maintained a specific activity nearly equal to that of the wild-type enzyme [104]. The enhanced

Saleh *et al. Journal of Biological Engineering*   (2019) 13:43
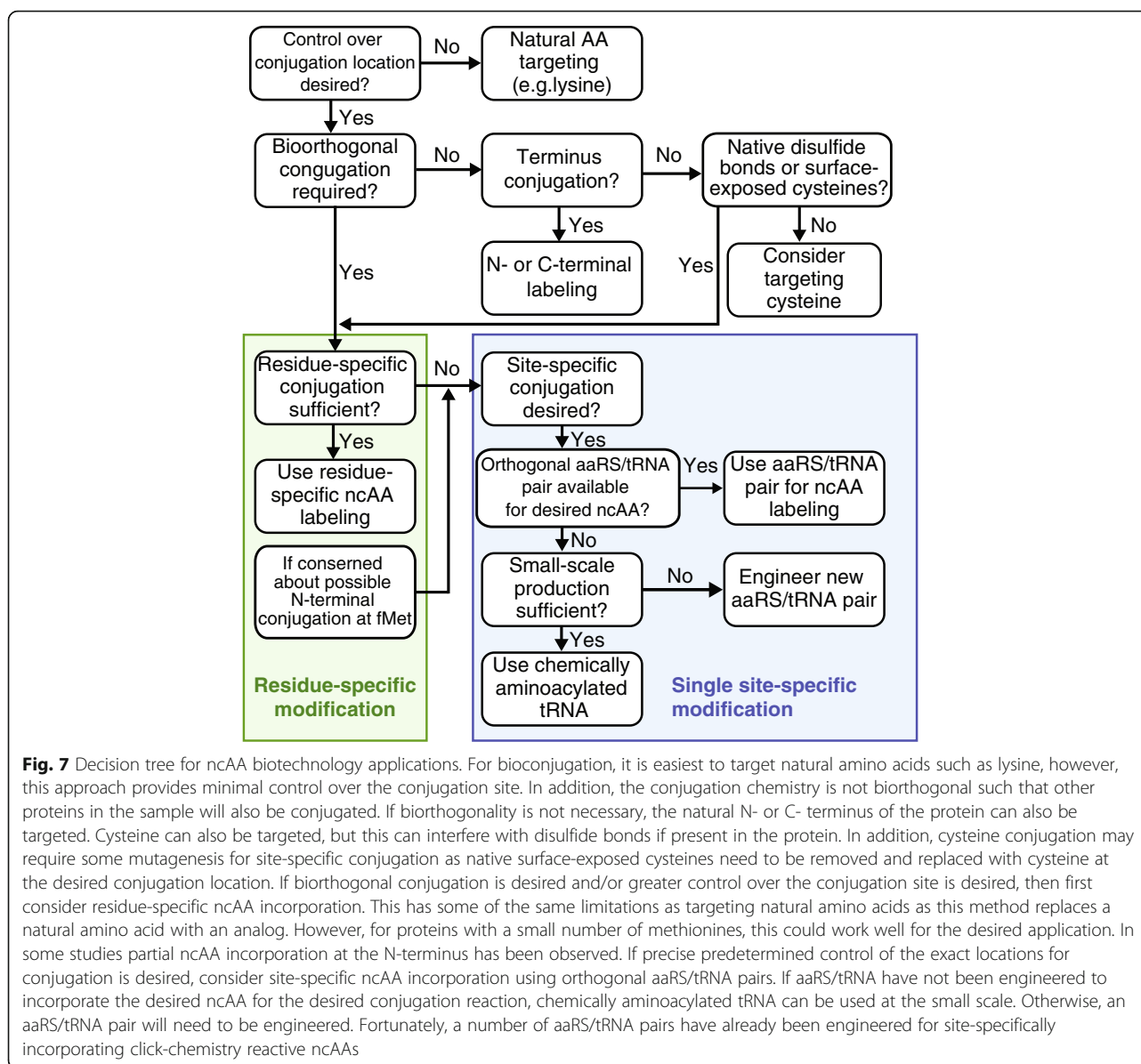
Page 9 of 14

potential for optimization of conjugated enzymes is further demonstrated in a study by Mu and coworkers, in which monoPEGylated derivatives of fibroblast growth factor 21 (FGF21) were prepared via site-specific incorporation of Acf (Fig. 3). This study identified multiple PEGylated derivatives of FGF21, including one where the substituted residue was originally a leucine, which maintained high activity and 15–30-fold increased half-lives [112]. By contrast, another leucine substitution in the same protein resulted in a conjugate which was entirely inactive, highlighting the necessity for site-specific versus residue-specific modifications in maintaining activity of some proteins [112]. These studies emphasize the importance of precise control over conjugation site selection for optimal design and production of biotechnology products such as therapeutic protein conjugates and biocatalysts.

Site-specific ncAA incorporation also allows for close control over the number of sites which are modified by conjugation, which is an important aspect of conjugate optimization. For example, Wilding and coworkers recently demonstrated that dual PEGylation of T4 lysozyme at two site-specifically incorporated Azf (Fig. 3) residues decreased the double Azf-incorporated T4 lysozyme variant's activity and did not increase its stability, despite increases in the stability and activity corresponding to PEGylation of each site individually [118]. Similarly, close control over conjugation extent for antibody-drug conjugates is necessary to ensure drug homogeneity and enhance therapeutic index [126, 127]. Motivated by this capacity to improve antibody-drug conjugates through close control of drug-antibody ratios (DAR), Zimmerman and coworkers engineered a high fidelity tRNA/aaRS pair to incorporate the highly click-reactive ncAA azido-methyl-phenylalanine (AMF) site-specifically into a Trastuzumab antibody fragment [126]. The researchers demonstrated drug-to-antibody ratios ranging from 1.2 to 1.9 depending on the AMF incorporation site, and potent cytotoxic activity which correlated with the DAR of each variant tested [126]. Recently, Oller-Salvia and coworkers further demonstrated the ability to closely control DAR by using site-specific incorporation of a cyclopropane derivative of lysine to achieve drug-conjugated Trastuzumab with a DAR of > 1.9, indicating high conjugation efficiency of the two ncAA sites within the fragment [127]. Together, these studies illustrate the utility of site-specific ncAA incorporation in biotechnology towards the production of optimized, controlled, and well-characterized conjugates for medicinal and biocatalytic applications.

Given the varied, site-dependent effects of ncAA incorporation and conjugation, a major challenge with ncAA incorporation is understanding and predicting the impact that the mutation will have on the protein. However, recent progress has demonstrated the potential for molecular simulations to inform site selection [118, 124, 128]. For example, simulations unexpectedly predicted a 3% solvent accessible location as being highly stabilizing to the protein if covalently immobilized at this site [128]. Common design heuristics would prevent this site from ever being considered; however, using the ncAA propargyloxyphenylalanine (Pxf, Fig. 3), this site was shown to be better than highly surface accessible sites [124]. Using the same protein, simulation screening was also effective in predicting optimal specific sites for PEGylation, which were different than those predicted for immobilization [118]. The predictions were validated with high correlation using copper-free click-chemistry reactive ncAA Azf (Fig. 3) [118]. Due to these recent successes using molecular simulation, it is anticipated that rapid simulation approaches will increasingly assist in determining the best locations for ncAA incorporation for both the bioconjugation application and for reducing or eliminating structural strain due to the ncAA mutation. As tools for ncAA incorporation continue to increase in efficiency and simplicity and costs continue to decrease, it is anticipated that ncAAs will become not only a research tool for bioconjugation optimization, but also an industrially viable therapeutics and biocatalysts production platform.

With the increasing variety of approaches to site-specific ncAA incorporation, it is important to identify which approaches are best suited to a given application. Figure 7 provides a decision tree to aid in tool selection based on the needs of a specific application. If bioorthogonal conjugation is not necessary, conjugation at the C-terminus, with cysteine or with other natural amino acids such as lysine, could be considered. However, significant mutagenesis may be necessary to enable site-specific conjugation. In contrast, ncAAs provide bioorthogonal conjugation and facilitate control over the conjugation location with minimal mutagenesis. For proteins in which there are a limited number of surface-accessible instances of a residue such as Met, residue-specific ncAA labeling may be the most effective as it can be done without orthogonal translation machinery. Nevertheless, potential incorporation of a ncAA at fMet must be considered and a site-specific approach should be taken if fMet labeling is a concern. For any site-specific application, orthogonal aaRS/tRNA pairs enable straightforward implementation of nonsense and frameshift suppression, especially for in vivo protein synthesis applications, and are ideal when available. When an aaRS has not been engineered for the desired ncAA, chemically aminoacylated tRNA may be used. However, for large-scale applications, the higher cost of this approach motivates engineering of an orthogonal aaRS/tRNA pair. Finally, as will be discussed in the future directions section, a cell-free protein synthesis approach

Fig. 7 Decision tree for ncAA biotechnology applications. For bioconjugation, it is easiest to target natural amino acids such as lysine, however, this approach provides minimal control over the conjugation site. In addition, the conjugation chemistry is not biorthogonal such that other proteins in the sample will also be conjugated. If biorthogonality is not necessary, the natural N- or C- terminus of the protein can also be targeted. Cysteine can also be targeted, but this can interfere with disulfide bonds if present in the protein. In addition, cysteine conjugation may require some mutagenesis for site-specific conjugation as native surface-exposed cysteines need to be removed and replaced with cysteine at the desired conjugation location. If biorthogonal conjugation is desired and/or greater control over the conjugation site is desired, then first consider residue-specific ncAA incorporation. This has some of the same limitations as targeting natural amino acids as this method replaces a natural amino acid with an analog. However, for proteins with a small number of methionines, this could work well for the desired application. In some studies partial ncAA incorporation at the N-terminus has been observed. If precise predetermined control of the exact locations for conjugation is desired, consider site-specific ncAA incorporation using orthogonal aaRS/tRNA pairs. If aaRS/tRNA have not been engineered to incorporate the desired ncAA for the desired conjugation reaction, chemically aminoacylated tRNA can be used at the small scale. Otherwise, an aaRS/tRNA pair will need to be engineered. Fortunately, a number of aaRS/tRNA pairs have already been engineered for site-specifically incorporating click-chemistry reactive ncAAs

should be considered in cases where high-throughput evaluation or on-demand production of conjugates is necessary.

Future directions

To expand the potential of ncAA labeling for research and industrial applications, additional studies are necessary to address key limitations in the efficiency of ncAA incorporation and optimal modification site selection. It is generally recognized that one limitation of residue-specific ncAA labeling is that it commonly requires prior depletion of a natural amino acid to achieve high proteome labeling. This practice can disturb normal biological functions and hence adapting methods that enable high levels of ncAA incorporation in the presence of the

canonical amino acid is an important advancement for applications in higher-order organisms [53, 55, 85–87].

Current challenges in obtaining the highest quality proteomic mapping lie in the optimization of the click chemistry reactions and enrichment protocols. Therefore, continued discovery of new click chemistries with faster kinetics and higher specificity will increase the potential for ncAAs in proteomics applications. In addition, development of techniques that allow cell- and tissue-specific labeling in mammalian systems with lower non-specific labeling and background noise will have a significant impact on resolving cellular proteomics maps with high resolution. This, combined with advances in engineering aaRS mutants that enable charging ncAAs at higher rates and promoters that can drive the expression of the mutant

synthetase with high cell specificity, will enhance our understanding of the spatial and temporal aspects of proteome dynamics.

A major hurdle for biotechnology applications where stoichiometric labeling is desired is that ncAA incorporation efficiency for site-specific protein modification often varies by the incorporation site. Elucidating factors determining site-dependence will enable the more effective design of ncAA-modified proteins, for example, by targeting bases that flank suppressed codons [129]. Additionally, investigation of mechanisms involved in ribosome stoppage, where polypeptide synthesis stalls or terminates prematurely, may also provide illumination towards efficient modification site selection. Development of novel cell strains lacking factors inhibitory to ncAA incorporation may also improve labeling efficiency. Such strains have already been developed in *E. coli* by knocking out release factor components responsible for competition with nonsense suppression at amber stop codons to reduce premature termination [125, 130, 131]. However, development of such strains for other organisms or ncAA incorporation methods may be challenging as the rarely used amber stop codon required significant mutation before a viable *E. coli* strain was produced [125, 130, 131].

Protein labeling, even site-specifically, can also have a dramatic effect on the properties of the protein in a manner that is highly dependent on modification site/sites. Currently, no complete set of parameters exists to identify sites amenable to labeling based on the primary, secondary, or tertiary structural context [118]. This limitation is compounded by a similar lack of knowledge regarding the effects of locational dependence on ncAA incorporation [118, 129]. In order to capitalize on the benefits of ncAA incorporation for biotechnology applications, tools that enable the rapid identification of sites most amenable to ncAA incorporation and post-translational modification are necessary. Such tools include high-throughput screens for modification site evaluation and development of accurate parameters for ncAA incorporation into coarse-grain molecular models to enable rapid in silico screening of modification sites. Development and refinement of such tools are critical to circumvent costly design/build/test cycles for advanced proteins in fields such as imaging, medicine, and biocatalysis.

Another potential solution to improve ncAA incorporation into particular proteins of interest is in vitro or 'cell-free' protein synthesis where some of the factors limiting ncAA incorporation can be overcome. For example, multiple labs have removed the native tRNAs and then added a minimal set of in vitro synthesized tRNAs, essentially emancipating most codons for competition-free ncAA incorporation [63, 132]. Additional advantages that in vitro or 'cell-free' protein synthesis provides over in vivo expression include direct access to the reaction environment, eliminating transport limitations of ncAAs across cell membranes and walls and allowing facile supplementation with exogenous components to improve incorporation efficiency [69, 133]. The flexibility of this system allows the incorporation of less-soluble ncAAs with click-compatible side chains, expanding the repertoire for protein labeling [133]. Importantly, cell-free systems can also be lyophilized for on-demand distributed use in an endotoxin-free format for point-of-care medicinal applications or for rapid response to market demands for biochemical products [134, 135].

In conclusion, ncAA labeling is a versatile tool that enables the identification of de novo protein synthesis and proteome dynamics and adds new functionality to proteins of interest. With the continued development of new technologies for ncAA incorporation, it is increasingly difficult to determine the best approach for a given application. To assist in the experimental design of new applications of ncAA labeling, decision tree diagrams are provided for proteomics and biotechnology applications in Figs. 6 and 7, respectively. It is expected that these technologies will continue to expand into other applications areas in proteomics and biotechnology and be used to increase insights into spatiotemporal protein expression patterns, protein structure-function relationships, and to open new avenues into engineering new protein functions.

Abbreviations
aaRS: aminoacyl tRNA synthetase; Acf: Acetylphenylalanine; Aha: Azidohomoalanine; Anl: Azidonorleucine; Azf: Azidophenylalanine; CuAAC: Copper(I)-catalyzed azide-alkyne cycloaddition; Hag: Homoallylglycine; Hpg: Homopropargylglycine; Met: Methionine; ncAAs: non-canonical amino acids; Pxf: Propargyloxyphenylalanine; SPAAC: Strain promoted azide-alkyne cycloaddition

Availability of data and materials
Not applicable

Authors' contributions
All authors contributed to the writing of this manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable

Consent for publication
All authors have consented to the publication of this manuscript.

Author details
[1]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA. [2]Department of Chemical Engineering, Brigham Young University, Provo, UT, USA.

References
1. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics. 2002;1(5):376–86.
2. Kolb HC, Finn MG, Sharpless KB. Click chemistry: diverse chemical function from a few good reactions. Angew Chem Int Ed. 2001;40(11):2004–21.
3. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proc Natl Acad Sci U S A. 2002;99(1):19–24.
4. Saxon E, Bertozzi CR. Cell surface engineering by a modified Staudinger reaction. Science. 2000;287(5460):2007–10.
5. Tsao ML, Tian F, Schultz PG. Selective Staudinger modification of proteins containing p-azidophenylalanine. ChemBioChem. 2005;6(12):2147–9.
6. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angew Chem Int Ed. 2002;41(14):2596–9.
7. Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG. Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. J Am Chem Soc. 2003;125(11):3192–3.
8. Agard NJ, Prescher JA, Bertozzi CR. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. J Am Chem Soc. 2004;126(46):15046–7.
9. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. A comparative study of bioorthogonal reactions with azides. ACS Chem Biol. 2006;1(10):644–8.
10. Mahmoodi MM, Rashidian M, Dozier JK, Distefano MD. Chemoenzymatic site-specific reversible immobilization and labeling of proteins from crude cellular extract without prior purification using oxime and hydrazine ligation. Curr Protoc Chem Biol. 2013;5(2):89–109.
11. Rashidian M, Dozier JK, Distefano MD. Enzymatic labeling of proteins: techniques and approaches. Bioconjug Chem. 2013;24(8):1277–94.
12. Sherratt AR, Chigrinova M, MacKenzie DA, Rastogi NK, Ouattara MTM, Pezacki AT, et al. Dual strain-promoted alkyne–Nitrone cycloadditions for simultaneous labeling of bacterial peptidoglycans. Bioconjug Chem. 2016; 27(5):1222–6.
13. MacKenzie DA, Sherratt AR, Chigrinova M, Cheung LLW, Pezacki JP. Strain-promoted cycloadditions involving nitrones and alkynes—rapid tunable reactions for bioorthogonal labeling. Curr Opin Chem Biol. 2014;21:81–8.
14. Blackman ML, Royzen M, Fox JM. Tetrazine ligation: fast bioconjugation based on inverse-Electron-demand Diels–Alder reactivity. J Am Chem Soc. 2008;130(41):13518–9.
15. Oliveira B, Guo Z, Bernardes G. Inverse electron demand Diels-Alder reactions in chemical biology. Chem Soc Rev. 2017;46(16):4895–950.
16. Sletten EM, Bertozzi CR. A bioorthogonal quadricyclane ligation. J Am Chem Soc. 2011;133(44):17570–3.
17. Tomlin FM, Gordon CG, Han Y, Wu TS, Sletten EM, Bertozzi CR. Site-specific incorporation of quadricyclane into a protein and photocleavage of the quadricyclane ligation adduct. Bioorg Med Chem. 2018;26(19):5280–90.
18. van Hest JC, Tirrell DA. Protein-based materials, toward a new level of structural control. Chem Commun. 2001;19:1897–904.
19. Kolb HC, Sharpless KB. The growing impact of click chemistry on drug discovery. Drug Discov Today. 2003;8(24):1128–37.
20. Johnson JA, Lu YY, Burts AO, Lim YH, Finn MG, Koberstein JT, et al. Core-clickable PEG-branch-azide bivalent-bottle-brush polymers by ROMP: grafting-through and clicking-to. J Am Chem Soc. 2011;133(3):559–66.
21. Speers AE, Adam GC, Cravatt BF. Activity-based protein profiling in vivo using a copper(i)-catalyzed azide-alkyne [3 + 2] cycloaddition. J Am Chem Soc. 2003;125(16):4686–7.
22. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc Natl Acad Sci U S A. 2006;103(25):9482–7.
23. Jewett JC, Sletten EM, Bertozzi CR. Rapid Cu-free click chemistry with readily synthesized biarylazacyclooctynones. J Am Chem Soc. 2010;132(11):3688–90.
24. Beatty KE, Fisk JD, Smart BP, Lu YY, Szychowski J, Hangauer MJ, et al. Live-cell imaging of cellular proteins by a strain-promoted azide-alkyne cycloaddition. ChemBioChem. 2010;11(15):2092–5.
25. Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, et al. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. Nat Neurosci. 2010;13(7):897–905.
26. Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. In vivo imaging of membrane-associated glycans in developing zebrafish. Science. 2008; 320(5876):664–7.
27. Neef AB, Schultz C. Selective fluorescence labeling of lipids in living cells. Angew Chem Int Ed. 2009;48(8):1498–500.
28. Ngo JT, Tirrell DA. Noncanonical amino acids in the interrogation of cellular protein synthesis. Acc Chem Res. 2011;44(9):677–85.
29. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science. 2001;292(5516):498–500.
30. Link AJ, Mock ML, Tirrell DA. Non-canonical amino acids in protein engineering. Curr Opin Chem Biol. 2003;14(6):603–9.
31. Xiao H, Schultz PG. At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code. Cold Spring Harb Perspect Biol. 2016;8(9):a023945.
32. Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG. A general method for site-specific incorporation of unnatural amino acids into proteins. Science. 1989;244(4901):182–8.
33. Palaniappan KK, Bertozzi CR. Chemical Glycoproteomics. Chem Rev. 2016; 116(23):14277–306.
34. Rexach JE, Clark PM, Hsieh-Wilson LC. Chemical approaches to understanding O-GlcNAc glycosylation in the brain. Nat Chem Biol. 2008;4(2):97–106.
35. Yang S, Rubin J, Eshghi ST, Zhang H. Chemoenzymatic method for glycomics: isolation, identification, and quantitation. Proteomics. 2016;16(2):241–56.
36. Ritzefeld M. Sortagging: a robust and efficient chemoenzymatic ligation strategy. Chemistry. 2014;20(28):8516–29.
37. Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL, Rexach JE, et al. Sortagging: a versatile method for protein labeling chemical approaches to understanding O-GlcNAc glycosylation in the brain. Nat Chem Biol. 2007;3(11):707–8.
38. Lanyon-Hogg T, Faronato M, Serwa RA, Tate EW. Dynamic protein acylation: new substrates, mechanisms, and drug targets. Trends Biochem Sci. 2017; 42(7):566–81.
39. Hang HC, Wilson JP, Charron G. Bioorthogonal chemical reporters for analyzing protein lipidation and lipid trafficking. Acc Chem Res. 2011;44(9):699–708.
40. Thinon E, Hang HC. Chemical reporters for exploring protein acylation. Biochem Soc Trans. 2015;43(2):253–61.
41. Hannoush RN. Synthetic protein lipidation. Curr Opin Chem Biol. 2015;28: 39–46.
42. Duckworth BP, Zhang Z, Hosokawa A, Distefano MD. Selective labeling of proteins by using protein farnesyltransferase. ChemBioChem. 2007;8(1):98–105.
43. Gao X, Hannoush RN. A decade of click chemistry in protein Palmitoylation: impact on discovery and new biology. Cell Chem Biol. 2018;25(3):236–46.
44. Hang HC, Linder ME. Exploring protein lipidation with chemical biology. Chem Rev. 2011;111(10):6341–58.
45. Kulkarni C, Kinzer-Ursem TL, Tirrell DA. Selective functionalization of the protein N terminus with N-myristoyl transferase for bioconjugation in cell lysate. ChemBioChem. 2013;14(15):1958–62.
46. Kulkarni C, Lo M, Fraseur JG, Tirrell DA, Kinzer-Ursem TL. Bioorthogonal Chemoenzymatic functionalization of calmodulin for bioconjugation applications. Bioconjug Chem. 2015;26(10):2153–60.
47. Ejendal KFK, Fraseur J, Kinzer-Ursem T. Protein labeling and bioconjugation using N-Myristoyltransferase. In: Massa S, Devoogdt N, editors. Methods in molecular biology series: bioconjugation. 1. Switzerland: Springer. (In Press).
48. Witten AJ, Ejendal KFK, Gengelbach LM, Traore MA, Wang X, Umulis DM, et al. Fluorescent imaging of protein myristoylation during cellular differentiation and development. J Lipid Res. 2017;58(10):2061–70.

49. Ho SH, Tirrell DA. Chemoenzymatic labeling of proteins for imaging in bacterial cells. J Am Chem Soc. 2016;138(46):15098–101.

50. van Hest JCM, Kiick KL, Tirrell DA. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. J Am Chem Soc. 2000;122(7): 1282–8.

51. Kiick KL, van Hest JC, Tirrell DA. Expanding the scope of protein biosynthesis by altering the Methionyl-tRNA Synthetase activity of a bacterial expression host. Angew Chem Int Ed. 2000;39(12):2148–52.

52. van Hest JC, Tirrell DA. Efficient introduction of alkene functionality into proteins in vivo. FEBS Lett. 1998;428(1–2):68–70.

53. Alvarez-Castelao B, Schanzenbacher CT, Hanus C, Glock C, Tom Dieck S, Dorrbaum AR, et al. Cell-type-specific metabolic labeling of nascent proteomes in vivo. Nat Biotechnol. 2017;35(12):1196–201.

54. Mahdavi A, Hamblin GD, Jindal GA, Bagert JD, Dong C, Sweredoski MJ, et al. Engineered aminoacyl-tRNA Synthetase for cell-selective analysis of mammalian protein synthesis. J Am Chem Soc. 2016;138(13):4278–81.

55. Yuet KP, Doma MK, Ngo JT, Sweredoski MJ, Graham RL, Moradian A, et al. Cell-specific proteomic analysis in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2015;112(9):2705–10.

56. Truong F, Yoo TH, Lampo TJ, Tirrell DA. Two-strain, cell-selective protein labeling in mixed bacterial cultures. J Am Chem Soc. 2012;134(20):8551–6.

57. Ngo JT, Babin BM, Champion JA, Schuman EM, Tirrell DA. State-selective metabolic labeling of cellular proteins. ACS Chem Biol. 2012;7(8):1326–30.

58. Furter R. Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. Protein Sci. 1998;7(2):419–26.

59. Sun SB, Schultz PG, Kim CH. Therapeutic applications of an expanded genetic code. ChemBioChem. 2014;15(12):1721–9.

60. Zemella A, Thoring L, Hoffmeister C, Kubick S. Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. ChemBioChem. 2015; 16(17):2420–31.

61. Chatterjee A, Sun SB, Furman JL, Xiao H, Schultz PG. A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. Biochemistry. 2013;52(10):1828–37.

62. Wan W, Huang Y, Wang Z, Russell WK, Pai PJ, Russell DH, et al. A facile system for genetic incorporation of two different noncanonical amino acids into one protein in Escherichia coli. Angew Chem Int Ed. 2010;122(18):3279–82.

63. Cui Z, Mureev S, Polinkovsky ME, Tnimov Z, Guo Z, Durek T, et al. Combining sense and nonsense codon reassignment for site-selective protein modification with unnatural amino acids. ACS Synth Biol. 2017;6(3): 535–44.

64. Zang Q, Tada S, Uzawa T, Kiga D, Yamamura M, Ito Y. Two site genetic incorporation of varying length polyethylene glycol into the backbone of one peptide. Chem Commun. 2015;51(76):14385–8.

65. 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):9778–9.

66. Dougherty DA, Van Arnam EB. In vivo incorporation of non-canonical amino acids by using the chemical aminoacylation strategy: a broadly applicable mechanistic tool. ChemBioChem. 2014;15(12):1710–20.

67. Bundy BC, Swartz JR. Site-specific incorporation of p-Propargyloxyphenylalanine in a cell-free environment for direct protein –protein click conjugation. Bioconjug Chem. 2010;21(2):255–63.

68. Smith MT, Wu JC, Varner CT, Bundy BC. Enhanced protein stability through minimally invasive, direct, covalent, and site-specific immobilization. Biotechnol Prog. 2013;29(1):247–54.

69. Rodriguez EA, Lester HA, Dougherty DA. In vivo incorporation of multiple unnatural amino acids through nonsense and frameshift suppression. Proc Natl Acad Sci U S A. 2006;103(23):8650–5.

70. Xiao H, Chatterjee A, Choi S, Bajjuri KM, Sinha SC, Schultz PG. Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. Angew Chem Int Ed. 2013;52(52):14080–3.

71. Smith MT, Hawes AK, Bundy BC. Reengineering viruses and virus-like particles through chemical functionalization strategies. Curr Opin Biotechnol. 2013;24(4):620–6.

72. Dieterich DC, Lee JJ, Link AJ, Graumann J, Tirrell DA, Schuman EM. Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. Nat Protoc. 2007;2(3): 532–40.

73. Beatty KE, Liu JC, Xie F, Dieterich DC, Schuman EM, Wang Q, et al. Fluorescence visualization of newly synthesized proteins in mammalian cells. Angew Chem Int Ed. 2006;45(44):7364–7.

74. Zhang MM, Tsou LK, Charron G, Raghavan AS, Hang HC. Tandem fluorescence imaging of dynamic S-acylation and protein turnover. Proc Natl Acad Sci USA. 2010;107(19):8627–32.

75. Zhang J, Wang J, Ng S, Lin Q, Shen HM. Development of a novel method for quantification of autophagic protein degradation by AHA labeling. Autophagy. 2014;10(5):901–12.

76. Choi KYG, Lippert DND, Ezzatti P, Mookherjee N. Defining TNF-α and IL-1β induced nascent proteins: combining bio-orthogonal non-canonical amino acid tagging and proteomics. J Immunol Methods. 2012;382(1-2):189–95.

77. Bagert JD, Van Kessel JC, Sweredoski MJ, Feng L, Hess S, Bassler BL, et al. Time-resolved proteomic analysis of quorum sensing in Vibrio harveyi. Chem Sci. 2016;7(3):1797–1806.

78. Mahdavi A, Szychowski J, Ngo JT, Sweredoski MJ, Graham RL, Hess S, et al. Identification of secreted bacterial proteins by noncanonical amino acid tagging. Proc Natl Acad Sci U S A. 2014;111(1):433–8.

79. Van Elsland DM, Bos E, De Boer W, Overkleeft HS, Koster AJ, Van Kasteren SI. Detection of bioorthogonal groups by correlative light and electron microscopy allows imaging of degraded bacteria in phagocytes. Chem Sci. 2016;7:752–8.

80. Hinz FI, Dieterich DC, Tirrell DA, Schuman EM. Non-canonical amino acid labeling in vivo to visualize and affinity purify newly synthesized proteins in larval zebrafish. ACS Chem Neurosci. 2012;3(1):40–9.

81. Ullrich M, Liang V, Chew YL, Banister S, Song X, Zaw T, et al. Bio-orthogonal labeling as a tool to visualize and identify newly synthesized proteins in Caenorhabditis elegans. Nat Protoc. 2014;9(9):2237–55.

82. Shen W, Liu HH, Schiapparelli L, McClatchy D, Hy H, Yates JR, et al. Acute Synthesis of CPEB Is Required for Plasticity of Visual Avoidance Behavior in Xenopus. Cell Rep. 2014;6:737–47.

83. Liu J, Xu Y, Stoleru D, Salic A. Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. Proc Natl Acad Sci U S A. 2012;109:413–8.

84. Schiapparelli LM, McClatchy DB, Liu HH, Sharma P, Yates JR, Cline HT. Direct detection of biotinylated proteins by mass spectrometry. J Proteome Res. 2014;13:3966–78.

85. McClatchy DB, Ma Y, Liu C, Stein BD, Martínez-Bartolomé S, Vasquez D, et al. Pulsed azidohomoalanine labeling in mammals (PALM) detects changes in liver-specific LKB1 knockout mice. J Proteome Res. 2015.

86. McClatchy DB, Ma Y, Liem DA, Ng DCM, Ping P, Yates JR. Quantitative temporal analysis of protein dynamics in cardiac remodeling. J Mol Cell Cardiol. 2018;121:163–72.

87. Calve S, Witten AJ, Ocken AR, Kinzer-Ursem TL. Incorporation of non-canonical amino acids into the developing murine proteome. Sci Rep. 2016;6:32377.

88. Link AJ, Vink MKS, Agard NJ, Prescher JA, Bertozzi CR, Tirrell DA. Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. Proc Natl Acad Sci U S A. 2006;103(27):10180–5.

89. Tanrikulu IC, Schmitt E, Mechulam Y, Goddard WA 3rd, Tirrell DA. Discovery of Escherichia coli methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo. Proc Natl Acad Sci U S A. 2009; 106(36):15285–90.

90. Ngo JT, Champion JA, Mahdavi A, Tanrikulu IC, Beatty KE, Connor RE, et al. Cell-selective metabolic labeling of proteins. Nat Chem Biol. 2009;5(10):715–7.

91. Grammel M, Zhang MM, Hang HC. Orthogonal alkynyl amino acid reporter for selective labeling of bacterial proteomes during infection. Angew Chem Int Ed. 2010;49(34):5970–4.

92. Wier GM, McGreevy EM, Brown MJ, Boyle JP. New Method for the Orthogonal Labeling and Purification of Toxoplasma gondii Proteins While Inside the Host Cell. mBio. 2015;6(2):e01628–14.

93. Erdmann I, Marter K, Kobler O, Niehues S, Abele J, Müller A, et al. Cell-selective labelling of proteomes in Drosophila melanogaster. Nat Commun. 2015;6:e01628–14.

94. Niehues S, Bussmann J, Steffes G, Erdmann I, Köhrer C, Sun L, et al. Impaired protein translation in Drosophila models for Charcot-Marie-tooth neuropathy caused by mutant tRNA synthetases. Nat Commun. 2015;6:7520.

95. Müller A, Stellmacher A, Freitag CE, Landgraf P, Dieterich DC. Monitoring Astrocytic Proteome Dynamics by Cell Type-Specific Protein Labeling. PloS One. 2015;10(12):e014545.

96. Liu Y, Conboy MJ, Mehdipour M, Liu Y, Tran TP, Blotnick A, et al. Application of bio-orthogonal proteome labeling to cell transplantation and heterochronic parabiosis. Nat Commun. 2017;8(1):643.

97. Sadlowski CM, Balderston S, Sandhu M, Hajian R, Liu C, Tran T-T, et al. Graphene-based biosensor for on-chip detection of Bio-orthogonally Labeled Proteins to Identify the Circulating Biomarkers of Aging during Heterochronic Parabiosis. Lab Chip. 2018;18(21):3230–38.

98.   Yang AC, du Bois H, Olsson N, Gate D, Lehallier B, Berdnik D, et al. Multiple click-selective tRNA Synthetases expand mammalian cell-specific proteomics. J Am Chem Soc. 2018;140(23):7046–51.

99.   Bianco A, Townsley FM, Greiss S, Lang K, Chin JW. Expanding the genetic code of Drosophila melanogaster. Nat Chem Biol. 2012;8(9):748–50.

100. Chang H, Han M, Huang W, Wei G, Chen J, Chen PR, et al. Light-induced protein translocation by genetically encoded unnatural amino acid in Caenorhabditis elegans. Protein Cell. 2013;4(12):883–6.

101. Levine PM, Craven TW, Bonneau R, Kirshenbaum K. Intrinsic bioconjugation for site-specific protein PEGylation at N-terminal serine. Chem Commun. 2014;50:6909–12.

102. Dozier JK, Distefano MD. Site-specific PEGylation of therapeutic proteins. Int J Mol Sci. 2015;16(10):25831–64.

103. Roberts MJ, Bentley MD, Harris JM. Chemistry for peptide and protein PEGylation. Adv Drug Del Rev. 2012;64:Supplement:116–27.

104. Deepankumar K, Nadarajan SP, Mathew S, Lee S-G, Yoo TH, Hong EY, et al. Engineering transaminase for stability enhancement and site-specific immobilization through multiple noncanonical amino acids incorporation. Chem Cat Chem. 2014;7(3):417–21.

105. Halder K, Dölker N, Van Q, Gregor I, Dickmanns A, Baade I, et al. MD simulations and FRET reveal an environment-sensitive conformational plasticity of importin-β. Biophys J. 2015;109(2):277–86.

106. Turcatti G, Nemeth K, Edgerton MD, Meseth U, Talabot F, Peitsch M, et al. Probing the structure and function of the tachykinin Neurokinin-2 receptor through biosynthetic incorporation of fluorescent amino acids at specific sites. J Biol Chem. 1996;271(33):19991–8.

107. Ravikumar Y, Nadarajan SP, Hyeon Yoo T, Lee C-S, Yun H. Incorporating unnatural amino acids to engineer biocatalysts for industrial bioprocess applications. Biotechnol J. 2015;10(12):1862–76.

108. Ngo TA, Nakata E, Saimura M, Morii T. Spatially organized enzymes drive cofactor-coupled Cascade reactions. J Am Chem Soc. 2016;138(9):3012–21.

109. Song J, Su P, Ma R, Yang Y, Yang Y. Based on DNA Strand displacement and functionalized magnetic nanoparticles: a promising strategy for enzyme immobilization. Ind Eng Chem Res. 2017;56(17):5127–37.

110. Liu M, Fu J, Qi X, Wootten S, Woodbury NW, Liu Y, et al. A three-enzyme pathway with an optimised geometric arrangement to facilitate substrate transfer. ChemBioChem. 2016;17(12):1097–101.

111. Cho H, Daniel T, Buechler YJ, Litzinger DC, Maio Z, Putnam A-MH, et al. Optimized clinical performance of growth hormone with an expanded genetic code. Proc Natl Acad Sci U S A. 2011;108(22):9060–5.

112. Mu J, Pinkstaff J, Li Z, Skidmore L, Li N, Myler H, et al. FGF21 analogs of sustained action enabled by orthogonal biosynthesis demonstrate enhanced antidiabetic pharmacology in rodents. Diabetes. 2012;61(2):505–12.

113. Gauba V, Grünewald J, Gorney V, Deaton LM, Kang M, Bursulaya B, et al. Loss of CD4 T-cell-dependent tolerance to proteins with modified amino acids. Proc Natl Acad Sci U S A. 2011;108(31):12821–6.

114. Lu Y, Chan W, Ko BY, VanLang CC, Swartz JR. Assessing sequence plasticity of a virus-like nanoparticle by evolution toward a versatile scaffold for vaccines and drug delivery. Proc Natl Acad Sci U S A. 2015;112(40):12360.

115. Fan Y, Su F, Li K, Ke C, Yan Y. Carbon nanotube filled with magnetic iron oxide and modified with polyamidoamine dendrimers for immobilizing lipase toward application in biodiesel production. Sci Rep. 2017;7:45643.

116. Wilding KM, Schinn S-M, Long EA, Bundy BC. The emerging impact of cell-free chemical biosynthesis. Curr Opin Biotechnol. 2018;53:115–21.

117. Jackson E, López-Gallego F, Guisan JM, Betancor L. Enhanced stability of l-lactate dehydrogenase through immobilization engineering. Process Biochem. 2016;51(9):1248–55.

118. Wilding KM, Smith AK, Wilkerson JW, Bush DB, Knotts TA, Bundy BC. The Locational Impact of Site-Specific PEGylation: Streamlined Screening with Cell-Free Protein Expression and Coarse-Grain Simulation. ACS Synth Biol. 2018;7(2):510–21.

119. Pelegri-O'Day EM, Lin E-W, Maynard HD. Therapeutic protein–polymer conjugates: advancing beyond PEGylation. J Am Chem Soc. 2014;136(41):14323–32.

120. Fuhrmann G, Grotzky A, Lukić R, Matoori S, Luciani P, Yu H, et al. Sustained gastrointestinal activity of dendronized polymer–enzyme conjugates. Nat Chem. 2013;5(7):582–9.

121. Fogarty JA, Swartz JR. The exciting potential of modular nanoparticles for rapid development of highly effective vaccines. Curr Opin Chem Eng. 2018;19:1–8.

122. Rohovie MJ, Nagasawa M, Swartz JR. Virus-like particles: next-generation nanoparticles for targeted therapeutic delivery. Bioeng Transl Med. 2016;2(1):43–57.

123. Strop P, Liu S-H, Dorywalska M, Delaria K, Dushin Russell G, Tran T-T, et al. Location matters: site of conjugation modulates stability and pharmacokinetics of antibody drug conjugates. Chem Biol. 2013;20(2):161–7.

124. Wu JC, Hutchings CH, Lindsay MJ, Werner CJ, Bundy BC. Enhanced enzyme stability through site-directed covalent immobilization. J Biotechnol. 2015;193:83–90.

125. Yin G, Stephenson HT, Yang J, Li X, Armstrong SM, Heibeck TH, et al. RF1 attenuation enables efficient non-natural amino acid incorporation for production of homogeneous antibody drug conjugates. Sci Rep. 2017;7(1):3026.

126. Zimmerman ES, Heibeck TH, Gill A, Li X, Murray CJ, Madlansacay MR, et al. Production of site-specific antibody–drug conjugates using optimized non-natural amino acids in a cell-free expression system. Bioconjug Chem. 2014;25(2):351–61.

127. Oller-Salvia B, Kym G, Chin JW. Rapid and efficient generation of stable antibody-drug conjugates via an encoded Cyclopropene and an inverse-Electron-demand Diels-Alder reaction. Angewandte Chemie (International ed in English). 2018;57(11):2831–4.

128. Wei S, Knotts TA. Effects of tethering a multistate folding protein to a surface. J Chem Phys. 2011;134(18):185101.

129. Schinn S-M, Bradley W, Groesbeck A, Wu JC, Broadbent A, Bundy BC. Rapid in vitro screening for the location-dependent effects of unnatural amino acids on protein expression and activity. Biotechnol Bioeng. 2017;114(10):2412–7.

130. Johnson DBF, Xu J, Shen Z, Takimoto JK, Schultz MD, Schmitz RJ, et al. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. Nat Chem Biol. 2011;7(11):779–86.

131. Lajoie MJ, Rovner AJ, Goodman DB, Aerni H-R, Haimovich AD, Kuznetsov G, et al. Genomically recoded organisms expand biological functions. Science. 2013;342(6156):357.

132. Salehi ASM, Smith MT, Schinn S-M, Hunt JM, Muhlestein C, Diray-Arce J, et al. Efficient tRNA degradation and quantification in *Escherichia coli* cell extract using RNase-coated magnetic beads: A key step toward codon emancipation. Biotechnol Prog. 2017;33(5):1401–7.

133. Bundy BC, Hunt JP, Jewett MC, Swartz JR, Wood DW, Frey DD, et al. Cell-free biomanufacturing. Curr Opin Chem Eng. 2018;22:177–83.

134. Wilding KM, Hunt JP, Wilkerson JW, Funk PJ, Swensen RL, Carver WC, et al. Endotoxin-free E. coli-based cell-free protein synthesis: Pre-expression endotoxin removal approaches for on-demand cancer therapeutic production. Biotechnol J. 2019;14(3):e1800271.

135. Smith MT, Berkheimer SD, Werner CJ, Bundy BC. Lyophilized Escherichia coli-based cell-free systems for robust, high-density, long-term storage. BioTechniques. 2014;56(4):186–93.

136. Datta D, Wang P, Carrico IS, Mayo SL, Tirrell DA. A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins. J Am Chem Soc. 2002;124(20):5652–3.