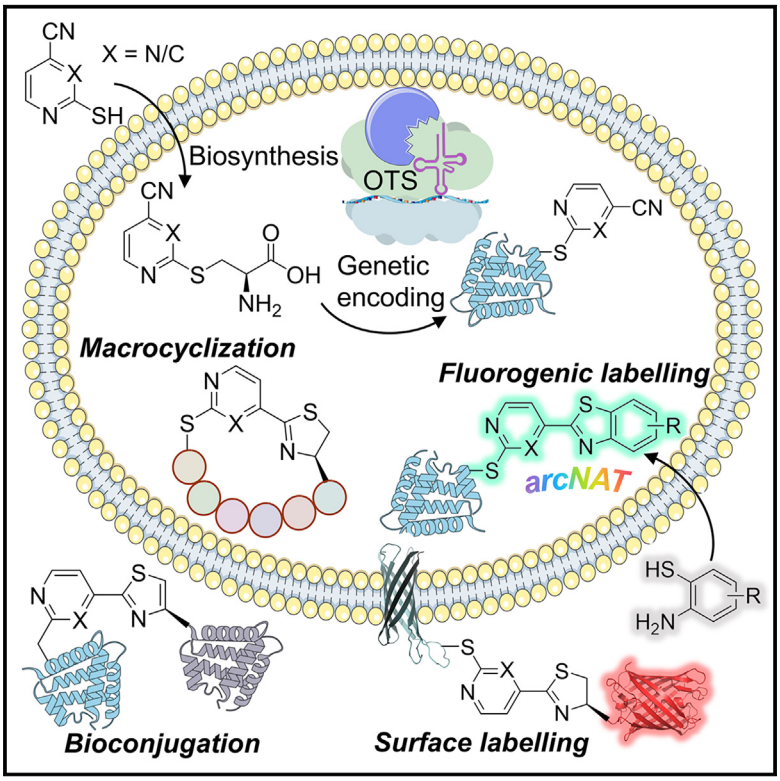
Chem

Article

Biosynthesis and genetic encoding of activated nitriles for fast protein conjugation and tunable fluorogenic labeling

Graphical abstract



Authors

Elwy H. Abdelkader, Haocheng Qianzhu, Gottfried Otting, Thomas Huber

Correspondence

gottfried.otting@anu.edu.au (G.O.), t.huber@anu.edu.au (T.H.)

In brief

A new platform for biosynthesizing and genetically encoding activated nitrile-bearing amino acids enables precise protein labeling and rapid post-translational modifications. Using the biocompatible nitrile-aminothiol (NAT) click reaction, this system supports tunable fluorogenic labeling and wash-free, live-cell imaging, simplifying workflows and unlocking new possibilities in biochemistry, chemical biology, and therapeutic development.

Highlights

- Cost-effective biosynthesis and genetic encoding of nitrile-bearing amino acids
- High-yield production of nitrile-bearing proteins through genetic code expansion
- Fast site-specific protein labeling via the nitrile-aminothiol (NAT) click reaction
- Wash-free, fluorogenic labeling for live-cell imaging via the arcNAT click reaction

Abdelkader et al., 2025, Chem *11*, 102385
May 8, 2025
https://doi.org/10.1016/j.chempr.2024.12.003

CellPress

Chem

CellPress

Article

Biosynthesis and genetic encoding of activated nitriles for fast protein conjugation and tunable fluorogenic labeling

Elwy H. Abdelkader,[1,3] Haocheng Qianzhu,[2,3] Gottfried Otting,[1,*] and Thomas Huber[2,4,*]

[1]ARC Centre of Excellence for Innovations in Peptide & Protein Science, Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

[2]Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

[3]These authors contributed equally

[4]Lead contact

*Correspondence: gottfried.otting@anu.edu.au (G.O.), t.huber@anu.edu.au (T.H.)

https://doi.org/10.1016/j.chempr.2024.12.003

THE BIGGER PICTURE Advancements in protein engineering are increasingly focused on sustainability and the development of resource-efficient chemical tools. Conventional methods for site-specific labeling and post-translational modifications of proteins often rely on expensive reagents and complex synthetic processes, limiting accessibility and scalability. Our work presents a sustainable platform that repurposes readily available chemicals for the *in vivo* biosynthesis and site-specific incorporation of reactive nitrile-bearing amino acids into proteins in *E. coli*. By harnessing bacterial biosynthetic pathways, this approach minimizes chemical waste and reduces the dependence on resource-intensive processes. Our system integrates the nitrile-aminothiol click reaction for tunable fluorogenic labeling and rapid protein conjugation, enabling efficient, wash-free, live-cell labeling under physiological conditions. This method simplifies workflows and broadens the scope of applications in fluorescence imaging, protein functionalization, and chemical biology. It empowers researchers with sustainable tools for drug discovery, diagnostics, and therapeutic applications while promoting responsible resource usage.

SUMMARY

Few chemistries are suitable for in-cell protein labeling, and the required reagents are costly. We present an approach for the coupled biosynthesis and genetic encoding of activated nitriles, delivering a facile way to furnish proteins with biocompatible reactive handles suitable for subsequent site-specific modifications both in cell and *in vitro*. The strategy utilizes the endogenous bacterial cysteine biosynthetic machinery to produce the nitrile-bearing non-canonical amino acids (ncAAs) *in situ* and then perform genetic encoding through an engineered orthogonal translation system. We demonstrate the utility of our system for rapid site-specific bioconjugation and macrocyclization through the nitrile-aminothiol (NAT) click reaction. In addition, we introduce the aromatic condensation NAT (arcNAT) click reaction as a tool for generating a diverse array of turn-on fluorophores. arcNAT achieves fluorogenic labeling of proteins for live-cell microscopy without requiring washing steps. Our approach provides a uniquely convenient, versatile, and cost-effective platform for the post-translational diversification of proteins.

INTRODUCTION

Modern biosciences critically depend on biocompatible reactions for the precise post-translational modification of proteins *in vivo* with minimal side reactions and toxicity. Only a few synthetic reactions meet the criteria of biocompatibility, where (1) the reaction proceeds efficiently under physiological conditions, (2) none of the educts react with endogenous molecules at their cellular concentrations, and (3) no potentially cytotoxic catalyst is required. Prominent examples include strain-promoted cycloadditions (SPCAs) and inverse electron-demand Diels-Alder (IEDDA) reactions, which have found wide use in numerous applications, including the tagging of biomolecules for structural investigations, the fluorescent labeling of proteins to track their intracellular localization, and antibody-drug conjugations for therapeutic applications.[1–11] The drawbacks of these reactions

are the instability of the reactive handles in a reducing *in vivo* environment[12,13] and the chemical complexity of their reactive groups, which often involves elaborate and costly chemical synthesis.[14–18]

The nitrile-aminothiol (NAT) click reaction presents a promising alternative to existing bioconjugation methods because it requires only relatively simple and inexpensive reactants. We previously introduced this condensation reaction by using a cyanopyridine moiety to present an electrophilic nitrile for a straightforward reaction with 1,2-aminothiol compounds, such as L-cysteine, under catalyst-free conditions.[19] The low complexity of the cyanopyridine moiety proved to be particularly beneficial given that non-canonical amino acids (ncAAs) with this group can be readily synthesized and was amenable to efficient genetic encoding for ribosomal protein synthesis by genetic code expansion (GCE).[20] Furthermore, peptides and proteins with an N-terminal cysteine residue present a 1,2-aminothiol functionality suitable for the NAT click reaction and are readily accessible through the proteolytic cleavage of a leader sequence.[21–23] Alternatively, the functionality can be introduced by 1,2-aminothiol-bearing ncAAs.[24,25] The NAT click reaction thus offers an exceptionally effective biocompatible strategy for ligating highly complex molecules. Here, we show that the reaction also offers a straightforward way to turn on fluorescent labeling.

Fluorescent labeling is fundamentally important to tracking biological macromolecules as well as analyzing their structure, function, and dynamics both *in vitro* and *in vivo*.[26–30] The exceptional sensitivity of fluorescence allows the single-molecule detection of biomolecules and microscopy at sub-micrometer resolution, making it an indispensable tool for bioanalytics.[31–34] Commonly, the fluorescent tagging of proteins employs either fusions with fluorescent proteins, which are large and usually limited to the N- or C-terminal ends of the target protein,[35] or small organic fluorescent dyes with high quantum yields that are covalently conjugated to the protein. Although straightforward in principle, the chemical labeling of native proteins often suffers from a lack of site specificity and target selectivity because complex cellular environments lead to off-target labeling.[36,37] GCE provides a solution to this problem by site-specifically installing a ncAA. These ncAAs can either contain a fluorescent moiety in their side chain[38–40] or serve as a reactive handle for subsequent conjugation with fluorescent groups in a biocompatible reaction.[41,42] Although GCE is a well-established technique capable of incorporating over 500 different ncAAs into proteins, the number of genetically encoded fluorescent amino acids remains small while their excitation wavelength tends to be short (<400 nm) and their brightness is relatively low.[43,44] By contrast, biocompatible reactions that conjugate ncAAs with fluorophores possessing better photochemical properties have proven to be highly successful in many applications.[45–47] As a downside, this approach typically requires special reactive groups that render the synthesis of the ncAA and the corresponding fluorescent tag expensive. Furthermore, because both the tagging reagents and the labeled target proteins share similar absorbance and fluorescence properties, excess reagents must be removed by dialysis or extensive washing of cells. As a result, GCE-based approaches for the fluorescent labeling of proteins are often perceived as technically challenging and resource intensive, hindering their widespread adoption in many research labs.[48] Recent studies have demonstrated the efficient biosynthesis and GCE of aromatic thioethers of cysteine.[49,50] These aromatic thioether ncAAs lack intrinsic fluorescence in the visible range, but we show that they can serve as versatile handles for subsequent conjugation with either fluorophores or aromatic compounds that create fluorophores with long wavelength emissions. The modular nature of this approach allows labeling with diverse fluorophores, expanding the scope for fluorescence-based applications.

This work is based on the coupled biosynthesis and genetic encoding of activated nitriles. Specifically, we demonstrate that the biosynthesis of two reactive nitrile-bearing ncAAs, *S*-(4-cyanopyridin-2-yl)-L-cysteine (CysCNP) and *S*-(4-cyanopyrimidin-2-yl)-L-cysteine (CysCNPym), by the endogenous cysteine biosynthetic pathway in *E. coli* is sufficiently fast to install the ncAAs by genetic encoding in high yield and purity. In this way, proteins containing site-specific activated nitrile handles are produced from inexpensive and commercially available precursors. Furthermore, the heteroaromatic moieties in the ncAAs undergo fast NAT click reactions in a facile and catalyst-free manner. We introduce the aromatic condensation NAT (arcNAT) click reaction by reacting the activated nitrile groups with 2-aminobenzenethiols (ABTs) to obtain fluorophores absorbing and emitting in the visible spectral range. The reaction proceeds efficiently under physiological conditions in complex biological environments, including live cells. With ON/OFF ratios of up to three orders of magnitude, this protocol provides highly tunable, background-free fluorescence easily detectable in complex biological matrices without time-consuming dialysis or washing of the labeled cells.

RESULTS AND DISCUSSION

Coupling the biosynthesis of activated nitrile-bearing ncAAs with GCE

The NAT click reaction depends on electron-deficient nitriles, but the corresponding ncAAs are costly because of time-consuming multistep syntheses.[19,20,51] By contrast, the approach described below uses readily available nitrile building blocks and harnesses bacterial biosynthesis to produce the ncAAs CysCNP and CysCNPym *in vivo* (Figure 1A).

Among the different biosynthetic pathways that could be exploited for producing a NAT-compatible nitrile-bearing ncAA,[25,52–61] we identified the cysteine biosynthetic pathway as the most promising candidate.[62] In *E. coli*, cysteine synthase A (CysK) and cysteine synthase B (CysM) are pyridoxal-5′-phosphate (PLP)-dependent enzymes, which catalyze cysteine synthesis from *O*-acetyl-L-serine and sulfide ions via a β-substitution reaction (Figure 1B). Previous studies showed that CysK and CysM are capable of producing ncAAs by substituting sulfide with thiol compounds as nucleophiles.[49,50,63] Hypothesizing that these enzymes could also accept an activated nitrile-bearing pyridine-thiol as substrate (Figure S1), we identified 2-sulfanylpyridine-4-carbonitrile (CNP-thiol) as a potential precursor for the biosynthetic transformation to CysCNP (Figure 1B).

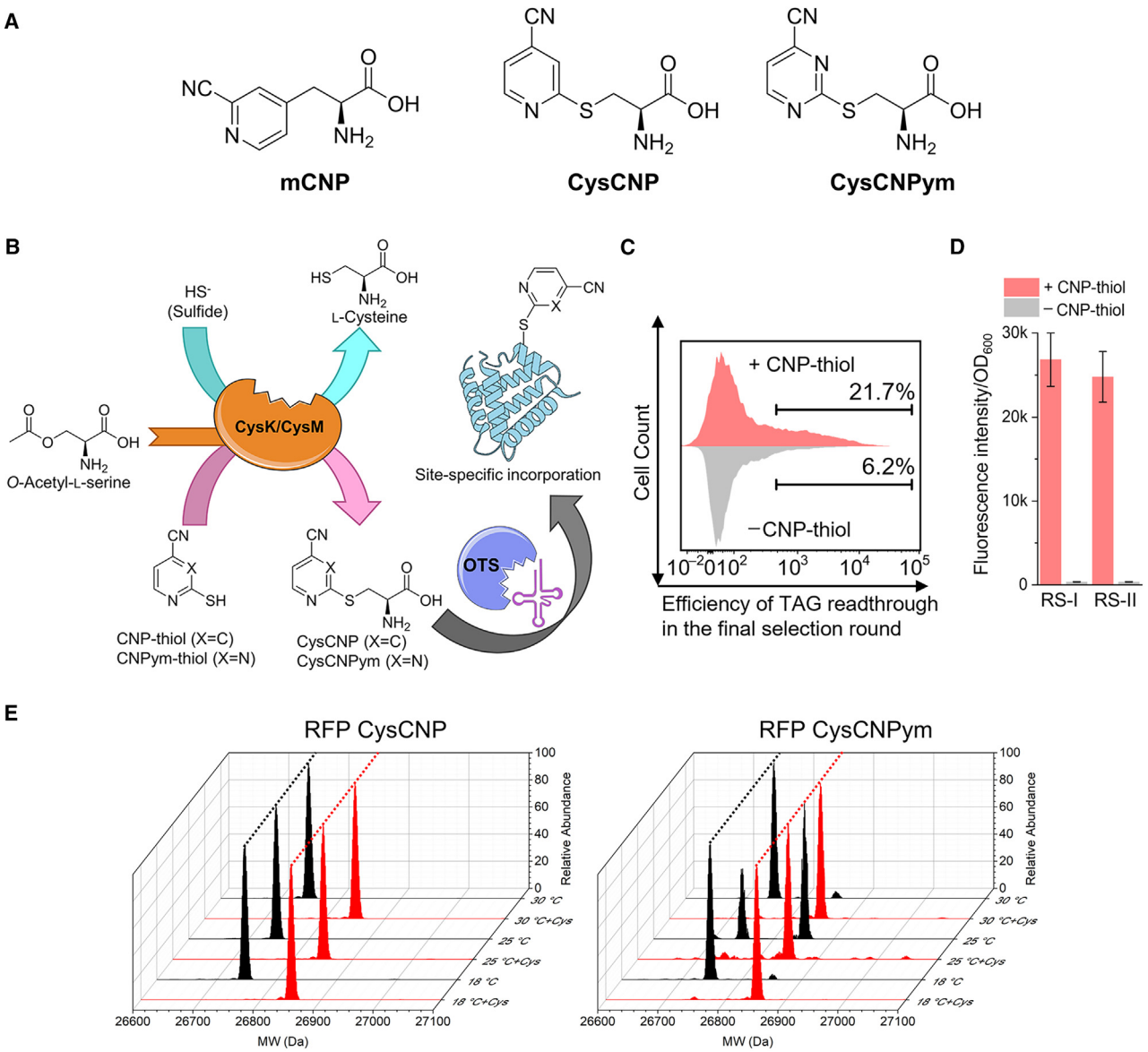


Figure 1. Coupling the biosynthesis of activated nitrile-bearing ncAAs with GCE

(A) Chemical structures of ncAAs with an activated nitrile group. mCNP is fully synthetic. CysCNP and CysCNPym are produced enzymatically *in vivo*.

(B) L-cysteine biosynthetic pathway in *E. coli*. CysK and CysM use sulfide ions (HS^-) to convert *O*-acetyl-L-serine to L-cysteine. HS^- can be replaced with CNP-thiol and CNPym-thiol for the production of CysCNP and CysCNPym, respectively. The ncAA biosynthesis is fast, allowing simultaneous selection of an orthogonal aminoacyl-tRNA synthetase for the incorporation of CysCNP and CysCNPym by suppression of the amber stop codon.

(C) Histogram representation of cells expressing RFP after the fifth selection round for CysCNP-specific G1PylRS mutants. The horizontal axis represents the level of red fluorescence, and the vertical axis indicates the number of cells with this fluorescence level. The difference in RFP fluorescence intensity between cells grown with and without CNP-thiol indicates the presence of CysCNP-specific G1PylRS mutants in the gene pool.

(D) Activity of the top-performing G1RS mutants, selected on the basis of their efficiency in suppressing an amber codon in the RFP reporter gene by CysCNP incorporation ($n \geq 2 \pm$ standard deviation).

(E) Intact protein mass spectrometry analysis of RFP CysCNP and RFP CysCNPym expressed under varying conditions: 25°C or 18°C for 16 h or 30°C for 5 h (black). Red spectra: after ligation with L-cysteine (100 μM protein, 10 mM L-Cys in PBS buffer containing 20 mM tris(2-carboxyethyl)phosphine [TCEP], 25°C for 1 h). The dashed lines identify the expected masses of the unreacted and reacted proteins. The individual mass spectra are shown in Figure S15.

Using CNP-thiol as the precursor for the biosynthesis of CysCNP, we screened a library of variants of the pyrrolysyl-tRNA synthase derived from the methanogenic archaeon ISO4-G1 (G1PylRS) for mutants capable of recognizing CysCNP as a specific substrate.[20,64–66] The library targeted mutations in the amino acid binding pocket by mutating the sites of Leu124, Tyr125, Ala221, and Trp237 to all 20 canonical amino acids; Asn165 to a set of eight amino acids (Gly, Ala, Val, Ser,

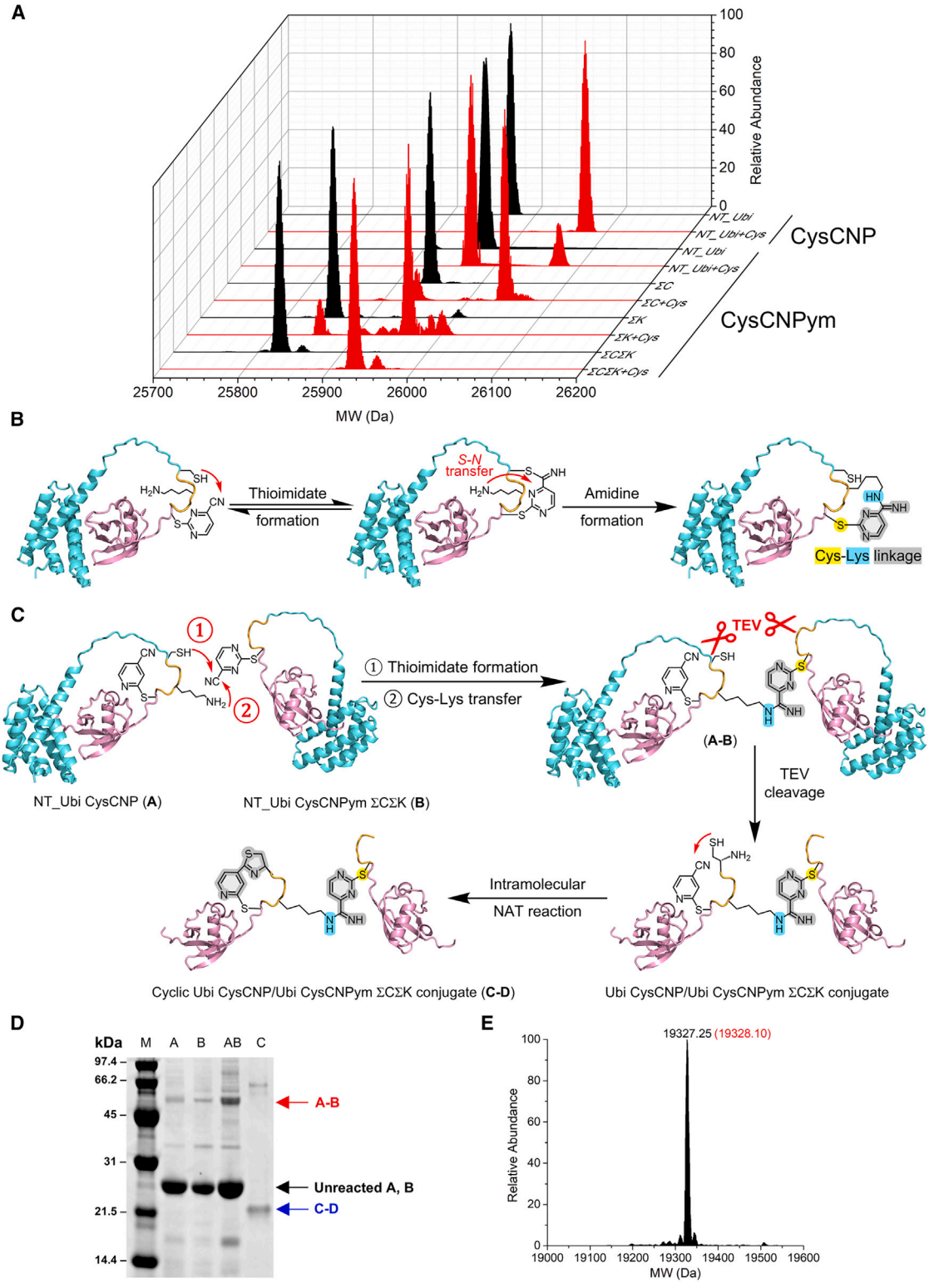


Figure 2. Reaction between CysCNPym and the cysteine-lysine motif for intermolecular bioconjugation via site-selective lysine ligation

(A) Intact protein mass spectrometry analysis of unreacted NT_Ubi CysCNP and NT_Ubi CysCNPym (black) and their respective NAT reaction products with L-Cys (red). The individual mass spectra are shown in Figure S15.

(legend continued on next page)

Asn, Thr, Ile, and Asp); Val167 to seven amino acids (Gly, Ala, Val, Ser, Cys, Leu, and Phe); and Tyr204 to Phe, Tyr, and Trp. *E. coli* DH10B cells were co-transformed with the library plasmid pBK-G1PylRS and the selection plasmid pBAD-H6RFP encoding mCherry red fluorescent protein (RFP) with an amber stop codon following an N-terminal His_6 tag. Using fluorescence-activated cell sorting (FACS), we conducted multiple rounds of alternating positive (with 0.5 mM CNP-thiol in the culture medium) and negative (without supplementing CNP-thiol) selection rounds to enrich for cells expressing functional CysCNP-specific G1PylRS (Figures 1C and S2). After five selection rounds, several colonies expressing high levels of RFP were isolated and sequenced, revealing two distinct G1PylRS mutant sequences: RS-I (124S/215F/165N/167F/204W/221G/237Y) and RS-II (124L/215L/165T/167F/204W/221G/237K) (Figures 1D and S3 and Table S1).

The top-performing RS-I enzyme, designated G1(CysCNP)RS, was cloned into the high-copy-number pRSF plasmid for use in our previously reported dual-plasmid system for efficient suppression of the amber stop codon.[67] The correct biosynthesis and genetic encoding of CysCNP was confirmed with the NT*_GDX construct, which consists of the NT* solubility tag followed by a TEV protease cleavage site, glycine, aspartic acid, and the ncAA encoded by the amber stop codon.[65] The NT*_GD(CysCNP) protein was cleaved with TEV protease, and the liberated GD(CysCNP) tripeptide containing the incorporated CysCNP was separated from the NT* tag and TEV protease by Ni-NTA resin purification. ^{1}H NMR spectra of the tripeptide confirmed near-complete incorporation of the ncAA and its identity as CysCNP (Figure S4).

The dual-plasmid system with G1(CysCNP)RS displayed exceptional efficiency and fidelity. It surpassed the expression levels achievable with mCNP supplemented at a concentration of 1 mM. Remarkably, the expression levels of proteins containing CysCNP were comparable to those of the wild-type proteins (Table S2). Intact protein mass spectra showed no evidence of truncation or misrecognition of the amber stop codon by glutaminyl-tRNA[68] (Figures S4 and S5). The reactivity toward intracellular levels of free L-Cys was negligible, whereas ligation proceeded readily and quantitatively at high L-Cys concentrations (Figure 1E). High cell-density fermentation produced up to 1.8 g/L of CysCNP-containing protein, which is the highest yield achieved in our lab for ncAA incorporation, without compromising the fidelity of amber suppression (Figures S5 and S6). The success of our strategy attests to the speed of biosynthesis of CysCNP, which delivers the amino acid quickly and at sufficiently high concentration to sustain a high rate of protein synthesis. The endogenous cysteine biosynthetic machinery of *E. coli* is ubiquitous in bacteria, eliminating the need for complex cellular metabolic engineering.[49]

Encouraged by these results, we further assessed the capacity of our system to utilize CNPym-thiol, which is the commercially available pyrimidine analog of CNP-thiol. Given their bioisosteric relationship, we hypothesized that the CNPym-thiol would also be accepted by the cysteine biosynthetic enzymes as well as the G1(CysCNP)RS selected for CysCNP.[69] The second nitrogen atom in the six-membered ring of CysCNPym (Figure 1B) enhances the electrophilicity of the nitrile group and consequently increases the rate of the NAT click reaction.[70]

By supplementing the culture media of *E. coli* cells harboring the pCDF-RFP amber/pRSF-G1(CysCNP)RS dual-plasmid system with 0.5 mM CNPym-thiol, we achieved RFP yields comparable to wild-type levels after 16 h expression at 25°C. However, the intact protein mass spectrum showed the expected mass for only one-third of the expressed RFP CysCNPym protein, whereas two-thirds showed the mass of the NAT reaction product with L-Cys (Figures 1E and S15). This indicates that the increased electrophilicity of the nitrile group in CysCNPym permits reaction with free L-Cys at the natural intracellular concentration in bacteria. Nonetheless, the desired protein was obtained in >95% purity when the expression time was reduced to 5 h at 30°C or the expression temperature was lowered to 18°C for overnight expression (Figure 1E and S15). Kinetic *in vitro* analyses showed that CysCNP displayed rapid reactivity in the intramolecular click NAT reaction with a half-life ($t_{1/2}$) of 5.8 min at 37°C (Figure S7). Compared with CysCNP, CysCNPym reacted approximately 40 times faster with 1,2-aminothiols in the intermolecular NAT click reaction and typically completed the reaction in less than 30 min (Figure S8).

Enhanced electrophilicity of the CysCNPym nitrile group enables site-selective lysine ligation

Using a fusion construct between the NT* domain and ubiquitin with a CysCNPym residue in the linker region (NT_Ubi CysCNPym), we compared the NAT click reaction with L-Cys with that of the corresponding CysCNP construct. Unexpectedly, the more reactive CysCNPym mutant produced the expected ligation product in less than 20% yield, whereas the CysCNP mutant under identical reaction conditions yielded complete ligation (Figure 2A and S15). Inspection of the polypeptide segment containing the ncAA (CGKRKS(ncAA)) showed a cysteine-lysine motif. Previous work by Bertozzi and co-workers showed that cysteine-lysine motifs can react irreversibly with

(B) Proposed mechanism for the intramolecular reaction between CysCNPym and the cysteine-lysine motif in NT_Ubi CysCNPym. The cysteine thiol group reversibly reacts with the nitrile group, forming a thioimidate intermediate, which subsequently reacts with a nearby lysine side-chain amine group via an *S*-to-*N* imido transfer, resulting in the formation of a stable amidine macrocyclic product. The NT_Ubi models were generated with AlphaFold Server.[72]

(C) Intermolecular ligation reaction between NT_Ubi CysCNP and NT_Ubi CysCNPym ΣCΣK illustrates sequential bioconjugation and macrocyclization exploiting the difference in electrophilicity of CysCNPym and CysCNP.

(D) Reducing SDS-PAGE analysis of the intermolecular ligation reaction for an equimolar mixture (lane AB) of 0.5 mM NT_Ubi CysCNP (lane A) and NT_Ubi CysCNPym ΣCΣK (lane B) after 24 h incubation in PBS buffer containing 20 mM TCEP at 37°C. The red arrow indicates the intermolecular ligation product **A-B**, and the blue arrow in lane C denotes the purified cyclic Ubi CysCNP/Ubi CysCNPym ΣCΣK conjugate **C-D** generated by TEV proteolytic cleavage of **A-B**. Lane M shows the molecular-weight markers.

(E) Intact protein mass spectrometry analysis of the **C-D** conjugate confirms its formation in addition to the intramolecular macrocyclization between the N-terminal cysteine and CysCNP (expected mass indicated in red).

electron-deficient nitriles to form an amidine adduct between 2-cyanobenzothiazole (CBT) derivatives and the amine group of the lysine side chain via a thioimidate intermediate.[71]

If the nitrile group in NT_Ubi CysCNPym were to undergo a similar intramolecular reaction, where the thioimidate intermediate is formed by nucleophilic attack of the nearby cysteine thiol group on the nitrile group followed by a subsequent *S*-to-*N* imido transfer involving the amine group of the adjacent lysine side chain, a stable intramolecular amidine macrocyclic product would form with no change in protein mass. The ligation between CysCNPym and the lysine ε-NH_2 group would result in a stapled product with a pyrimidinyl-amidine linker (Figure 2B), which would render the product unreactive toward the NAT click reaction with free L-Cys.

To validate this hypothesis, we designed and expressed three NT_Ubi CysCNPym mutants: NT_Ubi CysCNPym ΣC (where the cysteine residue in the motif is mutated to alanine), NT_Ubi CysCNPym ΣK (where the two lysine residues in the peptide motif are mutated to alanine), and NT_Ubi CysCNPym ΣCΣK (where the cysteine and the two lysine residues in the peptide motif are mutated to alanine). Subjecting these mutants to the NAT reaction with L-Cys produced the expected L-Cys adduct in conversion yields of 86%, 84%, and 100%, respectively (Figures 2A and S15). The restoration of CysCNPym nitrile reactivity in these mutants supports the activity of the cysteine-lysine motif in the original NT_Ubi CysCNPym construct and underscores the importance of the cysteine residue in the reaction.

Next, we investigated the feasibility of using the cysteine-lysine motif with CysCNPym for intermolecular bioconjugation without a 1,2-aminothiolate (Figure 2C). Incubation of an equimolar mixture of NT_Ubi CysCNP, which contains the cysteine-lysine motif, and NT_Ubi CysCNPym ΣCΣK at 37°C for 24 h resulted in a new band in the SDS-PAGE with the mass expected for the intermolecular conjugate (Figure 2D). To confirm the expected linkage, we used the TEV protease cleavage sites between the NT* and ubiquitin domains to remove the NT* domains, yielding the Ubi CysCNP/Ubi CysCNPym ΣCΣK conjugate, which underwent an intramolecular macrocyclization by the NAT reaction between the CysCNP residue and the N-terminal cysteine generated by the TEV cleavage. Intact protein mass spectrometry confirmed the expected mass (Figure 2E). The availability of the free thiol group in the cysteine-lysine motif after intermolecular ligation with CysCNPym strongly suggests the site-selective attachment of CysCNPym to the nearby lysine residue in the motif. As demonstrated by the reaction scheme of Figure 2C, the unique capacity of CysCNPym for the cysteine-lysine motif opens a route to controlling the order and sites of post-translational modifications in proteins, where CysCNPym is used for one ligation with lysine and CysCNP for another ligation with a 1,2-aminothiol.

Protein-protein ligation and cell surface labeling via CysCNPym reactivity with genetically encoded 1,2-aminothiols

The high purity and yield of genetically encoded CysCNPym-containing proteins and the high reactivity of the CysCNPym nitrile group supports efficient protein-protein ligation via the biocompatible NAT click reaction. To demonstrate this approach, we explored the capacity of this reaction for labeling a CysCNPym residue with an RFP mutant containing an N-terminal cysteine residue generated by TEV proteolytic cleavage of a leader sequence (C-RFP). Equimolar mixtures of C-RFP with RFP CysCNPym were incubated at 4°C, 25°C, and 37°C (Figure 3A), and the reaction was monitored by SDS-PAGE (Figures 3B and S9). A new band corresponding to the RFP CysCNPym/C-RFP conjugate appeared at the expected mass (Figures 3B and 3C). After 4 h, the reaction reached a conversion yield of 30%–34% with no significant difference between the different temperature conditions. The ability of the CysCNPym-mediated NAT reaction to proceed at 4°C renders this technique particularly advantageous for ligation of temperature-sensitive recombinant proteins.

The reactivity of the CysCNP and CysCNPym amino acids invites applications for selective cell surface labeling with 1,2-aminothiol-bearing molecules. To display the nitrile functionality on the cell surface, we introduced CysCNPym into the *E. coli* outer-membrane protein OmpX (OmpX CysCNPym) between residues S53 and S54 in the extracellular OmpX loop (Figure 3D).[73] Incubation of *E. coli* cells displaying OmpX CysCNPym with C-RFP at 25°C for 4 h resulted in successful fluorescence labeling of the cell surface, as confirmed by flow cytometry and confocal fluorescence microscopy. Flow cytometry analysis revealed that more than 94% of cells displaying OmpX CysCNPym were strongly fluorescently labeled even when they reacted with only 20 μM C-RFP, whereas less than 1% of cells devoid of the ncAAs were labeled after incubation with C-RFP under the same conditions (Figure 3E). Confocal fluorescence microscopy imaging clearly demonstrated the localization of C-RFP-labeled OmpX on the cell surface of the labeled cells (Figures 3F and S10).

Tunable fluorogenic protein labeling via the arcNAT click reaction between nitrile-bearing ncAAs and ABTs

Screening different 1,2-aminothiols for NAT click reactivity, we observed that not only aliphatic 1,2-aminothiols but also ABT readily reacted with mCNP residues in proteins under ambient, catalyst-free conditions (Figures 4A and S11). This arcNAT reaction yielded a 2-pyridinyl benzothiazole derivative, extending the number of conjugated double bonds and increasing the fluorescence excitation and emission wavelengths ($^{max}\lambda_{Ex}$ = 320 nm and $^{max}\lambda_{Em}$ = 386 nm; Figure S11). ABT has previously been reported to react with aromatic heterocyclic nitriles to form the benzothiazole derivative but only with catalysis by metal ions and under elevated temperatures or microwave irradiation.[74–77] The benzothiazole product presents a D-π-A push-pull-type chromophore, where the benzothiazole and pyridinyl rings function as electron-donor (D) and electron-acceptor (A) groups, respectively, linked by a π-conjugated system.[78–81] The UV spectroscopic properties of D-π-A chromophores are governed by intramolecular charge transfer (ICT) following photon absorption, allowing for tunable absorption and emission properties through variations in the D and A groups.[82,83]

We hypothesized that the fluorescent properties of the pyridinyl benzothiazole groups resulting from the arcNAT reaction could be tuned in a modular fashion by modifications of the electronic properties of the nitrile-bearing ncAA (Nit_x), the ABT

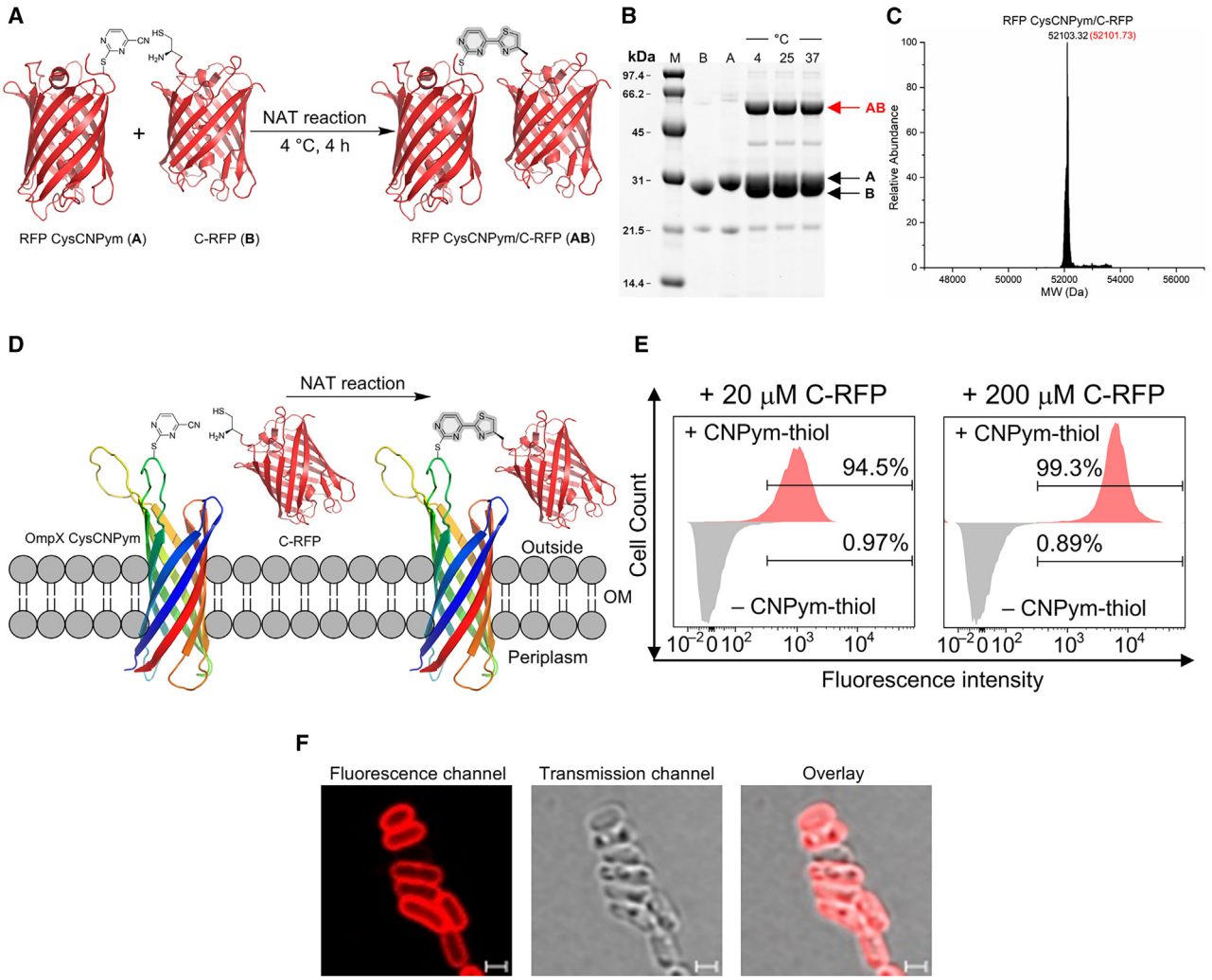


Figure 3. Site-specific protein-protein ligation and cell-surface labeling via the NAT click reaction

(A) Reaction scheme for site-specific conjugation. The CysCNPym residue of RFP CysCNPym is near the N terminus. C-RFP is produced by the cleavage of a modified TEV protease recognition sequence to expose an N-terminal cysteine residue. Incubation of both RFP versions in PBS buffer containing 20 mM TCEP at 4°C yields the RFP CysCNPym/C-RFP conjugate via the NAT click reaction. The RFP models were generated from PDB: 2QLG.

(B) Monitoring the reaction by reducing SDS-PAGE. The reaction mixture contained 300 μM RFP CysCNPym (lane A) and C-RFP (lane B). After 4 h incubation at different temperatures (4°C, 25°C, and 37°C), the new band of the RFP CysCNPym/C-RFP conjugate (**AB**; red arrow) demonstrates ligation. Lane M: markers of molecular weight.

(C) Intact protein mass spectrometry analysis of the purified conjugate **AB** confirms its successful formation. The expected mass is in red.

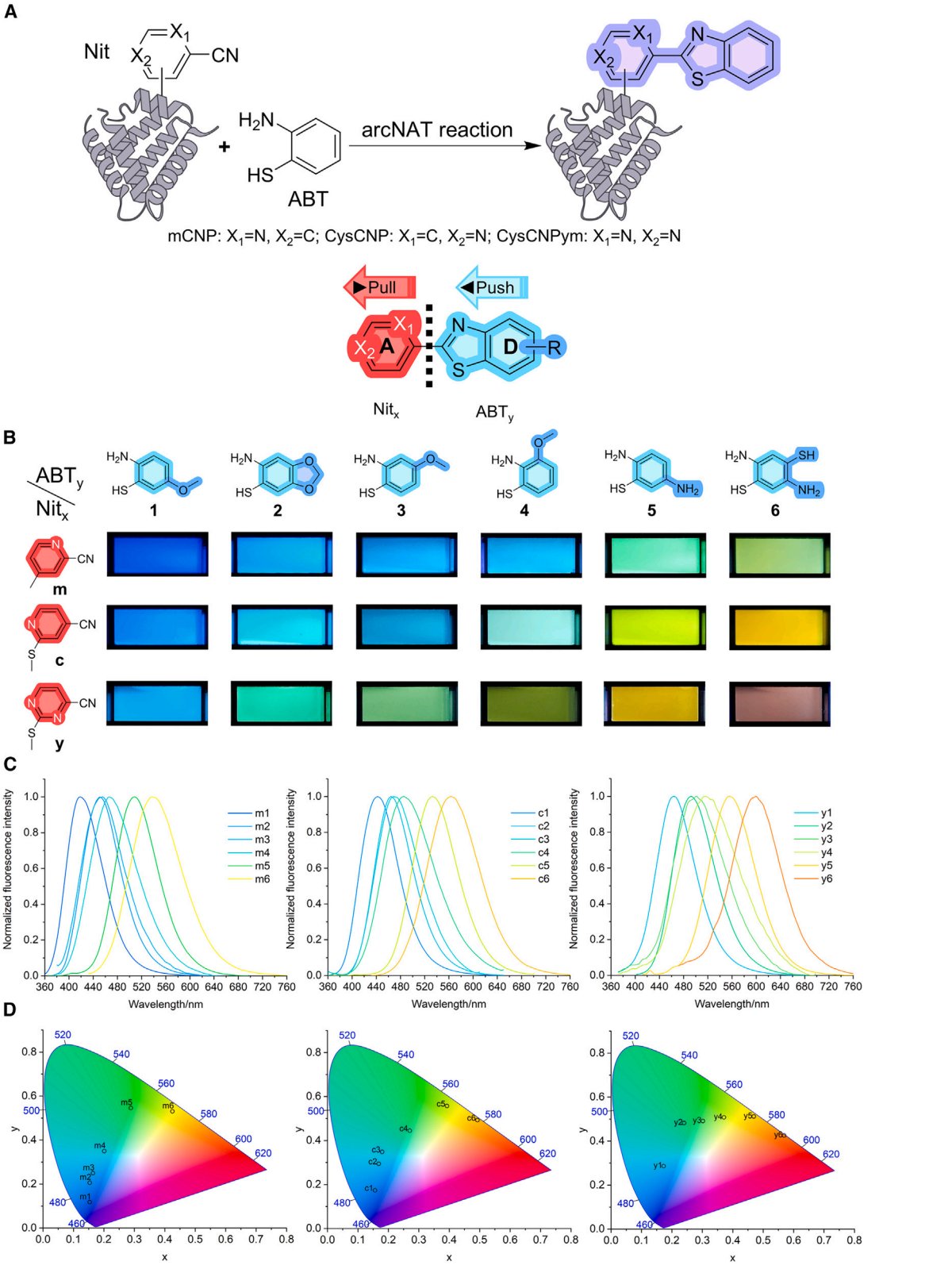
(D) Cell-surface labeling using the NAT click reaction. OmpX CysCNPym displays the CysCNPym residue in the extracellular space for the reaction with C-RFP. The OmpX model was generated from PDB: 2M06.

(E) Flow cytometry analysis of *E. coli* cells displaying OmpX CysCNPym (+CNPym-thiol) after 4 h incubation with 20 μM (left) or 200 μM (right) C-RFP at 25°C. Incubation of the negative control (*E. coli* cells-CNPym-thiol) with C-RFP under the same conditions shows no labeling (Figure S10).

(F) Confocal fluorescence microscopy images of live *E. coli* cells displaying OmpX CysCNPym after 4 h incubation with 200 μM C-RFP at 25°C confirm successful cell-surface fluorescent labeling. Fluorescence imaging used a 561 nm excitation laser (scale bar: 1 μm). The uncropped images are provided in Figure S10.

module (ABT_y), or both (Figure 4A). A library of commercially available aryl and heteroaryl 2-aminothiol building blocks (>100 compounds with diverse substituent types and numbers) facilitated extensive exploration of the D substituent effects on the fluorophores' photophysical properties. Regarding the Nit_x-ABT_y reaction products, we expected that differences in the electron-withdrawing ability between the pyridine and pyrimidine moieties, the position of the A center in relation to the D group, and the presence of the thioether linkage would increase the overall dipole moments of the chromophores and redshift the light absorption and fluorescent emission of the push-pull system (Figure 4A).[82,84,85]

To explore the tunability of fluorescent product in the arcNAT reaction, we investigated the reaction between mCNP, CysCNP, or CysCNPym incorporated into NT_Ubi and 13 commercially available aryl and heteroaryl ABTs (Table S4). Six of the tested



(legend on next page)

ABTs showed the formation of the fluorescent benzothiazole derivatives with diverse photophysical properties (Figures 4B–4D and S16; Table S3). As predicted, compared with the unsubstituted ABT, ABTs with strong electron-donating substituents (R = methoxy, methylenedioxy, amino) yielded redshifted chromophores with large Stokes shifts (82–178 nm) (Table S3). The large Stokes shifts of these chromophores are beneficial for fluorescence imaging because the substantial separation between the excitation and emission maxima reduces self-absorption, thereby minimizing the inner-filter effect.[86,87] In contrast, ABTs with electron-withdrawing substituents (R = cyano, carboxyl) produced non-fluorescent benzothiazole derivatives (Figure S16). No detectable arcNAT reaction products were observed with heteroaromatic ABTs (Table S4).

Modifications of the nitrile module significantly altered the fluorescence properties of the resulting chromophores, specifically by eliciting a bathochromic shift in $^{max}\lambda_{Em}$. This shift correlated with the increasing electron-withdrawing ability of the heteroaromatic ring bearing the nitrile group, as quantified by their Hammett σ_p substitution constants: 2-pyridyl = 0.17 (mCNP) < 4-pyridyl = 0.44 (CysCNP) < 4-pyrimidinyl = 0.63 (CysCNPym).[88] Most importantly, the fluorogenic arcNAT click reaction exhibited a remarkable fluorescence turn-on effect with a fluorescence ON/OFF ratio exceeding 1,000-fold in the formation of chromophore **y6** (Figure S12). Kinetic studies revealed that the arcNAT reaction proceeded at rates comparable to those of strain-promoted [3 + 2] azide-alkyne cycloaddition (SPAAC) bioorthogonal reactions, albeit at the slower end of the range (Figure S8).[89]

Application of the arcNAT click reaction for wash-free, live-cell fluorescent labeling

The combination of a high signal-to-noise ratio (ON/OFF ratio) and fast reaction kinetics (Figures S12 and S13 and Video S1), especially for chromophores **y2** and **y6**, renders the fluorogenic arcNAT click reaction an attractive candidate for live-cell fluorescent labeling without the need for washing off excess reagent, making it highly suitable for both flow cytometry and fluorescence microscopy. For extracellular surface labeling, *E. coli* cells displaying OmpX CysCNPym were incubated with ABT **2** or **6** at 25°C for 4 h, resulting in cell-surface labeling with chromophore **y2** or **y6**, respectively. Confocal fluorescence microscopy confirmed successful cell surface labeling without the removal of the tagging solution and without any washing steps with labeled cell fractions of 93.7% and 84.5%, respectively (Figure 5).

The arcNAT reaction can also be used for efficiently labeling intracellular proteins containing CysCNPym given the small size of ABTs and thiol-mediated cellular uptake.[90,91] *E. coli* cells expressing NT_Ubi CysCNPym ΣCΣK were incubated with ABT **2** or **6** at 25°C for 4 h. Flow cytometry analysis showed that the majority of cells (99.5% or 99.5%, respectively) had a distinctly different fluorescent profile from that of the control cells. Wash-free confocal fluorescence microscopy imaging confirmed the chromophore formation and intracellular localization of the arcNAT click reaction products, chromophores **y2** and **y6**, demonstrating the reaction efficiency and specificity in a live-cell context (Figure 5). Cell viability and morphology were not significantly affected by the ncAA-forming thiol compounds (CNP-thiol and CNPym-thiol) or ABTs (**2** and **6**) at concentrations used for protein expression and live-cell labeling (Figures S17–S19). These results highlight the potential of the fluorogenic arcNAT click reaction as a powerful tool for live-cell labeling by offering advantages in terms of simplicity, speed, tunability, and biocompatibility.

Conclusions

Coupling the biosynthesis of ncAAs with GCE enables the cost-effective production of proteins with a handle with selective chemical properties. The activated nitrile-bearing ncAAs of the present work afford a flexible and highly efficient platform for site-specific post-translational modifications of proteins. The approach expands the applications of NAT chemistry to include protein-protein cross-linking and turn-on fluorogenic labeling of proteins for live-cell imaging. CysCNPym stands out as a ncAA that strikes a balance between high reactivity toward 1,2-aminothiols and sufficient resilience toward intracellular concentrations of thiol and amine nucleophiles to permit genetic encoding of the unmodified ncAA. The purity and yield with which proteins can be produced with CysCNP and CysCNPym highlight the efficiency with which thiol compounds are taken up by cells and processed by the cysteine biosynthetic machinery,[90,91] as well as the extraordinary activity of the G1(CysCNP)RS variant identified in this work.

In addition, the aromatic condensation reaction between activated nitrile-bearing ncAAs and electron-rich ABT—arcNAT—readily proceeds under physiological conditions, enabling turn-on fluorogenic labeling of proteins for fluorescence-based applications including live-cell microscopy. As illustrated by the modularity of the arcNAT reaction, the *in situ* NAT click reaction presents a powerful platform for increasing the chemical diversity of peptides and proteins for applications in medicinal chemistry and chemical biology. Protein libraries constructed from activated nitriles and chemically diverse 2-aminothiols promise to significantly expand the chemical space available for protein chemistry.[92,93]

Figure 4. arcNAT click reaction for tunable fluorogenic protein labeling

(A) arcNAT click reaction between 2-aminobenzenethiol (ABT) and the nitrile group of an ncAA (Nit) to form a benzothiazole with longer fluorescence wavelengths. Modulating the electronic effects (both inductive and mesomeric) of the R group attached to the ABT in the electron-donor (D) module and the heterocyclic ring in the electron-acceptor (A) module can alter the push-pull effect on the fluorophores' excitation and emission properties.

(B) Photographic images of the arcNAT reaction products of NT_Ubi mCNP (**m**), NT_Ubi CysCNP (**c**), or NT_Ubi CysCNPym ΣCΣK (**y**) with various ABTs (**1**–**6**). The images are photographs of 5–50 μM solutions in a 400 μL quartz fluorescence cell and were taken with an ordinary mobile phone camera under 365 nm LED illumination. The intact protein mass spectrometry analysis of the reaction products is shown in Figure S16.

(C and D) Fluorescence emission spectra (C) and CIE 1931 xy chromaticity diagrams (D) of the samples shown in (B) measured at the corresponding $^{max}\lambda_{Ex}$ listed in Table S3.

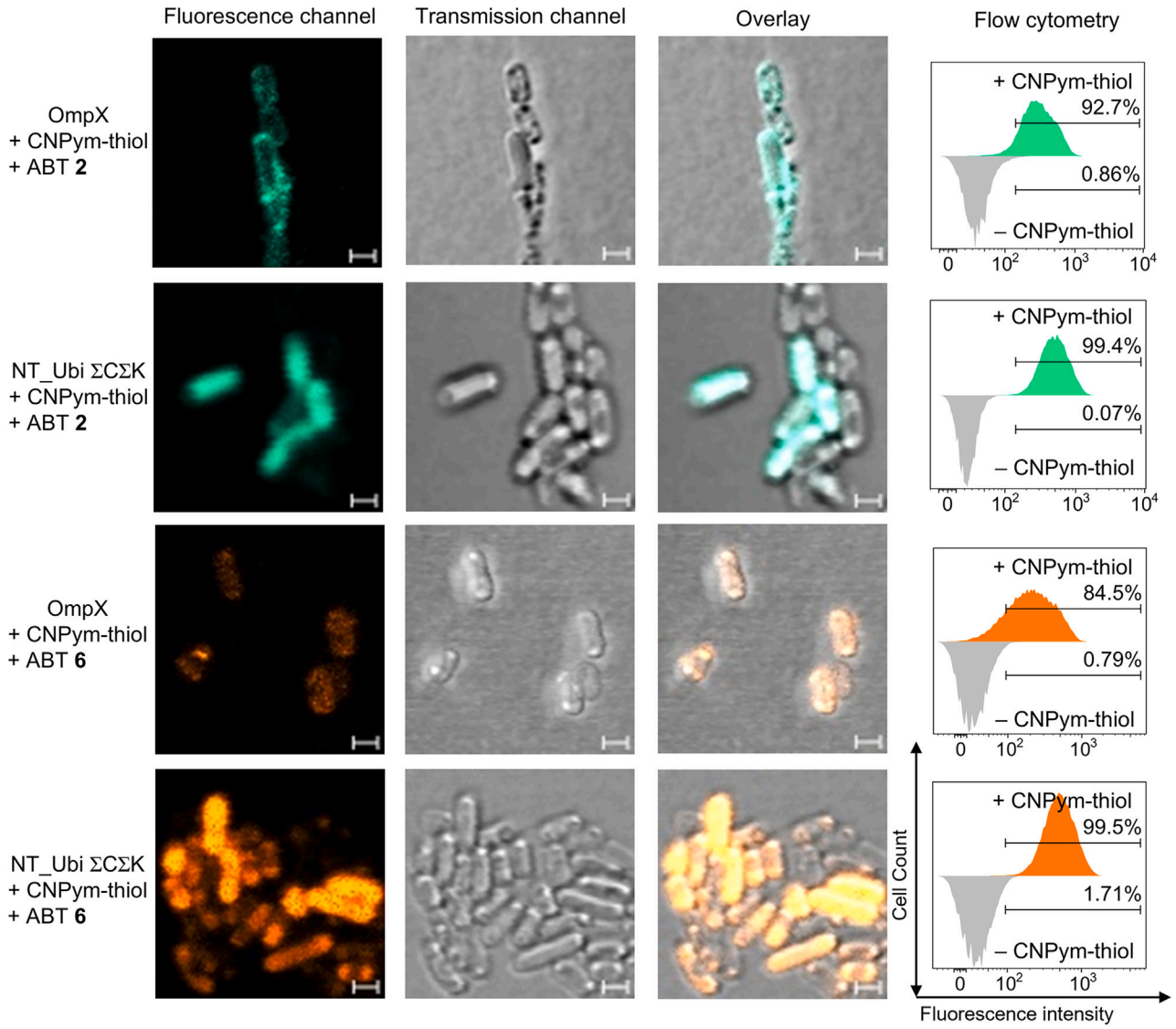


Figure 5. Wash-free, live-cell fluorescent labeling using the fluorogenic arcNAT click reaction

Confocal fluorescence microscopy images and FACS analysis of live *E. coli* cells displaying OmpX CysCNPym (surface labeling) or expressing NT_Ubi CysCNPym ΣCΣK (intracellular labeling) after 4 h incubation with 2 mM ABT **2** or **6** at 25°C demonstrate successful fluorescent labeling (scale bar: 1 μm). Fluorescent images were recorded with 405 and 458 nm excitation lasers. The uncropped images are shown in Figure S14.

METHODS

Detailed methods can be found in the supplemental information.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Thomas Huber (t.huber@anu.edu.au).

Materials availability

The pRSF-G1(CysCNP)RS and pCDF-RFP amber plasmids have been deposited at Addgene (Watertown, MA, USA) (plasmids 225054 and 229524, respectively).

All materials generated in this study are available from the lead contact with a completed materials transfer agreement.

Data and code availability

All data supporting this study are available within the main article and supplemental information. This study did not generate any datasets or original code.

ACKNOWLEDGMENTS

We thank Dr. Harpreet Vohra and Michael Devoy at the John Curtin School of Medical Research of the Australian National University for technical support on the FACS experiments. We acknowledge Microscopy Australia (ROR: 042mm0k03) at the Centre for Advanced Microscopy of the Australian National University, especially Daryl Webb for his knowledge and input. This work was supported by Australian Research Council grants CE200100012 and DP230100079.

AUTHOR CONTRIBUTIONS

E.H.A. conceived the work, designed the study, performed experiments, analyzed results, and wrote the paper. H.Q. designed the study, performed experiments, analyzed results, and wrote the paper. G.O. designed the study, supervised the research, and wrote the paper. T.H. conceived the work, designed the study, supervised the research, and wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chempr.2024.12.003.

Received: September 20, 2024
Revised: November 12, 2024
Accepted: December 5, 2024
Published: January 6, 2025

REFERENCES

1. Jana, S., Evans, E.G.B., Jang, H.S., Zhang, S., Zhang, H., Rajca, A., Gordon, S.E., Zagotta, W.N., Stoll, S., and Mehl, R.A. (2023). Ultrafast bioorthogonal spin-labeling and distance measurements in mammalian cells using small, genetically encoded tetrazine amino acids. J. Am. Chem. Soc. *145*, 14608–14620. https://doi.org/10.1021/jacs.3c00967.
2. Beliu, G., Kurz, A.J., Kuhlemann, A.C., Behringer-Pliess, L., Meub, M., Wolf, N., Seibel, J., Shi, Z.-D., Schnermann, M., Grimm, J.B., et al. (2019). Bioorthogonal labeling with tetrazine-dyes for super-resolution microscopy. Commun. Biol. *2*, 261. https://doi.org/10.1038/s42003-019-0518-z.
3. Seitchik, J.L., Peeler, J.C., Taylor, M.T., Blackman, M.L., Rhoads, T.W., Cooley, R.B., Refakis, C., Fox, J.M., and Mehl, R.A. (2012). Genetically encoded tetrazine amino acid directs rapid site-specific *in vivo* bioorthogonal ligation with *trans*-cyclooctenes. J. Am. Chem. Soc. *134*, 2898–2901. https://doi.org/10.1021/ja2109745.
4. Thomas, J.D., Cui, H., North, P.J., Hofer, T., Rader, C., and Burke, T.R., Jr. (2012). Application of strain-promoted azide–alkyne cycloaddition and tetrazine ligation to targeted Fc-drug conjugates. Bioconjugate Chem. *23*, 2007–2013. https://doi.org/10.1021/bc300052u.
5. Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D.J., Blackman, M.L., Fox, J.M., and Chin, J.W. (2012). Genetic encoding of bicyclononynes and *trans*-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels–Alder reactions. J. Am. Chem. Soc. *134*, 10317–10320. https://doi.org/10.1021/ja302832g.
6. Werther, P., Yserentant, K., Braun, F., Grußmayer, K., Navikas, V., Yu, M., Zhang, Z., Ziegler, M.J., Mayer, C., Gralak, A.J., et al. (2021). Bio-orthogonal red and far-red fluorogenic probes for wash-free live-cell and super-resolution microscopy. ACS Cent. Sci. *7*, 1561–1571. https://doi.org/10.1021/acscentsci.1c00703.
7. Wang, H., Zhang, Y., Zeng, K., Qiang, J., Cao, Y., Li, Y., Fang, Y., Zhang, Y., and Chen, Y. (2021). Selective mitochondrial protein labeling enabled by biocompatible photocatalytic reactions inside live cells. JACS Au *1*, 1066–1075. https://doi.org/10.1021/jacsau.1c00172.
8. Shi, Y., Fu, L., Yang, J., and Carroll, K.S. (2021). Wittig reagents for chemoselective sulfenic acid ligation enables global site stoichiometry analysis and redox-controlled mitochondrial targeting. Nat. Chem. *13*, 1140–1150. https://doi.org/10.1038/s41557-021-00767-2.
9. Lumen, D., Vugts, D., Chomet, M., Imlimthan, S., Sarparanta, M., Vos, R., Schreurs, M., Verlaan, M., Lang, P.A., Hippeläinen, E., et al. (2022). Pretargeted PET imaging with a TCO-conjugated anti-CD44v6 chimeric mAb U36 and [^{89}Zr]Zr-DFO-PEG$_5$-Tz. Bioconjugate Chem. *33*, 956–968. https://doi.org/10.1021/acs.bioconjchem.2c00164.
10. Oller-Salvia, B., Kym, G., and Chin, J.W. (2018). Rapid and efficient generation of stable antibody–drug conjugates via an encoded cyclopropene and an inverse-electron-demand Diels–Alder reaction. Angew. Chem. Int. Ed. *57*, 2831–2834. https://doi.org/10.1002/anie.201712370.
11. Baskin, J.M., Prescher, J.A., Laughlin, S.T., Agard, N.J., Chang, P.V., Miller, I.A., Lo, A., Codelli, J.A., and Bertozzi, C.R. (2007). Copper-free click chemistry for dynamic *in vivo* imaging. Proc. Natl. Acad. Sci. USA *104*, 16793–16797. https://doi.org/10.1073/pnas.0707090104.
12. Arranz-Gibert, P., Vanderschuren, K., Haimovich, A., Halder, A., Gupta, K., Rinehart, J., and Isaacs, F.J. (2022). Chemoselective restoration of *para*-azido-phenylalanine at multiple sites in proteins. Cell Chem. Biol. *29*, 1046–1052.e4. https://doi.org/10.1016/j.chembiol.2021.12.002.
13. Eddins, A.J., Bednar, R.M., Jana, S., Pung, A.H., Mbengi, L., Meyer, K., Perona, J.J., Cooley, R.B., Karplus, P.A., and Mehl, R.A. (2023). Truncation-free genetic code expansion with tetrazine amino acids for quantitative protein ligations. Bioconjugate Chem. *34*, 2243–2254. https://doi.org/10.1021/acs.bioconjchem.3c00380.
14. Jewett, J.C., and Bertozzi, C.R. (2011). Synthesis of a fluorogenic cyclooctyne activated by Cu-free click chemistry. Org. Lett. *13*, 5937–5939. https://doi.org/10.1021/ol2025026.
15. Wilkovitsch, M., Haider, M., Sohr, B., Herrmann, B., Klubnick, J., Weissleder, R., Carlson, J.C.T., and Mikula, H. (2020). A cleavable C_2-symmetric *trans*-cyclooctene enables fast and complete bioorthogonal disassembly of molecular probes. J. Am. Chem. Soc. *142*, 19132–19141. https://doi.org/10.1021/jacs.0c07922.
16. Sondag, D., Maartense, L., de Jong, H., de Kleijne, F.F.J., Bonger, K.M., Löwik, D.W.P.M., Boltje, T.J., Dommerholt, J., White, P.B., Blanco-Ania, D., and Rutjes, F.P.J.T. (2023). Readily accessible strained difunctionalized *trans*-cyclooctenes with fast click and release capabilities. Chemistry *29*, e202203375. https://doi.org/10.1002/chem.202203375.
17. Gonçalves, M.S.T. (2009). Fluorescent labeling of biomolecules with organic probes. Chem. Rev. *109*, 190–212. https://doi.org/10.1021/cr0783840.
18. Richardson, M.B., Brown, D.B., Vasquez, C.A., Ziller, J.W., Johnston, K.M., and Weiss, G.A. (2018). Synthesis and explosion hazards of 4-azido-L-phenylalanine. J. Org. Chem. *83*, 4525–4536. https://doi.org/10.1021/acs.joc.8b00270.
19. Nitsche, C., Onagi, H., Quek, J.-P., Otting, G., Luo, D., and Huber, T. (2019). Biocompatible macrocyclization between cysteine and 2-cyanopyridine generates stable peptide inhibitors. Org. Lett. *21*, 4709–4712. https://doi.org/10.1021/acs.orglett.9b01545.
20. Abdelkader, E.H., Qianzhu, H., George, J., Frkic, R.L., Jackson, C.J., Nitsche, C., Otting, G., and Huber, T. (2022). Genetic encoding of cyanopyridylalanine for in-cell protein macrocyclization by the nitrile–aminothiol click reaction. Angew. Chem. Int. Ed. *61*, e202114154. https://doi.org/10.1002/anie.202114154.
21. Hempfling, J.P., Sekera, E.R., Sarkar, A., Hummon, A.B., and Pei, D. (2022). Generation of proteins with free N-terminal cysteine by aminopeptidases. J. Am. Chem. Soc. *144*, 21763–21771. https://doi.org/10.1021/jacs.2c10194.
22. Tolbert, T.J., and Wong, C.-H. (2002). New methods for proteomic research: preparation of proteins with N-terminal cysteines for labeling and conjugation. Angew. Chem. Int. Ed. *41*, 2171–2174. https://doi.org/10.1002/1521-3773(20020617)41:12<2171::AID-ANIE2171>3.0.CO;2-Q.
23. Liu, D., Xu, R., Dutta, K., and Cowburn, D. (2008). N-terminal cysteinyl proteins can be prepared using thrombin cleavage. FEBS Lett. *582*, 1163–1167. https://doi.org/10.1016/j.febslet.2008.02.078.
24. Nguyen, D.P., Elliott, T., Holt, M., Muir, T.W., and Chin, J.W. (2011). Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. J. Am. Chem. Soc. *133*, 11418–11421. https://doi.org/10.1021/ja203111c.

25. Tai, J., Wang, L., Chan, W.S., Cheng, J., Chan, Y.H., Lee, M.M., and Chan, M.K. (2023). Pyrrolysine-inspired in cellulo synthesis of an unnatural amino acid for facile macrocyclization of proteins. J. Am. Chem. Soc. *145*, 10249–10258. https://doi.org/10.1021/jacs.3c01291.

26. Schneider, A.F.L., and Hackenberger, C.P.R. (2017). Fluorescent labelling in living cells. Curr. Opin. Biotechnol. *48*, 61–68. https://doi.org/10.1016/j.copbio.2017.03.012.

27. Pan, B., Gardner, S.M., Schultz, K., Perez, R.M., Deng, S., Shimogawa, M., Sato, K., Rhoades, E., Marmorstein, R., and Petersson, E.J. (2023). Semi-synthetic CoA-α-synuclein constructs trap N-terminal acetyltransferase NatB for binding mechanism studies. J. Am. Chem. Soc. *145*, 14019–14030. https://doi.org/10.1021/jacs.3c03887.

28. Zhao, J., Elgeti, M., O'Brien, E.S., Sár, C.P., Ei Daibani, A., Heng, J., Sun, X., White, E., Che, T., Hubbell, W.L., et al. (2024). Ligand efficacy modulates conformational dynamics of the μ-opioid receptor. Nature *629*, 474–480. https://doi.org/10.1038/s41586-024-07295-2.

29. Wang, H.-Y., Yang, H., Holm, M., Tom, H., Oltion, K., Al-Khdhairawi, A.A.Q., Weber, J.-F.F., Blanchard, S.C., Ruggero, D., and Taunton, J. (2022). Synthesis and single-molecule imaging reveal stereospecific enhancement of binding kinetics by the antitumour eEF1A antagonist SR-A3. Nat. Chem. *14*, 1443–1450. https://doi.org/10.1038/s41557-022-01039-3.

30. Feng, R.-R., Wang, M., Zhang, W., and Gai, F. (2024). Unnatural amino acids for biological spectroscopy and microscopy. Chem. Rev. *124*, 6501–6542. https://doi.org/10.1021/acs.chemrev.3c00944.

31. Reinhardt, S.C.M., Masullo, L.A., Baudrexel, I., Steen, P.R., Kowalewski, R., Eklund, A.S., Strauss, S., Unterauer, E.M., Schlichthaerle, T., Strauss, M.T., et al. (2023). Ångström-resolution fluorescence microscopy. Nature *617*, 711–716. https://doi.org/10.1038/s41586-023-05925-9.

32. Lincoln, R., Bossi, M.L., Remmel, M., D'Este, E., Butkevich, A.N., and Hell, S.W. (2022). A general design of caging-group-free photoactivatable fluorophores for live-cell nanoscopy. Nat. Chem. *14*, 1013–1020. https://doi.org/10.1038/s41557-022-00995-0.

33. Aktalay, A., Khan, T.A., Bossi, M.L., Belov, V.N., and Hell, S.W. (2023). Photoactivatable carbo- and silicon-thodamines and their application in MINFLUX nanoscopy. Angew. Chem. Int. Ed. *62*, e202302781. https://doi.org/10.1002/anie.202302781.

34. Mulhall, E.M., Gharpure, A., Lee, R.M., Dubin, A.E., Aaron, J.S., Marshall, K.L., Spencer, K.R., Reiche, M.A., Henderson, S.C., Chew, T.-L., et al. (2023). Direct observation of the conformational states of PIEZO1. Nature *620*, 1117–1125. https://doi.org/10.1038/s41586-023-06427-4.

35. Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide to choosing fluorescent proteins. Nat. Methods *2*, 905–909. https://doi.org/10.1038/nmeth819.

36. Marks, K.M., and Nolan, G.P. (2006). Chemical labeling strategies for cell biology. Nat. Methods *3*, 591–596. https://doi.org/10.1038/nmeth906.

37. Crivat, G., and Taraska, J.W. (2012). Imaging proteins inside cells with fluorescent tags. Trends Biotechnol. *30*, 8–16. https://doi.org/10.1016/j.tibtech.2011.08.002.

38. Speight, L.C., Muthusamy, A.K., Goldberg, J.M., Warner, J.B., Wissner, R.F., Willi, T.S., Woodman, B.F., Mehl, R.A., and Petersson, E.J. (2013). Efficient synthesis and in vivo incorporation of acridon-2-ylalanine, a fluorescent amino acid for lifetime and Förster resonance energy transfer/luminescence resonance energy transfer studies. J. Am. Chem. Soc. *135*, 18806–18814. https://doi.org/10.1021/ja403247j.

39. Wang, J., Xie, J., and Schultz, P.G. (2006). A genetically encoded fluorescent amino acid. J. Am. Chem. Soc. *128*, 8738–8739. https://doi.org/10.1021/ja062666k.

40. Summerer, D., Chen, S., Wu, N., Deiters, A., Chin, J.W., and Schultz, P.G. (2006). A genetically encoded fluorescent amino acid. Proc. Natl. Acad. Sci. USA *103*, 9785–9789. https://doi.org/10.1073/pnas.0603965103.

41. Lee, S., Kim, J., and Koh, M. (2022). Recent advances in fluorescence imaging by genetically encoded non-canonical amino acids. J. Mol. Biol. *434*, 167248. https://doi.org/10.1016/j.jmb.2021.167248.

42. Koch, N.G., and Budisa, N. (2024). Evolution of pyrrolysyl-tRNA synthetase: from methanogenesis to genetic code expansion. Chem. Rev. *124*, 9580–9608. https://doi.org/10.1021/acs.chemrev.4c00031.

43. Jann, C., Giofré, S., Bhattacharjee, R., and Lemke, E.A. (2024). Cracking the code: reprogramming the genetic script in prokaryotes and eukaryotes to harness the power of noncanonical amino acids. Chem. Rev. *124*, 10281–10362. https://doi.org/10.1021/acs.chemrev.3c00878.

44. Yi, H.B., Lee, S., Seo, K., Kim, H., Kim, M., and Lee, H.S. (2024). Cellular and biophysical applications of genetic code expansion. Chem. Rev. *124*, 7465–7530. https://doi.org/10.1021/acs.chemrev.4c00112.

45. Bednar, R.M., Jana, S., Kuppa, S., Franklin, R., Beckman, J., Antony, E., Cooley, R.B., and Mehl, R.A. (2021). Genetic incorporation of two mutually orthogonal bioorthogonal amino acids that enable efficient protein dual-labeling in cells. ACS Chem. Biol. *16*, 2612–2622. https://doi.org/10.1021/acschembio.1c00649.

46. Mattheisen, J.M., Wollowitz, J.S., Huber, T., and Sakmar, T.P. (2023). Genetic code expansion to enable site-specific bioorthogonal labeling of functional G protein-coupled receptors in live cells. Protein Sci. *32*, e4550. https://doi.org/10.1002/pro.4550.

47. Meineke, B., Heimgärtner, J., Eirich, J., Landreh, M., and Elsässer, S.J. (2020). Site-specific incorporation of two ncAAs for two-color bioorthogonal labeling and crosslinking of proteins on live mammalian cells. Cell Rep. *31*, 107811. https://doi.org/10.1016/j.celrep.2020.107811.

48. Tang, H., Zhang, P., and Luo, X. (2022). Recent technologies for genetic code expansion and their implications on synthetic biology applications. J. Mol. Biol. *434*, 167382. https://doi.org/10.1016/j.jmb.2021.167382.

49. Wang, Y., Chen, X., Cai, W., Tan, L., Yu, Y., Han, B., Li, Y., Xie, Y., Su, Y., Luo, X., and Liu, T. (2021). Expanding the structural diversity of protein building blocks with noncanonical amino acids biosynthesized from aromatic thiols. Angew. Chem. Int. Ed. *60*, 10040–10048. https://doi.org/10.1002/anie.202014540.

50. Exner, M.P., Kuenzl, T., To, T.M.T., Ouyang, Z., Schwagerus, S., Hoesl, M.G., Hackenberger, C.P.R., Lensen, M.C., Panke, S., and Budisa, N. (2017). Design of *S*-allylcysteine in situ production and incorporation based on a novel pyrrolysyl-tRNA synthetase variant. Chembiochem *18*, 85–90. https://doi.org/10.1002/cbic.201600537.

51. Tamura, T., Inoue, M., Yoshimitsu, Y., Hashimoto, I., Ohashi, N., Tsumura, K., Suzuki, K., Watanabe, T., and Hohsaka, T. (2022). Chemical synthesis and cell-free expression of thiazoline ring-bridged cyclic peptides and their properties on biomembrane permeability. Bull. Chem. Soc. Jpn. *95*, 359–366. https://doi.org/10.1246/bcsj.20210409.

52. Chen, Y., Jin, S., Zhang, M., Hu, Y., Wu, K.-L., Chung, A., Wang, S., Tian, Z., Wang, Y., Wolynes, P.G., and Xiao, H. (2022). Unleashing the potential of noncanonical amino acid biosynthesis to create cells with precision tyrosine sulfation. Nat. Commun. *13*, 5434. https://doi.org/10.1038/s41467-022-33111-4.

53. Reed, K.B., Brooks, S.M., Wells, J., Blake, K.J., Zhao, M., Placido, K., d'Oelsnitz, S., Trivedi, A., Gadhiyar, S., and Alper, H.S. (2024). A modular and synthetic biosynthesis platform for de novo production of diverse halogenated tryptophan-derived molecules. Nat. Commun. *15*, 3188. https://doi.org/10.1038/s41467-024-47387-1.

54. Rogerson, D.T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaite, A., Hancock, S.M., Huguenin-Dezot, N., Muqit, M.M.K., Fry, A.M., Bayliss, R., and Chin, J.W. (2015). Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. Nat. Chem. Biol. *11*, 496–503. https://doi.org/10.1038/nchembio.1823.

55. Zhang, M.S., Brunner, S.F., Huguenin-Dezot, N., Liang, A.D., Schmied, W.H., Rogerson, D.T., and Chin, J.W. (2017). Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. Nat. Methods *14*, 729–736. https://doi.org/10.1038/nmeth.4302.

56. Almhjell, P.J., Boville, C.E., and Arnold, F.H. (2018). Engineering enzymes for noncanonical amino acid synthesis. Chem. Soc. Rev. *47*, 8980–8997. https://doi.org/10.1039/C8CS00665B.

57. Zhu, P., Stanisheuski, S., Franklin, R., Vogel, A., Vesely, C.H., Reardon, P., Sluchanko, N.N., Beckman, J.S., Karplus, P.A., Mehl, R.A., and Cooley, R.B. (2023). Autonomous synthesis of functional, permanently phosphorylated proteins for defining the interactome of monomeric 14–3-3ζ. ACS Cent. Sci. *9*, 816–835. https://doi.org/10.1021/acscentsci.3c00191.

58. Butler, N.D., Sen, S., Brown, L.B., Lin, M., and Kunjapur, A.M. (2023). A platform for distributed production of synthetic nitrated proteins in live bacteria. Nat. Chem. Biol. *19*, 911–920. https://doi.org/10.1038/s41589-023-01338-x.

59. Chen, Y., Loredo, A., Chung, A., Zhang, M., Liu, R., and Xiao, H. (2022). Biosynthesis and genetic incorporation of 3,4-dihydroxy-L-phenylalanine into proteins in *Escherichia coli*. J. Mol. Biol. *434*, 167412. https://doi.org/10.1016/j.jmb.2021.167412.

60. Chen, Y., Tang, J., Wang, L., Tian, Z., Cardenas, A., Fang, X., Chatterjee, A., and Xiao, H. (2020). Creation of bacterial cells with 5-hydroxytryptophan as a 21st amino acid building block. Chem *6*, 2717–2727. https://doi.org/10.1016/j.chempr.2020.07.013.

61. Guo, Y., Cheng, L., Hu, Y., Zhang, M., Liu, R., Wang, Y., Jiang, S., and Xiao, H. (2024). Biosynthesis of halogenated tryptophans for protein engineering using genetic code expansion. Chembiochem *25*, e202400366. https://doi.org/10.1002/cbic.202400366.

62. Kredich, N.M. (2008). Biosynthesis of cysteine. EcoSal Plus *3*. https://doi.org/10.1128/ecosalplus.3.6.1.11.

63. Maier, T.H.P. (2003). Semisynthetic production of unnatural L-α-amino acids by metabolic engineering of the cysteine-biosynthetic pathway. Nat. Biotechnol. *21*, 422–427. https://doi.org/10.1038/nbt807.

64. Qianzhu, H., Abdelkader, E.H., Otting, G., and Huber, T. (2024). Genetic encoding of fluoro-L-tryptophans for site-specific detection of conformational heterogeneity in proteins by NMR spectroscopy. J. Am. Chem. Soc. *146*, 13641–13650. https://doi.org/10.1021/jacs.4c03743.

65. Abdelkader, E.H., Qianzhu, H., Huber, T., and Otting, G. (2023). Genetic encoding of 7-aza-L-tryptophan: isoelectronic substitution of a single CH-group in a protein for a nitrogen atom for site-selective isotope labeling. ACS Sens. *8*, 4402–4406. https://doi.org/10.1021/acssensors.3c01904.

66. Qianzhu, H., Abdelkader, E.H., Herath, I.D., Otting, G., and Huber, T. (2022). Site-specific incorporation of 7-fluoro-L-tryptophan into proteins by genetic encoding to monitor ligand binding by ^{19}F NMR spectroscopy. ACS Sens. *7*, 44–49. https://doi.org/10.1021/acssensors.1c02467.

67. Abdelkader, E.H., Qianzhu, H., Tan, Y.J., Adams, L.A., Huber, T., and Otting, G. (2021). Genetic encoding of N^6-(((trimethylsilyl)methoxy)carbonyl)-L-lysine for NMR studies of protein–protein and protein–ligand interactions. J. Am. Chem. Soc. *143*, 1133–1143. https://doi.org/10.1021/jacs.0c11971.

68. Roy, B., Leszyk, J.D., Mangus, D.A., and Jacobson, A. (2015). Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3. Proc. Natl. Acad. Sci. USA *112*, 3038–3043. https://doi.org/10.1073/pnas.1424127112.

69. Meanwell, N.A. (2011). Synopsis of some recent tactical application of bioisosteres in drug design. J. Med. Chem. *54*, 2529–2591. https://doi.org/10.1021/jm1013693.

70. Oballa, R.M., Truchon, J.-F., Bayly, C.I., Chauret, N., Day, S., Crane, S., and Berthelette, C. (2007). A generally applicable method for assessing the electrophilicity and reactivity of diverse nitrile-containing compounds. Bioorg. Med. Chem. Lett. *17*, 998–1002. https://doi.org/10.1016/j.bmcl.2006.11.044.

71. Keyser, S.G.L., Utz, A., and Bertozzi, C.R. (2018). Computation-guided rational design of a peptide motif that reacts with cyanobenzothiazoles via internal cysteine–lysine relay. J. Org. Chem. *83*, 7467–7479. https://doi.org/10.1021/acs.joc.8b00625.

72. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

73. Lee, Y.-J., Wu, B., Raymond, J.E., Zeng, Y., Fang, X., Wooley, K.L., and Liu, W.R. (2013). A genetically encoded acrylamide functionality. ACS Chem. Biol. *8*, 1664–1670. https://doi.org/10.1021/cb400267m.

74. Zadeh, F.G., Asadi, B., Mohammadpoor-Baltork, I., Tangestaninejad, S., Mirkhani, V., Moghadam, M., and Omidvar, A. (2023). Triazine diphosphonium tetrachloroferrate ionic liquid immobilized on functionalized halloysite nanotubes as an efficient and reusable catalyst for the synthesis of mono-, bis- and tris-benzothiazoles. RSC Adv. *13*, 31213–31223. https://doi.org/10.1039/D3RA05491H.

75. Sondhi, S.M., Arya, S., and Rani, R. (2012). Microwave-assisted conversion of aromatic heterocyclic nitriles to various heterocyclic molecules. Green Chem. Lett. Rev. *5*, 409–414. https://doi.org/10.1080/17518253.2011.643827.

76. Dhawale, K.D., Ingale, A.P., Shinde, S.V., Thorat, N.M., and Patil, L.R. (2021). ZnO-NPs catalyzed condensation of 2-aminothiophenol and aryl/alkyl nitriles: efficient green synthesis of 2-substituted benzothiazoles. Synth. Commun. *51*, 1–14. https://doi.org/10.1080/00397911.2021.1894577.

77. Nale, D.B., and Bhanage, B.M. (2015). N-substituted formamides as C_1-sources for the synthesis of benzimidazole and benzothiazole derivatives by using zinc catalysts. Synlett *26*, 2835–2842. https://doi.org/10.1055/s-0035-1560319.

78. Wang, H., Chen, G., Xu, X., Chen, H., and Ji, S. (2010). The synthesis and optical properties of benzothiazole-based derivatives with various π-electron donors as novel bipolar fluorescent compounds. Dyes Pigm. *86*, 238–248. https://doi.org/10.1016/j.dyepig.2010.01.010.

79. Zhang, L., Tang, B., and Ding, Y. (2005). Study of 2-(2-pyridyl) benzothiazoline as a novel fluorescent probe for the identification of superoxide anion radicals and the determination of superoxide dismutase activity in scallion genus foods. J. Agric. Food Chem. *53*, 549–553. https://doi.org/10.1021/jf049724a.

80. Hrobárik, P., Sigmundová, I., and Zahradník, P. (2005). Preparation of novel push-pull benzothiazole derivatives with reverse polarity: compounds with potential non-linear optic application. Synthesis *2005*, 600–604. https://doi.org/10.1055/s-2004-837309.

81. Hirose, W., Sato, K., and Matsuda, A. (2011). Fluorescence properties of 5-(5,6-dimethoxybenzothiazol-2-yl)-2′-deoxyuridine (d^{bt}U) and oligodeoxyribonucleotides containing d^{bt}U. Eur. J. Org. Chem. *2011*, 6206–6217. https://doi.org/10.1002/ejoc.201100818.

82. Bureš, F. (2014). Fundamental aspects of property tuning in push–pull molecules. RSC Adv. *4*, 58826–58851. https://doi.org/10.1039/C4RA11264D.

83. Rybczyński, P., Bousquet, M.H.E., Kaczmarek-Kędziera, A., Jędrzejewska, B., Jacquemin, D., and Ośmiałowski, B. (2022). Controlling the fluorescence quantum yields of benzothiazole-difluoroborates by optimal substitution. Chem. Sci. *13*, 13347–13360. https://doi.org/10.1039/D2SC05044G.

84. de Silva, A.P., Gunaratne, H.Q.N., and Lynch, P.L.M. (1995). Luminescence and charge transfer. Part 4. 'On–off' fluorescent PET (photoinduced electron transfer) sensors with pyridine receptors: 1,3-diaryl-5-pyridyl-4,5-dihydropyrazoles. J. Chem. Soc. Perkin Trans. 2 *4*, 685–690. https://doi.org/10.1039/P29950000685.

85. Wong, B.A., Friedle, S., and Lippard, S.J. (2009). Solution and fluorescence properties of symmetric dipicolylamine-containing dichlorofluorescein-based Zn^{2+} sensors. J. Am. Chem. Soc. *131*, 7142–7152. https://doi.org/10.1021/ja900980u.

86. Santos, E.M., Sheng, W., Esmatpour Salmani, R., Tahmasebi Nick, S., Ghanbarpour, A., Gholami, H., Vasileiou, C., Geiger, J.H., and Borhan, B. (2021). Design of large Stokes shift fluorescent proteins based on excited state proton transfer of an engineered photobase. J. Am. Chem. Soc. *143*, 15091–15102. https://doi.org/10.1021/jacs.1c05039.

87. Likhotkin, I., Lincoln, R., Bossi, M.L., Butkevich, A.N., and Hell, S.W. (2023). Photoactivatable large Stokes shift fluorophores for multicolor nanoscopy. J. Am. Chem. Soc. *145*, 1530–1534. https://doi.org/10.1021/jacs.2c12567.

88. Hansch, C., Leo, A., and Taft, R.W. (1991). A survey of Hammett substituent constants and resonance and field parameters. Chem. Rev. *91*, 165–195. https://doi.org/10.1021/cr00002a004.

89. Oliveira, B.L., Guo, Z., and Bernardes, G.J.L. (2017). Inverse electron demand Diels–Alder reactions in chemical biology. Chem. Soc. Rev. *46*, 4895–4950. https://doi.org/10.1039/C7CS00184C.

90. Torres, A.G., and Gait, M.J. (2012). Exploiting cell surface thiols to enhance cellular uptake. Trends Biotechnol. *30*, 185–190. https://doi.org/10.1016/j.tibtech.2011.12.002.

91. Laurent, Q., Martinent, R., Lim, B., Pham, A.-T., Kato, T., López-Andarias, J., Sakai, N., and Matile, S. (2021). Thiol-mediated uptake. JACS Au *1*, 710–728. https://doi.org/10.1021/jacsau.1c00128.

92. Iskandar, S.E., Pelton, J.M., Wick, E.T., Bolhuis, D.L., Baldwin, A.S., Emanuele, M.J., Brown, N.G., and Bowers, A.A. (2023). Enabling genetic code expansion and peptide macrocyclization in mRNA display via a promiscuous orthogonal aminoacyl-tRNA synthetase. J. Am. Chem. Soc. *145*, 1512–1517. https://doi.org/10.1021/jacs.2c11294.

93. Proj, M., Strašek, N., Pajk, S., Knez, D., and Sosič, I. (2023). Tunable heteroaromatic nitriles for selective bioorthogonal click reaction with cysteine. Bioconjugate Chem. *34*, 1271–1281. https://doi.org/10.1021/acs.bioconjchem.3c00163.