Electrochemical Labeling of Hydroxyindoles with Chemoselectivity for Site-Specific Protein Bioconjugation

Conor Loynd[1,2], Soumya Jyoti Singha Roy[1,2], Vincent J. Ovalle[1], Sarah E. Canarelli[1], Atanu Mondal[1], Delilah Jewel[1], Elise D. Ficaretta[1], Eranthie Weerapana[1], Abhishek Chatterjee[1*]

1. Department of Chemistry, Boston College, 2609 Beacon Street, Chestnut Hill, Massachusetts 02467, United States. Tel.: +1-617-552-1778
*Email: abhishek.chatterjee@bc.edu

2. These authors contributed equally

Abstract

Electrochemistry has recently emerged as a powerful approach in small molecule synthesis owing to numerous attractive features, including precise control over fundamental reaction parameters, mild reaction conditions, and innate scalability. Even though these advantages also make it an attractive strategy for chemoselective modification of complex biomolecules such as proteins, such applications remain poorly developed. Here we report an electrochemically-promoted coupling reaction between 5-hydroxytryptophan (5HTP) and simple aromatic amines—electrochemical labeling of hydroxyindoles with chemoselectivity (eCLIC) — that enables site-specific labeling of full-length proteins under mild conditions. Using genetic code expansion technology, the 5HTP residue can be incorporated into predefined sites of a recombinant protein expressed in either prokaryotic or eukaryotic hosts for subsequent eCLIC labeling. We used the eCLIC reaction to site-specifically label various recombinant proteins, including a full-length human antibody. Furthermore, we show that eCLIC is compatible with strain-promoted alkyne-azide and alkene-tetrazine click reactions, enabling site-specific modification of proteins at two different sites with distinct labels.

Introduction

Electrochemistry has long played an important role in preparative chemical processes, with some early applications dating back to the 19th century.[1,2] It offers many unique advantages, including exquisite control over the fundamental parameters driving reaction outcome through facile manipulation of electrode potential, mild reaction conditions, innate scalability, and the ability to generate reactive intermediates *in situ*.[3-6] Although the adoption of electrochemical approaches in synthetic organic chemistry was traditionally somewhat tepid, this field has experienced a renaissance over the last few decades, spawning numerous innovative synthetic methodologies that cleverly build on the unique strengths of electrochemistry.[3-6]

The ability to chemoselectively modify proteins at predefined sites is a powerful technology to understand and engineer protein function.[7-18] It has facilitated countless applications from novel strategies to probe complex biological questions, to the development of sophisticated diagnostics and biotherapeutics. Strategies have been developed for modifying proteins at the termini, at specific canonical amino acid side chains, as well as at the noncanonical amino acid (ncAA) side chains that are co-translationally incorporated at specific sites.[7-20] However, there is continuing demand for new protein labeling methods that are efficient, use readily accessible reagents, provide excellent chemoselectivity, and are compatible with other well-established labeling strategies for introducing multiple, different modifications onto proteins in a controlled manner.[21-24] Many of the unique advantages of electrochemistry, such as its tunability and mild reaction conditions, make it an attractive strategy for developing such chemoselective protein labeling methods.[25] Indeed, recent reports have used electrochemistry to activate small-molecule labeling reagents – such as a phenothiazine radical,[26] triazolinediones,[27] or luminols,[28,29] generated through anodic oxidation – which label tyrosine residues on target proteins (**Figure 1**). Similarly, *in situ* generated

oxoammonium reagents have also been used to label tryptophan residues.[30,31] Electrochemical strategies have also been employed to drive nitration[32] or iodination[33] of tyrosine residues on proteins. Although these examples corroborate the enormous untapped potential of using electrochemical strategies to develop protein labeling methods, two key limitations exist: targeting canonical amino residues such as tyrosine and tryptophan intrinsically compromises the ability to precisely define the site and stoichiometry of protein labeling. Furthermore, reactive radical/electrophilic reagents generated through such strategies may show off-target reactivity,[34,35] particularly in the context of proteins, where the reactivity of amino acid side chains can be significantly enhanced by their microenvironment.

Here we report electrochemical labeling of hydroxyindoles with chemoselectivity (eCLIC) (**Figure 1**), which enables selective labeling of recombinantly expressed proteins at predefined sites. Unlike previous approaches described above that rely on electrochemically activating exogenous pro-electrophilic reagents, which are limited by suboptimal control over the site and stoichiometry of protein labeling, eCLIC uses a two-pronged approach to establish exquisite site-control. First, the noncanonical amino acid (ncAA) 5-hydroxytryptophan (5HTP), which has a significantly lower oxidation potential relative to canonical aromatic amino acids, is introduced into a chosen site of the target protein through nonsense suppression using an engineered tryptophanyl-tRNA synthetase (EcTrpRS)/tRNA pair that we recently developed.[36] Next, the electron-rich 5HTP residue is selectively oxidized at the anode to generate an electrophilic intermediate, which can be efficiently trapped using various aromatic amines to generate stable conjugates. This 'Umpolung' approach[37] to chemoselectively convert the 5HTP residue into a reactive electrophile enables its coupling with nucleophilic aromatic amines, which are inherently unreactive toward proteins, making eCLIC highly chemoselective. The conditions for the eCLIC

reaction were optimized and used to site-specifically modify several different proteins, including a full-length antibody, with good efficiency and selectivity. We also show that the eCLIC reaction is mutually compatible with the strain-promoted azide-alkyne cycloaddition (SPAAC),[38] as well as the inverse-electron demand Diels-Alder (IEDDA) reaction between strained alkenes and tetrazines,[16,39] two of the most widely used bioorthogonal conjugation reactions. Furthermore, we expressed a protein containing both a 5HTP and an azide-ncAA residue at specific sites, and demonstrated that eCLIC and SPAAC can be used sequentially to selectively label each side chain with a distinct cargo to generate a homogeneous dual-labeled conjugate. The eCLIC reaction is compatible with a broad range of aromatic amine substrates, and these reagents, as well as 5HTP, are readily accessible and inexpensive. These advantages establish eCLIC as an attractive method for site-specific protein bioconjugation.

Results and discussion

Development, optimization, and scope of eCLIC

Genetic code expansion (GCE) technology enables site-specific incorporation of noncanonical amino acids (ncAAs) using engineered aminoacyl-tRNA synthetase/tRNA pairs.[40-42] This technology provides an attractive way to precisely introduce a unique chemical functionality into predefined sites of proteins, which can be subsequently activated using electrochemistry with high chemoselectivity for conjugation with suitable bioorthogonal reaction partners. We have recently engineered the *E. coli* tryptophanyl-tRNA synthetase (EcTrpRS)/$tRNA^{EcTrp}$ pair for the site-specific incorporation of 5-hydroxytryptophan (5HTP) into proteins expressed in both *E. coli* and eukaryotic cells with excellent fidelity and efficiency.[22,24,36,43-47] The electron-rich 5-hydroxyindole functionality is otherwise stable, but under mild oxidative conditions, it is known

to readily undergo coupling with itself or other electron-rich aromatic compounds.[48-51] We thought that the ability to selectively trigger an electrochemical oxidation of the 5HTP residue on a protein may provide a plausible path for chemoselective protein labeling. To explore this possibility, we analyzed the behavior of 5HTP using cyclic voltammetry (CV) relative to other redox-active aromatic amino acids such as tyrosine and tryptophan (**Figure 2a**). 5HTP was found to be oxidized at a significantly lower voltage than canonical aromatic amino acids, revealing room for its selective oxidative modification.

Using our engineered EcTrpRS/tRNA^{EcTrp} pair, we incorporated a 5HTP residue at the surface-exposed 151 site of superfolder GFP (sfGFP), and used the resulting protein to explore if the unique 5-hydroxyindole functionality could be chemoselectively modified using electrochemistry. An otherwise identical wild-type sfGFP protein, with a tyrosine residue at the 151 site instead of 5HTP, was used as a control. We screened a variety of reaction conditions and potential coupling partners using an IKA ElectraSyn 2.0 Microvial instrument to find that sfGFP-151-5HTP forms an adduct with N,N-dimethylaniline **1a**, as observed by whole-protein mass-spectrometry (**Figure 2b**). Using graphite electrodes at a controlled potential (0.79 V), stirring 20 μM sfGFP-151-5HTP and 10 mM N,N-dimethylaniline **1a** in 100 mM phosphate buffer resulted in complete protein modification within 30 minutes (**Figure 2b**). Treating the wild-type sfGFP protein under the same condition resulted in no detectable protein labeling, highlighting the chemoselectivity of the eCLIC reaction.

Next, we optimized the eCLIC reaction conditions in an attempt to reduce the concentration of the aniline substrate needed for efficient conjugation. These optimizations were performed using a synthetic dialkylaniline **1b** (**Figure 2c**; **Supplementary Figure S1**), which has a carboxylic acid group that improves solubility in aqueous buffers and serves as an attachment handle for other

cargo. We were able to achieve near-complete labeling of 20 μM sfGFP-151-5HTP using only 1.5 mM of **1b**, but an elevated electrode potential of 1V was necessary. Further reducing the concentration of **1b** under these conditions resulted in incomplete conversions, as well as protein overoxidation (**Figure 2c**; **Supplementary Figure S1**). Even at this elevated potential, the chemoselectivity of the eCLIC reaction was not compromised, as evidenced from the absence of protein modification in corresponding control reactions with wild-type sfGFP. Further optimization identified HEPES (N-2-hydroxyethylpiperazine-N′-2-ethanesulfonicacid)[52] buffer to be beneficial for eCLIC; the use of 100 mM HEPES (pH 7) instead of phosphate buffer facilitated near-complete protein labeling using only 0.5 mM **1b**, and also suppressed protein overoxidation, likely due to its known ability to scavenge reactive oxygen species.[53] Finally, the inclusion of TEMPO,[54] a known electron-transfer mediator, further improved the efficiency of the reaction, enabling near-complete labeling of sfGFP-151-5HTP using only 0.25 mM dialkylaniline **1b**, at 0.7 V .

Using these reaction conditions, we were able to functionalize sfGFP-151-5HTP with several other aromatic amines with diverse structures (**Figure 3a**; **Supplementary Figure S2**). Wild-type sfGFP was not labeled by these reagents under the same eCLIC condition (**Supplementary Figure S3**). We also synthesized anilines functionalized with additional cargo, such as a terminal alkyne, a biotin, and a fluorophore, attached through a short PEG linker, and demonstrated their successful conjugation to sfGFP-151-TAG using eCLIC (**Figure 3a**; **Supplementary Figure S2**). Most of the structurally diverse anilines tested demonstrated efficient protein labeling without further optimization of reaction conditions tailored to individual substrates. Not surprisingly, a labeling reaction with phenothiazine, which is known to facilitate the labeling of tyrosine residues, led to multiple labeling of both the wild-type and 5HTP-mutant of sfGFP (**Figure 3a; Supplementary**

Figures S2l, S3). In addition to sfGFP, we also site-specifically incorporated 5HTP into myoglobin (at site 99) and the 5F7 nanobody (at site 69) that binds the HER2 receptor.[55] In each case, we were able to efficiently modify the target protein using eCLIC (**Figure 3b** and **3c**), while the corresponding wild-type proteins lacking the 5HTP residue showed no conversion under identical reaction conditions (**Supplementary Figure S4**). These observations underscore the generality of eCLIC for site-specific protein bioconjugation. Its compatibility with a wide range of aromatic amine coupling partners – which are accessible and inexpensive – significantly enhances the flexibility and utility of this bioconjugation strategy. It should be noted that prolonged incubation in reaction buffer after eCLIC labeling may result in varying levels of non-specific protein oxidation in some cases, observed as a +16 Da species during MS analyses. However, promptly purifying the conjugate using buffer exchange largely eliminates the appearance of such species.

The specificity of the eCLIC reaction toward the 5HTP residue is evident by the fact that only 5HTP mutants of each protein underwent efficient modification, whereas the corresponding wild-type proteins did not. To further confirm the site-specificity of the eCLIC reaction, we subjected the conjugation product of sfGFP-151-5HTP and **1b** to tryptic digestion followed by sensitive HPLC-coupled MS-MS analysis, which confirmed that the **1b** label was exclusively attached to the 5HTP residue (**Supplementary Figure S5**). We also demonstrated that Fmoc-protected free amino acids tyrosine, cysteine, tryptophan, and histidine (the four redox-active natural amino acids) do not react with aromatic amines such as p-toluidine under standard eCLIC conditions (**Supplementary Figure S6**), further supporting the chemoselectivity of the eCLIC reaction.

Structure of the eCLIC product and mechanistic insights

MS analyses of the eCLIC labeling reactions show product masses that are consistent with the attachment of the aromatic amine to the protein substrate and a concomitant loss of 2 Da, which is indicative of an oxidative coupling reaction. To elucidate the structure of this conjugate, we attempted to model the reaction using free 5HTP and N,N-dimethylaniline **1a**. However, this was complicated by the high propensity of 5-hydroxyindoles to form complex oligomeric products under mild oxidative conditions through homocoupling.[45,48-50] Subjecting a mixture of 5HTP and **1a** to eCLIC conditions led to complex product mixtures, even when the latter was used in large excess. This is in stark contrast to the clean eCLIC labeling observed with 5HTP-containing proteins, which likely stems from a significantly reduced propensity of a protein-bound 5HTP residue to undergo homocoupling, due to the steric barriers imposed by neighboring protein residues and slower macromolecular diffusion. Interestingly, we found that the presence of a bulky protection group, such as tert-butyloxycarbonyl (Boc), on the amine functionality of 5HTP suppresses the homocoupling reaction, perhaps mimicking the behavior of a protein-associated 5HTP residue. Using the Boc-protected 5HTP, we were able to significantly promote the heterocoupling reaction with **1a** under eCLIC conditions and isolate the coupling product with the correct mass using HPLC (**Supplementary Figure S7**). A thorough structural characterization using NMR, X-ray crystallography, and mass-spectrometry revealed a C-C coupled product where 5HTP and the N.N-dimethylaniline **1a** are connected through the 4 positions of both aromatic rings (**Figure 4**). We also structurally characterized the coupling product formed between Boc-protected 5HTP (**18**) and a primary aniline substrate **4**. Intriguingly, the eCLIC product (**20**) in this case resulted in a different connectivity, where the aromatic amine is attached to the 4 position of 5HTP, as confirmed by NMR analyses (**Figure 4**, **Supplementary Figure S8**).

CV of both N,N-dimethylaniline **1a** and the primary aniline substrate **2** show a significantly higher potential needed for their oxidation relative to 5HTP (**Figure 2a**). Consequently, it is reasonable to propose a mechanism where the reaction is initiated through anodic oxidation of 5HTP side chain to create a radical cation. The aromatic amines react with this electrophilic radical either through a radical coupling or a nucleophilic attachment mechanism (**Supplementary Figure S9**). A further single-electron oxidation of this intermediate is necessary to rearomatize the coupled product and complete the reaction. We believe that this terminal electron transfer is facilitated by dissolved oxygen or oxidized TEMPO. Indeed, no protein labeling was observed when eCLIC was performed under oxygen-depleted condition and in the absence of TEMPO (**Supplementary Figure S10**). Including TEMPO under oxygen depleted condition partially restored protein labeling. These observations suggest that, unlike established electrochemical protein modification strategies that activate a pro-electrophilic small-molecule labeling reagent, eCLIC proceeds through direct chemoselective activation of the uniquely reactive 5HTP residue on the protein.

Synthesis of functional protein conjugates using eCLIC

In recent decades, monoclonal antibodies have been exploited for countless applications, from enabling reagents for biological discovery to powerful new biologics that offer novel therapeutic modalities. Such applications frequently rely on the ability to functionalize the antibody with additional entities. For example, antibodies labeled with fluorescent probes are widely used for numerous assays both for biological research and disease diagnosis. Additionally, antibody-drug conjugates (ADCs),[18,56,57] where a toxic drug molecule is covalently attached to a monoclonal antibody, provide a powerful strategy to deliver such payloads into target cells with high precision, reducing both the effective therapeutic dose, as well as undesirable side-effects. Traditionally, antibody modification has relied on functionalizing canonical nucleophilic side chains such as

lysine or cysteine, but such strategies are limited by poor control over the site and stoichiometry of antibody modification. Using the GCE technology, it has been possible to incorporate ncAAs with bioconjugation handles into full-length antibodies, which allows subsequent attachment of cargo with full site control.[58,59] Such site-specific labeling strategy provides significant advantage for developing both antibody-based reagents and therapeutics.[15,58,60,61] To explore the feasibility of using eCLIC for modifying full-length antibodies, we recombinantly expressed anti-HER2 antibody Trastuzumab site-specifically incorporating 5HTP at the 121 position of the heavy chain. Expi293 suspension cells were transiently transfected with two plasmids, one encoding the mutant heavy chain (HC-121-TGA) and the light chain of Trastuzumab, and the other a TGA-suppressing EcTrpRS/tRNA pair selective for 5HTP (**Figure 5a**), which produced Trastuzumab-HC-121-5HTP with good yield (8 mg/L) after purification. In comparison, the same expression/purification conditions yielded 12 mg/L wild-type Trastuzumab. Gratifyingly, we were able to homogeneously label Trastuzumab-HC-121-5HTP with aniline probes **2** or **1b** using eCLIC selectively on the heavy chain (**Figure 5c**). Next, we synthesized a primary aniline warhead coupled to a fluorescein (**13**; **Figure 5c**) and conjugated it to Trastuzumab-HC-121-5HTP using the optimized eCLIC condition. Fluorescence imaging following SDS-PAGE (**Figure 5b**) and mass-spectrometry analyses (**Figure 5c**) confirmed selective and efficient labeling of the heavy chain with the fluorophore, whereas the light chain and the corresponding wild-type antibody was not modified. When this antibody-fluorophore conjugate was incubated with the SK-BR-3 cell line that overexpresses the HER2 receptor, robust cell-binding was observed using fluorescence-activated cell sorting (FACS) analyses, while the same conjugate failed to stain HEK293T cells that do not overexpress this receptor (**Figure 5d**). Finally, we attached a primary aniline eCLIC handle to monomethyl auristatin-F (MMAF), a highly cytotoxic tubulin inhibitor, and used eCLIC to

conjugate this compound (**21**) to Trastuzumab-HC-121-5HTP. Successful production of this ADC was confirmed by mass-spectrometry (**Figure 5c**), which demonstrated potent cytotoxicity toward HER2-positive SK-BR-3 cells, but far lower toxicity toward HER2-negative MDA-MB-231 cells (**Figure 5e**). These observations confirm that eCLIC can be used to efficiently generate functional conjugates of full-length antibodies for various applications.

eCLIC is compatible with other click reactions

We envisioned that the eCLIC labeling strategy would be mutually compatible with the two other widely used bioorthogonal reactions: the strain-promoted azide-alkyne cycloaddition (SPAAC),[38] and the inverse-electron demand Diels-Alder (IEDDA) reaction between tetrazines and strained alkenes.[16,39] To test this hypothesis, we expressed the sfGFP protein separately incorporating three different ncAAs at the surface-exposed 151 site: 5HTP, AzK, and CpK (**Figure 6a**). The pyrrolysyl pair was used to incorporate AzK and CpK in response to a UAG codon, as described previously.[22] In each case, incorporation of the desired ncAA was confirmed by the MS analysis, and each sfGFP mutant was separately subjected to three different labeling reactions (**Figure 6a**): SPAAC (with BCN), IEDDA (with tetrazine), and eCLIC (with **1b**). We observed modification of each ncAA residue only upon treatment with its cognate reaction partner, confirming the mutual orthogonality of the three labeling reactions (**Figure 6a**). Next, we sought to demonstrate that eCLIC and SPAAC can be used together to selectively label a protein at two different sites. To this end, we generated an sfGFP protein incorporating 5HTP and AzK at sites 3 and 151, respectively (**Figure 6b**).[22] This was achieved by co-expressing the sfGFP-3-TGA-151-TAG reporter in ATMW-BL21 *E. coli*[43] with a TGA-suppressing EcTrpRS/tRNA pair that charges 5HTP, and a TAG-suppressing pyrrolysyl pair that charges AzK. The resulting protein was purified by immobilized metal-ion chromatography using a C-terminal polyhistidine tag, and was

characterized by mass-spectrometry (**Figure 6c**). This sfGFP double mutant was first labeled at the 5HTP residue through eCLIC using **1b**, followed by SPAAC modification of the AzK side chain with a cyclooctyne. MS analysis of these reactions confirmed efficient sequential labeling at intended sites to yield the dual-labeled product (**Figure 6c**). These results confirm that eCLIC is compatible with other established bioconjugation reactions and can be used together to facilitate precise multi-site protein labeling.

Conclusions

In conclusion, here we show that electrochemistry can be used to efficiently modify proteins at predefined sites using the unique reactivity of the 5HTP residue that can be introduced into proteins through the GCE technology. Previously reported electrochemical protein bioconjugation strategies rely on activating exogenous electrophilic reagents that functionalize a canonical amino acid residue, which intrinsically lack precise control over the site and stoichiometry of protein labeling. In contrast, eCLIC proceeds through chemoselective anodic oxidation of the protein-associated 5HTP residue, followed by its efficient capture by a wide array of aromatic amines. Notably, the reagents used in eCLIC are inexpensive and readily accessible. For example, 5HTP (<$20/g) is significantly less expensive relative to other commonly used ncAAs for bioconjugation ($250-$1000/g), whereas simple anilines such as 4-aminophenylacetic acid **2** (<$3/g) are orders of magnitude cheaper than labeling agents such as cyclooctynes, and trans-cyclooctenes (>$1000/g). Moreover, we recently showed that 5HTP can also be biosynthetically generated in cells,[62] obviating its exogenous supply altogether. A variety of aromatic amines can be used as coupling partners in eCLIC, which enhances the utility of this bioconjugation strategy. Even though we demonstrated this broad substrate scope with several different aromatic amines (**Figure 3a**), it is

likely that the full scope of potential coupling partners is much larger, an aspect we are actively exploring. The GCE technology has been used to express ncAA-containing relevant proteins, such as full-length antibodies, in industrial scale (grams/L),[61,63] providing an established path to express conjugation-ready proteins for eCLIC labeling in large scale. Finally, the compatibility of eCLIC with established bioconjugation strategies such as SPAAC and IEDDA will enable its use for site-specific labeling of proteins with distinct entities with precise site control. Work is currently in progress to further improve the conjugation efficiencies of the eCLIC reaction, particularly for aromatic amine equipped with larger cargo, which appear to require higher concentration to drive full protein modification. We are also exploring additional factors such as the impact of electrode surface properties on the efficiency of protein labeling.

Acknowledgements

This work was supported by National Science Foundation (2128185 to AC) and by NIH (R35GM134964 to EW, R35GM136437 to AC). We thank Prof. Matthias Waegele (Boston College) for insightful discussions, Dr. B. Li (Director, XRD facility, B.C.) and Dr. T. Jayasundera (Director, NMR facilities, B.C.) for their assistance. The funding agencies had no role in study design.

Author contributions

A.C., C.L., and S.J.S.R. designed the experiments and interpreted data; C.L. conducted bioconjugation experiments; S.J.S.R. generated the synthetic probes and performed structural characterizations; V.J.O. assisted with cyclic voltammetry and earlier electrochemistry experiments; S.E.C. and E.W. performed trypsin digestion/MS and associated analysis; A.M.

optimized the eCLIC modification of nanobody 5F7; D.J. assisted with flow cytometry analyses; E.D.F. assisted with bacterial expression of 5HTP-containing proteins; A.C., C.L., and S.J.S.R. prepared the manuscript; A.C. supervised the project.

Competing interests

A patent application on the eCLIC technology reported here has been submitted. AC is a cofounder and senior advisor of BrickBio, Inc. The remaining authors declare no competing interests.

Figures:

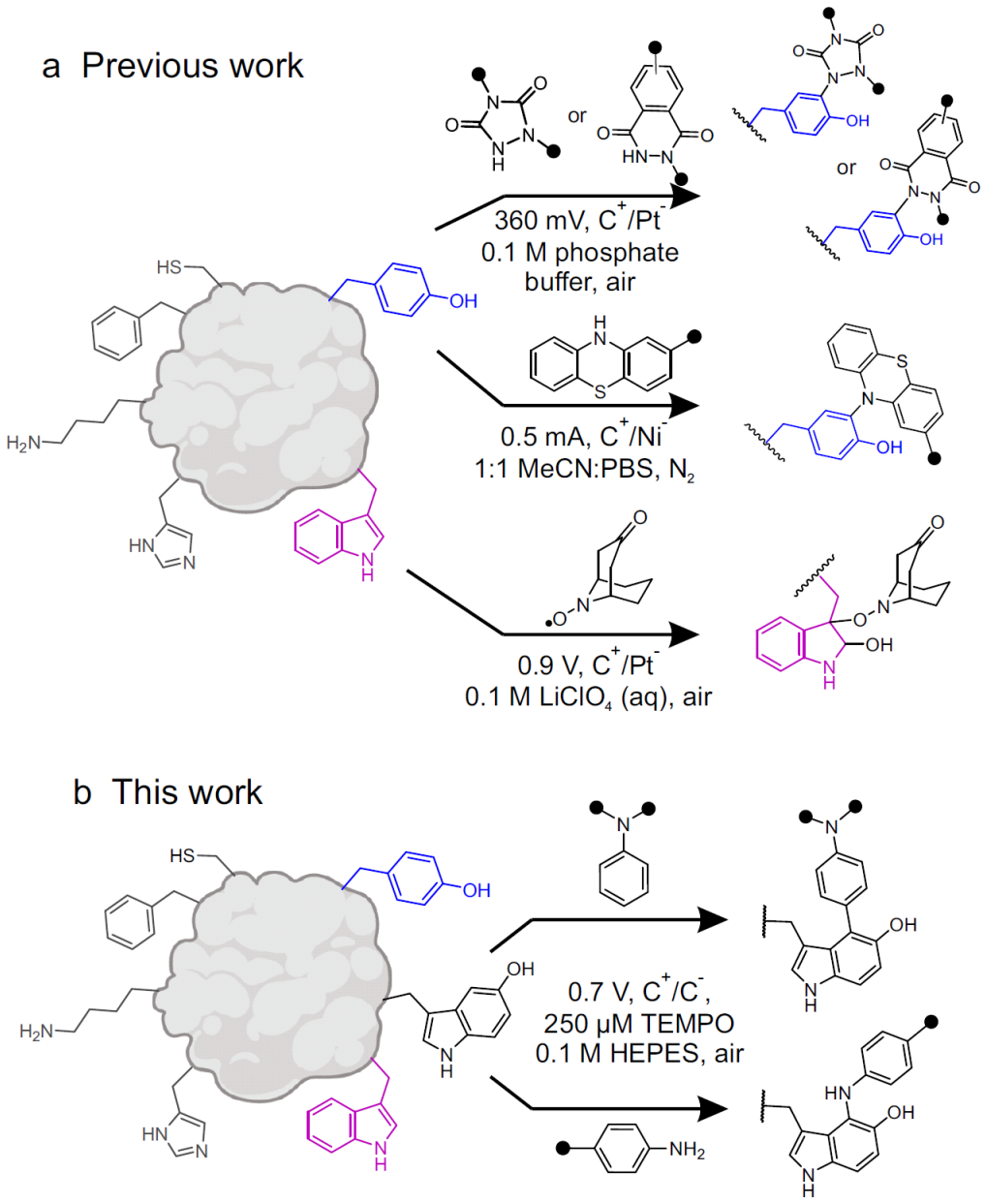
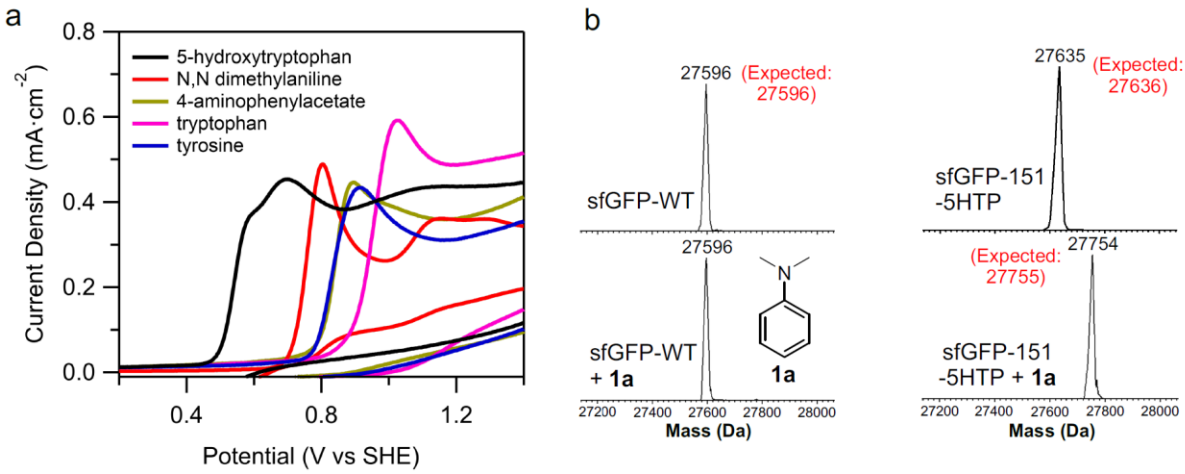
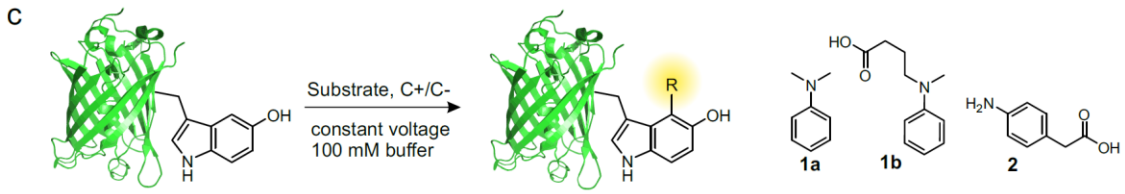


Figure 1. **Electrochemical protein labeling strategies**. **a,** Previous strategies rely on electrochemical activation of exogenous pro-electrophilic reagents, which target canonical amino acid residues such as tyrosine and tryptophan. **b,** In contrast, eCLIC directly oxidizes a site-specifically incorporated 5HTP residue (structure of the side chain shown in red), which then reacts with aromatic amines to generate stable conjugates.





Voltage (V)	Time (min)	Buffer	Substrate	Additive	% Conversion
0.79*	30	Phosphate	3 mM, **1a**	-	84
1	30	Phosphate	1.5 mM,**1b**	-	>95
1	30	Phosphate	0.5 mM,**1b**	-	75
1	30	HEPES	0.5 mM,**1b**	-	>95
0.7	80	HEPES	0.25 mM,**1b**	-	72
0.7	80	HEPES	0.25 mM,**1b**	250 μM TEMPO	>95
0.7	60	HEPES	0.25 mM,**2**	250 μM TEMPO	91

Figure 2. **Discovery and optimization of the eCLIC reaction. a**, CV analysis shows 5HTP is oxidized at a significantly lower potential relative to other canonical aromatic amino acids such as tyrosine and tryptophan, as well as the aromatic amine coupling partners used for eCLIC conjugation. **b**, N,N-dimethylaniline **1a** can be electrochemically attached to sfGFP-151-5HTP to generate a single homogeneous conjugate, whereas the corresponding wild-type sfGFP protein does not react under identical conditions. **c**, Optimization of eCLIC reaction (also see **Supplementary Figure S1**) revealed beneficial effect of HEPES buffer and the inclusion of TEMPO (*vs SHE; all others are ElectraSyn cell voltages).

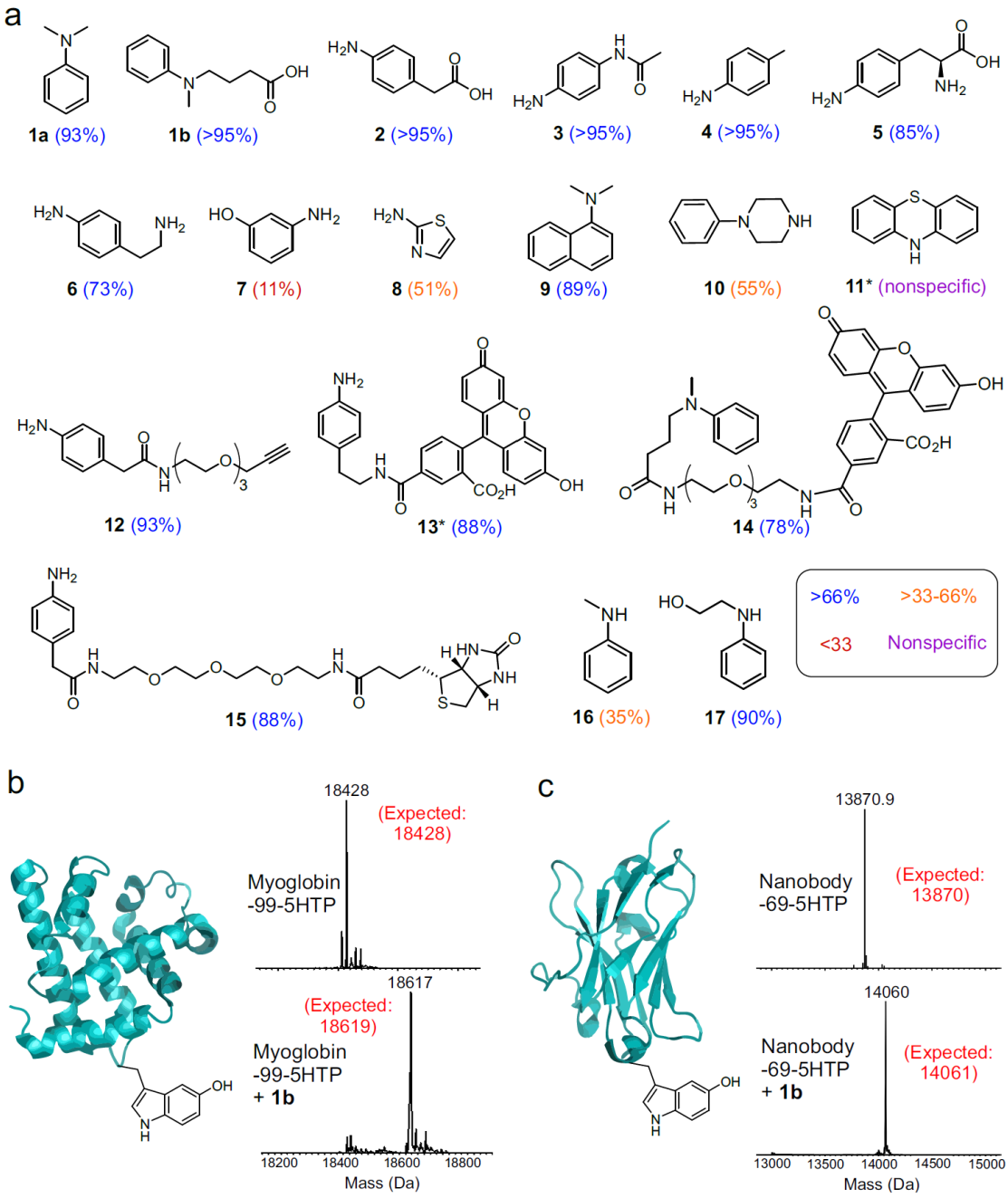


Figure 3. **The substrate scope for eCLIC. a**, Various aromatic amine substrates can successfully participate in the eCLIC reaction (also see **Supplementary Figure S2**). % conversion in brackets under the following reaction conditions: 20 µM sfGFP-151-5HTP + 1 mM substrate, 100 mM HEPES (pH 7), 250 µM TEMPO, 0.7 V, 60 min). *Phenothiazine (**11**) was used at 0.5 mM. Except for phenothiazine, which is known to react with tyrosine residues, none of the other substrates labeled wild-type sfGFP lacking the 5HTP residue (**Supplementary Figure S3**). Myoglobin-99-5HTP (**b**) and anti-HER2 nanobody-69-5HTP (**c**) can also be efficiently labeled using the optimized eCLIC condition. The corresponding wild-type proteins, lacking the 5HTP residue, show no labeling under identical conditions (**Supplementary Figure S4**).

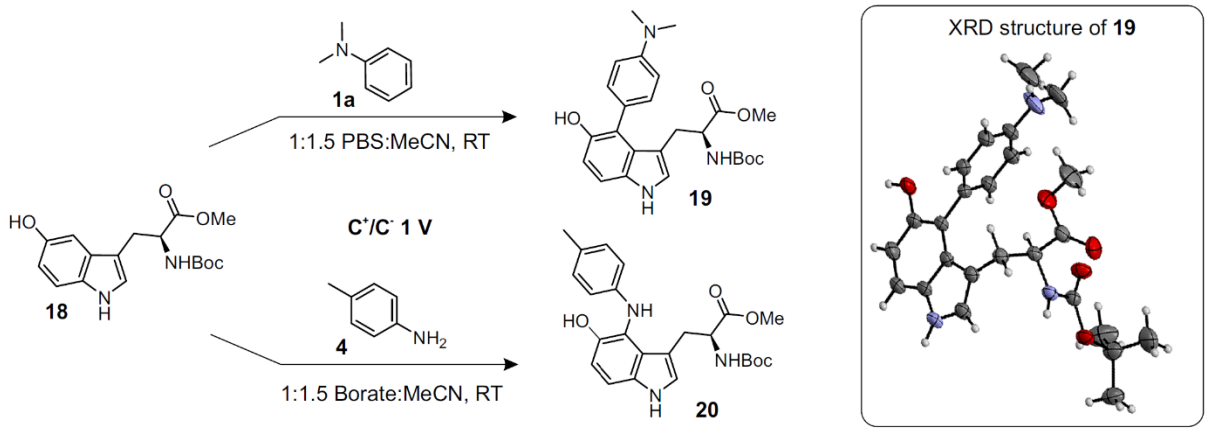


Figure 4. **Structural characterization of the eCLIC product.** A Boc-protected 5HTP (**18**) enabled the suppression of the interfering homocoupling reaction and formation of significant quantities of the aniline coupled product. These products were characterized by NMR and MS analysis (**Supplementary Figures S7, S8**). The coupling product (**19**) with N,N-dimethylaniline (**1a**) was further characterized using X-ray crystallography, which is shown above.

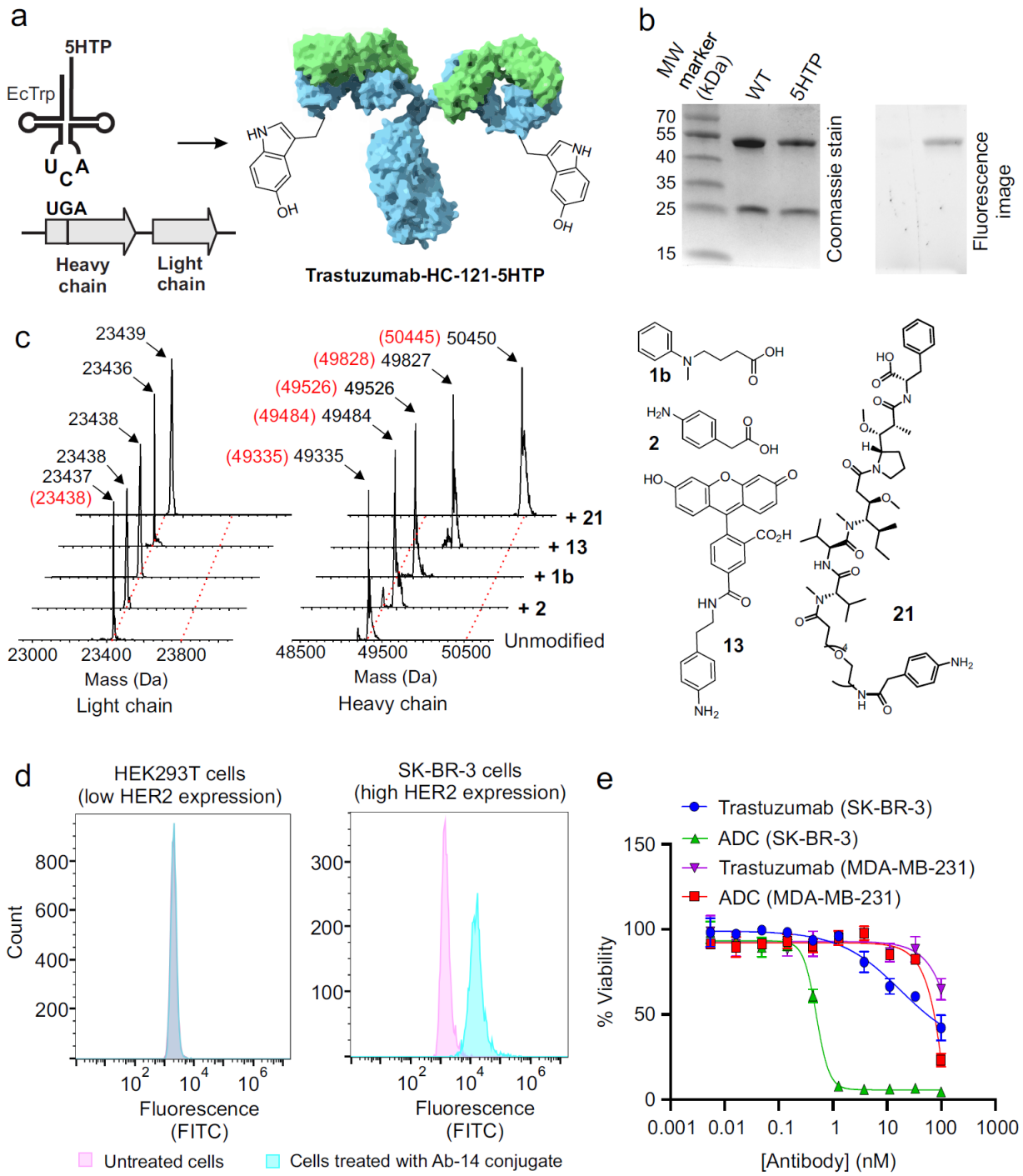


Figure 5. **Functional conjugates of a full-length antibody using eCLIC**. **a**, Using an engineered EcTrpRS/tRNA pair, 5HTP was site-specifically incorporated into the 121 site of the heavy chain of Trastuzumab. **b**, Analysis of the eCLIC labeling reaction of wild-type Trastuzumab and Trastuzumab-HC-121-5HTP with the fluorescent substrate **13** by SDS-PAGE followed by fluorescence imaging revealed selective labeling of the 5HTP-containing heavy chain. This experiment was performed twice with similar results. **c**, Subjecting Trastuzumab-HC-121-5HTP to eCLIC reaction using three different coupling partners (**1b**, **2**, **13**, **21**) led to efficient labeling

of the heavy chain in each case, while the light chain remained unmodified, as shown by MS analysis (red: theoretical mass; black: observed mass).**d**, Incubating HER2-expressing SK-BR-3 cells with **13**-labeled Trastuzumab (prepared by eCLIC) results in efficient cell staining, as shown by FACS analysis, while it failed to stain the HER2-negative HEK293T cells. Gating strategy for flow cytometry analysis is shown in **Supplementary Figure S12**. **e**, Trastuzumab-**21** (antibody-drug conjugate; ADC) conjugate generated through eCLIC shows potent and selective toxicity toward HER2+ SK-BR-3 cells, but not HER2-negative MDA-MB-231 cells. Each experiment was performed in biological triplicate with error bars representing +/- SEM.

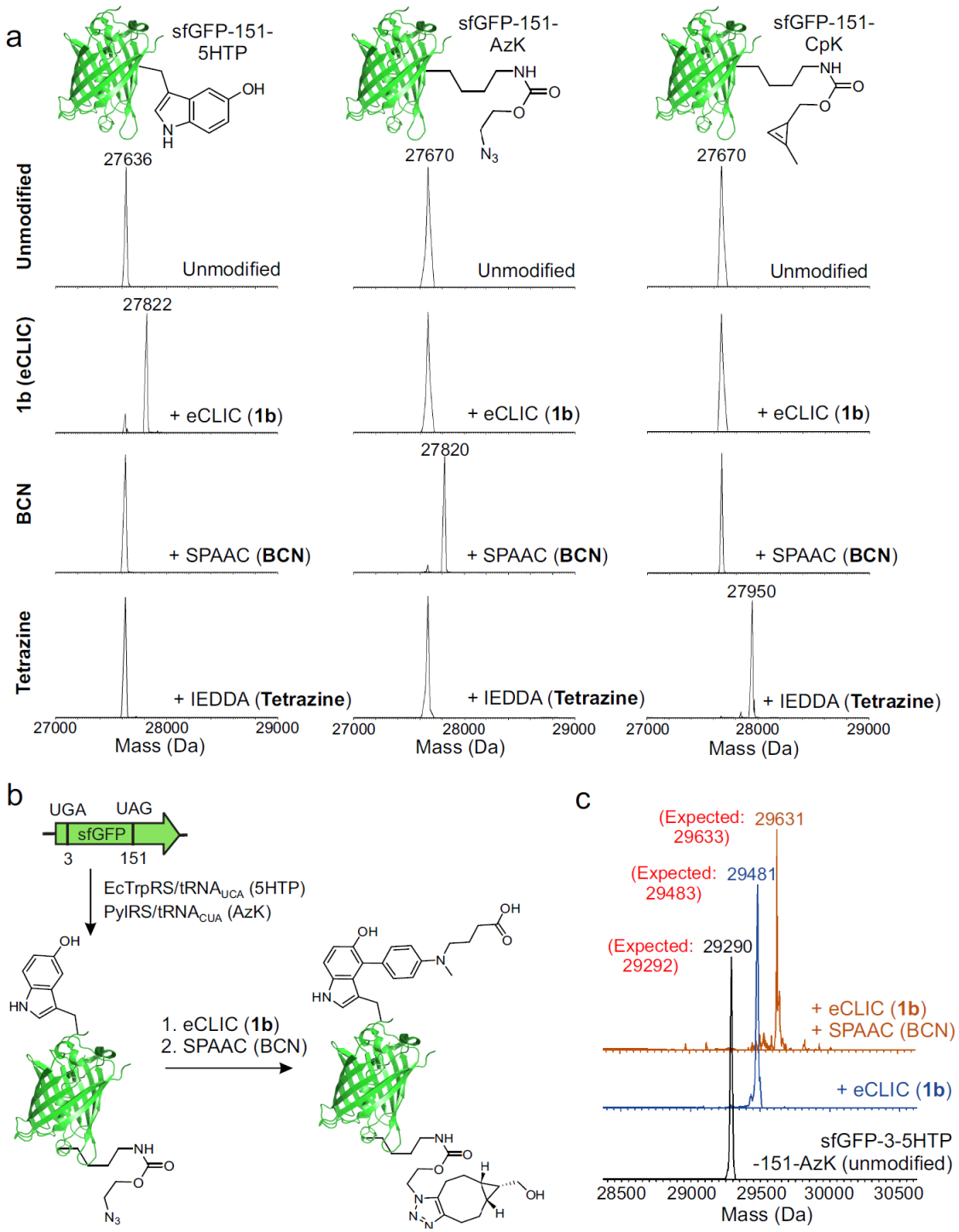


Figure 6. **The eCLIC reaction is compatible with SPAAC and IEDDA**. **a**, sfGFP reporters harboring either 5HTP, AzK, or CpK (the structure of each amino acid side chain is shown above the panel) were separately subjected to eCLIC conditions (with **1b**), BCN, or tetrazine, which shows that each reagent selectively labels its cognate substrate, but not the others. **b**, A sfGFP reporter harboring a 5HTP at site 3 and an AzK at site 151 was expressed using engineered EcTrp and pyrrolysyl tRNA-synthetase/tRNA pairs, respectively. Sequentially subjecting this protein to eCLIC (with **1b**) and SPAAC (with BCN) resulted in efficient site-specific dual-labeling.

References:

1 Faraday, M. Siebente Reihe von Experimental-Untersuchungen über Elektricität. *Annalen der Physik* **109**, 481-520 (1834).

2 Kolbe, H. Beobachtungen über die oxydirende Wirkung des Sauerstoffs, wenn derselbe mit Hülfe einer elektrischen Säule entwickelt wird. *Journal für praktische Chemie* **41**, 137-139 (1847).

3 Horn, E. J., Rosen, B. R. & Baran, P. S. Synthetic Organic Electrochemistry: An Enabling and Innately Sustainable Method. *ACS Central Science* **2**, 302-308, doi:10.1021/acscentsci.6b00091 (2016).

4 Yan, M., Kawamata, Y. & Baran, P. S. Synthetic Organic Electrochemical Methods Since 2000: On the Verge of a Renaissance. *Chemical Reviews* **117**, 13230-13319, doi:10.1021/acs.chemrev.7b00397 (2017).

5 Little, R. D. A Perspective on Organic Electrochemistry. *The Journal of Organic Chemistry* **85**, 13375-13390, doi:10.1021/acs.joc.0c01408 (2020).

6 Zhu, C., Ang, N. W. J., Meyer, T. H., Qiu, Y. & Ackermann, L. Organic Electrochemistry: Molecular Syntheses with Potential. *ACS Central Science* **7**, 415-431, doi:10.1021/acscentsci.0c01532 (2021).

7 Spicer, C. D. & Davis, B. G. Selective chemical protein modification. *Nature communications* **5**, ncomms5740 (2014).

8 Stephanopoulos, N. & Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nature chemical biology* **7**, 876-884, doi:10.1038/nchembio.720 (2011).

9 Boutureira, O. & Bernardes, G. a. J. Advances in chemical protein modification. *Chemical reviews* **115**, 2174-2195 (2015).

10 deGruyter, J. N., Malins, L. R. & Baran, P. S. Residue-Specific Peptide Modification: A Chemist's Guide. *Biochemistry* **56**, 3863-3873 (2017).

11 Lang, K. & Chin, J. W. Bioorthogonal reactions for labeling proteins. *ACS chemical biology* **9**, 16-20, doi:10.1021/cb4009292 (2014).

12 McKay, C. S. & Finn, M. G. Click chemistry in complex mixtures: bioorthogonal bioconjugation. *Chemistry & biology* **21**, 1075-1101, doi:10.1016/j.chembiol.2014.09.002 (2014).

13 Shih, H. W., Kamber, D. N. & Prescher, J. A. Building better bioorthogonal reactions. *Current opinion in chemical biology* **21**, 103-111, doi:10.1016/j.cbpa.2014.07.002 (2014).

14 Sletten, E. M. & Bertozzi, C. R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angewandte Chemie (International ed. in English)* **48**, 6974-6998, doi:10.1002/anie.200900942 (2009).

15 Kim, C. H., Axup, J. Y. & Schultz, P. G. Protein conjugation with genetically encoded unnatural amino acids. *Current opinion in chemical biology* **17**, 412-419, doi:10.1016/j.cbpa.2013.04.017 (2013).

16 Scinto, S. L. *et al.* Bioorthogonal chemistry. *Nature Reviews Methods Primers* **1**, 30, doi:10.1038/s43586-021-00028-z (2021).

17 Hoyt, E. A., Cal, P. M., Oliveira, B. L. & Bernardes, G. J. Contemporary approaches to site-selective protein modification. *Nature Reviews Chemistry* **3**, 147-171 (2019).

18 Walsh, S. J. *et al.* Site-selective modification strategies in antibody–drug conjugates. *Chemical Society Reviews* **50**, 1305-1353, doi:10.1039/D0CS00310G (2021).

19 Bloom, S. *et al.* Decarboxylative alkylation for site-selective bioconjugation of native proteins via oxidation potentials. *Nature chemistry* **10**, 205 (2018).
20 Lin, S. *et al.* Redox-based reagents for chemoselective methionine bioconjugation. *Science* **355**, 597-602, doi:doi:10.1126/science.aal3316 (2017).
21 Patterson, D. M. & Prescher, J. A. Orthogonal bioorthogonal chemistries. *Current opinion in chemical biology* **28**, 141-149, doi:10.1016/j.cbpa.2015.07.006 (2015).
22 Italia, J. S. *et al.* Mutually Orthogonal Nonsense-Suppression Systems and Conjugation Chemistries for Precise Protein Labeling at up to Three Distinct Sites. *Journal of the American Chemical Society* **141**, 6204-6212, doi:10.1021/jacs.8b12954 (2019).
23 Bednar, R. M., Karplus, P. A. & Mehl, R. A. Site-specific dual encoding and labeling of proteins via genetic code expansion. *Cell Chemical Biology* **30**, 343-361, doi:https://doi.org/10.1016/j.chembiol.2023.03.004 (2023).
24 Osgood, A. O. *et al.* An Efficient Opal-Suppressor Tryptophanyl Pair Creates New Routes for Simultaneously Incorporating up to Three Distinct Noncanonical Amino Acids into Proteins in Mammalian Cells. *Angewandte Chemie (International ed. in English)* **62**, e202219269, doi:10.1002/anie.202219269 (2023).
25 Mackay, A. S., Payne, R. J. & Malins, L. R. Electrochemistry for the Chemoselective Modification of Peptides and Proteins. *Journal of the American Chemical Society* **144**, 23-41, doi:10.1021/jacs.1c11185 (2022).
26 Song, C. *et al.* Electrochemical oxidation induced selective tyrosine bioconjugation for the modification of biomolecules. *Chemical science* **10**, 7982-7987 (2019).
27 Alvarez-Dorta, D. *et al.* Electrochemically Promoted Tyrosine-Click-Chemistry for Protein Labeling. *Journal of the American Chemical Society* **140**, 17120-17126, doi:10.1021/jacs.8b09372 (2018).
28 Sato, S. *et al.* Site-Selective Protein Chemical Modification of Exposed Tyrosine Residues Using Tyrosine Click Reaction. *Bioconjugate Chemistry* **31**, 1417-1424, doi:10.1021/acs.bioconjchem.0c00120 (2020).
29 Depienne, S. *et al.* Luminol anchors improve the electrochemical-tyrosine-click labelling of proteins. *Chem Sci* **12**, 15374-15381, doi:10.1039/d1sc04809k (2021).
30 Seki, Y. *et al.* Transition Metal-Free Tryptophan-Selective Bioconjugation of Proteins. *Journal of the American Chemical Society* **138**, 10798-10801, doi:10.1021/jacs.6b06692 (2016).
31 Toyama, E. *et al.* Electrochemical Tryptophan-Selective Bioconjugation. *ChemRxiv*, 10.26434/chemrxiv.7795484.v7795481, doi:10.26434/chemrxiv.7795484.v1 (2019).
32 Kendall, G. *et al.* Specific electrochemical nitration of horse heart myoglobin. *Arch Biochem Biophys* **392**, 169-179, doi:10.1006/abbi.2001.2451 (2001).
33 Iniesta, J. *et al.* Specific electrochemical iodination of horse heart myoglobin at tyrosine 103 as determined by Fourier transform ion cyclotron resonance mass spectrometry. *Archives of biochemistry and biophysics* **474**, 1-7 (2008).
34 Baran, P. S., Guerrero, C. A. & Corey, E. J. The First Method for Protection−Deprotection of the Indole 2,3-π Bond. *Organic Letters* **5**, 1999-2001, doi:10.1021/ol034634x (2003).
35 Decoene, K. *et al.* Protein Conjugation with Triazolinediones: Switching from a General Tyrosine-Selective Labeling Method to a Highly Specific Tryptophan Bioconjugation Strategy. *ChemRxiv* **10.26434/chemrxiv.13739320.v1**, doi:10.26434/chemrxiv.13739320.v1 (2021).

36 Italia, J. S. *et al.* An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. *Nature chemical biology* **13**, 446-450, doi:10.1038/nchembio.2312 (2017).
37 White, A. M., Palombi, I. R. & Malins, L. R. Umpolung strategies for the functionalization of peptides and proteins. *Chem Sci* **13**, 2809-2823, doi:10.1039/d1sc06133j (2022).
38 Jewett, J. C. & Bertozzi, C. R. Cu-free click cycloaddition reactions in chemical biology. *Chemical Society Reviews* **39**, 1272-1279 (2010).
39 Devaraj, N. K. & Weissleder, R. Biomedical applications of tetrazine cycloadditions. *Accounts of chemical research* **44**, 816-827 (2011).
40 Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53 (2017).
41 Italia, J. S. *et al.* Expanding the genetic code of mammalian cells. *Biochemical Society transactions* **45**, 555-562, doi:10.1042/bst20160336 (2017).
42 Young, D. D. & Schultz, P. G. Playing with the molecules of life. *ACS chemical biology* (2018).
43 Ficaretta, E. D. *et al.* A Robust Platform for Unnatural Amino Acid Mutagenesis in E. coli Using the Bacterial Tryptophanyl-tRNA synthetase/tRNA pair. *Journal of molecular biology* **434**, 167304, doi:10.1016/j.jmb.2021.167304 (2022).
44 Addy, P. S., Erickson, S. B., Italia, J. S. & Chatterjee, A. A chemoselective rapid azo-coupling reaction (CRACR) for unclickable bioconjugation. *Journal of the American Chemical Society* **139**, 11670-11673 (2017).
45 Addy, P. S., Italia, J. S. & Chatterjee, A. An Oxidative Bioconjugation Strategy Targeted to a Genetically Encoded 5-Hydroxytryptophan. *Chembiochem : a European journal of chemical biology* **19**, 1375-1378, doi:10.1002/cbic.201800111 (2018).
46 Addy, P. S., Zheng, Y., Italia, J. S. & Chatterjee, A. A "Quenchergenic" Chemoselective Protein Labeling Strategy. *Chembiochem : a European journal of chemical biology* **20**, 1659-1663, doi:10.1002/cbic.201800817 (2019).
47 Singha Roy, S. J. *et al.* Photoredox-Catalyzed Labeling of Hydroxyindoles with Chemoselectivity (PhotoCLIC) for Site-Specific Protein Bioconjugation. *Angewandte Chemie (International ed. in English)* **62**, e202300961, doi:10.1002/anie.202300961 (2023).
48 Humphries, K. A., Wrona, M. Z. & Dryhurst, G. Electrochemical and enzymatic oxidation of 5-hydroxytryptophan. *Journal of Electroanalytical Chemistry* **346**, 377-403, doi:https://doi.org/10.1016/0022-0728(93)85026-D (1993).
49 Napolitano, A., d'Ischia, M. & Prota, G. A profile of the oxidation chemistry of 5-hydroxyindole under biomimetic conditions. *Tetrahedron* **44**, 7265-7270 (1988).
50 Wrona, M. Z. & Dryhurst, G. Oxidation of serotonin by superoxide radical: implications to neurodegenerative brain disorders. *Chemical research in toxicology* **11**, 639-650 (1998).
51 Yamagishi, Y., Ashigai, H., Goto, Y., Murakami, H. & Suga, H. Ribosomal synthesis of cyclic peptides with a fluorogenic oxidative coupling reaction. *Chembiochem : a European journal of chemical biology* **10**, 1469-1472, doi:10.1002/cbic.200900021 (2009).
52 Ferguson, W. J. *et al.* Hydrogen ion buffers for biological research. *Analytical biochemistry* **104**, 300-310, doi:10.1016/0003-2697(80)90079-2 (1980).

53 Grady, J. K., Chasteen, N. D. & Harris, D. C. Radicals from "Good's" buffers. *Analytical biochemistry* **173**, 111-115, doi:10.1016/0003-2697(88)90167-4 (1988).
54 Nutting, J. E., Rafiee, M. & Stahl, S. S. Tetramethylpiperidine N-Oxyl (TEMPO), Phthalimide N-Oxyl (PINO), and Related N-Oxyl Species: Electrochemical Properties and Their Use in Electrocatalytic Reactions. *Chem Rev* **118**, 4834-4885, doi:10.1021/acs.chemrev.7b00763 (2018).
55 Sheikholeslami, F. *et al.* Isolation of a Novel Nanobody Against HER-2/neu Using Phage Displays Technology. *Laboratory Medicine* **41**, 69-76, doi:10.1309/lm0wxkm0r0dvuzwf (2010).
56 Drago, J. Z., Modi, S. & Chandarlapaty, S. Unlocking the potential of antibody-drug conjugates for cancer therapy. *Nature reviews. Clinical oncology* **18**, 327-344, doi:10.1038/s41571-021-00470-8 (2021).
57 Beck, A., Goetsch, L., Dumontet, C. & Corvaïa, N. Strategies and challenges for the next generation of antibody-drug conjugates. *Nature reviews. Drug discovery* **16**, 315-337, doi:10.1038/nrd.2016.268 (2017).
58 Axup, J. Y. *et al.* Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16101-16106, doi:10.1073/pnas.1211023109 (2012).
59 VanBrunt, M. P. *et al.* Genetically Encoded Azide Containing Amino Acid in Mammalian Cells Enables Site-Specific Antibody–Drug Conjugates Using Click Cycloaddition Chemistry. *Bioconjugate Chemistry* **26**, 2249-2260, doi:10.1021/acs.bioconjchem.5b00359 (2015).
60 Agarwal, P. & Bertozzi, C. R. Site-specific antibody–drug conjugates: the nexus of bioorthogonal chemistry, protein engineering, and drug development. *Bioconjugate chemistry* **26**, 176-192, doi:10.1021/bc5004982 (2015).
61 Tian, F. *et al.* A general approach to site-specific antibody drug conjugates. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 1766-1771, doi:10.1073/pnas.1321237111 (2014).
62 Chen, Y. *et al.* Creation of Bacterial cells with 5-Hydroxytryptophan as a 21(st) Amino Acid Building Block. *Chem* **6**, 2717-2727, doi:10.1016/j.chempr.2020.07.013 (2020).
63 Roy, G. *et al.* Development of a high yielding expression platform for the introduction of non-natural amino acids in protein sequences. *mAbs* **12**, 1684749, doi:10.1080/19420862.2019.1684749 (2020).

Methods:

Cyclic Voltammogram Measurements

Experiments were carried out in a home-made glass three electrode cell containing a carbon rod counter electrode (99.995% trace metals; Sigma Aldrich), a Ag/AgCl reference electrode (3 M NaCl; Basi Inc.; West Lafayette, IN), and a glassy carbon working electrode (3.0 mM OD; Pine Research; Durham, NC). Prior to each experiment, the glassy carbon electrode was polished for 5 minutes with a 1 μM polycrystalline diamond slurry (Ted Pella; Redding, CA). The electrode potential was controlled with a Versastat 3 potentiostat (AMETEK; Berwyn, PA).

The CV experiments were carried out with 5 mL of electrolyte solution containing a pH 7 PBS supporting electrolyte (20 mM phosphate; 150 mM sodium chloride) and 1 mM of the analyte. The nominal turning potentials for the CV were -0.3 V and +1.7 V vs SHE with a scan rate of 100 mV/s.

The CV experiments with TEMPO and N-Boc-OMe-5HTP were carried out in 5 mL of electrolyte solution containing a 1:1 mixture of pH 7 PBS supporting electrolyte (20 mM phosphate; 150 mM sodium chloride) and acetonitrile. Mixed into the electrolyte was 250 μM of the analyte of interest. The nominal turning potentials for the CVs were +0.2 V and +1.0 V vs SHE with a scan rate of 10 mV/s.

Various eCLIC bioconjugation reactions

General optimized condition for eCLIC

To a 1 mL ElectraSyn reaction vial was added 40 μL of 1 M HEPES, pH 7, 1 μL of 100 mM TEMPO, 8 μL of 1 mM sfGFP-151-5HTP (20 μM final concentration), a sufficient volume of substrate in DMSO for a 250 μM – 1000 μM final concentration, and molecular biology grade water to a final volume of 400 μL. Reactions were run under "Constant Voltage" settings at 0.7 V for 60 min at 250 μM – 1000 μM substrate (time and concentration may be optimized), using SK-50 microelectrodes as both the anode and cathode with no reference electrode. Due to the low current generated in the reaction, Vial Detection setting was turned off. The reaction mixture was buffer-exchanged using a Bio-Spin® P-30 column (Bio-Rad 7326006) pre-equilibrated with PBS to remove excess small-molecule substrate and HEPES buffer.

Reaction under oxygen-depleted conditions

Oxygen-free deionized water was prepared by sparging with nitrogen for 1 hour. The sparged water, solid HEPES powder, concentrated 151-5HTP-sfGFP stock (1.1 mM), 100 mM TEMPO in DMSO, and 100 mM dialkylaniline carboxylate (1b) were transferred to a N_2 glovebox after 3 purge/refill cycles and left for 24 hours. The water was used to prepare fresh 1 M HEPES buffer, pH 7. All components were mixed as in standard reaction conditions (20 μM 151-5HTP-sfGFP, 100 mM HEPES, 1 mM 1b) in the presence or absence of 250 μM TEMPO and sealed in an Electrasyn 2.0 microvial. Reactions were carried out at 700 mV for 80 minutes and processed as described above for mass spectrometry analysis outside of the glove box.

Testing orthogonality of eCLIC, SPAAC, and IEDDA

sfGFP-151-5HTP, sfGFP-151-AzK, and sfGFP-151-CpK were each diluted to 20 μM in 100 mM HEPES, pH 7 for all reactions. eCLIC was carried out on each protein with 250 μM dialkylaniline carboxylate **1b** and 250 μM TEMPO for 1 hour. Tetrazine cycloadditions were carried out with a final concentration of 200 μM tetrazineamine,[64,65] and allowed to react for 3 h. Strain-promoted azide-alkyne cycloadditions were carried out with a final concentration of 500 μM BCN-OH (Sigma-Aldrich), and allowed to react for 3 h. All labeling reactions were buffer exchanged into PBS by PD-10 spin columns and analyzed by ESI-MS.

Site-specific dual labelling of sfGFP-3-5HTP-151-AzK

A stock of sfGFP-3-5HTP-151-AzK was diluted to a 20 μM in 100 mM HEPES, pH 7, and reacted with 1 mM **1b** for 30 minutes at 1 V. The sample was buffer exchanged into PBS using a PD-10 spin column, then reacted with BCN-OH was added to a final concentration of 1 mM and was allowed to incubate for 1 hour at room temperature. The protein was analyzed after both steps by ESI-MS.

Antibody drug conjugate synthesis

121-5HTP-Trastuzumab was modified with aniline-MMAF (**20)** as described in the *general optimized condition for eCLIC* using V2100 microelectrodes instead of SK-50. To remove any remaining **20**, the resulting antibody-drug conjugate was buffer exchanged with a Bio-Spin® P-30 column followed by three rounds of concentration and dilution in PBS using an Amicon® Ultra 0.5mL filter (Millipore Sigma UFC501096).

Mass spectrometry analysis of purified proteins

Whole-protein MS analysis

Purified proteins and protein conjugates were analyzed on an Agilent 1260 Infinity ESI–TOF. Samples were separated on a Phenomenex Aeris 3.6 μm Widepore XB-C8 (100 x 4.5 mm) using a gradient from 5–99% buffer B in buffer A (buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; buffer B; 5% water, 95% acetonitrile, 0.1% formic acid), and analyzed in positive ion mode. Total protein masses were calculated by deconvolution using MagTran (Amgen Inc.). Theoretical masses of wild-type proteins were calculated on Benchling (Benchling.com), and theoretical masses for ncAA-containing proteins or conjugates thereof were manually calculated.

Tryptic MS/MS analysis of 1b-modified 151-5HTP-sfGFP

Modified 151-5HTP-sfGFP (100 μg) was precipitated by adding 100% trichloroacetic acid (TCA) in PBS (100 μg/100 μL). The solution was vortexed and then incubated at -80 ºC overnight. After thawing, the sample was centrifuged (15,000 x g, 10 minutes, 4 ºC) and the supernatant was removed. The protein pellet was washed with cold acetone (500 μL), resuspended by brief sonication (4 pulses, 30% amp) and pelleted by centrifugation (5,000 x g, 10 minutes, 4 ºC). The acetone was removed and the protein pellet was allowed to air dry. The protein pellet was

resuspended in 8 M Urea/PBS (30 µL) and then 100 mM ammonium bicarbonate (70 µL) was added to a total volume of 100 µL. Reductive alkylation of cysteines was performed with 50 mM TCEP (5 µL, 15 min, 65 ^{o}C) and 500 mM iodoacetamide (2.5 µL, 30 min, 25 ^{o}C). The sample was then diluted in PBS (120 µL) and digested overnight at 37 ^{o}C following the addition of trypsin (4 µL of 20 µg diluted in trypsin resuspension buffer) and 100 mM $CaCl_2$ (2.5 µL). After quenching with 10% formic acid, the trypsin-digested peptide sample was desalted using a SEP-Pak desalting columns (Waters), dried by speed-vac, and resuspended in Buffer A (200 µL, 100% H_2O, 0.1% formic acid) for LC-MS/MS analysis.

LC-MS/MS analysis was performed on an Orbitrap Exploris 240 mass spectrometer running Xcalibur v4.4 (Thermo Scientific) coupled to a Dionex Ultimate 3000 RSLCnano system. Samples (5 µL) were injected directly onto a 4 cm Acclaim PepMap 100 C18 column and peptides were eluted onto an Acclaim PepMap RSLC. Peptides were separated with a 1-hour gradient from 5% to 25% of Buffer B (20% H_2O, 80 % MeCN, 0.1% formic acid) in Buffer A (100% H_2O, 0.1% formic acid) at a flow rate of 0.3 µL/min. The spray voltage was set to 2.1 kV. One full MS1 scan (120,000 resolution, 350-1800 m/z, RF lens 65%, AGC target 300%, automatic maximum injection time, profile mode) was obtained every 2 secs with dynamic exclusion (repeat count 2, duration 10 s), isotopic exclusion (assigned), and apex detection (30% desired apex window) enabled. A variable number of MS2 scans (15,000 resolution, AGC 75%, maximum injection time 100 ms, centroid mode) were obtained between each MS1 scan based on the highest precursor masses, filtered for monoisotopic peak determination, theoretical precursor isotopic envelope fit, intensity (5E4), and charge state (2-6). MS2 analysis consisted of the isolation of precursor ions (isolation window 2 m/z) followed by higher-energy collision dissociation (HCD) (collision energy 30%). The tandem MS data was analyzed by the Thermo Proteome Discoverer V2.4 software package and searched using the SequestHT and Percolator algorithms against a UniprotKB database of *E. coli* K12 proteome amended to include the sequence for **151-W-sfGFP**. Trypsin was specified as the protease with a maximum of 2 missed cleavages. Peptide precursor mass tolerance was set to 10 ppm with a fragment mass tolerance of 0.02 Da. Oxidation of tryptophan (+15.995), **dialkylaniline (1b) labeling** of oxidized tryptophan (+207.090), oxidation of methionine (+15.995) as well as acetylation (+42.011) and/or methionine-loss (+131.040) of the protein N-terminus were set as dynamic modifications. Cysteine alkylation (+57.021) was set as a static modification. The false discovery rate (FDR) for peptide identification was set to 1%. All mass-spectrometry data were collected as two technical replicates each from two independent biological replicates.

Cell line sources

HEK293T (CRL-11268^{TM}) and SK-BR-3 (HTB-30^{TM}) cells were acquired from ATCC. MDA-MB-231 (HTB-26^{TM}) cells were a gift from EW, originally sourced from ATCC. Expi293F^{TM} cells were acquired from Thermo Fisher (Catalog # A14527). ATMW-BL21 cells were prepared in-house and reported previously.[3]

Protein expression and purification

Expression of sfGFP, myoglobin, and 5F7 nanobody in ATMW-BL21

Expression and protein purification of sfGFP, myoglobin, and 5F7 nanobody were performed as previously described, in the *E. coli* ATMW-BL21 strain.[43] All proteins were expressed from a pET22b plasmid, driven by a IPTG-inducible T5-lac promoter. In brief, ATMW-BL21 cells were co-transformed with the appropriate pET22b plasmid and a pEVOL-*tacI* EcTrpRS-h13 leuV[43] suppressor plasmid. For wild-type protein expression, the cells were transformed with only the appropriate pET22b plasmid. A single transformant was used to start a 5 mL overnight culture containing the appropriate antibiotics (100 μg/mL ampicillin, 20 μg/mL chloramphenicol, 95 μg/mL spectinomycin, 15 μg/mL zeocin, and 10 μg/mL gentamycin), 0.5 mL of which was used to inoculate a 50 mL LB culture (in a 250 mL sterile Erlenmeyer flask) supplemented with antibiotics. Cultures were grown to an OD_{600} of 0.5-0.6, followed by induction with a final concentration of 1 mM IPTG and 1 mM 5HTP. The cultures were expressed at 30 °C with shaking at 250 rpm for 16 h. The cultures were pelleted at 5,000 x g for 10 minutes at 4 °C, and the supernatant was removed. Cell pellets were resuspended in a ratio of 1 mL lysis buffer to 10 mL expression culture volume. Lysis buffer: B-PER Bacterial Protein Extraction Reagent (Thermo Scientific), 1X Halt Protease Inhibitor Cocktail (Thermo Scientific), 0.01% Pierce Universal Nuclease (Thermo Scientific). Cells were nutated for 30 minutes at room temperature, then clarified by centrifugation at 18,000 xg for 10 minutes at 4 °C. The sfGFP reporter proteins contain a C-terminal polyhistidine tag and were purified using a HisPur Ni-NTA resin (Thermo Scientific) following the manufacturer's protocol. Protein purity was confirmed by SDS-PAGE analysis and protein molecular mass was characterized using HPLC-coupled ESI-MS of intact proteins (Agilent Technologies, 1260 Infinity ESI–TOF).

To express the sfGFP-3TGA-151TAG reporter, ATMW-BL21 cells were co-transformed with a pGTEV-3TGA-151TAG reporter plasmid[22] and pEvoltac EcW.h14-MbPylRS TAG suppressor plasmid encoding both EcTrp and the pyrrolysyl pairs.[22] The sfGFP reporter was expressed and analyzed as described above, except supplemented with 1 mM 5HTP and 1 mM AzK upon 1 mM IPTG induction, and a larger culture volume of 500 mL (grown and expressed in a 2 L sterile Erlenmeyer flask) was used.

Expression of sfGFP-151-CpK and sfGFP-151-AzK were performed as previously described.[22]

Trastuzumab expression in Expi293 suspension cells and mass spectrometry analysis:

Expi293 cells were grown and cultivated under standard conditions as recommended by the supplier (ThermoFisher). Cells were grown in Expi293 medium supplemented with 0.5x antibiotic-antimyotic (Thermo Fisher, 15240062) at 37 ºC, 8% CO_2, in sterile plastic Erlenmeyer flasks with ventilated caps, and shaken using a CO_2-resistant orbital shaker at 125 rpm. Cultures were maintained by seeding at a 0.3 million viable cells/mL, then expanding cultures every 4th day when they reached a density of 3-5 million viable cells/mL. Cell density and viability were measured using the trypan blue exclusion method with a Bio-Rad TC20™ automated cell counter.

An adaptation of a previously reported PEI-based transfection method was used for antibody expression.[66] After reaching 3-5 million viable cells/mL at a minimum of 95% viability, cells were fuged at 200 x g for 6 min and resuspended in fresh media to 20 million viable cells/mL. To the concentrated cultures was added a 1:1 ratio of pCDNA3.1-HC-121-TGA-Trastuzumab and

pAcBac-EcWRS-H14-1x-EcWtR(TGA) at a final DNA concentration of 25 μg/mL. A concentrated stock solution of 40 kDa linear PEI (40 mg/mL) was added to a final concentration of 50 μg/mL. Cells were incubated with shaking under standard expansion conditions for 3 hours, then diluted to 3-4 million viable cells/mL, and sterile (0.22 μm filtered) 100 mM stocks of 5HTP and valproic acid were added to 1 mM and 2 mM final concentrations, respectively.

After 7-12 days of expression, Expi293 supernatant was harvested by pelleting cells at 800 x g for 8 minutes, then passing the decanted media through a sterile filter (0.22 μm). The media was adjusted to pH 5.4, 50 mM NaOAc using a 500 mM stock. After equilibrating 1 mL Pierce™ Protein G resin (Thermo Scientific, 20398) with 15 column volumes of wash buffer (pH 5.4, 50 mM NaOAc), the filtered media was passed through by gravity flow (300-600 mL). Protein G resin was washed with a volume of 50 mM NaOAc, pH 5.4 equal to approximately half the volume of added media (150-300 mL). Purified antibody was eluted by 16-20 stepwise additions of 1 mL pH 2.7, 100 mM glycine, collecting each elution in tubes containing 120 μL pH 7.4 1 M sodium phosphate for immediate neutralization. Eluted fractions were qualitatively tested for the presence of protein by adding 10 μL elution fraction to 10 μL Bradford reagent (Thermo Scientific 1856209). Fractions containing protein were pooled and was concentrated to ~ 1-2 mg/mL final protein concentration using Amicon® Ultra 15 (Millipore Sigma UFC901024) centrifugal filters. Concentrated antibody was buffer-exchanged using a PD midiTrap G-25 column (GE Healthcare 28918008) pre-equilibrated with 20 column volumes of PBS. Protein concentration was determined by absorption measurement using a NanoDrop 2000 spectrophotometer (Fisher Scientific ND2000).

Protein purity and protein conjugation state was determined by HPLC-coupled ESI-MS of fully reduced antibodies (Agilent Technologies, 1260 Infinity ESI–TOF). First, 20 μg antibody sample was adjusted to 50 mM phosphate pH 7.0, 10 mM fresh TCEP, then heated to 55 ºC for 10 minutes. After cooling to 37 ºC, antibody was incubated 1 μL Remove-iT® PNGase (New England Biolabs, P0706S) for 60 minutes. To remove PNGase, digested antibody was incubated with 25 μL of settled chitin resin (New England Biolabs, S6651L) on an orbital shaker for 15 minutes, then carefully removed as supernatant after centrifuging at 15,000 x g for 1 minute. 250 – 500 ng of protein was used for HPLC-coupled MS analysis as described above.

MTT Assay

SK-BR-3 or MDA-MB-231 cells were adjusted to $3*10^5$ cells/mL in DMEM, high glucose, supplemented with 10% FBS, then added in 100 μL aliquots to a 96 well plate. Cells were incubated for 24 hours at 37 ºC, 5 % CO_2 in a humidified mammalian cell incubator. For antibody drug conjugate toxicity testing, media was replaced with fresh DMEM media alone or containing 10 three-fold dilutions of 121-HC-5HTP-Traztusumab or 121-HC-5HTP-Traztusumab+MMAF. After five days, media was replaced with fresh DMEM high glucose (without phenol red, Corning 99-663-CV) containing 0.5 mg/mL MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) and incubated for 4 hours at 37 ºC. The wells were gently washed once with PBS at 37 ºC after removing media and then permeabilized with 200 μL DMSO to reveal intracellular formazan production. Absorbance was measured at 595 nm and viability was measured by

interpolating between average absorbance values of the positive (no treatment) and negative (no cells) control wells.

Flow cytometry

After reaching confluence, adherent cultures of SK-BR-3 and HEK293T cells were detached from the plate surfaces using 0.25% trypsin (HyClone) for 2 minutes at 37 ºC. Trypsin was quenched with two volumes of ice-cold DMEM + 10% FBS, centrifuged at 2,000 xg for 5 minutes, then resuspended in ice-cold PBS at 1-2 million cells/mL. Cells were treated with 10 nM Trastuzumab-HC-121-5HTP-fluorescein-aniline (**14)** conjugate and gently rocked for 15 minutes at 4 ºC. Cells were washed once and resuspended in ice-cold PBS, then analyzed by flow cytometry on a BD Accuri C6 Plus flow cytometer (BD Biosciences) using the FITC settings (excitation wavelength: 488 nm, standard filter: 533/30). 10,000 events were collected in BD Accuri C6 Plus. version 1.0.23.1, and data were processed using FlowJo, version 10.8.1. For gating, a plot of side scatter area vs. forward scatter area was used to exclude cell debris and large cell clumps; a plot of forward scatter height vs. forward scatter area was used to identify single cells and exclude doublets; FITC fluorescence of all single cells was plotted.

Additional information

Supplementary information is available for this paper.

Data availability statement

Data supporting the findings of this study are available within the paper and its Supplementary Information. Crystallographic data for the structure reported in this Article has been deposited at the Cambridge Crystallographic Data Centre, under deposition number CCDC 2179454 (**19**). Copies of the data can be obtained free of charge via https://www.ccdc.cam.ac.uk/structures

Methods-only References:

64 Lang, K. *et al.* Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat Chem* **4**, 298-304, doi:10.1038/nchem.1250 (2012).

65 Zheng, Y., Addy, P. S., Mukherjee, R. & Chatterjee, A. Defining the current scope and limitations of dual noncanonical amino acid mutagenesis in mammalian cells. *Chem Sci* **8**, 7211-7217, doi:10.1039/c7sc02560b (2017).

66 Backliwal, G. *et al.* Rational vector design and multi-pathway modulation of HEK 293E cells yield recombinant antibody titers exceeding 1 g/l by transient transfection under serum-free conditions. *Nucleic acids research* **36**, e96, doi:10.1093/nar/gkn423 (2008).