Site-specific bioconjugation through enzyme catalyzed tyrosine-cysteine bond formation

Marco J. Lobba[1], Christof Fellmann[2,3,4,*], Alan M. Marmelstein[1], Johnathan C. Maza[1], Elijah N. Kissman[1], Stephanie A. Robinson[1], Brett T. Staahl[2], Cole Urnes[2], Rachel J. Lew[3], Casey S. Mogilevsky[1], Jennifer A. Doudna[1,2,3,5,6,*], Matthew B. Francis[1,7,*]

[1]Department of Chemistry, University of California, Berkeley, CA 94720, USA.

[2]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.

[3]Gladstone Institutes, San Francisco, CA 94158, USA.

[4]Department of Cellular and Molecular Pharmacology, School of Medicine, University of California, San Francisco, CA 94158, USA.

[5]Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA.

[6]Innovative Genomics Institute, University of California, Berkeley, CA 94720, USA.

[7]Materials Sciences Division, Lawrence Berkeley National Laboratories, Berkeley, CA 94720, USA.

[*]Correspondence should be addressed to C.F. (christof.fellmann@gladstone.ucsf.edu), J.A.D. (doudna@berkeley.edu), or M.B.F (mfrancis@berkeley.edu).

Abstract: The synthesis of protein-protein and protein-peptide conjugates is an important capability for producing vaccines, immunotherapeutics, and targeted delivery agents. Herein we show that the enzyme tyrosinase is capable of oxidizing exposed tyrosine residues into *ortho*-quinones that react rapidly with cysteine residues on target proteins. This coupling reaction occurs under mild aerobic conditions and has the rare ability to join full-size proteins in under 2 h. The utility of the approach is demonstrated for the attachment of cationic peptides to enhance the cellular delivery of CRISPR-Cas9 by 20-fold, and for the coupling of reporter proteins to a cancer-targeting antibody fragment without losing its cell-specific binding ability. The broad applicability of this technique provides a new building block approach for the synthesis of protein chimeras.

Main Text:

The covalent modification of proteins while preserving native function has long been a goal of chemical biology research[1–3]. Unfortunately, it can be difficult to target chemically distinct functional groups in protein sequences, leading to heterogeneous product mixtures with many modification strategies. This challenge is particularly notable for the synthesis of protein-protein and protein-peptide conjugates, which can serve as highly valuable constructs for vaccines, immunotherapy agents, signaling peptides with enhanced circulation properties, and targeted delivery vehicles. Currently, the coupling of two biomolecules is usually achieved using heterobifunctional crosslinking agents (which are rarely site-selective for both reactive species)[1] or Click-type reactions[4] between bioorthogonal functional groups[5] that must be introduced using artificial amino acid incorporation, metabolic engineering, or additional synthetic steps. The coupling of native functional groups on proteins and peptides can be achieved through native chemical ligations and sortase-based techniques[6,7], but these approaches can only be used at the N- and C-termini and can limit the quantities of bioconjugates that can be obtained. In some cases it is possible to fuse multiple protein sequences at the genetic level for recombinant expression[8,9]. While this is a time-tested approach for producing chimeric proteins, such methods often require significant engineering to determine optimal linker lengths[10], can encounter folding problems during the expression of constructs of increasing size[11], and are again limited to attachment via the N and C termini.

In this report, a new method is described for the efficient generation of protein-peptide and protein-protein conjugates through the direct coupling of solvent-exposed tyrosine residues to cysteine sulfhydryl groups. This reaction requires only catalytic amounts of tyrosinase and adventitious oxygen, affording complex, multicomponent bioconjugates in under 1 h under mild reaction conditions and low stoichiometric excess. This strategy is demonstrated as a convenient method for the installation of cell penetrating peptides on proteins, such as Cas9, and the coupling of entire protein domains. Most importantly, this reaction demonstrates the very rare ability to couple full-size proteins together through positionally defined linkages using native amino acids, providing a powerful new tool for the construction of multifunctional biomolecular constructs.

Results and Discussion

In previous work, our lab has shown that proteins bearing aniline and N-terminal proline nucleophiles can be modified with *o*-quinone and *o*-iminoquinone electrophiles in high yields[12,13]. The oxidized intermediates in these reactions have typically been generated through the *in situ* oxidation of *o*-aminophenols or similar species using periodate[12] or ferricyanide anions[13]. While effective in many contexts, these stoichiometric reagents can exhibit collateral oxidation of methionine thioethers and surface exposed cysteines to varying degrees. A recent breakthrough in this chemistry has involved the use of the tyrosinase enzymes to generate the same reactive *o*-quinone electrophiles using simple phenols and oxygen[14]. The small amount of molecular oxygen that is dissolved in the buffers is sufficient for the reaction to proceed, and water is the only byproduct. This reaction addresses two important limitations of the previous versions of this chemistry. First, it is no longer necessary to prepare and install the oxygen-sensitive *o*-aminophenol and *o*-catechol groups; instead, simple tyramine and tyrosine derivatives can be used. Second, the need for an additional oxidant that must later be removed is avoided. Importantly, the enzymatic method for triggering these reactions is selective for phenols that are exposed in solution and does not cause off-target oxidation of sensitive functional groups.

It has been known for some time that tyrosinases can act upon tyrosine residues in proteins and peptides[15]. Typically these reactions have been used to achieve protein crosslinking at high substrate concentrations, involving indiscriminate conversion of tyrosine residues within the peptide backbone to *o*-quinines followed by reactions with a variety of nucleophiles[16–19]. Previous reports have followed reaction progress using UV or other spectroscopic analyses, and have explored the impact of tyrosine oxidation on enzyme activity; however the sequence specificities and sites of oxidation have not been characterized well. In terms of site-selective coupling chemistry, other labs have recently shown that tyrosine-tagged immunoglobulins can be oxidized and subjected to subsequent hetero-Diels-Alder reactions with cyclooctynes[20], and abTYR has been used to activate an ovalbumin peptide tag for coupling to substituted boronic acids[21]. Work from our lab has showed that abTYR and a new tyrosinase derived from *Bacillus megatarium* can activate phenols in both small molecule and protein substrates for coupling to N-terminal proline residues and anilines (Figure 1a, top and bottom reaction pathways)[14,22]. Taken together, these studies highlight the utility of tyrosinase enzymes to generate transient *o*-quinone electrophiles in dilute aqueous solutions, allowing highly chemoselective couplings that are not realized at the higher substrate concentrations evaluated previously.

Key to the present work is the observation that cysteine thiols on full-size biomolecules participate readily in this chemistry at low equivalencies, allowing the first general coupling reaction for two native and readily encodable functional groups on protein surfaces (Figure 1a, middle pathway). Although we have known that these residues can participate in oxidative coupling reactions to varying degrees, the reliance on stoichiometric oxidants in our previous studies led to numerous reaction byproducts that are completely avoided using the new tyrosinase oxidation system. Other reports have noted reactivity between small molecule thiols and *o*-quinones, but with minimal product characterization[23,24].

In the context of site-selective bioconjugation, this reactivity pathway was first tested on genome-free MS2 bacteriophage capsids bearing N87C mutations in each of the subunit proteins[25,26]. Each of these assemblies provides 180 cysteine residues facing the inside of a 27 nm sphere[27]. The capsid structures possess a series of 2 nm pores that are large enough for peptides and small molecules to access the inner surface (Figure 1b)[26,28,29]. Modification was achieved by exposing 10 μM (based on monomer) MS2 N87C to 200 μM α-endorphin peptide (Figure 1c) in the presence of 167 nM abTYR for 30 min at room temperature in 20 mM phosphate buffer, pH 6.5. Clean conversion of each capsid monomer to a single peptide conjugate was observed (Figure 1d**)**. This represents the installation of >170 peptide groups inside each capsid despite the high degree of steric crowding. Subsequent exposure to the common cysteine capping reagent *N*-ethylmaleimide (NEM) showed no further modification (Figure 1e, Figure S1), supporting the cysteine selectivity of the reaction. Similarly, no α-endorphin addition was observed for MS2 N87C that had been previously coupled to NEM, and there was no modification of wild type MS2 capsids lacking exposed cysteine residues (Figure S1). Equally importantly, there was no background oxidation observed on any of the MS2 proteins, which each contain four tyrosine residues (Figure 1c, Figure S1). This is presumably because tyrosine residues embedded in secondary and tertiary structural elements cannot access the deep binding pocket of the enzyme active site[30]. These results demonstrate that covalent bonds can be formed between these two native amino acid residues using an efficient and operationally simple procedure.

In order to characterize the chemical structure of the bioconjugation product, we conducted NMR experiments for a small molecule analog. Somewhat surprisingly, the product results from thiol addition to the 5 position of the *o*-quinone ring (numbering from the amino acid connection point as shown in Figure 1a), and it remains predominantly in the catechol form (Figure 2a, Figures S2, S3). Additional MS data further support the catechol structure. This product contrasts with the previously observed structures of proline and aniline additions to *o*-quinone species that add at the 6-position and tend to remain in the oxidized quinone form (Figure 1a)[12–14]. MS data confirmed that the thiol addition products on larger biomolecules also existed predominantly in the reduced catechol form.

To establish the degree of quinoid versus catechol nature of the product further, we exposed both thiol and N-terminal proline coupling adducts of Y200C GFP to CH_3ONH_2 (methoxyamine) overnight to effect oxime formation with any ketone groups that were present, as has been recently shown[31]. Notably, only the proline adducts exhibited any detectable reactivity to the methoxyamine, (Figure S4).

The energetics of the *o*-quinone-thiol coupling reaction were investigated using DFT calculations. Briefly, geometry optimized structures were evaluated using B3LYP-D3/6-31G** to obtain vibrational spectra, and ωB97M-V/6-311G^{++}3df-3pd was used to obtain accurate ground state electronic energies[32]. These results were combined to yield estimates of the thermodynamic parameters for the reaction. A CPCM solvent model[33–35] was used for all of the calculation results appearing in Figure 2b; gas phase versions of the same calculations are included in Figure S5. This study indicated that the reaction is highly favorable ($\Delta H^o = -38.0$ kcal/mol and $\Delta G^o = -25.3$ kcal/mol), which is a common feature associated with Click-type reactions and serves to validate why the reactions proceed despite the high degree of building steric bulk[4]. For comparison, the addition of a thiol nucleophile to a maleimide group affords more modest values ($\Delta H^o = -24.4$ kcal/mol and $\Delta G^o = -10.1$ kcal/mol) using the same computational approach (Figure S5). Calculations were also used to compare the relative oxidation propensities for the initial thiol- and aniline-addition products relative to a methyl-substituted *o*-quinone. While such an equilibrium was strongly favorable for the formation of the oxidized aniline product (Figure 2c), it was disfavored for the thiol addition product (Figure 2d). This relative comparison is in line with the observation that the thiol addition product remains predominantly in the reduced catechol form.

To test protein conjugation product stability, we subjected α-endorphin modified MS2 to buffers ranging from pH 5 to pH 8, potassium ferricyanide, 5 mM tris-carboxyethyl phosphine (TCEP), and 5 mM n-mercaptobutanol for up to 6 days at room temperature. None of the conditions led to greater than 5% loss of the attached peptides (Supplementary Table 1, Figure S6). Additionally, GFP modified with an RRRRY peptide showed no reversion when kept at 37 °C for up to 7 days, even when exposed to up to 4 mM glutathione (GSH). It should be noted that small amounts of GSH and other small molecule thiols were shown to add into the *o*-hydroquinone ring over the course of several days (Figure S7), presumably after aerobic oxidation to form small equilibrium amounts of the *o*-quinone; however, these adducts did not appear to affect the integrity of the linkage formed in the initial coupling reaction.

Stability tests were also conducted in human blood serum after coupling Y200C sfGFP to a biotin moiety through either maleimide- or tyrosinase-based conjugation. The resulting samples were kept in the serum at room temperature or at 37 °C for up to 7 days. At several time points during incubation, aliquots were removed and subjected to streptavidin bead capture, followed by a final elution using 80% acetonitrile (ACN), 5% formic acid (FA) and 2 mM biotin (Figure

S8)[36]. We found that the maleimide coupled product decreased over time, and at 7 days was no longer detectable in either sample (Figure S9). This result is consistent with literature observations for these types of adducts on antibodies and other proteins[37]. In contrast, the tyrosinase-coupled product was recovered throughout the course of the experiment, indicating that the tyrosinase coupling product has substantially greater stability in serum.

In addition to α-endorphin (Figure 3a), the tyrosine-cysteine coupling chemistry has proven remarkably successful for the site-selective attachment of additional peptides. Both super-folder GFP (sfGFP) bearing the Y200C mutation (GFP Y200C, Fig. 3a,b) and the N87C MS2 coat protein (Figure S10) were coupled to tyrosines at the ends of peptides derived from interleukin-13 (IL-13), the Simian Vacuolating Virus 40 (SV40) large T-antigen[38,39], HIV Tat[40], and polyarginine[40] with excellent conversion (Fig. 3b). It is important to note that N-terminal acylation or an additional glycine was needed in these cases to prevent dopachrome-like cyclization of the oxidized N-terminal tyrosine residues. Substrate compatibility tests using the tyrosinase activation/thiol addition protocol also showed that small molecule phenols, including biotin, rhodamine, and 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA) all resulted in high conversion of the starting protein to the expected products within 30 min (Figure S11). Some secondary adducts were observed with these small molecule substrates for sfGFP, which are believed to be the products of competing N-terminal modification (Figure S11)[14]. The diversity of small molecules and peptides compatible with this coupling strategy demonstrates the utility of the reaction as a general tool for protein modification. Our previous studies of aniline-based oxidation coupling reactions have indicated that *o*-quinone formation is rate-limiting[41]. This is likely the case for the thiol additions as well, complicating the calculation of the second-order rate constant for the key coupling step. Even considering the enzymatic activation requirement, these reactions still reach completion in 1 h with coupling partner concentrations well below 100 μM.

An application of this chemistry for the modification of large proteins and enzymes was demonstrated for CRISPR-Cas9 ribonucleoprotein (RNP) complexes. Cas9, an RNA-guided DNA endonuclease used widely for genome editing, contains two surface cysteines at positions 80 and 574, thus providing two native bioconjugation handles for peptide attachment. To generate a modified form of Cas9 bearing two peptides, coupling reactions were conducted using Cas9 bearing two copies of the SV40 large T-antigen nuclear localization sequence (NLS) to facilitate nuclear entry. Quantitative conversion was observed to the doubly modified protein in both the apo and RNA bound forms (Figure 3c). Importantly, tyrosinase-modified Cas9 RNPs fully retained their target cleavage activity when compared to unmodified Cas9 RNPs (Figure 3d). Building off of the success of peptide conjugates, we applied the tyrosinase conjugation method to the coupling of whole protein domains. We found that tyrosinase could be used to ligate Cas9 to green fluorescent protein with no detriment to DNA cleavage ability (MYGGS-GFP, Figure 2c,e**,** Figure S12). It is important to note that the mass spectral data revealed no evidence of Cas9 oxidation, nor did they indicate significant amounts of off-target modification. These experiments underscore the exquisite chemoselectivity of this reaction as the Cas9 constructs contain 158 lysine and 58 tyrosine residues in addition to representatives of the 18 other native amino acids, providing ample opportunities for off-target reactions. Despite having a highly reactive system, this method has exceptionally predictable conjugation outcomes.

With these results in hand, we examined whether we could attach peptides capable of mediating Cas9 delivery. A genetic fusion of four consecutive SV40 NLSs to the N-terminus of Cas9 bearing two additional C-terminal NLSs has been shown to enable delivery of Cas9 RNPs to human neural progenitor cells (NPCs) in culture and *in vivo*[38]. However, the production of such complexes is lengthy, requires dedicated cloning and protein purification, and limits the locations of the delivery peptides to the native N- and C-termini. Moreover, most commercially produced Cas9 RNPs demonstrate little-to-no cell penetrating capabilities[38]. Thus, we examined whether standard Cas9 RNP complexes could be chemically modified with peptides or proteins for editing in human NPCs[38]. Three different cell-penetrating peptides were coupled to the native cysteine residues of a Cas9 RNP bearing two C-terminal NLS sequences, which does not appreciably edit cells in the NPC assay on its own (Figure 4a). After the reaction, ESI-TOF MS analysis confirmed that two copies of each tested peptide coupled quantitatively to Cas9 (Figure S13). These constructs, plus Cas9RNP with sfGFP, were then assayed for their editing efficiencies[42–45]. Cas9 conjugates bearing +36 GFP proteins precipitated upon guide RNA addition, and thus could not be evaluated in this experiment. The Cas9RNP construct bearing two 2SV40 peptides (each peptide is composed of two consecutive SV40 NLSs) showed a 20-fold increase in editing efficiency (Figure 4a and Figure S14) relative to the unmodified 2NLS-Cas9RNP. This indicates that the delivery of Cas9 RNP can be mediated by direct peptide coupling to the protein surface of a readily available standard Cas9 protein. The addition of other peptides and sfGFP did not improve editing efficiency. In addition, experiments involving the original 4NLS Cas9 construct showed successful modification, but the added protein domains did not increase editing activity further (Figure S15**)**.

Though Cas9-GFP conjugates showed no change in cell uptake, the capability to couple two protein domains so efficiently presents an attractive prospect that has been pursued for some time by the field. To test the scope of the reaction, we generated versions of nano-luciferase (nanoLuc) and a HER2 binding scFv with tyrosine tags and combined them with Y200C sfGFP.

The nanoLuc was produced with either an N-terminal or C-terminal tag, and coupling was performed using 10 μM nanoLuc with 20 μM Y200C GFP in pH 6.5 phosphate buffer for 30 min at RT. Both constructs yielded over 80% conversion, (as determined using ESI-TOF MS) to the singly modified nanoLuc-GFP conjugates (Figure S16). The reaction was similarly successful for the coupling of Y200C sfGFP (~7 μM by absorbance at 485 nm using an extinction coefficient of 83,300 $M^{-1}cm^{-1}$) to a HER2-binding short chain variable fragment (scFv)-GGSY construct (5 μM) in pH 6.5 phosphate buffer at room temperature for 30 min. These conditions quantitatively converted the scFv into a tagged GFP (Figure 4b,c). Tagged GFP constructs were then incubated with resuspended $HER2^+$ and $HER2^-$ cells (SK-BR-3 and MDA-MB-468, respectively), at 2 μg per 100 μL of cell suspension. Binding was tracked via flow cytometry, where the scFv-GFP construct bound specifically to HER2 expressing cells, but not HER2-negative cells (Figure 4d). No cellular fluorescence was observed with either GFP alone or in controls where GFP was not coupled to the scFv (Figures S17, S18). These and corresponding control experiments showed that only highly exposed tyrosines near the N- or C-termini are capable of participating in the reaction due to the lack of further substrate oxidation[22]. This was quite important, as the scFv is known to have 15 native tyrosine residues, with 8 located in the antigen binding site[46]. Furthermore, these results demonstrate the utility of tyrosinase mediated conjugation for site-specific protein-protein coupling while preserving targeted immunoglobulin binding affinity.

Conclusion

This study establishes that tyrosinase activation can achieve site-specific protein-protein coupling through the formation of cysteine-tyrosine linkages. This simple approach allows for the targeted modification of proteins with a diverse array of molecules in potentially any desired location using an inexpensive, commercially available enzyme as the catalyst. It does not require additional modification steps or exposure to aggressive reaction conditions, conserving protein folding and function. It also offers the advantage of moving the attachment sites by changing the location of the introduced thiol component, which is an inherent advantage of cysteine-based modification chemistry. In particular, the success of Cas9 delivery via peptide conjugation highlights a potential application for tyrosinase mediated coupling to facilitate the screening of candidate peptides for the cellular entry of large biomolecules without the need to clone, express, and purify individual genetic fusions for each combination. Finally, its ability to couple full-sized proteins together with high yields and minimal stoichiometric excesses suggests broad potential for the generation of protein chimeras for biotherapeutic, vaccine, and bioanalytical applications.

Acknowledgments: The authors wish to thank Jamie Gleason for the kind gift of the plasmid containing GFP Y200C, Dr. Jeffrey Pelton for help with 900 MHz NMR data collection and analysis, and UCB Cell Culture Facility which is supported by The University of California Berkeley for providing cell cultures. **Funding:** MBF was supported by the NSF (CHE-1808189). Funds for the 900 MHz NMR spectrometer were provided by the NIH through grant GM68933. CF was supported by a US NIH Pathway to Independence Award (K99GM118909, R00GM118909) from the National Institute of General Medical Sciences (NIGMS). JAD receives funding from the William M. Keck Foundation, a Collaborative MS Research Center Award from the National Multiple Sclerosis Society, the Centers for Excellence in Genomic Science of the National Institutes of Health under award number RM1HG009490, the Somatic Cell Genome Editing Program of the Common Fund of the National Institutes of Health under award number U01AI142817-02, and the National Science Foundation under award number 1817593. JAD is an Investigator of the Howard Hughes Medical Institute and is a Paul Allen Distinguished Investigator. **Author Contributions:** MJL developed the project idea, conducted coupling reactions, ran NMR experiments, and prepared the manuscript and figures. CF ran cell culture assays with the help of RJL, assisted with writing, assisted with flow cytometry data analysis, and aided in experimental design. AMM synthesized and purified the small molecule product, assisted in the structural characterization, provided Trastuzumab scFV samples, and ran scFv-GFP cell coupling studies. ENK ran MS2 couplings, stability studies, and NEM capping experiments. JCM contributed to experimental design, provided tyrosinase stocks, and assisted in data analysis. CM assisted in reagent preparation and conducted stability studies. SAR assisted in Cas9 activity assays and the initial Cas9 coupling reactions. BTS provided initial samples of 4NLS Cas9 protein for NPC assays and established the NPC line and assay protocols. CJU ran NPC assays. JAD provided funding support, assisted in experimental design, and contributed to the text and figures. MBF provided funding support, assisted in data analysis, assisted with experimental design, and contributed to the text and figures. **Competing Interests:** J.A.D. is an Investigator of the Howard Hughes Medical Institute and executive director of the Innovative Genomics Institute at the University of California, Berkeley and the University of California, San Francisco. J.A.D. is a co-founder of Editas Medicine, Intellia Therapeutics, and Caribou Biosciences and a scientific adviser to Caribou, Intellia, eFFECTOR Therapeutics and Driver. The Regents of the University of California have patents pending for CRISPR technologies and the use of tyrosinase in the abovementioned manner on which the authors are inventors. **Data and materials availability:** All data are available in the main text or the supplementary materials.

Supplementary Materials:

Materials and Methods

Supplementary Figures S1-S18

Supplementary Table S1

External Database S1

References [14,22,38,43,47–51]

References and Notes

1. Hermanson, G. T. *Bioconjugate Techniques*. *Bioconjugate Techniques* (Elsevier, 2008). doi:10.1016/B978-0-12-370501-3.X0001-X
2. Witus, L. S. & Francis, M. B. Using synthetically modified proteins to make new materials. *Acc. Chem. Res.* **44**, 774–783 (2011).
3. Stephanopoulos, N. & Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nature Chemical Biology* **7**, 876–884 (2011).
4. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie Int. Ed.* **40**, 2004–2021 (2001).
5. Sletten, E. M. & Bertozzi, C. R. Bioorthogonal chemistry: Fishing for selectivity in a sea of functionality. *Angew. Chemie - Int. Ed.* **48**, 6974–6998 (2009).
6. Dawson, P., Muir, T., Clark-Lewis, I. & Kent, S. Synthesis of proteins by native chemical ligation. *Science (80-.).* **266**, 776–779 (1994).
7. Theile, C. S. *et al.* Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8**, 1800–1807 (2013).
8. Kim, Y. G. & Chandrasegaran, S. Chimeric restriction endonuclease. *Proc. Natl. Acad. Sci.* **91**, 883–887 (1994).
9. Bolukbasi, M. F. *et al.* Orthogonal Cas9-Cas9 chimeras provide a versatile platform for genome editing. *Nat. Commun.* **9**, 4856 (2018).
10. Oakes, B. L. *et al.* CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification. *Cell* **176**, 254-267.e16 (2019).
11. Schmidt, S. R. *Fusion Protein Technologies for Biopharmaceuticals: Applications and Challenges*. *Fusion Protein Technologies for Biopharmaceuticals: Applications and Challenges* (John Wiley & Sons, Inc., 2013). doi:10.1002/9781118354599
12. Behrens, C. R. *et al.* Rapid chemoselective bioconjugation through oxidative coupling of anilines and aminophenols. *J. Am. Chem. Soc.* **133**, 16398–16401 (2011).

13. Obermeyer, A. C., Jarman, J. B. & Francis, M. B. N-Terminal Modification of Proteins with o -Aminophenols. *J. Am. Chem. Soc.* **136**, 9572–9579 (2014).

14. Maza, J. C. *et al.* Enzymatic Modification of N-Terminal Proline Residues Using Phenol Derivatives. *J. Am. Chem. Soc.* **141**, 3885–3892 (2019).

15. Cory, J. G. & Frieden, E. Differential Reactivities of Tyrosine Residues of Proteins to Tyrosinase. *Biochemistry* **6**, 121–126 (1967).

16. Ito, S., Kato, T., Shinpo, K. & Fujita, K. Oxidation of tyrosine residues in proteins by tyrosinase. Formation of protein-bonded 3,4-dihydroxyphenylalanine and 5-S-cysteinyl-3,4-dihydroxyphenylalanine. *Biochem. J.* **222**, 407–411 (1984).

17. Selinheimo, E., Autio, K., Kruus, K. & Buchert, J. Elucidating the Mechanism of Laccase and Tyrosinase in Wheat Bread Making. *J. Agric. Food Chem.* **55**, 6357–6365 (2007).

18. Lantto, R., Puolanne, E., Kruus, K., Buchert, J. & Autio, K. Tyrosinase-Aided Protein Cross-Linking: Effects on Gel Formation of Chicken Breast Myofibrils and Texture and Water-Holding of Chicken Breast Meat Homogenate Gels. *J. Agric. Food Chem.* **55**, 1248–1255 (2007).

19. Thalmann, C. R. & Lötzbeyer, T. Enzymatic cross-linking of proteins with tyrosinase. *Eur. Food Res. Technol.* **214**, 276–281 (2002).

20. Bruins, J. J. *et al.* Inducible, Site-Specific Protein Labeling by Tyrosine Oxidation–Strain-Promoted (4 + 2) Cycloaddition. *Bioconjug. Chem.* **28**, 1189–1193 (2017).

21. Montanari, E. *et al.* Tyrosinase-Mediated Bioconjugation. A Versatile Approach to Chimeric Macromolecules. *Bioconjug. Chem.* **29**, 2550–2560 (2018).

22. Marmelstein, A. M. *et al.* Tyrosinase-Mediated Oxidative Coupling of Tyrosine Tags on Peptides and Proteins. *J. Am. Chem. Soc.* **142**, 5078–5086 (2020).

23. Ito, S. & Prota, G. *A facile one-step synthesis of cysteinyldopas using mushroom tyrosinase*. *Experientia* **33**, 1118–1119 (Birkhäuser-Verlag, 1977).

24. Agrup, G., Hansson, C., Rorsman, H. & Rosengren, E. The effect of cysteine on oxidation of tyrosine dopa, and cysteinyldopas. *Arch. Dermatol. Res.* **272**, 103–115 (1981).

25. Wu, W., Hsiao, S. C., Carrico, Z. M. & Francis, M. B. Genome-free viral capsids as multivalent carriers for taxol delivery. *Angew. Chemie - Int. Ed.* **48**, 9493–9497 (2009).

26. Capehart, S. L., Coyle, M. P., Glasgow, J. E. & Francis, M. B. Controlled integration of gold nanoparticles and organic fluorophores using synthetically modified MS2 viral capsids. *J. Am. Chem. Soc.* **135**, 3011–3016 (2013).

27. Valegård, K., Liljas, L., Fridborg, K. & Unge, T. The three-dimensional structure of the bacterial virus MS2. *Nature* **345**, 36–41 (1990).

28. Farkas, M. E. *et al.* PET Imaging and Biodistribution of Chemically Modified Bacteriophage MS2. *Mol. Pharm.* **10**, 69–76 (2013).

29. Carrico, Z. M., Romanini, D. W., Mehl, R. A. & Francis, M. B. Oxidative coupling of peptides to a virus capsid containing unnatural amino acids. *Chem. Commun.* **0**, 1205 (2008).

30. Muñoz-Muñoz, J. L. *et al.* Catalytic oxidation of o-aminophenols and aromatic amines by mushroom tyrosinase. *Biochim. Biophys. Acta - Proteins Proteomics* **1814**, 1974–1983 (2011).

31. Maza, J. C. *et al.* Secondary Modification of Oxidatively-Modified Proline N-termini for the Construction of Complex Bioconjugates. *Org. Biomol. Chem.* **18**, 1881–1885 (2020).

32. Mardirossian, N. & Head-Gordon, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Molecular Physics* **115**, 2315–2372 (2017).

33. Cossi, M., Rega, N., Scalmani, G. & Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **24**, 669–681 (2003).

34. Barone, V. & Cossi, M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A* **102**, 1995–2001 (1998).

35. Takano, Y. & Houk, K. N. Benchmarking the conductor-like polarizable continuum model (CPCM) for aqueous solvation free energies of neutral and ionic organic molecules. *J. Chem. Theory Comput.* **1**, 70–77 (2005).

36. Schiapparelli, L. M. *et al.* Direct detection of biotinylated proteins by mass spectrometry. *J. Proteome Res.* **13**, 3966–3978 (2014).

37. Fontaine, S. D., Reid, R., Robinson, L., Ashley, G. W. & Santi, D. V. Long-Term Stabilization of Maleimide–Thiol Conjugates. *Bioconjug. Chem.* **26**, 145–152 (2015).

38. Staahl, B. T. *et al.* Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes. *Nat. Biotechnol.* **35**, 431–434 (2017).

39. Pandya, H., Gibo, D. M., Garg, S., Kridel, S. & Debinski, W. An interleukin 13 receptor α 2-specific peptide homes to human Glioblastoma multiforme xenografts. *Neuro. Oncol.* **14**, 6–18 (2012).

40. Wender, P. A. *et al.* The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13003–13008 (2000).

41. ElSohly, A. M. *et al.* ortho-Methoxyphenols as Convenient Oxidative Bioconjugation Reagents with Application to Site-Selective Heterobifunctional Cross-Linkers. *J. Am. Chem. Soc.* **139**, 3767–3773 (2017).

42. Allen, J., Brock, D., Kondow-McConaghy, H. & Pellois, J.-P. Efficient Delivery of Macromolecules into Human Cells by Improving the Endosomal Escape Activity of Cell-Penetrating Peptides: Lessons Learned from dfTAT and its Analogs. *Biomolecules* **8**, 50 (2018).

43. Cronican, J. J. *et al.* Potent Delivery of Functional Proteins into Mammalian Cells in Vitro and in Vivo Using a Supercharged Protein. *ACS Chem. Biol.* **5**, 747–752 (2010).

44. Lönn, P. *et al.* Enhancing Endosomal Escape for Intracellular Delivery of Macromolecular Biologic Therapeutics. *Sci. Rep.* **6**, 32301 (2016).

45. Futaki, S. *et al.* Arginine-rich peptides. An abundant source of membrane-permeable

peptides having potential as carriers for intracellular protein delivery. *J. Biol. Chem.* **276**, 5836–40 (2001).

46. Eigenbrot, C., Randal, M., Presta, L., Carter, P. & Kossiakoff, A. A. X-ray structures of the antigen-binding domains from three variants of humanized anti-p185HER2 antibody 4D5 and comparison with molecular modeling. *J. Mol. Biol.* **229**, 969–995 (1993).

47. Fling, M., Horowitz, N. H. & Heinmann, S. F. The isolation and properties of crystalline tyrosinase from Neurospora. *J. Biol. Chem.* **238**, 2045–53 (1963).

48. Obermeyer, A. C., Jarman, J. B. & Francis, M. B. N-Terminal Modification of Proteins with o -Aminophenols. *J. Am. Chem. Soc.* **136**, 9572–9579 (2014).

49. Anders, C. & Jinek, M. In Vitro Enzymology of Cas9. in *Methods in enzymology* **546**, 1–20 (Elsevier Inc., 2014).

50. Tong, G. J., Hsiao, S. C., Carrico, Z. M. & Francis, M. B. Viral capsid DNA aptamer conjugates as multivalent cell-targeting vehicles. *J. Am. Chem. Soc.* **131**, 11174–11178 (2009).

51. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).

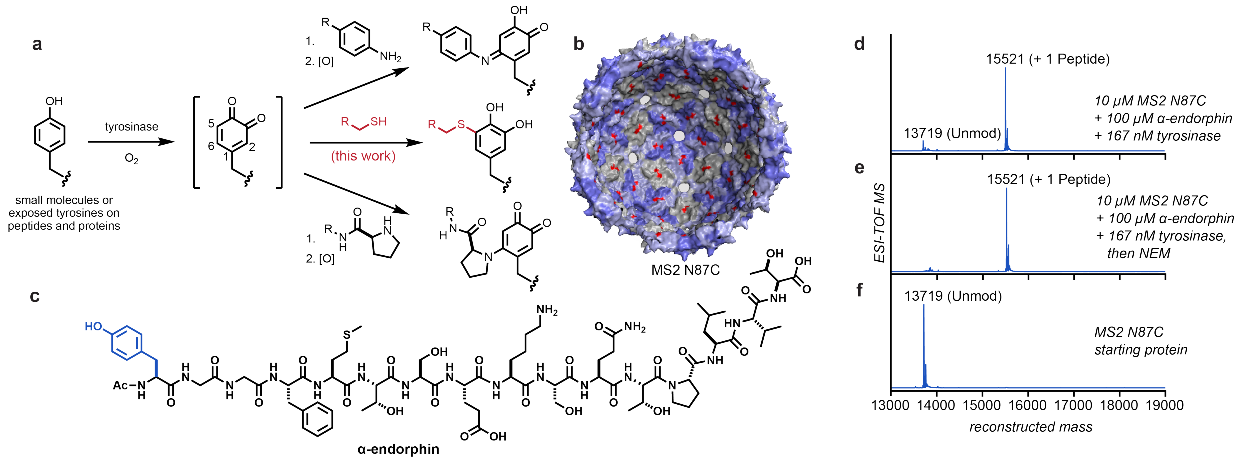


Fig. 1. Tyrosinase-mediated oxidative coupling reactions. (a) Phenols are oxidized by tyrosinase to yield *o*-quinone intermediates that couple with a variety of nucleophiles on biomolecules. In this work, the addition of cysteine thiolates is explored (shown in red). (b) An MS2 viral capsid is shown, with 180 interior cysteine residues (N87C) indicated in red. PDB ID: 2MS2. Pores in the capsid shell allow the entry of peptides and small molecules. (c) The structure of α-endorphin is shown, with the targeted tyrosine residue in blue. (d-f) Coupling reactions were screened using ESI-TOF MS, showing full modification of the MS2 N87C capsid without off-target oxidation. Reaction conditions: pH 6.5 phosphate buffer, RT, 30 min. Expected mass values: MS2 N87C $[M+H]^+ = 13719$; MS2 N87C–*N*-ethylmaleimide (NEM) adduct $[M+H]^+ = 13844$; MS2 N87C–endorphin *o*-hydroquinone product $[M+H]^+ = 15521$.

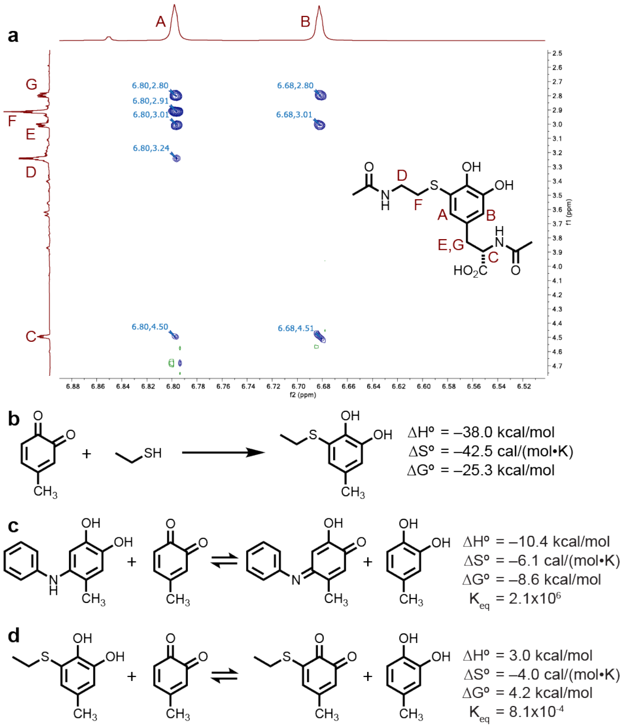


Fig. 2. Structural analysis of thiol-*o*-quinone coupling products. (a) ROESY ^{1}H NMR data (900 MHz) show a clear correlation between proton A and the corresponding signals from both the alkyl chain in *N*-acetylcysteamine (F) and the *N*-acetyltyrosine alkyl chain (E,G). (b) The energetics of the coupling reaction were estimated using DFT calculations. Geometry optimization and vibrational spectral calculations were conducted using B3LYP/6-31G**, and final electronic energies were calculated using ωB97M-V/6-311G^{++}3df-3pd. A CPCM solvation model was used in these calculations (see Supporting Information for details and molecular coordinates). These data were combined to obtain the reported thermodynamic values. For comparison purposes, the reaction energies and equilibrium constants were estimated using this approach for the oxidation of the initial (c) aniline and (d) thiol coupling products with a methyl-substituted *o*-quinone.

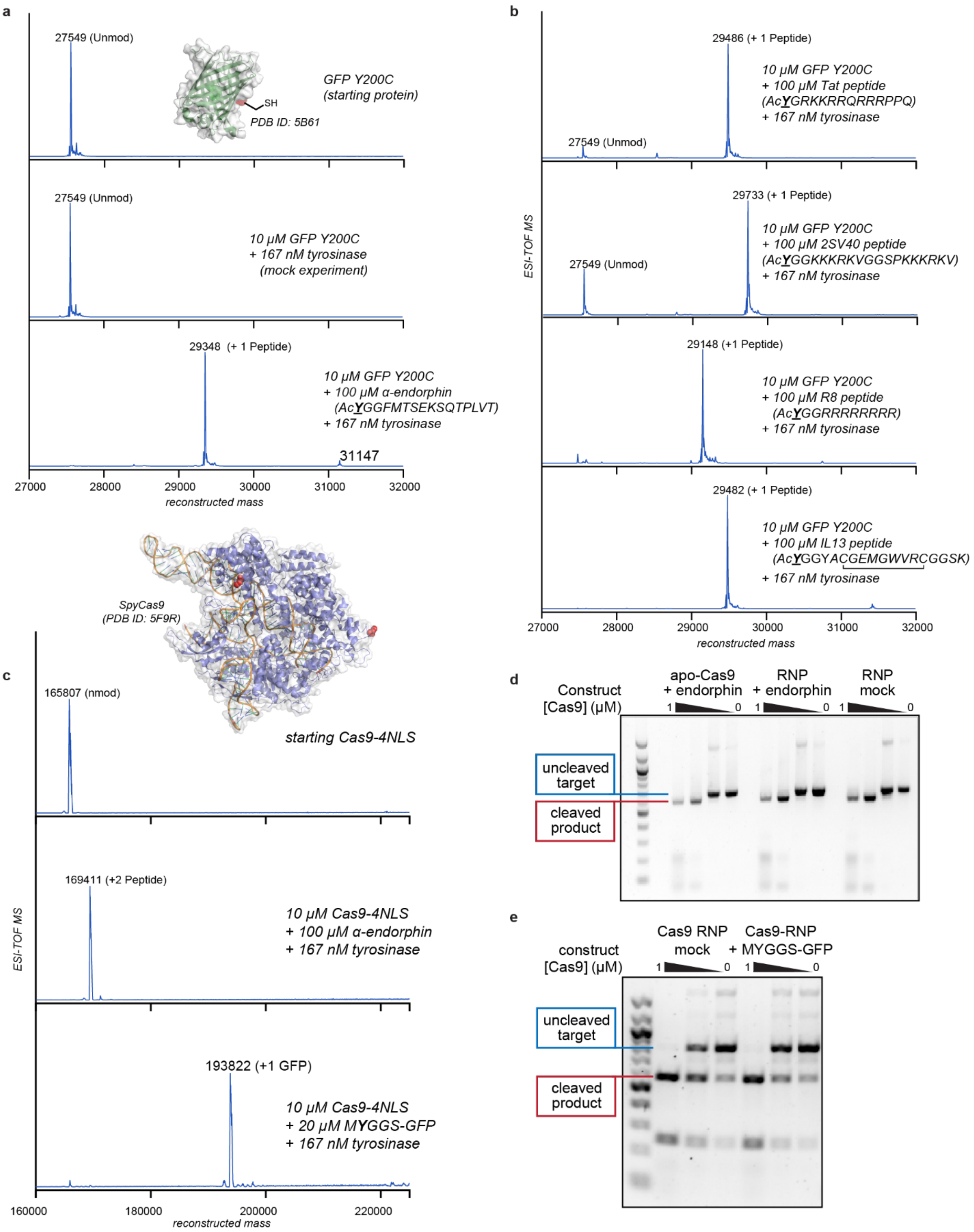
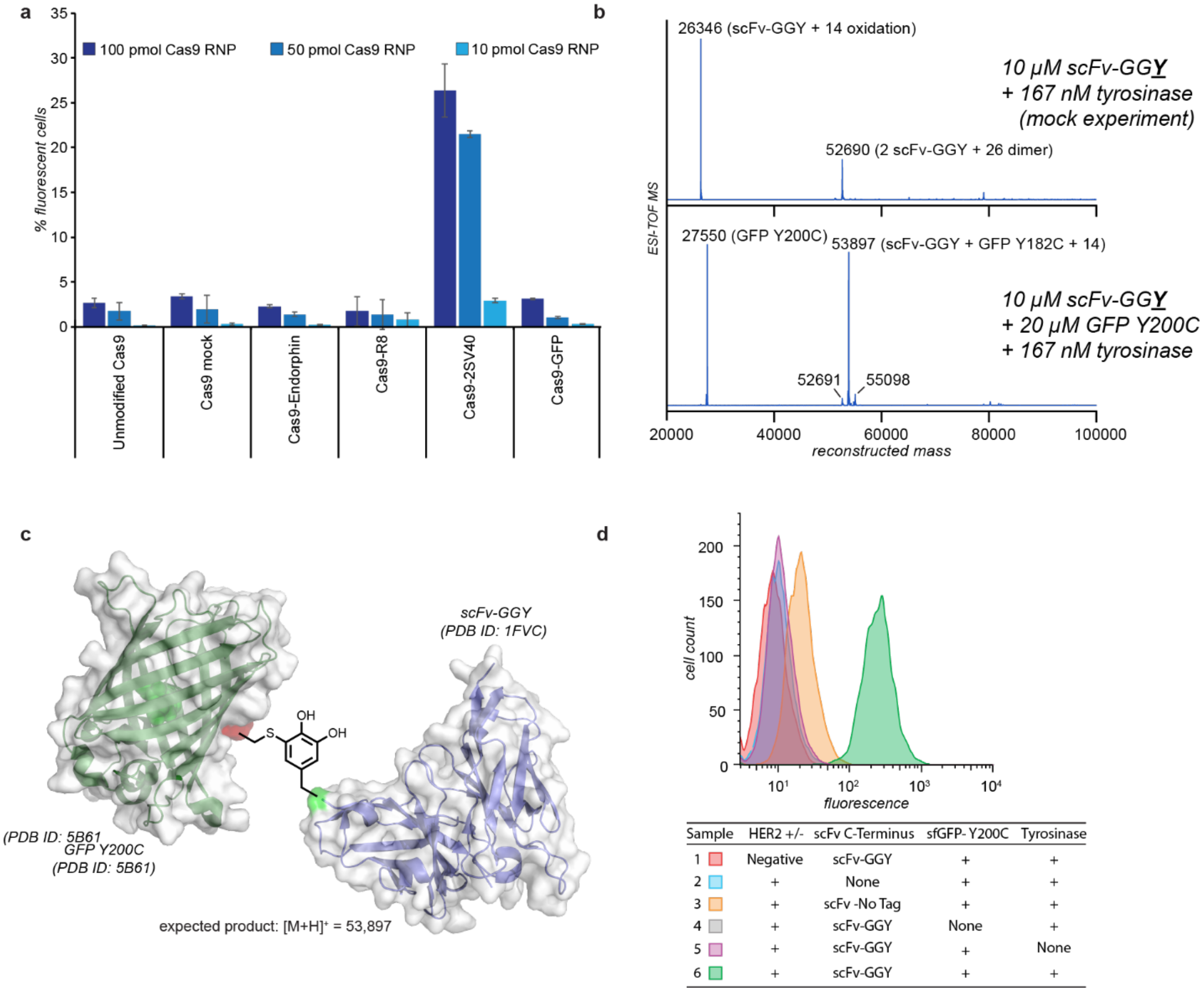

a
27549 (Unmod)
SH
PDB ID: 5B61
GFP Y200C
(starting protein)
27549 (Unmod)
10 µM GFP Y200C
+ 167 nM tyrosinase
(mock experiment)
29348 (+ 1 Peptide)
10 µM GFP Y200C
+ 100 µM α-endorphin
(AcYGGFMTSEKSQTPLVT)
+ 167 nM tyrosinase
31147
27000
28000
29000
30000
31000
32000
reconstructed mass
b
ESI-TOF MS
29486 (+ 1 Peptide)
27549 (Unmod)
10 µM GFP Y200C
+ 100 µM Tat peptide
(AcYGRKKRRQRRRPPQ)
+ 167 nM tyrosinase
29733 (+ 1 Peptide)
27549 (Unmod)
10 µM GFP Y200C
+ 100 µM 2SV40 peptide
(AcYGGKKKRKVGGSPKKKRKV)
+ 167 nM tyrosinase
29148 (+1 Peptide)
10 µM GFP Y200C
+ 100 µM R8 peptide
(AcYGGRRRRRRRR)
+ 167 nM tyrosinase
29482 (+ 1 Peptide)
10 µM GFP Y200C
+ 100 µM IL13 peptide
(AcYGGYACGEMGWVRCGGSK)
+ 167 nM tyrosinase
27000
28000
29000
30000
31000
32000
reconstructed mass
SpyCas9
(PDB ID: 5F9R)
c
ESI-TOF MS
165807 (nmod)
starting Cas9-4NLS
169411 (+2 Peptide)
10 µM Cas9-4NLS
+ 100 µM α-endorphin
+ 167 nM tyrosinase
193822 (+1 GFP)
10 µM Cas9-4NLS
+ 20 µM MYGGS-GFP
+ 167 nM tyrosinase
160000
180000
200000
220000
reconstructed mass
d
Construct
[Cas9] (µM)
apo-Cas9
+ endorphin
RNP
+ endorphin
RNP
mock
1
0
uncleaved
target
cleaved
product
e
construct
[Cas9] (µM)
Cas9 RNP
mock
Cas9-RNP
+ MYGGS-GFP
1
0
uncleaved
target
cleaved
product

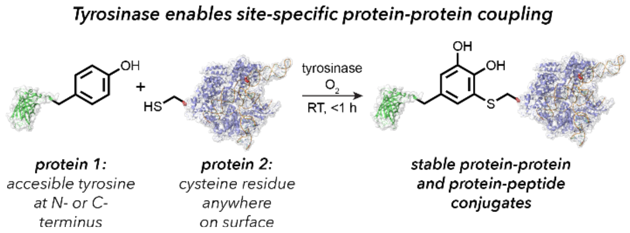
Fig. 3. Synthesis of protein-peptide and protein-protein conjugates using thiol-directed enzymatic oxidative coupling reactions. (a,b) A superfolder GFP thiol mutant (Y200C) was used as a protein component to evaluate peptide coupling reactions. All reactions were run for 30 min at RT and characterized using ESI-TOF MS. (c,d,e) A CRISPR-Cas9 variant with two C-terminal and four N-terminal nuclear localization sequences (Cas9-4NLS) was modified with peptide and protein coupling partners. Cas9 has two surface thiols (cys 80 and cys 574, red in the attached structure) in the native sequence. (c) ESI-TOF MS data indicated complete coupling at both sites with the α-endorphin peptide. The coupling reaction was successful even when the coupling partner was superfolder GFP bearing an exposed tyrosine residue near the N-terminus (MYGGS-GFP). All Cas9 couplings were run on ice for 1h. (d) The site-specific DNA cleaving ability of the Cas9-peptide conjugate was unchanged relative to an untreated control. This was true when the pre-assembled ribonucleoprotein (RNP) was modified directly, or when the RNA-free protein (apo-Cas9) was used for the modification reaction. In the latter case, the single guide RNA strand was added before the DNA cleavage experiment. (e) The Cas9-GFP conjugate retained site-specific DNA cleaving ability despite the high degree of added steric bulk.



Sample	HER2 +/-	scFv C-Terminus	sfGFP- Y200C	Tyrosinase
1 🟥	Negative	scFv-GGY	+	+
2 🟦	+	None	+	+
3 🟧	+	scFv -No Tag	+	+
4 ⬜	+	scFv-GGY	None	+
5 🟪	+	scFv-GGY	+	None
6 🟩	+	scFv-GGY	+	+

Fig. 4. Evaluation of tyrosine-cysteine bioconjugates in cell culture assays. (a) Evaluation of modified CRISPR-Cas9 ribonucleoproteins (RNP) for gene editing in neural progenitor cells. The assay measured the increase in fluorescent protein expression following successful editing at a tdTomato locus. Cas9 modified with two copies of the SV40 nuclear localization sequence (Cas9-2SV40) showed a 20-fold increase in editing efficiency compared to unmodified Cas9. Data were collected at 72 h post-treatment. Error bars show the standard deviation of triplicates. (b) The coupling reaction was successful for the coupling of GFP Y200C to a HER2-binding scFv-GGY construct, as measured by ESI-TOF. The observed product mass was consistent with the linkage depicted in (c). (d) Flow cytometry analysis of SK-BR-3 (HER2$^+$) and MDA-MB-468 (HER2$^-$) breast cancer cells treated with the Trastuzumab scFv-sfGFP construct showed HER2 specific cell binding. Gating and statistics are shown in Figure S18.

For Table of Contents Use Only



Synopsis:

Enzymatic oxidation of tyrosine residues followed by reaction with cysteine thiols allows the covalent coupling of proteins and peptides using only native amino acid side chains.