



Published in final edited form as:

*J Am Chem Soc.* 2017 August 30; 139(34): 11670–11673. doi:10.1021/jacs.7b05125.

## A Chemoselective Rapid Azo-Coupling Reaction (CRACR) for “Unclickable” Bioconjugation

Partha Sarathi Addy, Sarah B. Erickson, James S. Italia, and Abhishek Chatterjee\*

### Abstract

Chemoselective modification of complex biomolecules has become a cornerstone of chemical biology. Despite the exciting developments of the past two decades, the demand for new chemoselective reactions with unique abilities, and those compatible with existing chemistries for concurrent multisite-directed labeling, remains high. Here we show that 5-hydroxyindoles exhibit remarkably high reactivity toward aromatic diazonium ions and this reaction can be used to chemoselectively label proteins. We have previously genetically encoded the noncanonical amino acid 5-hydroxytryptophan in both *E. coli* and eukaryotes, enabling efficient site-specific incorporation of 5-hydroxyindole into virtually any protein. The 5-hydroxytryptophan residue was shown to allow rapid, chemoselective protein modification using the azocoupling reaction, and the utility of this bioconjugation strategy was further illustrated by generating a functional antibody–fluorophore conjugate. Although the resulting azo-linkage is otherwise stable, we show that it can be efficiently cleaved upon treatment with dithionite. Our work establishes a unique chemoselective “unclickable” bioconjugation strategy to site-specifically modify proteins expressed in both bacteria and eukaryotes.

The ability to chemoselectively label complex biomolecules has emerged as a powerful strategy to both study and engineer their structure and function.<sup>1</sup> To achieve this, uniquely reactive non-natural chemical functionalities are incorporated into biomolecules through synthetic or biosynthetic routes, followed by their selective functionalization using a chemoselective reaction.<sup>1</sup> A variety of such reactions have been developed to date, including the condensation between an aldehyde/ketone and an alkoxyamine/hydrazine,<sup>2</sup> the Staudinger ligation,<sup>3</sup> the Cu(I)-catalyzed<sup>4</sup> and strain promoted cycloaddition<sup>5</sup> between alkynes and azides, the inverse electron demand Diels–Alder reactions between tetrazines and strained alkenes,<sup>6</sup> various 1,3-dipolar cycloadditions,<sup>7</sup> etc. Many of these chemistries have been harnessed within a growing collection of genetically encoded noncanonical amino acids (ncAAs).<sup>1f,8</sup> Development of engineered nonsense-suppressing aminoacyl-tRNA synthetase (aaRS)/tRNA pairs that selectively charge such ncAAs has enabled their facile site-specific incorporation into proteins expressed in living cells, which can be subsequently used to precisely attach a variety of entities, from biophysical probes to therapeutic agents.

\*Corresponding Author: Department of Chemistry, Boston College, 2609 Beacon Street, 246B Merkert Chemistry Center, Chestnut Hill, Massachusetts 02467, United States. Tel: +1-617-552-1778 abhishek.chatterjee@bc.edu.  
ORCID Abhishek Chatterjee: 0000-0002-6231-5302

**Notes.** The authors declare no competing financial interest.

Supporting Information: The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b05125.

1f,8,9 However, many of the available chemistries are limited by their slow kinetics, the need for toxic catalysts, or a lack of compatibility with other chemistries to allow the simultaneous attachment of multiple distinct entities at different sites.<sup>10</sup> Consequently, there is continued interest in additional genetically encoded chemical functionalities that can be rapidly and chemoselectively labeled using catalyst-free conjugation reactions, and that are also compatible with other available chemistries to allow concurrent labeling at multiple sites.

The azo-coupling reaction between an aromatic diazonium ion and aromatic amino acid residues, such as tyrosine (Tyr, 1; Figure 1A), has been previously used to chemically modify proteins.<sup>11</sup> However, the inability to achieve site-specific labeling, owing to the abundance of surface exposed tyrosine residues commonly found on proteins, significantly limits the utility of this conjugation strategy. Even though reactive diazoniums, e.g., 4-nitrobenzenediazonium (4NDz, 3), can label tyrosine residues at physiological pH, less electrophilic counterparts are only able to react at a significantly elevated pH.<sup>11a</sup> We anticipated that the identification of aromatic groups that exhibit enhanced reactivity toward aromatic diazonium ions could form the basis of developing a new chemoselective reaction for bioconjugation.

5-Hydroxytryptophan (5HTP, 2; Figure 1A) is a stable nontoxic amino acid which is widely accessible as an over-the-counter dietary supplement.<sup>12</sup> It is also generated *in vivo* at low levels as a metabolic precursor to neurotransmitters serotonin and melatonin.<sup>13</sup> The 5-hydroxyindole ring of 5HTP is highly electron-rich, as indicated by its propensity to undergo oxidation under relatively mild conditions.<sup>14</sup> To evaluate if the electron-rich 5-hydroxyindole exhibits enhanced reactivity toward aromatic diazonium ions, we monitored the kinetics of the azo-coupling reaction between 4NDz with 5HTP and compared it to the corresponding reaction with tyrosine under ambient conditions, in an aqueous phosphate buffer at pH 7. The formation of the chromophoric azo-coupling product in each case was monitored using spectrophotometry (increased absorption at 450 nm; Figure S1 in the Supporting Information). 5HTP was found to react with 4NDz at a rate ( $k_2 = 63000 \pm 6500 \text{ M}^{-1} \text{ s}^{-1}$ ; Figures S2, S4) approximately 4500-fold faster than tyrosine ( $k_2 = 14.2 \pm 0.1 \text{ M}^{-1} \text{ s}^{-1}$ ; Figures S3, S4). The reaction between 5  $\mu\text{M}$  each of 5HTP and 4NDz in aqueous phosphate buffer (pH 7) was found to reach near-completion in less than a minute, while an identical reaction of 4NDz with tyrosine did not show noticeable progress in this time frame (Figure 1C). However, at elevated concentrations of Tyr, slow formation of the coupling product was observed (Figure 1C). Even though these observations indicate that 4NDz may be used to selectively label 5HTP in the presence of Tyr, its non-negligible reactivity toward the latter prompted us to explore less electrophilic diazonium compounds with attenuated reactivity.

We evaluated 4-carboxybenzenediazonium (4CDz, 4) and 4-methoxybenzenediazonium (4MDz, 5), which reacted with 5HTP with rate constants of  $193 \pm 37 \text{ M}^{-1} \text{ s}^{-1}$  (Figure 1D) and  $2.3 \pm 0.1 \text{ M}^{-1} \text{ s}^{-1}$  (Figure S5). In contrast, prolonged incubation of 4CDz with tyrosine at pH 7 resulted in significant levels of diazonium hydrolysis but no azo-coupling (Figures S6–S8), although some coupling was observed at pH 9, corroborating previous reports.<sup>11f</sup> Owing to its selective yet robust reactivity toward 5HTP at pH 7, we focused on the 4CDz

scaffold for further studies. The reaction between 5HTP (0.1 mM) and 4CDz (10  $\mu$ M) in 100 mM phosphate buffer (pH 7) was found to reach near-completion within minutes, while 4CDz failed to react with tyrosine and other canonical amino acids under similar conditions (Figures S9–S10). To characterize the product of the azo-coupling reaction with 5-hydroxyindole, three diazonium compounds were reacted with methyl 2-(5-hydroxy-1H-indol-3-yl)acetate (6; Figure 2A), a 5HTP analog, and the resulting products were isolated and characterized by NMR ( $^1\text{H}$  and  $^{13}\text{C}$ ) and mass spectrometry. The product with 4MDz was also characterized by X-ray crystallography (Figure 2B) to further confirm that the azo coupling occurs through the 4-position of the 5-hydroxyindole ring.

We have recently engineered the *E. coli* tryptophanyl-tRNA synthetase (EcTrpRS)/tRNA pair to efficiently charge 5HTP in response to the TGA nonsense codon in both eukaryotes, as well as in an engineered *E. coli* strain, where the endogenous EcTrpRS/tRNA<sup>Trp</sup> pair was functionally replaced with a eukaryotic counterpart.<sup>15</sup> Using this platform, we expressed superfolder green fluorescent protein incorporating 5HTP at a permissive surface exposed site (sfGFP-151-5HTP; Figure S11). Incubation of sfGFP-151-5HTP with 40  $\mu$ M 4CDz resulted in its rapid and complete covalent modification (Figure 3A, Figure S12) within 30 min, as observed by mass spectrometry analysis, while a nearly identical wild-type sfGFP (Tyr at 151 position) protein remained unmodified upon the same treatment (Figure 3B). We further showed that sfGFP-151-5HTP, but not wild-type sfGFP, can also be selectively modified using 4NDz (Figure S13) and 4MDz (Figure S14). As expected, the labeling with 4NDz was significantly faster, reaching near-completion within a minute (Figure S13C). These observations confirm that the genetically encoded 5HTP residue can be used to site-specifically label proteins using the chemoselective rapid azo-coupling reaction (CRACR).

We generated the fluorescent diazonium compound 7 (Figure 3C) from 6-aminofluorescein, which retains the 4CDz scaffold. Incubation with 7 resulted in selective and complete labeling of sfGFP-151-5HTP, but not wild-type sfGFP, as monitored by SDS-PAGE followed by fluorescence imaging (Figure 3D), and mass spectrometry (Figure S15). Recently, the aryl-triazabutadiene group has been established to be a facile precursor to the aryl-diazonium species, the release of which can be triggered by light or low pH.<sup>16</sup> We synthesized a conjugate (8, Figure 3C) between biotin and a triazabutadiene, which upon brief irradiation releases biotin-4CDz (Figure S16A). We showed selective biotinylation of sfGFP-151-5HTP, but not wild-type sfGFP, upon treatment with biotin-4CDz by a western-blot analysis using a streptavidin-HRP probe (Figure 3E), and MS analysis (Figure S16).

The ability to site-specifically label antibodies has been a valuable technology to generate therapeutically important antibody–drug conjugates, as well as covalent conjugates with biophysical probes that are important as diagnostics as well as research tools.<sup>9,17</sup> To evaluate if CRACR can be used to generate such antibody–conjugates in a site-specific manner, we expressed the previously described Fab fragment<sup>18</sup> of the anti-Her2 antibody Herceptin<sup>19</sup> in our engineered *E. coli* strain, site-specifically incorporating 5HTP at position 169. The wild-type Fab protein was also expressed as a control (lysine at position 169). SDS-PAGE and ESI-MS analysis was used to confirm successful expression of both proteins (Figure 4A–B). 80  $\mu$ M of the fluorescent diazonium compound 7 was incubated with both the wild-type and the 5HTP mutant anti-Her2 Fab for 15 min on ice. Subsequent ESI-MS analysis and SDS-

PAGE, followed by fluorescence imaging, confirmed complete labeling of the 5HTP mutant and no labeling of the wild-type antibody (Figure 4A–B). The resulting fluorescently labeled antibody was dialyzed to remove the unreacted fluorophore and was shown by fluorescence-activated cell sorting (FACS) analysis to successfully bind to the SK-BR-3 breast cancer cells, which overexpresses the Her2 receptor (Figure 4C).

Even though the azo-linkage is stable under physiologically relevant conditions, it is known to undergo facile cleavage upon treatment with dithionite,<sup>11a,20</sup> a strong but protein-compatible reducing agent. Given that the linkages generated by most available bioconjugation reactions are challenging to selectively undo under mild conditions, we aimed to explore if the conjugate made through CRACR can be “unclicked” upon a dithionite treatment. Indeed, treatment of the purified model azo molecule 9 (Figure 5A) with 1.5 mM dithionite led to its rapid reductive cleavage, as observed by HPLC-coupled MS analysis (Figure 5A, Figure S17). To demonstrate that the azo-linkage generated on a protein through a 5HTP residue can also be cleaved this way, we treated the aforementioned conjugate (Figure 3E) between sfGFP-151-5HTP and biotin-4CDz 8 with dithionite and confirmed the removal of the biotin by a western-blot analysis (Figure 5B). However, it should be noted that the 4-amino-5-hydroxytryptophan residue generated upon this reduction is readily oxidized under ambient conditions, which precluded the MS-analysis of this cleavage reaction on the protein. The remarkable speed and selectivity of the 5HTP-directed CRACR, and the ability to subsequently cleave the resulting azo-linkage, underscores unique advantages of this conjugation strategy.

In summary, we have established a general strategy to label recombinant proteins using 5HTP-directed CRACR. The ability to (1) site-specifically incorporate 5HTP, a stable and cheap nCAA, into any protein of interest expressed in both *E. coli* and eukaryotic cells, (2) chemoselectively functionalize this non-natural residue with readily generated diazonium ions through a rapid ( $>190 \text{ M}^{-1} \text{ s}^{-1}$  for the 4CDz scaffold) catalyst-free reaction, (3) systematically tune the coupling chemistry by rational substitution on the aryl-diazonium, and (4) unclick the resulting azo-linkage (if desired) using the relatively mild dithionite treatment will make this bioconjugation strategy a useful new tool in chemical biology. Another attractive feature is the small size of 5HTP, which is minimally different from the canonical amino acid tryptophan. It is also likely that this conjugation strategy will be compatible with existing bioorthogonal conjugation reactions (e.g., azide–alkyne click reaction)<sup>21</sup> – a possibility that we are actively exploring – which will enable site-specific labeling of proteins with multiple different entities.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

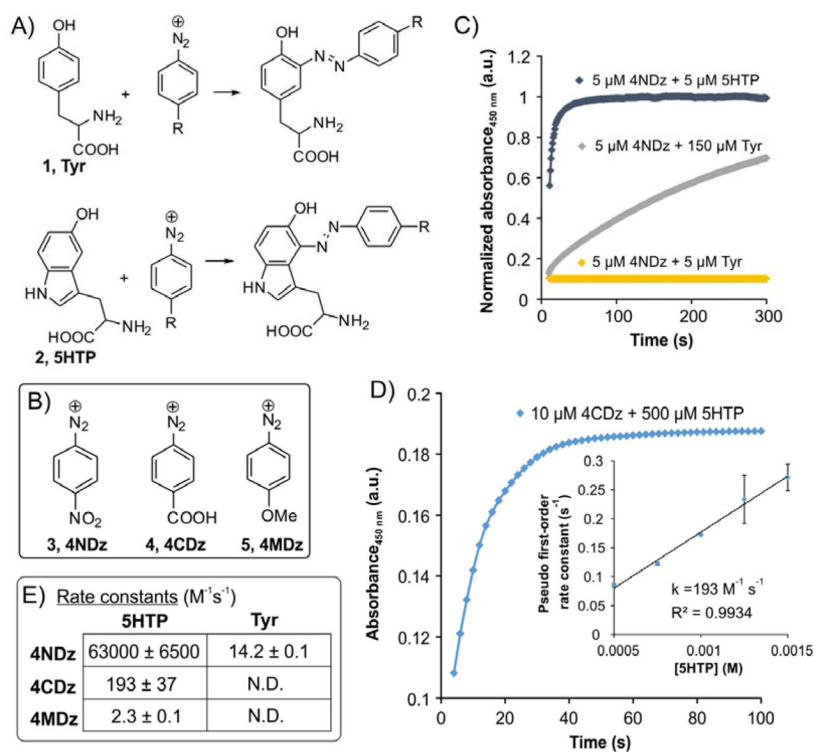
## Acknowledgments

We thank Prof. John Jewett (University of Arizona) for helpful discussions on triazabutadienes, and NIH (R01 GM124319 to A.C.) for financial support.

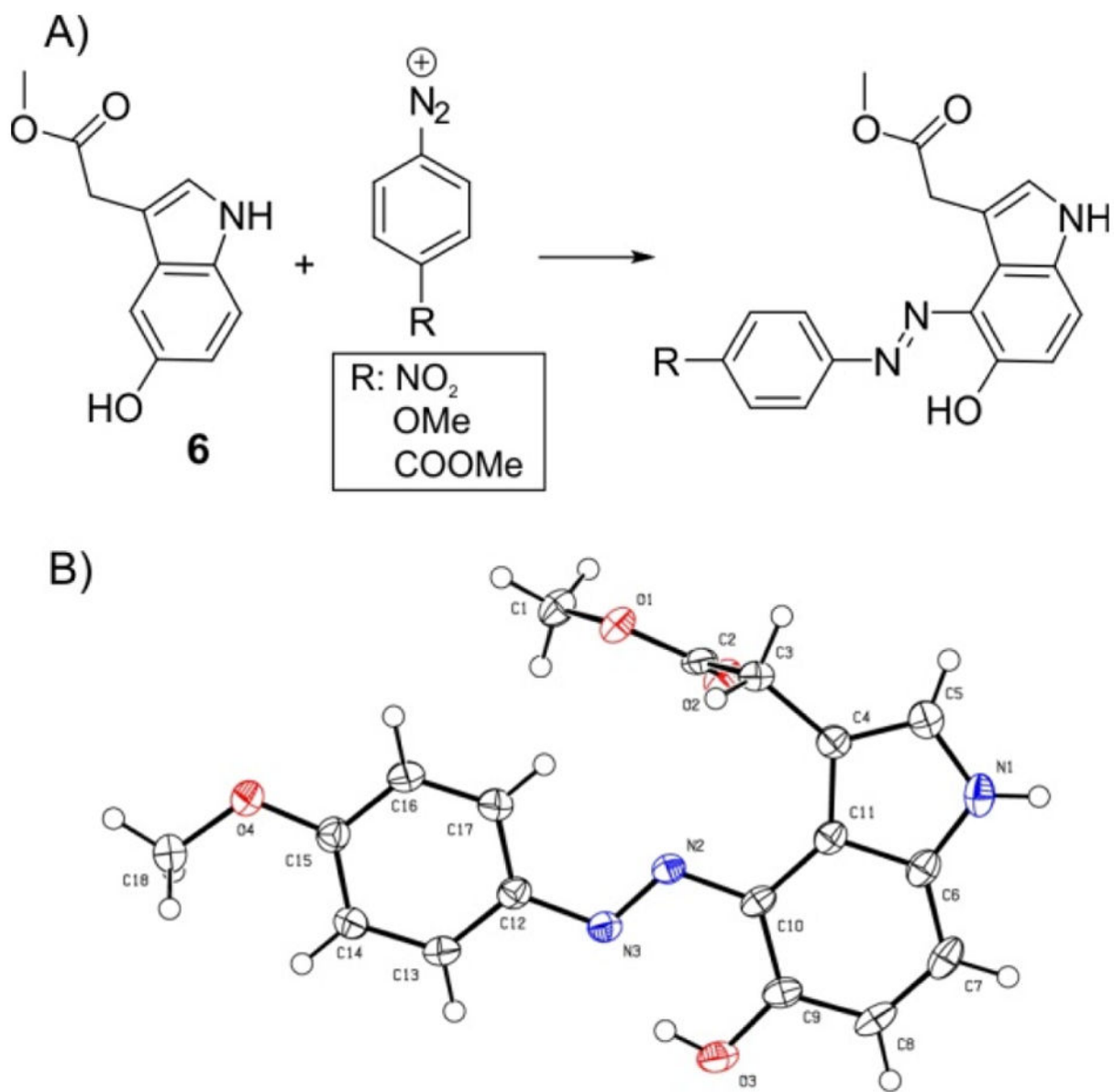
## References

1. (a) McKay CS, Finn MG. *Chem Biol.* 2014; 21:1075–101. [PubMed: 25237856] (b) Ramil CP, Lin Q. *Chem Commun.* 2013; 49:11007–22. (c) Shih HW, Kamber DN, Prescher JA. *Curr Opin Chem Biol.* 2014; 21:103–11. [PubMed: 25086220] (d) Sletten EM, Bertozzi CR. *Angew Chem, Int Ed.* 2009; 48:6974–98. (e) Stephanopoulos N, Francis MB. *Nat Chem Biol.* 2011; 7:876–84. [PubMed: 22086289] (f) Lang K, Chin JW. *ACS Chem Biol.* 2014; 9:16–20. [PubMed: 24432752]
2. Dirksen A, Dawson PE. *Bioconjugate Chem.* 2008; 19:2543–8.
3. Saxon E, Bertozzi CR. *Science.* 2000; 287:2007–10. [PubMed: 10720325]
4. (a) Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. *Angew Chem, Int Ed.* 2002; 41:2596–9. (b) Meldal M, Tornøe CW. *Chem Rev.* 2008; 108:2952–3015. [PubMed: 18698735]
5. (a) Agard NJ, Prescher JA, Bertozzi CR. *J Am Chem Soc.* 2004; 126:15046–15047. [PubMed: 15547999] (b) Jewett JC, Sletten EM, Bertozzi CR. *J Am Chem Soc.* 2010; 132:3688–3690. [PubMed: 20187640] (c) Jewett JC, Bertozzi CR. *Chem Soc Rev.* 2010; 39:1272–1279. [PubMed: 20349533]
6. (a) Blackman ML, Royzen M, Fox JM. *J Am Chem Soc.* 2008; 130:13518–13519. [PubMed: 18798613] (b) Devaraj NK, Weissleder R. *Acc Chem Res.* 2011; 44:816–827. [PubMed: 21627112] (c) Kamber DN, Liang Y, Blizzard RJ, Liu F, Mehl RA, Houk K, Prescher JA. *J Am Chem Soc.* 2015; 137:8388–8391. [PubMed: 26084312] (d) Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, Fox JM, Chin JW. *J Am Chem Soc.* 2012; 134:10317–20. [PubMed: 22694658] (e) Seitchik JL, Peeler JC, Taylor MT, Blackman ML, Rhoads TW, Cooley RB, Refakis C, Fox JM, Mehl RA. *J Am Chem Soc.* 2012; 134:2898–901. [PubMed: 22283158] (f) Taylor MT, Blackman ML, Dmitrenko O, Fox JM. *J Am Chem Soc.* 2011; 133:9646–9. [PubMed: 21599005]
7. (a) Song W, Wang Y, Qu J, Lin Q. *J Am Chem Soc.* 2008; 130:9654–5. [PubMed: 18593155] (b) Song W, Wang Y, Qu J, Madden MM, Lin Q. *Angew Chem, Int Ed.* 2008; 47:2832–5. (c) Kaya E, Vrabel M, Deiml C, Prill S, Fluxa VS, Carell T. *Angew Chem, Int Ed.* 2012; 51:4466–9.
8. (a) Dumas A, Lercher L, Spicer CD, Davis BG. *Chem Sci.* 2015; 6:50–69. [PubMed: 28553457] (b) Lang K, Chin JW. *Chem Rev.* 2014; 114:4764–4806. [PubMed: 24655057] (c) Liu CC, Schultz PG. *Annu Rev Biochem.* 2010; 79:413–44. [PubMed: 20307192] (d) Kim CH, Axup JY, Schultz PG. *Curr Opin Chem Biol.* 2013; 17(3):412–9. [PubMed: 23664497]
9. Sun SB, Schultz PG, Kim CH. *ChemBioChem.* 2014; 15:1721–9. [PubMed: 25044800]
10. Patterson DM, Prescher JA. *Curr Opin Chem Biol.* 2015; 28:141–9. [PubMed: 26276062]
11. (a) Hooker JM, Kovacs EW, Francis MB. *J Am Chem Soc.* 2004; 126:3718–9. [PubMed: 15038717] (b) Schlick TL, Ding Z, Kovacs EW, Francis MB. *J Am Chem Soc.* 2005; 127:3718–23. [PubMed: 15771505] (c) Bruckman MA, Kaur G, Lee LA, Xie F, Sepulveda J, Breitenkamp R, Zhang X, Joralemon M, Russell TP, Emrick T, Wang Q. *ChemBioChem.* 2008; 9:519–23. [PubMed: 18213566] (d) Gavriluk J, Ban H, Nagano M, Hakamata W, Barbas CF. *Bioconjugate Chem.* 2012; 23:2321–8. (e) Higgins H, Fraser D. *Aust J Chem.* 1952; 5:736–753. (f) Tabachnick M, Sobotka H. *J Biol Chem.* 1959; 234(7):1726–1730. [PubMed: 13672953] (g) Chen S, Tsao ML. *Bioconjugate Chem.* 2013; 24(10):1645–9.
12. Birdsall TC. *Alternative medicine review.* 1998; 3:271–80. [PubMed: 9727088]
13. Mockus SM, Vrana KE. *J Mol Neurosci.* 1998; 10:163–179. [PubMed: 9770640]
14. Wrona MZ, Dryhurst G. *Chem Res Toxicol.* 1998; 11:639–50. [PubMed: 9625732]
15. (a) Italia JS, Addy PS, Wrobel CJ, Crawford LA, Lajoie MJ, Zheng Y, Chatterjee A. *Nat Chem Biol.* 2017; 13:446–450. [PubMed: 28192410] (b) Italia JS, Zheng Y, Kelemen RE, Erickson SB, Addy PS, Chatterjee A. *Biochem Soc Trans.* 2017; 45:555–562. [PubMed: 28408495]
16. (a) He J, Kimani FW, Jewett JC. *J Am Chem Soc.* 2015; 137:9764–7. [PubMed: 26214020] (b) Jensen SM, Kimani FW, Jewett JC. *ChemBioChem.* 2016; 17:2216–2219. [PubMed: 27647786] (c) Kimani FW, Jewett JC. *Angew Chem, Int Ed.* 2015; 54:4051–4.
17. (a) Sievers EL, Senter PD. *Annu Rev Med.* 2013; 64:15–29. [PubMed: 23043493] (b) Beck A, Goetsch L, Dumontet C, Corva-a N. *Nat Rev Drug Discovery.* 2017; 16:315–337. [PubMed: 28303026] (c) Agarwal P, Bertozzi CR. *Bioconjugate Chem.* 2015; 26:176–192.
18. (a) Hutchins BM, Kazane SA, Staffin K, Forsyth JS, Felding-Habermann B, Schultz PG, Smider VV. *J Mol Biol.* 2011; 406:595–603. [PubMed: 21237172] (b) Hutchins BM, Kazane SA, Staffin

- K, Forsyth JS, Felding-Habermann B, Smider VV, Schultz PG. *Chem Biol.* 2011; 18:299–303. [PubMed: 21439474]
19. Hudis CA. *N Engl J Med.* 2007; 357:39–51. [PubMed: 17611206]
20. (a) Leriche G, Budin G, Brino L, Wagner A. *Eur J Org Chem.* 2010; 2010:4360–4364. (b) Verhelst SH, Fonovic M, Bogyo M. *Angew Chem, Int Ed.* 2007; 46:1284–6. (c) Jaffe CL, Lis H, Sharon N. *Biochemistry.* 1980; 19(19):4423–4429. [PubMed: 7190840]
21. Cornali BM, Kimani FW, Jewett JC. *Org Lett.* 2016; 18(19):4948–4950. [PubMed: 27619479]

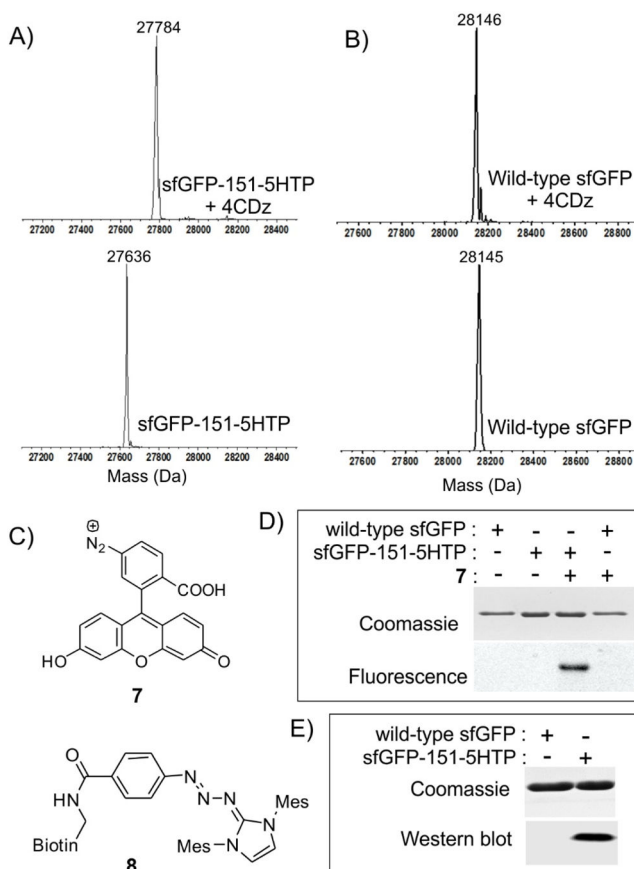


**Figure 1.** 5HTP **2** exhibit significantly higher reactivity toward aryl diazonium ions relative to Tyr **1**. (A) Azo-coupling reaction of aryl diazoniums with Tyr and 5HTP. (B) Structures of aryl diazonium ions used. (C) Observed rate of azo-coupling reaction of 4NDz **3** with 5HTP and Tyr at indicated concentrations in 100 mM phosphate buffer (pH 7, room temperature). (D) Measurement of the rate of the azo-coupling reaction between 4CDz and 5HTP under pseudo-first-order conditions. (E) Second-order rate constants of the azo-coupling reactions between the indicated partners ( $M^{-1} s^{-1}$ ). Each rate represents an average of three independent experiments, and error represents standard deviation.

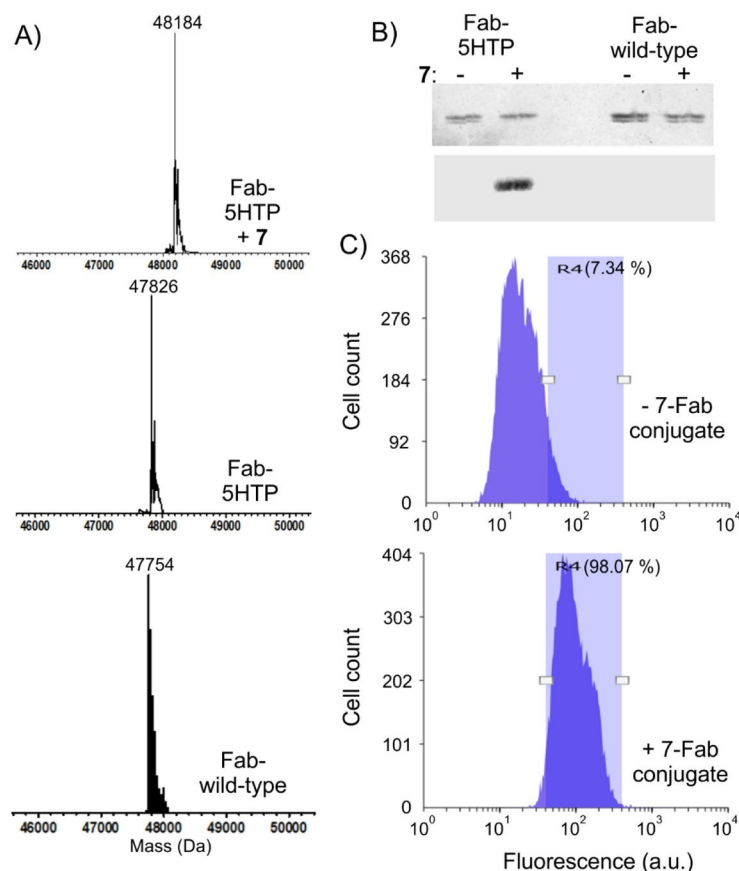


**Figure 2.** Characterizing the product of the azo-coupling reaction with 5HTP. (A) Products of the azo-coupling reaction between **6**, a 5HTP analog, and three aryl-diazoniums were purified and characterized. (B) Crystal structure of the azo-coupling product of **6** with 4MDz.





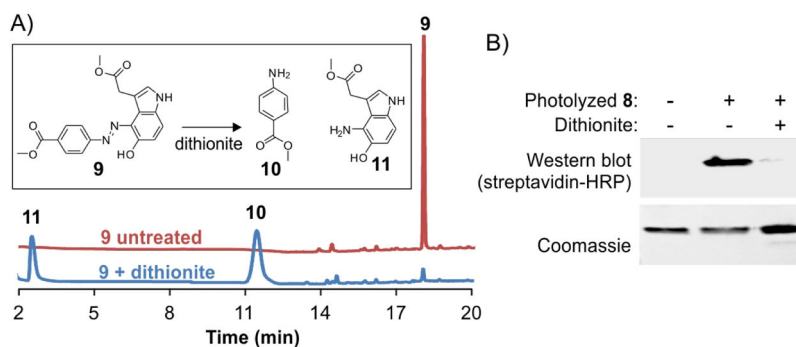
**Figure 3.** Chemoselective protein labeling using 5HTP-directed azo-coupling reaction. (A) Incubating 10  $\mu$ M sfGFP-151-5HTP with 40  $\mu$ M 4CDz at room temperature for 30 min results in complete protein labeling (expected mass 27784 Da). (B) The wild-type protein does not undergo modification upon the same treatment. (C) Structures of the fluorescent diazonium compound **7**, and the biotin-triazabutadiene conjugate **8**, synthesized for protein labeling through CRACR. (D) Selective fluorescence-labeling of sfGFP-151-5HTP by **7**, shown by fluorescence imaging following SDS-PAGE. (E) Selective biotinylation of sfGFP-151-5HTP using photolyzed **8**, revealed by Western blot (using a streptavidin-HRP probe) following SDS-PAGE.



**Figure 4.**

A functional Herceptin-Fab-fluorophore conjugate created using 5HTP-directed CRACR.

(A) ESI-MS analysis of the wild-type Fab (expected mass 47752 Da), Fab-169-5HTP (expected mass 47826 Da), and the conjugate between Fab-169-5HTP and **7** (expected mass 48186 Da). (B) SDS-PAGE followed by Coomassie staining (top) and fluorescence imaging (bottom) reveals selective fluorescence-labeling of Fab-169-5HTP, but not wild-type Fab by **7**. (C) The conjugate between **7** and Fab-169-5HTP conjugate can bind the Her2-overexpressing SK-BR-3 cells, as shown by FACS analysis. Percentage of cells fluorescing within the R4 gate is shown in each case.



**Figure 5.** Azo-linkage created through 5HTP-directed CRACR can be cleaved using dithionite. (A) Dithionite-mediated cleavage of the model azo-compound **9**. (B) HPLC-MS analysis of this reaction shows the production of the expected products (also see Figure S17). (C) Dithionite-mediated cleavage of the conjugate between photolyzed **8** and sfGFP-151-5HTP monitored by SDS-PAGE followed by Coomassie staining and Western blot (Streptavidin-HRP). Treating sfGFP-151-5HTP (left lane) with photolyzed **8** leads to its biotinylation (middle lane). Subsequent treatment of this purified conjugate with dithionite leads to the loss of the biotin group (right lane).