

# Genetic Incorporation of a 2-Naphthol Group into Proteins for Site-Specific Azo Coupling

Shuo Chen and Meng-Lin Tsao\*

Chemistry and Chemical Biology Group, School of Natural Sciences, University of California, Merced, 5200 North Lake Road, Merced, California 95343, United States

**Supporting Information** 



**ABSTRACT:** The 2-naphthol analogue of tyrosine, 2-amino-3-(6-hydroxy-2-naphthyl)propanoic acid (NpOH), has been genetically introduced into proteins in *Escherichia coli*. This is achieved through the directed evolution of orthogonal aminoacyl-tRNA synthetase/tRNA pairs that selectively charge the target amino acid in response to the amber stop codon, UAG. Moreover, chemoselective azo coupling reactions have been revealed between the 2-naphthol group and diazotized aniline derivatives that are substituted with an electron donating moiety. The coupling reactions required a very mild condition (pH 7) with great reaction rate (less than 2 h at 0  $^{\circ}$ C), high efficiency, and excellent selectivity.

## INTRODUCTION

Genetic incorporation of noncanonical amino acids (NCAAs) into proteins is a powerful method of introducing novel chemistries into proteins at chosen sites,<sup>1,2</sup> for which it enhances our ability to manipulate protein structures and functions as well as to study post-translational modifications. Several chemical transformations have been utilized at the sites of NCAAs for selective bioconjugations. For instance, the addition of an azido functionality into proteins provides a perfect position for modifications through the Staudinger ligation reaction, Cu(I) catalyzed azide-alkyne cycloaddition, or the strain-promoted cycloaddition to conjugate with triarylphosphines, terminal alkynes, or cyclooctynes, respectively.<sup>3-11</sup> Other examples include the hydrazone or oxime formation from a ketone or an aldehyde group,<sup>12-14</sup> the cyanobenzothiazole condensation initiated with an 1,2-aminothiol,<sup>15</sup> the photoinduced 1,3-dipolar cycloaddition between tetrazoles and alkenes,<sup>16</sup> the palladium catalyzed cross-coupling with an iodophenyl or a boronophenyl group,<sup>17–20</sup> and the cycloaddition of a norbornene moiety with tetrazines.<sup>21</sup> These chemical transformations, however, were sometimes hampered by limitations such as speed, toxicity, and synthetic accessibility.<sup>22-24</sup> Thus, this study aims to utilize the site selective azo coupling as an alternative biocompatible method for protein modifications.

Azo coupling, often used for producing azo dyes, has been historically applied to the modification of proteins for immune tolerance studies.<sup>25,26</sup> The coupling reaction reportedly took place very selectively at the phenol moieties of solvent-exposed

tyrosine residues.<sup>27,28</sup> Despite the high reaction yield (>90%) and fast reaction rate (2 h at 4 °C), the coupling reaction of tyrosine was restricted to a basic environment (pH 9) and required the presence of a strong electron withdrawing group at the para-position of the diazotized aniline.<sup>27–29</sup> Moreover, there is no selectivity among all of the solvent-exposed tyrosine residues. Thus, in order to utilize azo coupling for protein modifications in a generally site specific manner, we envisioned that diazotized aniline derivatives bearing an electron donating group would selectively react with a genetically introduced 2naphthol group under neutral pH (Scheme 1). This hypothesis is based on the coupling reaction's electrophilic aromatic substitution mechanism, for which electron-rich diazotized anilines would have a reduced reactivity toward tyrosine due to decreased electrophilicity; meanwhile, 2-naphthol would have a faster reaction rate than the phenol site chain of tyrosine since the former serves as a better nucleophile. Herein, we are pleased to report our findings for the selective azo coupling between the electron-rich diazonium compounds and a 2naphthol group that is genetically introduced into proteins.

# RESULTS AND DISCUSSION

The first step of our study was to add the 2-naphthol functionality into proteins. 2-Amino-3-(6-hydroxy-2-naphthyl)-propanoic acid (NpOH in Scheme 2), the 2-naphthol analogue

Received: April 4, 2013 Revised: August 22, 2013 Scheme 1. Azo Coupling at the Site of Phenol or 2-Naphthol



 $H_3 \bar{N}$ 

CO2

OH



Communication

To test the ability of the evolved synthetases for selective incorporation of NpOH into proteins, an amber stop codon was substituted at a permissive site (Lys7) in the gene for a mutant Z domain of protein A with a C-terminal 6 × His-tag.<sup>34</sup> Cells transformed with tRNA<sub>CUA</sub>, NpOH-RS1, and this Z domain gene were grown in the presence of 1 mM NpOH in LB medium, and the mutant protein was purified by Ni<sup>2+</sup> affinity column and subsequently analyzed by SDS-PAGE and ESI-MS (Figure 1). The yield of the Z-domain mutant was



**Figure 1.** Incorporation of NpOH into a target protein. (a) SDS-PAGE analysis of Lys7 $\rightarrow$ TAG amber mutant of Z-domain protein expressed under different conditions. Lane 1: molecular mass marker; lane 2: expression with WT MjTyrRS; lane 3: expression with NpOH-RS1 in the presence of NpOH; lane 4: expression with NpOH-RS1 in the absence of NpOH. The SDS-PAGE gel was stained with SimplyBlue SafeStain reagent. (b) ESI-MS deconvolution spectrum of NpOH incorporated Z-domain protein: peak a is corresponding to the Z domain protein without the first Met residue, peak b is the acetylated product of peak a, and peak c is the full length protein.

roughly 7 mg/L culture with the presence of NpOH, but was insignificant without NpOH (Figure 1a), indicating a very high fidelity in the incorporation of the unnatural amino acid. Moreover, the ESI-MS spectrum showed peaks at m/z = 7847.0, 7889.1, and 7978.3 (Figure 1b), which match the expected molecular weight for the full length NpOH-incorporated Z-domain mutant (m/z = 7978.6) and that for this mutant protein with the loss of its first methionine (m/z = 7847.4) and its subsequent post-translational acetylation product (m/z = 7889.5).<sup>5</sup> Furthermore, under 335 nm excitation, the purified Z-domain mutant displayed strong blue fluorescence with an emission maximum at 423 nm, a new feature acquired from the successful addition of NpOH (see Figure S1 in Supporting Information). Similar results were obtained when NpOH-RS1was replaced with NpOH-RS2.

After NpOH was successfully introduced into proteins, our next step was to uncover the proper conditions for selective azo couplings. We started at the small molecule level by using pcresol and 2-naphthol as the model molecules to mimic the reactions for tyrosine and NpOH, respectively. A series of aniline derivatives were converted to their corresponding diazonium salts after treatment with aqueous *p*-toluenesulfonic acid and sodium nitrite. The diazotized aniline derivatives were then mixed with *p*-cresol or 2-naphthol in a buffered solution at pH 7 or pH 9 under an ice bath for 5 min. To avoid complications due to the assumed instability of the diazonium compounds, which were used in excess quantities, the coupling reactions were subsequently stopped by adding sodium azide to quench the remaining diazonium salts. The progress of the reaction was easily observed from the appearance of the unique bright color derived from the azo products. As depicted in

Table 1, the azo coupling reaction selectively took place only for 2-naphthol but not p-cresol at pH 7 when an electron

#### Table 1. Percent Yields for the Azo Coupling of p-Cresol and 2-Naphthol with Diazotized Aniline Derivatives<sup>a</sup>



<sup>a</sup>Based on the weights of the precipitated azo adducts produced from 12  $\mu$ mol of *p*-cresol or 2-naphthol and 15  $\mu$ mol of a diazonium salt. <sup>b</sup>No precipitation is formed but a high yield is reported here judged by the appearance of the intense bright color from the azo adduct.

donating group was substituted at the para-position of the diazonium reagents. Additionally, we also examined the azo coupling of the electron-rich diazonium salts with phenylalanine, histidine, tryptophan, or deoxyribonucleotide triphosphate mixtures (dNTPs). None of them displayed the color of the azo product at pH 7. Unsurprisingly, the coupling reaction's preference toward 2-naphthol over p-cresol reduced drastically when the diazonium salts were substituted with an electron withdrawing group or when the reaction medium is kept at a basic environment of pH 9 (see Figure S2 and S3 in Supporting Information).

Ultimately, the selective azo coupling was proven to be effective at the protein level as well. For instance, the NpOH incorporated Z-domain was incubated with diazotized 4'aminoacetanilide (Figure 2a,  $R = CH_3$ ) for 90 min at 0 °C and neutral pH. The product was analyzed by ESI-MS spectroscopy and the result is shown in Figure 2b. Compared to the MS peaks from the protein measured before the coupling reaction (peaks a, b, and c in Figure 1b), a mass increment of +161.2 Da for each of them was expected at 8008.6, 8050.6, and 8139.8 Da due to the addition of the azo moiety. Indeed, the major peaks showed in Figure 2b at 8008.3 and 8050.0 Da can be attributed unambiguously to these assignments. Examining the MS spectra further, the mass peaks for the unmodified Z-domain mutant at 7847 and 7889 Da were insignificant in Figure 2b. This suggests that the coupling reaction is very efficient and has a near quantitative conversion yield (>96%, estimated from the heights of the MS peaks). In addition to the mass spectroscopic evidence, the modified protein displayed a distinct pink color with a broad UV-visible absorption band at 520 nm (see Figure S5 in Supporting Information). This strongly supports that the azo coupling took place at the NpOH sites. We further examined the kinetics of the reaction by monitoring at this unique visible band. A pseudo first-order kinetics was obtained with a rate constant of  $(8.4 \pm 0.4) \times 10^{-3} \text{ s}^{-1}$  under the conditions of 140  $\mu$ M protein and 30 mM diazonium salt at pH 7 in a 4 °C cold room (see Figure S6 in Supporting Information). To further demonstrate that the coupling reaction is triggered by the presence of NpOH, azo coupling reactions were tested for a Z-domain control protein (Figure 1a, lane 2), in which the NpOH residue was replaced with a tyrosine so that the control protein contains tyrosine residues

0°C nH7 (b)  $R = CH_3$ (c)  $R = PEG_{5000}$ 8008 3 3 2 4 kDa 80 60 50 40 30 20 15

10

3.5

(a)

Figure 2. Chemoselective azo coupling for NpOH-incorporated proteins. (a) NpOH incorporated Z-domain protein can be selectively modified through the indicated azo coupling reactions. (b) ESI-MS deconvolution spectrum for the azo coupling of NpOH incorporated Z-domain (8 nmol) with diazotized 4'-aminoacetanilide (1.5  $\mu$ mol). Expected mass at 8008.6, 8050.6, and 8139.8 Da. (c) SDS-PAGE analysis for the PEGylation of Z-domain proteins through azo coupling. Lane 1: molecular mass marker; lane 2: NpOH Z-domain mutant before the coupling reaction; lane 3: the azo coupling product for NpOH Z-domain (6 nmol) coupled with the  $PEG_{5000}$  containing reagent (2.5  $\mu$ mol), the target protein band is emphasized with an arrow sign; lane 4: the azo coupling between a control Z-domain protein (6 nmol) and the PEG<sub>5000</sub> containing reagent (2.5  $\mu$ mol). The SDS-PAGE gel was stained with SimplyBlue SafeStain reagent.

Mass (Da)

but no NpOH. Results from the ESI-MS analysis of the control reactions yielded only the mass peaks corresponding to the unmodified protein, meaning that the azo adduct formation was negligible for the control protein (see Figure S4 in Supporting Information).

A plausible application for site-specific PEGylation via azo coupling is demonstrated with a diazonium reagent bearing a poly(ethylene glycol) moiety (Figure 2a,  $R = PEG_{5000}$ ). As depicted in Figure 2c, the PEGylated product was obtained efficiently for the conjugation of the NpOH incorporated protein (lane 3), while applying the same condition to the control protein did not yield any adduct (lane 4). This indicates again that the coupling reaction is highly selective and also that it can be utilized to modify proteins with not only small molecules but also macromolecules.

#### CONCLUSION

In summary, we have evolved MjTyrRS synthetase mutants to genetically encode 2-amino-3-(6-hydroxy-2-naphthyl)propanoic acid (NpOH) in Escherichia coli with high fidelity and efficiency. We also demonstrated that the 2-naphthol side chain in NpOH can serve as the target for site specific protein modifications through the selective azo coupling reaction under neutral pH using diazonium reagents with an electron donating substituent. Our findings provide an alternative way for biocompatible chemical transformation of proteins. The selective azo coupling reported here could be potentially applied to reactions other than protein modifications and could extend to other aromatic compounds that undergo a faster electrophilic substitution than 2-naphthol does. Moreover, the

#### **Bioconjugate Chemistry**

chemoselective azo coupling could potentially bring forth a photoswitchable functionality into proteins, where the photoswitch can be activated by nonultraviolet light for which the excitation wavelength can be tuned in accordance with the diazonium compounds. Ongoing directions include the kinetic study of the azo coupling reactions, the transformation of several established procedures for selective tyrosine modifications<sup>35–41</sup> into reactions that are specific for NpOH in proteins, and immune tolerance study for varied azo adducts.

### ASSOCIATED CONTENT

#### **S** Supporting Information

Experimental details about the directed evolution, protein expression and purification, and the azo coupling reactions with small molecules and proteins. This material is available free of charge via the Internet at http://pubs.acs.org.

#### AUTHOR INFORMATION

#### **Corresponding Author**

\*E-mail: mtsao@ucmerced.edu.

#### Notes

The authors declare no competing financial interest.

#### ACKNOWLEDGMENTS

The authors are grateful for financial support from University of California, Merced, for DNA sequencing service and mass spectrometry service from UC Berkeley, and for materials from Prof. Peter Schultz's research group.

#### ABBREVIATIONS

NpOH, 2-amino-3-(6-hydroxy-2-naphthyl)propanoic acid; Tyr, tyrosine; RS, aminoacyl tRNA synthetase; PEG, poly(ethylene glycol)

#### REFERENCES

(1) Wang, L., and Schultz, P. G. (2005) Expanding the genetic code. Angew. Chem., Int. Ed. 44, 34–66.

(2) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* 79, 413–444.

(3) Saxon, E., and Bertozzi, C. R. (2000) Cell surface engineering by a modified Staudinger reaction. *Science* 287, 2007–2010.

(4) Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A.* 99, 19–24.

(5) Tsao, M. L., Tian, F., and Schultz, P. G. (2005) Selective staudinger modification of proteins containing p-azidophenylalanine. *ChemBioChem 6*, 2147–2149.

(6) Kolb, H. C., Finn, M. G., and Sharpless, K. B. (2001) Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem., Int. Ed.* 40, 2004–2021.

(7) Feng, T., Tsao, M. L., and Schultz, P. G. (2004) A phage display system with unnatural amino acids. *J. Am. Chem. Soc. 126*, 15962–15963.

(8) Link, A. J., Vink, M. K. S., and Tirrell, D. A. (2004) Presentation and detection of azide functionality in bacterial cell surface proteins. *J. Am. Chem. Soc.* 126, 10598–10602.

(9) Nguyen, D. P., Lusic, H., Neumann, H., Kapadnis, P. B., Deiters, A., and Chin, J. W. (2009) Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA synthetase/tRNA(CUA) pair and click chemistry. *J. Am. Chem. Soc.* 131, 8720–8721.

(10) Agard, N. J., Prescher, J. A., and Bertozzi, C. R. (2004) A strainpromoted [3 + 2] azide-alkyne cycloaddition for covalent modification Communication

of blomolecules in living systems. J. Am. Chem. Soc. 126, 15046-15047.

(11) Beatty, K. E., Fisk, J. D., Smart, B. P., Lu, Y. Y., Szychowski, J., Hangauer, M. J., Baskin, J. M., Bertozzi, C. R., and Tirrell, D. A. (2010) Live-cell imaging of cellular proteins by a strain-promoted azide-alkyne cycloaddition. *ChemBioChem* 11, 2092–2095.

(12) Zhang, Z. W., Smith, B. A. C., Wang, L., Brock, A., Cho, C., and Schultz, P. G. (2003) A new strategy for the site-specific modification of proteins in vivo. *Biochemistry* 42, 6735–6746.

(13) Huang, Y., Wan, W., Russell, W. K., Pai, P. J., Wang, Z. Y., Russell, D. H., and Liu, W. S. (2010) Genetic incorporation of an aliphatic keto-containing amino acid into proteins for their site-specific modifications. *Bioorg. Med. Chem. Lett.* 20, 878–880.

(14) Hudak, J. E., Yu, H. H., and Bertozzi, C. R. (2011) Protein glycoengineering enabled by the versatile synthesis of aminooxy glycans and the genetically encoded aldehyde tag. *J. Am. Chem. Soc.* 133, 16127–16135.

(15) Nguyen, D. P., Elliott, T., Holt, M., Muir, T. W., and Chin, J. W. (2011) Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *J. Am. Chem. Soc.* 133, 11418–11421.

(16) Song, W., Wang, Y., Qu, J., Madden, M. M., and Lin, Q. (2008) A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. *Angew. Chem., Int. Ed.* 47, 2832–2835.

(17) Kodama, K., Fukuzawa, S., Nakayama, H., Sakamoto, K., Kigawa, T., Yabuki, T., Matsuda, N., Shirouzu, M., Takio, K., Yokoyama, S., and Tachibana, K. (2007) Site-specific functionalization of proteins by organopalladium reactions. *ChemBioChem* 8, 232–238.

(18) Brustad, E., Bushey, M. L., Lee, J. W., Groff, D., Liu, W., and Schultz, P. G. (2008) A genetically encoded boronate-containing amino acid. *Angew. Chem., Int. Ed.* 47, 8220–8223.

(19) Spicer, C. D., and Davis, B. G. (2011) Palladium-mediated siteselective Suzuki-Miyaura protein modification at genetically encoded aryl halides. *Chem. Commun.* 47, 1698–1700.

(20) Li, N., Lim, R. K. V., Edwardraja, S., and Lin, Q. (2011) Copperfree Sonogashira cross-coupling for functionalization of alkyneencoded proteins in aqueous medium and in bacterial cells. *J. Am. Chem. Soc.* 133, 15316–15319.

(21) Lang, K., Davis, L., Torres-Kolbus, J., Chou, C. J., Deiters, A., and Chin, J. W. (2012) Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem.* 4, 298–304.

(22) Carrico, I. S. (2008) Chemoselective modification of proteins: hitting the target. *Chem. Soc. Rev.* 37, 1423–1431.

(23) Lim, R. K. V., and Lin, Q. (2010) Bioorthogonal chemistry: recent progress and future directions. *Chem. Commun.* 46, 1589–1600.
(24) Sletten, E. M., and Bertozzi, C. R. (2011) From mechanism to mouse: a tale of two bioorthogonal reactions. *Acc. Chem. Res.* 44, 666–

mouse: a tale of two bioorthogonal reactions. Acc. Chem. Res. 44, 666–676.
(25) Baker, M. C., Campbell, D. H., Epstein, S. I., and Singer, S. J.

(25) Baker, M. C., Campbell, D. H., Epstein, S. I., and Singer, S. J. (1956) Physical chemical studies of soluble antigen-antibody complexes. VII. Thermodynamics of the reaction between benzenear-sonic acid-azo-bovine serum albumin and rabbit antibodies to benzenearsonic acid. J. Am. Chem. Soc. 78, 312–316.

(26) Weigle, W. O. (1965) Induction of autoimmunity in rabbits following injection of heterologous or altered homologous thyroglobulin. J. Exp. Med. 121, 289–308.

(27) Hooker, J. M., Kovacs, E. W., and Francis, M. B. (2004) Interior surface modification of bacteriophage MS2. J. Am. Chem. Soc. 126, 3718–3719.

(28) Schlick, T. L., Ding, Z. B., Kovacs, E. W., and Francis, M. B. (2005) Dual-surface modification of the tobacco mosaic virus. *J. Am. Chem. Soc.* 127, 3718–3723.

(29) Gavrilyuk, J., Ban, H., Nagano, M., Hakamata, W., and Barbas, C. F., III (2012) Formylbenzene diazonium hexafluorophosphate reagent for tyrosine-selective modification of proteins and the introduction of a bioorthogonal aldehyde. *Bioconjugate Chem.* 23, 2321–2328.

#### **Bioconjugate Chemistry**

(30) Vela, M. A., Fronczek, F. R., Horn, G. W., and Mclaughlin, M. L. (1990) Syntheses of 1-naphthol and 2-naphthol analogs of dl-tyrosine. Potential fluorescent probes of peptide structure and dynamics in complex environments. *J. Org. Chem.* 55, 2913–2918.

(31) Turner, J. M., Graziano, J., Spraggon, G., and Schultz, P. G. (2006) Structural plasticity of an aminoacyl-tRNA synthetase active site. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6483–6488.

(32) Ryu, Y. H., and Schultz, P. G. (2006) Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat. Methods 3*, 263–265.

(33) Melancon, C. E., and Schultz, P. G. (2009) One plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases. *Bioorg. Med. Chem. Lett.* 19, 3845–3847.

(34) Wang, L., Zhang, Z. W., Brock, A., and Schultz, P. G. (2003) Addition of the keto functional group to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A. 100*, 56–61.

(35) Joshi, N. S., Whitaker, L. R., and Francis, M. B. (2004) A threecomponent Mannich-type reaction for selective tyrosine bioconjugation. J. Am. Chem. Soc. 126, 15942–15943.

(36) Romanini, D. W., and Francis, M. B. (2008) Attachment of peptide building blocks to proteins through tyrosine bioconjugation. *Bioconjugate Chem.* 19, 153–157.

(37) McFarland, J. M., Joshi, N. S., and Francis, M. B. (2008) Characterization of a three-component coupling reaction on proteins by isotopic labeling and nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* 130, 7639–7644.

(38) Tilley, S. D., and Francis, M. B. (2006) Tyrosine-selective protein alkylation using pi-allylpalladium complexes. *J. Am. Chem. Soc. 128*, 1080–1081.

(39) Meunier, S., Strable, E., and Finn, M. G. (2004) Crosslinking of and coupling to viral capsid proteins by tyrosine oxidation. *Chem. Biol.* 11, 319–326.

(40) Ban, H., Gavrilyuk, J., and Barbas, C. F. (2010) Tyrosine bioconjugation through aqueous ene-type reactions: a click-like reaction for tyrosine. *J. Am. Chem. Soc.* 132, 1523–1525.

(41) Minakawa, M., Guo, H. M., and Tanaka, F. (2008) Imines that react with phenols in water over a wide pH range. *J. Org. Chem.* 73, 8669–8672.