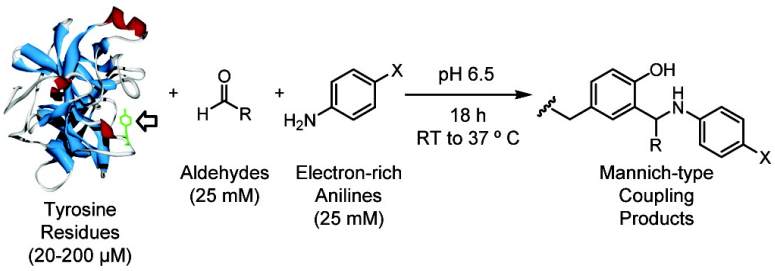
Communication

A Three-Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation

Neel S. Joshi, Leanna R. Whitaker, and Matthew B. Francis

J. Am. Chem. Soc., **2004**, 126 (49), 15942-15943• DOI: 10.1021/ja0439017 • Publication Date (Web): 19 November 2004



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 11/19/2004

A Three-Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation

Neel S. Joshi, Leanna R. Whitaker, and Matthew B. Francis*

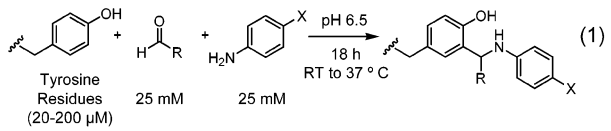
Department of Chemistry, University of California, Berkeley, California 94720-1460, and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Received October 6, 2004; E-mail: francis@cchem.berkeley.edu

The broad utility of bioconjugates for the study of biological systems and the construction of new materials[1] creates an ever-increasing need for new protein modification reactions. As a result, many strategies have appeared that can modify native protein functional groups (most commonly lysine and cysteine side chains) in a chemoselective fashion.[2] However, the high frequency of lysines on protein surfaces severely limits the site selectivity of reactions targeting this residue, typically leaving cysteine modification as the only effective way to functionalize specific locations.[3]

To complement this technique, we have focused on the development of new protein modification reactions that target aromatic amino acids,[4] as these residues are displayed with intermediate frequency on protein surfaces and can be introduced genetically without changing the overall charge state or redox sensitivity. Of these, tyrosine residues provide particularly attractive targets, as the reactivity of the phenolic side chains is largely orthogonal to that of cysteine, lysine, and carboxylate-containing residues. Furthermore, these residues are often partially "buried" in the surface of the proteins because of the amphiphilic nature of the phenolic group. This close association with the topography of protein surfaces results in varying levels of accessibility for the reactive sites, suggesting that one of several tyrosine residues could be targeted selectively through proper reagent design.

Existing methods for tyrosine modification typically occur through electrophilic aromatic substitution pathways, as exemplified by iodination[5] and diazonium coupling reactions.[4a,6,7] In terms of carbon−carbon bond-forming strategies, formaldehyde-based cross-linking techniques have been reported to proceed through the reaction of tyrosines with imines formed with lysine residues,[8] albeit with no control of site selectivity. These reactions are typically carried out using high concentrations of formaldehyde or with significant heating and thus are of limited utility for applications that require preservation of secondary and tertiary protein structure. As a more synthetically useful strategy, we have developed a new method that can modify proteins using imines formed in situ from aldehydes and electron-rich anilines (eq 1).[9] This three-component Mannich-type coupling reaction is highly selective for tyrosine side chains and proceeds under mild conditions.



As an example of this reactivity, a 20 μM solution of chymotrypsinogen A was exposed to 25 mM formaldehyde (**1a**) and 25 mM **2**[10] for 18 h at room temperature (Figure 1a). Subsequent analysis using ESI-MS indicated the formation of a new species into which both reactive components were incorporated with equal stoichiometry (Figure 1b). In an analogous experiment conducted in the absence of the aniline component, no modification of the protein was observed (Figure 1c). Optimization experiments determined that maximum reactivity can be obtained in a pH range of 5.5−6.5, with substantially lower product formation occurring at pH values above 8. Following removal of unreacted chromophore **2** using dialysis, UV analysis indicated that 60% conversion was achieved.

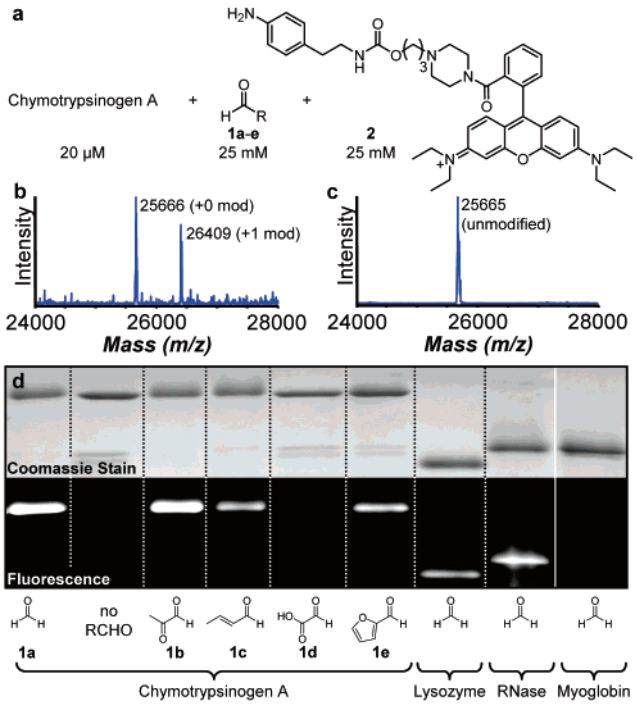


Figure 1. Modification of tyrosine residues with imine substrates. (a) A 20 μM solution of α-chymotrypsinogen A was exposed to **1a** and **2** (both at 25 mM) in aqueous buffer, pH 6.5, at room temperature for 18 h. (b) Following removal of the small molecules via gel filtration, ESI-MS analysis indicated the formation of the anticipated Mannich adduct (expected m/z = 26410). (c) In the absence of the aniline component, no reaction products were observed. (d) A gel-based assay was used to screen for reactivity using a variety of aldehyde and protein substrates. These reactions were carried out as described above, but at 37 °C. The top pane indicates the total protein content by Coomassie blue staining, and the bottom pane indicates the fluorescence of attached chromophore **2**.

To determine the reaction chemoselectivity, the products of this and analogous reactions employing 4-chloroaniline (**3e**) were subjected to trypsin digestion. Through analysis of the resulting peptide fragments by MALDI-MS, Y146 was determined to be the primary site of modification. As Friedel−Crafts reactions have rarely been demonstrated in aqueous solution,[11] we confirmed the product structure by carrying out an analogous reaction using *p*-cresol as a small molecule tyrosine mimic.[13]

The fluorescent nature of **2** allowed the use of an efficient gel-based assay to screen additional aldehyde components and protein targets (Figure 1d). In addition to formaldehyde, pyruvaldehyde

10.1021/ja0439017 CCC: $27.50

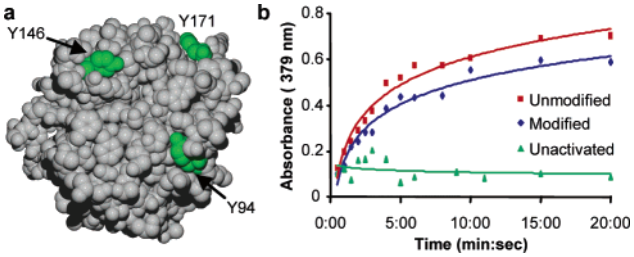


Figure 2. Modification sites and activity assay for chymotrypsinogen A. (a) Of the accessible tyrosines (green), Y146 was identified as the primary modification site. (b) Reaction profile for proteolysis of Suc-Gly-Gly-Phe-$NHC_6H_4NO_2$ for unmodified chymotrypsinogen (red) and the bioconjugate prepared from **1a** and **3a** (blue) after both enzymes were activated with trypsin.

Table 1. Modification of Tyrosine Residues on Chymotrypsinogen[a]

aniline	X-substituent	unmod(%)	+1mod[b]	+2mod[b]	+3mod[b]
2	−rhodamine dye	34	66	0	0
3a	$-CH_2CH_2OH$	20	45	35	0
3b	$-CH_2CH_2NH_2$	27	47	26	0
3c	$-CH_3$	38	46	16	0
3d	$-OCH_3$	57	43	0	0
3e	−Cl	53	36	11	0
3f	$-CO_2H$	57	33	10	0
3g	−F	91	9	0	0
3h	$-NO_2$	100	0	0	0

[a] Conditions: 200 μM α-chymotrypsinogen A, 25 mM formaldehyde, and 25 mM aniline in 100 mM phosphate buffer, pH 6.5, room temperature, 18 h. The X-substituents refer to the 4-position of the anilines, as indicated in eq 1. Product distributions were determined from ESI-MS analyses. [b] In percent.

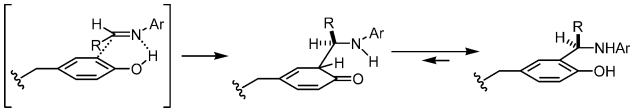
(**1b**), (*E*)-crotonaldehyde (**1c**), and 2-furaldehyde (**1e**) also yielded appreciable levels of reactivity; however, reactions carried out with glyoxylic acid (**1d**), benzaldehyde, and propionaldehyde did not afford modification products. In all cases, no cross-linking of the proteins was observed.

In addition to chymotrypsinogen A, other proteins bearing surface-accessible tyrosine residues have been modified, including lysozome and RNase A (Figure 1d). In contrast, no reactivity has been observed for horse heart myoglobin, a protein that lacks surface-accessible tyrosine residues.

In terms of reagent scope, successful reactions have been achieved with a range of electron-rich anilines (Table 1). As determined for the modification of chymotrypsinogen A, 4-alkyl-substituted anilines (**3a**−**c**) generally afforded high levels of reactivity and yielded both singly and doubly modified protein products. Chymotrypsinogen A bears three potentially reactive tyrosine residues on its surface (Figure 2a),[12] which can account for its overmodification. Efforts to quantify the reactivity levels at each site are underway. Aniline **3b** is of particular interest, as the aliphatic amine of this substrate can readily be coupled to NHS-esters, effectively converting the large number of commercially available lysine-reactive compounds into reagents for tyrosine modification. Other para-functionalized anilines (**3d**−**f**) also participated in the reaction, although lower levels of conversion were obtained. To date, reactions employing electron-deficient anilines and aliphatic amines have not been successful. This is significant, as the amino groups of lysine residues also do not participate in the reaction under these conditions.

To demonstrate that this bioconjugation method preserves the native function of biological targets, the catalytic activity of chymotrypsinogen A was evaluated after modification with **1a** and **3a**. It was found that little if any reactivity was lost even though 80% of the protein had been functionalized (Figure 2b).[13]

Scheme 1



One feature of this reaction that stands in contrast to other tyrosine modification strategies is the relatively low pH at which it takes place. We hypothesize that at pH 6.5, a cyclic transition state occurs in which the phenolic group donates a hydrogen bond to the imine nitrogen (Scheme 1). This would serve to increase the electrophilicity of the imine group while simultaneously activating the aromatic ring as a nucleophile. It is also possible that hydrophobic attractions between the imine substrate and the tyrosine ring accelerate the reaction by increasing the effective concentration. However, it should be noted that imine formation in water involves numerous equilibrating species, and thus the reaction could proceed through a more complex pathway.

As a result of these studies, an efficient new tyrosine modification method has emerged. The reaction is operationally simple to effect, and the mild conditions under which it proceeds ensure compatibility with a wide range of protein targets. The three-component nature of the reaction allows for the installation of multiple functional groups, and the selectivity of the reaction for tyrosine residues provides a complementary strategy to cysteine- and lysine-based methods. Current efforts are focused on the full exploration of the substrate scope as well as the elucidation of the mechanistic aspects. Experiments examining the ability of the reaction components to select between multiple tyrosines on a single protein surface are also underway.

Acknowledgment. We gratefully acknowledge the University of California, Berkeley, and the DOE Nanoscale Science, Engineering, and Technology Program for funding. N.S.J. acknowledges the Saegebarth family for generous fellowship support. We also thank Waters, Inc. for access to a Q-TOF Micro instrument, and Jacob M. Hooker for his analysis expertise.

Supporting Information Available: Full experimental procedures and characterization data are available for all compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) For lead references, see: (a) Niemeyer, C. M. *Angew. Chem., Int. Ed.* **2001**, *40*, 4128−4158. (b) Seeman, N. C.; Belcher, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6451−6455.
(2) For example, see: Hermanson, G. T. *Bioconjugate Techniques*, 1st ed.; Academic Press: San Diego, CA, 1996.
(3) For an example of site-selective binding to native protein functional groups, see: Griffin, B. A.; Adams, S. R.; Jones, J.; Tsien, R. Y. Fluorescent Labeling of Recombinant Proteins in Living Cells with FlAsH. In *Applications of Chimeric Genes and Hybrid Proteins*; Thorner, J., Emr, S. D., Abelson, J. N., Eds.; Methods in Enzymology 327; Academic Press: San Diego, CA; 2000, pp 565−578.
(4) (a) Hooker, J. M.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 3718. (b) Antos, J. M.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 10256.
(5) (a) Adams, H. R.; Paik, C. H.; Eckelman, W. C.; Reba, R. C. *J. Labelled Compd. Radiopharm.* **1982**, *19*, 1477−1478. (b) Eckelman, W. C.; Adams, H. R.; Paik, C. H. *Int. J. Nucl. Med. Biol.* **1984**, *11*, 163−166.
(6) (a) Higgins, H. G.; Fraser, D. *Aust. J. Sci. Res., Ser. A* **1952**, *5*, 736−753. (b) Shin, J. A. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2367.
(7) For alternative oxidative strategies, see: (a) Brown, K. C.; Yang, S. H.; Kodadek, T. *Biochemistry* **1995**, *34*, 4733−4739. (b) Meunier, S.; Strable, E.; Finn, M. G. *Chem. Biol.* **2004**, *11*, 319−326.
(8) Fraenkel-Conrat, H.; Olcott, H. S. *J. Biol. Chem.* **1948**, *174*, 827−843.
(9) For an example of a lanthanide-promoted phenol modification with imines in organic solvents, see: Huang, T. S.; Li, C. J. *Tetrahedron Lett.* **2000**, *41*, 6715.
(10) Nguyen, T.; Francis, M. B. *Org. Lett.* **2003**, *5*, 3245−3248.
(11) For a notable example, see: Breslow, R. *Acc. Chem. Res.* **2004**, *37*, 471.
(12) Pjura, P. E.; Lenhoff, A. M.; Leonard, S. A.; Gittis, A. G. *J. Mol. Biol.* **2000**, *300*, 235−239.
(13) See Supporting Information for details.

JA0439017