ORGANIC LETTERS

2006
Vol. 8, No. 1
131−134

An α-Formylglycine Building Block for Fmoc-Based Solid-Phase Peptide Synthesis

Jason Rush[‡] and Carolyn R. Bertozzi*[,†,‡,§]

Department of Chemistry, Department of Molecular and Cell Biology, and Howard Hughes Medical Institute, University of California, Berkeley, California 94720

crb@berkeley.edu

Received October 28, 2005

ABSTRACT

Nearly all known sulfatases share a common active site modification that is required for their activity: conversion of cysteine to α-formylglycine. We report the synthesis of an α-formylglycine building block suitable for Fmoc-based solid-phase peptide synthesis. The building block was incorporated into a synthetic peptide derived from the active site of a *Mycobacterium tuberculosis* sulfatase.

Sulfation of biomolecules is a common modification that is involved in both normal and pathological processes.[1] For example, sulfated oligosaccharides play important roles in inflammation,[2] tyrosine sulfation is linked to the recognition of target cells by HIV-1,[3] and an abundance of sulfated glycolipids is correlated with *Mycobacterium tuberculosis* strain virulence.[4]

Two classes of enzymes control sulfation. Sulfotransferases catalyze the formation of sulfate esters,[5] while sulfatases catalyze their hydrolysis.[6] Sulfatases utilize a unique mechanism that is facilitated by an unusual residue in the active site, α-formylglycine (FGly) (**1**) (Figure 1). In the proposed mechanism, FGly is hydrated in the active site[7] and the geminal diol (**2**) acts as a nucleophile to produce a sulfated enzyme intermediate (**3**) (Figure 1). This intermediate then collapses, expelling inorganic sulfate and regenerating FGly.

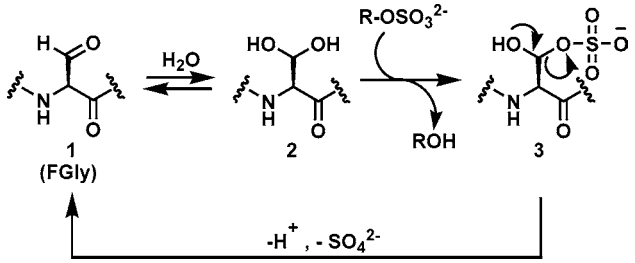


Figure 1. Catalytic cycle of sulfatases.

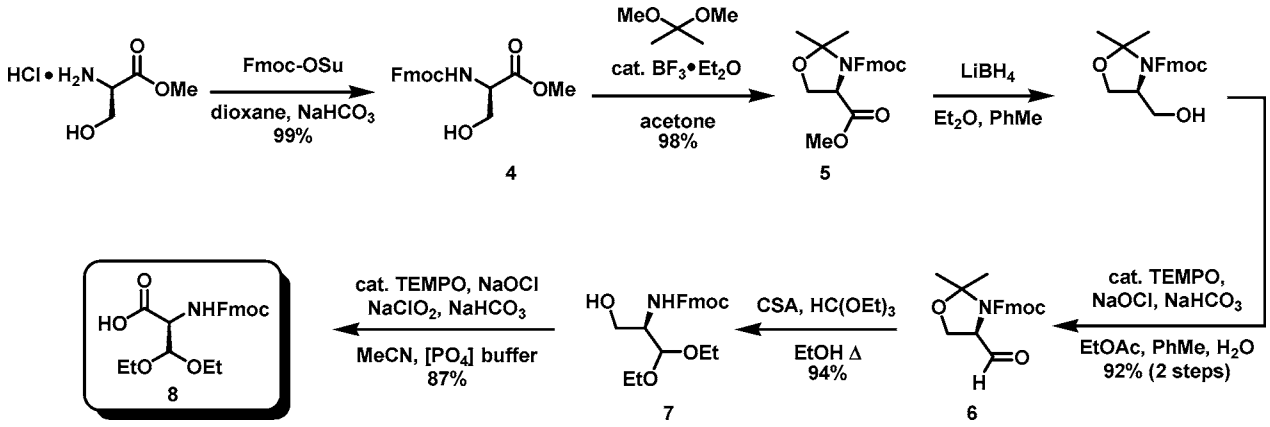
FGly is installed in the sulfatase active site via a post-translational modification prior to folding.[8] A genetically encoded cysteine residue is oxidized to form FGly by a recently characterized enzyme termed the formylglycine generating enzyme, or FGE.[9] FGE recognizes a consensus

* Address correspondence to this author.
† Department of Molecular and Cell Biology.
‡ Department of Chemistry.
§ Howard Hughes Medical Institute.

(1) Hemmerich, S.; Verdugo, D.; Rath, V. L. *Drug Discov. Today* **2004**, *9*, 967−975.
(2) Rosen, S. D. *Am. J. Pathol.* **1999**, *155*, 1013−1020.
(3) Farzan, M.; Mirzabekov, T.; Kolchinsky, P.; Wyatt, R.; Cayabyab, M.; Gerard, N. P.; Gerard, C.; Sodroski, J.; Choe, H. *Cell* **1999**, *96*, 667−676.
(4) Goren, M. B.; Brokl, O.; Schaefer, W. B. *Infect. Immun.* **1974**, *9*, 150−158.
(5) Chapman, E.; Best, M. D.; Hanson, S. R.; Wong, C. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 3526−3548.
(6) Hanson, S. R.; Best, M. D.; Wong, C. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 5736−5763.
(7) Boltes, I.; Czapinska, H.; Kahnert, A.; von Bulow, R.; Dierks, T.; Schmidt, B.; von Figura, K.; Kertesz, M. A.; Uson, I. *Structure* **2001**, *9*, 483−491.
(8) Dierks, T.; Schmidt, B.; von Figura, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11963−11968.
(9) Preusser-Kunze, A.; Mariappan, M.; Schmidt, B.; Gande, S. L.; Mutenda, K.; Wenzel, D.; von Figura, K.; Dierks, T. *J. Biol. Chem.* **2005**, *280*, 14900−14910.

10.1021/ol052623t CCC: $33.50
Published on Web 12/09/2005

Scheme 1. Synthesis of Fmoc-FGly Building-Block **8**



sequence around the target cysteine that is highly conserved across all species and is found in all sulfatases.[6] Interestingly, some bacterial sulfatases employ a serine residue in place of cysteine while retaining the overall consensus sequence. These sulfatases are modified by a distinct mechanism.[10]

Biochemical studies of sulfatases and FGEs could benefit tremendously from the availability of synthetic peptides bearing the central FGly residue. Currently, one scalemic synthesis of FGly has been reported, which utilizes the highly versatile Williams' glycine template.[11] However, an Fmoc-protected variant has not been reported and FGly analogues have not been incorporated into synthetic peptides. Here we report an efficient synthesis of a Fmoc-protected FGly building block and its incorporation into a synthetic peptide derived from a *Mycobacterium tuberculosis* sulfatase. The procedure requires no chromatography and produces Fmoc-FGly in 6 steps in >70% overall yield.

In designing a FGly analogue suitable for solid-phase peptide synthesis (SPPS), we chose to protect the aldehyde functionality as an acetal that could be removed alongside other side chain protecting groups during the acidic cleavage step of Fmoc-based SPPS. The diethyl variant was utilized as both literature precedent[12] and our own experience indicated that the dimethyl acetal was too resistant to acidic cleavage. The synthesis began with protection of commercially available D-Ser-OMe•HCl, using Fmoc-OSu to provide Fmoc-protected amino acid ester **4**[13,14] in nearly quantitative yield (Scheme 1). Exposure to catalytic BF_3•Et_2O and 2,2-dimethoxypropane afforded oxazolidine **5** in excellent yield.[15] Sequential reduction and TEMPO-mediated oxidation[16] yielded an Fmoc variant of Garner's aldehyde (**6**) in 92% yield over two steps. $LiBH_4$[17] is the optimal reductant in this sequence, as other conditions (i.e., DIBAL-H, $NaBH_4$ in protic media, or $LiAlH_4$) partially cleaved the Fmoc group. Simultaneous protection of the aldehyde and ring opening were effected in refluxing EtOH with catalytic CSA to furnish amino alcohol **7**. Finally, TEMPO-mediated oxidation[18] of the primary alcohol provided the desired SPPS building block **8** in high yield. Notably, all intermediates were either taken on crude or purified via recrystallization.

With building block **8** (Fmoc-FGly$(OEt)_2$-OH) in hand, we sought to determine its enantiomeric purity and coupling efficiency for incorporation into peptides. Thus, the FGly building block was coupled to both chiral and racemic α-methylbenzylamine with use of EDC/HOBt.[19] Following extractive workup, ^{1}H NMR analysis of the products formed from the chiral amine indicated that a single stereoisomer was formed. Analysis of the reaction containing racemic amine indicated two diasteriomers were formed in approximately equal amounts.[19] These results suggest that the coupling reaction proceeds without racemization at the α-carbon, and that the starting material was ≥95% enantiomerically pure.

With the coupling reaction characterized, we synthesized the consensus sequence derived from a *Mycobacterium tuberculosis* sulfatase, tetradecapeptide **9** (LFGlyTPSRGSLFTGRK). Given the sensitive nature of the aldehyde functionality, we expected that standard peptide cleavage cocktails containing silanes or thiols would not be compatible. Indeed, exposure of resin-bound model peptides to cleavage cocktails containing silanes (i.e., TIS or TES) resulted in reduction of FGly to serine, while addition of ethanedithiol resulted in quantitative formation of the dithioacetal. Fortunately, thioanisole and anisole were found to be satisfactory alternative scavengers. In addition to these anticipated issues, an unexpected problem arose when the full-length peptide was

(10) Szameit, C.; Miech, C.; Balleininger, M.; Schmidt, B.; von Figura, K.; Dierks, T. *J. Biol. Chem.* **1999**, *274*, 15375−15381.
(11) DeMong, D. E.; Williams, R. M. *Tetrahedon Lett.* **2002**, *43*, 2355−2357.
(12) DeMong, D. E.; Williams, R. M. *J. Am. Chem. Soc.* **2003**, *125*, 8561−8565.
(13) While the L isomer of **4** is commercially available and has been made synthetically (see ref 14), for cost effeciency and safety we adopted another route.
(14) Gomez-Vidal, J. A.; Forrester, M. T.; Silverman, R. B. *Org. Lett.* **2001**, *3*, 2477−2479.
(15) McKillop, A.; Taylor, R. J. K.; Watson, R. J.; Lewis, N. *Synthesis* **1994**, 31−33.
(16) Janusz, J.; Gryko, D.; Kobrzycka, E.; Gruza, H.; Prokopowicz, P. *Tetrahedon* **1998**, *54*, 6051−6064.
(17) Brown, H. C.; Narasimhan, S.; Choi, Y. M. *J. Org. Chem.* **1982**, *47*, 4702−4708.
(18) Zhao, M.; Li, J.; Mano, E.; Song, Z.; Tschaen, D.; Grabowski, E. J. J.; Reider, P. *J. Org. Chem.* **1999**, *64*, 2564−2566.
(19) See the Supporting Information.

cleaved, namely observation of an abundant byproduct (~50% by HPLC peak integration) with a mass 172 Da greater than the desired product. The side reaction occurred independently of the cleavage cocktail and the abundance of the byproduct increased as a function of time. Screening of model tripeptides indicated that the reaction was associated with the Fmoc-Arg(Pbf)-OH building block. Purification of the byproduct formed from a model tripeptide (Ac-Arg-FGly-Ala-OH) allowed its structure to be assigned as **10** (Figure 2) by HRMS and 2D NMR.[19]

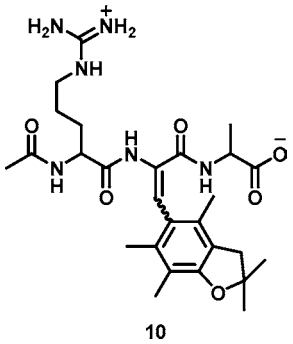


Figure 2. Modified tripeptide **10**.

This unique structure likely results from a Friedel−Crafts-type reaction between the cleaved 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) protecting group and FGly (Figure 3). The initial attack of the electron-rich aromatic

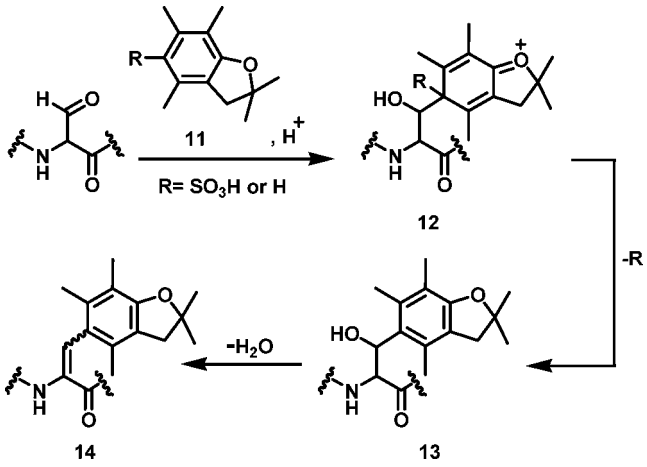


Figure 3. Proposed mechanism for the formation of byproduct.

ring (**11**) on the FGly aldehyde could occur either before or after desulfonation to form transient intermediate **12**. Loss of either a proton or SO_3 would provide **13**, an adduct we detected by LCMS analysis. Subsequent dehydration then results in the proposed modified peptide **14**.

This was an interesting development as most side reactions during peptide cleavage result from the nucleophilic nature

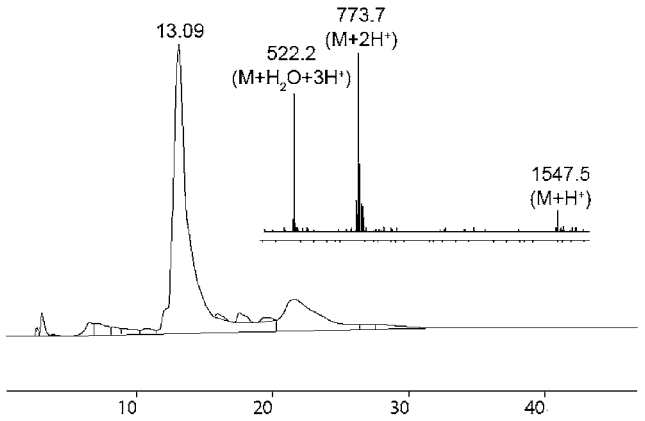


Figure 4. HPLC and mass spectrometry analysis of **9**.

of amino acid side chains. These are avoided by the inclusion of nucleophilic scavengers in the cleavage reaction (i.e., silanes, EDT, or anisole). However, the electrophilic nature of the FGly side chain produces a novel situation in which the scavengers have either no effect or, worse, react with the FGly residue. We attempted to avoid the side reaction during peptide cleavage using (1) various nucleophilic scavengers, (2) altered cleavage temperatures, or (3) electrophilic scavengers (i.e., aromatic aldehydes or 1,3-dicarbonyl-containing compounds). None of these conditions suppressed formation of the byproduct. We therefore conclude that the Pbf protecting group is incompatible with FGly.

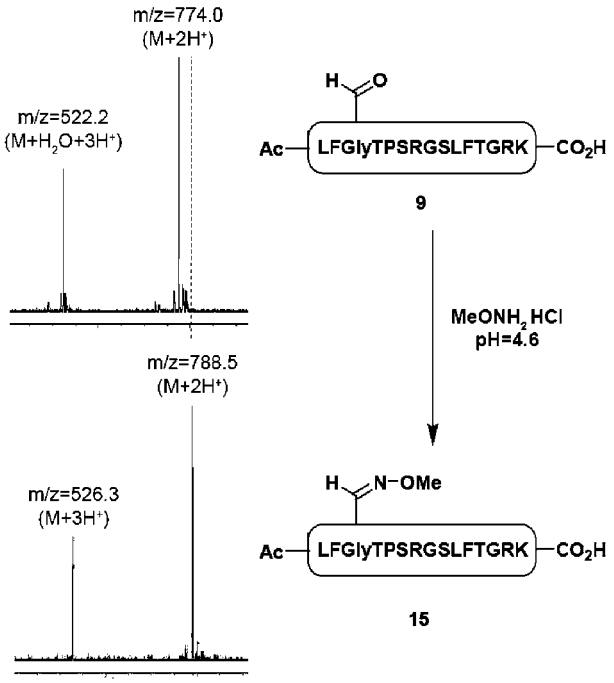


Figure 5. Mass spectrometry analysis of peptide **9** and its oxime adduct **15**.

Accordingly, an alternative arginine building block was employed. With use of commercially available Fmoc-Arg-$(Boc)_2$-OH, the synthesis and cleavage of tetradecapeptide **9** was accomplished without incident (Figure 4). Exposure of purified peptide to $MeONH_2$ afforded complete conversion of the aldehyde to oxime **15** (Figure 5), confirming that the aldehyde retains normal reactivity.

In summary, the Fmoc-FGly building block and SPPS methods for its use that we report here should facilitate studies of sulfatases and FGEs that activate them. In addition, the ability to synthesize peptides with internal electrophilic groups such as the aldehyde may find use in generating cyclic peptides or in conjugating unnatural epitopes to peptides for biological studies.

Acknowledgment. We thank Fiona Lin for NMR assistance. This work was supported by a grant to C.R.B. from the National Institutes of Health (AI51622).

Supporting Information Available: Experimental procedures, spectral and analytical data, and 2D-NMR spectra for **10**. This material is available free of charge via the Internet at http://pubs.acs.org.

OL052623T