Peptide Synthesis

Deutsche Ausgabe: DOI: 10.1002/ange.201509016
Internationale Ausgabe: DOI: 10.1002/anie.201509016

Synthesis of Sulfotyrosine-Containing Peptides by Incorporating Fluorosulfated Tyrosine Using an Fmoc-Based Solid-Phase Strategy

Wentao Chen[+], Jiajia Dong[+], Suhua Li, Yu Liu, Yujia Wang, Leonard Yoon, Peng Wu, K. Barry Sharpless, and Jeffery W. Kelly**

Abstract: *Tyrosine O-sulfation is a common protein post-translational modification that regulates many biological processes, including leukocyte adhesion and chemotaxis. Many peptides with therapeutic potential contain one or more sulfotyrosine residues. We report a one-step synthesis for Fmoc-fluorosulfated tyrosine. An efficient Fmoc-based solid-phase peptide synthetic strategy is then introduced for incorporating the fluorosulfated tyrosine residue into peptides of interest. Standard simultaneous peptide-resin cleavage and removal of the acid-labile side-chain protecting groups affords the crude peptides containing fluorosulfated tyrosine. Basic ethylene glycol, serving both as solvent and reactant, transforms the fluorosulfated tyrosine peptides into sulfotyrosine peptides in high yield.*

Tyrosine *O*-sulfation is a common enzymatic post-translational modification that occurs while the secreted and transmembrane proteome traffics through the Golgi compartment of the cell.[1] Phosphorylation and sulfation of tyrosine (Tyr) similarly modulate protein–protein interactions and affect conformational changes within a protein.[2]

Currently, one of several approaches can be used for the solid-phase peptide synthesis (SPPS) of protein fragments and polypeptides comprising sulfotyrosine (sY) residues. In the oldest approach, Fmoc-based (Tyr-OSO_3^-)$^+$Na (Fmoc = (9H-fluoren-9-ylmethoxy)-carbonyl) is simply coupled into the growing peptide;[3] however, subsequent couplings can be challenging and direct incorporation of more than one sY residue often compromises resin swelling, impeding further amino acid coupling steps.[3c,d] Moreover, acidic deprotection conditions often lead to desulfation. In an alternative method, the neopentyl (Np) group is used to protect the sY residue as a neutral sulfate diester, that is, as Tyr-*O*-SO_2-ONp.[2,4] The Fmoc-Tyr(OSO_3Np)-OH amino acid is obtained by a four-step synthesis in 66% overall yield and is currently commercially available. Fmoc-based SPPS is used to incorporate the Np-protected sY diester building block into the peptide of interest. After cleavage of the peptide from the 2-chlorotrityl resin and removal of the standard side-chain protecting groups, the Np group is removed in 1–2 M ammonium acetate at 37 °C over 6–12 h.[2] Another option employs a five-step strategy to synthesize a dichlorovinyl sulfate ester protected sY residue that is incorporated by SPPS into the desired peptide employing an Fmoc-based strategy.[5] Resin cleavage and cleavage of the side-chain protecting groups are performed under typical conditions (TFA:TIPS:H_2O = 95:2.5:2.5 volume ratio; or TFA:phenol:H_2O:thioanisole:EDT = 82.5:5:5:5:2.5; where TFA = trifluoroacetic acid, TIPS = triisopropylsilane, and EDT = 1,2-ethanedithiol) and then the dichlorovinyl sulfate ester protecting group(s) is removed in solution by hydrogenolysis using 10% Pd/C (30 wt.%), H_2, and ammonium formate (9 equiv) in methanol at 25 °C for 1 h, minimizing desulfation.[5] The hydrogenolysis step precludes the incorporation of cysteine (Cys) residues into the peptides. An even more recent strategy involves the Fmoc-based solid-phase synthesis of peptides containing Tyr residues with distinct phenol protecting groups.[6] These protecting groups are selectively removed while the peptide is still attached to the resin and the phenol is subjected to tyrosine *O*-sulfation employing sulfuryl imidazolium salt treatment (8 equiv/phenol functional group). Acidic cleavage of the peptide from the resin and removal of the standard side-chain protecting groups are followed by removal of the 2,2,2-trichloroethyl protecting group by means of catalytic hydrogenation using $Pd(OH)_2$ on carbon to afford the sY-containing peptide of interest.[6a]

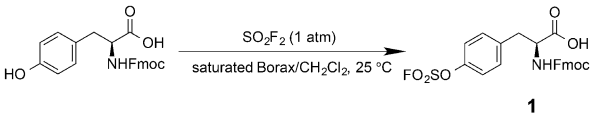
Herein, we report a short and efficient route to sY-containing peptides, wherein Fmoc-protected fluorosulfated tyrosine (Y(OSO_2F)) is incorporated into the peptide of interest through an Fmoc-based solid-phase synthetic strategy, either manually or by use of a peptide synthesizer. Like other sulfur(VI) fluorides, aromatic fluorosulfates are redox stable and hence do not serve as halogenation agents.[7] They are also very stable toward hydrolysis under neutral and acidic conditions, and moreover survive in basic milieu (for example, in phosphate buffer at pH 10).[7] However, the $ArOSO_2$-F linkage becomes reactive in the presence of an appropriate nucleophile only if the reaction conditions meet the stringent requirements for the departure of the "F^-" from its covalent link to the S^{VI} center.[7,8] In the case at hand, the contrast between high stability alongside the activatable sulfur(VI) fluoride exchange (SuFEx) reactivity pushes this latest click reaction to the very top; only the copper-catalyzed

[*] W. Chen,[+] Dr. J. Dong,[+] Dr. S. Li, Dr. Y. Liu, Dr. Y. Wang, L. Yoon, Dr. P. Wu, Dr. K. B. Sharpless, Dr. J. W. Kelly
Department of Chemistry
The Scripps Research Institute, La Jolla, CA 92037 (USA)
E-mail: sharples@scripps.edu
jkelly@scripps.edu
W. Chen,[+] Dr. J. W. Kelly
Department of Molecular and Experimental Medicine
The Scripps Research Institute, La Jolla, CA 92037 (USA)
Dr. Y. Liu, Dr. K. B. Sharpless, Dr. J. W. Kelly
The Skaggs Institute for Chemical Biology
The Scripps Research Institute, La Jolla, CA 92037 (USA)
[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201509016.

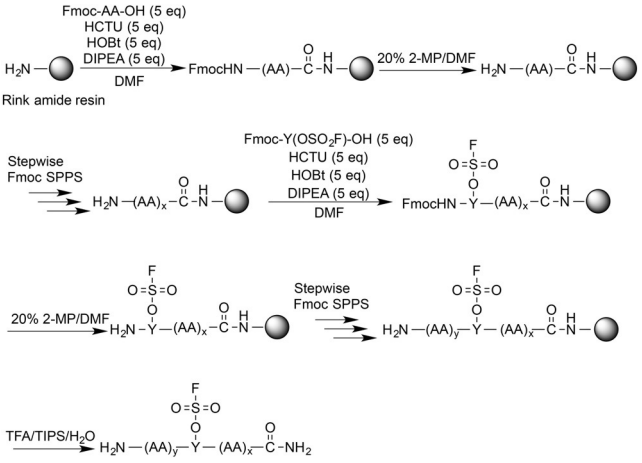
azide–alkyne cycloaddition (CuAAC) process is still standing in this rarified territory of click chemistry.[7–9] Click reactions are defined as processes that proceed under operationally simple conditions and generate products in high yields with minimal requirements for purification.[7,9]

The ease of obtaining the Fmoc-protected Y(OSO_2F) SPPS building block and the high stability of aromatic fluorosulfates enables the efficient synthesis of peptides containing the Ar-*O*-SO_2F side chain using a strategy based on Fmoc chemistry. The Fmoc-Y(OSO_2F)-OH amino acid (**1**) used in SPPS is prepared in one step in 96% yield by reacting commercially available Fmoc-protected Tyr and sulfuryl fluoride (gas) in a biphasic solvent system (CH_2Cl_2/saturated aqueous Borax buffer; Scheme 1). The synthesis was performed on a 5 g scale and is expected to be amenable to scale up. After removal of CH_2Cl_2 at reduced pressure and addition of 1M HCl, the precipitated product was removed by filtration, washed with water, dried, and used without further purification in SPPS. The Fmoc primary amine protecting group is removed during each SPPS cycle (Scheme 2) using 2-methylpiperidine (2-MP)[5b,10] to avoid a small but observable reaction between piperidine and the fluorosulfate functionality that lowered the yield and purity of the desired Y(OSO_2F)-containing peptides.[11] The fluorosulfate functional group is stable under the standard acidic Rink amide resin peptide cleavage conditions (TFA:TIPS:H_2O = 95:2.5:2.5 v/v/v) used to liberate the side-chain-deprotected peptide from the resin. The resin-free Y(OSO_2F) substructure(s) in the peptide of interest is then converted into the sY functionality by employing ethylene glycol as both reactant and solvent along with a base (Cs_2CO_3, or 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU); Scheme 3).



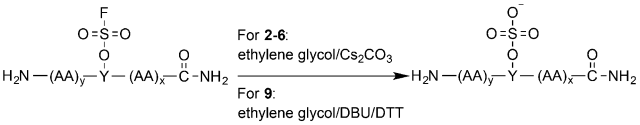
1

Scheme 1. Synthesis of Fmoc-fluorosulfated tyrosine **1**.



Scheme 2. Overview of the synthesis of Y(OSO_2F)-containing peptides. HCTU = O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt = 1-hydroxybenzotriazole; DIPEA = diisopropylethylamine; DMF = dimethylformamide; TIPS = triisopropylsilane.

Five sY-containing peptides **2–6** (Table 1) were prepared using this optimized Fmoc-based SPPS strategy followed by arylfluorosulfate ethylene glycolysis (i.e., ethylene glycol solvolysis and ethylene oxide extrusion, yielding sY from Y(OSO_2F); see Scheme 4). The peptides were purified by high performance liquid chromatography (HPLC) as explained in more detail below. The peptide DADE**sY**L-NH_2 (**2**) comprises a sequence in the epidermal growth factor receptor (EGFR), which when tyrosine *O*-sulfated is expected to be a good inhibitor of protein tyrosine phosphatase 1B.[12] The monosulfated peptide YE**sY**LDYDF-NH_2 (**3**) and the trisulfated peptide **sY**E**sY**LD**sY**DF-NH_2 (**4**) correspond to residues 5–12 of mature P-selectin glycoprotein ligand 1 (PGSL-1) that binds to P-selectin and plays an important role in the rolling adhesion of leukocytes on the vascular endothelium.[3c,13] Disulfated peptide TTPD**sY**GH-**sY**DDKDTLDLNTPVDK-NH_2 (**5**) is a substructure of C5aR, a classical G-protein-coupled receptor that is implicated in many inflammatory diseases.[14] The tetrasulfated peptide DADSENSSF**sY**s**Y**s**Y**D**sY**LDEVAF-NH_2 (**6**) corresponds to residues 14–33 of chemokine receptor D6, which scavenges extracellular pro-inflammatory CC chemokines and suppresses inflammation and tumorigenesis.[15] Peptides **4–6**, which contain multiple sY residues, could be isolated in respectable yields, reflecting the efficiencies of incorporating Y(OSO_2F) and arylfluorosulfate ethylene glycolysis in different peptide sequences (Table 1).



Scheme 3. Overview of arylfluorosulfate ethylene glycolysis to afford sY-containing peptides.

Table 1: Amino acid sequences and yields[a] of sY- and Y(OSO_2F)-containing peptides.

No.	Protein Subsequence	Amino Acid Sequence	Yield [%][a]
2	EGFR (988–993)	DADE**sY**L-NH_2	67
3	PGSL-1 (5–12)	YE**sY**LDYDF-NH_2	54
4	PGSL-1 (5–12)	**sY**E**sY**LD**sY**DF-NH_2	58
5	C5aR (7–28)	TTPD**sY**GH**sY**DDKDTLDLNTPVDK-NH_2	54
6	D6 (14–33)	DADSENSSF**sYsYsY**D**sY**LDEVAF-NH_2	36
7	EGFR (988–993)	DADE**Y(OSO$_2$F)**L-NH_2	64
8	CXCR4 (19–30)	GD**Y(OSO$_2$F)**DSMKEPCFR-NH_2	40
9	CXCR4 (19–30)	GD**sY**DSMKEPCFR-NH_2	35

[a] Yield of the isolated product.

For all SPPS couplings, including the coupling of amino acid **1**, we used 5 equiv of the appropriate side-chain-protected amino acid preactivated with HCTU/HOBt/DIPEA (1:1:1 v/v/v) for 30 min. The activated amino acid

was added to the resin-bound primary amine with stirring or shaking for a coupling period of 30–60 min. Every Fmoc protecting group was removed employing three applications of 20% 2-MP in dimethylformamide or *N*-methyl-2-pyrrolidone (alternative solvent) for 10 min. We used preferentially the TFA:TIPS:H_2O (95:2.5:2.5) deprotection solution (25 °C, 180 min) to cleave the peptide of interest off the Rink resin and to liberate the standard side-chain protecting groups; however, the other cleavage/deprotection combination mentioned above was used without a noticeable change in the purity of the crude peptide generated. The Y(OSO_2F)-containing peptides can be easily purified by reverse-phase HPLC (RP-HPLC), exemplified by DADEY(OSO_2F)L-NH_2 (**7**) which was isolated in 64% yield. Arylfluorosulfate ethylene glycolysis of **7** using Cs_2CO_3 revealed complete conversion into **2** without any noticeable additional peak in the analytical HPLC chromatogram (Figure 1A; see also Figure S1 and S2 in the Supporting Information). However, there is no need to purify the crude Y(OSO_2F) peptides before ethylene glycolysis. The crude Y(OSO_2F)-containing peptides can be directly subjected to arylfluorosulfate ethylene glycolysis using Cs_2CO_3 as the base.[16] The sY-containing peptides were then purified by semipreparative RP-HPLC using a C18 column and a 20 mM ammonium acetate/CH_3CN mobile phase gradient (minimizes desulfation by maintaining a near-neutral pH environment). Using this approach, sY-containing peptides **2**–**6** were obtained in 36–67% yield (Table 1) after RP-HPLC purification.

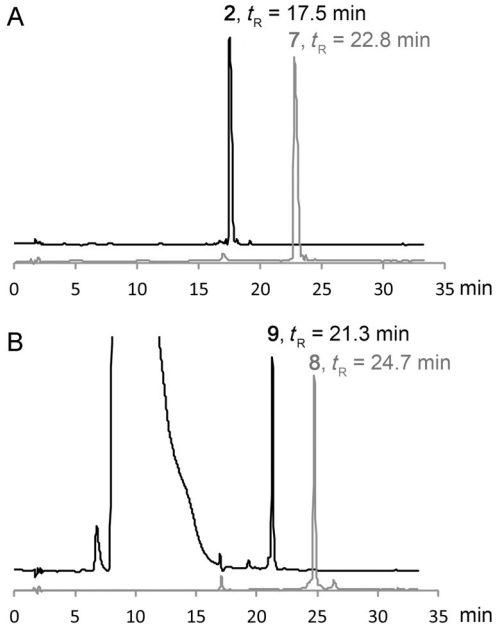
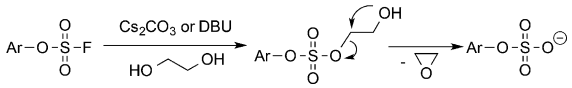


Figure 1. Monitoring the ethylene glycolysis of purified Y(OSO_2F)-containing peptides A) **7** and B) **8**, affording sY-containing peptides **2** and **9**, respectively, by RP-HPLC using 20 mM ammonium acetate/CH_3CN mobile phases. Purified **7** (gray line in (A)) and **8** (gray line in (B)) dissolved in ethylene glycol were analyzed by analytical RP-HPLC with retention times t_R = 22.8 and 24.7 min, respectively. After adding Cs_2CO_3 for **7** and DBU/DTT for **8** and stirring for 120 min, the samples were analyzed using the same gradient. sY-containing peptides **2** and **9** eluted at t_R = 17.5 and 21.3 min, respectively. The absorption in (B) between 5 and 17 min is from DBU and DTT.

In the optimization of the hydrolysis to convert Y-(OSO_2F)-containing peptide **7** into sY-containing peptide **2**, we detected significant desulfation of sY in the presence of base in aqueous solutions. In addition, upon treating peptide **7** with Cs_2CO_3 dissolved in methanol, we detected the apparent methylation of peptide **7**, presumably owing to the formation of a Tyr-*O*-SO_2-OCH_3 intermediate, which appears to transfer a methyl group to a neighboring carboxylate side chain (Figure S3). Neighboring acidic residues are common at sites of protein tyrosine sulfation, thus it is important to solve this problem.[1b,c,17] Although utilization of methanol/NH_3(2 M)/Cs_2CO_3 attenuated methylation, it was still observed. Utilizing Cs_2CO_3 dissolved in ethanol resulted in peptide ethylation, consistent with formation of a Tyr-*O*-SO_2-OCH_2CH_3 intermediate (Figure S4). With Cs_2CO_3 dissolved in isopropanol or tertiary butyl alcohol, no reaction occurred. Notably, although ethylene glycol/Cs_2CO_3 and 1,4-butanediol/Cs_2CO_3 combinations afforded quantitative lysis with no side reactions, 1,3-propanediol/Cs_2CO_3 afforded less than 50% yield of sY-containing peptide **2** and numerous side products (Figure S5). We hypothesized that efficient cyclic ether formation is key to the mechanism of arylfluorosulfate ethylene glycolysis (Scheme 4). To support this hypothesis, we explored the ethylene glycolysis of the small molecule Ph-*O*-SO_2-F (5 mmol), employing Cs_2CO_3 or DBU as the base. This relatively large-scale ethylene glycolysis reaction generated gaseous ethylene oxide, whose identity was confirmed by ^1H and ^{13}C NMR employing a distillation-like capture in cold $CDCl_3$ (Figure S6–S8).



Scheme 4. Proposed mechanism of arylfluorosulfate ethylene glycolysis.

The crude peptide GDY(OSO_2F)DSMKEPCFR-NH_2 (**8**) (Table 1), containing Cys and methionine residues, was successfully synthesized using the SPPS strategy and the side-chain deprotection/resin cleavage approach outlined above. Peptide **8** was then purified by HPLC and was isolated in a yield of 40% (based on resin loading) in order to optimize the hydrolysis strategy for affording sY-based peptides containing Cys, in this case peptide GD**sY**DSMKEPCFR-NH_2 (**9**; Table 1). This optimization was carried out because the ethylene glycolysis/Cs_2CO_3 method generated a SCH_2-CH_2-OH functional group on the Cys side chain of **9**, as discerned by high resolution mass spectrometry (Figure S9).[18] The optimized procedure for the ethylene glycolysis of peptide **8** employed 5% DBU as the base in ethylene glycol containing 0.5% dithiothreitol (DTT). This ethylene glycolysis solution converted **8** into **9**, without any discernable by-product based on monitoring the reaction by HPLC (Figure 1B; Figures S10 and S11). The addition of DTT was the key to minimizing the ethylene-oxide-derived thiol

alkylation mentioned above. This approach afforded sY-containing peptide **9** in 35% isolated yield (Table 1) based on resin loading. Peptide **9** comprises residues 19–30 of CXCR4, which is crucial for embryonic development and has been implicated in cancer metastasis and HIV infection.[19]

In summary, we have described a one-step synthesis of the Fmoc-Y(OSO_2F)-OH amino acid used without purification for the synthesis of Y(OSO_2F)-containing peptides. We demonstrate that the Fmoc synthesis of Y(OSO_2F)-containing peptides is both practical and efficient. Standard side-chain deprotection and resin cleavage solutions perform well. Two different fluorosulfate ethylene glycolysis procedures are introduced for the efficient production of sY-containing peptides depending on whether the peptide lacks or contains a Cys residue. The facile synthesis described herein takes advantage of the unique reactivity of sulfur(VI) fluorides. Our approach can easily be implemented by commercial and academic peptide synthesis facilities, since the Fmoc-Y(OSO_2F)-OH amino acid will become commercially available.

Acknowledgements

We thank the Skaggs Institute for Chemical Biology and NIH GM051105 (J.W.K.) for financial support.

Keywords: click chemistry · peptides · solid-phase synthesis · SuFEx · sulfotyrosine

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 1835–1838
Angew. Chem. **2016**, *128*, 1867–1870

[1] a) W. B. Huttner, *Annu. Rev. Physiol.* **1988**, *50*, 363–376; b) J. W. Kehoe, C. R. Bertozzi, *Chem. Biol.* **2000**, *7*, R57–61; c) C. Seibert, T. P. Sakmar, *Biopolymers* **2008**, *90*, 459–477.

[2] L. S. Simpson, J. Z. Zhu, T. S. Widlanski, M. J. Stone, *Chem. Biol.* **2009**, *16*, 153–161.

[3] a) T. Yagami, S. Shiwa, S. Futaki, K. Kitagawa, *Chem. Pharm. Bull.* **1993**, *41*, 376–380; b) K. Kitagawa, S. Futaki, T. Yagami, S. Sumi, K. Inoue, *Int. J. Pept. Protein Res.* **1994**, *43*, 190–200; c) K. M. Koeller, M. E. Smith, C. H. Wong, *Bioorg. Med. Chem.* **2000**, *8*, 1017–1025; d) T. Young, L. L. Kiessling, *Angew. Chem. Int. Ed.* **2002**, *41*, 3449–3451; *Angew. Chem.* **2002**, *114*, 3599–3601.

[4] L. S. Simpson, T. S. Widlanski, *J. Am. Chem. Soc.* **2006**, *128*, 1605–1610.

[5] a) A. M. Ali, B. Hill, S. D. Taylor, *J. Org. Chem.* **2009**, *74*, 3583–3586; b) A. M. Ali, S. D. Taylor, *Angew. Chem. Int. Ed.* **2009**, *48*, 2024–2026; *Angew. Chem.* **2009**, *121*, 2058–2060; c) A. M. Ali, S. D. Taylor, *J. Pept. Sci.* **2010**, *16*, 190–199.

[6] a) X. Liu, L. R. Malins, M. Roche, J. Sterjovski, R. Duncan, M. L. Garcia, N. C. Barnes, D. A. Anderson, M. J. Stone, P. R. Gorry, R. J. Payne, *ACS Chem. Biol.* **2014**, *9*, 2074–2081; b) D. Taleski, S. J. Butler, M. J. Stone, R. J. Payne, *Chem. Asian J.* **2011**, *6*, 1316–1320.

[7] J. Dong, L. Krasnova, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2014**, *53*, 9430–9448; *Angew. Chem.* **2014**, *126*, 9584–9603.

[8] J. Dong, K. B. Sharpless, L. Kwisnek, J. S. Oakdale, V. V. Fokin, *Angew. Chem. Int. Ed.* **2014**, *53*, 9466–9470; *Angew. Chem.* **2014**, *126*, 9620–9624.

[9] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021; *Angew. Chem.* **2001**, *113*, 2056–2075.

[10] J. Hachmann, M. Lebl, *J. Comb. Chem.* **2006**, *8*, 149.

[11] When piperidine was used in the SPPS of peptide **7**, we could isolate some sulfamate as a by-product of peptide **7**. See the Supporting Information for details.

[12] a) S. Desmarais, Z. Jia, C. Ramachandran, *Arch. Biochem. Biophys.* **1998**, *354*, 225–231; b) N. R. Glover, A. S. Tracey, *Biochem. Cell Biol.* **1999**, *77*, 469–486.

[13] a) T. Pouyani, B. Seed, *Cell* **1995**, *83*, 333–343; b) D. Sako, K. M. Comess, K. M. Barone, R. T. Camphausen, D. A. Cumming, G. D. Shaw, *Cell* **1995**, *83*, 323–331; c) S. D. Rodgers, R. T. Camphausen, D. A. Hammer, *Biophys. J.* **2001**, *81*, 2001–2009.

[14] P. N. Monk, A. M. Scola, P. Madala, D. P. Fairlie, *Br. J. Pharmacol.* **2007**, *152*, 429–448.

[15] a) P. E. Blackburn, C. V. Simpson, R. J. Nibbs, M. O'Hara, R. Booth, J. Poulos, N. W. Isaacs, G. J. Graham, *Biochem. J.* **2004**, *379*, 263–272; b) M. Weber, E. Blair, C. V. Simpson, M. O'Hara, P. E. Blackburn, A. Rot, G. J. Graham, R. J. Nibbs, *Mol. Biol. Cell* **2004**, *15*, 2492–2508; c) C. V. McCulloch, V. Morrow, S. Milasta, I. Comerford, G. Milligan, G. J. Graham, N. W. Isaacs, R. J. Nibbs, *J. Biol. Chem.* **2008**, *283*, 7972–7982.

[16] All peptides were soluble in ethylene glycol at the concentration we used herein (10 mg mL^{-1}). Addition of Cs_2CO_3 made the solubilization process faster. The amount of Cs_2CO_3 used for hydrolysis varied with the peptide sequence. See the Supporting Information for details.

[17] a) W. H. Lin, K. Larsen, G. L. Hortin, J. A. Roth, *J. Biol. Chem.* **1992**, *267*, 2876–2879; b) T. Teramoto, Y. Fujikawa, Y. Kawaguchi, K. Kurogi, M. Soejima, R. Adachi, Y. Nakanishi, E. Mishiro-Sato, M. C. Liu, Y. Sakakibara, M. Suiko, M. Kimura, Y. Kakuta, *Nat. Commun.* **2013**, *4*, 1572.

[18] The molecular weight of the peptide by-product generated using ethylene glycol/Cs_2CO_3 glycolysis is 44 Da heavier than the molecular weight of peptide **9**. Similarly, we used a tripeptide containing a Cys residue to capture the ethylene oxide generated during arylfluorosulfate ethylene glycolysis and obtained the ethylene oxide adduct as a by-product in the context of this peptide as well. See the Supporting Information for more details.

[19] a) C. T. Veldkamp, C. Seibert, F. C. Peterson, T. P. Sakmar, B. F. Volkman, *J. Mol. Biol.* **2006**, *359*, 1400–1409; b) C. Seibert, C. T. Veldkamp, F. C. Peterson, B. T. Chait, B. F. Volkman, T. P. Sakmar, *Biochemistry* **2008**, *47*, 11251–11262; c) C. T. Veldkamp, C. Seibert, F. C. Peterson, N. B. De la Cruz, J. C. Haugner 3rd, H. Basnet, T. P. Sakmar, B. F. Volkman, *Sci. Signaling* **2008**, *1*, ra4.

Received: September 25, 2015
Published online: December 22, 2015

