HHS Public Access

Author manuscript

Angew Chem Int Ed Engl. Author manuscript; available in PMC 2017 January 26.

Published in final edited form as:

Angew Chem Int Ed Engl. 2016 January 26; 55(5): 1835–1838. doi:10.1002/anie.201509016.

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

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

[‡]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA

[#]Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA

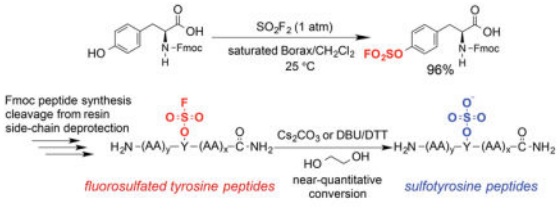
[§]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

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 also contain a sulfotyrosine residue(s). We report a one-step synthesis of Fmoc-fluorosulfated tyrosine. An efficient Fmoc solid-phase peptide synthesis 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 as solvent and reactant, transforms the fluorosulfated tyrosine peptides into sulfotyrosine peptides in high yield.

SuFEx reaction enables sulfotyrosine peptide synthesis

A facile Fmoc solid-phase synthesis of fluorosulfated tyrosine peptides is reported, enabled by a one-step synthesis of Fmoc-fluorosulfated tyrosine. An efficient ethylene glycolysis method for the transformation of side-chain deprotected fluorosulfated tyrosine peptides to sulfotyrosine peptides is central to the efficiency of this methodology.



[*]Corresponding authors: KBS: sharples@scripps.edu; JWK: jkelly@scripps.edu.

[†]These authors contributed equally to this work

Supporting information for this article is given via a link at the end of the document.

Keywords

fluorosulfate; peptides; SuFex; sulfotyrosine; sulfur

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 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, 3d] Moreover, acidic deprotection conditions often lead to desulfation. In an alternative method, the neopentyl (Np; $(CH_3)_3CCH_2$-) group is used to protect the **sY** residue as a neutral sulfate diester, i.e., as Tyr-O-SO_2-ONp.[2, 4] The Fmoc-Tyr(OSO_3Np)-OH building block is obtained by a 4-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 5-step strategy to synthesize a dichlorovinyl sulfate ester protected **sY** residue that is incorporated by SPPS into the desired peptide employing an Fmoc strategy.[5] Resin cleavage and side-chain protecting group cleavage are performed as usual (95:2.5:2.5 = trifluoroacetic acid (TFA):triisopropylsilane (TIPS):H_2O or 82.5:5:5:5:2.5 = TFA:phenol:H_2O:thioanisole:1,2-Ethanedithiol(EDT)) and then the dichlorovinyl sulfate ester protecting group(s) is removed in solution by hydrogenolysis (balloon) using 30 weight % of 10% Pd/C, H_2 and 9 equiv. of ammonium formate 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 via catalytic hydrogenation using $Pd(OH)_2$ on carbon to afford the **sY** 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 via a Fmoc solid phase synthesis 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 (e.g., 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 needs 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 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 an Fmoc chemistry strategy. 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 scaling up. After removal of CH_2Cl_2 at reduced pressure and addition of 1 M HCl, the precipitated product is 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-methyl-piperidine (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 (95:2.5:2.5 = TFA:TIPS:H_2O) 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 reactant and solvent along with a base (Cs_2CO_3, or 1,8-diazabicyclo[5.4.0]undec-7-ene(DBU)) (Scheme 3).

Five **sY**-containing peptides, **2**–**6** (Table 1), were prepared using this optimized Fmoc SPPS strategy followed by arylfluorosulfate ethylene glycolysis. 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 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] The respectable isolated yields of peptides **4**–**6**, which contain multiple **sY** residues, reflects the efficiencies of incorporating Y(OSO_2F) and arylfluorosulfate ethylene glycolysis in different peptide sequences (Table 1).

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) 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 3 applications of 20% 2-MP in dimethylformamide or N-methyl-2-pyrrolidone (alternative solvent) for 10 min. We used and prefer the 95:2.5:2.5 = TFA:TIPS:H_2O 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 cocktail 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 (RP)-HPLC, exemplified by DADEY(OSO_2F)L-NH_2 **(7)** in a 64% yield. Arylfluorosulfate ethylene glycolysis of **7** using Cs_2CO_3 revealed complete conversion to **2**, without any noticeable additional peak in the analytical HPLC chromatogram (Figure 1A, S1 and S2). 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** peptides were then purified by semi-preparative 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). Using this approach, **sY** peptides **2–6** were obtained in 36–67% yield (Table 1) after RP-HPLC purification.

In the optimization of the hydrolysis of Y(OSO_2F)-containing peptide **7** to **sY**-containing peptide **2**, we observed 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 observed 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, 1c, 17] While utilization of methanol/NH_3(2M)/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, while ethylene glycol/Cs_2CO_3 and 1,4 butanediol/Cs_2CO_3 combinations afforded quantitative lysis with no side reactions, 1,3 propanediol/Cs2CO_3 afforded < 50 % yield of **sY** 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 mmoles), 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).

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 side chain deprotection/resin cleavage approach outlined above. Peptide **8** was then HPLC purified in an isolated yield of 40 % (based on resin loading) in order to optimize the hydrolysis

strategy for affording **sY** peptides containing Cys, in this case peptide GD**sY**DSMKEPCFR-NH_2 (**9;** Table 1). This was done because the ethylene glycolysis/Cs_2CO_3 method generated a S CH_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 protocol 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 byproduct based on HPLC reaction monitoring (Figure 1B, S10 and S11). Adding DTT was the key to minimizing the ethylene oxide-derived thiol alkylation mentioned above. This approach afforded **sY** 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) 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 protocols are introduced for the efficient production of **sY** 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) amino acid is commercially available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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

References

1. a) Huttner WB. Annu Rev Physiol. 1988; 50:363–376. [PubMed: 3288098] b) Kehoe JW, Bertozzi CR. Chem Biol. 2000; 7:R57–61. [PubMed: 10712936] c) Seibert C, Sakmar TP. Biopolymers. 2008; 90:459–477. [PubMed: 17680702]
2. Simpson LS, Zhu JZ, Widlanski TS, Stone MJ. Chem Biol. 2009; 16:153–161. [PubMed: 19246006]
3. a) Yagami T, Shiwa S, Futaki S, Kitagawa K. Chem Pharm Bull. 1993; 41:376–380. [PubMed: 8500203] b) Kitagawa K, Futaki S, Yagami T, Sumi S, Inoue K. Int J Pept Protein Res. 1994; 43:190–200. [PubMed: 8200739] c) Koeller KM, Smith ME, Wong CH. Bioorg Med Chem. 2000; 8:1017–1025. [PubMed: 10882013] d) Young T, Kiessling LL. Angew Chemie. 2002; 41:3449–3451.
4. Simpson LS, Widlanski TS. J Am Chem Soc. 2006; 128:1605–1610. [PubMed: 16448132]
5. a) Ali AM, Hill B, Taylor SD. J Org Chem. 2009; 74:3583–3586. [PubMed: 19331341] b) Ali AM, Taylor SD. Angew Chemie. 2009; 48:2024–2026.c) Ali AM, Taylor SD. J Pept Sci. 2010; 16:190–199. [PubMed: 20196090]
6. a) Liu X, Malins LR, Roche M, Sterjovski J, Duncan R, Garcia ML, Barnes NC, Anderson DA, Stone MJ, Gorry PR, et al. ACS Chem Biol. 2014; 9:2074–2081. [PubMed: 24963694] b) Taleski D, Butler SJ, Stone MJ, Payne RJ. Chem Asian J. 2011; 6:1316–1320. [PubMed: 21509944]
7. Dong J, Krasnova L, Finn MG, Sharpless KB. Angew Chemie. 2014; 53:9430–9448.

8. Dong J, Sharpless KB, Kwisnek L, Oakdale JS, Fokin VV. Angew Chemie. 2014; 53:9466–9470.
9. Kolb HC, Finn MG, Sharpless KB. Angew Chemie. 2001; 40:2004–2021.
10. Hachmann J, Lebl M. J Comb Chem. 2006; 8:149. [PubMed: 16529506]
11. When we used piperidine in the SPPS of peptide 7 we could purify the piperidine addition byproduct for peptide 7. See the Supporting Information for details.
12. a) Desmarais S, Jia Z, Ramachandran C. Arch Biochem Biophys. 1998; 354:225–231. [PubMed: 9637730] b) Glover NR, Tracey AS. Biochem Cell Biol. 1999; 77:469–486. [PubMed: 10593610]
13. a) Pouyani T, Seed B. Cell. 1995; 83:333–343. [PubMed: 7585950] b) Sako D, Comess KM, Barone KM, Camphausen RT, Cumming DA, Shaw GD. Cell. 1995; 83:323–331. [PubMed: 7585949] c) Rodgers SD, Camphausen RT, Hammer DA. Biophysical J. 2001; 81:2001–2009.
14. Monk PN, Scola AM, Madala P, Fairlie DP. Br J Pharmacol. 2007; 152:429–448. [PubMed: 17603557]
15. a) Blackburn PE, Simpson CV, Nibbs RJ, O'Hara M, Booth R, Poulos J, Isaacs NW, Graham GJ. Biochemical J. 2004; 379:263–272.b) Weber M, Blair E, Simpson CV, O'Hara M, Blackburn PE, Rot A, Graham GJ, Nibbs RJ. Mol Biol Cell. 2004; 15:2492–2508. [PubMed: 15004236] c) McCulloch CV, Morrow V, Milasta S, Comerford I, Milligan G, Graham GJ, Isaacs NW, Nibbs RJ. J Biol Chem. 2008; 283:7972–7982. [PubMed: 18201974]
16. All peptides were soluble in ethylene glycol in the concentration we used (10 mg/ml). 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) Lin WH, Larsen K, Hortin GL, Roth JA. J Biol Chem. 1992; 267:2876–2879. [PubMed: 1737745] b) Teramoto T, Fujikawa Y, Kawaguchi Y, Kurogi K, Soejima M, Adachi R, Nakanishi Y, Mishiro-Sato E, Liu MC, Sakakibara Y, et al. Nat Commun. 2013; 4:1572. [PubMed: 23481380]
18. The molecular weight of the peptide byproduct generated using ethylene glycol/Cs_2CO_3 glycolysis is 44 Daltons heavier than the molecular weight of peptide 9. Similarly, we used a tripeptide containing Cys residue to capture the ethylene oxide generated during arylfluorosulfate ethylene glycolysis and obtained the ethylene oxide byproduct in the context of this peptide as well. See the Supporting Information for more details.
19. a) Veldkamp CT, Seibert C, Peterson FC, Sakmar TP, Volkman BF. J Mol Biol. 2006; 359:1400–1409. [PubMed: 16725153] b) Seibert C, Veldkamp CT, Peterson FC, Chait BT, Volkman BF, Sakmar TP. Biochemistry. 2008; 47:11251–11262. [PubMed: 18834145] c) Veldkamp CT, Seibert C, Peterson FC, De la Cruz NB, Haugner JC 3rd, Basnet H, Sakmar TP, Volkman BF. Sci Signal. 2008; 1:ra4. [PubMed: 18799424]

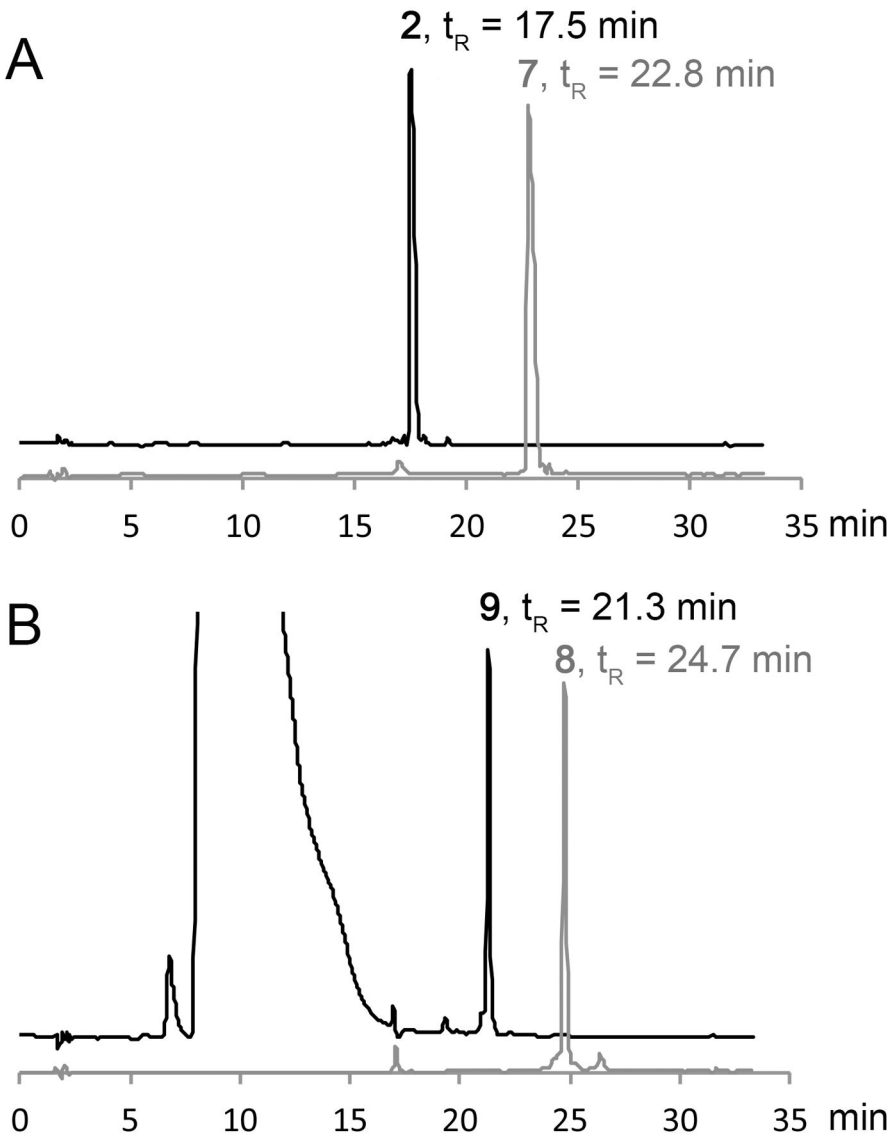
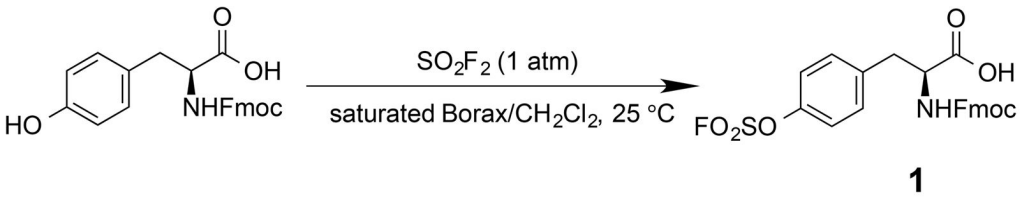


Figure 1.

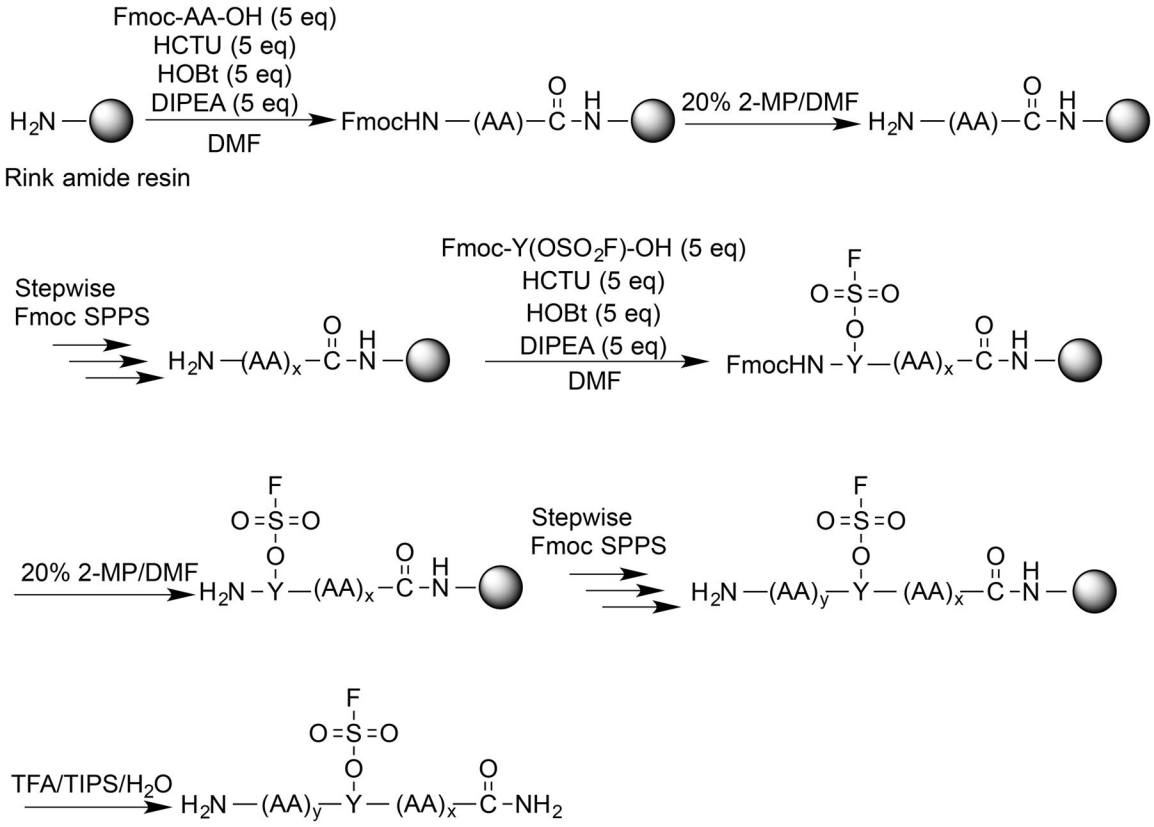
Monitoring the ethylene glycol-mediated hydrolysis of purified Y(OSO_2F) peptides **7** and **8**, affording **sY** peptides **2** and **9**, respectively, by RP-HPLC using 20 mM ammonium acetate/CH_3CN mobile phases. Purified **7** (gray line in A) and **8** (gray line in B) dissolved in ethylene glycol were analyzed by analytical RP-HPLC with 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** 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.



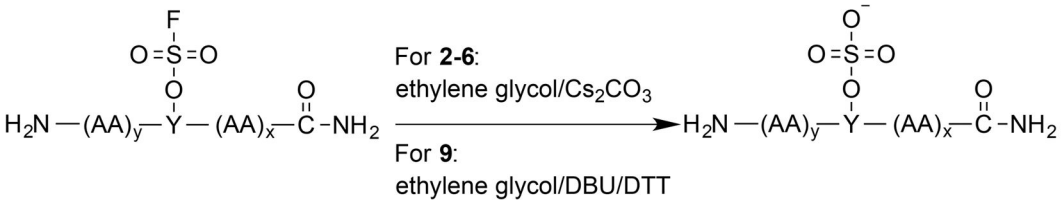
1

Scheme 1.

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

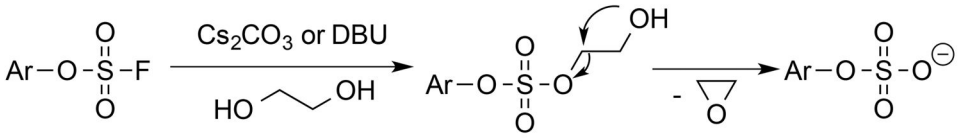


Scheme 2.
Overview of Y(OSO_2F) peptide synthesis



Scheme 3.

Overview of arylfluorosulfate hydrolysis to afford **sY** peptides.



Scheme 4.

Proposed mechanism of arylfluorosulfate ethylene glycolysis

Table 1

Amino acid sequences and isolated yields of **sY** and **Y(OSO$_2$F)** peptides.

No.	Protein Subsequence	Amino Acid Sequence	Isolated Yield
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**sY**s**Y**s**Y**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%