Accepted Article

Title: Chemical methods for N- and O-sulfation of small molecules, amino acids and peptides

Authors: Anna Mary Benedetti, Daniel M Gill, Chi W Tsang, and Alan M Jones

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

COMMUNICATION

Chemical methods for *N*- and *O*-sulfation of small molecules, amino acids and peptides

Anna Mary Benedetti,[a] Daniel M. Gill,[a] Chi W. Tsang,[b] and Alan M. Jones*[a]

Sulfation of the amino acid residues of proteins is an emerging post-translational modification, the functions of which are yet to be fully understood. Current sulfation methods are limited mainly to *O*-tyrosine (sY) which requires negatively charged species around the desired amino acid residue and a specific sulfotransferase enzyme. Alternatively, for solid phase peptide synthesis a *de novo* protected sY is required. Therefore, synthetic routes are required to go beyond *O*-sulfation. We have developed a novel route to *N*-sulfamation and can dial-in/out *O*-sulfation (without *S*-sulfurothiolation), mimicking the initiation step of the ping-pong sulfation mechanism identified in structural biology. This rapid, low temperature and non-racemizing method is applicable to a range of amines, amides, amino acids, and peptide sequences.

The sulfate group is a ubiquitous post-translational modification, accounting for approximately 1% of all known epigenetic markers, and plays a vital role in a variety of biological processes, including protein-protein and oligosaccharide interactions.[1-9]

Yet, sulfoproteins have been overlooked in favour of more common phosphorylation, limiting our exploration of their function. Recent developments in ultra-high resolution mass spectrometry have revealed a *m/z* difference of 9.5 mDa between kinase phosphorylation *versus* sulfation.[10-12] Therefore, many previously assigned phosphorylated proteins have been miss-assigned.

Protein sulfation may be more abundant and important than previously recognised. Similarly to phosphorylation, proteins have been discovered to be sulfated at tyrosine (sY), serine (sS), threonine (sT) and histidine (sH).[13] Opening up the possibility of further sulfation motifs that are yet to be discovered.

Methods to access sulfated proteins to date have focussed on the sulfated amino acid sY,[2-5] with very limited examples of sS and sT.[14-15] The pre-installation of a sulfate group onto Y is possible *via* alternative protecting group methods,[14,16-26] with the added complication of the instability of sY to solid phase peptide cleavage. These chemical methods have various merits and drawbacks, but do allow for pinpoint accuracy in installing a sulfate group. Alternatively, the use of SULT, TPST, ASST sulfotransferases with PAPS cofactor allows for selectivity between tyrosine groups, due to the ±5 amino acids around Y

and effects of pre-existing sulfate groups (**Scheme 1**, previous work).[2,27] A ping-pong mechanism has been discovered to operate with arylsulfotransferases *via* the intermediacy of a histidine *N*-sulfamate,[28] this spurred our interest in developing a general *N*-sulfamation strategy to small molecules and peptides.

To the best of our knowledge, limited non-enzymatic, methods to directly chemically sulfate a peptide have only been reported at phenolic tyrosine residues.[29] Inspired by the role of sY, we sought to expand the sulfation epigenetic tool box to other amino acids. To address this, we detail a mild betaine-mediated method using TBSAB for the novel and rapid preparation of *N*-sulfamates. We demonstrate scope on a diverse array of nitrogen containing substrates and investigate heteroatom selectivity (*N*-sulfamation *vs O*-sulfation *vs S*-sulfurothiolation). We surveyed the reactivity on nine canonical nucleophilic amino acid residues and demonstrate the ability to chemically modify peptide sequences (**Scheme 1**, this work).

There are limited reliable methods[30] to prepare *N*-sulfamates: (i) direct sulfamation with an SO_3-amine complex (amines used include: Me_3N, Et_3N, Me_2Bz, Py, *N*-methylmorpholine) or an SO_3 source (chlorosulfonic acid); (ii) sulfuric acid-mediated hydrosulfation of isocyanates; and (iii) (for *N*-acyl sulfamates only)[31] the reverse addition of a sulfamate through a carbonate intermediate. Limitations of these methods include the requirement for extensive chromatography, and ion-exchange methods, incompatibility with acid sensitive substrates, and multi-step sequences.

We have previously developed a new class of sulfating agent, tributyl sulfoammonium betaine (TBSAB) that overcomes the purification challenges,[15] and considered whether TBSAB would provide a solution to these issues associated with accessing an under-exploited but important functional group. Using benzylamine (**1**) as a model system to study *N*-sulfamation, the reaction progress was readily monitored by ^{1}H NMR spectroscopy (**Table 1**). Subtle changes in the shielding of the benzylic methylene group gave insight into the variety of reaction pathways that could operate, including the dynamic interplay between protonation, sulfamation, counterion exchange and imidosulfamation.

[a]   A. M. Benedetti, D. M. Gill, Dr A. M. Jones
      School of Pharmacy
      University of Birmingham
      United Kingdom
      E-mail: a.m.jones.2@bham.ac.uk
[b]   Dr C. W. Tsang
      School of Chemistry
      University of Birmingham
      United Kingdom

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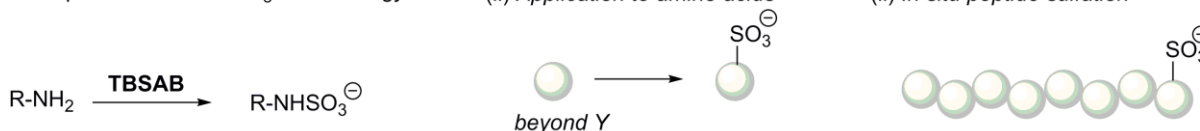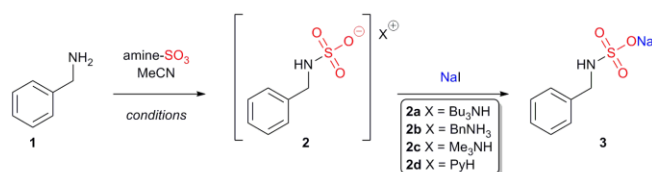
Previous work
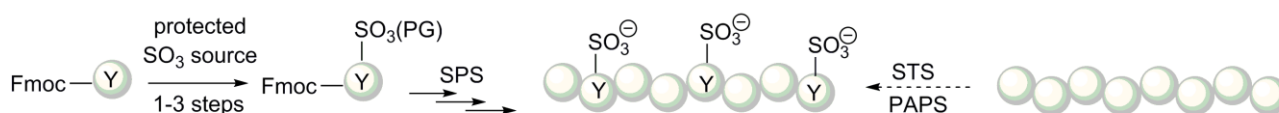
This work

(i) development of new N-SO₃ methodology

(ii) Application to amino acids

(ii) In situ peptide sulfation

Scheme 1. Previous approaches to sulfating tyrosine and sulfated peptides. *This work:* a new route to sulfating nitrogen, sulfation of a range of amino acids including tyrosine, and the non-enzymatic *in situ* sulfation of a peptide.

Entry	amine-SO₃	Eq	T (°C)	atm.	additive	Conv. (%)
1	Bu₃N	2.0	0	Ar	-	52[a]
2	Bu₃N	2.0	20	Ar	-	71
3	Bu₃N	2.0	30	Ar	-	>99 (98)
4	Bu₃N	2.0	60	Ar	-	87[b]
5	Bu₃N	2.0	80	Ar	-	84[b]
6	Bu₃N	0.5	30	Ar	-	>99[a]
7	Bu₃N	1.0	30	Ar	-	86[a]
8	Bu₃N	1.5	30	Ar	-	86
9	Bu₃N	3.0	30	Ar	-	81
10	Me₃N	2.0	60	Ar	-	>99 (23)
11	Py	2.0	60	Ar	-	47
12	Bu₃N	2.0	30	air	-	>99 (48)
13	Bu₃N	2.0	30	Ar	Bu₃N (0.5 eq)	>99
14	Bu₃N	2.0	80	Ar	Bu₃NHCl (0.1 eq)	>99
15	Bu₃N	2.0	80	Ar	Bu₃NHCl (0.5 eq)	78

Table 1. Optimization of benzylamine *N*-sulfamation. [a] formation of the benzylammonium salt (**2b**) observed; [b] miscellaneous side products observed. Reaction conversions reported as measured by ¹H NMR spectroscopy. Selected isolated yields are reported in parentheses.

Survey of reaction temperature (**Table 1**, entries 1-5) demonstrated that 30 °C was optimal with 2.0 equivalents of TBSAB. Varying the equivalents of TBSAB (**Table 1**, entry 3 *vs* 6-9) showed that 2.0 equivalents afforded high conversions. Unexpectedly, it was observed that the conversion for entry 6, with 0.5 eq. TBSAB, was quantitative. It was revealed that the tributylammonium counterion can exchange with unreacted benzylamine *in situ* to afford **2b**. We identified that longer reaction times gave rise to increased conversion to **2b**, with shorter reaction times, generally <30 min (depending on

substrate), avoiding this and obtaining **2a** exclusively. Although **2b** is an impurity, exchange to the sodium salt (**3**) delivers the identical product. The structures of **2a** and **2b** were confirmed by single crystal X-ray crystallography.[32]

We next considered a comparison of entry 3 with commercial sulfating agents Me₃N-SO₃ and Py-SO₃ (**Table 1**, entries 10 and 11, respectively). The opposite reactivity trend to *O*-sulfation[15] was observed with Me₃N-SO₃, affording a near-quantitative conversion (>99%) compared with Py-SO₃ (47%). Due to decreased lipophilicity of the trimethylammonium counterion (*c.f.* tributylammonium) an isolated yield of only 23% was obtained in the control experiments. These control experiments highlighted the challenging purification cascades to obtain analytically pure *N*-sulfated molecules with known strategies.

No deleterious effect of moisture was found (**Table 1**, entry 12). The use of additional base *e.g.* tributylamine was unnecessary (**Table 1**, entry 13). The effect of hydrochloride salt additives, at 10-50 mol% (**Table 1**, entries 14 and 15), showed none to modest (21%) decreases in conversion, respectively. In summary, robust optimised reaction conditions were identified using 2.0 equivalents of TBSAB at 30 °C afforded quantitative conversion and 98% isolated yield of **3** and in less than 30 min. Importantly, the formation of imidosulfates were not observed under these conditions.[33]
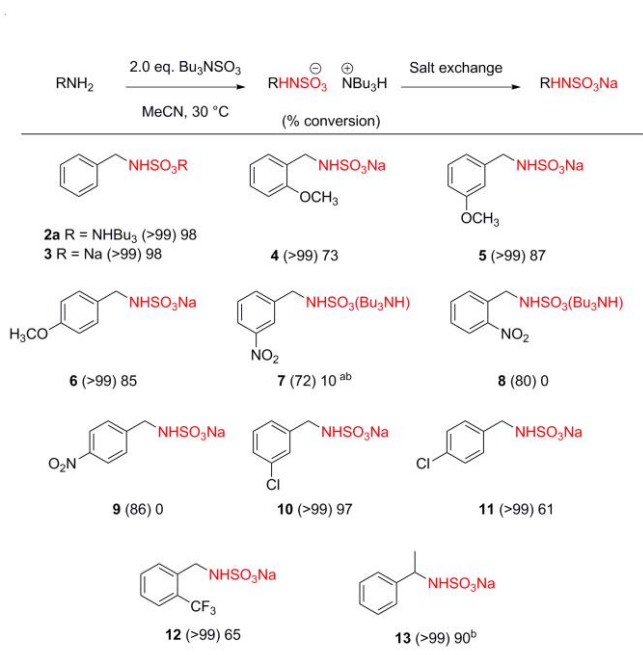
Chart 1. Reaction scope on benzylamines. [a] t = 60 min; [b] isolated as NBu₃H salt. (Parentheses indicate reaction conversion as measured by ¹H NMR spectroscopy.)

We then surveyed the generality of the developed conditions on a variety of benzylamines with varying stereoelectronic parameters (**Chart 1**). The parent benzylamine could be isolated as either the sodium (**3**) or tributylammonium salt (**2a**), in excellent isolated yields (both 98%). The incorporation of a strongly electron donating methoxy group (**4-6**) was tolerated in all positions (quantitative conversions) but a slight reduction in isolated yield for the *ortho* substituent (**4**) which is likely due to the isolation method (recrystallisation) rather than the effect of *intra*molecular hydrogen bonding. The incorporation of a strongly electron withdrawing nitro group (**7-9**) was tolerated, with good conversions in all cases (72-86%). Isolation as either the amine or sodium salt was complicated by the presence of the nitro group. However, the *meta* position (**7**) was isolated as the tributylammonium salt (10%). Next, we investigated the effect of medicinal chemistry relevant functionality, *e.g.* chloro (**10** and **11**) and trifluoromethyl (**12**) groups which proceeded in quantitative conversion and between 61-97% isolated yields. Finally, it was considered whether a steric block on the α-benzylic position (**13**) would impede the reaction. The reaction proceeded in high conversion (>99%) and excellent isolated yield (90%).
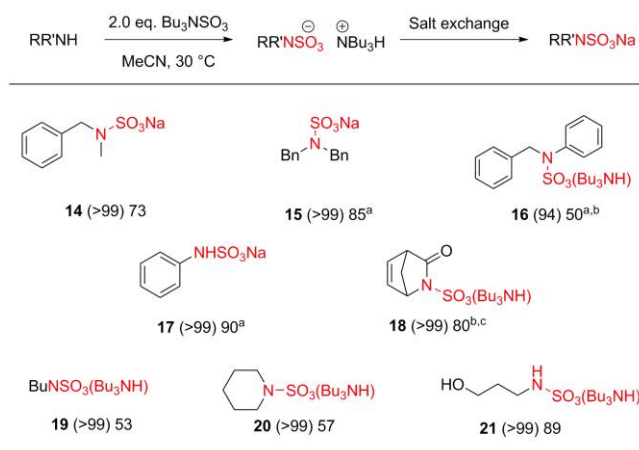

Chart 2. Reaction scope on diverse amines, amides and aminoalcohols. [a] t = 60 min; [b] isolated as NBu₃H salt; [c] t = 24 h. (Parentheses indicate reaction conversion as measured by ¹H NMR spectroscopy.)

N-sulfamation of secondary benzylamines with an *N*-methyl (**14**) or *N*-benzyl (**15**) group was well tolerated (**Chart 2**). The incorporation of a *N*-phenyl (**16**) group, which changed the amine from an sp³ to an sp² centre led to a noticeable drop in isolated yield and increased instability of the product. To investigate the *N*(sp²) effect the sulfamation of aniline was attempted with an extended reaction time. This led to an excellent yield of anti-pyretic **17**.[34] Finally, the sp² amide nitrogen reactivity of an *F*sp³-rich scaffold (vince lactam) was investigated; a modest conversion to the *N*-sulfamate **18** was observed after 30 min but improved upon extended reaction time to afford the desired compound in an excellent 80% yield.

The methodology is also readily applicable to alkyl amines (linear **19**, 53% and cyclic **20**, 57%) and amino alcohols (**21**, 89%) with exclusive selectivity for *N*-sulfation over oxygen. A control experiment was performed to determine the stability of this dual heteroatom containing molecule (**21**) for possible *N* to *O* sulfate transfer.[35-37] Warming the sample led to the partial conversion of the *N*-SO₃ containing **21** to the *O*-SO₃ containing molecule, presumably *via* a six-membered transition state, driven by the stability of the O-S bond over N-S.[38]
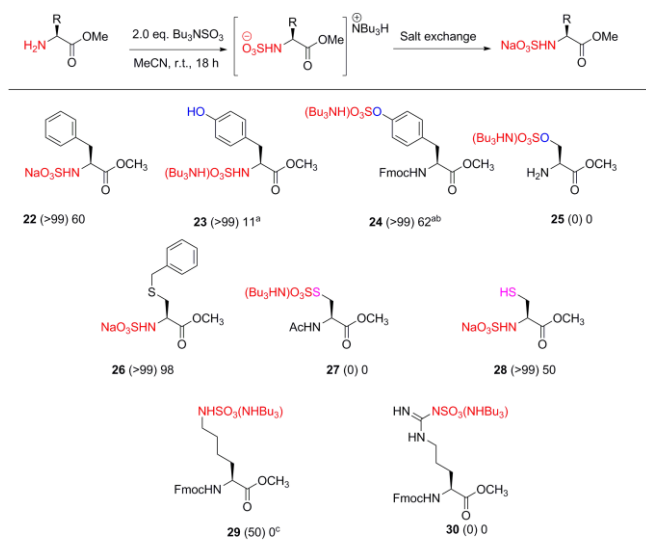
Chart 3. Canonical amino acid reactivity survey. [a] Isolated as NBu₃H salt; [b] 4.0 eq. TBSAB, reflux, 6 h; [c] An analytically pure, stable sample could not be obtained (Parentheses indicate reaction conversion as measured by ^{1}H NMR spectroscopy.)

We next sought to apply the methodology to amino acids (**Chart 3**). If the reaction is to have wide appeal, the *N*-sulfamation chemistry needs to be performed at a non-denaturing temperature. Therefore, we used low temperature and longer reaction times. In all cases the formation of a protonated amino acid counterion (*c.f.* **2b**) was not identified in the crude ^{1}H NMR spectra. We sought to explore the generality of the optimised one-pot sulfamation methodology and the potential for heteroatom selectivity (nitrogen, oxygen and sulfur). Our survey began with phenylalanine which under non-denaturing reaction conditions led to a quantitative conversion and a good isolated yield of **22** (60%) at the *N*-terminal position. The use of tyrosine raised the possibility of phenolic sulfation *versus* amine sulfamation. Under the optimised conditions a quantitative conversion of **23** was observed at the *N*-terminal position. However, a low isolated yield of the sulfamation product was obtained (11%). Importantly, no-detection of phenolic sulfation in the crude ^{1}H NMR was observed.

The ability to sulfate the phenolic oxygen of tyrosine (sY) is also possible with TBSAB but under different conditions which afforded the *N*-Fmoc protected tyrosine *O*-sulfate **24** in 62% isolated yield. **24** was found to be unstable in solution and desulfated to the starting material with time.[39]

We found that the amino acid serine led to no-detectable reaction *via* ^{1}H NMR spectroscopic studies at either the oxygen or nitrogen position (**25**, sulfation or sulfamation, respectively). The close proximity of an adjacent nucleophile shuts down the *N*-sulfamation reaction due to the steric bulk of TBSAB. Considering the result with tyrosine, the possibility of amino acid functional group selectivity is possible. We have shown a single example of serine *O*-sulfation is possible with TBSAB under markedly different reaction conditions (4.0 eq. of TBSAB, 38 °C for 52 h in DMF)[15] demonstrating the potential to dial-in different reaction outcomes with careful selection of conditions.

Next we considered the ability to obtain heteroatom selectivity in cysteine analogues, beginning with the fully protected *S*-benzyl analogue, gave exclusive *N*-terminal sulfamation in excellent conversion and isolated yield for **26** (98%). When the protecting group is transposed to the nitrogen (as an acyl amide), under mild condition no *S*-sulfurothiolation was observed in **27**, demonstrating the *N*-selectivity profile of TBSAB under these conditions. Using a molecule with an amine and thiol group, *N*-sulfamation was the exclusive product over *S*-sulfurothiolation with quantitative conversion and a modest isolated yield of **28** (50%). Next we considered nitrogen containing side chains in lysine and arginine. Lysine was partially sulfamated at the alkylamine position (**29**, 50% conversion) and arginine due to its poor guanidine *N*-nucleophilicity did not react with TBSAB (**30**) again demonstrating selectivity within the canonical amino acids. We also demonstrated that TBSAB does not racemise amino acids even under more forcing conditions.[15] Measurement of the $[\alpha]_D^{25}$ of the sulfamates (**22**-**24**, **26** and **28**) indicated that racemisation did not occur in these examples.
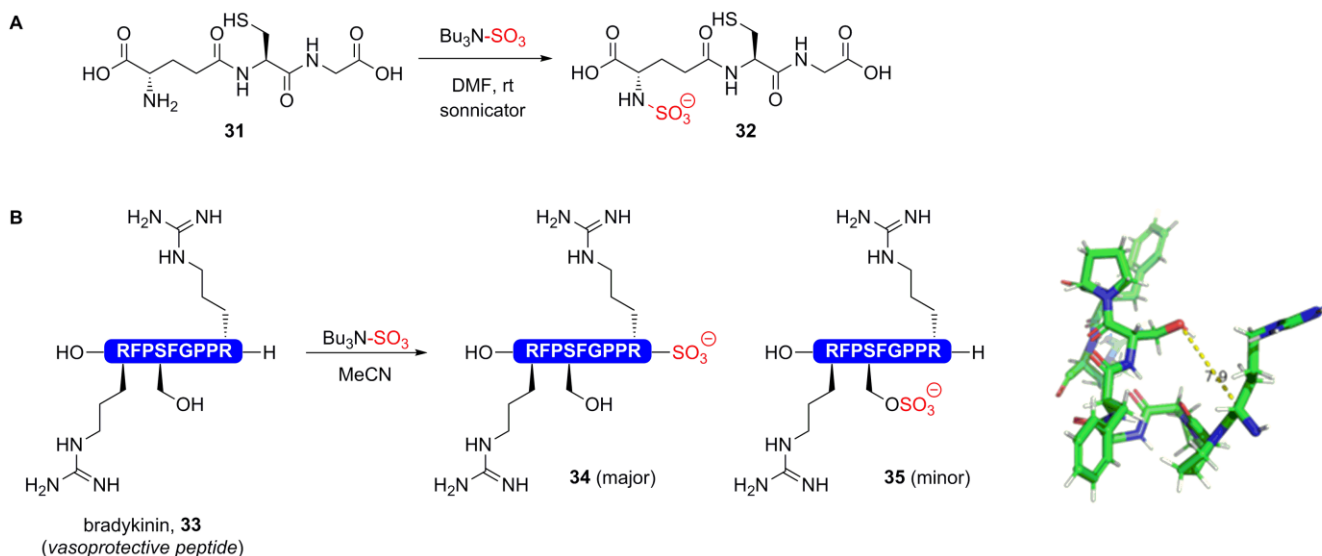
Scheme 2. *In situ* sulfation of glutathione and bradykinin. (inset. PDB: 6f3v crystal structure of Bradykinin with a 7.9 Å distance between Ser(OH) and Arg(NH$_2$).)

To test the hypothesis for site-selective sulfamation, a 3-mer (glutathione, **31**) and 9-mer (bradykinin, **33**) peptide sequence were selected (**Scheme 2**).[40] In particular, bradykinin, is a representative member of the vasoactive kinin class of cardioprotective peptides, and known to be secreted as sulfates in mammals.[41] Treatment of glutathione gave rise to a single *mono*-sulfate containing product, **32** at the *N*-terminal amine of glutamic acid. As a test of the methodology, it was envisioned that the sulfation of the 9-mer peptide, bradykinin should afford *N*-terminal sulfation (**34**). LCMS analysis revealed two *mono*-sulfated bradykinin derivatives (**34** and **35**). MS fragmentation patterns combined with the sulfation/sulfamation survey of canonical amino acids (**Chart 3**) indicated that Bradykinin contains two potential nucleophilic sites. The major product (**34**) being the expected *N*-terminal sulfamation product. The minor product (**35**) was shown to be serine *O*-sulfate. Analysis of the crystal structrure of **33** (**PDB:** 6f3v) revealed a 7.8 Å distance between the *i and i+5* residues (Arg and Ser) due to the presence of multiple proline residues in Bradykinin, this compresses the distance (approx. -20% per Pro) of the α-helix. This gives further credence to the previously reported *N* to *O*-sulfate shutttle in nature[28] and a rationalisation for the observed sS product.

Conclusion

In summary, we have developed a rapid, mild and operationally straightforward method to the currently underexplored *N*-sulfamate class of molecules, a key functional group in chemical biology. We have demonstrated wide scope and probed the heteroatom selectivity profile, showing that without protecting groups *N*-sulfamation preferentially occurs over *O*-sulfation and *S*-sulfurothiolation without racemisation of stereocentres. We have reported non-enzymatic examples of

chemically sulfating and sulfamating a peptide *in situ*. This holds promise as a method to perform site-selective sulfamation of amino acid residues. The ability to modify post-translational protein sequences opens the door to future applications in structure sensing and understanding the underlying biology of sulfated proteins.

[1] S. Gunal, R. Hardman, S. Kopriva, J. W. Mueller, *J. Biol. Chem.* **2019**, *294*, 12293-12312.
[2] C. Seibert, M. Cadene, A. Sanfiz, B. T. Chait, T. P. Sakmar, *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 11031-11036.
[3] C.-C. Wang, B.-H. Chen, L.-Y. Lu, K.-S. Hung, Y.-S. Yang, *ACS Omega*, **2018**, *3*, 11633-11642.
[4] M. J. Stone, S. Chuang, X. Hou, M. Shoham, J. Z. Zhu. *New Biotechnology*. **2009**, *25*, 299-317.
[5] J. P. Ludeman, M. J. Stone, *Brit. J. Pharmacol.*, **2014**, *171*, 1167-1179.
[6] G. Loers, Y. Liao, C. Hu, W. Xue, H. Shen, W. Zhao, M. Schachner, *Sci. Rep.*, **2019**, *9*, 1064.
[7] G. Chen, X. Chen, J. Biol. Chem. **2003**, 278, 36358-36364.
[8] A Hille, W. B. Huttner. *Eur. J. Biochem.* **1990**, *188*, 587-596.
[9] J.-C. Lee, X.-A. Lu, S. S. Kulkarni, Y.-S. Wen, S.-C. Hung, *J. Am. Chem. Soc.* **2004**, *126*, 476-477.
[10] D. P. Byrne, Y. Li, P. Ngamlert, K. Ramakrishnan, C. E. Eyers, C. Wells, D. H. Drewry, W. J. Zuercher, N. G. Berry, D. G. Fernig, P. A. Eyers. *Biochem. J.* **2018**, *475*, 2435.
[11] Y.-S. Yang, C.-C. Wang, B.-H. Chen, Y.-H. Hou, K.-S. Hung, Y.-C. Mao, *Molecules*, **2015**, *20*, 2138-2164.

[12] G. D. Balderrama, E. P. Meneses, L. H. Orihuela, O. V. Hernandez, R. C. Franco, V. P. Robles, C. V. F. Batista, *Rapid Commun. Mass Spectrom.,* **2011**, *25*, 1017-1027.

[13] T. Hunter, *Cell*, **1995**, *80*, 225-236.

[14] S. V. Campos, L. P. Miranda, M. Meldal, *J. Chem. Soc., Perkin Trans. 1*, **2002**, 682-686.

[15] D. M. Gill, L. Male, A. M. Jones, *Chem. Commun.* **2019**, *55*, 4319-4322.

[16] A. M Ali, S. D. Taylor, *Angew. Chem. Int. Ed.,* **2009**, *48*, 2024-2026.

[17] K. Kitagawa, C. Aida, H. Fujiwara, T. Yagami, S. Futaki, M. Kogire, J. Ida, K. Inoue, *J. Org. Chem.*, **2001**, *66*, 1-10.

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

[19] T. Young, L. L. Kiessling, *Angew. Chem. Int. Ed.*, **2002**, *41*, 3449-3451.

[20] L. S. Simpson, T. S. Widlanski, *J. Am. Chem. Soc.* **2006**, *5*, 1605-1610.

[21] A. Y. Desoky, J. Hendel, L. Ingram, S. D. Taylor, *Tetrahedron*, **2011**, *67*, 1281-1287.

[22] D. Taleski, S. J. Butler, M. J. Stone, R. J. Payne, *Chem. Asian J.* **2011**, *6*, 1316-1320

[23] X. Liu, L. R. Malins, M. Roches, 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.

[24] Y. S. Hsieh, L. C. Wijeyewickrema, B. L. Wilkinson, R. N. Pike, R. J. Payne, *Angew. Chem. Int. Ed.* **2014**, *53*, 3947-3951.

[25] M. J. Stone, R. J. Payne, *Acc. Chem. Res.* **2015**, *48*, 2251-2261.

[26] A. Benjdia, A. Guillot, P. Ruffie, J. Leprince, O. Berteau, *Nat. Chem.* **2017**, *9*, 698-707.

[27] A. S. Woods, H.-Y. J. Wang, S. N. Jackson, *J. Proteome Res.* **2007**, *6*, 1176-1182.

[28] G. Malojclc, R. L. Owen, J. P. A. Grimshaw, M. S. Brozzo, H. Dreher-Teo, R. Glockshuber, *Proc. Natl. Acad. Sci. USA,* **2008**, *105*, 19217-19222.

[29] T. Nakahara, M. Waki, H. Uchimura, M. Hirano, J. S. Kim, T. Matsumoto, K. Nakamura, K. Ishibashi, H. Hirano, A. Shiraishi, *Anal. Biochem.* **1986**, *154*, 194-199

[30] R. A. Al-Horani, U. R. Desai, *Tetrahedron,* **2010**, *66*, 2907-2918.

[31] H. Rogers, G. Humphrey, A. Chiu, T. Pei, *Synthesis,* **2008**, 2298.

[32] CCDC 1936116 (**2a**) and CCDC 1935974 (**2b**) contain the supplementary crystallographic data for this paper.

[33] F. Kanetani, H. Yamaguchi, *Bull. Chem. Soc. Jpn.* **1974**, *47*, 2713-2716.

[34] L. F. Audrieth, M. Sveda, *J. Org. Chem.* **1944**, *9*, 89-101.

[35] See supporting information: additional data section **S5** for ^{1}H-^{1}H COSY NMR spectra of the *in situ N*- to *O*-sulfate shuttle control experiment.

[36] For a report of *N*-desulfation kinetics see: K. Nagasawa, H. Yoshidome, *Chem. Pharm. Bull.* **1969**, *17*, 1316-1323.

[37] Related examples of *N* to *O* sulfate shuttling see: (a) A. Hopkins, A. Williams, *J. Org. Chem.* **1982**, *47*, 1745-1750; (b) A. Hopkins, R. A. Day, A. Williams, *J. Am. Chem. Soc.* **1983**, *105*, 6062-6070.

[38] Y. Inoue, K. Nagasawa. *J. Org. Chem.* **1973**, *38*, 1810-1813.

[39] See supporting information: additional data section **S6** for tyrosine *O*-sulfate stability studies.

[40] See supporting information: additional data section **S7** for sulfated peptide LCMS traces

[41] X. Xi, B. Li, T. Chen, H. F. Kwok, *Toxins*, **2015**, *7*, 951-970.

COMMUNICATION

Entry for the Table of Contents (Please choose one layout)

Layout 1:

COMMUNICATION

Text for Table of Contents

((Insert TOC Graphic here))

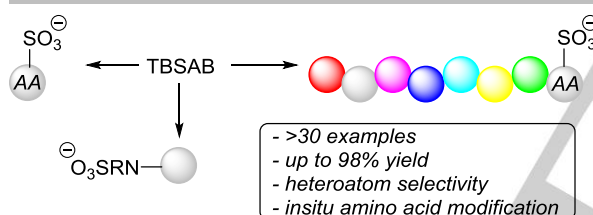*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title

Layout 2:

COMMUNICATION

- >30 examples
- up to 98% yield
- heteroatom selectivity
- insitu amino acid modification

A general sulfation/sulfamation strategy: A new method to selectively *N*-sulfamate or *O*-sulfate molecules is reported. A rapid, low temperature, non-racemizing approach to functionalise amino acids and peptides is demonstrated.

Anna Mary Benedetti,[a] Daniel M. Gill,[a] Chi W. Tsang,[b] and Alan M. Jones*[a]

Page No. – Page No.
Chemical methods for *N*- and *O*-sulfation of small molecules, amino acids and peptides