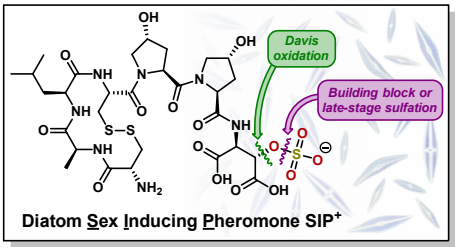
Total synthesis of a peptide diatom sex pheromone bearing an unusual sulfated aspartic acid

Andrew M. White,*[1,2] Brett D. Schwartz,[1,2] Michael G. Gardiner,[1] Lara R. Malins*[1,2]

[1]Research School of Chemistry, The Australian National University, Canberra, ACT 2601, Australia

[2]Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The Australian National University, Canberra, ACT 2601, Australia



ABSTRACT: The peptide Sex Inducing Pheromone SIP^+ (**1**) bearing an unusual sulfated aspartic acid residue induces sexual reproduction in diatom populations. Herein, we report the first total synthesis of SIP^+ using both a sulfated building block approach and a SPPS compatible late-stage sulfation strategy to assemble the natural product. The modular approaches provide concise routes to useful quantities of the natural product for future SAR investigations examining the role of SIP^+ in diatom biology.

Diatoms are a class of single celled algae that contain a characteristic silica (SiO_2) cell wall and are the second most common life form on earth after bacteria. As a critical primary producer in marine and freshwater environments, diatoms form the foundations of the ecological food chain, and their photosynthetic outputs are responsible for sequestration of approximately 20% of the world's atmospheric CO_2.[1] This equates to approximately 10–20 billion tons of CO_2 per year, and is equivalent to the carbon capture of all of the world's rainforests combined.[1] Moreover, diatoms are of critical importance to industrial processes including their use as feedstocks and water quality management in aquaculture[2] as well as utilization of their fossilized remains (diatomaceous earth, SiO_2) as filtration media, insecticides, and in healthcare.[3] The broad spectrum of ecological impact and industrial applications emphasizes the critical importance of diatoms, particularly in a climate with increasing concentrations of atmospheric CO_2[4] and a growing human population.

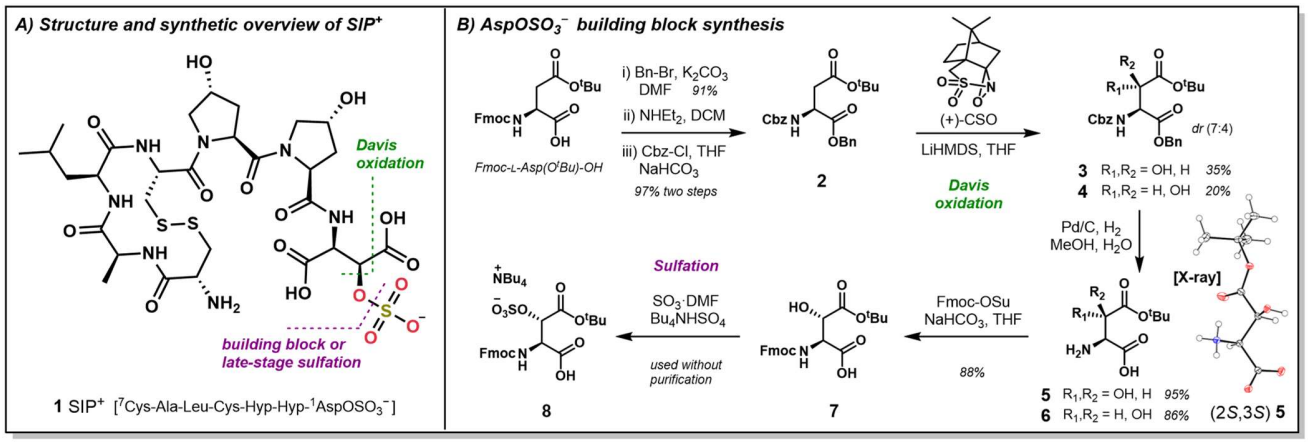
Despite such a central role in ecological and industrial processes, there is still a limited understating of the diatom reproduction cycle, which is characterized by asexual population growth with periods of sexual replication. Interestingly, the sexual reproduction cycle of diatoms is controlled by peptidic attractant pheromones that include diproline[5] and the sex inducing pheromone SIP^+ (**1**), characterized from populations of *Seminavis robusta*.[6] Specifically, SIP^+ has been shown to initiate reproduction signaling at low femtomolar concentrations and has potential to synchronize entire algal communities.[6]

The structure of SIP^+ (**1**, Scheme 1) was characterized by Pohnert and coworkers in 2023[6] and revealed a surprisingly complex peptide, bearing a macrocyclic disulfide linkage and an unusual sulfated β-hydroxylated aspartic acid residue at the C-terminus. Not only is this sulfated residue unique to diatom pheromones, but to our knowledge, there are no alternative examples of natural products bearing this sulfated aspartic acid motif from any other organisms, despite the broader prevalence of sulfopeptides, particularly those bearing tyrosine *O*-sulfation.[7] Given its structural novelty, limited isolable quantities and important ecological context, SIP^+ presents a prime target for total synthesis which we endeavored to explore.

To synthesize SIP^+, we first devised a synthetic strategy to access a solid phase peptide synthesis (SPPS) compatible β-OH-aspartic acid building block (**7**) to provide the basis for the exploration of sulfation and peptide elongation on solid support. There are a few examples of natural products containing β-OH-aspartic acid derivatives and thus strategies to access these building blocks using varied methods have been reported.[8] Our efforts focused on obtaining the desired (2*S*,3*S*) β-OH-aspartic acid, and we envisaged a regio- and diastereoselective hydroxylation at the β-position (C-3), following enolization of an appropriately protected Asp side-chain carboxylate. Work reported by Li[9] on the total synthesis of malacidin A demonstrated such a strategy using Davis' (+)-camphorylsulfonyl oxaziridine ((+)-CSO)[10] to access, diastereoselectively, Fmoc-(2*S*,3*S*) *β*-OH-aspartic acid derivatives and accordingly provided us with a robust starting point for our synthetic endeavors.

With this strategy in mind, we utilized the readily available starting material Fmoc-L-Asp(O^{t}Bu)-OH to prepare the benzyl carbamate (Cbz) and benzyl protected (Bn) building block **2**. In

Scheme 1. (A) Structure and synthetic strategy to access SIP$^+$; (B) Preparation of *β*-hydroxy aspartic acid SPPS building blocks.



contrast to the literature,[9] we selected the Cbz protection strategy as our investigations found that bulky α-amine protecting groups (e.g. Trt) limited β-alkylation of aspartic acid-derived enolates. Accordingly, enolization of building block **2** with LiHMDS followed by treatment with (+)-CSO afforded the diastereomers **3** and **4** in a ratio of 7:4, respectively. The diastereomers were separable by flash column chromatography and were individually subjected to a palladium on carbon catalyzed hydrogenolysis to afford the deprotected amino acids **5** and **6**. Crystallization, followed by single-crystal X-ray diffraction analyses determined unequivocally, their configurations to be (2*S*,3*S*)-**5** and (2*S*,3*R*)-**6** as depicted in Scheme 1.[11] Pleasingly, the configuration of the major diastereomer (2*S*,3*S*)-**5** corresponded to that assigned to the β-OH-aspartic acid residue of SIP$^+$.[6]

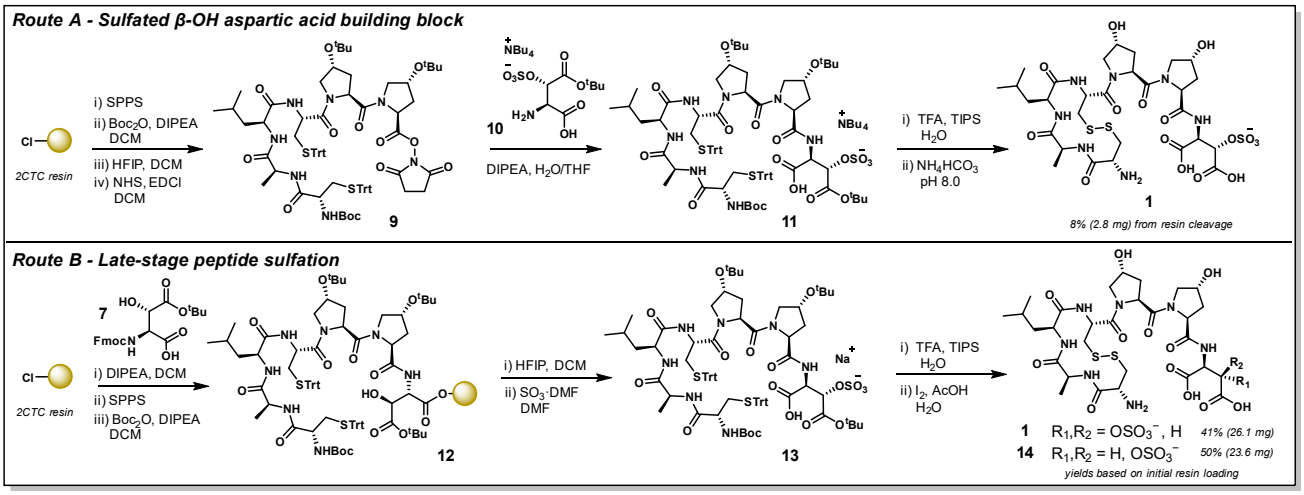
Our initial efforts to access SIP$^+$ focused on devising a sulfated building block that could be incorporated into the Fmoc-SPPS pipeline. Accordingly, compound **5** was Fmoc-protected under standard conditions to afford **7**. Reports have shown that *O*-sulfated Fmoc-amino acids (Ser or Thr) can be prepared with sulfate counterions for improved stability (e.g. tetrabutylammonium (TBA) or tributylammonium)[12] and TBA stabilized amino acids have been shown to withstand Fmoc-SPPS conditions.[12a, 13] With this in mind, we exposed **7** to the highly reactive sulfating agent SO_3·DMF, prepared freshly in house,[14] and subsequently introduced TBA in the work-up to provide **8**, primed for SPPS incorporation.

To synthesize SIP$^+$ using our sulfated building block, **8** was first loaded onto 2-chlorotrityl chloride resin (2CTC) using stoichiometric amounts to limit wastage of the custom-building block. However, initial loading attempts resulted in poor attachment to the resin (~12% loading efficiency), presumably due to the sterically crowded β-substituent and the corresponding TBA counterion. Given the poor loading efficiency and our desire to preserve the precious building block, we revised our strategy with the aim to incorporate the sulfated building block as a late-stage C-terminal modification (Scheme 2, Route A). To achieve this, we first synthesized a C-terminally truncated, protected SIP$^+$ precursor on resin, bearing a Boc-protected N-terminus. Following HFIP cleavage and purification, the C-terminal acid was activated with *N*-hydroxysuccinimide to yield **9** and the Fmoc deprotected AspOSO$_3^-$ building block **10** (see SI for preparation) was directly added to the crude peptide mixture to obtain the full-length linear peptide **11**. Pleasingly, this peptide was stable to global deprotection under acidic conditions (TFA) and the final disulfide oxidation with ammonium bicarbonate yielded the natural product **1**, albeit in modest yield (8% from resin cleavage).

Despite successfully establishing a route to SIP$^+$, the low yields associated with route A prompted exploration of a second strategy incorporating sulfation at a late-stage in the synthesis, after cleavage of the peptide from the resin (Scheme 2, Route B). Prior work has demonstrated that resin-bound protected peptides bearing free alcohols (Ser and Thr) can be sulfated,[13] suggesting that a selective peptide sulfation strategy could be a viable approach to access SIP$^+$. To this end, the unsulfated Fmoc-β-OH-Asp building block **7** was loaded onto 2CTC resin using stoichiometric amounts to conserve the custom-building block. Unlike the sulfated variant **8**, the loading of **7** proceeded with 64% incorporation onto the resin (see SI for further details). We then proceeded with standard SPPS conditions to obtain the full-length protected peptide **12** that was subject to HFIP-mediated resin cleavage to retain the orthogonal protecting groups. The crude peptide was then sulfated with SO_3·DMF[14] to obtain **13** with excellent conversion as judged by UPLC-MS analysis. In this instance the TBA counterion was not employed following sulfation, rather retaining the peptide as the sodium sulfate salt, where the sodiated ion was introduced through work up of the sulfation reaction with sodium hydrogen carbonate. In our hands, the sodiated form was found to be stable to TFA-mediated side-chain deprotection, as well as disulfide formation, without any notable degradation of the sulfate moiety. This strategy provided SIP$^+$ (**1**) in excellent yield via an efficient, stepwise pathway (41% yield based on initial resin loading, 26 mg).

To further exemplify the synthetic efficiency of the late-stage sulfation strategy and to support the structural assignment of SIP$^+$,[6] we decided to synthesize a peptide diastereomer capitalizing on the availability of the (2*S*,3*R*)-β-OH-Asp building block **6**. Following Fmoc protection, the building block was incorporated into a peptide synthesis workflow analogous to Route B (Scheme 2, see SI for full synthetic details). The strategy successfully yielded SIP$^+$ epimer **14** in excellent yield (24 mg and 50% yield based on resin loading). Considering the challenges associated with characterization of the natural product on very small scale,[6] access to **14** provided another tool to

Scheme 2. Synthesis of SIP$^+$ established by two synthetic routes A and B.



support the original structural assignment. Moreover, peptide **14** provides an opportunity to begin to establish structure activity relationships around the peculiar sulfated residue.

With significant quantities of synthetic SIP$^+$ (**1**) in hand, a comprehensive set of NMR experiments was performed in D_2O to confirm alignment with the data reported for the natural product.[6] Well resolved sets of signals were observed in both the ^{1}H and ^{13}C NMR spectra (see SI for spectra), however upon chemical shift assignment, it was evident that there was a significant disparity in the observed NMR shifts relative to those reported for naturally occurring **1**. Closer inspection of the chemical shift differences (Table S2) showed that the ^{7}Cys and ^{1}AspOSO$_3^-$ resonances differed considerably to the reported data set, while the remaining resonances closely matched the reported natural product. The mismatch of NMR data led us to consider the possibility of an epimerization event during chemical synthesis or a misassignment in the original characterization. To further explore these ideas, we assigned the NMR chemical shifts of peptide diastereomer **14**, which also showed significant differences to the natural product data (Table S4), including some minor changes around the ^{1}AspOSO$_3^-$ residue.

We noted that the major chemical shift differences between the natural and synthetic data sets were in the ionizable residues ^{7}Cys and ^{1}AspOSO$_3^-$ of **1**, so we shifted our focus to the protonation state of the peptide. Studies have shown that the protonation of ionizable residues in random coil peptides can cause significant differences in chemical shifts, varying up to 0.6 ppm and 8.5 ppm in the ^{1}H and ^{13}C spectra, respectively.[15] As the reported data did not specify a pH for NMR data acquisition, we decided to conduct an NMR titration to explore this variable for SIP$^+$. The initial pH of **1** in D_2O, was found to be 2.97, where presumably all ionizable sites would be expected to be protonated, with the exception of the sulfate group (predicted pK_a <−1).[16] The pH of the NMR sample was slowly adjusted using DCl and/or NaOD and NMR spectra were acquired at intervals between pH 2 – 11 (Figure 1). Interestingly, the signals observed in the ^{1}H spectrum which corresponded to the β-oxymethine and α-hydrogen of the ^{1}AspOSO$_3^-$ residue showed significant changes in their chemical shifts between pH 2 – 5. Additionally, α- and β-hydrogens corresponding to the ^{7}Cys residue, moved considerably between pH 5 – 10. At a pH >9 the ^{1}H NMR presented sharp, well-defined signals, and thus a second set of 2D experiments was acquired for both **1** and diastereomer **14** at this basic pH (see SI).

Pleasingly, assignment of the ^{1}H and ^{13}C chemical shifts of **1** and **14** at pH ~9.5 yielded data that closely resembled the reported natural product (Table 1, S3, S4, S7 and S8) and the substantial differences in NMR data observed at pH 2 were no longer apparent. The only appreciable chemical shift differences were observed at ^{1}AspOSO$_3^-$ for **14**, where there was divergence at the site of stereochemical inversion. Specifically, the β-Asp oxymethine signal of **14** was shifted 0.15 ppm upfield

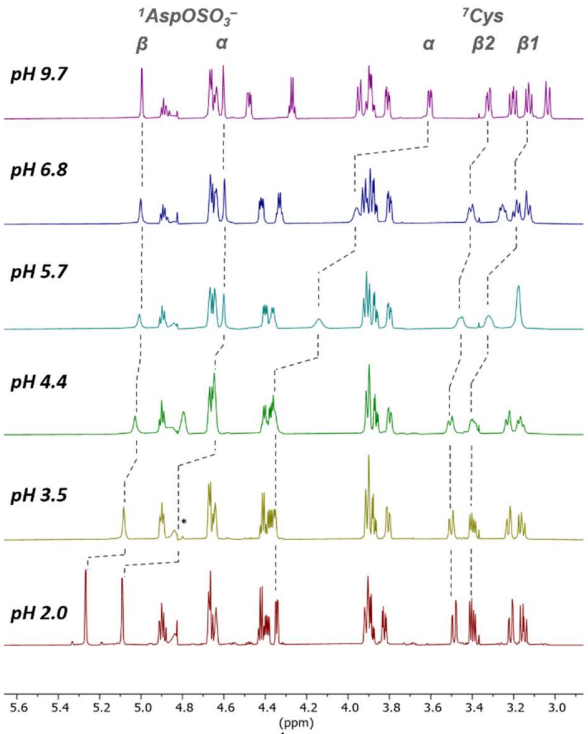


Figure 1. A selected region of ^{1}H NMR spectra (3.0 – 5.6 ppm, 800 MHz) of SIP$^+$ (**1**) at varying pH 2.0 – 9.7. Tracked in grey are the ^{1}H signals of the ionizable residues of SIP$^+$ (^{1}AspOSO$_3^-$ and ^{7}Cys), which demonstrate the chemical shift variances observed at different pH levels. *At pH 3.5, the α-^{1}AspOSO$_3^-$ signal is obscured by the water suppression within the ^{1}H NMR pulse sequence.

Table. 1. ^{1}H NMR chemical shift comparison of synthetic SIP$^+$ and natural SIP$^+$

		^{1}H Chemical Shifts δH [ppm]		
Residue	Position	Natural 1[a]	Synthetic 1 (pH 9.86)[b]	Difference
^{1}AspOSO$_3^-$	α (CH)	4.6	4.60	0.00
	β (CH)	5.00	4.99	-0.01
^{2}Hyp	α (CH)	4.65	4.65	0.00
	β1 (CH)	2.4	2.42	0.02
	β2 (CH)	2.2	2.19	-0.01
	γ (CH)	4.64	4.63	-0.01
	δ1 (CH)	3.9	3.89	-0.01
	δ2 (CH)	3.8	3.89	0.09
^{3}Hyp	α (CH)	4.9	4.89	-0.01
	β1 (CH)	2.4	2.43	0.03
	β2 (CH)	2.1	2.12	0.02
	γ (CH$_3$)	4.64	4.66	0.02
	δ1 (CH)	3.9	3.94	0.04
	δ2 (CH)	3.9	3.81	-0.09
^{4}Cys	α (CH)	4.87	4.87	0.00
	β1 (CH)	3.17	3.20	0.03
	β2 (CH)	3.07	3.04	-0.03
^{5}Leu	α (CH)	4.45	4.47	0.02
	β (CH$_2$)	1.7	1.72	0.02
	γ (CH)	1.6	1.61	0.01
	δ1 (CH$_3$)	0.95	0.94	-0.01
	δ2 (CH$_3$)	0.9	0.89	-0.01
^{6}Ala	α (CH)	4.3	4.27	-0.03
	β (CH$_3$)	1.45	1.45	0.00
^{7}Cys	α (CH)	3.7	3.61	-0.09
	β1 (CH)	3.35	3.32	-0.03
	β2 (CH)	3.15	3.13	-0.02

[a]Chemical shifts of natural SIP$^+$ as reported in reference,[6] data acquired on 600 MHz spectrometer in D_2O, at 297 K. [b]Data acquired on a 800 MHz spectrometer in D_2O, sample pH 9.86 at 298 K and the ^{6}Ala CH$_3$ was referenced to δH 1.45 ppm and δC 18.8 ppm for alignment with the reported data.

compared to **1**, suggesting that the (2*S*,3*S*) ^{1}AspOSO$_3^-$ configuration of **1** is the correct configuration as reported in the initial characterization of SIP$^+$. Synthetic **1** also displayed comprehensive agreement across all other ^{1}H NMR chemical shifts (Table 1). The ^{13}C chemical shifts were in accord with those of the natural isolate, save for minor variations (< 1.1 ppm), likely due to differences in the precise pH of the acquired data sets. The broad agreement in chemical shift comparisons confirms that synthetic **1** is consistent with the assigned structure of SIP$^+$.

In conclusion, this work has established the first synthetic strategy to access the diatom sex inducing pheromone SIP$^+$. A Davis oxidation in combination with two distinct SO$_3$·DMF-mediated sulfation strategies—a building block approach and a late-stage peptide modification strategy—provided efficient access to the unusual AspOSO$_3^-$ residue. The modular synthetic strategy is a salient feature of this approach. Analogues can conceivably be prepared for SAR studies, and the approach can be readily adapted to access new members of this natural product class as they arise. Importantly, this work provides appreciable quantities of SIP$^+$, not readily obtained from its natural source, and in doing so will advance the understanding of the sexual signaling pheromone and its role in diatom biology.

AUTHOR INFORMATION

Corresponding Authors

*Andrew M. White – andrew.white@anu.edu.au
*Lara R. Malins – lara.malins@anu.edu.au

Author Contributions

A.M.W and L.R.M conceptualized the investigation. A.M.W performed synthesis, NMR and wrote the manuscript. B.D.S performed synthesis and characterization. M. G. G performed X-ray analyses and associated data processing. All authors contributed to manuscript preparation and have approved the final version of the manuscript.

ACKNOWLEDGMENT

We would like to acknowledge the Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (CIPPS) (CE200100012), a CIPPS Fresh Ideas Grant (A.M.W) and the Westpac Research Fellowship (L.R.M.) for financial support. We also acknowledge Dr Doug Lawes for technical assistance with NMR spectroscopy. This research was undertaken with the assistance of the MX1 beamline[17] at the Australian Synchrotron, part of ANSTO.

REFERENCES

(1) (a) Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., Bittner, L., Dugdale, R., Finkel, Z., Iudicone, D., Jahn, O., Guidi, L., Lasbleiz, M., Leblanc, K., Levy, M., Pondaven, P., Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.,* **2018**, 11, 27-37. (b) Pierella Karlusich, J. J., Bowler, C., Biswas, H., Carbon dioxide concentration mechanisms in natural populations of marine diatoms: Insights from Tara Oceans. *Front. Plant Sci.,* **2021**, 12.
(2) Bhattacharjya, R., Marella, T. K., Kumar, M., Kumar, V., Tiwari, A., Diatom-assisted aquaculture: Paving the way towards sustainable economy. *Rev. Aquac.,* **2024**, 16, 491-507.
(3) (a) Maher, S., Kumeria, T., Aw, M. S., Losic, D., Diatom silica for biomedical applications: Recent progress and advances. *Adv. Healthcare Mater.,* **2018**, 7, 1800552. (b) Zeni, V., Baliota, G. V., Benelli, G., Canale, A., Athanassiou, C. G., Diatomaceous earth for arthropod pest control: Back to the future. *Molecules,* **2021**, 26, 7487.
(4) Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., Johnson, C. A., *Climate change 2001: the scientific basis*, Vol. 881, No. 9, Cambridge: Cambridge university press, **2001**
(5) Gillard, J., Frenkel, J., Devos, V., Sabbe, K., Paul, C., Rempt, M., Inzé, D., Pohnert, G., Vuylsteke, M., Vyverman, W., Metabolomics enables the structure elucidation of a diatom sex pheromone. *Angew. Chem. Int. Ed.,* **2013**, 52, 854-857.
(6) Klapper, F. A., Kiel, C., Bellstedt, P., Vyverman, W., Pohnert, G., Structure elucidation of the first sex-inducing pheromone of a diatom. *Angew. Chem. Int. Ed.,* **2023**, 62, e202307165.
(7) (a) Yang, Y.-S., Wang, C.-C., Chen, B.-H., Hou, Y.-H., Hung, K.-S., Mao, Y.-C., Tyrosine sulfation as a protein post-translational modification. *Molecules,* **2015**, 20, 2138-2164. (b) Stone, M. J., Payne, R. J., Homogeneous sulfopeptides and sulfoproteins: Synthetic approaches and applications to characterize the effects of tyrosine sulfation on biochemical function. *Acc. Chem. Res.,* **2015**, 48, 2251-2261.
(8) (a) Fernandez-Megia, E., Paz, M. M., Sardina, F. J., On the stereoselectivity of the reaction of N-(9-phenylfluoren-9-yl)aspartate enolates with electrophiles. Synthesis of enantiomerically pure 3-hydroxy-, 3-amino-, and 3-hydroxy-3-methylaspartates. *J. Org. Chem.,* **1994**, 59, 7643-7652. (b) Esgulian, M., Buchotte, M., Guillot, R., Deloisy, S., Aitken, D. J., Reversal of diastereoselectivity in a masked acyl cyanide (MAC) reaction: Synthesis of protected erythro-β-hydroxyaspartate derivatives. *Org. Lett.,* **2019**, 21, 2378-2382. (c) France, B., Bruno, V., Nicolas, I., Synthesis of a protected derivative of (2R,3R)-β-hydroxyaspartic acid suitable for Fmoc-based solid phase synthesis. *Tetrahedron Lett.,* **2013**, 54, 158-161. (d) Moreira, R., Taylor, S. D., Asymmetric synthesis of Fmoc-protected β-hydroxy and β-methoxy amino acids via a Sharpless aminohydroxylation reaction using FmocNHCl. *Org. Lett.,* **2018**, 20, 7717-7720. (e) Liu, L., Wang, B., Bi, C., He, G., Chen, G., Efficient preparation of β-hydroxy aspartic acid and its derivatives. *Chin. Chem. Lett.,* **2018**, 29, 1113-1115.

(9) Sun, Z., Shang, Z., Forelli, N., Po, K. H. L., Chen, S., Brady, S. F., Li, X., Total synthesis of malacidin A by β-hydroxyaspartic acid ligation-mediated cyclization and absolute structure establishment. *Angew. Chem. Int. Ed.,* **2020**, 59, 19868-19872.
(10) Davis, F. A., Chen, B. C., Asymmetric hydroxylation of enolates with N-sulfonyloxaziridines. *Chem. Rev.,* **1992**, 92, 919-934.
(11) The absolute configuration of **5** was determed by X-ray diffraction, while only the relative stereochemistry was discerned for compound **6**. Nevertheless, given the known stereochemistry (*S*) of the aspartic acid, the configuration at C-3 could be deduced for **6**.
(12) (a) Campos, S. V., Miranda, L. P., Meldal, M., Preparation of novel O-sulfated amino acid building blocks with improved acid stability for Fmoc-based solid-phase peptide synthesis. *J. Chem. Soc., Perkin Trans. 1,* **2002**, 682-686. (b) Gill, D. M., Male, L., Jones, A. M., Sulfation made simple: a strategy for synthesising sulfated molecules. *Chem. Commun.,* **2019**, 55, 4319-4322.
(13) Vázquez-Campos, S., St. Hilaire, P. M., Damgaard, D., Meldal, M., GAG mimetic libraries: Sulphated peptide as heparin-like glycosaminoglycan mimics in their interaction with FGF-1. *QSAR Comb. Sci.,* **2005**, 24, 923-942.
(14) Vo, Y., Schwartz, B. D., Onagi, H., Ward, J. S., Gardiner, M. G., Banwell, M. G., Nelms, K., Malins, L. R., A rapid and mild sulfation strategy reveals conformational preferences in therapeutically relevant sulfated xylooligosaccharides. *Chem. Eur. J.,* **2021**, 27, 9830-9838.
(15) Platzer, G., Okon, M., McIntosh, L. P., pH-Dependent random coil 1H, 13C, and 15N chemical shifts of the ionizable amino acids: A guide for protein pKa measurements. *J. Biomol. NMR,* **2014**, 60, 109-129.
(16) (a) Gutherie predicted the pKa of methyl hydrogen sulfate (CH_3OSO_3H) to be -3.4: Guthrie, J. P., Hydrolysis of esters of oxy acids: pKa values for strong acids; Brønsted relationship for attack of water at methyl; free energies of hydrolysis of esters of oxy acids; and a linear relationship between free energy of hydrolysis and pKa holding over a range of 20 pK units. *Can. J. Chem.,* **1978**, 56, 2342-2354. (b) $^{1}AspOSO_3H$ has a predicted pKa of -1.74 calculated using ChemAxon MarvinSketch 16.10.24.0 pKa plugin.
(17) Cowieson, N. P., Aragao, D., Clift, M., Ericsson, D. J., Gee, C., Harrop, S. J., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Williamson, R., Caradoc-Davies, T., MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.,* **2015**, 22, 187-190.