Glycopeptide Synthesis

DOI: 10.1002/ange.201404625

Chemical Synthesis of Syndecan-3 Glycopeptides Bearing Two Heparan Sulfate Glycan Chains**

*Keisuke Yoshida, Bo Yang, Weizhun Yang, Zeren Zhang, Jicheng Zhang, and Xuefei Huang**

Abstract: *Despite the ubiquitous presence of proteoglycans in mammalian systems, methodologies to synthesize this class of glycopeptides with homogeneous glycans are not well developed. Herein, we report the first synthesis of a glycosaminoglycan family glycopeptide containing two different heparan sulfate chains, namely the extracellular domain of syndecan-3. With the large size and tremendous structural complexity of these molecules, multiple unexpected obstacles were encountered during the synthesis, including high sensitivity to base treatment and the instability of glycopeptides with two glycan chains towards catalytic hydrogenation conditions. A successful strategy was established by constructing the partially deprotected single glycan chain containing glycopeptides first, followed by union of the glycan-bearing fragments and cleavage of the ester-type protecting groups. This work lays the foundation for preparing other members of this important class of molecules.*

Glycopeptides and glycoproteins play important roles in many biological events, such as cellular proliferation, neuron development, and inflammation.[1] There are two major classes of glycopeptides/glycoproteins: *N*-linked glycans and *O*-linked glycans. *O*-glycans can be further divided into two main classes: the mucin type and the glycosaminoglycan family proteoglycans. Many innovative strategies have been designed to synthesize glycopeptides bearing *N*-glycans and mucin type *O*-glycans,[2,3] with successful preparation of molecules approaching the complexity of native glycoproteins.[4,5]

In sharp contrast to peptides bearing *N*-glycans and mucin type *O*-glycans, glycosaminoglycan glycopeptide syntheses are much less developed, despite the ubiquitous presence and many important biological functions of this class of glycans.[6] Synthesis in this area has mainly focused on the glycosaminoglycan oligosaccharides[7–23] or the tetrasaccharide linker.[24–27] Recently, we reported the first synthesis of a proteoglycan family glycopeptide; a syndecan-1 glycopeptide bearing one heparan sulfate chain.[28] An additional level of complexity of proteoglycans is that many carry more than one glycosaminoglycan chain.[6] To address this, we have begun to establish a viable route towards homogeneous glycopeptides bearing multiple heparan sulfate moieties with syndecan-3 as the target. Serious difficulties were encountered since the presence of multiple glycan chains significantly increases the structural complexity and instability of the target molecules. Herein, we report the lessons we have learned and the eventual establishment of a successful approach to access these highly complex molecules.

Our synthetic target is the syndecan-3 extracellular domain glycopeptide (**1**), which displays the typical structural features of heparan sulfate proteoglycans, including a peptide backbone, different heparan sulfate chains, the full tetrasaccharide linkers, 2-*O* sulfation, 6-*O* sulfation, glucosamine α-linked to both glucuronic acid and iduronic acid, and *N*-acetylation. In order to prepare this molecule, we adapted a cassette approach[29] in which the glucuronic acid containing octasaccharide cassette **2** and iduronic acid cassette **3**[28] were produced first and then incorporated into the glycopeptide (Figure 1).

There are multiple possible reaction sequences to connect the glycosyl units in the octasaccharide modules. After much exploration, we established a 3+2+3 strategy that involves building blocks consisting of the ABC trisaccharide, the DE disaccharide, and the FGH trisaccharide to access the octasaccharide modules **2** and **3**. To prepare the ABC trisaccharide, the glucoside donor **4** was pre-activated with *p*-TolSCl/AgOTf[30] and subsequently glycosylated disaccharide **5**[28] to generate the ABC trisaccharide **6** in 85% yield (Scheme 1). The 3+2 glycosylation between trisaccharide **6** and the DE disaccharide **5** went smoothly to produce pentasaccharide **7**. Pentasaccharide **7** reacted with the trisaccharide serine unit **8**[28] to generate the octasaccharide cassette **9** in an excellent 87% yield. The TBDPS silyl ether groups in **9** were removed by HF/pyridine to expose the three primary hydroxyls, which were oxidized to carboxylic acids[31] and subsequently converted to methyl esters to give compound **12**. The two azide groups in **12** were transformed into *N*-acetyl moieties through a one-pot reduction/acetylation procedure with zinc, copper sulfate, and acetic anhydride to afford octasaccharide **2.**

A serious challenge in the assembly of heparan sulfate glycopeptides is the compatibility of the protecting group removal conditions with the sulfated glycopeptide. Owing to the high sensitivity of sulfates to acid, commonly used acid-cleavable amino acid side chain protecting groups such as Boc and trityl are to be avoided. Furthermore, caution needs to be taken since the glycoside–serine linkage is prone to base-promoted β-elimination.[28,32] The sequence of deprotection

[*] K. Yoshida,[+] B. Yang,[+] W. Yang,[+] Z. Zhang, J. Zhang, X. Huang
Department of Chemistry, Michigan State University
578 S. Shaw Lane, East Lansing, MI 48824 (USA)
E-mail: xuefei@chemistry.msu.edu

[+] These authors contributed equally to the work.

[**] This work was supported by the National Science Foundation (CHE 1111550) and the National Institute of General Medical Sciences NIH (R01GM072667).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201404625.

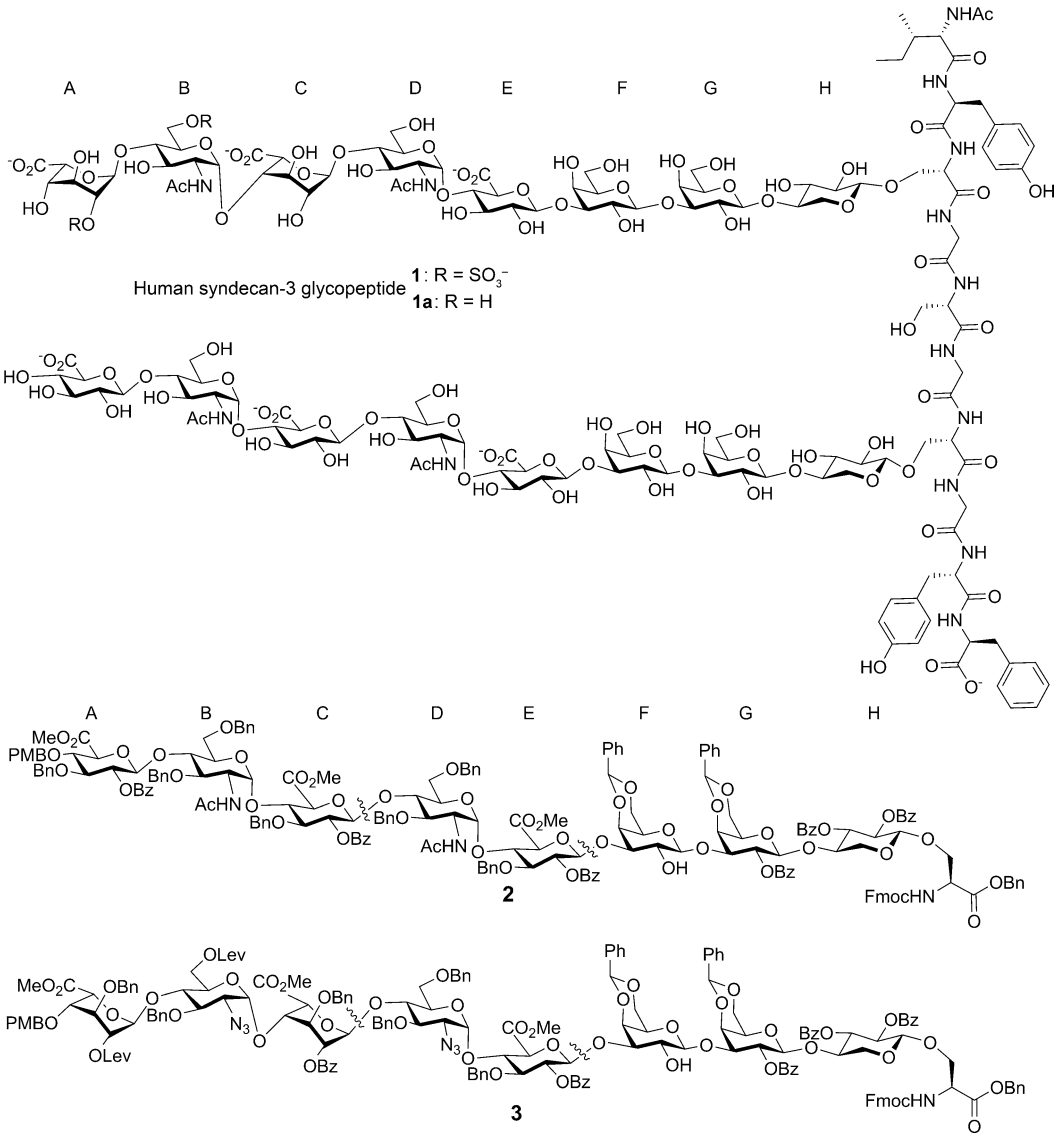


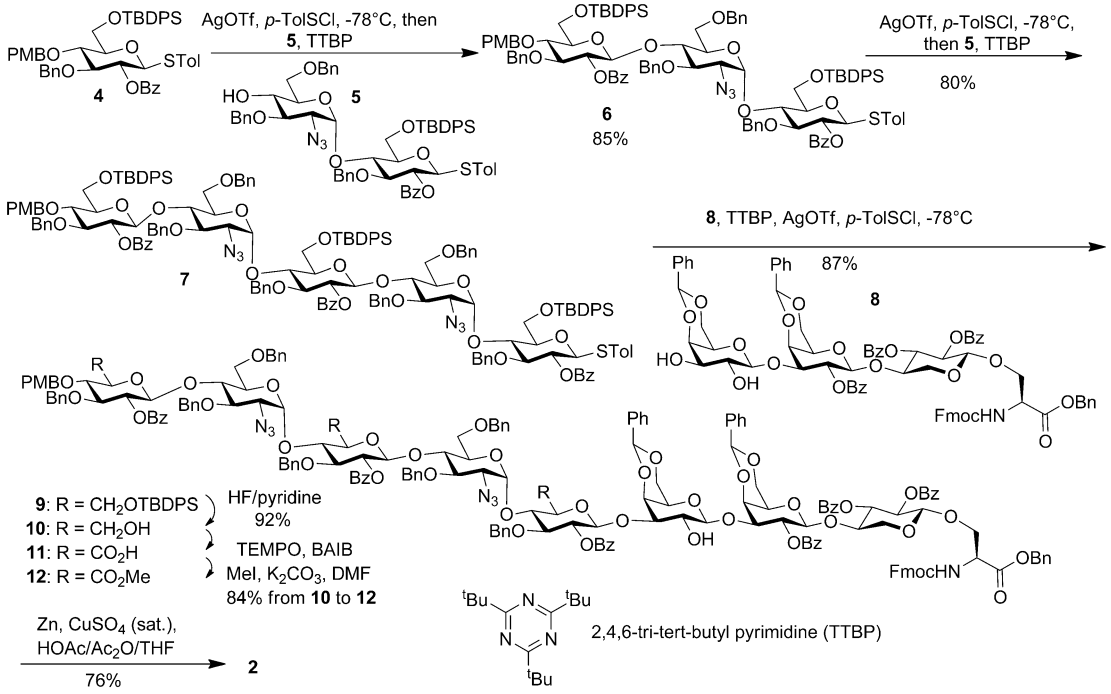
Figure 1. Syndecan-3 glycopeptide (**1**) and the octasaccharide cassettes **2** and **3**.

and reagents applied thus needs to be carefully designed and established.

Previously, we showed that the ester protecting groups (Ac, Bz) on glycopeptide **13** (Figure 2) could be successfully removed under transesterification conditions by using NaOMe.[28] The free C terminus of the glycopeptide **13** was crucial to prevent base-promoted β-elimination of the glycan chain. This route was applied to the glucoside-containing octasaccharide cassette **15**, which was produced from octasaccharide module **2** (Scheme 2). However, NaOMe treatment of **15** at room temperature led to multiple products as a result of backbone cleavage at the glucuronic acid sites, with only trace amounts of the desired product obtained. Lowering the pH value or reaction temperature led to incomplete removal of the Bz groups. The high lability of glycopeptide **15** to base treatment compared to glycopeptide **13** is possibly due to neighboring group assisted glycan cleavages[11,33] (Scheme S1 in the Supporting Information).

The failure of the previously established acyl removal strategy prompted us to examine alternatives. We envisioned that a less basic yet strong nucleophile such as hydrazine[34] could potentially remove the Ac and Bz groups without damaging the glycopeptide linkage. To incorporate hydrazinolysis, the full-length glycopeptide **17** was designed, which would be assembled from glycopeptides **18** and **19** (Figure 3). The uronic acid groups in the glycan chains of **18** and **19** are protected as methyl esters, which can be converted into free carboxylic acids by mild base treatment, thus setting the stage for hydrazinolysis to cleave all of the acyl protective groups.

Synthesis of glycopeptide **18** started with the acetylation of **3**. Subsequent conversion of azides into acetamides, Lev group removal by hydrazine acetate, and sulfation afforded octasaccharide **20** (Scheme 3). The Fmoc group in **20** was removed and the resulting free amine was coupled to dipeptide **21** to produce glycopeptide **22** in 56% yield over two steps. Selective removal of the benzyl ester in glycopep-



Scheme 1. TBDPS = *tert*-butyldiphenylsilyl, Tf = trifluoromethanesulfonyl, Bn = benzyl, Bz = benzoyl, PMB = *para*-methoxybenzyl, Fmoc = 9-fluorenylmethyloxycarbonyl, TEMPO = 2,2,6,6-tetramethylpiperidin-1-yloxyl, BAIB = bis(acetoxy)iodobenzene, DMF = *N*,*N*-dimethylformamide, THF = tetrahydrofuran.

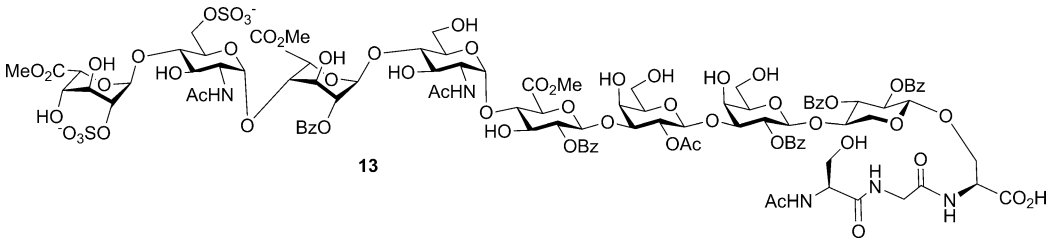
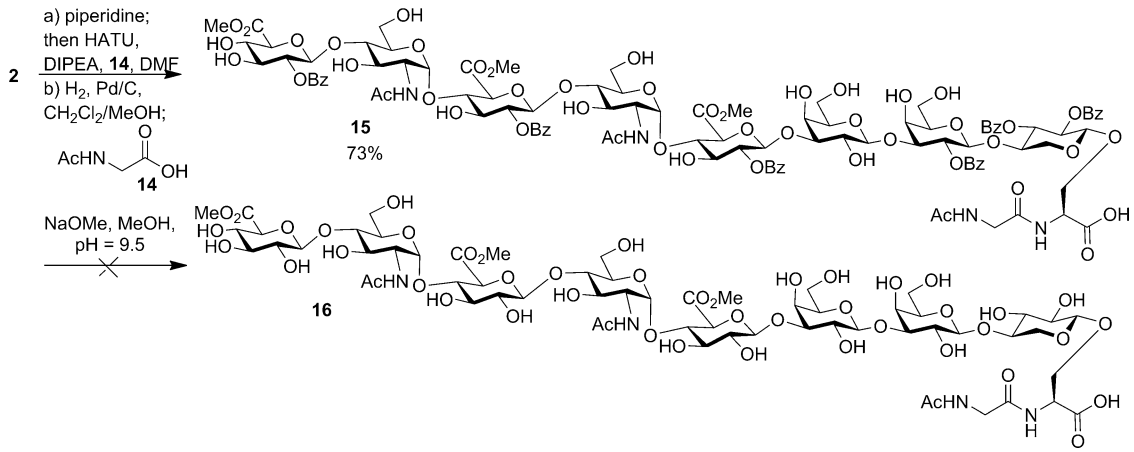


Figure 2. The structure of glycopeptide **13**.



Scheme 2. HATU = *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate, DIPEA = *N*,*N*-diisopropylethylamine.

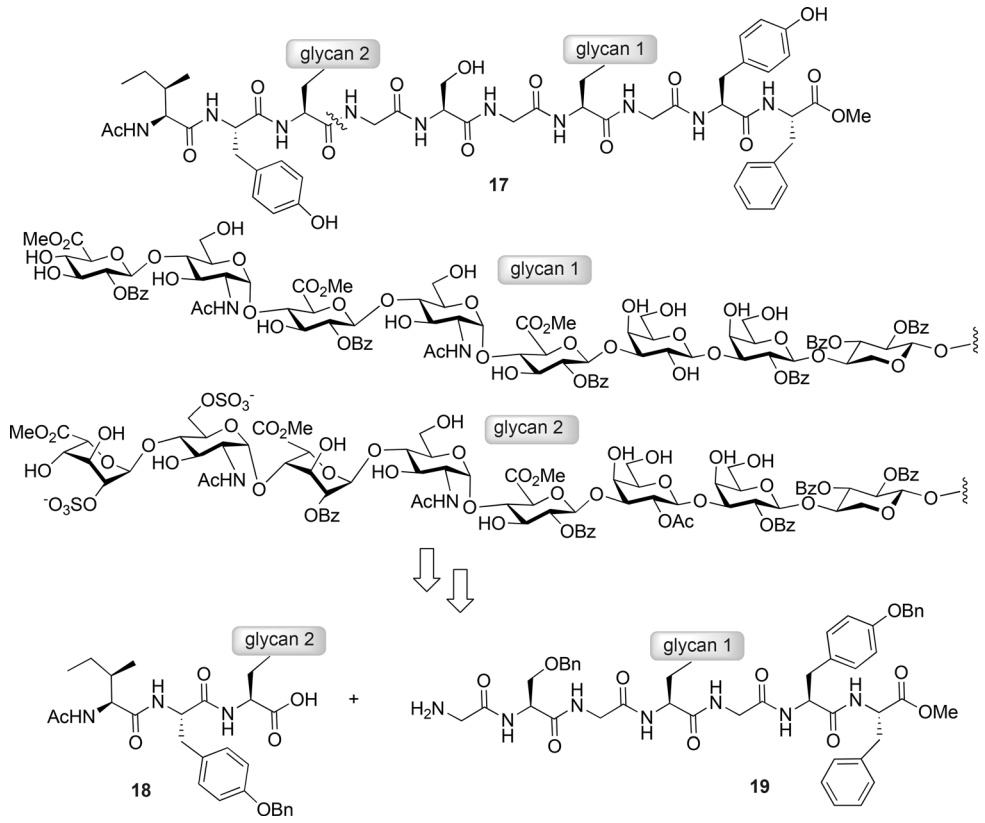
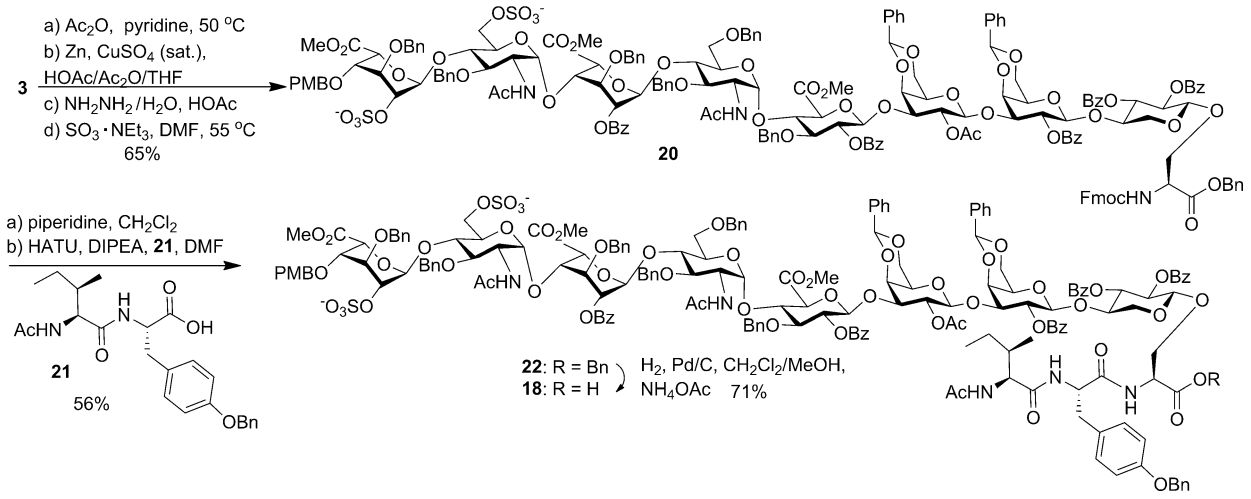


Figure 3. The full-length glycopeptide **17** is assembled from glycopeptides **18** and **19**.

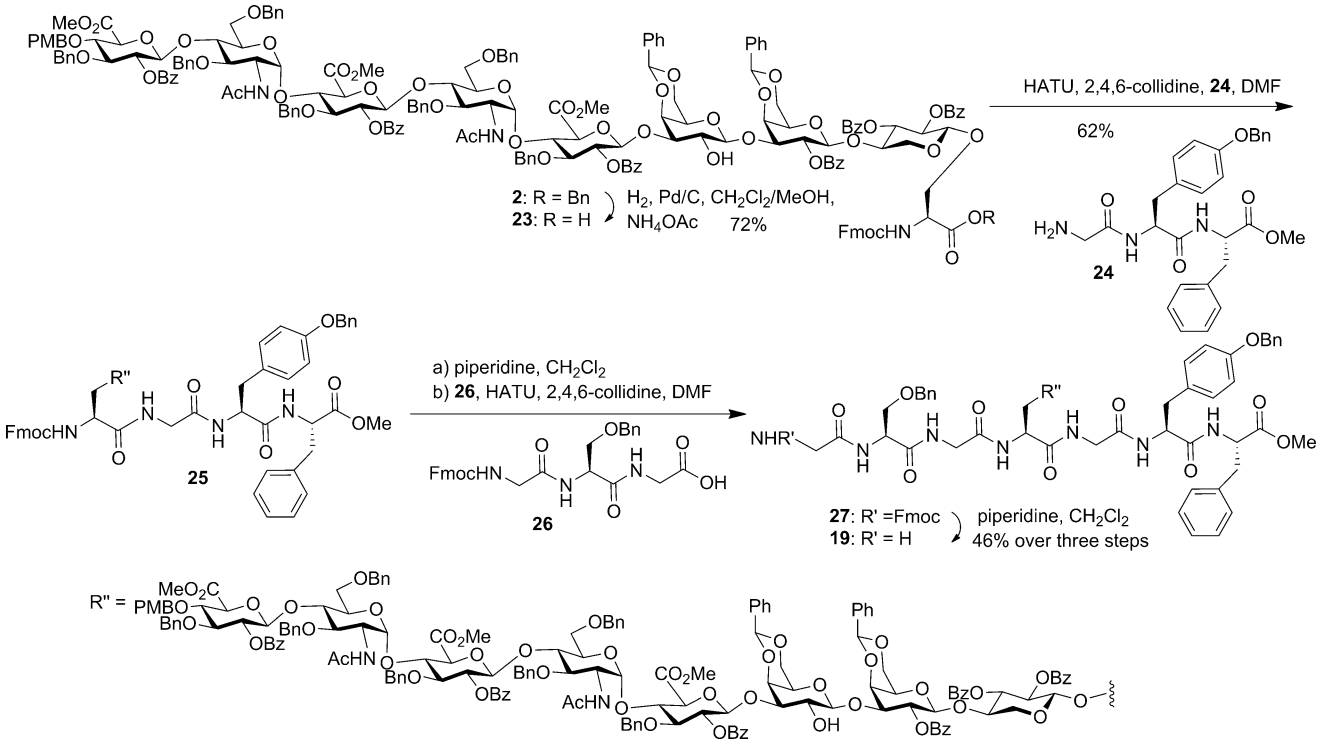


Scheme 3.

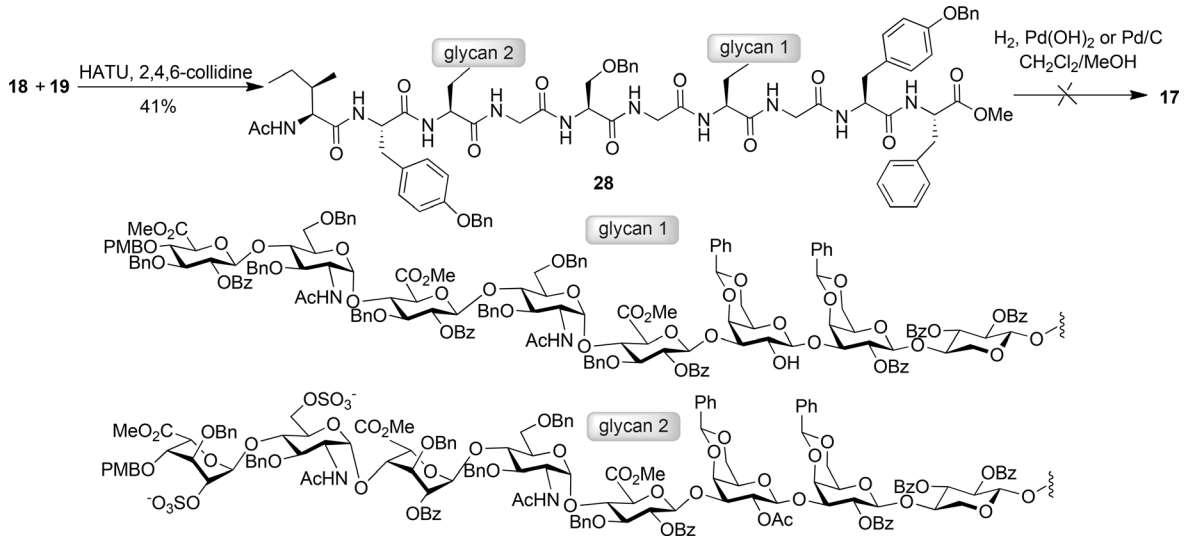
tide **22** under hydrogenation in the presence of NH_4OAc generated glycopeptide **18** with a free carboxylic acid terminus (Scheme 3).

The synthesis of glycopeptide **19** began with the hydrogenation of octasaccharide **2** in the presence of NH_4OAc to afford glycopeptide **23**, which contains a free C-terminal carboxylic acid. Coupling **23** with tripeptide **24** gave glycopeptide **25**, which was treated with piperidine and then joined to tripeptide **26**. The resulting product **27** was deprotected by piperidine to give the free amine bearing glycopeptide **19** (Scheme 4).

With glycopeptides **18** and **19** in hand, these two fragments were joined together in a reaction promoted by HATU to produce glycopeptide **28**, which bears two different glycans. To remove the benzyl ethers, PMB, and benzylidene groups, glycopeptide **28** was subjected to global hydrogenation with Pearlman's reagent or Pd/C under atmospheric pressure of hydrogen in mixed solvents of CH_2Cl_2 and methanol. How-



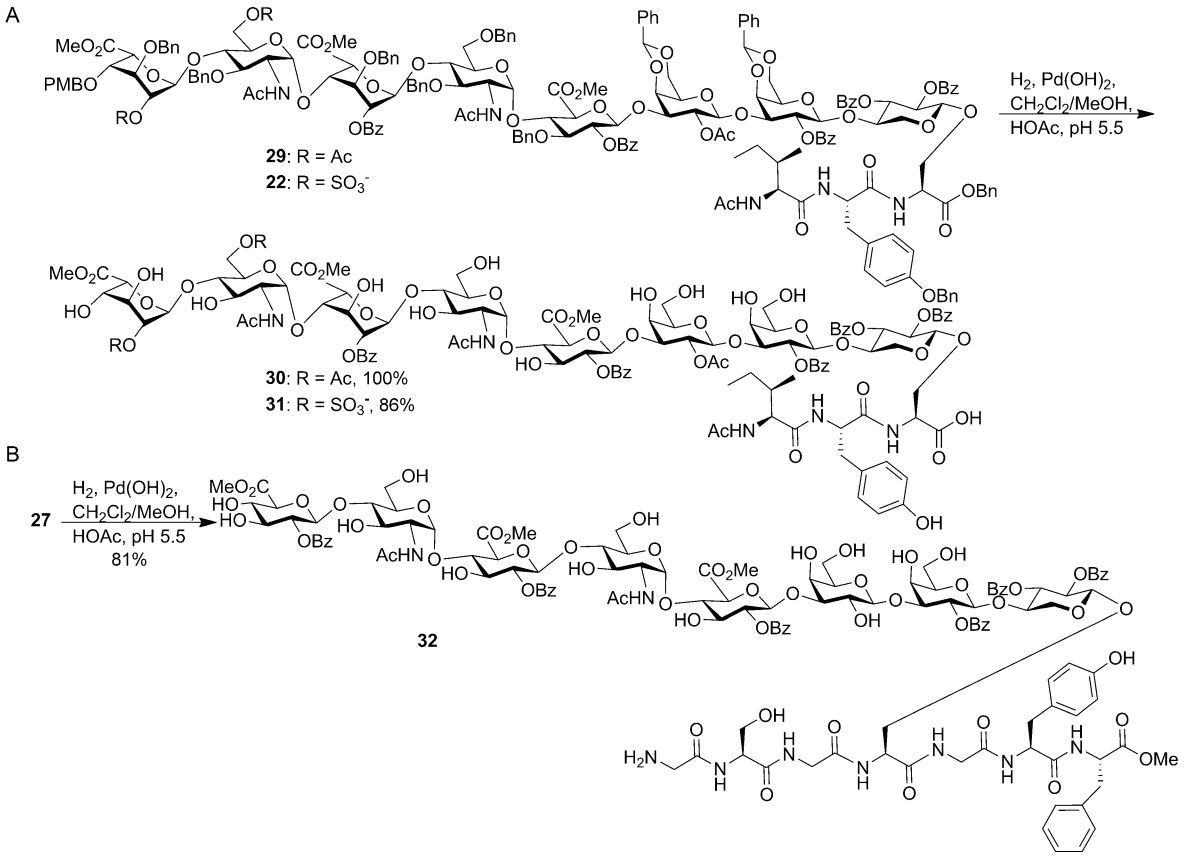
Scheme 4.



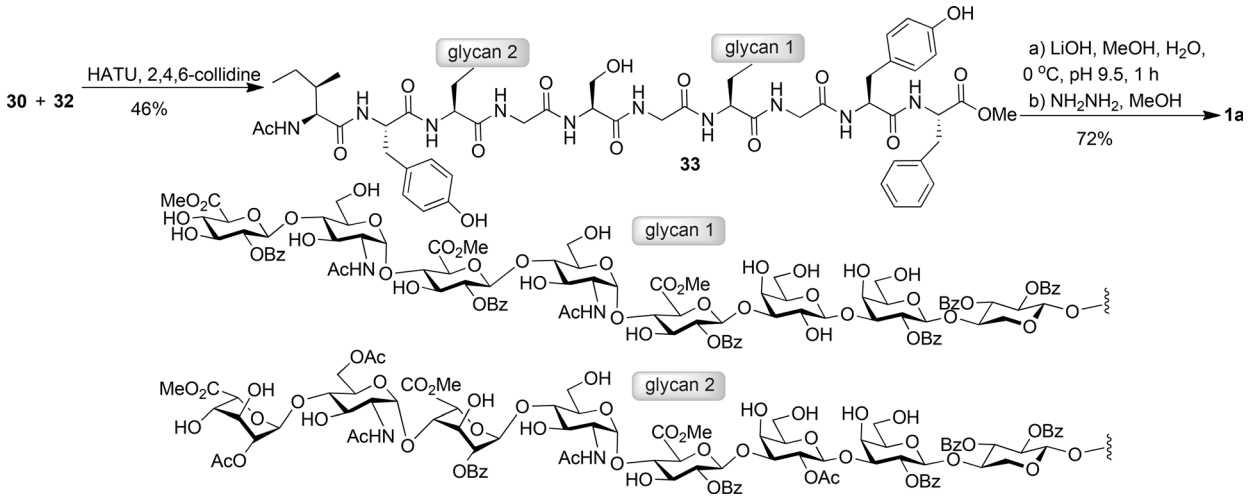
Scheme 5.

ever, this reaction failed to yield any of the desired product **17** (Scheme 5). Given that it is possible that the partially deprotected glycopeptides formed during the hydrogenation of **28** could undergo aggregation, thereby preventing access to the palladium catalyst, colloidal Pd nanoparticles were tested for the hydrogenation of **28** since these nanoparticles could remove benzyl ethers from solid phase bound glycans.[35] Various solvents, pH values, and reaction temperatures; elevated hydrogen pressure; and low acidity triflic acid conditions for benzyl removal[36] were also examined. However, none of these endeavors yielded the desired compounds and decomposition products were observed in mass spectrometry analysis.

Owing to the unexpected difficulty in the hydrogenation of **28**, an alternative route was envisioned in which hydrogenation was to be performed on glycopeptides **22** and **27**, which bear a single heparan sulfate chain. The iduronic acid containing nonsulfated glycopeptide **29** was tested first and was prepared in a manner analogous to the preparation of **22**. Catalytic hydrogenation of compound **29** under slightly acidic conditions (pH ≈ 5.5) with Pearlman's catalyst went smoothly and gave the desired product **30** in quantitative yield



Scheme 6.

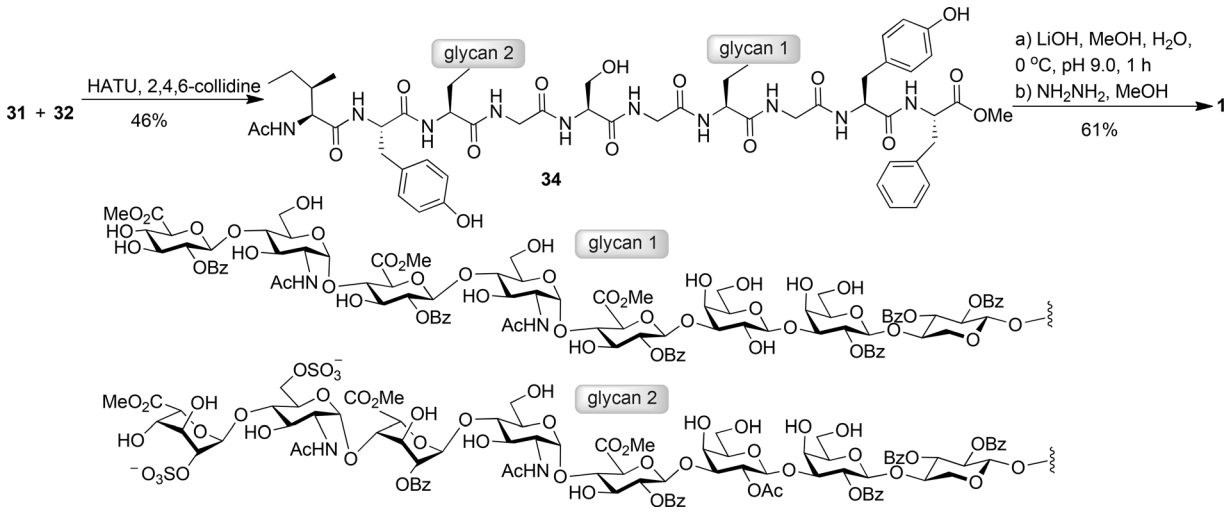


Scheme 7.

(Scheme 6A). In a similar manner, sulfated glycopeptide **22** and glycopeptide **27** were successfully hydrogenated to give **31** and **32** in quantitative yield (Scheme 6B). The reason for the increase in sensitivity of the glycopeptides bearing two heparan sulfate chains to hydrogenation is not clear.

The two partially deprotected glycopeptides **30** and **32** were joined through HATU-mediated coupling to afford glycopeptide **33** (Scheme 7). To complete the deprotection, the methyl esters in **33** were cleaved by LiOH (pH 9.5), which was followed by hydrazine treatment. The hydrazinolysis procedure successfully removed all of the acyl protecting groups, thereby affording the fully deprotected glycopeptide **1a**.

With the successful preparation of the nonsulfated glycopeptide **1a**, we moved on to the synthesis of the sulfated glycopeptide **1**. The coupling of glycopeptides **31** and **32**



Scheme 8.

afforded compound **34** (Scheme 8). The sulfated glycan chain turned out to be very sensitive to base since pH 9.5 LiOH led to partial chain cleavage. Instead, the methyl ester removal was performed at pH 9.0 with frequent monitoring by mass spectrometry, followed by hydrazinolysis to afford glycopeptide **1** in 61 % yield.

In conclusion, a successful strategy was developed for the assembly of syndecan-3 glycopeptides bearing two heparan sulfate chains. Many obstacles were encountered during the syntheses and the previous synthetic route established for syndecans-1 with one glycan chain could not be directly applied to glycopeptide **1**. The reason for this is that as the size of the molecule increases, unique reactivity and stability problems emerge. To overcome these challenges, the hydrogenation reaction was performed on glycopeptides bearing a single glycan chain followed by union of the partially deprotected fragments. The final deprotection required the cleavage of all of the ester protective groups, which was accomplished by mild base treatment followed by hydrazinolysis. The hydrazinolysis procedure was critical to ensure complete removal of the benzoyl moieties without undesired β-elimination or cleavage of the highly sensitive glycan chain. The knowledge gained in the current study will be valuable in the synthesis of other glycosaminoglycan family glycopeptides bearing multiple heparan sulfate chains.

Received: April 23, 2014
Published online: June 30, 2014

Keywords: glycopeptides · glycosylation · proteoglycans · sulfation · total synthesis

[1] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579 – 601, and references therein.
[2] C. Unverzagt, Y. Kajihara, *Chem. Soc. Rev.* **2013**, *42*, 4408 – 4420.
[3] P. Siman, A. Brik, *Org. Biomol. Chem.* **2012**, *10*, 5684 – 5697.
[4] P. Wang, S. Dong, J.-H. Shieh, E. Peguero, R. Hendrickson, M. A. S. Moore, S. J. Danishefsky, *Science* **2013**, *342*, 1357 – 1360.
[5] I. Sakamoto, K. Tezuka, K. Fukae, K. Ishii, K. Taduru, M. Maeda, M. Ouchi, K. Yoshida, Y. Nambu, J. Igarashi, N. Hayashi, T. Tsuji, Y. Kajihara, *J. Am. Chem. Soc.* **2012**, *134*, 5428 – 5431.
[6] J. R. Bishop, M. Schuksz, J. D. Esko, *Nature* **2007**, *446*, 1030 – 1037.
[7] S. U. Hansen, G. J. Miller, C. Cole, G. Rushton, E. Avizienyte, G. C. Jayson, J. M. Gardiner, *Nat. Commun.* **2013**, *4*, 2016.
[8] G. J. Sheng, Y. I. Oh, S.-K. Chang, L. C. Hsieh-Wilson, *J. Am. Chem. Soc.* **2013**, *135*, 10898 – 10901.
[9] Y. Xu, S. Masuko, M. Takeiddin, H. Xu, R. Liu, J. Jing, S. A. Mousa, R. J. Linhardt, J. Liu, *Science* **2011**, *334*, 498 – 501.
[10] Y.-P. Hu, S.-Y. Lin, C.-Y. Huang, M. M. L. Zulueta, J.-Y. Liu, W. Chang, S.-C. Hung, *Nat. Chem.* **2011**, *3*, 557 – 563.
[11] G. Tiruchinapally, Z. Yin, M. El-Dakdouki, Z. Wang, X. Huang, *Chem. Eur. J.* **2011**, *17*, 10106 – 10112.
[12] P. Czechura, N. Guedes, S. Kopitzki, N. Vazquez, M. Martin-Lomas, N.-C. Reichardt, *Chem. Commun.* **2011**, *47*, 2390 – 2392.
[13] S. Arungundram, K. Al-Mafraji, J. Asong, F. E. Leach, I. J. Amster, A. Venot, J. E. Turnbull, G. J. Boons, *J. Am. Chem. Soc.* **2009**, *131*, 17394 – 17405.
[14] F. Baleux, L. Loureiro-Morais, Y. Hersant, P. Clayette, F. Arenzana-Seisdedos, D. Bonnaffé, H. Lortat-Jacob, *Nat. Chem. Biol.* **2009**, *5*, 743 – 748.
[15] J. Chen, Y. Zhou, C. Chen, W. Xu, B. Yu, *Carbohydr. Res.* **2008**, *343*, 2853 – 2862.
[16] J. Tatai, P. Fügedi, *Tetrahedron* **2008**, *64*, 9865 – 9873.
[17] T. Polat, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 12795 – 12800.
[18] C. Noti, J. L. de Paz, L. Polito, P. H. Seeberger, *Chem. Eur. J.* **2006**, *12*, 8664 – 8686.
[19] J. D. C. Codée, B. Stubba, M. Schiattarella, H. S. Overkleeft, C. A. A. van Boeckel, J. H. van Boom, G. A. van der Marel, *J. Am. Chem. Soc.* **2005**, *127*, 3767 – 3773.
[20] J. L. de Paz, M. Martin-Lomas, *Eur. J. Org. Chem.* **2005**, 1849 – 1858.
[21] R.-H. Fan, J. Achkar, J. M. Hernandez-Torres, A. Wei, *Org. Lett.* **2005**, *7*, 5095 – 5098.
[22] L. Poletti, M. Fleischer, C. Vogel, M. Guerrini, G. Torri, L. Lay, *Eur. J. Org. Chem.* **2001**, 2727 – 2734.
[23] M. Petitou, J.-P. Herault, A. Bernat, P.-A. Driguez, P. Duchaussoy, J.-C. Lormeau, J.-M. Herbert, *Nature* **1999**, *398*, 417 – 422.
[24] K. Aït-Mohand, C. Lopin-Bon, J. C. Jacquinet, *Carbohydr. Res.* **2012**, *353*, 33 – 48, and references therein.
[25] T.-Y. Huang, M. M. L. Zulueta, S.-C. Hung, *Org. Lett.* **2011**, *13*, 1506 – 1509.

[26] J.-i. Tamura, T. Nakamura-Yamamoto, Y. Nishimura, S. Mizumoto, J. Takahashi, K. Sugahara, *Carbohydr. Res.* **2010**, *345*, 2115 – 2123.
[27] K. Shimawaki, Y. Fujisawa, F. Sato, N. Fujitani, M. Kurogochi, H. Hoshi, H. Hinou, S.-I. Nishimura, *Angew. Chem. Int. Ed.* **2007**, *46*, 3074 – 3079; *Angew. Chem.* **2007**, *119*, 3134 – 3139.
[28] B. Yang, K. Yoshida, Z. Yin, H. Dai, H. Kavunja, M. H. El-Dakdouki, S. Sungsuwan, S. B. Dulaney, X. Huang, *Angew. Chem. Int. Ed.* **2012**, *51*, 10185 – 10189; *Angew. Chem.* **2012**, *124*, 10332 – 10336.
[29] X.-T. Chen, D. Sames, S. J. Danishefsky, *J. Am. Chem. Soc.* **1998**, *120*, 7760 – 7769.
[30] X. Huang, L. Huang, H. Wang, X.-S. Ye, *Angew. Chem. Int. Ed.* **2004**, *43*, 5221 – 5224; *Angew. Chem.* **2004**, *116*, 5333 – 5336.
[31] L. J. Van den Bos, J. D. C. Codee, J. C. Van der Toorn, T. J. Boltje, J. H. Van Boom, H. S. Overkleeft, G. A. Van der Marel, *Org. Lett.* **2004**, *6*, 2165 – 2168.
[32] P. Sjölin, M. Elofsson, J. Kihlberg, *J. Org. Chem.* **1996**, *61*, 560 – 565.
[33] J. Gao, D. A. Thomas, C. H. Sohn, J. L. Beauchamp, *J. Am. Chem. Soc.* **2013**, *135*, 10684 – 10692.
[34] P. W. Glunz, S. Hintermann, L. J. Williams, J. B. Schwarz, S. D. Kuduk, V. Kudryashov, K. O. Lloyd, S. J. Danishefsky, *J. Am. Chem. Soc.* **2000**, *122*, 7273 – 7279.
[35] O. Kanie, G. Grotenbreg, C.-H. Wong, *Angew. Chem. Int. Ed.* **2000**, *39*, 4545 – 4547; *Angew. Chem.* **2000**, *112*, 4719 – 4721.
[36] K. Kawahira, H. Tanaka, A. Ueki, Y. Nakahara, H. Hojo, Y. Nakahara, *Tetrahedron* **2009**, *65*, 8143 – 8153.

