

Showcasing research from Professors Xuefei Huang and Angela Wilson's laboratories, Departments of Chemistry and Biomedical Engineering, Michigan State University, Michigan, USA as well as Professor Jian Liu's laboratory, Division of Chemical Biology and Medicinal Chemistry, University of North Carolina, North Carolina, USA.

Chemical synthesis of human syndecan-4 glycopeptide bearing O-, N-sulfation and multiple aspartic acids for probing impacts of the glycan chain and the core peptide on biological functions

Heparan sulfate proteoglycan glycopeptides bearing multiple aspartic acids in the core peptide and acid sensitive O- and N-sulfations in the glycan chain have been synthesized for the first time. The availability of such a well-defined structure enabled analysis of the respective effects of the glycan and the core peptide on functions of the glycopeptide. The glycopeptide exhibited context dependent changes in biological activities, which suggests linking the glycan to core proteins as in proteoglycans may be an additional approach to modulate biological functions of heparan sulfate in nature.

As featured in:



See Xuefei Huang *et al.*,
Chem. Sci., 2020, 11, 6393.

EDGE ARTICLE

Cite this: *Chem. Sci.*, 2020, 11, 6393

All publication charges for this article have been paid for by the Royal Society of Chemistry

Chemical synthesis of human syndecan-4 glycopeptide bearing O-, N-sulfation and multiple aspartic acids for probing impacts of the glycan chain and the core peptide on biological functions†

Weizhun Yang,^a Yigitcan Eken,^a Jicheng Zhang,^a Logan Emerson Cole,^a Sherif Ramadan,^{ab} Yongmei Xu,^c Zeren Zhang,^a Jian Liu,^c Angela K. Wilson^a and Xuefei Huang^{id}*^{ade}

Proteoglycans are a family of complex glycoproteins with glycosaminoglycan chains such as heparan sulfate (HS) attached to the core protein backbone. Due to the high structural heterogeneity of HS in nature, it is challenging to decipher the respective roles of the HS chain and the core protein on proteoglycan functions. While the sulfation patterns of HS dictate many activities, the core protein can potentially impact HS functions. In order to decipher this, homogeneous proteoglycan glycopeptides are needed. Herein, we report the first successful synthesis of proteoglycan glycopeptides bearing multiple aspartic acids in the core peptide and O- and N-sulfations in the glycan chain, as exemplified by the syndecan-4 glycopeptides. To overcome the high acid sensitivities of sulfates and base sensitivities of the glycopeptide during synthesis, a new synthetic approach has been developed to produce a sulfated glycan chain on a peptide sequence prone to the formation of aspartimide side products. The availability of the structurally well-defined synthetic glycopeptide enabled the investigation of their biological functions including cytokine, growth factor binding and heparanase inhibition. Interestingly, the glycopeptide exhibited context dependent enhancement or decrease of biological activities compared to the peptide or the glycan alone. The results presented herein suggest that besides varying the sulfation patterns of HS, linking the HS chain to core proteins as in proteoglycans may be an additional approach to modulate biological functions of HS in nature.

Received 26th February 2020

Accepted 8th May 2020

DOI: 10.1039/d0sc01140a

rsc.li/chemical-science

Heparan sulfate (HS) and its more highly sulfated derivative heparin are a class of highly sulfated polysaccharides.^{1,2} They are known to have multifaceted biological functions,^{3–5} which range from growth factor and chemokine binding, enzyme inhibition, to reducing blood clot formation. HS in nature is highly heterogeneous varying both in terms of the number and the location of sulfates as well as the backbone uronic acid structures. Many outstanding studies have been carried out synthesizing diverse HS structures to decode the structure and

activity relationship,^{6–12} which has led to novel therapeutics such as the synthetic anti-coagulant fondaparinux.³

In nature, HS is covalently linked to a core protein forming heparan sulfate proteoglycans (HSPGs).^{13,14} In contrast to the rich activities of HS,^{3,15,16} core proteins of HSPGs were previously thought to mainly serve as carriers of HS. However, there has been evidence suggesting that the core protein itself can be biologically active.^{17–19} For example, the core protein of a HSPG syndecan-1 can regulate invasive properties of cancer cells with the mutation of a 5 amino acid sequence in the core resulting in a loss of invasive migration abilities of these cells.²⁰ Another HSPG, syndecan-4,²¹ can activate kinases²² and facilitate the assembly of focal adhesions.²³ Interestingly, syndecan-4 mutants without any HS chains were found to be equally effective as the wild type glycoprotein in promoting focal adhesion suggesting that the core protein dictates the activity in this case.²³ As both HS and the core protein can be biologically active, the core protein may potentially modulate HS activities, adding another dimension to the functional complexity of HS.

To decipher how the core protein impacts HS functions, HSPG glycopeptides bearing well-defined homogeneous glycans

^aDepartment of Chemistry, Michigan State University, 578 South Shaw Lane, East Lansing, MI 48824, USA

^bChemistry Department, Faculty of Science, Benha University, Benha, Qalibiya 13518, Egypt

^cDivision of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA

^dDepartment of Biomedical Engineering, Michigan State University, East Lansing, MI 48824, USA

^eInstitute for Quantitative Health Science and Engineering, Michigan State University, East Lansing, MI 48824, USA. E-mail: huangxu2@msu.edu

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0sc01140a



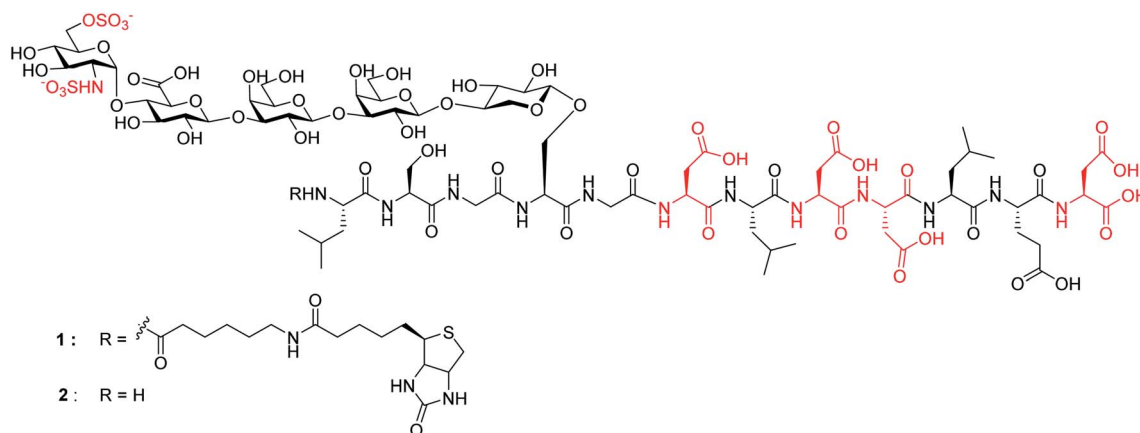


Fig. 1 Synthetic targets 1 and 2.

are critically needed, and total synthesis is an important strategy to access these complex molecules. While many innovative methods to produce HS and heparin oligosaccharides have been established,^{6–12} strategies for HS glycopeptide synthesis are underdeveloped.^{24–26} Currently, structural features common in naturally existing HSPGs, such as aspartic acids in the core peptide and glycan chains bearing both O- and N-sulfates, are not accessible with the existing synthetic methods.

Herein, we report the first successful synthesis of a human syndecan-4²¹ (amino acids 60–71) HS glycopeptides **1** and **2** bearing an N- and O-sulfated glycan chain and four aspartic acid (Asp) residues in the peptide backbone (Fig. 1). The availability of such structurally well-defined glycopeptides enabled the analysis of the roles of glycan and core peptide in their interactions with biological targets. Interestingly, the functions of the glycan chain can be modulated by the peptide, which

indicates that besides varying HS structures, the attachment of HS onto protein may be another avenue to direct HS functions.

Results and discussion

Synthetic design of syndecan-4 glycopeptides

Among many obstacles faced in HSPG synthesis, a major hurdle is that common amino acid side chain protective groups such as *t*-butyl (*t*Bu) ester and trityl requiring strong acids for deprotection cannot be applied in synthesis of glycopeptides bearing highly acid labile O- or N-sulfates. When alternative protective groups such as benzyl and allyl esters were used to protect the Asp side chain carboxylic acid, aspartimide formation is a major side reaction significantly lowering the yield of the desired peptide (Fig. 2a). Amide backbone protection by pseudoproline^{27,28} or 2,4-dimethoxybenzyl²⁹ could effectively suppress the aspartimide formation during glycopeptide synthesis. However,

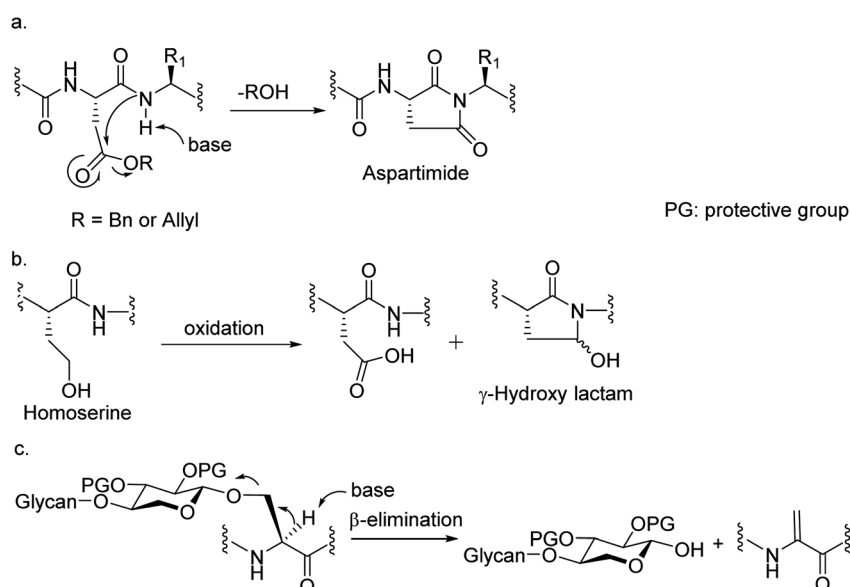


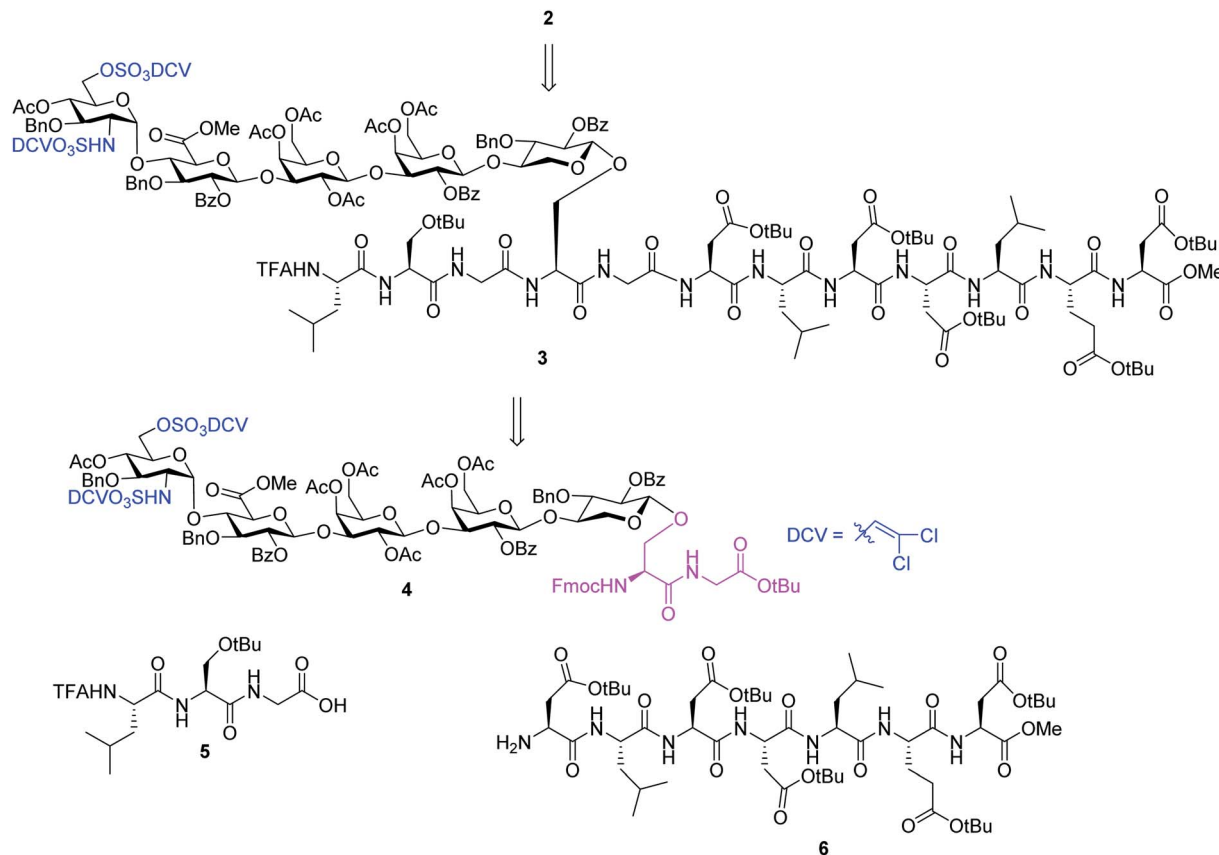
Fig. 2 Obstacles encountered in HS glycopeptide synthesis. (a) The formation of aspartimide during peptide synthesis; (b) the formation of γ -hydroxy lactam during the conversion of homoserine to Asp; (c) the potential elimination of glycan chain under basic conditions.

the strong acidic conditions required to cleave the pseudoproline or 2,4-dimethoxybenzyl group suggests they are not suitable for HS glycopeptides with free O- or N-sulfates. Recently, the cyanosulfurylide moiety was developed as a novel carboxylic acid protecting group to prevent aspartimide formation during polypeptide synthesis,³⁰ the applicability of which is yet to be demonstrated in sulfated glycopeptide synthesis. We reported a homoserine method where homoserine was used as the Asp surrogate, which could be converted to Asp after assembly of the full glycopeptide (Fig. 2b).³¹ While this strategy gave good yields of glycopeptides containing a single Asp, the yield was significantly lowered when the condition was extended to targets bearing more than one Asp due to the formation of γ -hydroxy lactam side products (Fig. 2b). As the core proteins of HSPGs can contain multiple aspartic acids flanking the serine linkage site of glycans,³² an effective strategy enabling access to these glycopeptides is needed.

Another significant hurdle in HSPG glycopeptide synthesis is N-sulfation. For the synthesis of HS oligosaccharides, N-sulfates are traditionally introduced by treating the free amines with a sulfation agent such as SO_3 , pyridine under strong basic conditions.^{8,33–36} However, when amine bearing HS glycopeptide precursors were subjected to sulfation under the traditional condition, β -elimination of the glycan chain from the peptide backbone was observed as a major side reaction presumably due to the high basicity of the reaction media needed to

promote N-sulfation (Fig. 2c). Decreasing the basicity of reaction led to incomplete sulfation with low yields of the desired glycopeptides.

We hypothesize that if acid stabilities of sulfates could be enhanced, the aforementioned difficulties could be potentially overcome. Innovative research by the Taylor,^{37–39} Linhardt^{40,41} and Widlanski⁴² groups and later extended by others^{43–46} demonstrated that sulfate esters can be attractive protective groups for sulfates. As the utility of sulfate ester chemistry has not been established in HS glycopeptide synthesis, we began to examine various sulfate esters. The dichlorovinyl (DCV) sulfate ester group³⁸ was found to be attractive due to its high stability to acids such as trifluoroacetic acid (TFA).⁴⁷ We envision that DCV sulfates can be installed first into protected glycopeptides followed by peptide elongation. Upon acidic cleavage of side chain protecting groups such as *t*Bu to free Asp, the DCV moieties can be removed under mild conditions to release the sulfate groups. With this consideration in mind, we designed the synthesis of syndecan-4 glycopeptide **2** from the fully protected precursor **3**. In turn, **3** can be generated using a cassette approach from glycosyl serine–glycine **4** building block followed by peptide chain elongation (Scheme 1). While previous synthesis utilized glycosyl serine as the cassette for peptide synthesis,^{26,48} we found that the glycosylated serine was prone to base promoted glycan elimination (Fig. 2c) during coupling with peptides more than five amino acids long. This was



Scheme 1 Retrosynthetic design of HS glycopeptide **2**.

presumably due to the reduced nucleophilicity of peptide amine slowing down the amide formation reaction and resulting in a higher relative rate of elimination. To overcome this, we chose to explore the utility of glycosyl dipeptide **4** (glycan–serine–glycine) as the cassette, as in all HSPGs, the glycosylated serine is linked to a glycine residue at its C-terminus.

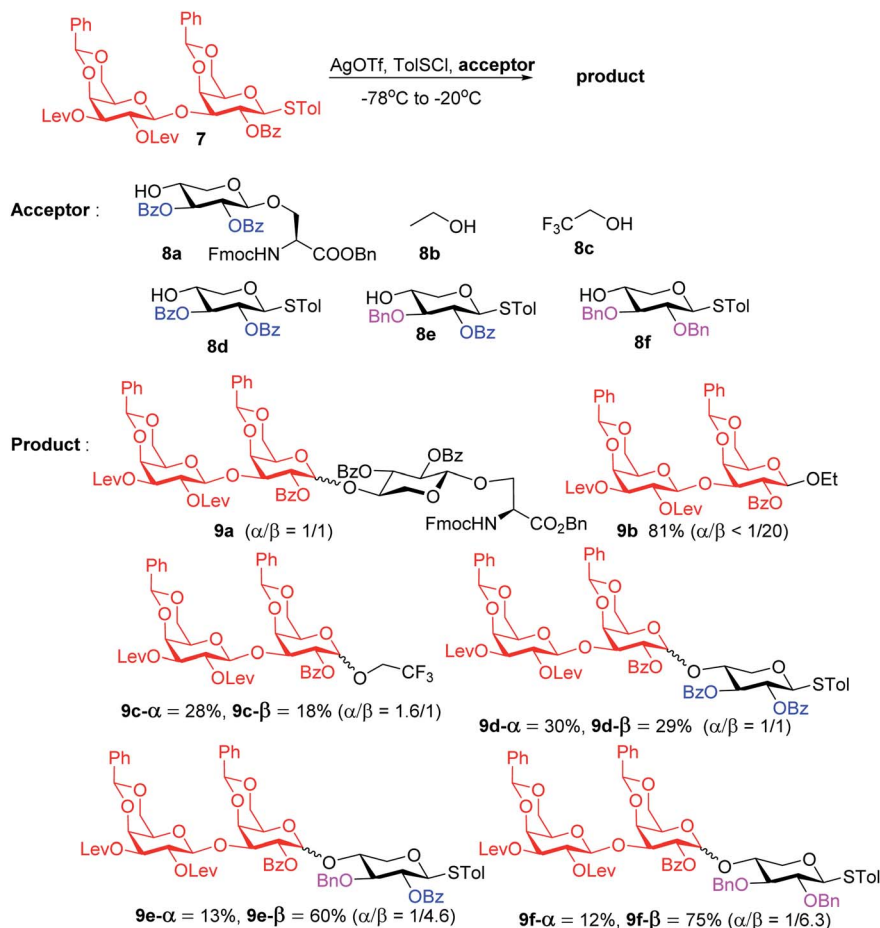
Synthesis of the glycosyl dipeptide cassette **4**

To form the glycosyl dipeptide cassette **4**, the galactose (gal)–gal–xylose (xyl) trisaccharide at the reducing end is an important module. Previous synthesis of the trisaccharide utilized gal–gal disaccharide donor **7** to glycosylate xyloside acceptor **8a** bearing two benzoyl (Bz) groups.²⁴ However, this reaction gave no selectivity ($\alpha : \beta = 1 : 1$) towards the desired β anomer **9a– β** (Scheme 2). The loss of 1,2-*trans* selectivity on galactosyl donors bearing 2-*O*-acyl groups has been reported by several groups.^{49–52} The lack of β -selectivity was presumably due to the presence of 4,6-*O*-benzylidene in the donor, conformationally reducing the propensity of neighboring group participation by the 2-*O* acyl group. This was supported by the reports that similar glycosylations using galactosyl donors without the 4,6-*O*-benzylidene moiety gave β -anomers only.^{50,51} However, in our hands, the replacement of the two 4,6-*O*-benzylidenes in **7** with

four *O*-acetate groups did not lead to any improvements of β -selectivity in glycosylation of **8a**.

To address this stereoselectivity problem, we turned our attention to the influence of glycosyl acceptor reactivity, as acceptor reactivities can significantly impact stereoselectivities.^{53–55} Two nucleophiles, *i.e.*, ethanol **8b** and trifluoroethanol (TFE) **8c**, were tested (Scheme 2). Glycosylation between **7** and the more reactive acceptor ethanol provided **9b** with high yield and β -selectivity ($\alpha/\beta < 1:20$), while less reactive acceptor TFE gave a slight preference for the α -anomer product ($\alpha/\beta = 1.6/1$). Inspired by this result, three thioxyloside acceptors (**8d–8f**) with different patterns of Bz and benzyl (Bn) groups on *O*-2 and *O*-3 positions were subsequently probed. Pre-activation based glycosylations⁵⁶ between donor **7** and the di-Bz acceptor **8d** gave product **9d** with no stereoselectivity ($\alpha/\beta = 1/1$), which was consistent with the result of **9a**. A significant increase of β -selectivity was observed upon switching 3-*O*-Bz (**8d**) to 3-*O*-Bn (**8e**),⁵⁷ while the di-*O*-Bn protected xylosyl acceptor **8f** further enhanced stereoselectivity towards the β -product (**9e– β** and **9f– β**). The results revealed that the β -selectivity of this 2 + 1 glycosylation reaction critically depends on the acceptor reactivity.

A plausible pathway to explain the outcome about the stereoselectivity is depicted in Fig. 3. Activation of **7** would lead to



Scheme 2 Stereochemical outcomes of glycosylation of disaccharide **7** with various acceptors.

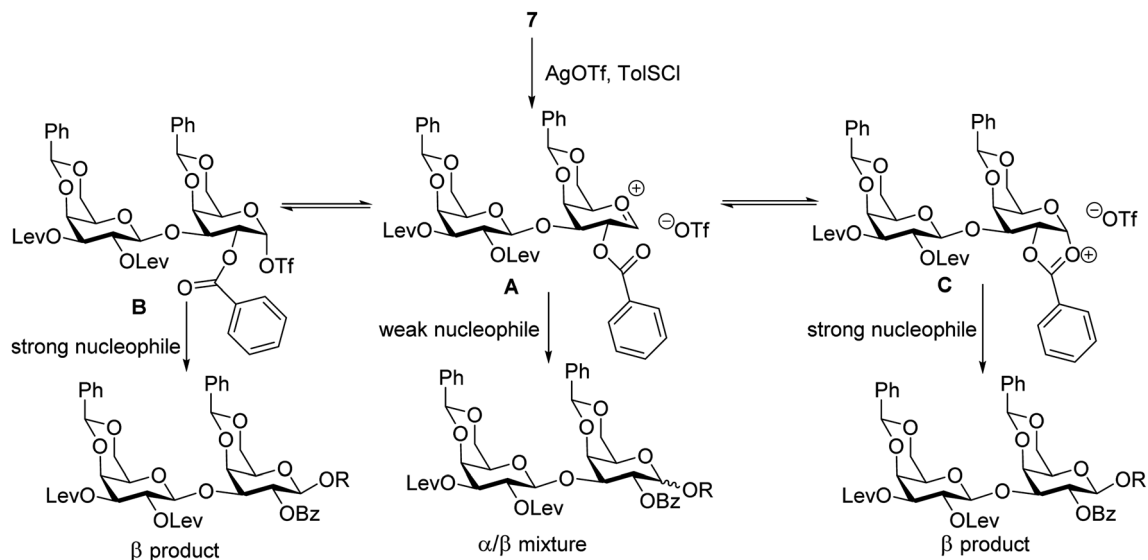
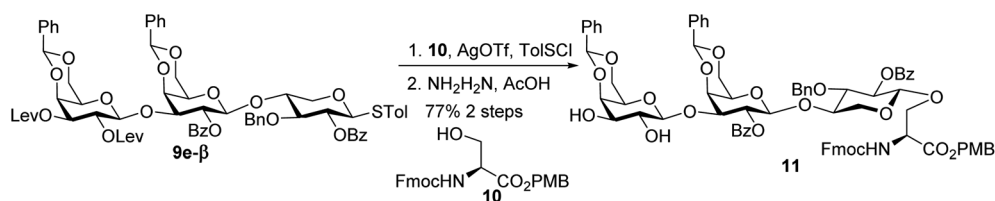


Fig. 3 Proposed pathway for stereochemical outcome of disaccharide **7** glycosylation.

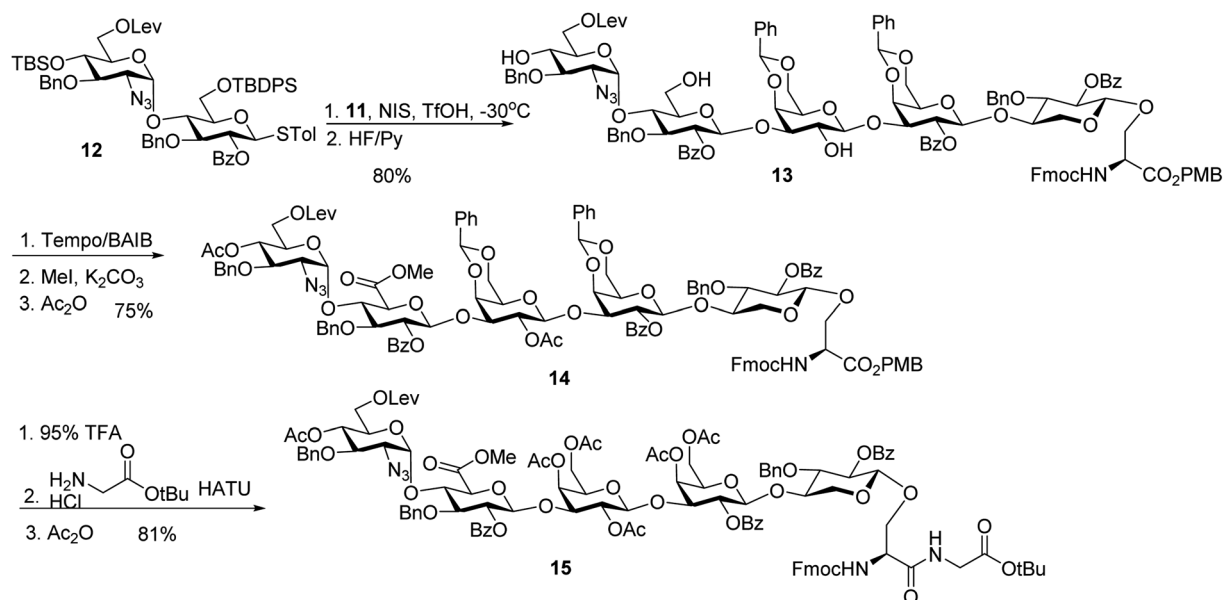


Scheme 3 Synthesis of trisaccharide serine **11**.

the oxocarbenium ion **A**, which can equilibrate with the glycosyl triflate **B**^{58,59} and/or dioxalenium ion **C**. The formation of the β -anomer can be explained through an $\text{S}_{\text{N}}2$ -like substitution on intermediate **B** or **C** by a reactive acceptor, while the less reactive

acceptor requires more reactive oxocarbenium ion **A**^{60,61} leading to α,β -mixtures in $\text{S}_{\text{N}}1$ -like fashion.

With the 2-*O*-Bz moiety on xylose of trisaccharide **9e- β** , it is a suitable donor for formation of 1,2-*trans* xylosyl linkage



Scheme 4 Synthesis of pentasaccharide – dipeptide module **15**.

(Scheme 3). The glycosylation between **9e-β** and serine **10** went smoothly followed by the removal of Lev groups to afford diol **11**. The PMB ester was selected as it could be selectively removed for peptide chain elongation.

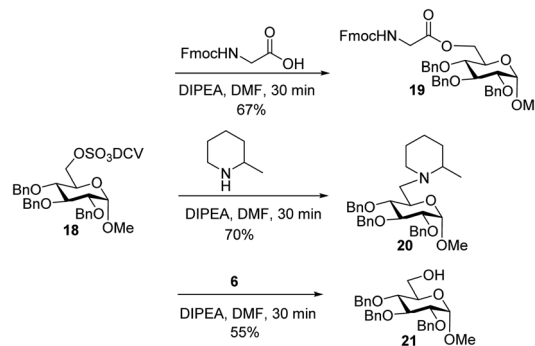
With the trisaccharide serine **11** in hand, the 2 + 3 coupling between **12** and **11** produced the pentasaccharide followed by removal of silyl ether groups to afford compound **13** in 80% yield for the two steps (Scheme 4). The primary hydroxyl group of **13** was selectively oxidized to a carboxylic acid, and subsequent methylation and acetylation generated compound **14** in a good yield. **14** was treated with 95% TFA to cleave the two benzylidenes and PMB ester. The resulting carboxylic acid was coupled with glycine *t*Bu ester followed by acetylation to give glycopeptide **15**.

Synthesis of sulfate ester bearing glycopeptides

In order to introduce sulfates to the glycan chain, the azide group in **15** was converted to a free amine by Zn reduction (Scheme 5). The sulfation of amine with reagent **16**³⁸ in the presence of 1,2-dimethylimidazole as the base provided **17** in 71% yield for the two steps. With the mild reaction condition, no elimination products were observed. Upon removal of Lev from **17**, sulfation of the free 6-OH proceeded to produce di-DCV protected glycopeptide **4**. However, attempts to elongate the peptide chain following Fmoc or *t*Bu removal from **4** failed to give any desired glycopeptides with multiple un-identifiable side products.

To better understand the failure to extend glycopeptide **4**, *O*-DCV sulfate **18** was tested as a model compound. When **18** was subjected to the typical condition (DIPEA, DMF, 30 min) for peptide coupling in the presence of Fmoc-glycine or 2-methylpiperidine, the substituted products **19** and **20** were produced. Mixing peptide **6** with **18** led to sugar **21** free of sulfates (Scheme 6). These model studies suggest that the failure of peptide elongation in glycopeptide **4** could result from the competitive substitution and hydrolysis reactions on the 6-*O* DCV sulfate ester.^{45,62}

To address this problem, we explored the alternative of installing the *O*-DCV sulfate group after peptide coupling. Compound **17** was treated with TFA to generate the carboxylic

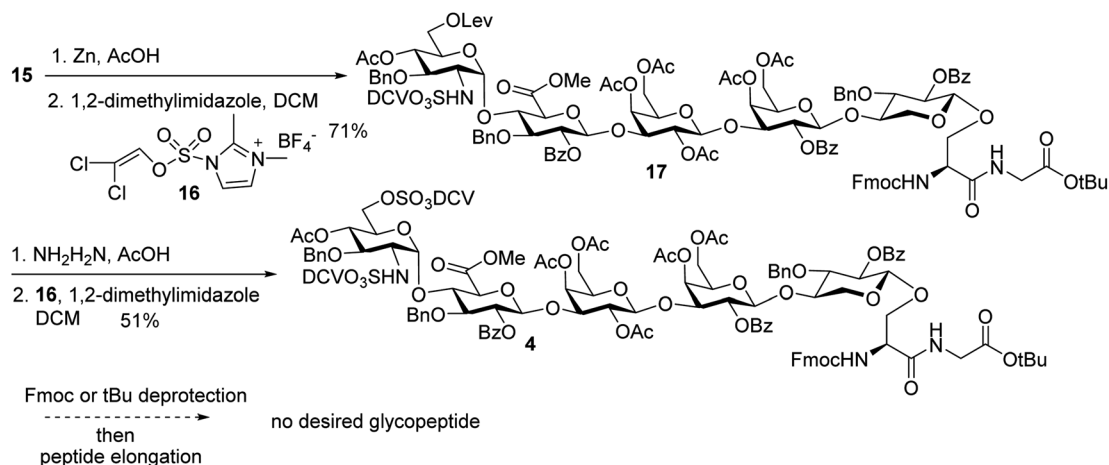


Scheme 6 Stability test of 6-*O*-DCV sulfate in model compound **18**.

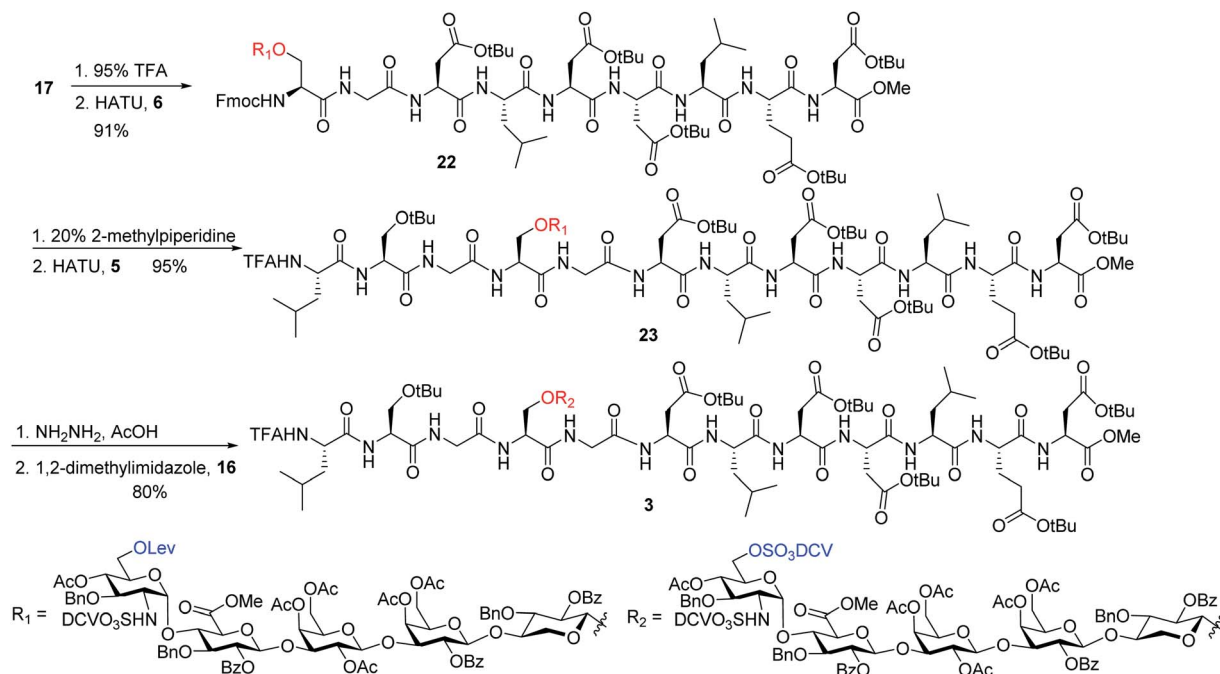
acid followed by coupling with heptapeptide **6** bearing four *t*Bu ester protected Asp residues (Scheme 7). The desired product **22** was obtained in 91% yield, with no β-elimination side product (Fig. 1c) observed. These results confirmed the hypothesis that the failure of **4** in glycopeptide synthesis was likely because of the instability of the 6-*O*-DCV sulfate ester under peptide coupling condition. Upon removal of Fmoc by 2-methylpiperidine and coupling with tripeptide **5**, compound **23** was formed in 95% yield. The Lev group on the glycan of **23** was cleaved by hydrazine acetate and the resulting primary OH was sulfated to afford the desired di-DCV protected glycopeptide **3**.

Successful synthesis of glycopeptide **1** and the corresponding peptide and glycan fragments

Next, we focused on the global deprotection. *t*Bu groups on the peptide side chain of **3** were deprotected by TFA without affecting DCV groups (Scheme 8). Subsequently, the DCV esters were removed by hydrogenolysis in the presence of Pd(OH)₂/C and ammonium formate to produce the free sulfates. Subsequent removal of the benzyl ethers, methyl esters and acyl groups by hydrogenolysis and a mild two step base treatment successfully produced the fully deprotected glycopeptide **2**. Finally, the free amine **2** was biotinylated to afford the target molecule **1**.



Scheme 5 Failure in peptide elongation of glycopeptide **4**.



Scheme 7 Synthesis of glycopeptide containing two DCV sulfate groups and full length core peptide.

To aid in biological studies, pentasaccharide **30** bearing the same glycan sequence as glycopeptide **1** was synthesized from trisaccharide **9e-β** in an analogous manner as **2** followed by biotinylation (Scheme 9a). The corresponding biotinylated peptide **31** was also synthesized (Scheme 9b).

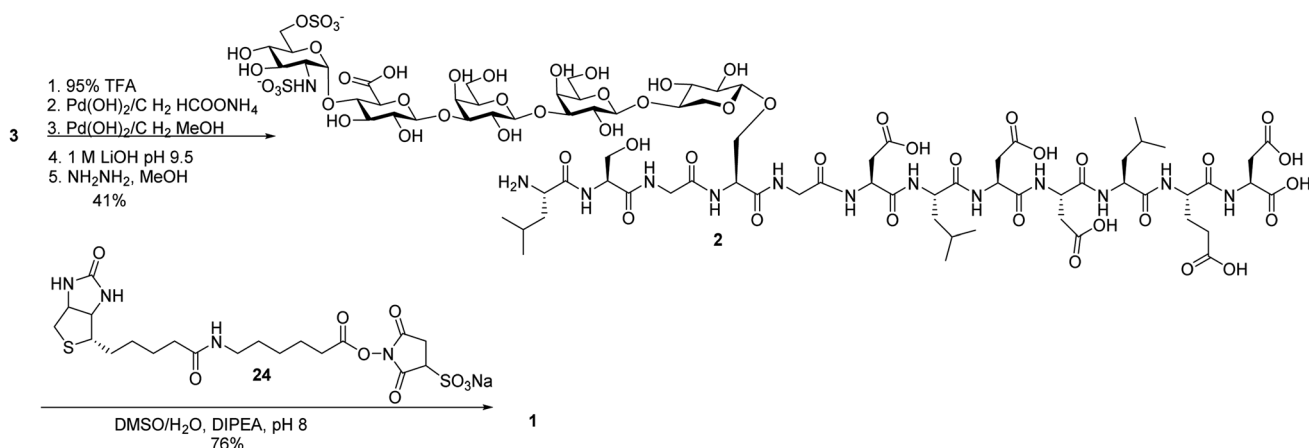
Glycan **30** significantly inhibited the activities of heparanase while glycopeptide **1** had little effects

With the synthetic glycopeptide and the corresponding peptide and glycan in hand, we began to probe the impacts of glycan and peptide on biological functions of the glycopeptide. The endoglycosidase heparanase is an important enzyme for HS metabolism, which can cleave HS polysaccharide chains at the cell surface or in the extracellular matrix.⁶³ Heparanase activities are important for a variety of biological processes, including

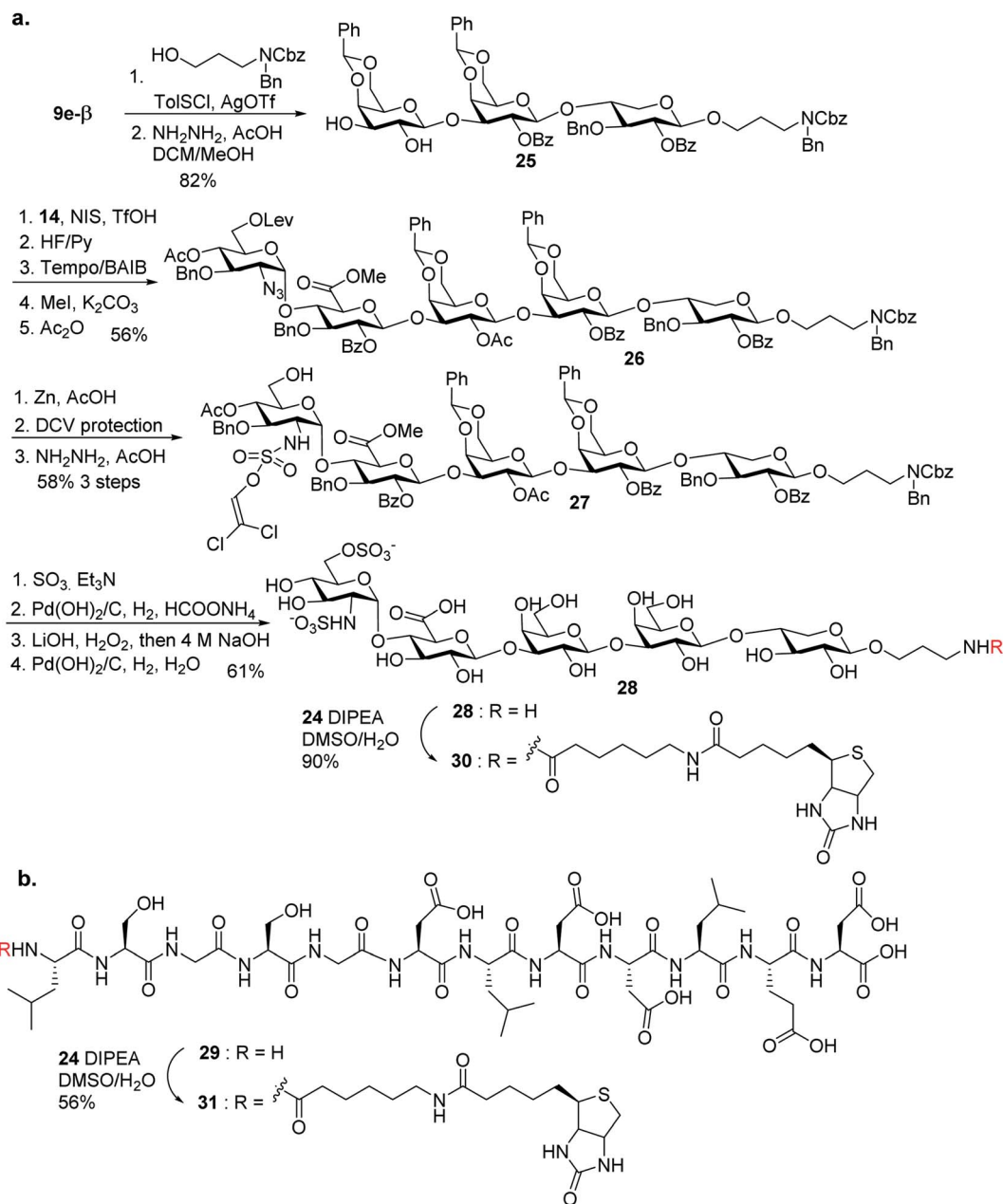
tumor metastasis and cell proliferation.⁶⁴ Some oligosaccharides can function as inhibitors of heparanase,⁶⁵ which was the case for glycan **30** as it exhibited 32% and 61% of heparanase activities at 10 and 33 μM (Table 1), respectively, using the activity assay developed by Hammond and co-workers.⁶⁶ Interestingly, neither peptide **31** nor glycopeptide **1** were able to inhibit heparanase at concentrations up to 100 μM. These results suggested that attaching the glycan to the peptide backbone blocked the inhibitory activities of the glycan.

Glycan chain and the core peptide could be synergistic in growth factor binding by the glycopeptide

Cytokines and growth factors are important proteins modulating cellular functions, many of which are known to interact with HS.^{1,8,34,67} To probe the impacts of the glycan and the core



Scheme 8 Global deprotection and biotinylation to form glycopeptide **1**.

Scheme 9 Synthesis of (a) compound **30**; and (b) compound **31**.

peptide, the binding of glycopeptide **1**, pentasaccharide **30**, and peptide **31** to several representative growth factors and cytokines including fibroblast growth factor 2 (FGF-2), interleukin 8

(IL-8) and chemokine (C-C motif ligand 13 (CCL-13)) respectively was tested using the biolayer interferometry (BLI). The biotinylated glycopeptide **1**, pentasaccharide **30**, and peptide **31**

Table 1 Inhibitory activities towards heparanase (5 nM) by glycopeptide **1**, **30** and peptide **31**. (NA denotes that no inhibitions were observed)

| Compound # | % of inhibition at | | |
|------------|--------------------|------------------|------------------|
| | 3.3 μM | 10 μM | 33 μM |
| 1 | NA | NA | NA |
| 30 | NA | 32% | 61% |
| 31 | NA | NA | NA |

Table 2 BLI experiments determined dissociation constant of synthetic compounds **1**, **30** and **31** that bind to protein CCL-13, IL-8 and FGF-2. Inactive denotes a K_D value greater than 400 μM

| Compound | K_D (nM) | | |
|-----------|------------|------|------|
| | CCL-13 | IL-8 | FGF2 |
| 1 | 498 | 28 | 5 |
| 30 | Inactive | 39 | 14.5 |
| 31 | 92 | 75 | 17 |

were immobilized on streptavidin coated biosensors, and their dissociation constants (K_D) with the proteins were measured (Table 2).

For CCL-13, glycan **30** did not exhibit much binding, while peptide **31** has a K_D of 92 nM. The biotin tag itself did not show any significant binding to the protein indicating the binding of **31** resulted from the peptide (data not shown). The glycopeptide **1** bound with CCL-13 with a K_D of 498 nM, suggesting the glycan interfered with peptide binding with CCL-13. For IL-8, glycopeptide **1** ($K_D = 28$ nM) exhibited a modest enhancement in binding compared to glycan **30** and peptide **31** alone ($K_D = 39$ and 75 nM respectively). The enhancement was more pronounced for FGF-2, as the binding is strongest for glycopeptide **1** ($K_D = 5$ nM), which is about 3-fold higher than pentasaccharide **30** ($K_D = 14.5$ nM) and peptide **31** ($K_D = 17$ nM). The FGF-2 binding results reveal that the glycan chain and the core peptide can be synergistic in enhancing glycopeptide-protein interactions, in contrast to heparanase inhibition and CCL-13 binding.

Computation analyses of the interactions of glycopeptide 2 with FGF-2 and heparanase provide insights into the impacts of peptide on HS functions

To better understand the binding of glycopeptide with FGF-2, modeling studies were performed on the FGF-2 complexes with the glycopeptide, peptide and glycan respectively, using crystal structures of the FGF2 protein (PDB⁶⁸ ID: 4OEE).⁶⁹ The potential ligand binding sites on the protein were detected by the Site Finder program implemented in Molecular Operating Environment (MOE). The results showed three potential ligand binding sites on FGF-2 with a positive Propensity of Ligand Binding (PLB) score (Fig. S4[†]). Glycopeptide **2**, pentasaccharide **28** and peptide **29** structures were docked individually into each of these potential binding sites. Molecular dynamics (MD) simulations and binding free energy calculations were performed on the distinct binding poses with highest GBVI/WSA ΔG scores. The results showed that site 1 had highest affinities for not only glycan **28** also for glycopeptide **2** and peptide **29** (Table S1[†]). X-ray crystal structure of complexes of FGF-2 and a heparin oligosaccharide showed that the glycan resided in site 1,^{69,70} which is consistent with our computation results.

The average binding energies and their experimental K_D values for FGF2 are listed in Table S2[†], and energies calculated from individual poses can be found in Table S3[†]. A strong correlation was observed between the experimental and the calculated binding free energy values (Table S2[†]).

Binding site 1 of FGF-2 is lined with many basic residues including Asn27, Arg44, Lys 119, Arg120, Lys125, Lys129, Gln134 and Lys135 (Fig. 4). MD simulations of FGF-2 complex with glycopeptide **2** showed that these residues formed hydrogen bonds with glycopeptide **2**. The distances between the side chains of Lys125 and Lys119 are within 5 Å from the sulfates on the glycan, indicating potential electrostatic interactions. In all glycopeptide poses, the glycan is located within binding site 1 while the peptide extends out of the pocket and towards the protein surface. Glycan **28** was found to reside in

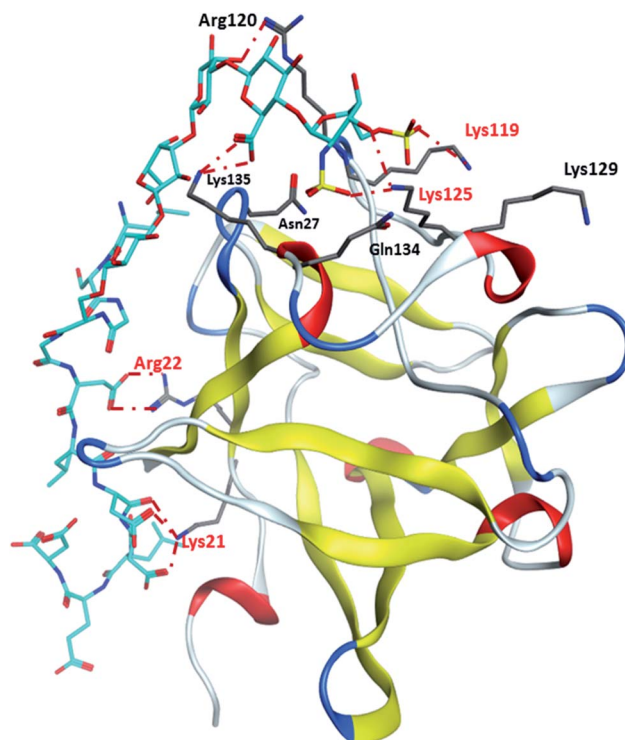


Fig. 4 Representative binding poses of glycopeptide 2 with site 1 of FGF2.

binding site 1 with an analogous conformation as that of glycopeptide **2** (Fig. S5[†]). The peptide portion of glycopeptide **2** extends out of site 1 towards the surface of FGF-2. This leads to the formation of additional salt bridges with the basic residues outside of binding site 1 including Arg22 and Lys21 (Fig. 4). These additional salt bridges are presumably responsible for improved binding to FGF-2 as observed in glycopeptide **2** as compared to glycan **28**.

The binding behavior of glycopeptide **2**, glycan **28** and peptide **29** on heparanase has also been investigated by computational studies. Glycan **28**, peptide **29**, glycopeptide **2**, and the biotin tag were docked into the heparin binding site of the heparanase (PDB ID: 5E9C)⁷¹ using MOE. The distinct poses with highest GBVI/WSA ΔG scores were further studied with molecular dynamics and binding free energy calculations (Fig. S6[†]). The average binding energies and energies calculated from individual poses can be found in Table S4.[†] The binding energy results show that glycan **28** has a higher affinity to heparanase than peptide **29** and glycopeptide **2**, respectively. The biotin tag gave little binding energy with heparanase indicating the major contributor of the binding energy of glycan **30** with heparanase should be from the interactions with the glycan.

Heparanase binding site consists of many basic residues including Lys159, Arg272, Lys231, Lys232, Arg303. Glycan **28** is oriented within the binding site by interacting with these basic residues through hydrogen bonds and ionic bonds (Fig. 5a). In glycopeptide **2** complex with heparanase, the glycan is situated within the binding site, while the peptide is extended toward

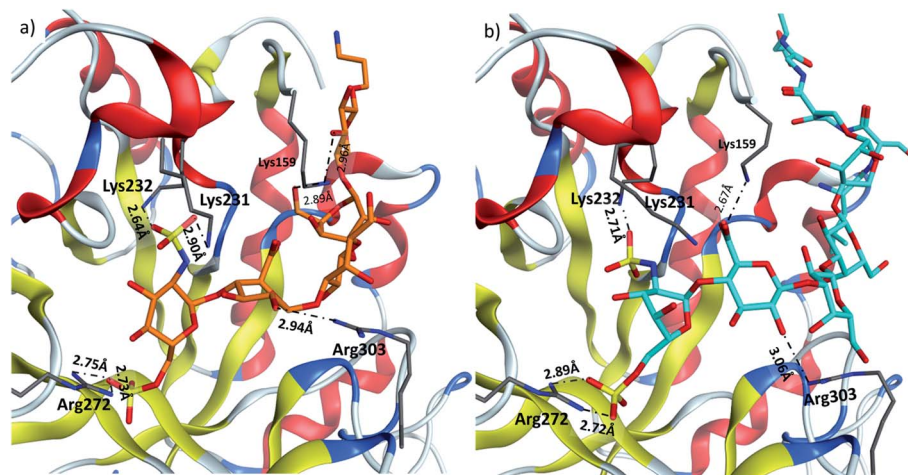


Fig. 5 Comparison of (a) glycan 28 and (b) glycopeptide 2 in heparin binding site of heparanase.

the solvent (Fig. 5b). The comparison of glycan 28 and glycopeptide 2 shows that core H-bonds and ionic interactions in the binding pocket are weakened in the glycopeptide complex. For example, interaction between Lys231 and N-sulfate group observed in glycan 28/heparanase is lost in the glycopeptide 2/heparanase complex. Furthermore, in glycan 28/heparanase complex *vs.* glycopeptide 2/heparanase, the distance between Lys232 and N-sulfate group increased from 2.64 Å to 2.71 Å, the distance between Arg272 and O-sulfate group increased from 2.75 Å to 2.89 Å, and H-bond distance between Arg303 and a hydroxyl group increased from 2.94 Å to 3.06 Å (Fig. 5). This weakening of glycan/protein interactions can be explained by the peptide backbone of glycopeptide 2 not fitting in the pocket, thus disrupting the glycan interactions with heparanase, which presumably leads to reduced affinity and inhibitory activity of glycopeptide 2 on heparanase.

Conclusion

For the first time, HSPG glycopeptides bearing multiple Asp residues in the peptide backbone and O- and N-sulfation on the glycan chain represented by syndecan-4 glycopeptides 1 and 2 have been successfully synthesized. Acceptor reactivity has been found to play a critical role in the synthesis of gal-gal-xyl trisaccharide module. A key factor for the glycopeptide synthesis is the judicious choice of reaction sequences to install O- and N-sulfates as DCV esters into the fully protected glycopeptides bearing multiple aspartic acids. With the significantly enhanced acid stabilities of DCV sulfate esters, *t*Bu esters could be utilized to protect the side chain carboxylic acids of the Asp, which enabled the successful synthesis of glycopeptide 1.

The availability of well-defined HS glycopeptide such as 1 provided unique opportunities to analyze the roles of the glycan chain and the core peptide in biological functions of the HS glycopeptide. While the glycan 30 inhibited the activities of heparanase, the glycopeptide 1 did not have much an effect on heparanase suggesting the peptide backbone could be antagonistic to glycan functions. In contrast, with the same glycan

attached to the same core peptide, the peptide backbone enhanced FGF-2 interactions with the glycan. Molecular modeling results have provided important structural insights on how the peptide backbone impacts HS functions.

While it is well known that HS backbone structure and sulfation pattern can be critical to its biological properties,^{8,34,69} our findings suggest that attaching HS to the core protein as in proteoglycans may be an additional approach to modulate functions. During some biological transformations, HS can be cleaved off HSPGs by enzymes releasing free HS, which may have very different biological properties from the parent HSPG molecules. Understanding the differences between free HS and HS in the context of HSPG can open up a new avenue of investigation into the multi-faceted biological roles of HS.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We are grateful for the financial support from the National Institute of General Medical Sciences, NIH (R01GM072667), as well as the support from the University of North Texas Academic Computing Services for the use of the UNT Research Clusters. Computational resources were provided *via* the NSF Major Research Instrumentation program supported by the National Science Foundation under Grant No. CHE-1531468.

References

- I. Capila and R. J. Linhardt, *Angew. Chem. Int. Ed.*, 2002, **41**, 390–412.
- A. K. Powell, E. A. Yates, D. G. Fernig and J. E. Turnbull, *Glycobiology*, 2004, **14**, 17R–30R.
- M. Petitou and C. A. A. van Boeckel, *Angew. Chem., Int. Ed.*, 2004, **43**, 3118–3133.

- 4 M. Bernfield, M. Götte, P. W. Park, O. Reizes, M. L. Fitzgerald, J. Lincecum and M. Zako, *Annu. Rev. Biochem.*, 1999, **68**, 729–777.
- 5 U. Lindahl, M. Kusche-Gullberg and L. Kjellen, *J. Biol. Chem.*, 1998, **273**, 24979–24982.
- 6 M. Mende, C. Bednarek, M. Wawryszyn, P. Sauter, M. B. Biskup, U. Schepers and S. Bräse, *Chem. Rev.*, 2016, **116**, 8193–8255.
- 7 P. Xu, S. Laval, Z. Guo and B. Yu, *Org. Chem. Front.*, 2016, **3**, 103–109.
- 8 C. Zong, A. Venot, X. Li, W. Lu, W. Xiao, J.-S. L. Wilkes, C. L. Salanga, T. M. Handel, L. Wang, M. A. Woolfert and G.-J. Boons, *J. Am. Chem. Soc.*, 2017, **139**, 9534–9543.
- 9 S. dey, H. J. Lo and C. H. Wong, *J. Am. Chem. Soc.*, 2019, **141**, 10309–10314.
- 10 W. Lu, C. Zong, P. Chopra, L. E. Pepi, Y. Xu, I. J. Amster, J. Liu and G.-J. Boons, *Angew. Chem. Int. Ed.*, 2018, **57**, 5340–5344.
- 11 B. F. Cress, U. Bhaskar, D. Vaidyanathan, A. Williams, C. Cai, X. Liu, L. Fu, V. M-Chari, F. Zhang, S. A. Mousa, J. S. Dordick, M. A. G. Koffas and R. J. Linhardt, *Angew. Chem. Int. Ed.*, 2019, **58**, 5962–5966.
- 12 S. B. Dulaney and X. Huang, *Adv. Carbohydr. Chem. Biochem.*, 2012, **67**, 95–136 and references cited therein.
- 13 U. Hacker, K. Nybakken and N. Perrimon, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 530–541.
- 14 J. R. Bishop, M. Schuksz and J. D. Esko, *Nature*, 2007, **446**, 1030–1037.
- 15 J. D. Esko and S. B. Selleck, *Annu. Rev. Biochem.*, 2002, **71**, 435–471.
- 16 A. J. Kvist, A. E. Johnson, M. Morgelin, E. Gustafsson, E. Bengtsson, K. Lindblom, A. Aszodi, R. Fassler, T. Sasaki, R. Timpl and A. Aspberg, *J. Biol. Chem.*, 2006, **281**, 33127–33139.
- 17 S. Choi, H. Lee, J. R. Choi and E. S. Oh, *BMB Rep.*, 2010, **4**, 305–310.
- 18 M. R. Morgan, M. J. Humphries and M. D. Bass, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 957–969.
- 19 R. V. Iozzo, *J. Clin. Invest.*, 2001, **108**, 165–167.
- 20 J. K. Langford, Y. Yang, T. Kieber-Emmons and R. D. Sanderson, *J. Biol. Chem.*, 2005, **280**, 3467–3473.
- 21 G. David, B. van der Schueren, P. Marynen, J.-J. Cassiman and H. van den Berghe, *J. Cell Biol.*, 1992, **118**, 961–969.
- 22 A. Elfenbein and M. Simons, *J. Cell Sci.*, 2013, **126**, 3799–3804.
- 23 R. L. Longley, A. Woods, A. Fleetwood, G. J. Cowling, J. T. Gallagher and J. R. Couchman, *J. Cell Sci.*, 1999, **112**, 3421–3431.
- 24 B. Yang, K. Yoshida, Z. Yin, H. Dai, H. Kavunja, M. H. El-Dakdouki, S. Sungsuwan, S. B. Dulaney and X. Huang, *Angew. Chem., Int. Ed.*, 2012, **51**, 10185–10189.
- 25 K. Yoshida, B. Yang, W. Yang, Z. Zhang, J. Zhang and X. Huang, *Angew. Chem., Int. Ed.*, 2014, **53**, 9051–9058.
- 26 W. Yang, K. Yoshida, B. Yang and X. Huang, *Carbohydr. Res.*, 2016, **435**, 180–194.
- 27 V. Ullmann, M. Radisch, I. Boos, J. Freund, C. Pohner, S. Schwarzinger and C. Unverzagt, *Angew. Chem., Int. Ed.*, 2012, **51**, 11566–11570.
- 28 P. Wang, B. Aussedat, Y. Vohra and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2012, **51**, 11571–11575.
- 29 T. Conroy, K. A. Jolliffe and R. J. Payne, *Org. Biomol. Chem.*, 2010, **8**, 3723–3733.
- 30 K. Neumann, J. Farnung, S. Baldauf and J. W. Bode, *Nat. Commun.*, 2020, **11**, 982.
- 31 W. Yang, S. Ramadan, B. Yang, K. Yoshida and X. Huang, *J. Org. Chem.*, 2016, **81**, 12052–12059.
- 32 J. D. Esko and L. Zhang, *Curr. Opin. Struct. Biol.*, 1996, **6**, 663–670.
- 33 H. A. Orgueira, A. Bartolozzi, P. Schell, R. E. J. N. Litjens, E. Palmacci and P. H. Seeberger, *Chem. - Eur. J.*, 2003, **9**, 140–169.
- 34 Y. P. Hu, Y. Q. Zhong, Z. G. Chen, C. Y. Chen, Z. Shi, M. M. L. Zulueta, C. C. Ku, P. Y. Lee, C. C. Wang and S. C. Hung, *J. Am. Chem. Soc.*, 2012, **134**, 20722–20727.
- 35 G. J. Miller, S. U. Hansen, E. Avizienyte, G. Rushton, C. Cole, G. C. Jayson and J. M. Gardiner, *Chem. Sci.*, 2013, **4**, 3218–3222.
- 36 S. U. Hansen, G. J. Miller, M. J. Cliff, G. C. Jayson and J. M. Gardiner, *Chem. Sci.*, 2015, **6**, 6158–6164.
- 37 L. J. Ingram and S. D. Taylor, *Angew. Chem., Int. Ed.*, 2006, **45**, 3503–3506.
- 38 A. M. Ali and S. D. Taylor, *Angew. Chem., Int. Ed.*, 2009, **48**, 2024–2026.
- 39 A. Y. Desoky, J. Hendel, L. Ingram and S. D. Taylor, *Tetrahedron*, 2011, **67**, 1281–1287.
- 40 N. A. Karst, T. F. Islam and R. J. Linhardt, *Org. Lett.*, 2003, **5**, 4839–4842.
- 41 N. A. Karst, T. F. Islam, F. Y. Avcı and R. J. Linhardt, *Tetrahedron Lett.*, 2004, **45**, 6433–6437.
- 42 L. S. Simpson and T. S. Widlanski, *J. Am. Chem. Soc.*, 2006, **128**, 1605–1610.
- 43 J. Chen and B. Yu, *Tetrahedron Lett.*, 2008, **49**, 1682–1685.
- 44 G. Tiruchinapally, Z. Yin, M. El-Dakdouki, Z. Wang and X. Huang, *Chem. - Eur. J.*, 2011, **17**, 10106–10112.
- 45 K. Matsushita, T. Nakata and J. I. Tamura, *Carbohydr. Res.*, 2015, **406**, 76–85.
- 46 R. E. Thompson, X. Liu, J. Ripoll-Rozada, A. Alonso-Garcia, B. L. Parker, P. J. B. Pereira and R. J. Payne, *Nat. Chem.*, 2017, **9**, 909–917.
- 47 D. Taleski, S. Butler, M. J. Stone and R. J. Payne, *Chem.-Asian J.*, 2011, **6**, 1316–1320.
- 48 S. Ramadan, W. Yang, Z. Zhang and X. Huang, *Org. Lett.*, 2017, **19**, 4838–4841.
- 49 L. Chen and F. Kong, *Tetrahedron Lett.*, 2003, **44**, 3691–3695.
- 50 J.-C. Jacquinet, *Carbohydr. Res.*, 2004, **339**, 349–359.
- 51 K. Marie, H. Go and S. Mikiko, *Chem. Lett.*, 2011, **40**, 877–879.
- 52 A. Imamura, N. Matsuzawa, S. Sakai, T. Udagawa, S. Nakashima, H. Ando, H. Ishida and M. Kiso, *J. Org. Chem.*, 2016, **81**, 9086–9104.
- 53 S. Xiang, K. L. M. Hoang, J. He, Y. J. Tan and X.-W. Liu, *Angew. Chem., Int. Ed.*, 2015, **54**, 604–607.
- 54 B. Schumann, S. G. Parameswarappa, M. P. Lisboa, N. Kottari, F. Guidetti, C. L. Pereira and P. H. Seeberger, *Angew. Chem., Int. Ed.*, 2016, **55**, 14431–14434.

- 55 S. van der Vorm, T. Hansen, J. M. A. van Hengst, H. S. Overkleeft, G. A. van der Marel and J. D. C. Codée, *Chem. Soc. Rev.*, 2019, **48**, 4688–4706.
- 56 X. Huang, L. Huang, H. Wang and X. S. Ye, *Angew. Chem., Int. Ed.*, 2004, **43**, 5221–5224.
- 57 T. Y. Huang, M. M. L. Zulueta and S. C. Hung, *Org. Lett.*, 2011, **13**, 1506–1509.
- 58 D. Crich and S. Sun, *J. Am. Chem. Soc.*, 1997, **119**, 11217–11223.
- 59 Y. Zeng, Z. Wang, D. Whitefield and X. Huang, *J. Org. Chem.*, 2008, **73**, 7952–7962.
- 60 M. G. Beaver and K. A. Woerpel, *J. Org. Chem.*, 2010, **75**, 1107–1118.
- 61 S. van der Vorm, T. Hansen, H. S. Overkleeft, G. v. d. Marel and J. D. C. Codée, *Chem. Sci.*, 2017, **8**, 1867–1875.
- 62 K. Matsushita, Y. Sato, S. Funamoto and J. i. Tamura, *Carbohydr. Res.*, 2014, **396**, 14–24.
- 63 I. Vlodaysky, Y. Friedmann, M. Elkin, H. Aingorn, R. Atzmon, R. Ishai-Michaeli, M. Bitan, O. Pappo, T. Peretz, I. Michal, L. Spector and I. Pecker, *Nat. Med.*, 1999, **5**, 793–802.
- 64 S. Rivara, F. M. Milazzo and G. Giannini, *Future Med. Chem.*, 2016, **8**, 647–680.
- 65 R. S. Loka, F. Yu, E. T. Sletten and H. M. Nguyen, *Chem. Commun.*, 2017, **53**, 9163–9166.
- 66 E. Hammond, C. P. Li and V. Ferro, *Anal. Biochem.*, 2010, **396**, 112–116.
- 67 J. L. de Paz, E. A. Moseman, C. Noti, L. Polito, U. H. von Andrian and P. H. Seeberger, *ACS Chem. Biol.*, 2007, **2**, 735–744.
- 68 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
- 69 Y. C. Li, I. H. Ho, C. C. Ku, Y. Q. Zhong, Y. P. Hu, Z. G. Chen, W. C. Lin, M. M. L. Zulueta, S. C. Hung, M. G. Lin, C. C. Wang and C. D. Hsiao, *ACS Chem. Biol.*, 2014, **9**, 1712–1717.
- 70 S. Faham, R. E. Hileman, J. R. Fromm, R. J. Linhardt and D. C. Rees, *Science*, 1996, **271**, 1116–1120.
- 71 L. Wu, C. M. Viola, A. M. Brzozowski and G. J. Davies, *Nat. Struct. Mol. Biol.*, 2015, **22**, 1016–1022.