



Cite this: *Org. Biomol. Chem.*, 2019, **17**, 2646

Received 1st February 2019,  
Accepted 12th February 2019

DOI: 10.1039/c9ob00270g

rsc.li/obc

## Glycoconjugate synthesis using chemoselective ligation†

Shuihong Cheng,<sup>a</sup> Paeton L. Wantuch,<sup>b</sup> Megan E. Kizer,<sup>b</sup> Dustin R. Middleton,<sup>b</sup> Ruitong Wang,<sup>c</sup> Mikaela DiBello,<sup>c</sup> Mingli Li,<sup>a</sup> Xing Wang,<sup>b</sup> Xuebing Li,<sup>a</sup> Vasanthi Ramachandiran,<sup>d</sup> Fikri Y. Avci,<sup>b</sup> Fuming Zhang,<sup>c</sup> Xing Zhang<sup>\*c</sup> and Robert J. Linhardt<sup>b</sup>

**Chemoselective ligation of carbohydrates and polypeptides was achieved using an adipic acid dihydrazide cross-linker. The reducing end of a carbohydrate is efficiently attached to peptides in two steps, constructing a glycoconjugate in high yield and with high regioselectivity, enabling the production of homogeneous glycoconjugates.**

Carbohydrates are often found linked to other biomolecules to form glycoconjugates. Glycoproteins and glycolipids, commonly found on the surface of cells, act as receptors for cell-cell recognition and are involved in cell growth, repair, adhesion and migration as well as tumor metastasis in cancer.<sup>1</sup> Vaccination is a key strategy for the control of infectious diseases caused by viruses, bacteria and parasites.<sup>2</sup> Oligosaccharides and polysaccharides can be antigenic,<sup>3</sup> and capsular polysaccharides (CPSs) of bacterial pathogens, including *Streptococcus pneumoniae*, *Haemophilus influenzae* type b (Hib), and *Neisseria meningitidis*, are highly antigenic and recognizable by mammalian B cell receptors. CPSs are major targets for eliciting carbohydrate specific antibody responses to confer protection from these pathogens. However, CPSs are poorly immunogenic only eliciting a short-term immune response without immunological memory. The conjugation of the carbohydrates of these pathogens to protein carriers can induce long-lasting protection against encapsulated bacteria opening a path to the development of glycoconjugate vac-

cines.<sup>3,4</sup> Glycoconjugate synthesis generally involves the random linking of carbohydrate and protein without regards to sites, leading to an incomplete understanding of mechanism of action.<sup>3</sup> The cellular and molecular mechanisms for adaptive immune activation mediated by glycoconjugate vaccines have been elucidated.<sup>5</sup> Demystifying T cell activation mechanisms of glycoconjugate vaccines represents a key step towards designing a knowledge-based, structurally-defined, generation of new vaccines. This study offers an efficient conjugation strategy.

Different carrier proteins can remarkably impact immunogenicity and the efficacy of glycoconjugate vaccines. The ability to synthesize glycoconjugates of greater structural diversity should afford an improved understanding of vaccine mechanism and result in the development of more effective vaccines.<sup>6</sup> In addition, the controlled glycosylation of peptide-based therapeutics can help protect against proteolytic degradation, denaturation and premature clearance, modulating their biophysical and physiological properties.<sup>7</sup> Improved methods for glycoconjugate synthesis are urgently needed.

There are many strategies for glycopeptide synthesis relying on conventional chemical utilization of aldehydes, thiols, activated esters or hydrazides, carboxylic acids and amines, and even new bacterial protein glycan coupling technologies (PGCT).<sup>2a,6b,c,8</sup> These approaches provide low yields and complex product mixtures, particularly when the reactants involve proteins and glycans with multiple reactive sites or with high levels of steric hindrance.<sup>2</sup>

We describe a high-yield ligation chemistry affording homogenous glycoconjugates, in which the reducing sugar is first reacted with adipic acid dihydrazide (ADH) to form a carbohydrate bearing a linker at its reducing-end. The remaining acyl hydrazide is oxidatively converted to an acyl azide and then captured as a thioester and transesterified with the cysteine residue of a peptide to obtain a thioester-linked glycoconjugate. When this cysteine is at the N-terminus of the peptide chain, the thioester rapidly rearranges to form a stable amide linkage between the carbohydrate and peptide

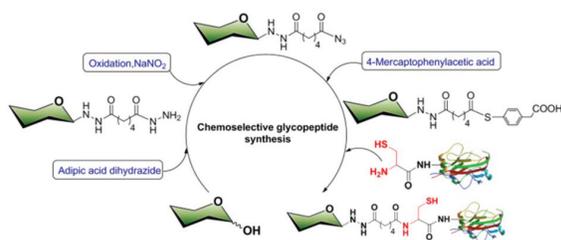
<sup>a</sup>CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, Chaoyang, China

<sup>b</sup>Department of Biochemistry and Molecular Biology, Center for Molecular Medicine, and Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA

<sup>c</sup>Department of Chemistry and Chemical Biology, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York 12180, USA. E-mail: xingzhang-cas@hotmail.com, linhar@rpi.edu

<sup>d</sup>NIH Tetramer Core Facility, Emory University, Atlanta, GA 30322, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9ob00270g



**Scheme 1** Glycopeptide preparation. Peptides where cysteine is not at the N-terminus could also be synthesized but instead result in a thio-ester linkage.

(Scheme 1). The installation of the ADH linker at the carbohydrate's reducing end and its selective reaction with the cysteine residue of the peptide affords homogeneous constructs with compositional control of the carbohydrate–protein conjugate and preserves the integrity of carbohydrate epitopes.<sup>9</sup> In addition, ADH provides a 10-atom bridge between the carbohydrate and peptide after conjugation, potentially increasing detection limits of immunoassays by reducing steric hindrance.<sup>10</sup>

Our study began by re-investigating the previously published reaction between heparin and adipic acid dihydrazide for the directional immobilization of heparin onto surfaces.<sup>11</sup> A heparin dodecasaccharide, prepared through the controlled enzymatic depolymerization of heparin,<sup>12</sup> and ADH were reacted to form a hydrazone bond (Scheme 1 and Table 1). This chemistry is particularly challenging with this glycan since the heparin dodecasaccharide is a polyanion (−48 charge) with a molecular weight 3990 Da. Its single reducing end is comprised of a relatively unreactive sugar with an anionic *N*-sulfo group adjacent to its anomeric center. Several common conjugation conditions including NaCNBH<sub>3</sub>, PhNH<sub>2</sub>/MeOH/H<sub>2</sub>O<sup>13</sup> and CeCl<sub>3</sub>·7H<sub>2</sub>O<sup>14</sup> were initially tested and resulted in very low yields (entries 1–3). Interestingly, conjugation in AcOH–DMSO<sup>15</sup> afforded a high yield of multiple ADH conjugated to the heparin dodecasaccharide (entry 4). HCOONa buffer<sup>16</sup> afforded the single desired product at 75% yield (entry 5) and PhNH<sub>2</sub>/acetate buffer<sup>17</sup> provided the same desired product in 80% yield (entry 6). Based on these results, PhNH<sub>2</sub>/acetate buffer was selected as it provided high yields in the shortest reaction time.

**Table 1** Optimization for conjugation of heparin dodecasaccharide with adipic acid dihydrazide

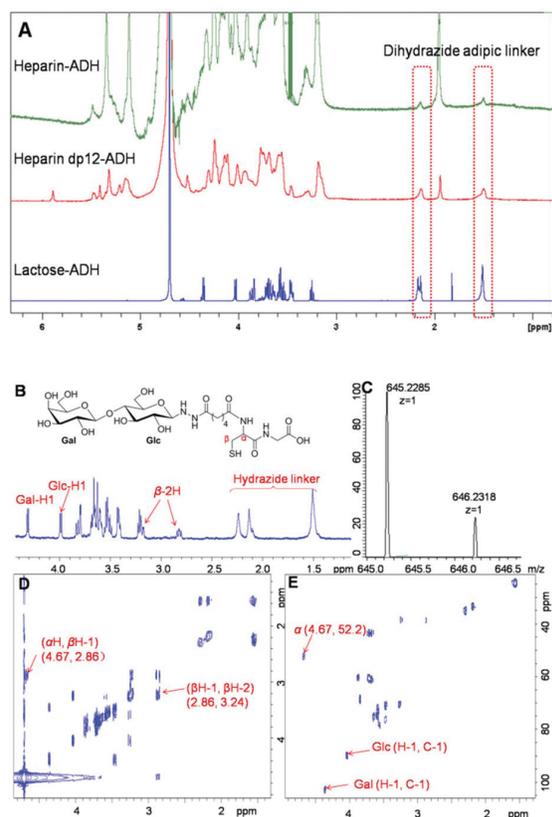
Entry	Reagents/solvents	Temp. (°C)	Time (h)	Yield
1	NaCNBH <sub>3</sub> , H <sub>2</sub> O	70	48	10%
2	PhNH <sub>2</sub> , MeOH/H <sub>2</sub> O	75	24	Trace
3	CeCl <sub>3</sub> ·7H <sub>2</sub> O, EtOH/H <sub>2</sub> O	50	1	Trace
4	AcOH/DMSO	45	48	Multiple conjugates
5	HCOONa buffer pH 4.75	37	72	75%
6	PhNH <sub>2</sub> , acetate buffer pH 5.5	50	24	80%

With the optimized conditions in hand, we next turned our attention to investigate the scope of this reaction by treating various reducing disaccharides, oligosaccharides and polysaccharides, with adipic acid dihydrazide to prepare a small library of carbohydrate-ADH conjugates (Table S1, ESI†). Coupling either simple neutral or acidic oligosaccharides with dihydrazide under these optimized conditions, such as lactose, maltotriose, maltotetraose, maltooligosaccharide and 6'-sialyllactose afforded the expected products in nearly quantitative yields, and the yields were greatly improved compared to those obtained using previous methods<sup>16</sup> (entries 1–5). Two tetrasaccharides, Glc(1→3)GlcA(1→4)Glc(1→3)GlcA<sup>18a</sup> and GlcA(1→4)Glc(1→3)GlcA(1→4)Glc,<sup>18b</sup> having glucuronic acid or glucose at the reducing end, prepared through controlled depolymerization of type-3 CPS of *Streptococcus pneumoniae* (Pn3P), also proved excellent substrates to conjugation with ADH, and are being investigated in vaccine development (entries 6 and 7). Notably, more structurally complex glycosaminoglycans, having both short or long carbohydrate chains, and low or high levels of sulfation, also furnished good results suggesting the adaptability of this method to a variety substrate (entries 8–15). Representative 1D <sup>1</sup>H NMR data for the analysis of heparin-ADH, depolymerized heparin dodecasaccharide-ADH and lactose-ADH, are presented in Fig. 1A. All the other compounds were also confirmed by NMR analyses (ESI†).

We next investigated the conjugation of different carbohydrate-ADH derivatives with a variety of peptides containing N-terminal cysteine residues (Table 2). Initially, we utilized conventional carbodiimide chemistry, involving EDC/NHS, to react lactose-ADH with the Cys-Gly dipeptide. This resulted in the recovery of starting materials with only trace amounts of conjugate (entry 1). 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM) has been reported to be highly efficient for the activation of hydrazide chemistry in protein–protein ligation.<sup>19</sup> Thus, we undertook the reaction of lactose-ADH with the Cys-Gly dipeptide using DMTMM and obtained the desired glycopeptide conjugate in moderate 37% yield (entry 2).

Liu and co-workers<sup>20</sup> recently developed a protein ligation method for the chemoselective reaction between a C-terminal peptide hydrazide and an N-terminal Cys-peptide affording a peptide bond. Inspired by this study, we investigated using sodium nitrite and thiol to form glycopeptides (Scheme 1 and Table 2). We were pleased that the reaction of lactose-ADH with the Cys-Gly dipeptide in the presence of NaNO<sub>2</sub>/MPAA (4-mercaptophenylacetic acid) generated the glycopeptide in excellent 85% yield (entry 3).

A representative characterization of lactose-dipeptide (Cys-Gly) by HRMS and NMR spectroscopy is presented in Fig. 1B–E. The signals corresponding to the anomeric protons of Gal and Glc residues are 4.36 and 4.04 ppm, respectively. The signals at 1.52, 2.15 and 2.28 ppm correspond to the adipic linker. The <sup>3</sup>J<sub>HH</sub> coupling constant of the Glc anomeric proton was 9.0 Hz, suggesting a β-linkage with the adipic linker. Proton signal of α-position at 4.67 ppm was overlapped by



**Fig. 1** NMR and HRMS analysis. Panel (A) shows 1D  $^1\text{H}$  NMR spectra of heparin-, heparin dodecasaccharide- and lactose-ADH. Two sets of peaks at 1.51 and 2.16 ppm in the  $^1\text{H}$  NMR correspond to dihydrazide adipic linker. Panel (B), (D), (E) and (C) show the 1D  $^1\text{H}$ , 2D  $^1\text{H}$ - $^{13}\text{C}$  COSY,  $^1\text{H}$ - $^{13}\text{C}$  HSQC NMR and HRMS (positive-mode) spectra of lactose-dipeptide (Cys-Gly) conjugate.

**Table 2** Conjugation of carbohydrate-adipic acid dihydrazide (ADH) with peptides

Entry	Carbohydrate-ADH	Aglycones	Conditions	Yield (%)
1	Lactose	Cys-Gly dipeptide	EDC/NHS	Trace
2	Lactose	Cys-Gly dipeptide	DMTMM	37
3	Lactose	Cys-Gly dipeptide	$\text{NaNO}_2$ , MPAA	85
4	Lactose	Cys-Laminin peptide	$\text{NaNO}_2$ , MPAA	75
5	Heparin dp12	Cys-Laminin peptide	$\text{NaNO}_2$ , MPAA	57
6	Lactose	Cys-OVA-323-339	$\text{NaNO}_2$ , MPAA	68
7	GlcA-Glc	Cys-OVA-323-339	$\text{NaNO}_2$ , MPAA	67
8	(GlcA-Glc) $_2$	Cys-OVA-323-339	$\text{NaNO}_2$ , MPAA	64
9	6'-Sialyllactose	AcHNCys-T20	$\text{NaNO}_2$ , MPAA	62

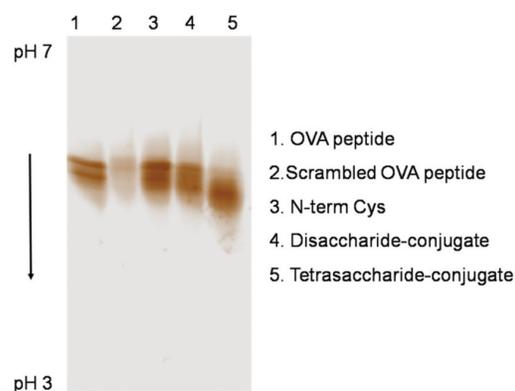
HDO in  $^1\text{H}$  NMR spectrum but could be detected (4.67, 52.2) in 2D HSQC spectrum. The H-H correlations between  $\alpha$  and  $\beta$  position of cysteine residue were also observed. The structure was characterized by HRMS with  $[\text{M} + \text{H}]^+$  645.2285 (Fig. 1C).

We next evaluated the substrate sensitivity of this reaction. Using the more complex cysteine-containing tridecapolypeptide CRKRLQVLSIRT, corresponding to the syndecan-1

binding sequence within laminin,<sup>21</sup> we coupled either simple lactose-ADH or the more complex heparin dp12-ADH (entries 4 and 5).

Cysteine-hydrazide coupling chemistry involves three steps based on a mechanistic analysis. Carbohydrate-ADH is first oxidized to the acyl azide, which subsequently reacts with the thiol group-containing reagent 4-mercaptophenylacetic acid and is then transesterified by the N-terminal cysteine residue generating an acetyl-thiol ester intermediate. Proceeding through a five-membered ring transition state, the vicinal amino group of this cysteine residue then attacks the acetyl-thiol ester forming a more stable amide linkage (Scheme S1, ESI<sup>†</sup>).

Ovalbumin (OVA) 323-339 is a major MHC class II-restricted T-cell epitope of OVA, which has been used for studies of immunological allergic responses. The N-cysteine-containing OVA 323-339, CISQAVHAAHAEINEAGR, was used in the preparation of glycopeptides with stable amide linkages to lactose and a disaccharide and tetrasaccharide derived from the type-3 CPS of *Streptococcus pneumoniae* derived (entries 6–8). The binding of the major histocompatibility complex class II protein (MHCII) with the two synthetic glycopeptides (entries 7 and 8) was demonstrated (Fig. 2). The OVA peptide (positive control, lane 1) and the two synthetic OVA-glycopeptides (lanes 4 and 5) bound to MHCII with comparable affinities. This is in contrast to the scrambled OVA peptide (negative control, lane 2), which only afforded a faint band due to aggregation of I-A<sup>d</sup> protein in the absence of a binding peptide. MHCII molecules with an empty binding groove reportedly aggregate during prolonged incubation periods.<sup>22</sup> Binding was observed through the downward shift in pI resulting from the introduction of the negative charge associated with the glycan that is especially noticeable in the tetrasaccharide-peptide conjugate. T-cells have long been believed to be non-responsive to carbohydrate antigens. Recently, however, a model for T-cell recognition of a glycoconjugate vaccine has been advanced in which a carbohydrate epitope is presented on the surface of an antigen-presenting cell through MHCII.



**Fig. 2** Novex™ pH 3–7 protein IEF gel analysis of MHCII binding with peptide and glycopeptides visualized using silver stain.

This presented epitope is then recognized by, and stimulates, carbohydrate specific CD4+ T cells.<sup>5a-c</sup> Carbohydrate does not interfere with peptide binding MHCII (Fig. 2). This suggests that in the case of glycoconjugate vaccines, the carrier peptide binds MHCII exposing the covalently linked carbohydrate for T-cell recognition.

Finally, we examined if coupling could occur with thiol-functionalized biomolecules in place of ones with an unprotected N-terminal cysteine. Enfuvirtide (ENF), a 36-residue synthetic peptide drug, is a potent inhibitor of HIV-1 infection used in the treatment of AIDS. Sialic acid conjugated ENF has a dramatically extended circulating half-life.<sup>23</sup> AcHNCys-T20, Ac-CYTSLIHSLEESQNQQEKNEQELLELDKWASLWNWF-NH<sub>2</sub>, a modified version of ENF with an acetyl protected cysteine at its N-terminus, was conjugated with 6'-sialyllactose-ADH and obtained the thioester linked glycoconjugate in good yield (entry 9). Additional glycoconjugates, to N-terminal protected, C-terminal, or internal cysteine residues also afford thioester-linked products (Table S2, ESI†).

In summary, we have successfully implemented a cross-linking strategy targeting glycoconjugate preparation. This chemistry involves ADH as cross-linking reagent combining the reducing end of carbohydrates with the thiol groups of target peptides. Based on Liu and coworkers,<sup>20</sup> who used a synthetic mono-hydrazide modified peptide as precursor in polypeptide synthesis, we introduced a bifunctional dihydrazide linker that could be directly introduced in high yield into the reducing end of the carbohydrate. The second hydrazide group was then oxidatively converted to an acyl azide, captured as a thioester and then chemoselectively transesterified with a Cysteine to afford the target glycopeptide. The strength of this approach lies in its ability to produce structurally defined, homogeneous glycan-peptide conjugates. Based on our mechanistic understanding of the activation of the immune response to glycoconjugates, this approach holds promise for the synthesis of glycopeptide vaccine.<sup>5</sup> This approach also facilitates the synthesis of glycosylated peptide-based therapeutics, protecting polypeptides against proteolysis, denaturation and improving their pharmacokinetics by modulating the biophysical and physiological properties. Finally, this chemistry has been extended to the chemoselective ligation of carbohydrate-hydrazide to biomolecules containing a single thiol group further broadening the scope of this reaction.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was supported by the National Key R&D Program of China (2017YFD0500201), National Natural Science Foundation of China (21778074) and the Open Project Program of the State Key Laboratory of Microbial Resources,

Institute of Microbiology, Chinese Academy of Sciences (SKLMR-20160604) and NIH (1R01AI123383, 5T32GM107004 (FA) and DK111958, HL125371, NS088496 (RJL).

## Notes and references

- 1 *Essentials of Glycobiology*, ed. A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger, Cold Spring Harbor Press, NY, 2017.
- 2 (a) B. Kuberan and R. J. Linhardt, *Curr. Org. Chem.*, 2000, **4**, 653–677; (b) J. Hütter and B. Lepenies, in *Carbohydrate-Based Vaccines: Methods and Protocols*, ed. B. Lepenies, Springer New York, New York, NY, 2015, pp. 1–10.
- 3 F. Y. Avci and D. L. Kasper, *Annu. Rev. Immunol.*, 2010, **28**, 107–130.
- 4 O. T. Avery and W. F. Goebel, *J. Exp. Med.*, 1931, **54**, 437–447.
- 5 (a) F. Y. Avci, X. M. Li, M. Tsuji and D. L. Kasper, *Nat. Med.*, 2011, **17**, 1602–1610; (b) F. Y. Avci, X. M. Li, M. Tsuji and D. L. Kasper, *Nat. Protoc.*, 2012, **7**, 2180–2192; (c) D. R. Middleton, L. Sun, A. V. Paschall and F. Y. Avci, *J. Immunol.*, 2017, **199**, 598–603.
- 6 (a) R. Dagan, J. Poolman and C.-A. Siegrist, *Vaccine*, 2010, **28**, 5513–5523; (b) K. M. Koeller and C.-H. Wong, *Chem. Rev.*, 2000, **100**, 4465–4494; (c) B. G. Davis, *J. Chem. Soc., Perkin Trans. 1*, 1999, 3215–3237.
- 7 R. J. Solá and K. Griebenow, *BioDrugs*, 2010, **24**, 9–21.
- 8 (a) V. S. Terra, D. C. Mills, L. E. Yates, S. Abouelhadid, J. Cuccui and B. W. Wren, *J. Med. Microbiol.*, 2012, **61**, 919–926; (b) P.-A. Ashford and S. P. Bew, *Chem. Soc. Rev.*, 2012, **41**, 957–978; (c) H. Hojo and Y. Nakahara, *Pept. Sci.*, 2007, **88**, 308–324; (d) H. Herzner, T. Reipen, M. Schultz and H. Kunz, *Chem. Rev.*, 2000, **100**, 4495–4538.
- 9 R. Adamo, *Acc. Chem. Res.*, 2017, **50**, 1270–1279.
- 10 A. Basu, T. G. Shrivastav and K. P. Kariya, *Clin. Chem.*, 2003, **49**, 1410–1412.
- 11 V. D. Nadkarni, A. Pervin and R. J. Linhardt, *Anal. Biochem.*, 1994, **222**, 59–67.
- 12 A. Pervin, C. Gallo, K. A. Jandik, X.-J. Han and R. J. Linhardt, *Glycobiology*, 1995, **5**, 83–95.
- 13 T. P. Coxon, T. W. Fallows, J. E. Gough and S. J. Webb, *Org. Biomol. Chem.*, 2015, **13**, 10751–10761.
- 14 J. M. dos Santos, *Eur. J. Org. Chem.*, 2014, 6411–6417.
- 15 G. Li, L. Li, C. Xue, D. Middleton, R. J. Linhardt and F. Y. Avci, *J. Chromatogr. A*, 2015, **1397**, 43–51.
- 16 N. S. Flinn, M. Quibell, T. P. Monk, M. K. Ramjee and C. J. Urch, *Bioconjugate Chem.*, 2005, **16**, 722–728.
- 17 K. Godula and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2010, **132**, 9963–9965.
- 18 (a) D. R. Middleton, X. Zhang, P. L. Wantuch, A. Ozdilek, X. Liu, R. Lopilato, N. Gangasani, R. Bridger, L. Wells, R. J. Linhardt and F. Y. Avci, *Glycobiology*, 2018, **28**, 90–99; (b) D. J. Lefeber, R. Gutiérrez Gallego, C. H. Grün,

- D. Proietti, S. D'Ascenzi, P. Costantino, J. P. Kamerling and J. F. G. Vliegthart, *Carbohydr. Res.*, 2002, **337**, 819–825.
- 19 A. Leitner, L. A. Joachimiak, P. Unverdorben, T. Walzthoeni, J. Frydman, F. Forster and R. Aebersold, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 9455–9460.
- 20 G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui and L. Liu, *Angew. Chem., Int. Ed.*, 2011, **50**, 7645–7649.
- 21 L. R. Balaoing, A. D. Post, A. Y. Lin, H. Tseng, J. L. Moake and K. J. Grande-Allen, *PLoS One*, 2015, **10**, e0130749.
- 22 G. M. Grotenbreg, M. J. Nicholson, K. D. Fowler, K. Wilbuer, L. Octavio, M. Yang, A. K. Chakraborty, H. L. Ploegh and K. W. Wucherpfennig, *J. Biol. Chem.*, 2007, **282**, 21425–21436.
- 23 S. Cheng, X. Chang, Y. Wang, G. F. Gao, Y. Shao, L. Ma and X. Li, *J. Med. Chem.*, 2015, **58**, 1372–1379.