


 Cite this: *RSC Adv.*, 2025, 15, 18010

# Synthesis, conformational analysis and biological activity of xylopyranosyl sulfur-containing glycosides: dependence of sulfur atom configuration†

 Pilar Blasco,<sup>a</sup> Jonas Stähle,<sup>a</sup> Karin Thorsheim,<sup>b</sup> Axel Furevi,<sup>a</sup> Anna Siegbahn,<sup>b</sup> Emil Tykesson,<sup>c</sup> Gunilla Westergren-Thorsson,<sup>c</sup> Ulf Ellervik<sup>b</sup> and Göran Widmalm<sup>\*a</sup>

Proteoglycans (PGs) consist of a core protein with covalently bound glycosaminoglycan (GAG) chains that are linked *via* a tetrasaccharide. PGs are important macromolecules that are involved in biological processes such as cell growth and differentiation. A key enzyme in the biosynthesis of PG GAG chains is  $\beta$ -1,4-galactosyltransferase 7 ( $\beta$ 4GalT7) that catalyzes the transfer of galactose to a xylose residue in the formation of the linker tetrasaccharide. It is well known that the addition of xylosides containing naphthyl aglycones can initiate the biosynthesis of GAG chains by acting as substrates for  $\beta$ 4GalT7. Previous studies have shown that its galactosylation ability is increased by using bioisosters, in which the anomeric oxygen is replaced with sulfur or sulfur-containing functional groups. Thus, 2-naphthyl xyloosyl sulfoxides were synthesized and characterized by  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectroscopy relying on both one- and two-dimensional experiments to differentiate the stereochemistry at the sulfur atom. Notably, the conformationally dependent  $^3J_{\text{CH}}$  coupling constants between the anomeric proton and the C2' atom of the naphthyl group were large and significant,  $\geq 3.3$  Hz, for the (*S*)<sub>S</sub>-configured compound as well as for the *O*-glycoside and the thio-derivative whereas the corresponding coupling for the (*R*)<sub>S</sub>-configured compound and the sulfone derivative had  $^3J_{\text{C2',H1}} < 0.6$  Hz and  $^3J_{\text{C2',H1}} < 0.5$  Hz, respectively. Quantum mechanical calculations of the  $^3J_{\text{C2',H1}}$  coupling constant corroborated the experimentally observed trends at the  $\phi$  torsion angle. The galactosylation by  $\beta$ 4GalT7 of the different acceptor substrates showed the highest affinity for the (*R*)<sub>S</sub>-configured compound and the sulfone derivative whereas an intermediate affinity was present for the (*S*)<sub>S</sub>-configured compound and the thio-derivative. The enzyme efficiency exhibited with the latter substrate was more than three times higher than with any other of the thio-derivatives. From molecular docking of the acceptor substrates to the UDP-galactose: $\beta$ 4GalT7 complex specific intermolecular interactions were identified. The binding affinity correlates with stacking to a tyrosine residue and a weak C–H $\cdots$ O hydrogen bond between the indole group of tryptophan in the enzyme and a proximate oxygen atom of sulfone and sulfinyl derivatives of 2-naphthyl xylosides.

 Received 21st January 2025  
 Accepted 23rd May 2025

DOI: 10.1039/d5ra00498e

[rsc.li/rsc-advances](https://rsc.li/rsc-advances)

## Introduction

Glycosyltransferases constitute a large family of enzymes that are involved in the biosynthesis of oligosaccharides, polysaccharides and glycoconjugates. These saccharide-containing enzyme products mediate a wide range of functions from

structure and storage to signaling.  $\beta$ -1,4-galactosyltransferase 7 ( $\beta$ 4GalT7) was identified in 1999, by cloning and expression, as having galactosyltransferase I activity, *i.e.*, catalyzing the transfer of galactose from uridine-5'-diphosphogalactose (UDP-Gal) to a xylose residue in the biosynthesis of proteoglycans (PGs).<sup>1</sup> Qasba and co-workers later reported the crystal structure of  $\beta$ 4GalT7 from both human and *Drosophila*, where a conformational change upon  $\text{Mn}^{2+}$  and UDP-Gal binding is observed, forming an acceptor binding pocket for the substrate xylose.<sup>2,3</sup> Crystallization of the ternary enzyme complex showed a Michaelis complex where the coupling of galactose to xylose takes place *via* an  $\text{S}_{\text{N}}2$ -type mechanism.<sup>3</sup> In a previous work, we further investigated the active site of  $\beta$ 4GalT7 and observed a very narrow pocket consisting of a precise set of hydrogen

<sup>a</sup>Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91 Stockholm, Sweden. E-mail: [goran.widmalm@su.se](mailto:goran.widmalm@su.se)

<sup>b</sup>Center for Analysis and Synthesis, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

<sup>c</sup>Department of Experimental Medical Science, Lund University, BMC, SE-221 00 Lund, Sweden

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d5ra00498e>



bond acceptor residues, where the xylose residue was found to be the ideal substrate.<sup>4</sup>

Proteoglycans (PGs) constitute the majority of the extracellular matrix and they are involved in a variety of biological processes, such as cell growth, brain development, differentiation, and cell adhesion.<sup>5–7</sup> These macromolecules are important regulators of tumor progression through their interactions with several proteins.<sup>8–13</sup> These glycoconjugates consist of one or several glycosaminoglycan (GAG) chains that are covalently bound to a core protein *via* a linker tetrasaccharide. The GAG chains are essential for the biological activity of PGs; GAGs interact with growth factors, cytokines, enzymes, and other signaling molecules which are found to be involved in cancer processes.<sup>14,15</sup> The biosynthesis of PGs is mediated by a cascade of different glycosyltransferases; the process begins with the linking of xylose<sup>16</sup> to a serine residue of the protein core, followed by the addition of two galactose moieties and one glucuronic acid to form the linker region. This is then elongated by the addition of alternating disaccharide units and modified, through *e.g.* epimerization, *N*-deacetylation, and *O*- and *N*-sulfation, to form different types of GAG chains, *i.e.* heparan sulfate (HS) and chondroitin sulfate/dermatan sulfate (CS/DS).<sup>4,17,18</sup>

It is known that exogenously added xylosides carrying hydrophobic aglycones could induce free GAG chain formation by acting as substrate for  $\beta$ 4GalT7. This competition approach was found to cause growth inhibition of tumor cells.<sup>19–23</sup> Moreover, the modified xylosides could increase the galactosyltransferase activity of the enzyme. Recent results have shown that xylosides carrying aromatic aglycones, *e.g.* a naphthyl moiety, generate efficient  $\beta$ 4GalT7 acceptors.<sup>4,19</sup> Moreover, the positioning of the aromatic moiety and the xylose residue further apart by using a linker was well tolerated by the enzyme without diminishing the activity.<sup>24</sup> Even more, high bulky aglycones were good acceptors.<sup>25</sup>

The anomeric oxygen atom in 2-naphthyl  $\beta$ -D-xylopyranoside (1) can be replaced by a sulfur atom (Fig. 1), generating thioxyloside 2, without losing its activity,<sup>19</sup> whereas exchanging it to a methylene group, significantly reduced the galactosylation activity.<sup>20</sup> These results agree with earlier investigations of GAG priming of *S*- and *C*-xylosides.<sup>14,17</sup> In addition, *N*-xylosides and triazolyl including xylosides have been evaluated in terms of GAG priming ability, where the structure of the aglycone strongly affects their activity.<sup>26–29</sup>

Interestingly, xylosyl sulfone 3, where a sulfonyl group is present at the glycosidic linkage, is several times more efficient in terms of galactosylation compared to xyloside 1 as well as thioxyloside 2.<sup>25</sup> To further investigate the effect of xyloside modifications, we have synthesized xylosyl sulfoxides (4 and 5) and studied conformational preferences of all the analogues in solution in order to gain insight into their ability to act as substrates for  $\beta$ 4GalT7. Sulfoxides possess a stereogenic center at the sulfur atom;<sup>30</sup> hence diastereomers can be formed with different configuration, (*R*)<sub>S</sub> and (*S*)<sub>S</sub>, at the sulfinyl sulfur atom (Fig. 1). With these compounds in hand, we investigated the activity of  $\beta$ 4GalT7 in galactosylation of xylosyl sulfoxides as acceptors. Moreover, molecular docking of the sulfoxide

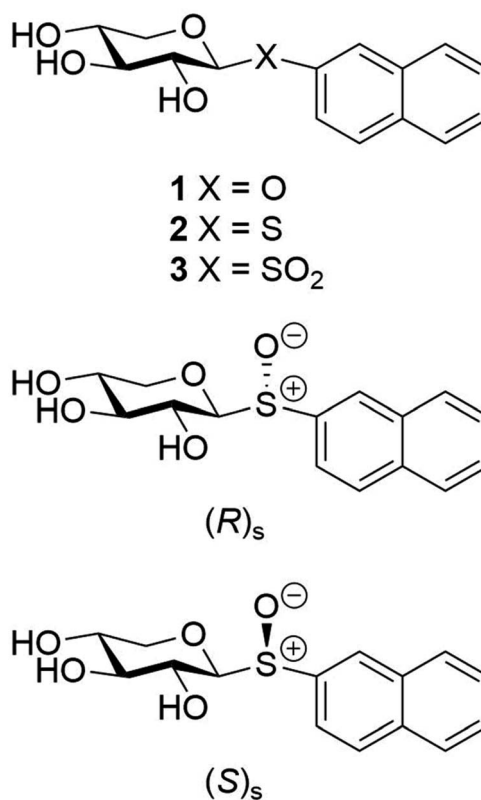


Fig. 1 Schematic of 2-naphthyl  $\beta$ -D-xylopyranoside compounds 1–3 and the diastereomeric xylosyl sulfoxides having the (*R*)<sub>S</sub>- and (*S*)<sub>S</sub>-configuration at the sulfur atom.

molecules was also performed to illuminate similarities and differences to 1–3 in the active site of  $\beta$ 4GalT7.

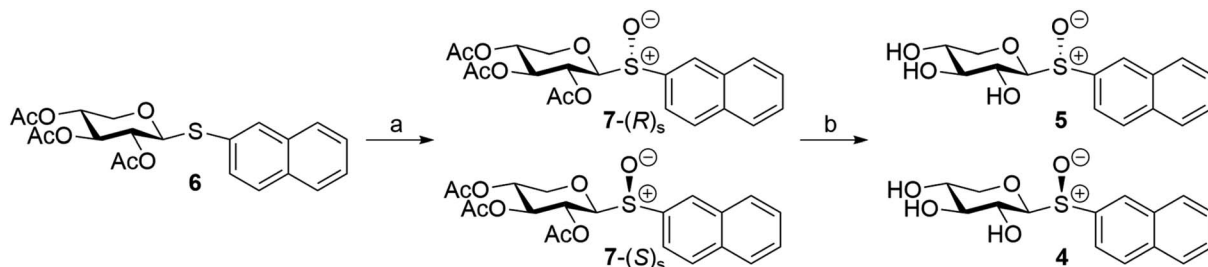
## Results and discussion

### Synthesis and diastereomeric differentiation

The diastereomeric xylosyl sulfoxides with the (*R*)<sub>S</sub>- and (*S*)<sub>S</sub>-configuration at the sulfur atom (Fig. 1) were synthesized by a two-step procedure from thioxylopyranoside 6 by oxidation using *m*CPBA at  $-78$  °C, which generated sulfoxides 7-(*R*)<sub>S</sub> and 7-(*S*)<sub>S</sub> in an excellent 95% yield, but as an inseparable mixture (Scheme 1). The relative ratio between the major and the minor product was 7 : 3 based on <sup>1</sup>H NMR analysis of the mixture. A diastereomeric product was anticipated since equatorial thioxylosides are known to show poor selectivity compared to the axial analogues in the oxidation reaction forming sulfoxides.<sup>31,32</sup> This may be explained by conformational preferences due to the *exo*-anomeric effect where predominantly the *S*<sub>S</sub> configuration is obtained when oxidizing equatorial thioglycosides and *R*<sub>S</sub> is almost exclusively generated when oxidizing a thioglycoside with an axially oriented C–S bond.

For equatorial glycosyl *S*<sub>S</sub> sulfoxides the <sup>13</sup>C NMR chemical shift of C1 is observed more downfield than the analogous *R*<sub>S</sub> sulfoxides.<sup>33,34</sup> In the product mixture of 7-(*R*)<sub>S</sub> and 7-(*S*)<sub>S</sub> the major diastereomer has for its anomeric carbon an NMR chemical shift of 93.4 ppm whereas the minor isomer shows the





Scheme 1 Synthesis of target sulfoxides 4 and 5. Reagents and conditions: (a) *m*CPBA, CH<sub>2</sub>Cl<sub>2</sub>, –78 °C, 2.5 h, 95% as a mixture ((*R*)<sub>S</sub>-7 and (*S*)<sub>S</sub>-7); (b) G/GHNO<sub>3</sub>, MeOH/CH<sub>2</sub>Cl<sub>2</sub> (9 : 1), rt, 30 min, 64% (4) and 21% (5).

corresponding chemical shift upfield at 91.8 ppm, *i.e.*  $\Delta\delta_C^{R-S} -1.6$ , which, according to literature data of similar compounds, makes it possible to assign the major isomer of the xylosyl sulfoxide to the (*S*)<sub>S</sub>-configuration and, consequently, the minor isomer to the (*R*)<sub>S</sub>-configuration. Furthermore, the two diastereomers of phenyl 2,3,4,6-*tetra-O*-benzoyl-1-sulfinyl- $\beta$ -D-glucopyranoside differ, *inter alia*, in the <sup>1</sup>H NMR chemical shift of their anomeric protons,<sup>35</sup> being 4.87 ppm and 4.62 ppm for the (*S*)<sub>S</sub>- and (*R*)<sub>S</sub>-isomers, respectively, *i.e.*  $\Delta\delta_H^{R-S} -0.25$ . Likewise, the <sup>1</sup>H NMR chemical shift of H1 for the herein *per-O*-acetylated 2-naphthyl derivative of the major isomer is observed at 4.57 ppm whereas that of the minor isomer resonates at 4.32 ppm, *i.e.*  $\Delta\delta_H^{R-S} -0.25$ , consistent with the conclusion made based on <sup>13</sup>C NMR data.

De-*O*-acetylation of the mixture of 7-(*R*)<sub>S</sub> and 7-(*S*)<sub>S</sub> using standard Zemplén conditions generated partial cleavage of the glycosidic linkage forming methyl  $\beta$ -D-xylopyranoside, which was difficult to separate from the desired target sulfoxides. However, when using a solution of guanidine/guanidinium nitrate<sup>36</sup> the hydrolysis of the glycosidic bond could be prevented and both diastereomers were separated by column chromatography. The major compound (4) was eluted first followed by the minor compound (5). Based on the fact that *m*CPBA oxidation of phenyl 2,3,4-*tri-O*-benzoyl-1-thio- $\beta$ -D-xylopyranoside gives a 7 : 3 ratio of the (*S*)<sub>S</sub>- and (*R*)<sub>S</sub>-configuration at the sulfur atom, we tentatively assign compound 4 as having the former and compound 5 as having the latter stereochemistry (Scheme 1).<sup>37</sup> In addition, the UV spectra of compounds 4 and 5 were closely similar and showed absorption band maxima at ~225 nm while the CD spectra at this wavelength were of opposite signs, consistent with different chirality at the sulfur atom<sup>38,39</sup> in the two compounds.

### Xyloside conformation by NMR and QM

**NMR chemical shift comparison.** 1D <sup>1</sup>H and <sup>13</sup>C NMR experiments in combination with 2D <sup>1</sup>H,<sup>1</sup>H- and <sup>1</sup>H,<sup>13</sup>C-correlated experiments were used to corroborate and assign <sup>1</sup>H and <sup>13</sup>C NMR chemical shifts of compounds 1–5 in methanol-*d*<sub>4</sub>. The <sup>1</sup>H chemical shifts were subsequently refined aided by NMR spin-simulation of spectra based on a total line-shape analysis.<sup>40</sup> However, a residual pentet CHD<sub>2</sub>OD peak from the solvent overlapped with a multiplet resonances in 4, which interfered with the total line-shape analysis. This was resolved

by a WEFT NMR experiment<sup>41</sup> in which the solvent peak was removed from the spectrum by choosing a suitable delay time of the inversion-recovery pulse sequence that is the basis for the solvent suppression experiment. The chemical shift difference for the anomeric carbon in 4 vs. 5 is smaller, 0.5 ppm, than that observed for the protected compounds (*vide supra*). For the major compound 4 the C1 chemical shift is higher than for the minor compound 5, like for 7-(*S*)<sub>S</sub> vs. 7-(*R*)<sub>S</sub>, indicating that 4 has the (*S*)<sub>S</sub>-configuration and that 5 has the (*R*)<sub>S</sub>-configuration; consequently  $\Delta\delta_C^{5-4} -0.5$ . Likewise, for the anomeric protons,  $\Delta\delta_H^{5-4} -0.34$ . However, for the H2 proton the chemical shift displacement is reversed<sup>32</sup> and larger,  $\Delta\delta_H^{5-4} 0.42$ , further corroborating the assigned stereochemistry.

**NMR and QM based conformational analysis.** NMR experiments suitable for addressing conformational preferences at the glycosidic linkage of the 2-naphthyl xylosyl glycosides include, *inter alia*, 1D <sup>1</sup>H,<sup>1</sup>H-NOESY<sup>42</sup> and heteronuclear one-dimensional long-range (1DLR)<sup>43,44</sup> techniques. In the former experiment correlations from the anomeric proton to H1' and/or H3' in the naphthyl group can be anticipated being dependent on the conformation(s) at the glycosidic linkage; intra-residue NOEs from H1 to H3 or alternatively to H5<sub>pro-S</sub> may be used (Fig. 2), in conjunction with a molecular model, to obtain a proton–proton reference distance  $r_{ref}$  for determination of an effective distance  $r_{ij}$  across the glycosidic linkage using the isolated spin-pair approximation (ISPA)<sup>45</sup> and the relationship  $r_{ij} = r_{ref}(\sigma_{ref}/\sigma_{ij})^{1/6}$  where  $r_{ref}$  is the known reference distance from the molecular model,  $\sigma_{ref}$  and  $\sigma_{ij}$  are the cross-relaxation rates for the reference interaction and the interaction between spins *i* and *j*, respectively. This reasoning is based on the fact that the  $\beta$ -D-xylopyranosyl residue is present in the <sup>4</sup>C<sub>1</sub> chair conformation.<sup>46,47</sup> In the latter technique the <sup>1</sup>H,<sup>13</sup>C-heteronuclear spin–spin coupling constant related to the torsion angle  $\phi$  (Fig. 2) is to be extracted. The torsion angle dependence can be described by Karplus-type relationships<sup>48,49</sup> for <sup>3</sup>J<sub>CH</sub> where the atom-sequences C–O–C–H and C–S–C–H are present in compounds 1 and 2, respectively. To proceed with the conformational analysis quantum mechanics (QM) geometry optimized models were computed, which for  $\phi$  are referred to as *exo-syn*, *non-exo* and *anti- $\phi$* ,<sup>50</sup> exemplified for the (*R*)<sub>S</sub> and (*S*)<sub>S</sub> sulfoxide isomers (Fig. 3), whereas for the torsion angle  $\psi$  (Fig. 2) defined by the atom sequence C1–X–C2'–C1' where X = O, S the conformational states are referred to as *syn*-periplanar (sp) for which  $\psi \approx 0^\circ$  and *anti*-periplanar (ap) where  $\psi \approx 180^\circ$ .



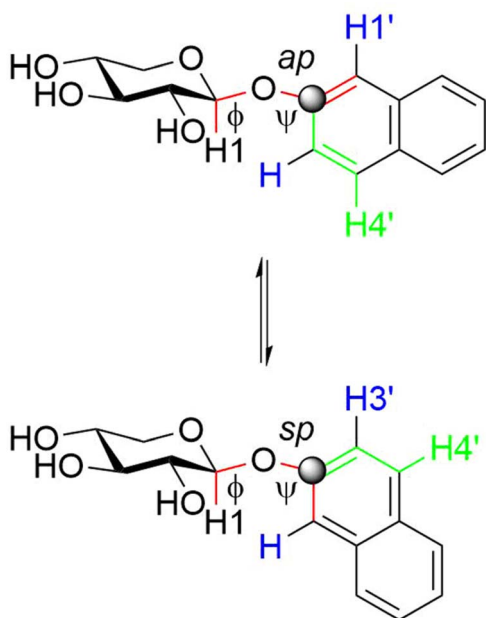


Fig. 2 Schematic of compound 1 illustrating  $^3J_{\text{CH}}$  coupling pathways from C2' to H1 and to H4'. In 1D  $^1\text{H}, ^1\text{H}$ -NOESY experiments irradiation at the H1 resonance frequency results in NOE peaks due to cross-relaxation to, *inter alia*, H1' and H3' (the latter two atoms are shown in blue color).

1D  $^1\text{H}, ^1\text{H}$ -NOESY buildup curves for compounds 1–5 were obtained by selective excitation of the anomeric proton and cross-relaxation rates were extracted using the PANIC approach;<sup>51</sup>

subsequent analysis using ISPA and quantum mechanical (QM) geometry optimized structures resulted in effective proton-proton distances (Table 1).

The glycosidic torsion angles of the geometry optimized compound 1 are quite similar to that of the crystal structure<sup>52</sup> of the compound (Table 1), *i.e.*,  $\phi$  has the *exo*-anomeric conformation and  $\psi$  an *sp* conformation. In this arrangement the solid state structure shows a short H1, H1' distance of 2.2 Å whereas a long H1, H3' distance of 4.3 Å is observed. In solution the population averaged distance of the former is slightly longer, 2.3 Å, whereas the latter is notably decreased to 2.9 Å. By noting that the barrier to rotation for an aryl methoxy group such as in anisole is low, on the order of 3 kcal mol<sup>-1</sup>,<sup>53</sup> and that the barrier to rotation of the naphthyl group at the  $\psi$  torsion angle has been estimated to be of similar magnitude<sup>47</sup> the rotation of the aglycone at  $\psi$  is considered to be unrestricted, but the conformational preference of the  $\psi$  torsion is still favoring an *sp* conformation, since the distance H1, H1' < H1, H3'. In the thioxyloside 2 and the sulfone derivative 3 the population-averaged distance from the anomeric proton to those adjacent to the linkage position C2' are similar, *i.e.*, H1, H1'  $\approx$  H1, H3' supporting unrestricted rotation and similar populations of *sp* and *ap* conformations at the  $\psi$  torsion angle. The (*S*)<sub>s</sub>- and (*R*)<sub>s</sub>-sulfinyl derivatives 4 and 5, respectively, contrast the conformational preference in 1; both have the distance for H1, H1' > H1, H3' indicating a small preference for the *ap* conformation at  $\psi$ .

$^1\text{H}, ^{13}\text{C}$ -Heteronuclear coupling constants were obtained by the 1DLR experiment. In the rigid planar naphthyl group

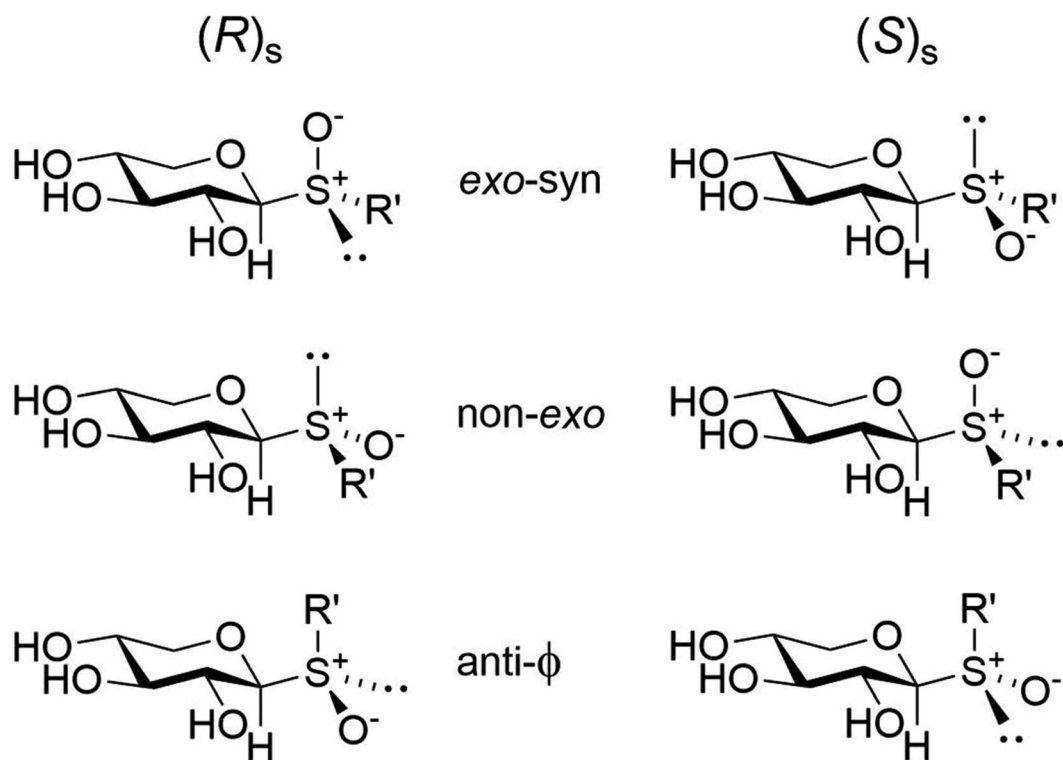


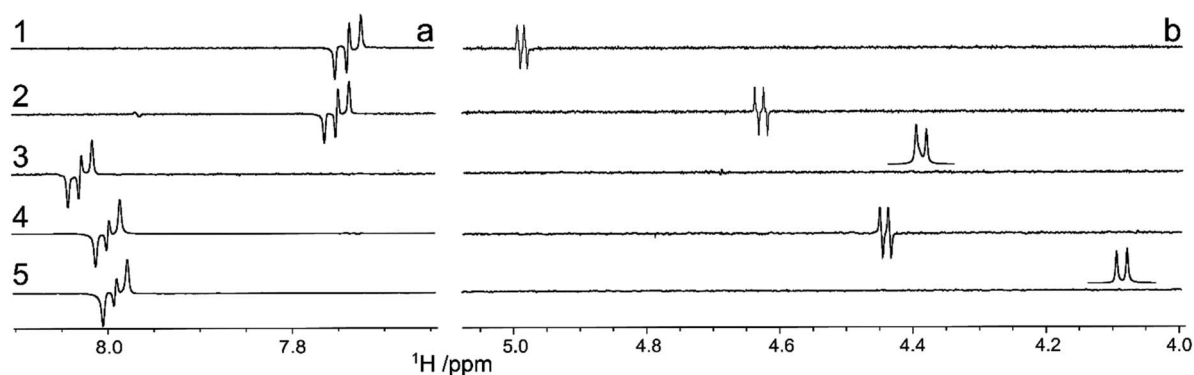
Fig. 3 Schematic of the three rotamers at the glycosidic  $\phi$  torsion angle for the diastereomeric xylosyl sulfoxides and pertinent description of their conformational states.



**Table 1** Proton–proton cross-relaxation rates ( $\sigma_{ij}$ ), experimentally derived distances ( $r_{\text{NMR}}$ ) and reference distances ( $r_{\text{ref}}$ ) from QM geometry optimized structures with the xylopyranoside in the  ${}^4\text{C}_1$  chair conformation for compounds 1–5 together with their glycosidic torsion angles

Proton pair		Compound				
		1	2	3	4	5
H1,H3	$\sigma_{\text{NMR}}/\text{s}^{-1}$		0.0298		0.0296	0.0345
	$r_{\text{ref}}/\text{\AA}$		2.62		2.64	2.70
H1–H5 <sub>pro-S</sub>	$\sigma_{\text{NMR}}/\text{s}^{-1}$	0.0295		0.0376		
	$r_{\text{ref}}/\text{\AA}$	2.44 (2.49) <sup>a</sup>		2.53		
H1,H1'	$\sigma_{\text{NMR}}/\text{s}^{-1}$	0.0434	0.0100	0.0071	0.0048	0.0093
	$r_{\text{NMR}}/\text{\AA}$	2.29 (2.21) <sup>a</sup>	3.14 (2.24) <sup>b</sup>	3.34	3.58	3.36
H1,H3'	$\sigma_{\text{NMR}}/\text{s}^{-1}$	0.0110	0.0076	0.0086	0.0087	0.0194
	$r_{\text{NMR}}/\text{\AA}$	2.88 (4.32) <sup>a</sup>	3.29 (4.76) <sup>b</sup>	3.24	3.24	2.97
<b>Torsion angle</b>						
$\phi$	H1–C1–X–C2'/ $^\circ$	42 (46) <sup>a</sup>	33	50	34	–63
$\psi$	C1–X–C2'–C1'/ $^\circ$	17 (1) <sup>a</sup>	40	96	87	–107

<sup>a,b</sup>Proton–proton distances from crystal structures of 1 and of 2-naphthyl 1,5-dithio- $\beta$ -D-xylopyranoside, respectively.<sup>68</sup>



**Fig. 4**  ${}^1\text{H}$ -detected spectra from 1D long-range experiments of compounds 1–5 where upon selective excitation at the C2' resonance frequency heteronuclear correlations are observed to (a) H4' in all compounds and (b) to H1 in compounds 1, 2 and 4. Inserts show the anomeric proton resonance for 3 and 5 at their respective chemical shift, since for these compounds the long-range correlations were not detected; consequently it was concluded that  ${}^3J_{\text{C2}',\text{H1}} < 0.5$  Hz and  ${}^3J_{\text{C2}',\text{H1}} < 0.6$  Hz, respectively.

${}^3J_{\text{C2}',\text{H4}'}$  coupling constants are large (Fig. 4a) and of similar magnitude in all five compounds (Table 2). The  ${}^3J_{\text{C2}',\text{H1}}$  in compound 1, as seen from the anti-phase peak separation in the NMR spectrum (Fig. 4b), is for a transglycosidic heteronuclear coupling constant of intermediate magnitude (Table 2) and the corresponding one in 2 is similar to that of 2-naphthyl 1,5-

dithio- $\beta$ -D-xylopyranoside, with a sulfur atom at the glycosidic linkage, being 4.1 Hz.<sup>52</sup> However, peaks from scalar coupled C2' and H1 nuclei are absent for the sulfone-containing compound 3 as well as for the (*R*)<sub>S</sub>-sulfoxide derivative 5, whereas in compound 4 having the (*S*)<sub>S</sub>-chirality at the sulfur atom the magnitude of the coupling constant is similar to those of 1 and 2 (Fig. 4b and Table 2). The absence of detectable  ${}^3J_{\text{C2}',\text{H1}}$  couplings in 3 and 5 makes it possible to determine an upper limit of these coupling constants. The natural line-width of these two samples under the experimental conditions employed was  $\nu_{1/2} \approx 1.0$  Hz. For an anti-phase peak a limiting separation is reached at  $\sim 0.576$  times the line-width beyond which the only effect of a smaller *J* coupling is to reduce the overall intensity of the anti-phase peak.<sup>54</sup> Thus, in compounds 3 and 5 the scalar interaction  ${}^3J_{\text{C2}',\text{H1}} < |0.5|$  Hz and  ${}^3J_{\text{C2}',\text{H1}} < |0.6|$  Hz, respectively (Table 2).

This conspicuous observation was further investigated by QM-based calculations of  ${}^3J_{\text{C2}',\text{H1}}$  as a function of the torsion angle  $\phi$ . The  ${}^3J_{\text{C2}',\text{H4}'}$  coupling constants in geometry optimized 1–5 were well reproduced at the level of theory used and so were

**Table 2** Experimental NMR heteronuclear  ${}^3J_{\text{CH}}$  (in Hz, absolute values) from 1DLR spectra and calculated using QM geometry optimized structures ( ${}^4\text{C}_1$ ) for compounds 1–5

Atom pair		Compound				
		1 <sup>a</sup>	2	3	4	5
C2',H4'	Expt	10.7	10.1	9.8	9.7	9.8
	Calc	11.2	10.8	10.9	10.2	10.1
C2',H1	Expt	3.6	4.2	<0.5	3.3	<0.6
	Calc	3.4	4.3	–0.8	3.4	0.4

<sup>a</sup>  ${}^3J_{\text{C2}',\text{H1}} = 3.3$  Hz from IPAP-selHSQMBC NMR experiments.



$^3J_{C2',H1}$  related to  $\phi$  (Table 2). Subsequent scanning of the  $\phi$  torsion angle in  $5^\circ$  increments and geometry optimization at each torsion angle followed by calculation of  $^3J_{C2',H1}$  resulted in

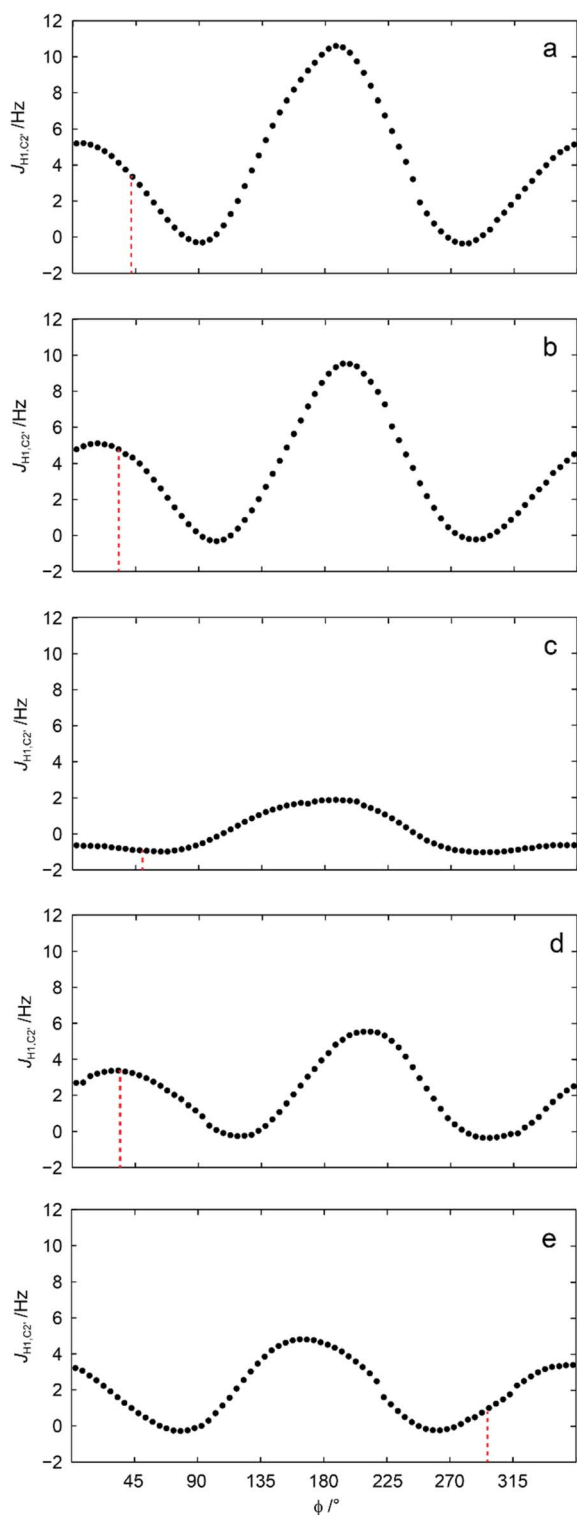


Fig. 5 Calculated heteronuclear  $^3J_{C2',H1}$  vs. the glycosidic  $\phi$  torsion angle for compounds 1–5 (a–e, respectively); the vertical dashed lines correspond to the  $\phi$  torsion angle of the geometry optimized structures described in Table 1. A conductor-like polarizable continuum model was used for methanol as solvent.<sup>69,70</sup>

computed values (Fig. 5) unveiling its conformational dependence, especially, that  $^3J_{C2',H1}$  is small for all conformations related to the torsion angle  $\phi$  in the sulfone derivative, *i.e.*, compound 3 (Fig. 5c).

### $\beta$ 4GalT7 assay

A truncated version of  $\beta$ 4GalT7 fused with glutathione *S*-transferase was expressed in *E. coli* and used in an enzymatic assay to measure the galactosylation by  $\beta$ 4GalT7.<sup>19</sup> Various concentrations of compounds 1–5 were incubated with the recombinant  $\beta$ 4GalT7 enzyme and the donor substrate, UDP-Gal, for 30 min at  $37^\circ\text{C}$ . The reaction progress was analyzed using HPLC with fluorescence detection and the identities of the galactosylated products were corroborated by LCMS analysis (Table S1†). The kinetic parameters were calculated using nonlinear regression to the Michaelis–Menten model (Table 3).

Between all substrates, the highest affinity was observed for analogs 3 and 5; compounds 2 and 4 showed an intermediate affinity, while the oxygen-containing compound 1 showed the lowest affinity, as measured by the  $K_m$  values. Interestingly, the (*S*)<sub>S</sub>-configuration sulfoxide 4 showed the highest turnover, as indicated by the  $k_{\text{cat}}$ , whereas the (*R*)<sub>S</sub>-sulfoxide 5 showed the lowest. Analog 1 showed similar  $k_{\text{cat}}$  as 5.

Moreover, the highest enzyme efficiency ( $k_{\text{cat}}/K_m$ ) was observed for sulfone 3 whereas the sulfoxides 4 and 5 showed very similar values. The lowest efficiency was present towards 1. In the case of analog 4 ((*S*)<sub>S</sub>-chirality) the turnover ( $k_{\text{cat}}$ ) was high whereas for analog 5 having the (*R*)<sub>S</sub> configuration the turnover was low, but the affinity high. The combined and potentially additive effects may thus result in stereochemical arrangements for donor molecule 3 in the active site of the enzyme thereby favoring the high efficiency.

### Molecular docking of xylosyl acceptors

Molecular docking was used to obtain further insights on atomic interactions in the binding pocket of  $\beta$ 4GalT7 and to shed light on the results of the biochemical assay. The protein crystal structures of *Drosophila*  $\beta$ -1,4-galactosyltransferase 7 mutant D211N (PDB ID: 4M4K and 4LW3) devoid of water molecules and xylose-containing acceptors but containing manganese and UDP-galactose were used for the computational studies as well as the enzyme containing the catalytic base

Table 3 Galactosylation of xylosides by  $\beta$ 4GalT7

Compound	$K_m/\text{mM}$	$V_{\text{max}}/\text{pmol s}^{-1}$	$k_{\text{cat}}/\text{s}^{-1}$	$k_{\text{cat}}/K_m/\text{mol}^{-1} \text{s}^{-1}$
1 <sup>a</sup>	0.70	1.6	1.2	1.6
2 <sup>a,b</sup>	0.30	1.7	1.8	5.8
3 <sup>b</sup>	0.10	2.1	2.2	21
4 ( <i>S</i> ) <sub>S</sub>	0.37	2.5	2.5	6.9
5 ( <i>R</i> ) <sub>S</sub>	0.14	1.1	1.1	7.8

<sup>a</sup> Slightly different kinetic parameters for 1 and 2 are obtained compared to previously published data since different batches of enzymes were used, which affect the parameters slightly. <sup>b</sup> Kinetic parameters were obtained for concentrations up to the highest observed reaction rate.



Asp211 instead of the mutant having Asn211. The previous molecular docking procedures resulted in identification of hydrogen bonding networks between the amino acids of the enzyme and the three hydroxyl groups in the xylose moiety, which was oriented above Tyr177, whereas the naphthyl aglycone adopted a face-to-face parallel stacking to Tyr179, and consistent with a product-forming nucleophilic attack from the acceptor molecule on the anomeric carbon of the nucleotide sugar the O4 atom of the xylose moiety was in proximity to the C1 atom of UDP-galactose.<sup>4,19,52</sup> In the present molecular modeling approach the number of sidechains treated as flexible was increased to encompass eight amino acids, *viz.*, K176, Y177, H178, Y179, L209, E210, N211/D211 and D212.

From the molecular docking of compounds 1–5 to the galactosyl transferase based on two crystal structures containing the donor UDP-galactose in the active site the binding affinities were calculated for the highest ranked poses for the mutant containing N211 and the wild type having instead the D211 residue (Table 4). Irrespective of the mutated amino acid 211 or the crystal structures the binding affinities showed a consistent pattern with the lowest affinity for XylONap (1), followed by an increased affinity for XylSNap (2), and further increased closely similar affinities for the sulfone XylSO<sub>2</sub>Nap (3), sulfoxide XylSOSNap (4), and XylSORNap (5) derivatives.

The valence angle C1–O–C2' at the glycosidic linkage is ~119° in XylONap (1), whereas for the sulfur-containing compounds it is lower, *viz.*, in XylSNap (2) the C1–O–C2' angle is ~103°, in XylSO<sub>2</sub>Nap (3) it is ~106°, and in 4 and 5 the valence angle is ~99°. The aglycon stacking to Y179 was larger for 1 with an angle in the range 8–14° compared to 1–7° in 2–5 for the four different systems, (*i.e.*, amino acid residues D211 and N211 in the protein and crystal structures 4LW3 and 4M4K), consistent with a more favorable stacking interaction (Fig. 6a) and higher binding affinity (Table 4). These results are in line with the  $K_m$  values measured, being highest for 1. Atom-atom interactions that differ between docked complexes are those between the hydrogen atom on the C6 carbon of the indole group of tryptophan 207 (W207). In the 4LW3-based complexes with either D211 or N211 the distance to the glycosidic oxygen in 1 is long being 3.6 Å, whereas it is shorter to the glycosidic sulfur atom in 2 being 2.9 Å, as well as to the oxygen atom at the glycosidic linkage of the (*S*)<sub>s</sub>-sulfinyl derivative 4 being 3.2–3.3 Å. Notably, the shortest distance of 2.4 Å is

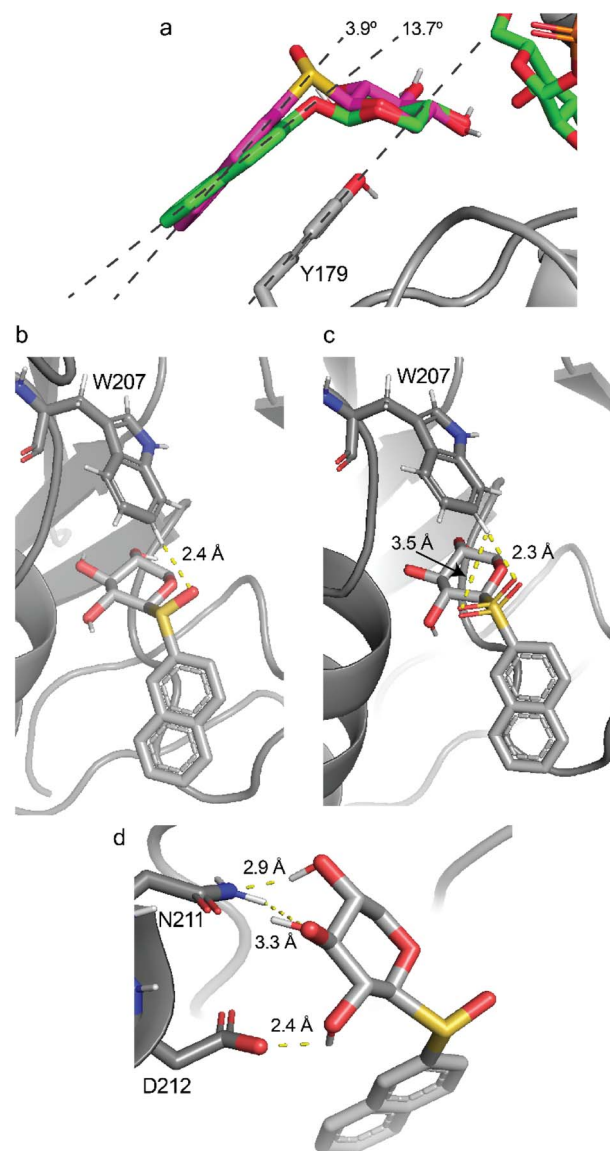


Fig. 6 Molecular docking complex of  $\beta$ 4GalT7 mutant 4LW3 N211 (a) XylONap (1) stacked to Y179, (b) (*R*)<sub>s</sub>-sulfoxide derivative 5 interaction with W207, (c) the corresponding (*R*)<sub>s</sub> oxygen atom in the sulfone-containing compound 3 interaction with W207. Molecular docking complex of  $\beta$ 4GalT7 mutant 4M4K N211, (d) compound 5 N211 interacts with O2 and O3 of the xylosyl residue and D212 interacts with O4 of the xylosyl residue.

Table 4 Binding affinities (kcal mol<sup>-1</sup>) of the highest ranked pose of compounds 1–5 docked using Auto-dock Vina to protein structures containing manganese and UDP-galactose based on the crystal structures from PDB entries 4M4K and 4LW3. The sidechains of K176, Y177, H178, Y179, L209, E210, N211/D211 and D212 were chosen as being flexible

Ligand	4M4K-N211	4M4K-D211	4LW3-N211	4LW3-D211
1 XylONap	-7.9	-8.0	-7.9	-8.0
2 XylSNap	-8.3	-8.2	-8.3	-8.3
3 XylSO <sub>2</sub> Nap	-8.5	-8.6	-8.5	-8.5
4 XylSOSNap	-8.5	-8.5	-8.5	-8.5
5 XylSORNap	-8.5	-8.6	-8.5	-8.5

observed for the oxygen atom of (*R*)<sub>s</sub>-sulfoxide derivative 5 (Fig. 6b) and the corresponding (*R*)<sub>s</sub> oxygen atom in the sulfone-containing compound 3, being 2.3 Å (Fig. 6c) and 2.4 Å in the D211 or N211 complexes, respectively. These short distances in 3 and 5 are contrasted with long distances between the hydrogen atom on the C6 carbon of the oxygen atom at the glycosidic linkage of the (*S*)<sub>s</sub>-sulfinyl derivative 4 and the corresponding (*S*)<sub>s</sub> oxygen atom in the sulfone-containing compound 3, in a range of 3.2–3.5 Å. The corresponding distances in the 4M4K-based complexes were similar or slightly longer by a few tens of an Ångström. Thus, the interatomic distances in the docked complexes of 3.6, ~3.0 and 2.4 Å



correlate with the  $K_m$  values of 0.70 (observed for **1**),  $\sim 0.35$  (**2** and **4**) and  $\sim 0.10$  mM (**3** and **5**) respectively, suggesting that an additional atom–atom interaction may be of importance in determining the binding affinity. These findings are consistent with the presence of a weak C–H $\cdots$ O hydrogen bond,<sup>55,56</sup> with a favorable interaction energy of  $<1$  kcal mol<sup>-1</sup> between W207 and a suitably oriented oxygen atom of the ligand. The less favorable interactions for **1** may consequently lead to a lower efficiency of the enzyme with the glycosidic oxygen-containing compound as measured by  $k_{cat}/K_m$  (Table 3). We also illustrate interatomic interactions in one of the complexes (N211/4M4K) where for compound **5** D212 interacts with O2 of the xylosyl residue and N211 interacts with O3 and O4 of the xylosyl residue, with short distances of 2.4, 3.3 and 2.9 Å, respectively (Fig. 6d), similar to those observed in the crystal structure being 2.7, 3.2 and 2.7 Å, respectively.<sup>3</sup>

## Conclusions

The diastereomeric pair of 2-naphthyl  $\beta$ -D-xylopyranoside sulfonides were synthesized, purified and characterized, in particular by NMR spectroscopy and quantum mechanical calculations, identifying them as having the (*R*)<sub>S</sub>- and (*S*)<sub>S</sub>-configuration at the sulfur atom for compounds **5** and **4**, respectively. Transglycosidic coupling constants at  $\phi$  in oligosaccharides often show  $^3J_{CH} \approx 4$  Hz<sup>57</sup> as a result of having an *exo*-anomeric conformation at this torsion angle, also observed for the *O*- and *thio*-glycosides **1** and **2**, respectively, as well as for the (*S*)<sub>S</sub>-configured sulfoxide **4**. The glycosidic torsion angles in these QM geometry optimized structures have  $\phi \approx 40^\circ$  and computed coupling constants revealed  $^3J_{C2',H1} \approx 4$  Hz for the conformers. However, the magnitude of  $^3J_{C2',H1}$  was small for all conformations of the sulfone derivative (**3**), and not detectable by the 1DLR NMR experiment, even though the glycosidic conformation at  $\phi = 50^\circ$  of the low-energy conformer. Furthermore, the  $^3J_{C2',H1}$  coupling was not detectable in (*R*)<sub>S</sub>-configured sulfoxide **5**, for which the low-energy conformer had a *gauche*<sup>-</sup> conformation at the  $\phi$  torsion angle. The importance of stereoelectronic effects in sulfide-, sulfoxide- and sulfone-containing compounds<sup>58,59</sup> on conformational preferences and the observed differences in  $^3J_{CH}$  coupling constants on these glycosides and similar compounds await further detailed studies to clarify the relative importance of steric and electronic effects.

The activity of  $\beta$ 4GalT7 in galactosylation reactions was analyzed and the reaction progress was monitored by identification of the products containing a galactosyl residue. The highest affinity was observed for the (*R*)<sub>S</sub>-configured compound and the sulfone derivative and the highest enzyme efficiency ( $k_{cat}/K_m$ ) was observed for sulfone-containing compound. Molecular docking of the acceptor substrates to the UDP-galactose: $\beta$ 4GalT7 complex identified aglycon stacking to Y179 as being more favorable in the derivatives of the 2-naphthyl xylosides than in the parent compound. Furthermore, the binding affinity correlated with the stacking to the tyrosine residue and a weak C–H $\cdots$ O hydrogen bond between the indole group of a tryptophan residue. To elucidate if an oxygen atom of sulfone

and sulfinyl derivatives of these 2-naphthyl xylosides is geometrically arranged to participate in a C–H $\cdots$ O hydrogen bond to the enzyme, X-ray crystallography studies of these derivatives and  $\beta$ 4GalT7 would be of great interest to perform in future studies, thereby facilitating additional specific information on protein:ligand complexes in  $\beta$ 4GalT7,<sup>60,61</sup> as a stepping-stone to further studies on  $\beta$ -1,4-GalT from bovine milk<sup>62</sup> or on the human galactosyltransferase  $\beta$ 3GalT5.<sup>63</sup>

## Materials and methods

### Synthesis

All moisture- and air-sensitive reactions were carried out under an atmosphere of dry nitrogen using oven-dried glassware. Solvents were dried prior to use. Purchased reagents were used without further purification. Organic phases were dried using Biotage ISOLUTE Phase separators. Chromatographic separations were performed on a Biotage Isolera One flash purification system using Biotage SNAP KP-Sil silica cartridges. Thin-layer chromatography was performed on precoated TLC alumina plates coated with silica gel 60 F<sub>254</sub> 0.25 mm (Merck). Spots were visualized with UV light or by staining with *para*-anisaldehyde. NMR spectra of synthetic intermediates were recorded at room temperature using a Bruker Avance II 400 MHz spectrometer in CDCl<sub>3</sub> from which chemical shifts (in ppm) were referenced to residual  $\delta_H$  7.26 and to  $\delta_C$  77.16. Mass spectra were recorded on Micromass Q-TOF micro<sup>TM</sup>. Synthesis and physical characterization of compounds **1**,<sup>46</sup> **2** (ref. 19) and **3** (ref. 19) have been previously reported.

### 2-Naphthyl 1-thio- $\beta$ -D-xylopyranoside (*S*)<sub>S</sub>-oxide **4** and 2-naphthyl 1-thio- $\beta$ -D-xylopyranoside (*R*)<sub>S</sub>-oxide **5**

2-Naphthyl 2,3,4-tri-*O*-acetyl-1-thio- $\beta$ -D-xylopyranoside (**6**) (307 mg, 0.734 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (6 mL) and cooled to  $-78^\circ\text{C}$ . *m*CPBA (180 mg, 0.78 mmol) dissolved in CH<sub>2</sub>Cl<sub>2</sub> (3 mL) was added dropwise and after 2.5 h, the reaction mixture was allowed to reach rt. Satd aq NaHCO<sub>3</sub> (10 mL) was added and the organic phase was dried before concentrated under reduced pressure. Column chromatography (SiO<sub>2</sub>, 25  $\rightarrow$  50% EtOAc in heptane) gave a mixture of the diastereomers (*R*)<sub>S</sub>-**7** and (*S*)<sub>S</sub>-**7**. The ratio between the major and the minor isomer was 7 : 3, based on the <sup>1</sup>H NMR spectrum. Evaporation of the solvent furnished the products as a white solid (304 mg, 95%). HRMS calcd for C<sub>21</sub>H<sub>22</sub>O<sub>8</sub>SNa [M + Na]<sup>+</sup>: 457.0933; found: 457.0930. Major isomer (relative intensity 0.70): <sup>1</sup>H NMR:  $\delta$  8.18 (1H, H-1'), 7.97–7.90 (m, 3H, H-4', H-5', H-7'), 7.64 (dd,  $J_{H3',H4'}$  8.64,  $J_{H3',H1'}$  1.78, 1H, H-3'), 7.59–7.58 (m, 2H, H-8', H-6'), 5.34 (dd,  $J_{H1,H2}$  8.26,  $J_{H2,H3}$  8.04, 1H, H-2), 5.24 (dd,  $J_{H2,H3}$  8.04,  $J_{H3,H4}$  8.35, 1H, H-3), 4.85 (ddd,  $J_{H3,H4}$  8.35,  $J_{H4,H5eq}$  4.89,  $J_{H4,H5ax}$  8.49, 1H, H-4), 4.57 (d,  $J_{H1,H2}$  8.26, 1H, H-1), 4.25 (dd,  $J_{H4,H5eq}$  4.89,  $J_{H5ax,H5eq}$   $-11.76$ , 1H, H-5<sub>eq</sub>), 3.48 (dd,  $J_{H5ax,H4}$  8.49,  $J_{H5ax,H5eq}$   $-11.76$ , 1H, H-5<sub>ax</sub>), 2.04, 2.00, 1.98 (3 s, 9H, CH<sub>3</sub>). <sup>13</sup>C NMR:  $\delta$  170.00, 169.75, 169.28 (3 CO), 136.64, 134.79, 132.76 (3 Ar), 129.22 (C-4'), 128.76 (C-7'), 128.22 (C-8'), 128.11 (C-5'), 127.45 (C-6'), 126.22 (C-1'), 93.43 (C-1), 71.79 (C-3), 67.96 (C-4), 66.50 (C-5), 66.29 (C-2), 20.75, 20.74, 20.51 (3 Me). Minor isomer (relative



intensity 0.30):  $^1\text{H}$  NMR:  $\delta$  8.18 (1H, H-1'), 7.94–7.90 (m, 3H, H-4', H-5', H-7'), 7.60–7.58 (m, 3H, H-3', H-6', H-8'), 5.37 (dd,  $J_{\text{H}2,\text{H}1}$  9.24,  $J_{\text{H}2,\text{H}3}$  8.84, 1H, H-2), 5.26 (dd,  $J_{\text{H}3,\text{H}2}$  8.84,  $J_{\text{H}3,\text{H}4}$  8.87, 1H, H-3), 4.98 (ddd,  $J_{\text{H}4,\text{H}3}$  8.87,  $J_{\text{H}4,\text{H}5\text{eq}}$  5.26,  $J_{\text{H}4,\text{H}5\text{ax}}$  9.26, 1H, H-4), 4.32 (d,  $J_{\text{H}1,\text{H}2}$  9.24, 1H, H-1), 4.19 (dd,  $J_{\text{H}5\text{eq},\text{H}4}$  5.26,  $J_{\text{H}5\text{eq},\text{H}5\text{ax}}$  –11.58, 1H, H-5<sub>eq</sub>), 3.28 (dd,  $J_{\text{H}5\text{ax},\text{H}4}$  9.26,  $J_{\text{H}5\text{ax},\text{H}5\text{eq}}$  –11.58, 1H, H-5<sub>ax</sub>), 2.04, 2.00, 1.98 (3 s, 9H, CH<sub>3</sub>).  $^{13}\text{C}$  NMR:  $\delta$  170.31, 169.63, 169.12 (3 CO), 136.20, 134.84, 132.83 (3 Ar), 129.35 (C-4'), 128.72 (C-7'), 128.19 (C-5'), 128.15 (C-8'), 127.45 (C-6'), 126.54 (C-1'), 120.77 (C-3'), 91.78 (C-1), 72.85 (C-3), 68.34 (C-4), 67.35 (C-2), 67.11 (C-5), 20.79, 20.71, 20.66 (3 Me).

The mixture of compounds (*R*)<sub>5</sub>-7 and (*S*)<sub>5</sub>-7 (152 mg, 0.349 mmol) was dissolved in a solution of guanidine/guanidinium nitrate<sup>36</sup> in MeOH/CH<sub>2</sub>Cl<sub>2</sub> (9 : 1, 25 mL) at room temperature and after 30 min, the mixture was neutralized with AcOH and concentrated under reduced pressure. Column chromatography (SiO<sub>2</sub>, 4% MeOH in CH<sub>2</sub>Cl<sub>2</sub>), in which the major product eluted prior to the minor product, gave **4** (69 mg, 64%) and **5** (23 mg, 21%) as white solids. Compound **4**: HRMS calcd for C<sub>15</sub>H<sub>16</sub>O<sub>5</sub>SNa [M + Na]<sup>+</sup>: 331.0616; found: 331.0615; compound **5**: HRMS calcd for C<sub>15</sub>H<sub>16</sub>O<sub>5</sub>SNa [M + Na]<sup>+</sup>: 331.0616; found: 331.0618.  $^{13}\text{C}$  and  $^1\text{H}$  NMR data ( $\delta_{\text{C}}/\delta_{\text{H}}$ ) of **4** (methanol-*d*<sub>4</sub>): 137.13 (C-2'), 136.46, 134.01, 130.06/8.00 (C-4'/H-4'), 129.72/7.95, 129.30/7.59, 129.16/7.92, 128.42/7.57, 128.22/8.19 (C-1'/H-1'), 122.83/7.73 (C-3'/H-3'), 96.56/4.43 (C-1/H-1), 78.37/3.38 (C-3/H-3), 71.30/3.33 (C-2/H-2), 71.15/3.95/3.25 (C-5/H-5<sub>pro-R</sub>/H-5<sub>pro-S</sub>), 70.37/3.23 (C-4/H-4).  $^{13}\text{C}$  and  $^1\text{H}$  NMR data ( $\delta_{\text{C}}/\delta_{\text{H}}$ ) of **5** (methanol-*d*<sub>4</sub>): 137.66 (C-2'), 136.14, 134.32, 130.14/7.98 (C-4'/H-4'), 129.49/7.93, 129.14/7.91, 129.03/7.57, 128.38/7.55, 127.05/8.14 (C-1'/H-1'), 121.98/7.62 (C-3'/H-3'), 96.06/4.09 (C-1/H-1), 79.26/3.39 (C-3/H-3), 71.34/3.79/3.02 (C-5/H-5<sub>pro-R</sub>/H-5<sub>pro-S</sub>), 70.54/3.47 (C-4/H-4), 70.07/3.75 (C-2/H-2).

## UV and CD spectroscopy

Ultraviolet (UV) and circular dichroism (CV) spectra of the sulfoxide-containing compounds **4** and **5** were recorded on an Applied Photophysics Chirascan instrument at 20.0 °C with the wavelength range set to 260 to 190 nm.

## NMR spectroscopy for conformational analysis

NMR experiments for conformational analysis were performed on the following spectrometers: a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm PFG triple resonance probe, a Bruker AVANCE 500 MHz spectrometer and a Bruker AVANCE III 700 MHz spectrometer; the latter two equipped with 5 mm TCI Z-Gradient CryoProbes.  $^1\text{H}$  NMR chemical shifts were referenced to external 3-trimethylsilyl-(2,2,3,3-<sup>2</sup>H<sub>4</sub>)-propanoate (TSP,  $\delta_{\text{H}}$  0.00) in D<sub>2</sub>O and  $^{13}\text{C}$  to external dioxane ( $\delta_{\text{C}}$  67.40) in D<sub>2</sub>O. The temperature was set to 37 °C with a neat deuterated methanol sample.<sup>64</sup>

For NMR experiments compounds **1**–**5** (5 mg) were dissolved in methanol-*d*<sub>4</sub> (0.5 mL) and transferred to 5 mm NMR tubes.  $^1\text{H}$  NMR spectra were recorded with 24k data points over a 12 ppm spectral width and using 400 scans.  $^{13}\text{C}$  NMR spectra were recorded using 65k data points over a 197 ppm spectral

width, zero-filled to 262k points and a 3 Hz line broadening was applied prior to Fourier transformation.

$^1\text{H}$  NMR chemical shifts were assigned using  $^1\text{H}$ ,  $^1\text{H}$ -TOCSY experiments (mixing times of 10, 40, 60 and 80 ms) and a  $^1\text{H}$ ,  $^1\text{H}$ -NOESY experiment (mixing time of 300 ms) with 2048 × 256 data points over a spectral width of 5 ppm.  $^{13}\text{C}$  NMR chemical shifts assignments were performed using multiplicity-edited  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC,  $^1\text{H}$ ,  $^{13}\text{C}$ -HMBC experiments.  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC experiments were recorded with 1024 × 512 data points and a spectral width of 10 and 120 ppm for  $^1\text{H}$  and  $^{13}\text{C}$ , respectively. The  $^1\text{H}$ ,  $^{13}\text{C}$ -HMBC experiments were carried out using 2048 × 256 data points over 10 and 120 ppm for  $^1\text{H}$  and  $^{13}\text{C}$ , respectively.

Selective 1D  $^1\text{H}$ ,  $^1\text{H}$ -NOESY experiments<sup>42</sup> and selective 1DLR experiments<sup>43,44</sup> were performed to study the torsional angle preferences at the glycosidic linkage in **1**–**5**, where  $\phi = \text{H}1-\text{C}1-\text{X}2'-\text{C}2'$  and  $\psi = \text{C}1-\text{X}2'-\text{C}2'-\text{C}1'$  with X = O in **1** or S in **2**–**5**. 1D  $^1\text{H}$ ,  $^1\text{H}$ -NOESY experiments were acquired at 700 MHz for compounds **1**–**3** and at 600 MHz for compounds **4** and **5** with 8k data points, a spectral width of 8 ppm and a relaxation delay of 2 s using eight different mixing times varying from 60 to 600 ms. Selective 1DLR experiments were acquired on a 700 MHz Bruker AVANCE III spectrometer with 50k data points, 11 200 scans and a spectral width of 12 ppm. The nominal value of the long-range coupling constant was set to 8 Hz and a 160 ms selective  $^{13}\text{C}$  excitation pulse (Gauss1\_90.1000) centered at the chemical shift of the C2' resonance. Heteronuclear *J* coupling constants were extracted using the *J*-doubling method<sup>65</sup> by deconvolution of the anti-phase peaks in the 1DLR spectra. The heteronuclear  $^3J_{\text{C}2',\text{H}1}$  of compound **1** was also determined from 2D IPAP-selHSQMBC NMR experiments<sup>66</sup> using a frequency-selective Gaussian shaped pulse with a duration of 40 ms, a nominal value of 5 Hz for the delay of the evolution of the long-range coupling and an acquisition time of 1.5 s. The numerical value of  $^3J_{\text{C}2',\text{H}1}$  was extracted from IP and AP peak separations in 1D spectra extracted from rows of the 2D NMR spectra.

## Quantum mechanical calculations

Quantum mechanical geometry optimizations and calculations of NMR scalar spin–spin coupling constants were carried out as previously described.<sup>32</sup>

## Molecular docking simulations of the Michaelis complex

The geometries obtained for compounds **1**–**5** from the QM calculations carried out herein were used as initial input for the docking studies. Molecular docking with Auto-dock VINA was carried out essentially as previously described.<sup>19</sup> The protein crystal structures of *Drosophila*  $\beta$ -1,4-galactosyltransferase 7 mutant D211N (PDB ID: 4M4K and 4LW3) were used to prepare complexes devoid of water molecules and xylose-containing acceptors but containing manganese and UDP-galactose. Furthermore, corresponding molecular models were created containing instead the catalytic base Asp211, referred to as D211. In the docking protocol used the sidechains of K176, Y177, H178, Y179, L209, E210, N211/D211 and D212 were chosen as being flexible. The top-ranked poses from the 20



docking simulations were analyzed in further detail with respect to atom–atom interactions and their geometrical arrangements.

### Biological testing

The  $\beta$ 4GalT7 assay was carried out according to protocols described in detail previously.<sup>4,67</sup>

### Data availability

The datasets supporting this article have been uploaded as part of the ESI.†

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

This work was supported by grants from the Swedish Research Council (no. 2022-03014) and the Knut and Alice Wallenberg Foundation, the Carl Trygger Foundation, the Crafoord Foundation, Lars Hiertas Memorial Foundation, the Magnus Bergwall Foundation, Lund University, and the Royal Physiographic Society in Lund. Dr J. Jarvet is thanked for recording UV and CD spectra.

### References

- R. Almeida, S. B. Levery, U. Mandel, H. Kresse, T. Schwientek, E. P. Bennett and H. Clausen, *J. Biol. Chem.*, 1999, **274**, 26165–26171.
- B. Ramakrishnan and P. K. Qasba, *J. Biol. Chem.*, 2010, **285**, 15619–15626.
- Y. Tsutsui, B. Ramakrishnan and P. K. Qasba, *J. Biol. Chem.*, 2013, **288**, 31963–31970.
- A. Siegbahn, S. Manner, A. Persson, E. Tykesson, K. Holmqvist, A. Ochocinska, J. Rönnols, A. Sundin, K. Mani, G. Westergren-Thorsson, G. Widmalm and U. Ellervik, *Chem. Sci.*, 2014, **5**, 3501–3508.
- S. O. Kolset, K. Sakurai, I. Ivhed, A. Overvatn and S. Suzuki, *Biochem. J.*, 1990, **265**, 637–645.
- J. R. Bishop, M. Schuksz and J. D. Esko, *Nature*, 2007, **446**, 1030–1037.
- N. B. Schwartz and M. S. Domowicz, *FEBS Lett.*, 2018, **592**, 3791–3805.
- R. V. Iozzo, *Annu. Rev. Biochem.*, 1998, **67**, 609–652.
- F. H. Blackhall, C. L. R. Merry, E. J. Davies and G. C. Jayson, *Br. J. Cancer*, 2001, **85**, 1094–1098.
- R. V. Iozzo and R. D. Sanderson, *J. Cell. Mol. Med.*, 2011, **15**, 1013–1031.
- I. J. Edwards, *Nat. Rev. Urol.*, 2012, **9**, 196–206.
- A. D. Theocharis, S. S. Skandalis, T. Neill, H. A. B. Mulhaupt, M. Hubo, H. Frey, S. Gopal, A. Gomes, N. Afratis, H. C. Lim, J. R. Couchman, J. Filmus, R. D. Sanderson, L. Schaefer, R. V. Iozzo and N. K. Karamanos, *Biochim. Biophys. Acta, Rev. Cancer*, 2015, **1855**, 276–300.
- S. Pellegatta, B. Savoldo, N. Di Ianni, C. Corbetta, Y. Chen, M. Patané, C. Sun, B. Pollo, S. Ferrone, F. DiMeco, G. Finocchiaro and G. Dotti, *Sci. Transl. Med.*, 2018, **10**, eaao2731.
- E. Ampofo, B. M. Schmitt, M. D. Menger and M. W. Laschke, *Cell. Mol. Biol. Lett.*, 2017, **22**, 4.
- S. A. Ibrahim, R. Gadalla, E. A. El-Ghonaimy, O. Samir, H. T. Mohamed, H. Hassan, B. Greve, M. El-Shinawi, M. M. Mohamed and M. Götte, *Mol. Cancer*, 2017, **16**, 57.
- K. Thorsheim, A. Siegbahn, R. E. Johnsson, H. Stålbrand, S. Manner, G. Widmalm and U. Ellervik, *Carbohydr. Res.*, 2015, **418**, 65–88.
- N. B. Schwartz, L. Galligani, P.-L. Ho and A. Dorfman, *Proc. Natl. Acad. Sci. U. S. A.*, 1974, **71**, 4047–4051.
- H. C. Robinson, M. J. Brett, P. J. Tralaggan, D. A. Lowther and M. Okayama, *Biochem. J.*, 1975, **148**, 25–34.
- A. Siegbahn, K. Thorsheim, J. Stähle, S. Manner, C. Hamark, A. Persson, E. Tykesson, K. Mani, G. Westergren-Thorsson, G. Widmalm and U. Ellervik, *Org. Biomol. Chem.*, 2015, **13**, 3351–3362.
- H. C. Robinson and U. Lindahl, *Biochem. J.*, 1981, **194**, 575–586.
- S. O. Kolset, K. Sakurai, I. Ivhed, A. Overvatn and S. Suzuki, *Biochem. J.*, 1990, **265**, 637–645.
- T. A. Fritz, F. N. Lugemwa, A. K. Sarkar and J. D. Esko, *J. Biol. Chem.*, 1994, **269**, 300–307.
- M. Jacobsson, K. Mani and U. Ellervik, *Bioorg. Med. Chem.*, 2007, **15**, 5283–5299.
- K. Holmqvist, A. Persson, R. Johnsson, J. Löfgren, K. Mani and U. Ellervik, *Bioorg. Med. Chem.*, 2013, **21**, 3310–3317.
- J. Malmberg, K. Mani, E. Säwén, A. Wirén and U. Ellervik, *Bioorg. Med. Chem.*, 2006, **14**, 6659–6665.
- M. Sobue, H. Habuchi, K. Ito, H. Yonekura, K. Oguri, K. Sakurai, S. Kamohara, Y. Ueno, R. Noyori and S. Suzuki, *Biochem. J.*, 1987, **241**, 591–601.
- L. Wang, C. A. Maniglia, S. L. Mella and A. C. Sartorelli, *J. Med. Chem.*, 1983, **26**, 629–632.
- B. Kuberan, M. Ethirajan, X. V. Victor, V. Tran, K. Nguyen and A. Do, *ChemBioChem*, 2008, **9**, 198–200.
- X. V. Victor, T. K. N. Nguyen, M. Ethirajan, V. M. Tran, K. V. Nguyen and B. Kuberan, *J. Biol. Chem.*, 2009, **284**, 25842–25853.
- R. Bentley, *Chem. Soc. Rev.*, 2005, **34**, 609–624.
- D. Crich, J. Mataka, S. Sun, D. J. Wink, K.-C. Lam and A. L. Rheingold, *Chem. Commun.*, 1998, 2763–2764.
- N. Khiar, *Tetrahedron Lett.*, 2000, **41**, 9059–9063.
- N. Khiar, I. Fernández, C. S. Araújo, J.-A. Rodríguez, B. Suárez and E. Álvarez, *J. Org. Chem.*, 2003, **68**, 1433–1442.
- J. P. Colomer, V. E. Manzano and O. Varela, *Eur. J. Org. Chem.*, 2013, 7343–7353.
- T. Taniguchi, M. Asahata, A. Nasu, Y. Shichibu, K. Konishi and K. Monde, *Cirality*, 2016, **28**, 534–539.
- U. Ellervik and G. Magnusson, *Tetrahedron Lett.*, 1997, **38**, 1627–1628.
- D. Crich, J. Mataka, L. N. Zakharov, A. L. Rheingold and D. J. Wink, *J. Am. Chem. Soc.*, 2002, **124**, 6028–6036.



- 38 C. A. Sanhueza, R. L. Dorta and J. T. Vázquez, *J. Org. Chem.*, 2011, **76**, 7769–7780.
- 39 C. A. Sanhueza, R. L. Dorta and J. T. Vázquez, *Tetrahedron:Asymmetry*, 2013, **24**, 582–593.
- 40 G. F. Pauli, S.-N. Chen, D. C. Lankin, J. Bisson, R. J. Case, L. R. Chadwick, T. Gödecke, T. Inui, A. Kronic, B. U. Jaki, J. B. McAlpine, S. Mo, J. G. Napolitano, J. Orjala, J. Lehtivarjo, S.-P. Korhonen and M. Niemitz, *J. Nat. Prod.*, 2014, **77**, 1473–1487.
- 41 S. L. Patt and B. D. Sykes, *J. Chem. Phys.*, 1972, **56**, 3182–3184.
- 42 K. Stott, J. Keeler, Q. N. Van and A. J. Shaka, *J. Magn. Reson.*, 1997, **125**, 302–324.
- 43 T. Nishida, G. Widmalm and P. Sandor, *Magn. Reson. Chem.*, 1995, **33**, 596–599.
- 44 T. Nishida, G. Widmalm and P. Sandor, *Magn. Reson. Chem.*, 1996, **34**, 377–382.
- 45 P. D. Thomas, V. J. Basus and T. L. James, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 1237–1241.
- 46 A. Siegbahn, U. Aili, A. Ochocinska, M. Olofsson, J. Rönnols, K. Mani, G. Widmalm and U. Ellervik, *Bioorg. Med. Chem.*, 2011, **19**, 4114–4126.
- 47 P. N. Goulart, C. O. da Silva and G. Widmalm, *J. Phys. Org. Chem.*, 2017, **30**, e3708.
- 48 E. Säwén, T. Massad, C. Landersjö, P. Damberg and G. Widmalm, *Org. Biomol. Chem.*, 2010, **8**, 3684–3695.
- 49 I. Tvaroska, K. Mazeau, M. Blanc-Muesser, S. Lavaitte, H. Driguez and F. R. Taravel, *Carbohydr. Res.*, 1992, **229**, 225–231.
- 50 M. Yang, T. Angles d'Ortoli, E. Säwén, M. Jana, G. Widmalm and A. D. MacKerell Jr, *Phys. Chem. Chem. Phys.*, 2016, **18**, 18776–18794.
- 51 S. Macura, B. T. Farmer II and L. R. Brown, *J. Magn. Reson.*, 1986, **70**, 493–499.
- 52 K. Thorsheim, D. Willén, E. Tykesson, J. Stähle, J. P. Praly, S. Vidal, M. T. Johnson, G. Widmalm, S. Manner and U. Ellervik, *Chem.–Eur. J.*, 2017, **23**, 18057–18065.
- 53 T. B. Grindley, A. R. Katritzky and R. D. Topsom, *J. Chem. Soc., Perkin Trans. 2*, 1974, 289–293.
- 54 D. Neuhaus, G. Wagner, M. Vasak, J. H. R. Kägi and K. Wüthrich, *Eur. J. Biochem.*, 1985, **151**, 257–273.
- 55 Z. S. Derewenda, L. Lee and U. Derewenda, *J. Mol. Biol.*, 1995, **252**, 248–262.
- 56 S. Scheiner, T. Kar and J. Pattanayak, *J. Am. Chem. Soc.*, 2002, **124**, 13257–13264.
- 57 T. Rundlöf, A. Kjellberg, C. Damberg, T. Nishida and G. Widmalm, *Magn. Reson. Chem.*, 1998, **36**, 839–847.
- 58 E. Juaristi, R. Martinez, R. Mendez, R. A. Toscano, M. Soriano-Garcia, E. L. Eliel, A. Petsom and R. S. Glass, *J. Org. Chem.*, 1987, **52**, 3806–3811.
- 59 T. Wedel, M. Müller, J. Podlech, H. Goesmann and C. Feldmann, *Chem.–Eur. J.*, 2007, **13**, 4273–4281.
- 60 M. Saliba, N. Ramalanjaona, S. Gulberti, I. Bertin-Jung, A. Thomas, S. Dahbi, C. Lopin-Bon, J.-C. Jacquinet, C. Breton, M. Ouzzine and S. Fournel-Gigleux, *J. Biol. Chem.*, 2015, **290**, 7658–7670.
- 61 S. Dahbi, J.-C. Jacquinet, I. Bertin-Jung, A. Robert, N. Ramalanjaona, S. Gulberti, S. Fournel-Gigleux and C. Lopin-Bon, *Org. Biomol. Chem.*, 2017, **15**, 9653–9669.
- 62 J. Jiang and G. K. Wagner, *Carbohydr. Res.*, 2017, **450**, 54–59.
- 63 J. M. Lo, C.-C. Kung, T.-J. R. Cheng, C.-H. Wong and C. Ma, *J. Am. Chem. Soc.*, 2025, **147**, 10875–10885.
- 64 M. Findeisen, T. Brand and S. Berger, *Magn. Reson. Chem.*, 2007, **45**, 175–178.
- 65 F. del Río-Portilla, V. Blechta and R. Freeman, *J. Magn. Reson., Ser. A*, 1994, **111**, 132–135.
- 66 J. Saurí, P. Nolis and T. Parella, *Magn. Reson. Chem.*, 2020, **58**, 363–375.
- 67 A. Siegbahn, K. Thorsheim, J. Stähle, S. Manner, C. Hamark, A. Persson, E. Tykesson, K. Mani, G. Westergren-Thorsson, G. Widmalm and U. Ellervik, *Org. Biomol. Chem.*, 2015, **13**, 3351–3362.
- 68 K. Thorsheim, D. Willén, E. Tykesson, J. Stähle, J. P. Praly, S. Vidal, M. T. Johnson, G. Widmalm, S. Manner and U. Ellervik, *Chem.–Eur. J.*, 2017, **23**, 18057–18065.
- 69 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
- 70 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.

