CrossMark
click for updates

Glycosyldiselenides as lectin ligands detectable by NMR in biofluids†

Ignacio Pérez-Victoria,‡ Omar Boutureira,‡ Tim D. W. Claridge* and Benjamin G. Davis*

The ability of glycosyldiselenides to act as lectin ligands and their selective detection in plasma by ^{77}Se NMR is reported.

Glycosyldisulfides are potentially interesting non-hydrolysable oligosaccharide mimetics that have attracted the attention of the Chemical Glycobiology community.[1] The synthesis of glycosyl disulfides has been motivated not only to make effective glycosyl donors[2] or cytotoxic agents[3] but also to exploit their synthetic flexibility in the development of site-selective protein modification methods.[4] Additionally, the reversible formation of disulfide linkages has made possible the discovery of new lectin binders through the generation and screening of dynamic libraries prepared in the presence of protein receptors.[1,5] It has been suggested that hits found in these libraries have potential as chemical platforms for lectin inhibitor design.[6] Despite the flexibility and topological differences among *O*-glycosides and *S*-glycosides,[7] experimental evidence has shown that both *S*-glycosides and glycosyldisulfides bind to lectins similarly to the corresponding *O*-glycosides.[5c,6,8] Therefore, a step forward in the field would be the replacement of the 'untraceable' *O* and *S* atoms by [77]Se as a label atom which possesses similar chemical properties but also an NMR-active nucleus.[9] This proxy atom might ultimately work as a privileged spectroscopic handle that would report key structural information with minimal steric constraints. As such, oxygen substitution in a glycosidic linkage by Se is compatible, and the binding of *Se*-glycosides to lectins has been recently detected by STD and ^{77}Se NMR methods.[10] Together this suggests that the respective glycosyldiselenides,

whose conformational flexibility can resemble that of glycosyldisulfides, would also act as lectin ligands,[6] with sufficient sensitivity for detection even in complex biological environments, such as blood. Given the emerging importance of *O*-GlcNAcylation in biology,[11] we chose here to address this question by probing for the first time the binding of glycosyldiselenides to lectins by STD-NMR and molecular docking studies, as well as demonstrating the advantage of Se for the selective detection of such carbohydrates in complex biological fluids by ^{77}Se NMR spectroscopy. The interaction of the widely-employed, GlcNAc-binding protein wheat-germ agglutinin (WGA) with model ligand probe bis(β-D-GlcNAc)diselenide **1**[12] is reported and its binding mode determined by molecular docking, STD-NMR and CORCEMA-ST calculations. Among the NMR methods employed to identify and characterize the binding of ligands to proteins, the saturation transfer difference (STD) experiment[13] is very robust and has been widely employed to characterize carbohydrate–protein interactions.[14] When surveyed for binding to WGA *in vitro*, **1** showed a clear STD effect (Fig. 1). Additional experimental evidence of binding was also demonstrated by both the line broadening of the ligand resonances after addition of the receptor and from 1D transferred NOESY (Tr-NOESY) experiments (see ESI†). The equilibrium binding constant of the WGA:**1** complex was further determined by classical titration of the ligand into the protein, following the changes in the linewidths of the *N*-acetyl resonance.[15] Using this method a dissociation constant (K_D) of 1.6 mM was obtained (see ESI†), similar to that determined for GlcNAc (K_D = 2.2 mM) by other classical methods,[15b] suggesting representative binding is maintained in the diselenide, at least with WGA.

Next, the theoretical prediction of the binding mode was approached by molecular docking,[16] as described for the interaction of WGA with GlcNAc and a number of its derivatives.[17] The most recently reported WGA·β-GlcNAc X-ray crystallographic structure[18] was chosen and the receptor was prepared for the docking study as already described.[19] Ligand **1** was modelled starting from the crystal structure of β-GlcNAc in complex with the lectin[18] and the crystal structure of the peracetylated

Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford OX1 3TA, UK. E-mail: ben.davis@chem.ox.ac.uk, tim.claridge@chem.ox.ac.uk; Fax: +44 (0) 1865 275674; Tel: +44 (0) 1865 275652

† Electronic supplementary information (ESI) available: Details on sample preparation, NMR experiments, molecular docking simulations and CORCEMA-ST calculations. See DOI: 10.1039/c5cc03952e

‡ Present address: Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Avda. del Conocimiento 3, Parque Tecnológico de Ciencias de la Salud, E-18160 Armilla, Granada, Spain (I.P.-V.) and Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK (O.B.).
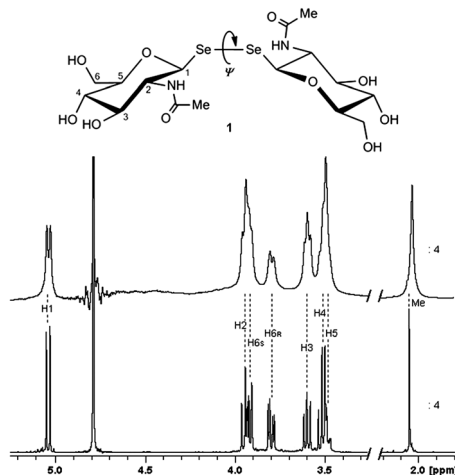
Fig. 1 ^1H NMR reference spectrum of **1** free in solution (bottom). STD-NMR spectrum of **1** (5 mM) in the presence of WGA (100 μM) after 2.5 s saturation (top). Methyl peak intensity reduced to 1/4 for better display. Spectra acquired at 500 MHz and 298 K.

bis(β-D-glucopyranosyl)diselenide[20] (see ESI†). Docking simulations were performed with the program AutoDock Vina.[21] The lectin was treated as a non-flexible receptor and the pyranoside ring bonds of **1** were kept rigid, while torsional rotation freedom was conferred to the remaining bonds. Glycosyldiselenides are expected to have torsional interglycosidic angles Ψ (C–Se–Se–C) close to $+90°$ or $-90°$, as described for glycosyldisulfides,[22] thus, those docking-generated binding modes (poses) where **1** displayed unreasonable values of Ψ were discarded. The top three poses (Fig. 2) had calculated binding affinities of -6.7 kcal mol^{-1} (pose 1) and -6.3 kcal mol^{-1} (poses 2 and 3). All three located one of the GlcNAc residues in the same site where the corresponding monosaccharide is found in the crystal structure of its complex with WGA.[18] Additional contact points created by the

second carbohydrate moiety appear to strengthen binding, suggesting additional possible benefits of sugar diselenides as lectin ligands over their monosaccharide counterparts. There is no direct contact of either selenium atom with the protein surface, with the closest *ca.* 4–5 Å away.

In order to determine which of the predicted poses corresponded to the actual binding mode, the experimental STD effects were compared with those calculated for each docking model (see ESI†) by the program CORCEMA-ST.[23] This protocol for validating binding modes generated by a combination of STD-NMR and molecular docking has been successfully applied since its first introduction.[24] Due to the symmetry of ligand **1**, the calculated STD effects for the same relative proton in each GlcNAc residue of diselenide **1** were mean-averaged as already described for the symmetrical disaccharide trehalose as a ligand of *E. coli* repressor protein TreR.[25] As shown in Fig. 3, the experimental STD effects and those calculated for pose 1 are in excellent agreement and consequently a remarkably low *R*-factor (0.07; see ESI†) was obtained for this model. The matching for the other two poses was significantly worse presenting higher *R*-factors (0.5 for pose 2 and 0.2 for pose 3). In view of these results it can be concluded that pose 1 is the model which better reflects the binding mode of **1** in the primary saccharide-binding site of WGA. The secondary binding site of the lectin was not considered for either the docking or CORCEMA-ST calculations since it is generally assumed to bind carbohydrates with an affinity too weak to be detected in solution.[26]

Finally, the potential of Se as a label for the selective NMR detection of glycosyldiselenides in complex biofluids was tested for **1** in rabbit plasma using detection by ^{77}Se NMR spectroscopy (see Fig. 4 and ESI†). Fig. 4D shows how such heteronuclear detection allows clean observation of **1** without interference from 'matrix' signals, something unavoidable by standard ^1H NMR spectroscopy (Fig. 4B and C). Despite the moderate nuclear
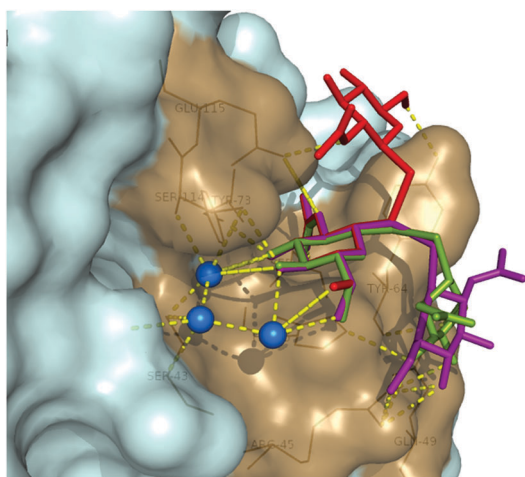

Fig. 2 Three main binding modes of **1** in the primary binding site of WGA derived from docking calculations: pose 1 (red), pose 2 (green) and pose 3 (purple). Interacting amino acids are represented in sand color and polar contacts as yellow dashed lines. Water molecules included in the calculations are represented as blue spheres.
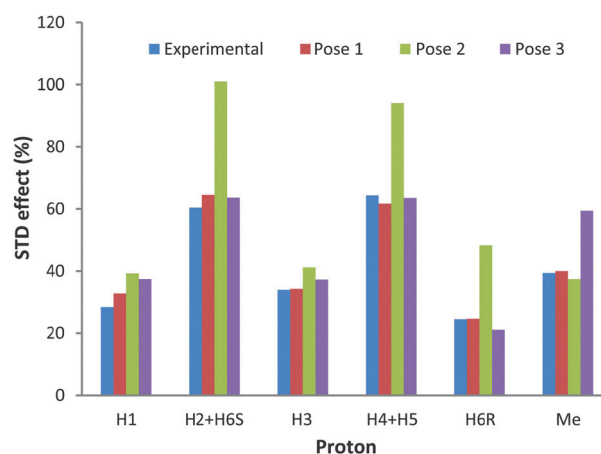

Fig. 3 Comparison of experimental and predicted STD values calculated with the CORCEMA-ST protocol for the three binding poses of **1** in the primary binding site of WGA as obtained with AutoDock Vina. Experimental STD values (coloured in blue) were calculated as $[(I_{0(k)} - I(t)_{(k)})/I_{0(k)} \times 100]$, with $I_{0(k)}$ being the intensity of the signal of the proton k without saturation transfer at time zero and $I(t)_{(k)}$ being the intensity of proton k after saturation transfer during the saturation time t. Pose colours match those of Fig. 2.
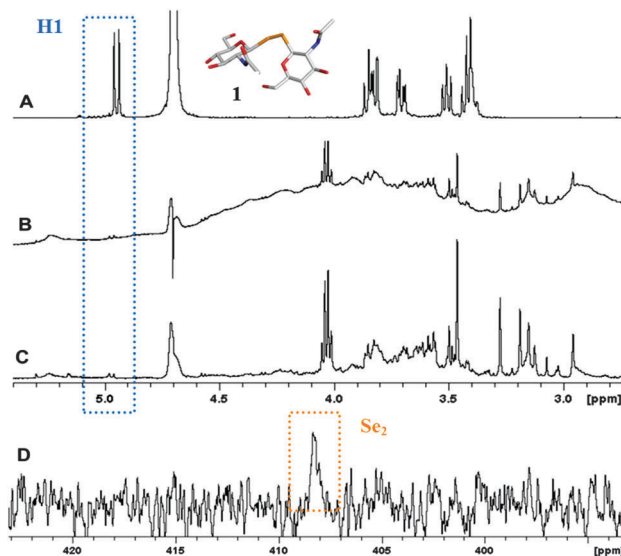
Fig. 4 Expansions of reference ^1H NMR spectrum of **1** in D_2O (A) and NMR spectra of a solution of **1** (1 mM) in rabbit plasma: water suppressed ^1H spectrum (B), water suppressed T_2 filtered ^1H spectrum (C), and ^{77}Se NMR spectrum (D).

sensitivity of ^{77}Se (similar to that of ^{13}C),[9] isotopic enrichment in glycosyldiselenides, hyperpolarization and the use of broadband cryoprobes may each facilitate this selective, easy detection method for this class of compounds in biological samples. Whilst such *in vivo* studies would require proper toxicological evaluation given the presence of selenium, our previous studies, and those of others, demonstrate that glycosylselenides are processed as part of the human selenide metabolic and catabolic pathways.[12b,27] We speculate that this may therefore render such probes reasonably tolerable and processable, and we suggest this proof-of-principle study now highlights their use as powerful novel tools.

In summary, the binding of a glycosyldiselenide probe to a GlcNAc-binding protein has been demonstrated experimentally for the first time. Molecular docking, STD-NMR and CORCEMA-ST calculations allowed an accurate determination of the binding mode of bis(β-D-GlcpNAc)diselenide **1** in the primary saccharide-binding site. Moreover, the use of glycosyldiselenides allows the exploitation of Se as a privileged tag for selective detection of these glycoconjugates in complex biological fluids (*e.g.* plasma) by ^{77}Se NMR spectroscopy. Following this vital proof-of-principle study, we anticipate use of this and related probes to detect GlcNAc-binding proteins *in cellulo* and even *in vivo*. Examples of such experiments measured directly on samples of biological complexity are rare thus far in glycobiology.[28] This, in turn, may allow us to directly identify, for example, so-called "readers" of GlcNAc in epigenetics,[29] where it has recently been implicated as a key protein modification.[30]

Notes and references

1 (a) S. André, Z. Pei, H.-C. Siebert, O. Ramström and H.-J. Gabius, *Bioorg. Med. Chem. Lett.*, 2006, **14**, 6314; (b) Z. Pei, R. Larsson, T. Aastrup, H. Anderson, J.-M. Lehn and O. Ramström, *Biosens. Bioelectron.*, 2006, **22**, 42.
2 (a) B. G. Davis, S. J. Ward and P. M. Rendle, *Chem. Commun.*, 2001, 189; (b) L. Szilágyi, T.-Z. Illyés and P. Herczegh, *Tetrahedron Lett.*, 2001, **42**, 3901; (c) E. J. Grayson, S. J. Ward, A. L. Hall, P. M. Rendle, D. P. Gamblin, A. S. Batsanov and B. G. Davis, *J. Org. Chem.*, 2005, **70**, 9740; (d) L. Szilágyi and O. Varela, *Curr. Org. Chem.*, 2006, **10**, 1745.
3 M. Adinolfi, D. Capasso, S. Di Gaetano, A. Iadonisi, L. Leone and A. Pastore, *Org. Biomol. Chem.*, 2011, **9**, 6278.
4 E. J. Grayson, G. J. L. Bernardes, J. M. Chalker, O. Boutureira, J. R. Koeppe and B. G. Davis, *Angew. Chem., Int. Ed.*, 2011, **50**, 4127.
5 (a) O. Ramström and J.-M. Lehn, *ChemBioChem*, 2000, **1**, 41; (b) T. Hotchkiss, H. B. Kramer, K. J. Doores, D. P. Gamblin, N. J. Oldham and B. G. Davis, *Chem. Commun.*, 2005, 4264; (c) R. Caraballo, M. Sakulsombat and O. Ramström, *Chem. Commun.*, 2010, **46**, 8469.
6 S. Martín-Santamaría, S. André, E. Buzamet, R. Caraballo, G. Fernández-Cureses, M. Morando, J. P. Ribeiro, K. Ramírez-Gualito, B. de Pascual-Teresa, F. J. Cañada, M. Menéndez, O. Ramström, J. Jiménez-Barbero, D. Solís and H.-J. Gabius, *Org. Biomol. Chem.*, 2011, **9**, 5445.
7 F. Strino, J.-H. Lii, C. A. K. Koppisetty, P.-G. Nyholm and H.-J. Gabius, *J. Comput.-Aided Mol. Des.*, 2010, **24**, 1009.
8 (a) Z. Pei, T. Aastrup, H. Anderson and O. Ramström, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 2707; (b) I. Cumpstey, C. Ramstadius, T. Akhtar, I. J. Goldstein and H. C. Winter, *Eur. J. Org. Chem.*, 2010, 1951.
9 H. Duddeck, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1995, **27**, 1.
10 C. Hamark, J. Landström and G. Widmalm, *Chem. – Eur. J.*, 2014, **20**, 13905.
11 (a) J. A. Hanover, *Chem. Biol.*, 2010, **17**, 1272; (b) C. Slawson and G. W. Hart, *Nat. Rev. Cancer*, 2011, **11**, 678.
12 (a) M. Fernández-González, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey and B. G. Davis, *Chem. Sci.*, 2010, **1**, 709; (b) O. Boutureira, G. J. L. Bernardes, M. Fernández-González, D. C. Anthony and B. G. Davis, *Angew. Chem., Int. Ed.*, 2012, **51**, 1432.
13 (a) M. Mayer and B. Meyer, *Angew. Chem., Int. Ed.*, 1999, **38**, 1784; (b) M. Mayer and B. Meyer, *Angew. Chem.*, 1999, **111**, 1902.
14 (a) J. Angulo, C. Rademacher, T. Biet, A. J. Benie, A. Blume, H. Peters, M. Palcic, F. Parra and T. Peters, *Methods Enzymol.*, 2006, **416**, 12; (b) V. Roldós, F. J. Cañada and J. Jiménez-Barbero, *ChemBioChem*, 2011, **12**, 990.
15 (a) F. Jordan, E. Bassett and W. R. Redwood, *Biochem. Biophys. Res. Commun.*, 1977, **75**, 1015; (b) A. Kristiansen, Å. Nysæter, H. Grasdalen and K. M. Vårum, *Carbohydr. Polym.*, 1999, **38**, 23.
16 A. R. Leach, B. K. Shoichet and C. E. Peishoff, *J. Med. Chem.*, 2006, **49**, 5851.
17 D. Neumann, O. Kohlbacher, H.-P. Lenhof and C.-M. Lehr, *Eur. J. Biochem.*, 2002, **269**, 1518.
18 D. Schwefel, C. Maierhofer, J. G. Beck, S. Seeberger, K. Diederichs, H. M. Möller, W. Welte and V. Wittmann, *J. Am. Chem. Soc.*, 2010, **132**, 8704.
19 K. Lycknert, M. Edblad, A. Imberty and G. Widmalm, *Biochemistry*, 2004, **43**, 9647.
20 M. J. Potrzebowski, M. Michalska, J. Blaszczyk, M. W. Wieczorek, W. Ciesielski, S. Kazmierski and J. Pluskowski, *J. Org. Chem.*, 1995, **60**, 3139.
21 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455.
22 K. Fehér, R. P. Matthews, K. E. Kövér, K. J. Naidoo and L. Szilágyi, *Carbohydr. Res.*, 2011, **346**, 2612.
23 (a) V. Jayalakshmi and N. R. Krishna, *J. Magn. Reson.*, 2004, **168**, 36; (b) N. R. Krishna and V. Jayalakshmi, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2006, **49**, 1.
24 X. Wen, Y. Yuan, D. A. Kuntz, D. R. Rose and B. M. Pinto, *Biochemistry*, 2005, **44**, 6729.
25 (a) I. Pérez-Victoria, S. Kemper, M. K. Patel, J. M. Edwards, J. C. Errey, L. F. Primavesi, M. J. Paul, T. D. W. Claridge and B. G. Davis,

Chem. Commun., 2009, 5862; (*b*) S. Kemper, M. K. Patel, J. C. Errey, B. G. Davis, J. A. Jones and T. D. W. Claridge, *J. Magn. Reson.*, 2010, **203**, 1.

26 G. Bains, R. T. Lee, Y. C. Lee and E. Freire, *Biochemistry*, 1992, **31**, 12624.

27 (*a*) C. W. Nogueira and J. B. T. Rocha, *Arch. Toxicol.*, 2011, **85**, 1313; (*b*) A. P. Fernandes and V. Gandin, *Biochim. Biophys. Acta, Gen. Subj.*, 2015, **1850**, 1642.

28 S. Mari, D. Serrano-Gómez, F. J. Cañada, A. L. Corbí and J. Jiménez-Barbero, *Angew. Chem., Int. Ed.*, 2005, **44**, 296.

29 J. A. Hanover, M. W. Krause and D. C. Love, *Nat. Rev. Mol. Cell Biol.*, 2012, **13**, 312.

30 (*a*) K. Sakabe, Z. Wang and G. W. Hart, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 19915; (*b*) B. A. Lewis and J. A. Hanover, *J. Biol. Chem.*, 2014, **289**, 34440.