Carbohydrate Research xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Note

Development of a ^{1}H NMR structural-reporter-group concept for the analysis of prebiotic galacto-oligosaccharides of the [β-D-Gal*p*-(1→*x*)]$_n$-D-Glc*p* type

Sander S. van Leeuwen [a], Bas J. H. Kuipers [b], Lubbert Dijkhuizen [a,*], Johannis P. Kamerling [a]

[a] Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, NL-9747 AG Groningen, The Netherlands
[b] FrieslandCampina, Stationsplein 4, NL-3818 LE Amersfoort, The Netherlands

ARTICLE INFO

Article history:
Received 17 July 2014
Received in revised form 11 August 2014
Accepted 13 August 2014
Available online xxxx

Keywords:
Galacto-oligosaccharides
NMR spectroscopy
Structural-reporter-group concept
Prebiotics

ABSTRACT

Some β-galactosidases (EC 3.2.1.23) are capable of producing mixtures of linear and branched galacto-oligosaccharides (GOS) with various types of glycosidic linkages [degree of polymerization (DP) 2–8; mainly Gal$_n$Glc] when incubated under specific conditions with lactose. These products are generally applied in infant formula. However, for most galacto-oligosaccharide products only major components (low DP) or linkage patterns have been described. To build up a library of ^{1}H and ^{13}C NMR data, a detailed NMR study on commercially available GOS di- and trisaccharides, and some larger GOS oligosaccharides was carried out. Based on the fully assigned ^{1}H and ^{13}C chemical shifts of these model compounds, a ^{1}H NMR structural-reporter-group concept was formulated to function as a tool in the structural analysis of single GOS components and GOS mixtures.

Galacto-oligosaccharides (GOS), produced enzymatically from lactose by incubation with β-galactosidase enzymes (EC 3.2.1.23),[1] are used as prebiotic additives in food industry,[2–6] particularly in infant formula. Various β-galactosidase enzymes from a range of microbial sources have been found capable of producing GOS.[7] Most GOS products are mixtures of linear and branched oligosaccharides with a degree of polymerization (DP) between 2 and 8 and different types of glycosidic linkages. Most structural information is available for the GOS products made with *Bacillus circulans* β-galactosidase, although mainly for lower DP fractions.[8–12]

Structural characterization of a complex mixture of highly similar oligosaccharides, with different linkage types often poses a challenge. Although combinations of 1D and 2D NMR spectroscopic techniques are capable of elucidating the exact structure of purified compounds, interpretation of data is often difficult and requires a high level of expertise. To facilitate structure elucidation of carbohydrates by NMR spectroscopy, structural-reporter-group concepts have been developed for the analysis of glycoprotein N- and O-glycans,[13–15] (oligosaccharides of) arabinoxylans,[15] and (oligosaccharides of) α-glucans.[16] These structural-reporter-group concepts allow for structural annotations on complex and mixed samples, based on 1D ^{1}H NMR spectroscopy.[17–19] Moreover, libraries of fully assigned chemical shift systems for elucidated structures allow for a more rapid assignment of chemical shifts for newly isolated structures.[20]

To get a better insight into the whole ensembles of GOS products, produced by β-galactosidases of different origin, after isolation on the individual component level or as obtained mixtures, we have carried out a detailed NMR analysis on commercially available pure GOS di- and trisaccharides, and some larger GOS oligosaccharides as a first step. Here, we describe the development of a ^{1}H NMR structural-reporter-group concept that forms the basis for further detailed structural investigations of GOS.[21]

1. Experimental

1.1. Materials

The structures of reference compounds **1**–**13** used in this study are shown in Scheme 1. Compound **1** was kindly supplied by Professor Nicola Pohl (Indiana University, Bloomington, IN), compounds **2**–**10** were bought from Carbosynth Ltd (Compton, UK) and compounds **11**–**13** were kindly donated by Professor Tadasu Urashima (Obihiro University of Agriculture & Veterinary Medicine, Obihiro, Japan). D_2O (99.9 atom %) was acquired from Cambridge Isotope Laboratories Ltd (Andover, MA).

* Corresponding author. Tel.: +31 503632153.
E-mail address: l.dijkhuizen@rug.nl (L. Dijkhuizen).

http://dx.doi.org/10.1016/j.carres.2014.08.011
0008-6215/

1.2. NMR spectroscopy

^{1}H NMR spectra, including ^{1}H–^{1}H and ^{13}C–^{1}H correlation spectra were recorded at a probe temperature of 298 K on a Varian Inova 600 spectrometer (NMR Department, University of Groningen, The Netherlands). Samples were exchanged twice with 99.9 atom % D_2O with intermediate lyophilization, finally dissolved in 650 μL D_2O. ^{1}H and ^{13}C chemical shifts are expressed in ppm by reference to internal acetone (δ^1H 2.225; δ^{13}C 31.08). 1D 600-MHz ^{1}H NMR spectra were recorded with 5000 Hz spectral width at 16 k complex data points, using a WET1D pulse to suppress the HOD signal. 2D ^{1}H–^{1}H COSY spectra were recorded in 256 increments in 4000 complex data points with a spectral width of 5000 Hz. 2D ^{1}H–^{1}H TOCSY spectra were recorded with MLEV17 mixing sequences with 30, 50 and 150 ms spin-lock times. 2D ^{13}C–^{1}H gHSQC spectra with multiplicity editing, inverting CH_2 signals, were recorded with a spectral width of 5000 Hz for t_2 and 10,000 Hz for t_1. 2D ^{1}H–^{1}H ROESY spectra with a mixing time of 300 ms were recorded in 128 increments of 4000 complex data points with a spectral width of 5000 Hz. All spectra were processed using MestReNova 5.3 (Mestrelabs Research SL, Santiago de Compostela, Spain), using Whittaker Smoother baseline correction.

2. Results and discussion

Samples of reference compounds **1**–**13** (Scheme 1) were measured with ^{1}H (1D, 2D ^{1}H–^{1}H COSY, TOCSY and ROESY) NMR spectroscopy and 2D ^{1}H–^{13}C gHSQC spectroscopy. From the collected data, all ^{1}H and ^{13}C chemical shifts were assigned (Table 1). Inspection of the ^{1}H and ^{13}C chemical shifts afforded specific (combinations of) δ-values, suitable as structural-reporter-group signals for specific structural elements in GOS components. In the following, →*x*)-D-Glc*p* and →*x*)-D-Gal*p* stand for a reducing monosaccharide unit, β-D-Gal*p*-(1→*x*)- for a non-reducing terminal monosaccharide unit and -(1→*x*)-β-D-Gal*p*-(1→*y*)- for an internal monosaccharide unit.

2.1. Substitution of the reducing D-glucopyranose unit in GOS components

Comparison of the ^{1}H chemical shifts of structures **1**–**4** [β-D-Gal*p*-(1→2/3/4/6)-D-Glc*p*; **B**→**A**] (Scheme 1, Table 1, Supplementary information Fig. S1A–D) with those of free D-Glc*p* shows that the introduction of a β-D-Gal*p* substituent at D-Glc*p* O-2, O-3, O-4 or O-6 is reflected by the set of α-D-Glc*p* H-1, β-D-Glc*p* H-1 and β-D-Glc*p* H-2 δ-values.

The →2)-D-Glc*p* unit (**1**) is characterized by the **Aα** H-1 signal at δ 5.448 ($\Delta\delta$ + 0.224 ppm) and the **Aβ** H-1 and H-2 signals at δ 4.727 ($\Delta\delta$ + 0.09 ppm) and δ 3.53 ($\Delta\delta$ + 0.294 ppm), respectively. In case of the →3)-D-Glc*p* unit (**2**), the **Aα** H-1 signal is detected at δ 5.239 ($\Delta\delta$ + 0.015 ppm) and the **Aβ** H-1 and H-2 signals at δ 4.679 ($\Delta\delta$ + 0.042 ppm) and δ 3.435 ($\Delta\delta$ + 0.199 ppm), respectively.

To differentiate between the →4)-D-Glc*p* (**3**) and →6)-D-Glc*p* (**4**) residues, the δ values of the **Aα** H-1 signals are not suitable (**3**, δ 5.222; **4**, δ 5.227). However, here the combination of the **Aβ** H-1 and H-2 signals at δ 4.662 ($\Delta\delta$ + 0.025 ppm) and δ 3.287 ($\Delta\delta$ + 0.051 ppm), respectively, for 4-substitution (**3**), and δ 4.655 ($\Delta\delta$ + 0.018 ppm) and δ 3.255 ($\Delta\delta$ + 0.019 ppm), respectively, for 6-substitution (**4**) makes the difference.

Each type of linkage also influences the ^{1}H chemical shift values of directly involved and neighbouring H-atoms. For instance, in case of a →3)-D-Glc*p* residue, the H-2, H-3 and H-4 resonances show strong shifts compared to free D-Glc*p*. Most noticeable are the downfield shifts of the H-6a signals in a →6)-D-Glc*p* residue (**4**), that is, from inside the bulk region (δ 4.00–3.50) to outside the bulk region (**Aα** H-6a, δ 4.160; **Aβ** H-6a, δ 4.216). It should be noted that signals outside the bulk region are particularly useful as structural-reporter-group signals.

Nr.	Structure
1	β-D-Gal*p*-(1→2)-D-Glc*p* **B**1→2**A**
2	β-D-Gal*p*-(1→3)-D-Glc*p* **B**1→3**A**
3	β-D-Gal*p*-(1→4)-D-Glc*p* **B**1→4**A**
4	β-D-Gal*p*-(1→6)-D-Glc*p* **B**1→6**A**
5	β-D-Gal*p*-(1→3)-D-Gal*p* **B**1→3**A**
6	β-D-Gal*p*-(1→4)-D-Gal*p* **B**1→4**A**
7	β-D-Gal*p*-(1→6)-D-Gal*p* **B**1→6**A**
8	β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc*p* **C**1→3**B**1→4**A**
9	β-D-Gal*p*-(1→4)-β-D-Gal*p*-(1→4)-D-Glc*p* **C**1→4**B**1→4**A**
10	β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→4)-D-Glc*p* **C**1→6**B**1→4**A**
11	β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-D-Glc*p* **D**1→3**C**1→3**B**1→4**A**
12	β-D-Gal*p*-(1→3)-[β-D-Gal*p*-(1→3)-]$_2$β-D-Gal*p*-(1→4)-D-Glc*p* **D**1→3**C**1→3**C**1→3**B**1→4**A**
13	β-D-Gal*p*-(1→3)-[β-D-Gal*p*-(1→3)-]$_3$β-D-Gal*p*-(1→4)-D-Glc*p* **D**1→3**C**1→3**C**1→3**C**1→3**B**1→4**A**

Scheme 1. Structural overview of reference compounds **1**–**13**, residue labels correspond with those used in text and Table 1.

Compounds **8**–**10** [β-D-Gal*p*-(1→3/4/6)-β-D-Gal*p*-(1→4)-D-Glc*p*] and **11**–**13** {β-D-Gal*p*-(1→3)-[β-D-Gal*p*-(1→3)]$_{1-3}$-β-D-Gal*p*-(1→4)-D-Glc*p*} (Scheme 1, Table 1, Supplementary information Figs. S2A–C and S3A–C), all with a →4)-D-Glc*p* unit, follow the same ^{1}H structural reporters as deduced above for compound **3**: **Aα** H-1, δ 5.223–5.226; **Aβ** H-1, δ 4.663–4.668; **Aβ** H-2, δ 3.281–3.294.

Inspection of the ^{13}C chemical shifts of compounds **1**–**4**, **8**–**13**, as deduced from ^{1}H–^{13}C gHSQC measurements (Table 1), shows clear downfield shifts of the δ values of the substituted carbon atoms of the reducing D-Glc*p* residues, as described earlier in detail for oligosaccharides in general.[22,23] Summarizing, →2)-α/β-D-Glc*p* gives C-2 signals at δ 82.2/82.4, →3)-α/β-D-Glc*p* gives C-3 signals at δ 83.4/85.5, →4)-α/β-D-Glc*p* gives C-4 signals at δ 79.1–80.2/79.1–80.2, and →6)-α/β-D-Glc*p* gives C-6 signals at δ 69.7/69.9 (Table 1).

Please cite this article in press as: van Leeuwen, S. S.; et al. *Carbohydr. Res.* (2014), http://dx.doi.org/10.1016/j.carres.2014.08.011