Development of a $^1$H NMR structural-reporter-group concept for the analysis of prebiotic galacto-oligosaccharides of the $[(\beta-p\text{-Galp-(1\rightarrow x)})_n-p\text{-Glcp}]$ type

Sander S. van Leeuwen $^a$, Bas J. H. Kuipers $^b$, Lubbert Dijkhuizen $^{a,n}$, 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

**Abstract**

Some $\beta$-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$_p$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.

**Keywords:**
Galacto-oligosaccharides, $^1$H NMR spectroscopy, Structural-reporter-group concept, Prebiotics

Galacto-oligosaccharides (GOS), produced enzymatically from lactose by incubation with $\beta$-galactosidase enzymes (EC 3.2.1.23), are used as prebiotic additives in food industry, particularly in infant formula. Various $\beta$-galactosidase enzymes from a range of microbial sources have been found capable of producing GOS. 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 $\beta$-galactosidase, although mainly for lower DP fractions.

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 (oligosaccharides of) arabinoxylans, and (oligosaccharides of) $\alpha$-glucans. These structural-reporter-group concepts allow for structural annotations on complex and mixed samples, based on 1D $^1$H NMR spectroscopy. Moreover, libraries of fully assigned chemical shift systems for elucidated structures allow for a more rapid assignment of chemical shifts for newly isolated structures.

To get a better insight into the whole ensembles of GOS products, produced by $\beta$-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.

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$_2$O (99.9 atom %) was acquired from Cambridge Isotope Laboratories Ltd (Andover, MA).
1.2. NMR spectroscopy

1H NMR spectra, including 1H–1H and 13C–1H 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 % D2O with intermediate lyophilization, finally dissolved in 650 μL D2O. 1H and 13C chemical shifts are expressed in ppm by reference to internal acetone (δ1H 2.225; δ13C 31.08). 1D 600-MHz 1H 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 1H–1H COSY spectra were recorded in 256 increments in 4000 complex data points with a spectral width of 5000 Hz. All spectra were processed with 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 1H (1D, 2D 1H–1H COSY, TOCSY and ROESY) NMR spectroscopy and 2D 1H–13C gHSQC spectroscopy. From the collected data, all 1H and 13C chemical shifts were assigned (Table 1). Inspection of the 1H and 13C 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)-α-D-Glc and →x)-α-D-Gal stand for a reducing monosaccharide unit, β-α-D-Galp-(1→x) for a non-reducing terminal monosaccharide unit and (1→x)-β-α-D-Galp-(1→y) for an internal monosaccharide unit.

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

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

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

To differentiate between the →4)-α-D-Glc (3) and →6)-α-D-Glc (4) residues, the δ values of the Az 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 (Δδ + 0.025 ppm) and δ 3.287 (Δδ + 0.051 ppm), respectively, for 4-substitution (3), and δ 4.655 (Δδ + 0.018 ppm) and δ 3.255 (Δδ + 0.019 ppm), respectively, for 6-substitution (4) makes the difference.

Each type of linkage also influences the δ1H chemical shift values of directly involved and neighbouring H-atoms. For instance, in case of a →3)-3)-α-D-Glc residue, the H-2, H-3 and H-4 resonances show strong shifts compared to free α-D-Glc. Most noticeable are the downfield shifts of the H-6a signals in a →6)-α-D-Glc residue, that is, from inside the bulk region (δ 4.00–3.50) to outside the bulk region (3α H-6a, δ 4.160; 3β-H-6a, δ 4.216). It should be noted that signals outside the bulk region are particularly useful as structural-reporter-group signals.

Comounds 8–10 [β-α-D-Galp-(1→3/4/6)-β-α-D-Galp-(1→4)-α-D-Glc] and 11–13 [β-α-D-Galp-(1→3)-[β-α-D-Galp-(1→3)]-β-α-D-Galp-(1→4)-β-D-Glc] (Scheme 1, Table 1, Supplementary information Figs. S2A–C and S3A–C), all with a →4)-α-D-Glc unit, follow the same 1H structural reporters as deduced above for compound 3: Az H-1, δ 5.223–5.226; Aβ H-1, δ 4.668–4.668; Aβ H-2, δ 3.281–3.294.

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

<table>
<thead>
<tr>
<th></th>
<th>1H</th>
<th>13C</th>
<th>1H</th>
<th>13C</th>
<th>1H</th>
<th>13C</th>
<th>1H</th>
<th>13C</th>
<th>1H</th>
<th>13C</th>
<th>1H</th>
<th>13C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Av-1</td>
<td>5.224</td>
<td>93.5</td>
<td>5.448</td>
<td>93.4</td>
<td>5.329</td>
<td>93.0</td>
<td>5.222</td>
<td>92.8</td>
<td>5.227</td>
<td>93.2</td>
<td>5.258</td>
<td>93.2</td>
</tr>
<tr>
<td>Av-2</td>
<td>3.57</td>
<td>73.1</td>
<td>3.64</td>
<td>82.2</td>
<td>3.72</td>
<td>72.1</td>
<td>3.58</td>
<td>72.4</td>
<td>3.54</td>
<td>71.9</td>
<td>3.80</td>
<td>69.5</td>
</tr>
<tr>
<td>Av-3</td>
<td>3.71</td>
<td>44.5</td>
<td>3.86</td>
<td>74.0</td>
<td>3.91</td>
<td>58.4</td>
<td>3.83</td>
<td>58.3</td>
<td>3.75</td>
<td>74.7</td>
<td>3.71</td>
<td>74.3</td>
</tr>
<tr>
<td>Av-4</td>
<td>3.40</td>
<td>70.1</td>
<td>3.47</td>
<td>70.9</td>
<td>3.52</td>
<td>69.2</td>
<td>3.66</td>
<td>79.8</td>
<td>3.49</td>
<td>70.4</td>
<td>3.98</td>
<td>70.6</td>
</tr>
<tr>
<td>Av-5</td>
<td>3.84</td>
<td>72.9</td>
<td>3.84</td>
<td>72.9</td>
<td>3.89</td>
<td>72.2</td>
<td>3.95</td>
<td>71.2</td>
<td>3.986</td>
<td>71.6</td>
<td>4.083</td>
<td>71.7</td>
</tr>
<tr>
<td>Av-6a</td>
<td>3.83</td>
<td>62.2</td>
<td>3.84</td>
<td>61.8</td>
<td>3.86</td>
<td>61.4</td>
<td>3.87</td>
<td>61.5</td>
<td>4.160</td>
<td>69.7</td>
<td>3.76</td>
<td>62.2</td>
</tr>
<tr>
<td>Av-6b</td>
<td>0.74</td>
<td>3.77</td>
<td>0.80</td>
<td>3.84</td>
<td>0.80</td>
<td>3.84</td>
<td>0.80</td>
<td>3.87</td>
<td>3.7</td>
<td>3.72</td>
<td>3.83</td>
<td></td>
</tr>
<tr>
<td>Av-1</td>
<td>4.637</td>
<td>97.3</td>
<td>4.727</td>
<td>96.1</td>
<td>4.679</td>
<td>96.7</td>
<td>4.662</td>
<td>96.9</td>
<td>4.655</td>
<td>97.1</td>
<td>4.577</td>
<td>97.5</td>
</tr>
<tr>
<td>Av-2</td>
<td>3.236</td>
<td>75.7</td>
<td>3.51</td>
<td>82.4</td>
<td>3.43</td>
<td>74.8</td>
<td>3.28</td>
<td>75.0</td>
<td>3.255</td>
<td>75.1</td>
<td>3.485</td>
<td>72.8</td>
</tr>
<tr>
<td>Av-3</td>
<td>3.40</td>
<td>73.2</td>
<td>3.68</td>
<td>73.8</td>
<td>3.74</td>
<td>75.5</td>
<td>3.64</td>
<td>75.9</td>
<td>3.63</td>
<td>76.0</td>
<td>3.70</td>
<td>76.2</td>
</tr>
<tr>
<td>Av-4</td>
<td>3.39</td>
<td>71.2</td>
<td>3.43</td>
<td>70.9</td>
<td>3.51</td>
<td>69.2</td>
<td>3.66</td>
<td>79.8</td>
<td>3.50</td>
<td>70.4</td>
<td>3.92</td>
<td>69.8</td>
</tr>
<tr>
<td>Av-5</td>
<td>3.44</td>
<td>77.4</td>
<td>3.46</td>
<td>77.3</td>
<td>3.50</td>
<td>76.5</td>
<td>3.50</td>
<td>76.5</td>
<td>3.75</td>
<td>75.6</td>
<td>3.75</td>
<td>75.6</td>
</tr>
<tr>
<td>Av-6a</td>
<td>3.89</td>
<td>62.3</td>
<td>3.89</td>
<td>61.8</td>
<td>3.92</td>
<td>61.6</td>
<td>3.95</td>
<td>61.6</td>
<td>4.216</td>
<td>69.9</td>
<td>3.76</td>
<td>62.2</td>
</tr>
<tr>
<td>Av-6b</td>
<td>0.71</td>
<td>3.70</td>
<td>0.78</td>
<td>3.80</td>
<td>0.78</td>
<td>3.80</td>
<td>0.78</td>
<td>3.73</td>
<td>3.7</td>
<td>3.72</td>
<td>3.85</td>
<td></td>
</tr>
</tbody>
</table>

a From Ref. 16.
b From Ref. 23, δ values adjusted to internal acetone at δ 31.08.

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

2.2. Substitution of the reducing α-galactopyranose unit in GOS components

Comparison of the chemical shifts of structures 5–7 [β-0-GalP-(1→3/4/6)0-GalP; B→A] (Scheme 1, Table 1, Supplementary information Fig. S4A–C) with free α-GlcP shows that the introduction of a β-0-GalP substituent at α-GlcP O-3, O-4 or O-6 gives rise to specific sets of H-1,2,3,4,5,6ab δ-values.

When focusing on the anomic signals outside the bulk region, a (3→)-GalP unit (5) gives rise to shifts of Ax H-1 (δ 5.284, Δδ 0.014 ppm) and Ap H-1 (δ 4.635, Δδ 0.045 ppm). For the (4→)-GalP and (6→)-GalP residues in 6 and 7, respectively, the H-1 shifts are very minor, except for Ap H-1 in 6 (δ 4.613, Δδ 0.023). However, shifts of other 1H signals outside the bulk region can be used for differentiation. In structures 5 and 6 both reducing α-GalP and β-0-GalP H-4 signals are shifted outside the bulk region to δ 4.26 (Δδ 0.025 ppm) and δ 4.20 (Δδ 0.220 ppm) in case of (3→)-GalP, respectively, and to δ 4.233 (Δδ 0.020 ppm) and δ 4.174 (Δδ 0.224 ppm) in case of (4→)-GalP, respectively. Due to the axial position of O-4 of α-GalP residues, the J3,4 and J4,5 coupling constants are small, resulting in easily recognizable peak shapes for Gal H-4. In case of a (6→)-GalP residue (7), the Ax H-5 signal with its characteristic multiplet peak pattern shifts outside the bulk region (δ 4.28; Δδ 0.18 ppm); Ap H-5 gives a shift inside the bulk region (δ 3.90; Δδ 0.17 ppm). Furthermore, the Ax and Ap H-6a signals are observed just outside the bulk region at δ 4.04 (Δδ 0.023 ppm) and δ 4.07 (Δδ 0.26 ppm), respectively.

Comparison, (β-0-GalP-(1→4)-[β-0-GalP-(1→4)-]0,β-0-GalP-(1→4)-GalP), all with a (→)-α-GalP unit, follow the same structural reporters as deduced above for compound 6: (4→)-α-GalP H-1, δ 5.271–5.273; (4→)-α-GalP H-4, δ 4.227–4.230; (4→)-β-0-GalP H-1, δ 4.609–4.610; (4→)-β-0-GalP H-4, δ 4.165–4.170. Inspection of the 13C chemical shifts of compounds 5–7, as deduced from 13C gHSQC measurements (Table 1), shows clear downfield shifts of the δ values of the substituted carbon atoms of the reducing α-GalP residues, as described earlier in detail for oligosaccharides in general.22,23 Summarizing, (3→)-αβ-0-GalP gives C-3 signals at δ 80.4/83.5, (4→)-αβ-0-GalP gives C-4 signals at δ 79.3/78.4, and (6→)-αβ-0-GalP gives C-6 signals at δ 69.8/69.8 (Table 1).


2.3. Non-reducing and internal α-galactopyranose units in GOS components

Trisaccharides 8, 9 and 10 are elongations of compound 3 (β-0-GalP-(1→4)-α-GlcP) with a β-0-GalP residue linked (1→3), (1→4) and (1→6), respectively (C→B→A). The (→)-α-GlcP residue has already been discussed above.

Comparison of the δ-values of the β-0-GalP-(1→4) unit in 3 with those of the internal β-0-GalP unit of 8, 9 and 10 show clear shifts for the anomic signals outside the bulk region, that is, B H-1 (3), δ 4.447; B H-1 (8), δ 4.513; B H-1 (9), δ 4.486; and B H-1 (10), δ 4.483 and combined with their H-2 and H-3 resonances (obtained via 2D 1H–1H NMR experiments), it demonstrates the effect of the linkage position of the terminal β-0-GalP unit. Although both the B H-4 signals of the 3- (8) and 4-substituted (9) internal β-0-GalP unit appear outside the bulk region, they cannot be used for discrimination [B H-4 (8), δ 4.198; B H-4 (9), δ 4.193; in case of 10: B H-4, δ 3.940]; a differentiation is possible on the basis of the 13C δ-values [B C-3, δ 82.8 and B C-4, δ 69.4 (8); B C-3, δ 73.9 and B C-4, δ 78.3 (9)]. Furthermore, the presence of a β-0-GalP-(1→3)-β-0-GalP-(1→4)-β-0-GalP (1→-6)-α-GlcP, δ 4.248/4.446.

2.3. Non-reducing and internal α-galactopyranose units in GOS components

Trisaccharides 8, 9 and 10 are elongations of compound 3 (β-0-GalP-(1→4)-α-GlcP) with a β-0-GalP residue linked (1→3), (1→4) and (1→6), respectively (C→B→A). The (→)-α-GlcP residue has already been discussed above.

Comparison of the δ-values of the β-0-GalP-(1→4) unit in 3 with those of the internal β-0-GalP unit of 8, 9 and 10 show clear shifts for the anomic signals outside the bulk region, that is, B H-1 (3), δ 4.447; B H-1 (8), δ 4.513; B H-1 (9), δ 4.486; and B H-1 (10), δ 4.483 and combined with their H-2 and H-3 resonances (obtained via 2D 1H–1H NMR experiments), it demonstrates the effect of the linkage position of the terminal β-0-GalP unit. Although both the B H-4 signals of the 3- (8) and 4-substituted (9) internal β-0-GalP unit appear outside the bulk region, they cannot be used for discrimination [B H-4 (8), δ 4.198; B H-4 (9), δ 4.193; in case of 10: B H-4, δ 3.940]; a differentiation is possible on the basis of the 13C δ-values [B C-3, δ 82.8 and B C-4, δ 69.4 (8); B C-3, δ 73.9 and B C-4, δ 78.3 (9)]. Furthermore, the presence of a β-0-GalP-(1→3)-β-0-GalP-(1→4)-β-0-GalP (1→-6)-α-GlcP, δ 4.248/4.446.

2.4. Conclusion

For the structural analysis of GOS products, this study has identified a variety of 1H NMR structural reporters. Identities that are not easily made by 1D 1H NMR spectroscopy alone, can be further distinguished by H-1 and H-2 combinations, derived from 2D 1H–1H COSY spectra. This may be useful, when minor sample amounts are available and full 2D NMR spectroscopy, particularly 2D 1H–1H TOCSY spectra with long mixing times, cannot be acquired. Despite the possibility of identifying structural elements with only partial 1H chemical shift assignments, the 13C chemical shifts derived from 2D 13C–1H gHSQC spectroscopy are useful for verification of structural assignments. The availability of basic NMR data, as worked out in full detail in this model study for low-DP GOS components, is of high importance in the NMR characterization of linear and branched higher-DP GOS components, as described in the accompanying study.22

Acknowledgments

This work was financially supported by the European Union/European Regional Development Fund and by the Dutch Ministry of Economic Affairs, Agriculture and Innovation, the Dutch innovation program ‘Peaks in the Delta’, the Municipality of Groningen, the Provinces of Groningen, Fryslân and Drenthe, and the Dutch Carbohydrate Competence Center (CCC Project WP6a).
Supplementary data

Supplementary data (1D $^1$H NMR spectra of compounds 1–13) associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.carres.2014.08.011.

References