



University of Groningen

## Biochemical characterization of $\beta$ -galactosidases and engineering of their product specificity

Yin, Huifang

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Yin, H. (2017). Biochemical characterization of  $\beta$ -galactosidases and engineering of their product specificity. [Groningen]: University of Groningen.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Chapter 2

**Reaction kinetics and galactooligosaccharide product profiles of the  $\beta$ -galactosidases from *Bacillus circulans*, *Kluyveromyces lactis* and *Aspergillus oryzae***

**Huifang Yin<sup>†</sup>, Jelle B. Bultema<sup>†</sup>, Lubbert Dijkhuizen<sup>†,\*</sup> and Sander S. van Leeuwen<sup>†</sup>**

<sup>†</sup>Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

*Food Chemistry* (2017), **225**, 230-238

## Abstract

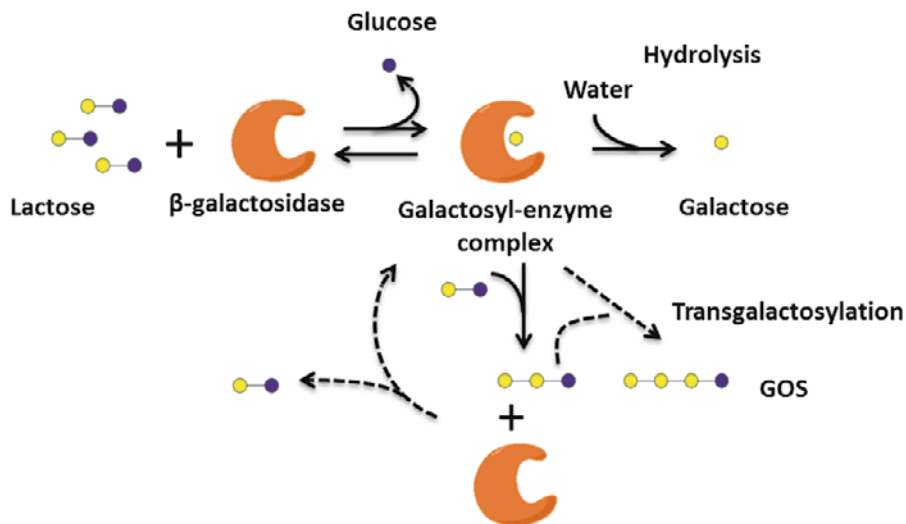
$\beta$ -Galactosidase enzymes are used in the dairy industry to convert lactose into galactooligosaccharides (GOS) that are added to infant formula to mimic the molecular sizes and prebiotic functions of human milk oligosaccharides. Here we report a detailed analysis of the clearly different GOS profiles of the commercial  $\beta$ -galactosidases from *Bacillus circulans*, *Kluyveromyces lactis* and *Aspergillus oryzae*. Also the GOS yields of these enzymes differed, varying from 48.3% (*B. circulans*) to 34.9% (*K. lactis*), and 19.5% (*A. oryzae*). Their incubation with lactose plus the monosaccharides Gal or Glc resulted in altered GOS profiles. Experiments with  $^{13}\text{C}_6$  labeled Gal and Glc showed that both monosaccharides act as acceptor substrates in the transgalactosylation reactions. The data shows that the lactose isomers  $\beta\text{-D-Galp-(1}\rightarrow\text{2)-D-Glcp}$ ,  $\beta\text{-D-Galp-(1}\rightarrow\text{3)-D-Glcp}$  and  $\beta\text{-D-Galp-(1}\rightarrow\text{6)-D-Glcp}$  are formed from acceptor reactions with free Glc and not by rearrangement of Glc in the active site.

## Introduction

More than 200 human milk oligosaccharides (hMOS) have been identified in human mother milk, and they fulfill many functions in the health and development of the neonate [1]. In addition to providing nutrients for brain development, hMOS modulate intestinal immunity, block the binding of pathogens, and promote the growth of beneficial gut bacteria [2], [3]. Nowadays many babies receive infant formula based on bovine milk, which has a more limited abundance and complexity of oligosaccharides [3]. As an alternative route to provide beneficial oligosaccharides, analogues have been developed and added to infant formula to mimic the molecular sizes and prebiotic functions of hMOS [4]. These prebiotic analogues consists of short chain galactooligosaccharides (GOS), long chain fructooligosaccharides (FOS), polydextrose, and mixtures of these in different ratios [4], [5], [6]. Babies who received these analogues had significantly more Bifidobacteria in their gut microbiome than those in the placebo group [7]. In addition, the species distribution of Bifidobacteria was more similar to that of the group receiving human mother milk [7].

$\beta$ -Galactosidase enzymes are widely used in the dairy industry to convert lactose into GOS. They attack the anomeric center of the galactose residue in lactose, forming a galactosyl-enzyme complex while releasing the Glc molecule [8], [9], [10]. The subsequent step depends on the acceptor substrate: if the acceptor is water, the galactosyl-enzyme complex undergoes hydrolysis and releases the Gal molecule as well; if lactose, monosaccharide or oligosaccharide serves as acceptor, GOS are formed as the transgalactosylation product. The previously formed disaccharide or oligosaccharide can either serve as a new acceptor substrate yielding GOS products with a higher degree of polymerization (DP), or bind to the enzyme to be used as a donor substrate (Figure 1). The linkage types and the DP of GOS produced depend on the specific enzyme and reaction conditions [9], [10].

Several commercial GOS products currently are available, such as Oligomate 55, Bimuno, and Vivinal® GOS [9], [11]. Various microbial  $\beta$ -galactosidases are used for GOS synthesis, using relatively high lactose concentrations, yielding GOS mixtures with different structural compositions which are likely to result in different prebiotic effects [11], [12].



**Figure 1.** Reaction scheme of the  $\beta$ -galactosidase enzyme with lactose.

$\beta$ -Galactosidases from the bacterium *B. circulans*, the yeast *K. lactis* and the fungus *A. oryzae* are used in the dairy industry because of their high transgalactosylation activity and different ranges of products [13], [14], [15], [16]. Other formulations are produced using two enzymes, e.g. Oligomate 55 by the *A. oryzae* and *Streptococcus thermophilus*  $\beta$ -galactosidase enzymes [9]. *B. circulans*  $\beta$ -galactosidase production yields 4 isoforms (BgaD-A, BgaD-B, BgaD-C, and BgaD-D), caused by truncation at the C-terminus of the BgaD protein (full length 1737 amino acids) [17]. The shortest isoform (BgaD-D) contains 812 amino acids, and the crystal structure is a dimer [18], [19]. At high lactose

concentration, BgaD-D has a similar ability to produce GOS as the other isoforms [20]. NMR analysis of Vivinal<sup>®</sup> GOS, the commercial GOS product made with the *B. circulans*  $\beta$ -galactosidase, has revealed more than 40 structures, covering 99% of the products [21,22]. The  $\beta$ -galactosidase from *K. lactis* consists of 1024 amino acids and the crystal structure is a tetramer [23]. HPAEC-PAD analysis has revealed 6 structures in the *K. lactis* Lactozyme 3000 HG GOS product mixture, with a preference for ( $\beta$ 1 $\rightarrow$ 6) linkages [24]. The  $\beta$ -galactosidase from *A. oryzae* is a monomer with 985 residues [25]. The GOS molecules identified as products of this enzyme constitutes a mixture of 9 structures, among which  $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\beta$ -D-Galp-(1 $\rightarrow$ 4)-D-Glcp accounts for nearly one-third of the total GOS [26].

Over the years, much effort has been devoted to optimize the reaction conditions of these 3  $\beta$ -galactosidases to obtain a higher GOS yield. Several immobilization methods have been tested to enhance the stability of the enzymes [15], [27], [28], [29]. Different reaction temperatures and pH values have been used to improve the GOS yield [30], [31]. It has been suggested that the monosaccharides produced from lactose (Gal and Glc) inhibit the activity of the  $\beta$ -galactosidase enzymes from *A. oryzae*, *K. lactis*, and *B. circulans*, resulting in failure to reach the highest GOS yield [10], [32], [33]. Several studies also have explored the transgalactosylation products of the 3 enzymes individually [34,30,26,24], and made a comparison between the 3 enzymes [35].

The reaction kinetic changes [34] of GOS fractions of these 3  $\beta$ -galactosidase enzymes have not been studied yet. The aim of the present study is a comprehensive comparison of the complex GOS-synthesis process for three of the most prominent  $\beta$ -galactosidase enzymes currently used in industry. In this paper, the dynamic changes of the major GOS fractions produced by these 3 enzymes during the GOS synthesis process are presented. Also the influence of

the monosaccharides (Gal and Glc) on GOS synthesis is studied in detail and compared among the 3 enzymes.

## **Materials and methods**

### **Materials**

$\beta$ -Galactosidase from *K. lactis* (Lactozyme 2600L) and  $\beta$ -galactosidase from *A. oryzae*, Fucose, D-Glucose- $^{13}\text{C}_6$  ( $\geq 99$  atom %  $^{13}\text{C}$ ), and D-Galactose- $^{13}\text{C}_6$  (98 atom %  $^{13}\text{C}$ ) were purchased from Sigma-Aldrich (St. Louis, USA). Gal, lactose, Glc, sodium chloride, sodium hydroxide, sodium hydrogen phosphate, sodium dihydrogen phosphate, and sodium acetate were from Merck (Darmstadt, Germany).

### **Recombinant protein expression and purification**

The C-terminally truncated *B. circulans* ATCC 31382 recombinant  $\beta$ -galactosidase (rBgaD-D) protein was constructed previously [8] and used in this study. PCR amplification was performed in order to add a 6 $\times$ His tag at the N-terminus of rBgaD-D. The template was plasmid pET-15b containing the rBgaD-D encoding gene. A forward primer (5'-CAGGGACCCGGTATGGAAACAGTGTGAGC-3') and reverse primer (5'-CGAGGAGAAGCCCGTTATGGCGTTACCGTAAATAC-3') were used for PCR amplification; the PCR products were purified on an agarose gel. Vector pET-15b-LIC was digested by FastDigest KpnI (Thermo Scientific) and purified with a PCR purification kit (GE Healthcare). Subsequently, the PCR product was treated with T4 DNA polymerase (New England BioLabs) in the presence of 2.5 mM dATP, while the vector was digested with T4 DNA polymerase in the presence of 2.5 mM dTTP. Both reactions were incubated at room temperature for 60 min, followed by 20 min at 75 °C to inactivate the enzymes. The reaction mixture containing 2  $\mu\text{L}$  of the target DNA and 1  $\mu\text{L}$  vector was incubated at room temperature for 15 min to allow ligation. Then the mixture was transformed

into *Escherichia coli* DH5 $\alpha$  competent cells (Phabagen) for DNA amplification. The DNA sequence was verified by sequencing.

The plasmid containing the gene encoding the His-tagged rBgaD-D protein was transformed into *E. coli* BL21\* DE3 competent cells (Invitrogen, Carlsbad, USA). Precultures of *E. coli* BL21\* DE3 harboring the plasmid were grown overnight at 37 °C. Then 1% preculture was inoculated into fresh LB medium containing 100  $\mu$ g/ml ampicillin. When the cell density reached 0.6 at 600 nm, the expression of His-tagged rBgaD-D was induced with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside. Subsequently, the cells were cultured overnight at 30 °C and 220 rpm/min, and harvested by centrifugation. Cell pellets were washed with 20 mM Tris-HCl (pH 8.0) buffer and resuspended in B-PER lysis solution (ThermoScientific, Pierce). After incubation at room temperature for 30 min, the cell debris was removed by centrifugation. To purify the protein, the cell-free extract was mixed with HIS-Select<sup>®</sup> Nickel Affinity Gel (Sigma, USA), which was previously equilibrated with 20 mM Tris-HCl (pH8.0), 50 mM NaCl (buffer A), and incubated at 4 oC overnight. The unbound protein was washed away with 20 column volumes of buffer A, and then the rBgaD-D protein was eluted with buffer A containing 100 mM imidazole. The protein was centrifuged with a centrifugal filter with a cutoff of 30 kDa (Merck, Darmstadt, Germany) to remove the imidazole.

### **Enzyme incubations**

#### *Enzyme activity assays*

The  $\beta$ -galactosidase activity towards lactose under relevant GOS production conditions was measured using an oxidase/peroxidase method (GOPOD Format, Megazyme, Ireland). One unit (U) of (total) activity was defined as the enzyme amount required to release 1  $\mu$ mol Glc per min. For rBgaD-D, 50% lactose (w/w) in 0.1 M sodium phosphate buffer (pH 6.0) was used as substrate [40], incubated at 60 °C for 5 min. For Lactozyme 2600L, 30% lactose (w/w) in 0.1 M sodium



phosphate buffer (pH 7.0) was used and the mixture was incubated at 40 °C for 5 min [15]. For the  $\beta$ -galactosidase from *A. oryzae*, 30% lactose (w/w) in 0.1 M sodium acetate buffer (pH 4.5) was used and incubation was carried out at 45 °C for 5 min [32]. Different lactose concentrations were used reflecting the solubility of lactose at the respective enzyme optimum temperature. After 5 min, the reactions were stopped immediately by adding 1.5 M NaOH, followed by neutralization with 1.5 M HCl. The amount of released Glc was determined in the GOPOD assay by measuring the absorbance at 510 nm.

### *GOS synthesis*

For the production of GOS from lactose, incubations of all 3 enzymes contained 37 U enzyme activity per gram lactose. The incubation conditions were: 50% lactose (w/w) in 0.1 M sodium phosphate buffer (pH 6.0), 60 °C for rBgaD-D; 30% lactose (w/w) in 0.1 M sodium acetate buffer (pH 4.5), 45 °C for  $\beta$ -galactosidase from *A. oryzae*; 30% lactose (w/w) in 0.1 M sodium phosphate buffer (pH 7.0), 40 °C for Lactozyme 2600L. At specified time intervals, 50  $\mu$ L aliquots of reaction mixture were withdrawn and heated at 100 °C for 5 min to stop the reaction.

To analyze the synthesis of GOS products in time by the 3  $\beta$ -galactosidases, 3.75 U enzyme activity per gram lactose was used for all 3 enzymes and incubated at their optimal conditions (see above). Aliquot samples were taken at 5 min, 15 min, 30 min, 1 h, 2 h, 4 h, and 8 h and immediately incubated at 100 °C for 5 min to stop the reaction.

The final GOS profiles were obtained by using 370 U enzyme activity per gram lactose for all 3 enzymes. The reactions were incubated at their optimal conditions for 72 h, 48 h, and 48 h for rBgaD-D, Lactozyme 2600L, and the  $\beta$ -galactosidase from *A. oryzae*, respectively.

### *Effects of Gal and Glc on GOS synthesis*

To investigate the influence of the presence of the monosaccharides Gal and Glc on GOS synthesis, a mixture of 30% (w/w) lactose plus 20% (w/w) Gal or Glc was used for incubations with rBgaD-D. For Lactozyme 2600L and *A. oryzae*  $\beta$ -galactosidases, a mixture containing 20% (w/w) lactose plus 10% (w/w) Gal or Glc was used. The reactions were carried out using 37 U of the enzymes per gram lactose, incubated at their respective optimal conditions. The rBgaD-D enzyme also was incubated with 30% (w/w) lactose plus 20% (w/w)  $^{13}\text{C}_6$  labelled Gal or 20% (w/w)  $^{13}\text{C}_6$  labelled Glc as described above. All reactions were stopped after 2 h by incubation at 100 °C for 5 min. Structures of interest were isolated using preparative HPAEC-PAD separations.

### HPAEC-PAD

For analytical HPAEC-PAD the reaction samples were diluted 3000 times with Milli-QTM water, resulting in samples of  $\sim 0.10$ - $0.17$  mg/mL, containing 200  $\mu\text{M}$  fucose as internal reference for the HPAEC-PAD analysis and quantification. The analysis of the samples was carried out with a Dionex ICS-3000 work station (ThermoScientific, Amsterdam, the Netherlands), coupled to a CarboPac PA-1 column (250  $\times$  4 mm, Dionex) and an ICS-3000 ED pulsed amperometric detector (PAD). The separation conditions were the same as used previously [21]. Briefly, the elution buffer consisted of a complex gradient of A) 100 mM sodium hydroxide, B) 600 mM sodium acetate in 100 mM sodium hydroxide, C) Milli-Q water, and D) 50 mM sodium acetate, details are provided in Supplementary Figure S1. The quantification of GOS fractions is determined by the peak intensities of HPAEC-PAD profiles. GOS yield (%) = peak intensities of total GOS fractions / peak intensities of (galactose + glucose + lactose + total GOS fractions) \* 100%. Percentage of specific GOS fraction (%) = peak intensities of specific GOS fraction / peak intensities of (galactose + glucose + lactose + total GOS fractions) \* 100%. Preparative separations were performed on an ICS-5000 work station (ThermoScientific), coupled to a CarboPac PA-1 column (250  $\times$  9

mm), using the same gradient as used for analytical separations at 4 mL/min. Samples were diluted to 4 mg/mL concentration and 250  $\mu$ L was injected per separation. After separation samples were immediately neutralized using 4 M acetic acid, followed by desalting on Carbograp SPE (Grace, Breda, The Netherlands), eluting with 3 x 1 mL 40% acetonitrile in Milli-Q water.

### **MALDI-TOF-MS analysis**

Positive-ion mass spectra were recorded on an Axima<sup>TM</sup> Performance mass spectrometer (Shimadze Kratos Inc., Manchester, UK) equipped with a nitrogen laser (337 nm, 3 ns pulse width). Spectra were recorded in reflectron mode at a resolution of at least 5000 FWHM and acquired with software controlled delayed-extraction optimized for  $m/z$  800. Spectra were recorded with a range of 1-5000  $m/z$  with ion-gate blanking set to 300  $m/z$ . Samples were prepared by mixing on the target plate 1  $\mu$ L sample solution ( $\sim$ 1 mg/mL) with 1  $\mu$ L matrix solution, consisting of 2,5-dihydroxybenzoic acid (10 mg/mL) in 40% acetonitrile.

### **NMR spectroscopy**

The isolated oligosaccharide samples and the reaction mixtures were lyophilized and exchanged twice with 99.9% atom D<sub>2</sub>O (Cambridge Isotope Laboratories Ltd, Andover, MA). Finally, samples were dissolved in 650  $\mu$ L 99.9% atom D<sub>2</sub>O, containing 25 ppm acetone ( $\delta$ <sup>1</sup>H 2.225,  $\delta$ <sup>13</sup>C 31.08) as internal standard. All NMR spectra, including 1D <sup>1</sup>H, as well as 2D <sup>1</sup>H-<sup>1</sup>H and <sup>13</sup>C-<sup>1</sup>H correlation spectra were recorded at a probe temperature of 298K on a Varian Inova 600 spectrometer (NMR Department, University of Groningen, The Netherlands). 1D 600-MHz <sup>1</sup>H NMR spectra were recorded with 5000 Hz spectral width at 16k complex data points, using a WET1D pulse to suppress the HOD signal. 2D <sup>1</sup>H-<sup>1</sup>H spectra were recorded in 200 increments of 4000 complex data points with a spectral width of 5000 Hz. 2D <sup>1</sup>H-<sup>1</sup>H TOCSY spectra were recorded with MLEV17 mixing sequences with 30, 60, and 150 ms spin-lock times. 2D <sup>13</sup>C-<sup>1</sup>H

HSQC spectra were recorded with a spectral width of 5000 Hz in  $t_2$  and 10,000 Hz in  $t_1$  direction. 2D  $^1\text{H}$ - $^1\text{H}$  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 [36].

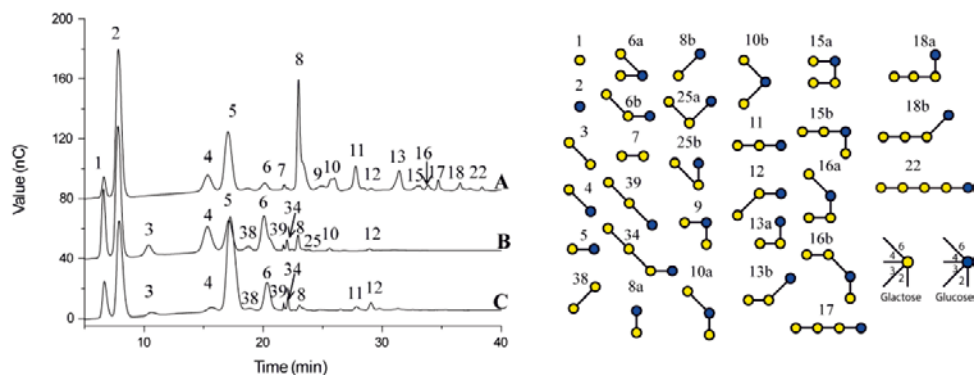
## Results and Discussion

### Structures of GOS products

For each  $\beta$ -galactosidase enzyme a product profile was obtained after 4 h incubation at their respective optimal conditions (Figure 2). GOS synthesized by rBgaD-D presents a mixture matching that of the commercial product of the native *B. circulans* enzyme, Vivinal<sup>®</sup> GOS [21], [22], here we only list the 21 major structures in Figure 2. We also identified a total of 12, and 11 structures from the GOS profiles of Lactozyme 2600L and *A. oryzae*  $\beta$ -galactosidase, respectively. The GOS structures of peak **1-25a**, **34** and **38** were identified according to the GOS library described previously (Table S1) [12], [21], [22]. Structure **39** [ $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\beta$ -D-Galp-(1 $\rightarrow$ 6)-D-Glcp] was identified by matching the 1D  $^1\text{H}$  NMR spectrum with that previously published [37]. Structure **25b** [ $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\beta$ -D-Galp-(1 $\rightarrow$ 2)-D-Glcp] was identified by a combination of HPAEC-PAD profiling, MALDI-TOF-MS and NMR spectroscopy (Supplementary).

GOS synthesized by rBgaD-D presents a mixture showing ( $\beta$ 1 $\rightarrow$ 4), ( $\beta$ 1 $\rightarrow$ 2), ( $\beta$ 1 $\rightarrow$ 3), ( $\beta$ 1 $\rightarrow$ 6) linked Gal on the reducing Glc, with both linear and branched structures. All of these structures can be further elongated with ( $\beta$ 1 $\rightarrow$ 4)-linked Gal residues. Structure **6b** (6`GalLac) with a ( $\beta$ 1 $\rightarrow$ 6)-linked elongation of lactose, and structure **12** (3`GalLac) with a ( $\beta$ 1 $\rightarrow$ 3)-linked elongation of lactose, were present only in trace amounts (Figure 2).

In comparison, the GOS structures of the Lactozyme 2600L and *A. oryzae*  $\beta$ -galactosidase present much less complexity and variety. For *K. lactis* (Lactozyme 2600L) structures **3**, **4**, **6b**, **8b**, **12**, and **38** (Table 1) were previously identified,











































**Figure 2.** HPAEC-PAD profiles (left) and GOS structures (right) produced from lactose by the 3  $\beta$ -galactosidases, (A) rBgAD-D from *B. circulans*, (B) Lactozyme 2600L from *K. lactis*, (C)  $\beta$ -galactosidase from *A. oryzae*. The profiles were obtained after incubation for 4 h with the enzymes supplied at 37 U/g lactose. The determination of the GOS product structures was based on the retention times, NMR spectroscopy, and MALDI-TOF-MS (see Supplementary).

using Lactozyme 3000 HG [24]. Structures **6a**, **8a**, **10a**, **10b**, **25a**, **25b**, and **39** (Table S1) had not been identified for the *K. lactis*  $\beta$ -galactosidase (Figure 2), representing various peaks in the previous work that remained to be assigned [24]. No structure **11** (4'GalLac) was observed here; possibly the Lactozyme 3000 HG preparation used by Rodriguez-Colinas *et al.* [24] has slightly different activity and/or specificity.

For *A. oryzae*  $\beta$ -galactosidase structures **3**, **4**, **6b**, **8b**, **11**, **12**, **34** and **38** (Table 1) were observed both in this study as well as in previous work [26]. Here, also **6a**, **8a**, and **39** were identified (Table 1, Figure 2). Structure **7** [ $\beta$ -D-Galp-(1 $\rightarrow$ 4)-D-Galp] was not observed here but in the previous work [26].

**Table 1.** Structures observed in the GOS produced by  $\beta$ -galactosidases from *B. circulans* (BC), *K. lactis* (KL) and *A. oryzae* (AO).

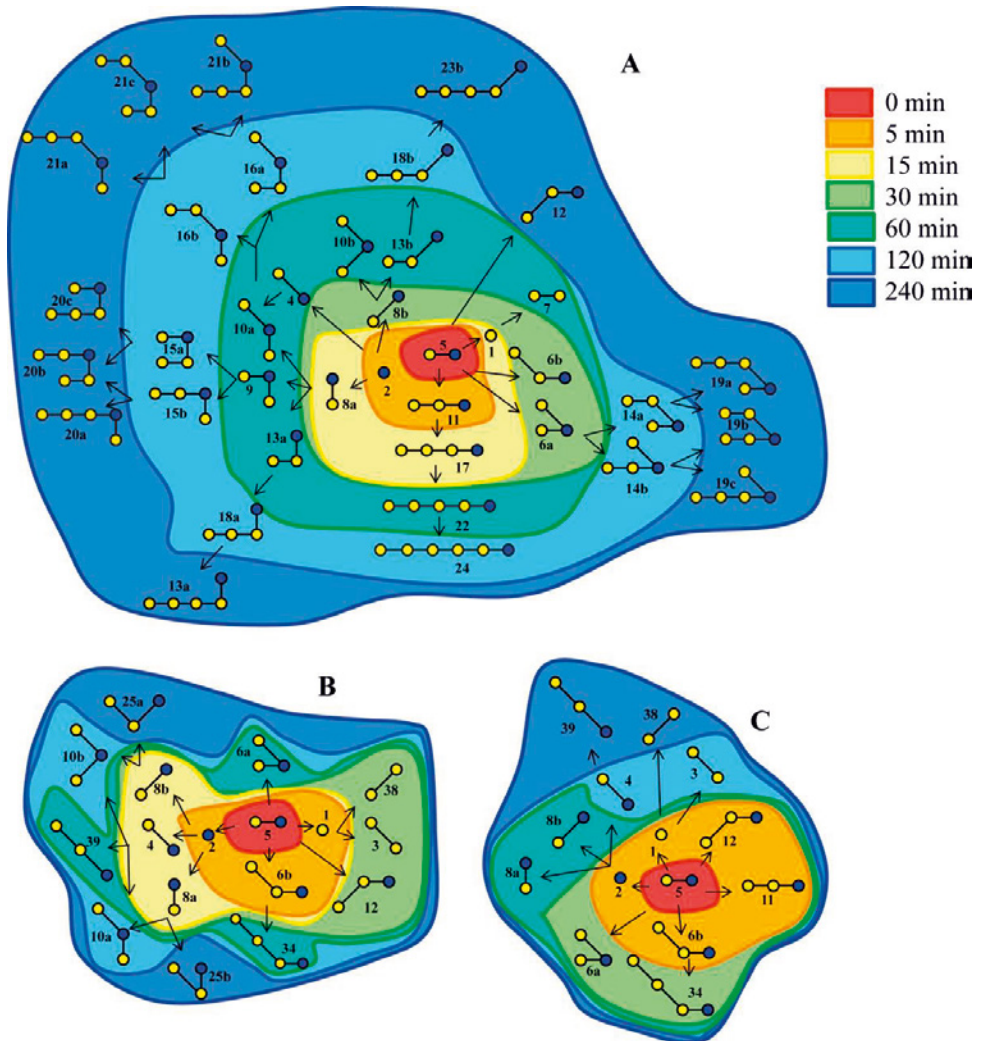
Nr	Structure	BC	KL	AO	Nr	Structure	BC	KL	AO
1		●	●	●	17		●	○	○
2		●	●	●	18a		●	○	○
3		●	●	●	18b		●	○	○
4		●	●	●	19a		●	○	○
5		●	●	●	19b		●	○	○
6a		●	●	●	19c		●	○	○
6b		●	●	●	20a		●	○	○
7		●	○	○	20b		●	○	○
8a		●	●	●	20c		●	○	○
8b		●	●	●	21a		●	○	○
9		●	○	○	21b		●	○	○
10a		●	●	○	21c		●	○	○
10b		●	●	○	22		●	○	○
11		●	○	●	23a		●	○	○
12		●	●	●	23b		●	○	○
13a		●	○	○	24		●	○	○
13b		●	○	○	25a		○	●	○
14a		●	○	○	25b		○	●	○
14b		●	○	○	34		○	●	●
15a		●	○	○	38		○	●	●

Nr	Structure	BC	KL	AO	Nr	Structure	BC	KL	AO
<b>15b</b>		●	○	○	<b>39</b>		○	●	●
<b>16a</b>		●	○	○	<b>X</b>		●	○	○
<b>16b</b>		●	○	○					
<b>16c</b>		●	○	○					

The  $\beta$ -galactosidases from *K. lactis* and *A. oryzae* clearly showed a preference for ( $\beta$ 1 $\rightarrow$ 6) elongation, reflected in structures **4**, **6a**, **6b**, **25a**, **25b**, **34** and **38** (Table 1, Figure 2). There are some differences, however. GOS produced by Lactozyme 2600L showed two branched structures (peak **10a**, **10b**) stemming from ( $\beta$ 1 $\rightarrow$ 6) substitution of the reducing Glc in structures **8a** and **8b**, respectively. GOS from *A. oryzae*  $\beta$ -galactosidase exhibited also ( $\beta$ 1 $\rightarrow$ 4) elongating activity (peak **11**; 4'GalLac), which was not found in *K. lactis*. Overall, *K. lactis*  $\beta$ -galactosidase produced mainly ( $\beta$ 1 $\rightarrow$ 6) elongations (peaks **3**, **4**, **6**, Figure 2) while  $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\beta$ -D-Galp-(1 $\rightarrow$ 4)-D-Glcp is the major structure produced by *A. oryzae*  $\beta$ -galactosidase, as previously reported by Urrutia *et al.* [26].

### Formation routes of GOS structures

The synthesis of GOS products was followed in time for all 3  $\beta$ -galactosidase enzymes, incubated at 3.75 U/g lactose (Figure 3). From an analysis of the Vivinal product mixture it was previously observed that the commercially used *B. circulans*  $\beta$ -galactosidase is able to introduce all types of substitution on the reducing Glc residue [21], [22]. When rBgaD-D was incubated with lactose and the reaction followed in time (Figure 3A), first ( $\beta$ 1 $\rightarrow$ 4) elongation of lactose occurred, resulting in structure **11** (4'GalLac). At 15 min ( $\beta$ 1 $\rightarrow$ 4)-elongation of



**Figure 3.** Kinetic changes of major GOS fractions produced by the 3  $\beta$ -galactosidases. (A) rBgaD-D from *B. circulans*, (B) Lactozyme 2600L from *K. lactis*, (C)  $\beta$ -galactosidase from *A. oryzae*. The enzyme amount and reaction conditions are the same as in Figure 2. The numbers correspond to the peak numbers in Figure 2. Data were obtained in three parallel experiments.

structure **11** to structure **17** occurred, followed by further ( $\beta$ 1 $\rightarrow$ 4)-elongation to structure **22** at 60 min and to structure **24** at 120 min (Figure 3A, Table 1). The first substitution structure of released Glc, by ( $\beta$ 1 $\rightarrow$ 2)-elongation, i.e. [ $\beta$ -D-Galp-(1 $\rightarrow$ 2)-D-Glcp] (structure **8a**) was observed at 15 min. Hydrolysis, as evidenced



by free Gal in minor amounts (structure **1**), was also observed at 15 min. Further substitution of released Glc by ( $\beta$ 1 $\rightarrow$ 3)-linked Gal, i.e. [ $\beta$ -D-Galp-(1 $\rightarrow$ 3)-D-Glcp] (structure **8b**) was found at 30 min, and elongation of free Glc by ( $\beta$ 1 $\rightarrow$ 6)-linked Gal [ $\beta$ -D-Galp-(1 $\rightarrow$ 6)-D-Glcp] (allolactose; structure **4**) was observed at 60 min. Interestingly, the occurrence of 4,6-branched Glc in structure **6a** [ $\beta$ -D-Galp-(1 $\rightarrow$ 4)-{ $\beta$ -D-Galp-(1 $\rightarrow$ 4)-}D-Glcp] occurred already at 30 min, indicating it was not initially formed from allolactose (structure **4**). This branched structure could already be formed from ( $\beta$ 1 $\rightarrow$ 6)-substitution of the Glc in lactose. At 60 min, all branched structures were found (**9**, **10a**, and **10b**, besides **6a** that was formed earlier), as well as the products of ( $\beta$ 1 $\rightarrow$ 4) elongation of structures **8a** and **8b**, i.e. structures **13a** [ $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ 2)-D-Glcp] and **13b** [ $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ 3)-D-Glcp], respectively (Figure 3A, Table 1). All these trisaccharides were further elongated in a ( $\beta$ 1 $\rightarrow$ 4) manner in time up to the full DP5 spectrum at 240 min. The ( $\beta$ 1 $\rightarrow$ 3) elongation of lactose (structure **12**) was only observed at 240 min, indicating that this reaction is less favorable for the enzyme.

With *K. lactis*  $\beta$ -galactosidase (Figure 3B), ( $\beta$ 1 $\rightarrow$ 6) elongation on the Gal unit of lactose (structure **6b**) was first observed at 5 min, followed by further ( $\beta$ 1 $\rightarrow$ 6) elongation forming structure **34** at 60 min. The free Glc and Gal were released at 5 min. Then Gal was further elongated by ( $\beta$ 1 $\rightarrow$ 3) and ( $\beta$ 1 $\rightarrow$ 6) linkages, forming structures **38** and **3** at 30 min, respectively. Meanwhile, the free Glc was used as acceptor, forming structures **4**, **8a**, **8b** at 15 min. These 3 structures were used for synthesis of structures **39**, **10a**, **10b**, **25a**, and **25b** (Figure 3B, Table 1) from 60 min to 240 min. Structure **12** is a ( $\beta$ 1 $\rightarrow$ 3) elongation of lactose on the Gal unit, it formed at 30 min. Structure **6a** is a ( $\beta$ 1 $\rightarrow$ 6) branching of lactose on the Glc unit, it was found at 60 min.

In case of *A. oryzae*  $\beta$ -galactosidase (Figure 3C), the ( $\beta$ 1 $\rightarrow$ 6) elongation on lactose (structure **6b**) was first observed at 5 min, then further elongated with

( $\beta 1 \rightarrow 6$ ), forming structure **34** at 30 min. Structures **11** and **12** were also formed at 5 min, by ( $\beta 1 \rightarrow 3$ ) and ( $\beta 1 \rightarrow 4$ ) elongation on lactose (Table 1). The free Gal and Glc were released at 5 min, then the Gal was elongated with ( $\beta 1 \rightarrow 6$ ) and ( $\beta 1 \rightarrow 3$ ) linkages forming structures **3** at 120 min, and structure **38** at 240 min, respectively (Figure 3C, Table 1). Free Glc was used as acceptor substrate forming structures **8a** and **8b** at 60 min, and structure **4** at 120min. Structure **4** was further elongated with ( $\beta 1 \rightarrow 6$ ) linkage forming structure **39** at 240 min (Figure 3C, Table 1). Structure **6a** was formed at 30 min by ( $\beta 1 \rightarrow 6$ ) elongation on the Glc unit of lactose.

Analysis of the GOS synthesis profiles in time thus allowed deduction of the formation routes of the different GOS structures in time. This clearly revealed that rBgaD-D has a ( $\beta 1 \rightarrow 4$ ) linkage preference while Lactozyme 2600L and *A. oryzae*  $\beta$ -galactosidases show a preference of ( $\beta 1 \rightarrow 6$ ) linkage formation .

### Changes of major GOS composition

The maximum GOS yields were calculated as follows (Table 2): rBgaD-D, 48.3 $\pm$ 1.2% GOS yield with 88.4 $\pm$ 0.4% lactose consumption at 8 h; Lactozyme 2600L, 34.9 $\pm$ 1.8% GOS yield with 91.8 $\pm$ 0.8% lactose consumption at 6 h;  $\beta$ -galactosidase from *A. oryzae*, 19.5 $\pm$ 2.2% GOS yield with 69.6 $\pm$ 1.1% lactose consumption at 8 h. The GOS composition and yield of  $\beta$ -galactosidase enzymes depend on the ratio between its transgalactosylation and hydrolysis. A high transgalactosylation/hydrolysis ratio is beneficial for GOS yield. When the GOS yield reached the highest points, the GOS/Gal factor was 10.9, 2.1, and 1.5 for rBgaD-D, Lactozyme 2600L, and *A. oryzae*  $\beta$ -galactosidase, respectively. This also explains why rBgaD-D has the highest GOS yield (Table 2).

The kinetic changes of the major GOS fractions produced by the 3 enzymes (Figure 4) were followed by HPAEC-PAD integrations. Due to lack of suitable calibration standards, the assumption was used that all GOS structures have the

same response on the PAD detector. Structure **11** is a ( $\beta$ 1 $\rightarrow$ 4) elongation of lactose and the first GOS structure produced by rBgaD-D; it reached its highest yield at 45 min, then decreased quickly (Figure 4A). Gal is only released in small amounts, indicating that the decline in **11** is the result of further transgalactosylation rather than of hydrolysis. Meanwhile structure **8** increased quickly in 5 h, and then stayed almost stable at longer incubation times. This may reflect the continued release of Glc from lactose, used as acceptor substrate for the formation of **8a** and **8b**; both these structures also are used as acceptor substrate for ( $\beta$ 1 $\rightarrow$ 4) elongation at similar rates. Structure **4**, allolactose, first increased gradually in 5 h, and then increased quickly. As a result, structures **8** and **4** became the major GOS structures upon longer incubation times with rBgaD-D (Figure 4A). Figure 4B showed that structures **4** and **6** are the major GOS fractions produced by Lactozyme 2600L, which reflects its strong preference for formation of ( $\beta$ 1 $\rightarrow$ 6) linkages. Figure 4C clearly showed that structure **6** is the major GOS fraction produced by *A. oryzae*  $\beta$ -galactosidase. It decreased gradually upon longer incubation times, while structures **3** and **4** increased gradually.

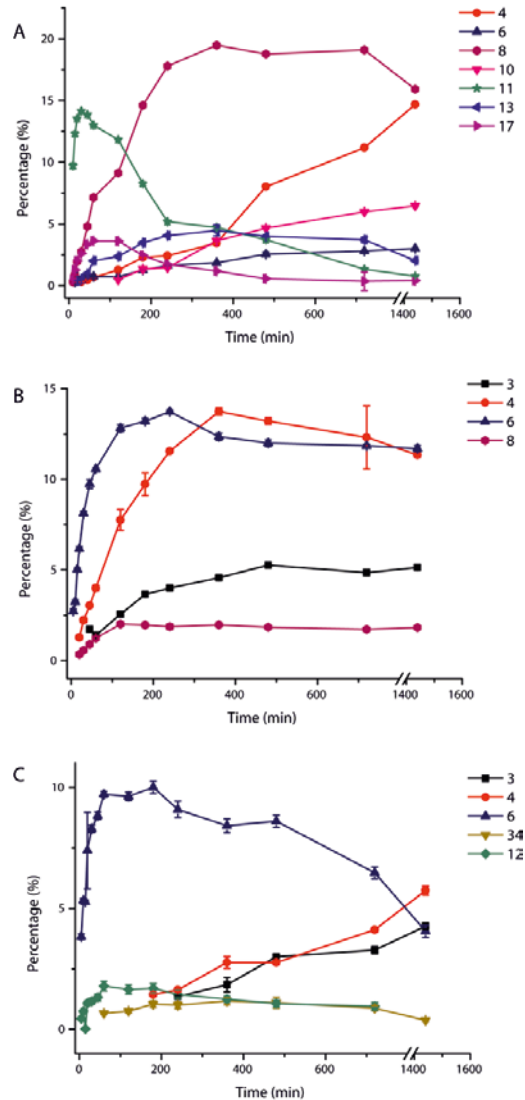
**Table 2.** Reaction time, lactose consumption, and GOS/galactose factor at the highest GOS yield of the three enzymes. The enzyme amounts and reaction conditions used are as shown in Figure 2.

Enzyme source	GOS yield <sup>1</sup> (%)	GOS/Gal Factor <sup>2</sup>	Lactose consumption (%)	Reaction time <sup>3</sup> (h)
<i>B. circulans</i>	48.3 $\pm$ 1.2	10.9	88.4 $\pm$ 0.4	8
<i>K. lactis</i>	34.9 $\pm$ 1.8	2.1	91.8 $\pm$ 0.8	6
<i>A. oryzae</i>	19.5 $\pm$ 2.2	1.5	69.3 $\pm$ 1.1	8

<sup>1</sup>Determined from the peak intensities, data obtained by three parallel experiments. Only the peaks labeled in Figure 2 (except lactose, Gal, and Glc) were used to determine the GOS yield.

<sup>2</sup>Representing the transgalactosylation/hydrolysis ratio, determined by the peak intensities of GOS and galactose.

<sup>3</sup>Reaction time at which the maximum GOS yield was obtained.

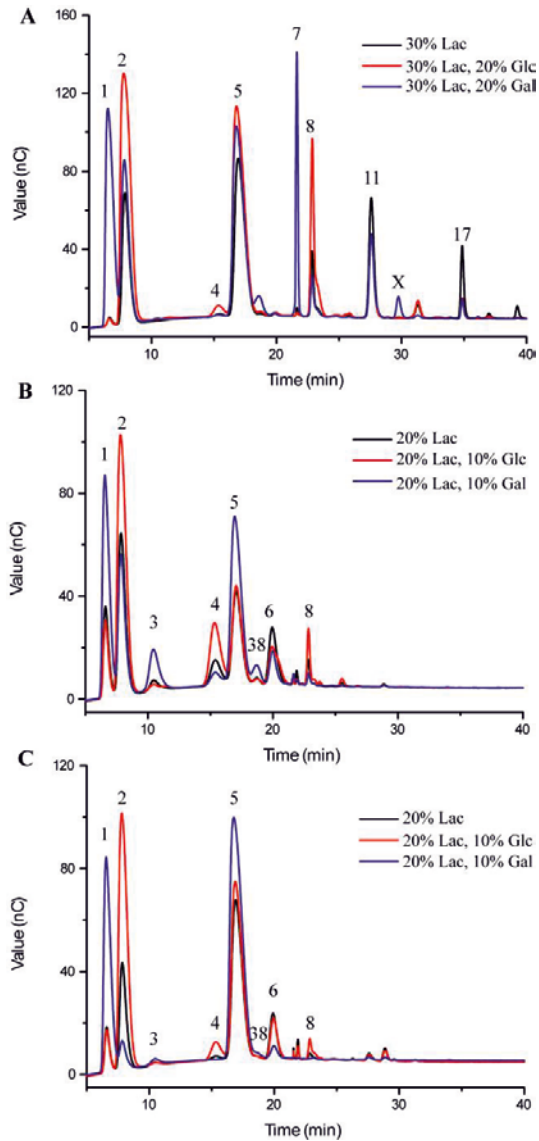


**Figure 4.** The effects of Gal and Glc on the GOS profiles of the 3  $\beta$ -galactosidases incubated with lactose: (A) rBgaD-D from *B. circulans*, (B) Lactozyme 2600L from *K. lactis*, (C)  $\beta$ -galactosidase from *A. oryzae*. Profiles were obtained using 37 U of the enzymes per g of lactose, incubated for 2 h at their respective optimal conditions. A mixture of 30% (w/w) lactose (plus 20% (w/w) Gal or Glc) was used for rBgaD-D. For the Lactozyme 2600L and *A. oryzae*  $\beta$ -galactosidases, a mixture containing 20% (w/w) lactose (plus 10% (w/w) Gal or Glc) was used.

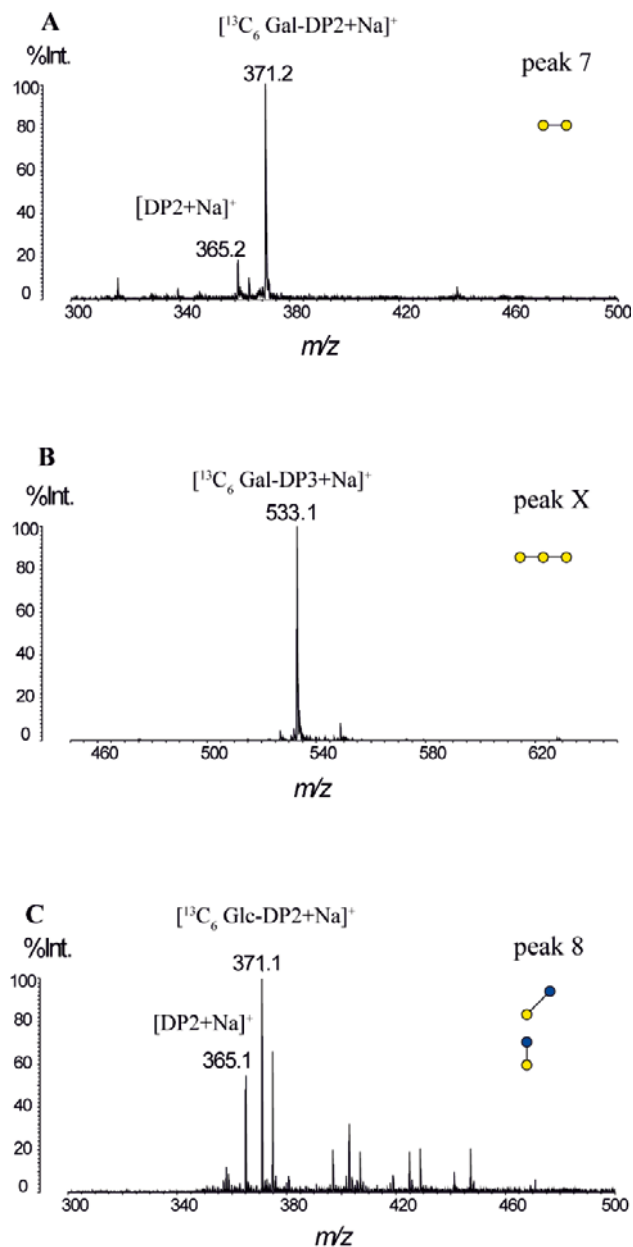
The kinetic changes in the major GOS fractions of these three enzymes showed that the GOS composition strongly depends on the enzyme source and is dynamically related to the reaction time.

### **The effects of Gal and Glc on GOS synthesis**

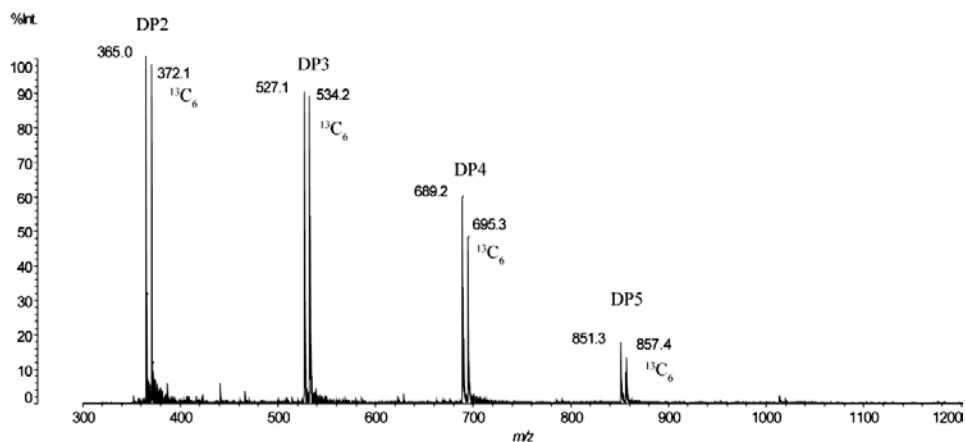
To test whether Gal and Glc were just acceptor substrates, or whether they inhibited enzyme activity, reactions were performed with a high concentration of either monosaccharide from the start. The addition of Gal or Glc significantly changed the GOS profiles of all 3  $\beta$ -galactosidase enzymes. For example, when Gal was added in the reaction with rBgaD-D, disaccharide  $\beta$ -D-Galp-(1 $\rightarrow$ 4)-D-Galp (structure **7**, Figure 5A) increased significantly. This also resulted in appearance of an additional peak (structure **X**) in the product profile. This peak was isolated for NMR analysis and characterized as the ( $\beta$ 1 $\rightarrow$ 4) elongation of structure **7** (trisaccharide  $\beta$ -D-Galp-(1 $\rightarrow$ 4)-D-Galp-(1 $\rightarrow$ 4)-D-Galp) (Supplementary, Table S1). When Glc was used in the reaction, structure **8** had a much higher yield (Figure 5A); also structure **4** increased. The changed GOS profiles clearly suggested that Gal and Glc were used as additional acceptor substrates. To investigate this in more detail,  $^{13}\text{C}_6$  labeled Gal or Glc together with lactose were used in incubations with rBgaD-D. Structures **7**, **X** and **8** were isolated to check their masses using MALDI-TOF-MS (Figure 6). The intensities of the peaks and the  $m/z$  showed that majority of structure **7** is formed from the  $^{13}\text{C}_6$  labelled Gal, only a minority is formed from the normal Gal released during the reaction with lactose (Figure 6A). Figure 6B showed that the DP3 structure **X** contains mostly  $^{13}\text{C}_6$  labelled Gal. Structures **8a** and **8b** contain both the  $^{13}\text{C}_6$  labelled Glc and the normal Glc released during the reaction with lactose (Figure 6C). These results confirm that Gal and Glc were used as acceptor substrates in the formation of GOS. The MALDI-TOF-MS analysis of the reaction mixture with  $^{13}\text{C}_6$  Glc showed that the  $^{13}\text{C}_6$  Glc was incorporated up to DP5 structures (Figure 7).



**Figure 5.** The effects of Gal and Glc on the GOS profiles of the 3  $\beta$ -galactosidases incubated with lactose: (A) rBgaD-D from *B. circulans*, (B) Lactozyme 2600L from *K. lactis*, (C)  $\beta$ -galactosidase from *A. oryzae*. Profiles were obtained using 37 U of the enzymes per g of lactose, incubated for 2 h at their respective optimal conditions. A mixture of 30% (w/w) lactose (plus 20% (w/w) Gal or Glc) was used for rBgaD-D. For the Lactozyme 2600L and *A. oryzae*  $\beta$ -galactosidases, a mixture containing 20% (w/w) lactose (plus 10% (w/w) Gal or Glc) was used.



**Figure 6.** MALDI-TOF-MS spectra of (A) structure 7, (B) structure X, and (C) structure 8. Compared to Figure 5A,  $^{13}\text{C}_6$  Gal or  $^{13}\text{C}_6$  Glc were used in the incubations with lactose and rBgaD-D, then the peak fractions were isolated for MALDI-TOF-MS analysis.



**Figure 7.** MALDI-TOF-MS analysis of the reaction mixture of rBgaD-D, incubated with 30% (w/w) lactose plus 20% (w/w)  $^{13}\text{C}_6$  Glc.

These results indicate that structures **3** and **7** are only formed when enough Gal is released; this is also supported by the formation process followed in time (Figure 3). Structures **4**, **8a** and **8b** are only formed when a certain Glc threshold is reached, as was also evident when following the reaction in time (Figure 3). Previously it was suggested that allolactose (structure **4**) particularly was formed from lactose by a rearrangement of the Glc in the active site of  $\beta$ -galactosidase from *E. coli* [38]. Our data, combined with the observations following the reaction in time, indicate that allolactose (structure **4**) as well as structures **8a** and **8b** were formed from Glc, acting as acceptor substrate in the transgalactosylation reaction catalyzed by the  $\beta$ -galactosidase from *B. circulans*.

When *K. lactis*  $\beta$ -galactosidase Lactozyme 2600L was incubated with 10% w/w Gal peaks **3** and **38** increased (Figure 5B), showing that Gal can be used as an acceptor substrate as well. When incubated with Glc, peaks **4** and **8** increased, indicating that also Glc can be used as acceptor substrate by Lactozyme 2600L of *K. lactis*.

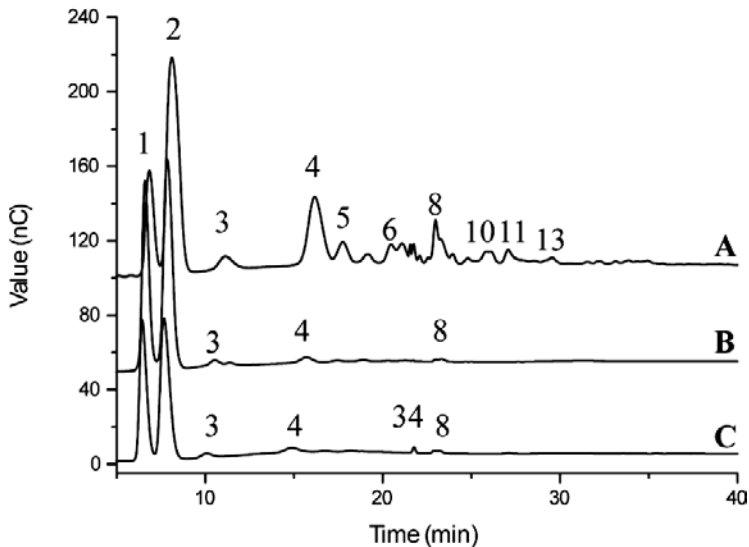


Addition of 10% w/w Gal to the incubation with  $\beta$ -galactosidase of *A. oryzae* resulted in a much lower GOS yield (Figure 5C), indicating that Gal is an inhibitor for this enzyme; all peaks were reduced, also **3** and **38**, the (putative) products of the reaction with the Gal acceptor substrate. Adding Glc to the reaction resulted in increased levels of structures **4** and **8**, but to a lower extent than for the other two  $\beta$ -galactosidase enzymes.

### Reaction end-point profiles

GOS generally are regarded as intermediate reaction products that ultimately will be hydrolyzed into Gal and Glc. In the above we have studied the GOS product profiles at maximum yields. To study the final, end-point GOS profiles of each of the 3 enzymes, a high enzyme dose of 370 U/g lactose was incubated at the respective optimum conditions (Figure 8). GOS initially produced by the Lactozyme 2600L and *A. oryzae*  $\beta$ -galactosidases were almost entirely hydrolyzed after prolonged incubation for 48 h (Figure 8 B and C). Peaks **3** and **4** represent the remaining disaccharides  $\beta$ -D-Galp-(1 $\rightarrow$ 6)-D-Galp and  $\beta$ -D-Galp-(1 $\rightarrow$ 6)-D-Glcp. In contrast, there were still many GOS disaccharides and trisaccharides left in the rBgaD-D incubation, even after 72 h (Figure 8 A). An earlier study also reported that there were still large amounts of GOS and lactose left even after 400 h of incubation, and this was suggested to be caused by the inactivation of the Biolactase (commercial preparation of  $\beta$ -galactosidase from *B. circulans*) [34]. A study on the stability of  $\beta$ -galactosidase from *B. circulans* showed that the enzyme retained 27% activity after incubation for 24 h at 60 °C in 30% (w/w) lactose solution [39]. The authors also indicated that a high initial lactose concentration had a large positive effect on the enzyme activity and stability. Our results showed that, when incubated with 50% (w/w) initial lactose at 60 °C, the enzyme retained 65% activity after 24 h, 37% activity after 48 h, and 28% activity after 72 h. In view of these results, it is unlikely that inactivation of the rBgaD-D enzyme can explain the remaining GOS and lactose in our study.

Moreover, addition of a fresh dose of this enzyme after 48 h did not induce further changes in the final GOS profile (not shown).



**Figure 8.** HPAEC-PAD analysis of the final GOS profiles in incubations with 370 U/g lactose of (A) rBgaD-D of *B. circulans*, incubated for 72 h; (B) Lactozyme 2600 L of *K. lactis*, incubated for 48 h; (C)  $\beta$ -galactosidase from *A. oryzae*, incubated for 48 h. All 3  $\beta$ -galactosidases were incubated at their respective optimum conditions.

## Conclusions

This study has identified several parameters that influence the final GOS profiles of the  $\beta$ -galactosidases from *B. circulans*, *K. lactis*, and *A. oryzae*. Firstly, the  $\beta$ -galactosidase from *B. circulans* has a relatively high transgalactosylation/hydrolysis ratio (the GOS/Gal factor is 10.9, Table 2) in comparison with the other two enzymes, resulting in a high yield of GOS. The formation process of the rBgaD-D, showed a complex formation pattern (Figure 3A), in which lactose analogues **8a**, **8b** and **4** were formed from acceptor reactions with released Glc, rather than rearrangement reactions with Glc retained in the active site. This was also evident from the observation that the lactose analogues were not formed until a certain level of free Glc was formed. This was

further confirmed by reactions with added Glc, showing higher yields when Glc was present in high levels at the start of the reaction. Similar results were found for the *A. oryzae* and *K. lactis* enzymes, i.e. first free Glc was formed, and only later the lactose analogues were observed. The GOS formed by rBgaD-D are much more complex than the GOS from *A. oryzae* and *K. lactis* enzymes. Although the latter enzymes both had a preference for ( $\beta 1 \rightarrow 6$ )-elongation, there is a difference also. The *A. oryzae* enzyme can also synthesize ( $\beta 1 \rightarrow 3$ ) and ( $\beta 1 \rightarrow 4$ )-elongations. The *K. lactis* enzyme showed a higher yield of complex GOS, but was unable to introduce ( $\beta 1 \rightarrow 4$ )-elongations.

Following the reactions to the end-point, it was observed that structures containing ( $\beta 1 \rightarrow 6$ ) linkages, like  $\beta$ -D-Galp-(1 $\rightarrow$ 6)-D-Glcp are difficult for rBgaD-D to hydrolyze, thus these structures are likely to accumulate during the reaction. For rBgaD-D the reaction ends in a dynamic equilibrium with relatively high amounts of structures with  $\beta$ -D-Galp-(1 $\rightarrow$ 6)- elements. For *A. oryzae* and *K. lactis* enzymes the reaction ends with almost complete hydrolysis, showing only trace amounts of residual GOS after 48 h incubation with high enzyme activity.

Finally, the produced Gal and Glc did not inhibit the transgalactosylation reaction of rBgaD-D as shown in Figure 5A; instead Gal and Glc can be used as acceptor substrates to synthesize GOS derivatives. Incubation with high initial Gal and Glc also showed that *K. lactis* is capable of effectively using both monosaccharides as an acceptor substrate and found no inhibitory effect. In the case of *A. oryzae* it was clear that Gal had an inhibitory effect on the enzyme, resulting in a lower GOS yield when incubated with a high starting concentration of Gal. In case of added Glc, similar GOS yields were achieved.

In summary, it is concluded that *A. oryzae* and *K. lactis* are enzymes with similar properties, but also showing minor differences in GOS production capacities. The *B. circulans* derived rBgaD-D enzyme produces a more complex product profile,

and has a significantly higher transgalactosylation capacity, resulting in a more dynamic GOS profile, also at longer reaction times.

### Abbreviations used

GOS, galactooligosaccharides; HPAEC-PAD, high-pH anion-exchange chromatography coupled with pulsed amperometric detection; NMR, nuclear magnetic resonance. MALDI-TOF-MS, Matrix Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry; Gal, galactose; Glc, glucose; Lac, lactose.

### Acknowledgments

This work was financially by China Scholarship Council (to HY) and the University of Groningen (to LD and SL). We thank Prof. Johannis P. Kamerling for his knowledge and advice, and Bas J. Kuipers (FrieslandCampina) for his support in the early stages of this work.

### References

- 1 Barile, D. and Rastall, R. A. (2013). Human milk and related oligosaccharides as prebiotics. *Curr. Opin. Biotechnol.* **24**, 214–219.
- 2 Petherick, A. (2010). Development: Mother's milk: A rich opportunity. *Nature*, **468**, S5–S7.
- 3 Bode, L. (2012). Human milk oligosaccharides: Every baby needs a sugar mama. *Glycobiology*, **22**, 1147–62.
- 4 Boehm, G., Fanaro, S., Jelinek, J., Stahl, B. and Marini, A. (2003). Prebiotic concept for infant nutrition. *Acta. Paediatr. Suppl.* **91**, 64–67.
- 5 Vandenplas, Y., De Greef, E. and Veereman, G. (2015). Prebiotics in infant formula. *Gut Microbes*, **5**, 681–687.
- 6 Fanaro, S., Boehm, G., Garssen, J., Knol, J., Mosca, F., Stahl, B. and Vigi, V. (2005). Galacto-oligosaccharides and long-chain fructo-oligosaccharides as prebiotics in infant formulas: a review. *Acta. Paediatr. Suppl.* **94**, 22–26.

- 7 Oozeer, R., Van Limpt, K., Ludwig, T., Amor, K. Ben, Martin, R., Wind, R. D., Boehm, G. and Knol, J. (2013). Intestinal microbiology in early life: Specific prebiotics can have similar functionalities as human-milk oligosaccharides. *Am. J. Clin. Nutr.* **98**, 561S-71S.
- 8 Bultema, J. B., Kuipers, B. J. H. and Dijkhuizen, L. (2014). Biochemical characterization of mutants in the active site residues of the  $\beta$ -galactosidase enzyme of *Bacillus circulans* ATCC 31382. *FEBS Open Bio.* **4**, 1015–20.
- 9 Torres, D. P. M., Gonçalves, M. D. P. F., Teixeira, J. A. and Rodrigues, L. R. (2010). Galacto-oligosaccharides: Production, properties, applications, and significance as prebiotics. *Compr. Rev. Food Sci. Food Saf.* **9**, 438–454.
- 10 Gosling, A., Stevens, G. W., Barber, A. R., Kentish, S. E. and Gras, S. L. (2010). Recent advances refining galactooligosaccharide production from lactose. *Food Chem.* **121**, 307–318.
- 11 Otieno, D. O. (2010). Synthesis of  $\beta$ -galactooligosaccharides from lactose using microbial  $\beta$ -galactosidases. *Compr. Rev. Food Sci. Food Saf.* **9**, 471–482.
- 12 van Leeuwen, S. S., Kuipers, B. J. H., Dijkhuizen, L. and Kamerling, J. P. (2016a). Comparative structural characterization of 7 commercial galacto-oligosaccharide (GOS) products. *Carbohydr. Res.* **425**, 48–58.
- 13 Neri, D. F. M., Balcão, V. M., Costa, R. S., Rocha, I. C. A. P., Ferreira, E. M. F. C., Torres, D. P. M., Rodrigues, L. R. M., Carvalho Jr., L. B. and Teixeira, J. A. (2009). Galacto-oligosaccharides production during lactose hydrolysis by free *Aspergillus oryzae*  $\beta$ -galactosidase and immobilized on magnetic polysiloxane-polyvinyl alcohol. *Food Chem.* **115**, 92–99.
- 14 Cataldi, T. R., Campa, C. and De Benedetto, G. E. (2000). Carbohydrate analysis by high-performance anion-exchange chromatography with pulsed amperometric detection: the potential is still growing. *Fresenius J. Anal. Chem.* **368**, 739–758.
- 15 Zhou, Q. Z. K. and Chen, X. D. (2001). Effects of temperature and pH on the catalytic activity of the immobilized  $\beta$ -galactosidase from *Kluyveromyces lactis*. *Biochem. Eng. J.* **9**, 33–40.
- 16 Warmerdam, A., Zisopoulos, F. K., Boom, R. M. and Janssen, A. E. M. (2014). Kinetic characterization of galacto-oligosaccharide (GOS) synthesis by three commercially important  $\beta$ -galactosidases. *Biotechnol. Prog.* **30**, 38–47.
- 17 Song, J., Abe, K., Imanaka, H., Imamura, K., Minoda, M., Yamaguchi, S. and Nakanishi, K. (2011). Causes of the production of multiple forms of  $\beta$ -galactosidase by *Bacillus circulans*. *Biosci. Biotechnol. Biochem.* **75**, 268–78.
- 18 Song, J., Imanaka, H., Imamura, K., Minoda, M., Katase, T., Hoshi, Y., Yamaguchi, S. and Nakanishi, K. (2011). Cloning and expression of a  $\beta$ -galactosidase gene of *Bacillus circulans*. *Biosci. Biotechnol. Biochem.* **75**, 1194–7.

- 19 Ishikawa, K., Kataoka, M., Yanamoto, T., Nakabayashi, M., Watanabe, M., Ishihara, S. and Yamaguchi, S. (2015). Crystal structure of  $\beta$ -galactosidase from *Bacillus circulans* ATCC 31382 (BgaD) and the construction of the thermophilic mutants. *FEBS J.* **282**, 2540–2552.
- 20 Warmerdam, A., Paudel, E., Jia, W., Boom, R. M. and Janssen, A. E. M. (2013). Characterization of  $\beta$ -galactosidase isoforms from *Bacillus circulans* and their contribution to GOS production. *Appl. Biochem. Biotechnol.*, **170**, 340–58.
- 21 van Leeuwen, S. S., Kuipers, B. J. H., Dijkhuizen, L. and Kamerling, J. P. (2014). <sup>1</sup>H NMR analysis of the lactose/ $\beta$ -galactosidase-derived galacto-oligosaccharide components of Vivinal® GOS up to DP5. *Carbohydr. Res.* **400**, 59–73.
- 22 van Leeuwen, S. S., Kuipers, B. J. H., Dijkhuizen, L. and Kamerling, J. P. (2016b). Corrigendum to “<sup>1</sup>H NMR analysis of the lactose/ $\beta$ -galactosidase-derived galacto-oligosaccharide components of Vivinal® GOS up to DP5” [*Carbohydr. Res.* **400** (2014) 59–73]. *Carbohydr. Res.* **419**, 69-70
- 23 Pereira-Rodríguez, A., Fernández-Leiro, R., González-Siso, M. I., Cerdán, M. E., Becerra, M. and Sanz-Aparicio, J. (2012). Structural basis of specificity in tetrameric *Kluyveromyces lactis*  $\beta$ -galactosidase. *J. Struct. Biol.* **177**, 392–401.
- 24 Rodríguez-Colinas, B., De Abreu, M. A., Fernández-Arrojo, L., De Beer, R., Poveda, A., Jimenez-Barbero, J., Haltrich, D., Olmo, A. O. B., Fernández-Lobato, M. and Plou, F. J. (2011). Production of galacto-oligosaccharides by the  $\beta$ -galactosidase from *kluyveromyces lactis*: Comparative analysis of permeabilized cells versus soluble enzyme. *J. Agric. Food Chem.* **59**, 10477–10484.
- 25 Maksimainen, M. M., Lampio, A., Mertanen, M., Turunen, O. and Rouvinen, J. (2013). The crystal structure of acidic  $\beta$ -galactosidase from *Aspergillus oryzae*. *Int. J. Biol. Macromol.* **60**, 109–115.
- 26 Urrutia, P., Fernandez-arrojo, L., Ballesteros, A. O., Wilson, L. and Plou, F. J. (2013). Detailed analysis of galactooligosaccharides synthesis with  $\beta$ -galactosidase from *Aspergillus oryzae*. *J. Agric. Food Chem.* **61**, 1081–1087.
- 27 Gaur, R., Pant, H., Jain, R. and Khare, S. K. (2006). Galacto-oligosaccharide synthesis by immobilized *Aspergillus oryzae*  $\beta$ -galactosidase. *Food Chem.* **97**, 426–430.
- 28 Benjamins, E. (2014) *Galacto-oligosaccharide synthesis using immobilized  $\beta$ -galactosidase*. PhD thesis, University of Groningen
- 29 Kovács, Z., Benjamins, E., Grau, K., Rehman, A. U., Ebrahimi, M. and Czermak, P. (2014). Recent developments in manufacturing oligosaccharides with prebiotic functions. *Adv. Biochem. Eng. Biotechnol.* **143**, 257–295.
- 30 Vera, C., Guerrero, C., Conejeros, R. and Illanes, A. (2012). Synthesis of galacto-oligosaccharides by  $\beta$ -galactosidase from *Aspergillus oryzae* using partially

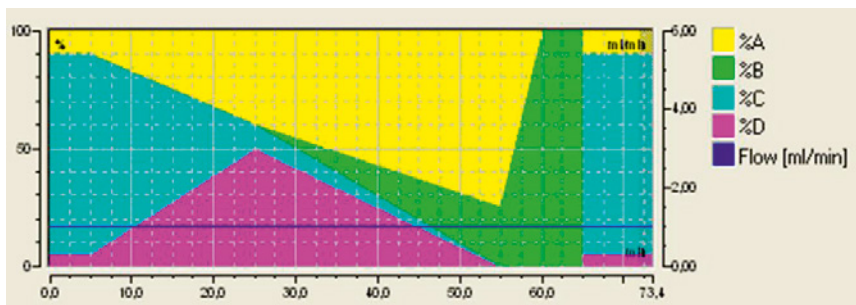
- dissolved and supersaturated solution of lactose. *Enzyme Microb. Technol.* **50**, 188–194.
- 31 Palai, T., Mitra, S. and Bhattacharya, P. K. (2012). Kinetics and design relation for enzymatic conversion of lactose into galacto-oligosaccharides using commercial grade  $\beta$ -galactosidase. *J. Biosci. Bioeng.* **114**, 418–423.
- 32 Vera, C., Guerrero, C. and Illanes, A. (2011). Determination of the transgalactosylation activity of *Aspergillus oryzae*  $\beta$ -galactosidase: Effect of pH, temperature, and galactose and glucose concentrations. *Carbohydr. Res.* **346**, 745–752.
- 33 Warmerdam, A., Wang, J., Boom, R. M. and Janssen, A. E. M. (2013). Effects of carbohydrates on the *o*NPG converting activity of  $\beta$ -galactosidases. *J. Agric. Food Chem.* **61**, 6458–6464.
- 34 Rodriguez-Colinas, B., Poveda, A., Jimenez-Barbero, J., Ballesteros, A. O. and Plou, F. J. (2012). Galacto-oligosaccharide synthesis from lactose solution or skim milk using the  $\beta$ -galactosidase from *Bacillus circulans*. *J. Agric. Food Chem.* **60**, 6391–6398.
- 35 Frenzel, M., Zerge, K., Clawin-Rädecker, I. and Lorenzen, P. C. (2015). Comparison of the galacto-oligosaccharide forming activity of different  $\beta$ -galactosidases. *LWT - Food Sci. Technol.* **60**, 1068–1071.
- 36 Kamerling, J. P. and Vliegenthart, J. F. G. (1992). High-resolution  $^1\text{H}$ -nuclear magnetic resonance spectroscopy of oligosaccharide-alditols released from mucin-type O-glycoproteins. *Biol. Magn. Reson.* **10**, 1–194.
- 37 Beccati, D. (2009). *Investigations of prebiotics and of inter- and intra-molecular glycan-protein interactions*. PhD thesis, University of Utrecht
- 38 Huber, R. E., Kurz, G. and Wallenfels, K. (1976). A quantitation of the factors which affect the hydrolase and transgalactosylase activities of beta-galactosidase (*E. coli*) on lactose. *Biochemistry*, **15**, 1994–2001.
- 39 Warmerdam, A., Boom, R. M. and Janssen, A. E. (2013).  $\beta$ -Galactosidase stability at high substrate concentrations. *SpringerPlus*, **2**, 402.
- 40 Park, A. R. and Oh, D. K. (2010). Galacto-oligosaccharide production using microbial beta-galactosidase: current state and perspectives. *Appl. Microbiol. Biotechnol.* **85**, 1279–86.

## Supplemental data

**Table S1.** NMR  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts determined from 1D and 2D NMR spectroscopy of **25b** (Main paper, Figures 2, 3) and peak **X** (Main paper, Figure 5A).

	<b>25b</b> $^1\text{H}$	$^{13}\text{C}$	<b>X</b> $^1\text{H}$	$^{13}\text{C}$
A $\alpha$ -1	5.447	93.0	5.271	93.3
A $\alpha$ -2	3.69	80.6	3.89	70.0
A $\alpha$ -3	3.85	72.3	3.95	71.0
A $\alpha$ -4	3.46	70.7	4.226	79.9
A $\alpha$ -5	3.84	72.3	4.129	71.0
A $\alpha$ -6a	3.84	61.8	3.85	61.5
A $\alpha$ -6b	3.77		3.75	
A $\beta$ -1	4.722	nd	4.611	97.4
A $\beta$ -2	3.53	81.0	3.56	73.4
A $\beta$ -3	3.68	73.7	3.75	74.4
A $\beta$ -4	3.44	70.7	4.165	79.1
A $\beta$ -5	3.48	76.6	3.75	75.4
A $\beta$ -6a	3.89	62.0	3.80	62.0
A $\beta$ -6b	3.70		3.77	
B-1	4.657/720	104.4	4.640	105.5
B-2	3.60/53	72.1	3.67	72.7
B-3	3.67	73.7	3.78	74.1
B-4	3.93/91	69.8	4.175	78.3
B-5	3.95	75.3	3.72	75.6
B-6a	4.05	70.2	3.80	62.0
B-6b	3.91		3.77	
C-1	4.44-4.47	104.4	4.596	104.4
C-2	3.55	71.9	3.60	72.6
C-3	3.69	73.7	3.66	73.9
C-4	3.97	69.7	3.904	69.9
C-5	3.70	76.3	3.68	76.4
C-6a	3.82	62.1	3.80	62.0
C-6b	3.75		3.77	





**Figure S1.** Separation conditions for HPAEC-PAD analysis

### Structure **25b**

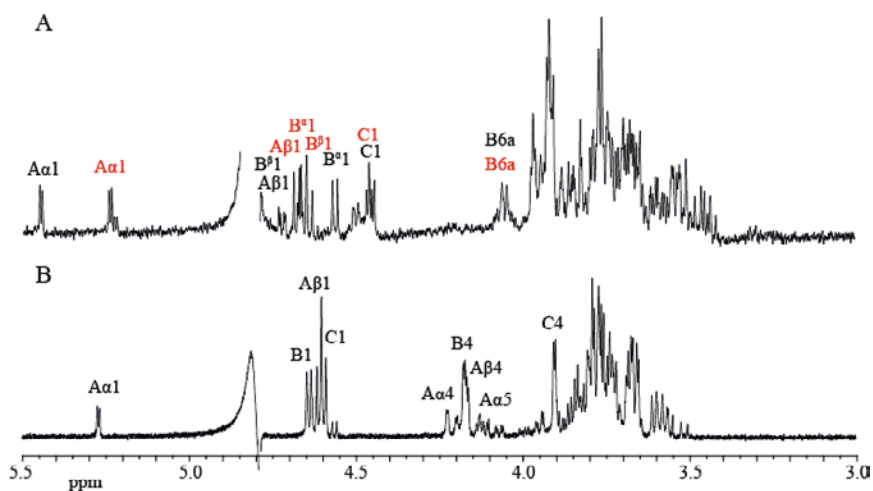
MALDI-TOF-MS analysis of the fraction containing structure **25b** showed only one peak at 527.3  $m/z$  fitting the sodium adduct of a trisaccharide structure. The 1D  $^1\text{H}$  NMR spectrum of **25** showed anomeric peaks fitting two structures in approximately equal abundance. Further separation by HPAEC-PAD successfully isolated structure **25a**, the 1D  $^1\text{H}$  NMR matched that of structure **25**,<sup>1</sup> indicating  $\beta\text{-D-Galp-(1}\rightarrow\text{6)-}\beta\text{-D-Galp-(1}\rightarrow\text{3)-D-Glcp}$ . For structure **25b** the separation was not successful.

The second structure, **25b**, showed anomeric signals at  $\delta$  5.447 (**A $\alpha$**  H-1), 4.722 (**A $\beta$**  H-1), 4.675 (**B $^a$**  H-1), 4.720 (**B $^b$**  H-1), and 4.44-4.47 (C H-1). The anomeric signals at  $\delta$  5.447 and 4.722 are indicative of a 2-substituted Glc residue. The anomeric signal at  $\delta$  4.44-4.47, which overlaps with C H-1 of **25a**, is deduced from the signal intensity of the mixture. The position of the anomeric signal fits the occurrence of a  $\beta\text{-D-Galp-(1}\rightarrow\text{6)-}$  residue. Since the isolation of pure **25b** was not successful, all  $^1\text{H}$  and most  $^{13}\text{C}$  chemical shifts were deduced from 2D spectroscopy on the peak **25** mixture, since all peaks belonging to **25a** are already known.<sup>1</sup> From the chemical shift pattern of residue **B**, particularly H-6a/C-6, it was found that residue **B** has to be 6-substituted. Moreover, the clear anomeric track of **A $\alpha$**  H-1 showed no evidence of 6-substitution. Together with ROESY interresidual correlations between C H-1 and **B** H-6a and H-6b, and between **B $^a$**  H-1 and **A $\alpha$**  H-2, and between **B $^b$**  H-1 and **A $\beta$**  H-2 these data indicate a  $\beta\text{-D-Galp-(1}\rightarrow\text{6)-}\beta\text{-D-Galp-(1}\rightarrow\text{2)-D-Glcp}$  structure for **25b** (Main paper; Figure 2).

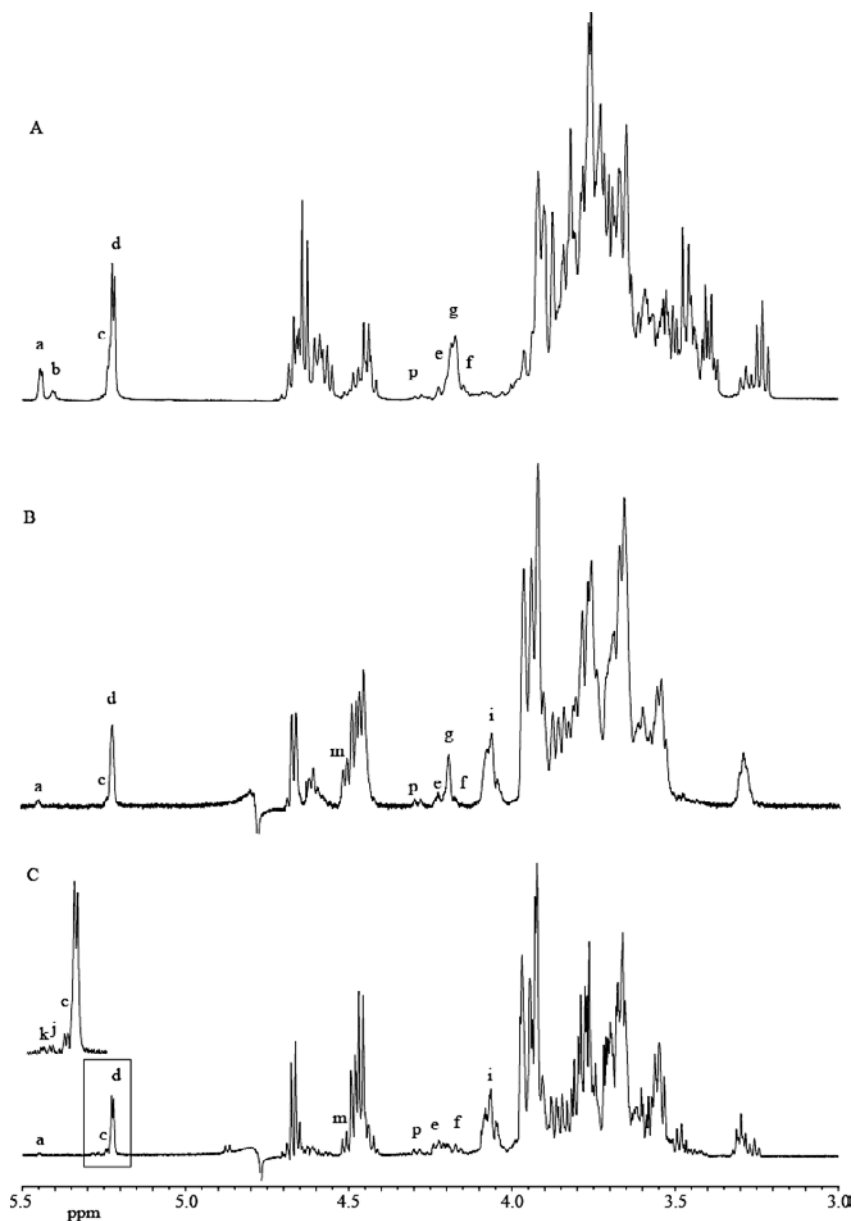
Considering the elution position of **25b**, together with **25a**, fits observations for elongations of **8a** and **8b** with ( $\beta$ 1 $\rightarrow$ 4) (**13a** and **13b**) and also for structures elongated with ( $\beta$ 1 $\rightarrow$ 3) found previously (structures **29** and **30**) in reference 1.

### Structure X

The 1D  $^1\text{H}$  NMR spectrum (Figure S1) showed anomeric signals at  $\delta$  5.271 (**A $\alpha$**  H-1), 4.611 (**A $\beta$**  H-1), 4.640 (**B** H-1) and 4.596 (**C** H-1). From 2D NMR spectra all  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts could be assigned (Table S1). The chemical shift pattern of residue **A $\alpha$**  matches with that of 4-substituted reducing Gal.<sup>2</sup> Particularly H-4:C-4 at  $\delta$  4.226:79.9 and H-5 at  $\delta$  4.129 are indicative of a 4-substitution. The 4-substitution of the reducing Gal residue is also reflected in the **A $\beta$**  H-4:C-4 at  $\delta$  4.165:79.1 ppm. Residue **B** showed the pattern of a 4-substituted residue as well, reflected in the H-4:C-4 at  $\delta$  4.175:78.3 ppm. Finally, residue **C** showed the chemical shift pattern of a terminal residue, linked via ( $\beta$ 1 $\rightarrow$ 4) bond, as indicated by the H-4 at  $\delta$  3.904 ppm. Furthermore, ROESY spectroscopy showed interresidual correlations between **C** H-1 and **B** H-4 and between **B** H-1 and **A $\alpha$**  and **A $\beta$**  H-4 signals. These data result in a structure of  $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ 4)-D-Galp for **X** (Main paper; Figure 5A).



**Figure S2.** 1D  $^1\text{H}$  NMR spectra of (A) mixture **25**, and (B) structure **X**. Distinctive peaks are marked. In spectrum (A) Red labels belong to **25a**, whereas black labels belong to **25b**.



**Figure S4.** 1D  $^1\text{H}$  NMR spectra of GOS produced by incubation of (A) rBgaD-D with 50 % (w/w) lactose at 60 °C and pH 6.0, (B) Lactozyme L2600 with 30 % (w/w) lactose at 45 °C and pH 4.5 and (C) *A. oryzae*  $\beta$ -galactosidase incubated with 30 % (w/w) lactose at 40 °C and pH 7.0. Structural-reporter-group signals **a-p** are explained in Table S1, adapted from reference 1.

Table S2. <sup>1</sup>H NMR structural-reporter-group signals used for the evaluation of 1D <sup>1</sup>H NMR spectra of GOS preparations (references 1, 2, 3; this study). Structure numbers used here and throughout the main paper match those used in reference 1. Table is adapted from reference 1.

Signal Code	Chemical shift	Explanation	Structures
<b>A</b>	5.45	H-1 of a 2-substituted $\alpha$ -D-Glcp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 2)- $\alpha$ -D-Glcp]	<b>8a, 9, 10a, 15b, 16b, 20a, 21a, 30, 33, yb</b>
<b>B</b>	5.42-5.40	H-1 of a 2-substituted $\alpha$ -D-Glcp unit, which bears at O-2 a $\beta$ -D-Galp unit, elongated at O-4 with another $\beta$ -D-Galp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-Galp-(1 $\rightarrow$ 2)- $\alpha$ -D-Glcp]	<b>13a, 15a, 16a, 18a, 20bc, 21bc, 23a, ya</b>
<b>C</b>	5.25-5.23	H-1 of a 3-substituted $\alpha$ -D-Glcp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\alpha$ -D-Glcp]	<b>8b, 10b, 13b, 18b, 23b, 25, 27, 29, 32</b>
<b>D</b>	5.23-5.22	H-1 of a 4-, 6-, and/or 1-substituted $\alpha$ -D-Glcp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 4/6)- $\alpha$ -D-Glcp; $\alpha$ -D-Glcp-(1 $\leftrightarrow$ 1)- $\beta$ -D-Galp]	<b>4, 5, 6ab, 11, 12, 14ab, 16c, 17, 19a, 22, 24, 26, 28, 31, 34, 35, 36, 37, 39, xabc, yab, zabc</b>
<b>E</b>	4.22-4.21	H-1 of free $\alpha$ -D-Glcp	<b>2</b>
<b>F</b>	4.17-4.16	H-6a of a 6-substituted $\beta$ -D-Glcp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\beta$ -D-Glcp]	<b>4, 10ab, 16abc, 21abc, 28, 39</b>
<b>G</b>	4.21-4.17	H-6a of a 6-substituted $\alpha$ -D-Glcp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\alpha$ -D-Glcp]	<b>4, 10ab, 16abc, 21abc, 28, 39</b>
<b>H</b>	5.26	H-4 of a 3- and/or 4-substituted (reducing) $\beta$ -D-Galp unit; { $\beta$ -D-Galp-(1 $\rightarrow$ 3/4)- $\beta$ -D-Galp[-(1 $\rightarrow$ )]}	<b>7, 11, 12, 13ab, 14ab, 15ab, 16abc, 17, 18ab, 19abc, 20abc, 21abc, 22, 23ab, 24, 26, 27, 28, 29, 30, 31, 32, 33, xabc, yab, zabc</b>
<b>I</b>	4.08-4.05	H-1 of free $\alpha$ -D-Galp	<b>1</b>
<b>J</b>	5.27	H-6a of a 6-substituted (reducing) $\beta$ -D-Galp unit; { $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\beta$ -D-Galp[-(1 $\rightarrow$ )]}	<b>3, 6b, 25, 26, 27</b>
<b>K</b>	5.28	H-1 of a 4- and/or 6-substituted reducing $\alpha$ -D-Galp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 4/6)- $\alpha$ -D-Galp]	<b>3, 7, X</b>
<b>L</b>	5.28	H-1 of a 3-substituted reducing $\alpha$ -D-Galp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\alpha$ -D-Galp]	<b>38</b>
<b>M</b>	4.51-4.52	H-1 of a 3-substituted $\beta$ -D-Galp unit in a $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp sequence	<b>12, 31</b>
<b>N</b>	4.23	H-4 of a 3,6-disubstituted $\beta$ -D-Galp-(1 $\rightarrow$ 4)- unit, H-4 of a 3-substituted $\beta$ -D-Galp-(1 $\rightarrow$ 4)- unit in a $\beta$ -D-Galp-(1 $\rightarrow$ 6)- $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp sequence	<b>36, 37</b>
<b>P</b>	4.29-4.28	H-6a of a 4,6-disubstituted $\beta$ -D-Glcp unit; [ $\beta$ -D-Galp-(1 $\rightarrow$ 4,6)- $\beta$ -D-Glcp]	<b>6a, 14ab, 19abc</b>

## References

- 1 van Leeuwen, S. S., Kuipers, B. J. H., Dijkhuizen, L. and Kamerling, J. P. (2016a). Comparative structural characterization of 7 commercial galacto-oligosaccharide (GOS) products. *Carbohydr. Res.* **425**, 48–58.
- 2 Van Leeuwen S. S., Kuipers, B. J. H., Dijkhuizen, L. and Kamerling, J. P. (2014) Development of a  $^1\text{H}$  NMR structural-reporter-group concept for the analysis of prebiotic galacto-oligosaccharides of the  $[\beta\text{-D-Galp-(1}\rightarrow\text{x)}]_n\text{-D-Glcp}$  type. *Carbohydr. Res.* **400**, 54-58.
- 3 van Leeuwen, S. S., Kuipers, B. J. H., Dijkhuizen, L. and Kamerling, J. P. (2014).  $^1\text{H}$  NMR analysis of the lactose/ $\beta$ -galactosidase-derived galacto-oligosaccharide components of Vivinal® GOS up to DP5. *Carbohydr. Res.*, **400**, 59–73. Corrigendum 2016, *Carbohydr. Res.* **419**, 69-70.