Contents lists available at ScienceDirect

Food and Bioproducts Processing

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

# Study of Galactooligosaccharides production from dairy waste by FTIR and chemometrics as Process Analytical Technology

*Fabián Rico-Rodriguez*[a], *Lorenzo Strani*[b], *Silvia Grassi*[b,*], *Ruth Lancheros*[c], *Juan Carlos Serrato*[c], *Ernestina Casiraghi*[b]

[a] *Program of Food Engineering, Faculty of Engineering, Universidad de Cartagena – Cartagena de Indias D.T. y C., Colombia*

[b] *Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, Milano, Italy*

[c] *Department of Chemical and Environmental, Faculty of Engineering, Universidad Nacional de Colombia, Bogotá D.C, Colombia*

ARTICLE INFO

*Article history:*
Received 16 October 2020
Received in revised form 12 December 2020
Accepted 20 December 2020
Available online 30 December 2020

*Keywords:*
Prebiotic synthesis
Cheese by-products
Infrared spectroscopy
Process Analytical Technology
Chemometrics
Multivariate Curve Resolution
Partial Least Square regression

ABSTRACT

Galactooligosaccharides (GOS) production from whey, a relevant by-product of dairy industry, answers to the Circular Economy principle of extending the life cycle of products. Indeed, it allows the reuse of dairy waste to produce prebiotics to be used in functional food preparations. For this purpose, the effective monitoring of GOS production should be performed in real time and by environmentally friendly techniques. Thus, FTIR spectroscopy, combined with different chemometric approaches, has been tested to assess a Process Analytical Technology to follow GOS production from cheese whey. Partial Least Square regression models were reliable for lactose, glucose and galactose determination (Root Mean Square Error of Prediction of 21.9, 11.1 and 12.4 mg $mL^{-1}$, respectively). Furthermore, Multivariate Curve Resolution – Alternating Least Square models were proposed to describe trends of the reaction components along the process being an interesting alternative to chromatographic determinations. The real time implementation of the proposed approach will provide the dairy industry with a reliable and green Process Analytical Technology for dairy waste reallocation, avoiding sample pre-processing, large use of organic solvents and long times of analysis.



## 1. Introduction

Galactooligosaccharides (GOS) comprise a group of oligomers, derived from lactose, possessing functional properties since they act as prebiotics and promote benefits in microbial gut and human health (Rodrigues Mano et al., 2019). Indeed, GOS are one of the main prebiotics used for functional food preparations. Among their functional properties, it is possible to mention their ability to increase beneficial gut microbiota, immune system modulation and/or antipathogenic effect (Byfield et al., 2010; Sangwan et al., 2011). Besides their functional characteristics, GOS possess desired technological properties which allows them to be used on different food matrixes like dairy, bakery or beverages, among others (Sangwan et al., 2011).

GOS are produced by lactose transgalactosylation. This reaction is catalysed by β-galactosidase (βgal) enzymes, obtained from a wide variety of microorganisms: bacteria, yeast and/or fungi (Byfield et al., 2010). Two of the most used βgal in industry are obtained from *Kluyveromyces lactis* and *Aspergillus oryzae* (Fischer and Kleinschmidt, 2015). As βgal

* Corresponding author at: Università degli Studi di Milano, Department of Food, Environmental and Nutritional Sciences (DeFENS), via Celoria, 2, 20133, Milano, Italy.
E-mail address: silvia.grassi@unimi.it (S. Grassi).

https://doi.org/10.1016/j.fbp.2020.12.009
0960-3085/

possesses double activity (hydrolytic and transgalactosylation), it is necessary to control the thermodynamics of the reaction by increasing lactose concentration in the media, which favours GOS formation over monosaccharide release (González-Delgado et al., 2016). GOS production depends on many factors like substrate and initial sugar concentration, pH and temperature of reaction, presence of enzyme inhibitors and source of enzyme (Chockchaisawasdee et al., 2004).

It has been suggested that a variation in the GOS degree of polymerisation and linkages plays an important role on the effect of gut microbiota (Akiyama et al., 2015). Different assays have shown that GOS produced by βgal from *K. lactis* and *A. oryzae* differs in yields and composition. Whilst *K. lactis* GOS are mainly composed of di- and trisaccharides, *A. oryzae* βgal produces GOS from di- to hexa-saccharides (Gosling et al., 2010). However, yields for lactose transgalactosylation with *K. lactis* βgal are around 45–50% of total products (Rico-Rodríguez et al., 2018). Nonetheless, with *A. oryzae* βgal this value is lower than 35% (Otieno, 2010).

On the other hand, GOS quantification is generally performed by high performance liquid chromatography (HPLC) (Hernández-Hernández et al., 2012) or gas chromatography (GC) (Ruiz-Matute et al., 2012), two of the most reliable analytical techniques for the purpose. However, these methods require sample pre-processing (GC), large volumes of organic solvents (HPLC) and long analysis time per sample.

Nowadays, the research for quick-environmentally friendly analytical techniques has focused on vibrational spectroscopic technologies, such as mid-infrared, near-infrared and Raman spectroscopy (Moros et al., 2010).

Fourier transformed mid-infrared (FTIR) has been used in enzymatic studies for structural characterization from β-galactosidase immobilization, encapsulation and conjugation (Gennari et al., 2019; Li et al., 2019; Eskandarloo and Abbaspourrad, 2018; Misson et al., 2016) to derivatives examination (Kumar et al., 2020). However, little investigation has been proposed for process monitoring in this field, mainly represented by the in-line anomer concentration measurements in solution proposed by Schiele et al. (2020). Furthermore, Romano et al. (2016) evaluated, by FTIR and Partial Least Squares (PLS) regression, the effect of sucrose concentration on the composition of enzymatically synthesized short-chain fructo-oligosaccharides; similarly, the lactose hydrolysis in milk has been studied by Cocciardi et al. (2004). On the other hand, FTIR proved to be reliable to control processes in a fast, real time and non-destructive way (Grassi et al., 2014), or as a tool for evaluating quality parameters in dairy products (Mohamed et al., 2020) among others.

Despite this, FTIR technique has few drawbacks. Among them, the difficult interpretation of the signals, which are often composed of overlapping spectral bands resulting from absorption of multiple compounds in the sample. In this context, chemometric techniques can help to overcome interpretation problems and they allow to extract relevant information. Spectral data can be modelled by both hard- and soft-modelling methods. Among hard modelling methods, Partial Least Square (PLS) regression is one of the most used for predicting quality parameters from a single spectrum after a proper model calibration, thanks to its ability to handle highly overlapping and colinear data (Martens and Næs, 1989). Instead, Multivariate Curve Resolution – Alternating Least Square (MCR-ALS) (De Juan and Tauler, 2006), a soft-modelling method, is useful to describe trends of the reaction components along the process.

This work aims to develop an analytical technology process based on FTIR spectroscopy, coupled with multivariate analysis, for at-line monitoring of GOS production from whey.

## 2. Material and methods

### 2.1. Reagents

Cheese whey was purchased from CIMPA SAS (Bogotá, Colombia) with lactose content of 71.12% and <1% of monosaccharides. Commercial β-galactosidase Enzeco fungal lactase from *A. oryzae* (Ao) was provided by Enzyme Development Corporation (New York, USA); HA-Lactase5200 from *K. lactis* (Kl) was provided by CHR Hansen (Bogotá, Colombia). Enzymes had a total protein content, determined by Bradford method (Bradford, 1976), of 38.9 ± 0.1% (Ao) and 4.8 ± 0.1% (Kl). Enzyme activities were measured as reported by Rico-Rodríguez et al. (2018) giving enzyme activity of 15,045 U/g (Ao) and 5172 U/g (Kl). Bovine serum albumin (BSA), sodium dihydrogen phosphate monohydrate, disodium hydrogen phosphate dihydrate, and sodium hydroxide were purchased from Sigma–Aldrich (St. Louis, MO). Glucose (Glu), galactose (Gal), lactose (Lac) and raffinose (DP3), were purchased from Panreac® (Barcelona, Spain).

**Table 1 – Reaction arrays and dose for each assay.**

| Assay | Alias | Hal dose (U) | Enz dose (U) |
|---|---|---|---|
| Enz:Hal (100:0) | WO | 37.8 | 0 |
| Enz:Hal (75:25) | W1/4 | 28.4 | 0.9 |
| Enz:Hal (50:50) | W1/2 | 18.9 | 1.9 |
| Enz:Hal (25:75) | W3/4 | 9.9 | 2.8 |
| Enz:Hal (0:100) | WK | 0 | 3.8 |

### 2.2. Enzyme reactions

All assays were carried out in 50 mL falcon tubes with an effective volume of 20 mL (370 mg $mL^{-1}$ of lactose equivalent) in sodium phosphate buffer solution 0.01 M at pH 6.0 and temperature of 42 ± 1 °C. Enzyme doses for each assay are reported in Table 1. Sampling was done by duplicate after 0, 20, 40, 60, 120, 180, 240, 300 min of reaction under continuous stirring at 800 rpm. Samples were immersed in boiling water (5 min) for enzyme inactivation; afterwards, samples were cooled and stored at −20 °C until carbohydrate quantification.

### 2.3. Carbohydrate quantification

Substrates and products of enzyme reactions were analysed in an Ultimate 3000 HPLC instrument (Themo Fisher Scientific, USA). For GOS3, GOS4, GOS5, lactose and monosaccharides a BP-100 $Ca^{++}$ (300 mm × 7.8 mm – Benson polymerics, USA) was set at 80 °C and 0.5 mL/min of deionised water flowing rate. GOS2 were separated from lactose with a Spherisorb S5-$NH_2$ (250 mm × 4.0 mm – Waters, USA) with acetonitrile:water (75:25) as eluent at 30 °C and 0.75 mL/min flowing rate. Both measurements were done in a Transgenomic® RI detector. Calibration curves were performed with appropriate sugar standards in the linear range 0.05 mg $mL^{-1}$ to 5 mg $mL^{-1}$.

### 2.4. FTIR measurements

IR spectra, from different enzyme mixture reaction samples, were collected in transmission mode, with a FT-IR spectrometer (IR-Prestige21, Shimatsu Corporation, Japan) equipped with an ATR cell. The spectra information was collected in the range 4000–400 $cm^{-1}$, with a resolution of 4 $cm^{-1}$ and 25 scans for both background and samples. For each duplicate of the enzymatic rection, the same samples were analysed twice; the total number of collected spectra was 160. Besides, samples of carbohydrate standards (400 mg $mL^{-1}$) were analysed following the same procedure.

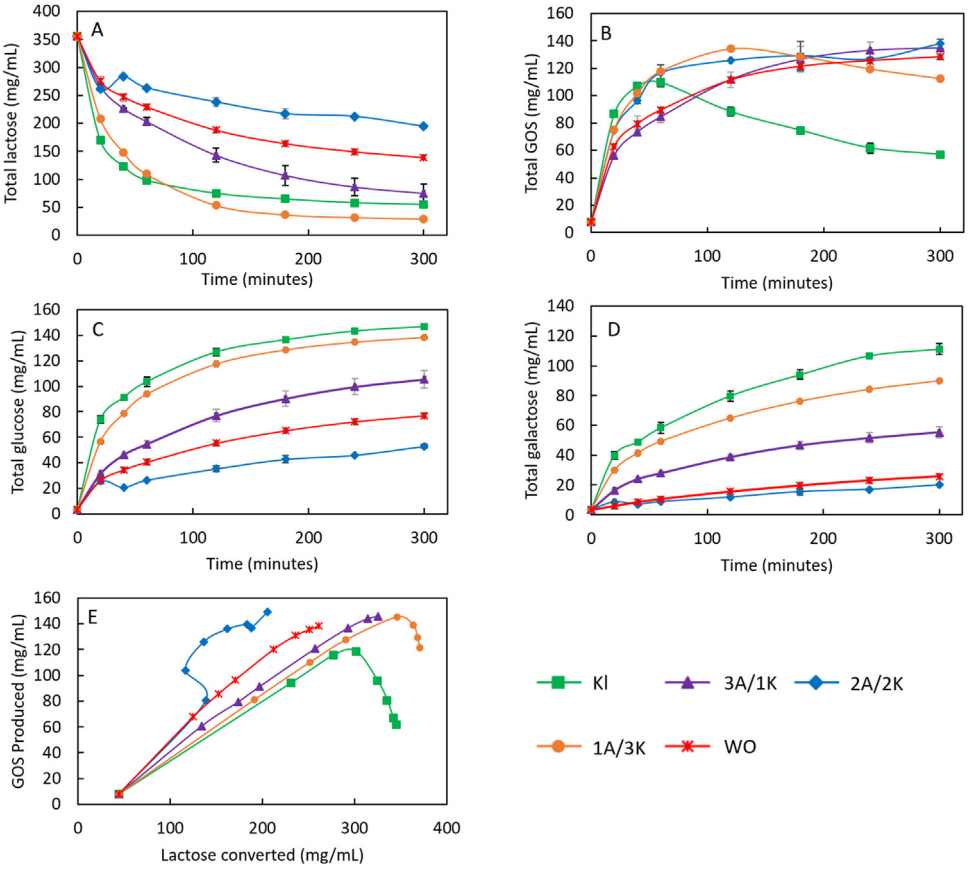


**Fig. 1 – Lactose transgalactosylation in presence of β-galactosidase mixture: (A) lactose consumption; (B) GOS production; (C) glucose production; (D) galactose production; (E) lactose conversion to GOS. Error bars represent the standard deviations between.**

### 2.5. *Data processing*

The whole wavenumber range (4000–400 $cm^{-1}$) was reduced to the fingerprint region (1200–900 $cm^{-1}$) and spectra were averaged on sample base. Before applying multivariate analysis methods, several pre-treatments were performed in order to assess which one could efficiently correct scattering effects and reduce noise: Standard Normal Variate (SNV), first and second derivatives, baseline correction (automatic weighted least squares) and smoothing (Savitzky–Golay, filter width 7 points; polynomial order 1).

To predict the amounts of lactose, GOS, glucose and galactose in whey samples, PLS regression was performed by using the PLS Toolbox software (Eigenvector Research, Inc., Wenatchee, Washington, USA) implemented in Matlab (MathWorks, Natick, MA, USA). For each evaluated parameter, data were partitioned in a calibration (including 64 averaged spectra) and a validation set (including 16 averaged spectra). The validation sets were constructed on reaction bases. This means that one out of five enzymatic reactions (16 averaged spectra) was kept out from the calibration set in an iterative way, thus developing five different models, each of them tested by one different reaction test set. Models were also internally validated by Venetian Blind Cross-Validation, with 10 splits and 2 samples per split, being sure to keep duplicates together. To evaluate model reliability, Root Mean Square Error of Prediction (RMSEP) was compared with the Root Mean Square Error of Laboratory (RMSEL). Furthermore, residual prediction deviation (RPD), i.e. ratio of standard error of performance to standard deviation, was calculated to compare the precision of the prediction with the average composition of all the samples, as stated by Camacho-Tamayo et al. (2014).

MCR-ALS method was applied on spectra using MCR-ALS 2.0 Toolbox (Jaumot et al., 2015) implemented in Matlab (MathWorks, Natick, MA, USA). Spectral data were arranged in five matrices, one for each enzymatic reaction considering together the duplicates, and analysed separately. For each model the number of components was selected by Pure Variable detection method. Model optimization was achieved by iterative implementation of ALS algorithm, i.e. until the Lack of Fit (LOF) difference in two consecutive iterative cycles was lower than 0.1%. A systematic investigation on any constraint in spectral modes of data matrix has be performed via calculation of feasible solutions and interpretation of obtained results. Then, non-negativity constraint was imposed on both concentration and spectral profiles to solve MCR-ALS ambiguities. For further details about the MCR-ALS procedure and optimization the reader can refer to Grassi et al. (2019).

## 3. Results and discussion

### 3.1. *Kinetics of GOS production by HPLC*

Transgalactosylation reactions by βgal from two sources (*K. lactis* and *A. oryzae*) and their mixtures followed the characteristic lactose consumption (Fig. 1). Lactose conversion for βgal from the yeast *Kluyveromyces lactis* results in values over 90% with a GOS production above 40% – mainly disaccharides and trisaccharides – and few tetrasaccharides (<20% of the total GOS). However, combination with βgal from *A. oryzae* could lead to better profiles of oligomers present in the final mix-

ture. In Fig. 1A it is observed that lactose consumption is highly affected by βgal mixture, i.e. by the enzyme source. When *K. lactis* enzyme is predominant in the mixture, Kl and 1A/3K trials, lactose concentration dropped faster than in other cases to values below 50 mg $mL^{-1}$. Nonetheless, kinetic behaviour of sample 2A/2K, which had equal dose of both enzymes, showed an unusual trend. Lactose consumed (50% of initial lactose) by 2A/2K reaction was lower than Ao (63% of initial lactose), which corresponded to *A. oryzae* pure enzyme.

In this set of reactions, monosaccharides (Fig. 1C and D) had one of the most significant variations. Though in all reactions there was a typical increase in the concentration of monosaccharides (glucose and galactose), there was also, in each treatment, a significant variation in their amount. *K. lactis* enzyme is known by its poor ability to include glucose in GOS molecule, hence, the concentration of this monosaccharide is above 40% of total carbohydrates present in media (Jenab et al., 2017; Rico-Rodríguez et al., 2018). Moreover, galactose release depends on the ability of the enzyme to transform lactose into GOS at defined conditions of reaction as the later are mainly composed of this monosaccharide. In this case, it is possible to observe that a gradually decreasing in final concentration of those monosaccharides occurs with the reduction in the concentration of initial *K. lactis* enzyme or with the increase of *A. oryzae* enzyme source.

Fig. 1B shows GOS kinetic concentration for each reaction. Samples Ao, 2A/2K and 3A/1K show a similar kinetics with GOS concentration increasing over the 5 h of reaction. A typical trend is shown by Kl; in this reaction, GOS hydrolysis started after 1 h. For 1A/3K the reaction exhibits similar trend, nonetheless, maximum concentration was found at 120 min and GOS hydrolysis is not as fast as Kl. Highest concentration of GOS was 36% of total carbohydrates and it was reached after 60 min for 1A/3K and after 300 min for 2A/2K. In Fig. 1B it is notable a clear differentiation in the kinetics of the pure enzyme *K. lactis*, if compared to all enzyme combinations.

On the other hand, a graph of lactose conversion to GOS is presented in Fig. 1E. This graph shows the ability of βgal to transform lactose to GOS in reaction media. Although in Fig. 1E the conversion of lactose is well differentiated for pure enzymes and their mixtures, 2A/2K has a bulging trend, an unexpected behaviour possibly linked to the interaction between the enzymes.

Results suggest that βgal from *K. lactis* consumes high amounts of lactose (73%) to reach 30% of GOS in media with initial lactose of 370 mg $mL^{-1}$. However, the amount of required lactose changes gradually when enzyme doses change in different reactions. Reaction Ao, in which only *A. oryzae* enzyme is present, required the lowest amount of lactose to produce the maximum concentration of GOS when initial lactose was 370 mg $mL^{-1}$. Fischer and Fischer and Kleinschmidt (2018) performed a similar assay, however, they evaluated time of enzyme addition instead of enzyme concentration. The yield they found was significantly lower than those obtained in the present work, even for pure enzymes. Another difference, very important for this kind of reactions, is the source of lactose employed in both works. While Fischer and Fischer and Kleinschmidt (2018) evaluated pure lactose as reaction media, our work assayed cheese whey, which as reported by same authors, might interfere significantly with enzyme transgalactosylation ability (Fischer and Kleinschmidt, 2015). Same results were found in a previous work from our group (Rico-Rodríguez et al., 2020).

### 3.2. *FTIR spectroscopy*

FTIR data obtained from enzymatic reactions followed a similar trend. As the composition of total carbohydrates in the media changed with time, so FTIR spectra changed. In Fig. 2 it is possible to observe data collected for treatments Kl (Fig. 2A), 1A/3K (Fig. 2B) and 3A/1K (Fig. 2C) as well as spectra of standards of lactose, glucose, galactose and GOS (370 mg $mL^{-1}$) (Fig. 2D).

Changes in spectra profiles of GOS reactions are more evident in the first minutes of the reactions, especially for Kl and 3A/2K reactions. These changes are linked to the fast drop of lactose concentration after 40 min of reaction and to the rise in glucose, galactose and GOS signals.

Indeed, the FTIR spectra of cheese whey (beginning of the reactions) are characterised by a broad absorption band from 1200 to 964 $cm^{-1}$, with two maxima at 1074 and 1034 $cm^{-1}$ and a shoulder at 995 $cm^{-1}$. Those signals are mainly related to lactose presence (the main carbohydrate present in cheese whey – >98% of total carbohydrates). Pure lactose (Fig. 2D) showed a profile characterised by a band with a maximum at 1157 $cm^{-1}$, followed by a broad band between 1110 and 970 $cm^{-1}$ with two maxima with similar absorbance at 1074 $cm^{-1}$ and 1034 $cm^{-1}$ and a shoulder at 995 $cm^{-1}$.

Spectra collected along reactions showed slightly different profiles from those collected at the beginning. Particularly, with the progress of time, the peak at 1100 $cm^{-1}$ disappears, the difference between the absorbance at 1074 and 1034 $cm^{-1}$ increases and a small signal becomes visible at 1050 $cm^{-1}$. Those changes can be attributed to the consumption of lactose and the increase of glucose, galactose and GOS. Indeed, pure glucose and galactose FTIR spectra (Fig. 2D) showed a spectral profile slightly different from the one of lactose. Even if the maxima are similar, glucose FTIR profile differs from the other pure components due to the presence of a shoulder at 1110 $cm^{-1}$ and a high difference between the maxima at 1074 and 1034 $cm^{-1}$; whereas the spectra collected for pure galactose is characterised by a higher absorbance at 1148 $cm^{-1}$ and the absence of the shoulder at 995 $cm^{-1}$. For pure GOS spectra (Fig. 2D) the band between 1110 and 970 $cm^{-1}$ does not show double maxima but a higher absorbance at 1050 $cm^{-1}$. Moreover, the shoulder observed at 995 $cm^{-1}$ for lactose and glucose is a distinct peak followed by a peak at 927 $cm^{-1}$.

Even if the contribution of the pure compounds can be inferred by visual inspection, the complexity of FTIR data could be better managed by multivariate approaches such as PLS and MCR-ALS, with the aim of monitoring GOS formation.

### 3.3. *PLS regression*

The PLS models were calibrated using 64 averaged spectra and internally validated by Venetian Blind Cross-Validation. Furthermore, they were tested for prediction in an iterative way, i.e. by testing they prediction capability by a validation set consisting of 16 averaged spectra of one of the enzymatic reactions performed. In Table 2 are reported the best results obtained applying PLS regression on pre-treated spectral data, they were obtained when the validation set was constructed by Kl, 3A/1K and 1A/3K data. Less promising results (data not shown) were obtained when Ao and 2A/2K enzymes were used for validation. These enzymes showed an unexpected behaviour, more specifically the reactions presented a lactose consumption and its conversion to GOS significantly different from the other three enzymes (Fig. 1A and C).

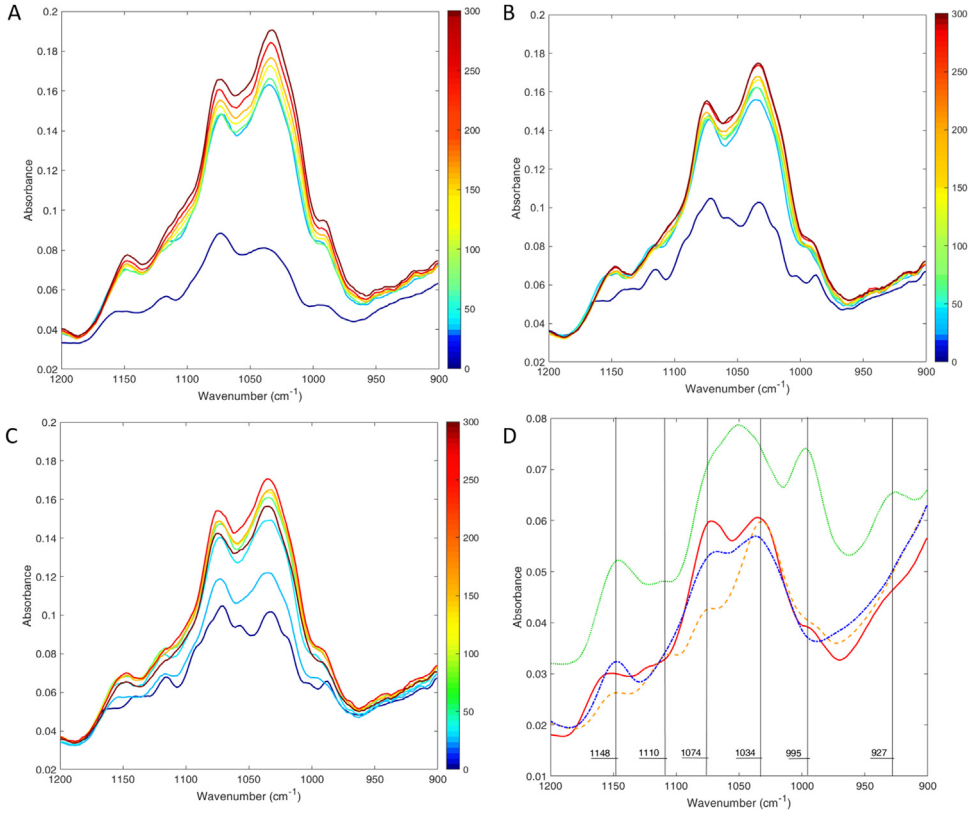


**Fig. 2 – FTIR spectra in the region 1200–900 cm$^{-1}$ collected from different enzymatic reactions (A, Kl; B, 3A/1K; C, 1A/3K) and pure components spectra (D). In A, B and C spectra are coloured according to reaction time (min). In D spectra correspond to lactose (red, -), glucose (orange, —), galactose (blue, -.-.) and GOS (green, ▼▼▼). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)**

**Table 2 – PLS regression models for lactose, GOS, glucose and galactose content prediction from FTIR spectra.**

| | Test set | LV | Pre-treatment | RMSEL (mg mL$^{-1}$) | RMSEC (mg mL$^{-1}$) | RMSECV (mg mL$^{-1}$) | RMSEP (mg mL$^{-1}$) | $R^2_{Cal}$ | $R^2_{CV}$ | $R^2_P$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Kl** | **6** | **Baseline correction** | | **14.9** | **19.2** | **21.9** | **0.98** | **0.97** | **0.98** |
| Lactose | 3A/1K | 3 | 1st derivative | 8.2 | 24.7 | 28.3 | 28.4 | 0.94 | 0.92 | 0.98 |
| | 1A/3K | 5 | Baseline correction | | 16.6 | 20.3 | 22.3 | 0.98 | 0.97 | 0.96 |
| | Kl | 7 | SNV | | 9.4 | 17.9 | 18.7 | 0.95 | 0.84 | 0.76 |
| GOS | **3A/1K** | **7** | **Baseline correction** | 3.2 | **10.2** | **14.3** | **16.0** | **0.94** | **0.89** | **0.88** |
| | 1A/3K | 7 | 1st Derivative | | 11.7 | 17.3 | 19.5 | 0.92 | 0.84 | 0.92 |
| | Kl | 2 | 1st derivative | | 12.6 | 13.7 | 14.5 | 0.91 | 0.89 | 0.96 |
| Glucose | **3A/1K** | **6** | **Raw** | 8.0 | **8.1** | **10.6** | **11.1** | **0.97** | **0.94** | **0.99** |
| | 1A/3K | 5 | Baseline correction | | 8.7 | 10.3 | 12.1 | 0.97 | 0.96 | 0.93 |
| | Kl | 2 | 1st derivative | | 11.9 | 12.9 | 23.6 | 0.8 | 0.76 | 0.95 |
| Galactose | **3A/1K** | **6** | **SNV** | 8.1 | **9.6** | **11.5** | **12.4** | **0.91** | **0.88** | **0.97** |
| | 1A/3K | 5 | Baseline correction | | 10.1 | 12.5 | 12.6 | 0.92 | 0.88 | 0.75 |

LV, latent variables; SNV, Standard Normal Variate; RMSEL, Root Mean Square Error of Laboratory; RMSEC, Root Mean Square Error of Calibration; RMSECV, Root Mean Square Error of Cross-Validation; RMSEP, Root Mean Square Error of Prediction; $R^2$, coefficient of determination in calibration (Cal), cross-validation (CV) and prediction (P).

Considering reactions with Kl, 3A/1K and 1A/3K enzymes PLS results were good both in terms of RMSEP and $R^2_P$. In particular, lactose models (range 31.50–384.44 mg mL$^{-1}$) presented very high $R^2$ values in prediction ($R^2_P > 0.96$) and a RMSEP value approximately lower than 3 times the RMSEL, an acceptable error for a NIR application (Shenk and Westerhaus, 1996). Similarly, Cocciardi et al. (2004) predicted lactose during hydrolysis reaction in milk. They obtained a reliable PLS models (SEP of 0.20% w/v) considering a lactose variability between 0 and 5% w/v. Our results, converted into percentage, are comparable being our lowest RMSEP for lactose prediction 6.2% (i.e. 21.9 mg/mL), that is comparable in magnitude with Cocciardi et al. (2004) SEP expressed in percentage (i.e. 4%).

Furthermore, in Table 2 are reported the models to predict in a non-destructive, fast and at-line approach the amount of GOS (range 8.06–151.62 mg mL$^{-1}$). The model validated with 3A/1K spectra, pre-processed with the baseline correction, using 7 LV gave a RMSEP of 16.0 mg mL$^{-1}$ and a $R^2_P$ of 0.88. With the same validation set, good prediction ability was obtained for both glucose (range 3.90–160.25 mg mL$^{-1}$) and galactose (range 3.60–122.97 mg mL$^{-1}$). In these cases, RMSEP values are extremely close to RMSEL values, indicat-

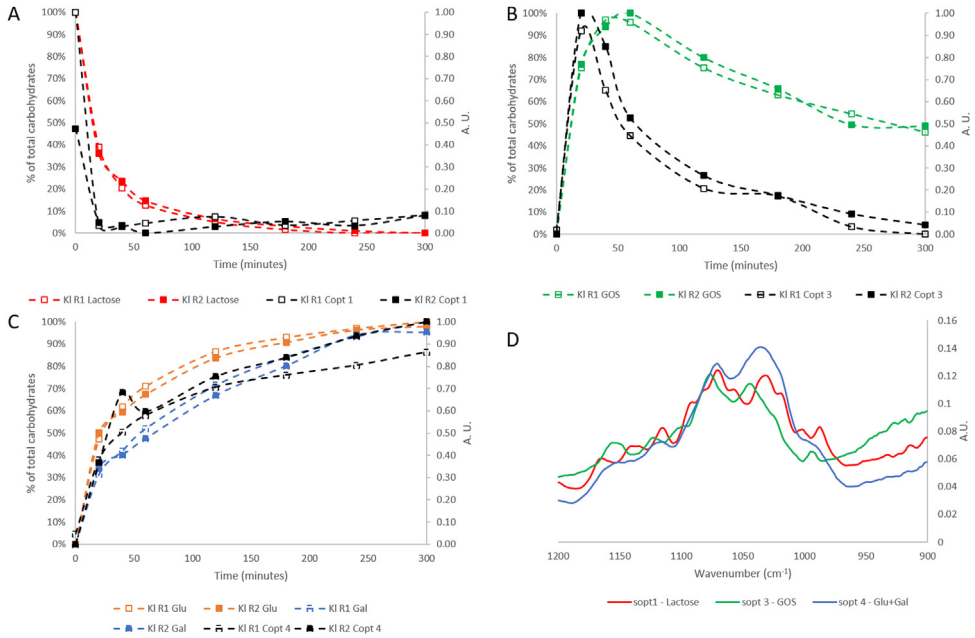


**Fig. 3 – MCR-ALS results for the Kl enzymatic reaction: (A) concentration profiles of lactose and lactose concentration as determined by HPLC; (B) concentration profiles of GOS and GOS concentration as determined by HPLC; (C) concentration profiles of glucose-galactose and glucose and galactose concentrations as determined by HPLC; (D) spectral profiles of lactose, GOS and glucose-galactose.**

ing an excellent estimation of these parameters (Shenk and Westerhaus, 1996).

Referring to the work by Cocciardi et al. (2004), minimum SEP values of 0.13% w/v and 0.075% w/v for glucose and galactose prediction were reached, respectively. However, the authors performed a leave-one-out cross-validation within each enzymatic reaction, thus considering a small inter-enzyme variability and a small range of sugar content variability (0–2.5% w/v). Another relevant work in the field, by Romano et al. (2016), focused on the development of PLS models to predict the effect of sucrose concentration on fructo-oligosaccharides (FOS). They reached optimal results when predicting glucose production by a PLS model with a $R^2_P$ of 0.98 and an RMSEP of 2.028 g of glucose/100 g of sucrose. Being their variability in a range of 0–56 g of glucose/100 g of sucrose, the %RMSEP will correspond to 3.6, which is lower than our (i.e. %RMSEP = 7.1).

Furthermore, our model performances were evaluated in terms of precision of the prediction with the average composition of all the samples, by means of residual prediction deviation (RPD). GOS and galactose models have an RPD of 2.8 and 2.7 respectively, suggesting a good predictive performance of the models, whereas lactose and glucose presented an RPD > 3 (5.9 and 4.9, respectively), indicating an excellent predictive performance (Saeys et al., 2005).

### 3.4. *MCR-ALS models*

The promising results obtained by PLS regression, suggested the possibility of building MCR-ALS models to follow the enzymatic reaction in vision of a Process Analytical Technology able to predict the process kinetics. Thus, an MCR-ALS model for the duplicates of each enzymatic combination has been developed.

The application of ALS procedure to the FTIR spectra collected allowed the resolution of both spectral and concentration profiles by four-component models, retaining at least 99.9% of the total variance, and a LOF lower than 0.44%.

Three out of four MCR-ALS profiles were assigned to lactose, GOS and a combination of Glu and Gal; whereas the fourth profile was assumed to be an interference used to isolate possible noise in the process as reported by Ahmadi et al. (2015). This is in agreement with previous works, reporting that the quantification of monosaccharides, Glu and Gal, as one group (Palai et al., 2012) or as a ratio (González-Delgado et al., 2016) can ease the analysis of their effect on the kinetic model for GOS prediction. According to these works, it is reasonable that one of the MCR resulting profiles is associated to the sum of the monosaccharides.

The profile identity was confirmed by comparing the spectral features of the pure compounds (Fig. 2D) and their kinetics measured by HPLC (Fig. 1). Fig. 3 reports the comparison between MCR-ALS components and the measured trends for Kl. Lactose MCR-ALS spectral profile (Fig. 3D) showed the characteristics observed for pure lactose FTIR spectra (Fig. 2D) with a broad band between 1110 and 970 $cm^{-1}$ with two maxima at similar absorbances (1074 $cm^{-1}$ and 1034 $cm^{-1}$) and a shoulder at 995 $cm^{-1}$. The related MCR-ALS concentration profile described a fast decrease in lactose within the 40 min followed by a plateau until the end of the process monitoring. This behaviour is similar to the lactose content quantified by HPLC analysis, even though by the chromatography determination lactose reached the plateau less fast, being the concentration close to 0% after 120 min (Fig. 3A). The MCR-ALS profile attributed to GOS showed a peak at 1148 $cm^{-1}$, a broad band between 1110 and 970 $cm^{-1}$, a distinct peak at 950 $cm^{-1}$ and a small feature around 927 $cm^{-1}$ (Fig. 3D). All these characteristics were observed for GOS pure spectra reported in Fig. 2D. Even if there was a spectral correspondence, the GOS concentration profile failed in describing the GOS behaviour as observed by HPLC analysis. Indeed, even if the GOS concentration profile described a fast increase in relative concentration

within the first 20 min, its drop has been much higher than the one recorded by HPLC (Fig. 3B).

Different reasons could explain the failure in kinetic prediction. First of all, it should be considered that in our approach GOS were evaluated as a unique product from the transgalactosylation reaction. However, the biochemical reactions for GOS production are rather a complex mechanism in which lactose is converted to one of a wide group of oligosaccharides derived from lactose (Gosling et al., 2010). Indeed, GOS are a carbohydrate-based molecule group, thus, they can be considered as a class, but the different proportion of di-, tri-, tetra-, penta- and hexa- saccharides might affect the kinetic modelling results. For this reason, trying to explain the reaction through a mathematical or a statistical model is not as simple as one can expect if the total amount of GOS is considered. Several attempts to model the kinetics have been assayed (Jenab et al., 2017; Palai et al., 2012; Warmerdam et al., 2014; Yin et al., 2017), nonetheless, uncertainness is always present.

Furthermore, the considered reaction media, i.e. cheese whey, is very complex. Indeed, it contains salts and whey proteins, as well as a mixture of sugars of different size and linkages; all of them leading to bands overlapping in the FTIR fingerprint region (1200–950 $cm^{-1}$). Even more, the total content of saccharides does not change over time, but saccharides are transformed by the transgalactosylation reaction; so, FTIR measurements could be less reliable in detecting changes in GOS production. Moreover, to improve the efficacy of this technique, it would be necessary to test standards of every possible carbohydrate formed during the reaction.

Even if the MCR-ALS models failed in assessing GOS production, model reliability was confirmed for monosaccharides released by the reaction, i.e. glucose and galactose. In fact, the spectral profile (Fig. 3D) exhibits features resulting from the combination of pure glucose and pure galactose FTIR spectra (Fig. 2D). In this case, high correspondence was found between the MCR-ALS concentration profile and the monosaccharides concentrations measured by the chromatographic method (Fig. 3C).

Similar results were obtained by MCR-ALS models developed for 1A/3K (Fig. 1S) and 3A/1K (Fig. 2S). Indeed, lactose decrease, as well as the glucose and galactose production, has been detected with a faster drop by the MCR-ALS models, especially for 1A/3K duplicates. However, the GOS production was not systematically modelled by MCR algorithm, especially for 1A/3K trials. Regardless GOS typical behaviour after reaching their maximum value is to follow a reduction in their concentration due to hydrolytic enzyme property, this depletion was depicted faster than the observed HPLC data.

## 4. Conclusions

Different chemometric approaches have been tested to assess a Process Analytical Technology to follow GOS production from lactose naturally present in dairy industry waste by FTIR spectroscopy. PLS regression models demonstrated to be reliable to assess lactose, glucose and galactose content in a non-destructive, fast and at-line approach, which could be used for real time applications in the future. Indeed, their prediction by Kl or 3A/1K enzymatic reaction led to good fit ($R^2_P$ > 0.97) and acceptable errors (RMSEP of 21.9 mg $mL^{-1}$ for lactose, 11.1 mg $mL^{-1}$ for glucose and 12.4 mg $mL^{-1}$ for galactose). Unfortunately, GOS prediction by regression models did not reach the same performance. This could be linked to the development of a model considering GOS as a unique class, whereas it is quite heterogeneous accounting for di-, tri-, tetra-, penta- and hexa- saccharides.

Furthermore, an MCR-ALS model for each enzymatic combination has been developed focusing more on the process kinetics rather than the prediction of each component. Three MCR-ALS profiles were assigned to lactose, GOS and a combination of glucose and galactose. The corresponding MCR-ALS concentration profiles gave trends similar to the sugar concentrations measured by the chromatographic method, however they demonstrated not to be enough accurate to be implemented. Thus, this work should be considered as a preliminary analysis to demonstrate the potential of FTIR spectroscopy, combined with Chemometrics, to follow the kinetics of GOS production from lactose naturally present in dairy industry waste.

The future implementation of this approach to real time systems will answer the need of Circular Economy by providing a reliable monitoring of dairy waste reallocation, avoiding sample pre-processing, large volumes of organic solvents and long-time of analysis. In any case, further investigation is suggested to better assess specific GOS prediction by a single FTIR analysis.

## Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

F. Rico-Rodriguez thanked to Ministry of science, technology and innovation of Colombia – MINCIENCIAS by the grant for Colombian posdoctoral contract No. 811/2018.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fbp.2020.12.009.

## References

Akiyama, T., Kimura, K., Hatano, H., 2015. Diverse galactooligosaccharides consumption by bifidobacteria: implications of β-galactosidase—LacS operon. Biosci. Biotechnol. Biochem. 79, 664–672, http://dx.doi.org/10.1080/09168451.2014.987204.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254, http://dx.doi.org/10.1016/0003-2697(76)90527-3.

Byfield, J., Cardenas, S., Alméciga-Diaz, C.J., Sánchez, O.F., 2010. β-Galactosidase and galactooligosaccharides production and applications. Recent Patents Chem. Eng. 3, 17–29.

Chockchaisawasdee, S., Athanasopoulos, V.I., Niranjan, K., Rastall, R.A., 2004. Synthesis of galacto-oligosaccharide from lactose using beta-galactosidase from *Kluyveromyces lactis*: studies on batch and continuous UF membrane-fitted bioreactors. Biotechnol. Bioeng. 89, 9, http://dx.doi.org/10.1002/bit.20357.

Cocciardi, R.A., Ismail, A.A., Van De Voort, F.R., Sedman, J., 2004. Monitoring of lactose hydrolysis in milk by single-bounce attenuated total reflectance Fourier transform infrared spectroscopy. Milchwissenschaft 59 (7–8), 403–407.

Eskandarloo, H., Abbaspourrad, A., 2018. Production of galacto-oligosaccharides from whey permeate using β-galactosidase immobilized on functionalized glass beads. Food Chem. 251, 115–124, http://dx.doi.org/10.1016/j.foodchem.2018.01.068.

Fischer, C., Kleinschmidt, T., 2015. Synthesis of galactooligosaccharides using sweet and acid whey as a substrate. Int. Dairy J. 48, 15–22, http://dx.doi.org/10.1016/j.idairyj.2015.01.003.

Fischer, C., Kleinschmidt, T., 2018. Combination of two β-galactosidases during the synthesis of galactooligosaccharides may enhance yield and structural diversity. Biochem. Biophys. Res. Commun. 506 (1), 211–215, http://dx.doi.org/10.1111/1541-4337.12344.

Gennari, A., Mobayed, F.H., Catto, A.L., Benvenutti, E.V., Volpato, G., de Souza, C.F.V., 2019. Kluyveromyces lactis β-galactosidase immobilized on collagen: catalytic stability on batch and packed-bed reactor hydrolysis. React. Kinet. Mech. Catal. 127 (2), 583–599, http://dx.doi.org/10.1007/s11144-019-01598-6.

González-Delgado, I., López-Muñoz, M.-J., Morales, G., Segura, Y., 2016. Optimisation of the synthesis of high galacto-oligosaccharides (GOS) from lactose with β-galactosidase from *Kluyveromyces lactis*. Int. Dairy J. 61, 211–219, http://dx.doi.org/10.1016/j.idairyj.2016.06.007.

Gosling, A., Stevens, G.W., Barber, A.R., Kentish, S.E., Gras, S.L., 2010. Recent advances refining galactooligosaccharide production from lactose. Food Chem. 121, 307–318, http://dx.doi.org/10.1016/j.foodchem.2009.12.063.

Grassi, S., Amigo, J.M., Lyndgaard, C.B., Foschino, R., Casiraghi, E., 2014. Assessment of the sugars and ethanol development in beer fermentation with FT-IR and multivariate curve resolution models. Food Res. Int. 62, 602–608, http://dx.doi.org/10.1016/j.foodres.2014.03.058.

Grassi, S., Strani, L., Casiraghi, E., Alamprese, C., 2019. Control and monitoring of milk renneting using FT-NIR spectroscopy as a process analytical technology tool. Foods 8, 405, http://dx.doi.org/10.3390/foods8090405.

Hernández-Hernández, O., Calvillo, I., Lebrón-Aguilar, R., Moreno, F.J., Sanz, M.L., 2012. Hydrophilic interaction liquid chromatography coupled to mass spectrometry for the characterization of prebiotic galactooligosaccharides. J. Chromatogr. A 1220, 57–67, http://dx.doi.org/10.1016/j.chroma.2011.11.047.

Jaumot, J., de Juan, A., Tauler, R., 2015. MCR-ALS GUI 2.0: new features and applications. Chemometr. Intell. Lab. Syst. 140, 1–12, http://dx.doi.org/10.1016/j.chemolab.2014.10.003.

Jenab, E., Omidghane, M., Mussone, P., Hernandez Armada, D., Cartmell, J., Montemagno, C., 2017. Enzymatic conversion of lactose into galacto-oligosaccharides: the effect of process parameters, kinetics, foam architecture, and product characterization. J. Food Eng. 222, 63–72, http://dx.doi.org/10.1016/j.jfoodeng.2017.11.015.

Kumar, V., Sharma, D.K., Sandhu, P.P., Jadaun, J., Sangwan, R.S., Yadav, S.K., 2020. Sustainable process for the production of cellulose by an *Acetobacter pasteurianus* RSV-4 (MTCC 25117) on whey medium. Cellulose, 1–14, http://dx.doi.org/10.1007/s10570-020-03519-6.

Li, H., Cao, Y., Li, S., Jiang, Y., Chen, J., Wu, Z., 2019. Optimization of a dual-functional biocatalytic system for continuous hydrolysis of lactose in milk. J. Biosci. Bioeng. 127, 38–44, http://dx.doi.org/10.1016/j.jbiosc.2018.07.009.

Misson, M., Du, X., Jin, B., Zhang, H., 2016. Dendrimer-like nanoparticles based β-galactosidase assembly for enhancing its selectivity toward transgalactosylation. Enzyme Microb. Technol. 84, 68–77, http://dx.doi.org/10.1016/j.enzmictec.2015.12.008.

Mohamed, H., Nagy, P., Agbaba, J., Kamal-Eldin, A., 2020. Use of near and mid infra-red spectroscopy for analysis of protein, fat, lactose and total solids in raw cow and camel milk. Food Chem. 334, 127436, http://dx.doi.org/10.1016/j.foodchem.2020.127436.

Moros, J., Garrigues, S., de la Guardia, M., 2010. Vibrational spectroscopy provides a green tool for multi-component analysis. Trend. Anal. Chem. 29 (7), 578–591, http://dx.doi.org/10.1016/j.trac.2009.12.012.

Otieno, D.O., 2010. Synthesis of β-galactooligosaccharides from lactose using microbial β-galactosidases. Compr. Rev. Food Sci. Food Saf. 9, 471–482, http://dx.doi.org/10.1111/j.1541-4337.2010.00121.x.

Palai, T., Mitra, S., Bhattacharya, P.K., 2012. Kinetics and design relation for enzymatic conversion of lactose into galacto-oligosaccharides using commercial grade β-galactosidase. J. Biosci. Bioeng. 114, 418–423, http://dx.doi.org/10.1016/j.jbiosc.2012.05.012.

Rico-Rodríguez, F., Serrato, J.C.J.C., Montilla, A., Villamiel, M., 2018. Impact of ultrasound on galactooligosaccharides and gluconic acid production throughout a multienzymatic system. Ultrason. Sonochem. 44, 177–183, http://dx.doi.org/10.1016/j.ultsonch.2018.02.022.

Rico-Rodríguez, F., Villamiel, M., Ruiz-Aceituno, L., Serrato, J.C., Montilla, A., 2020. Effect of the lactose source on the ultrasound-assisted enzymatic production of galactooligosaccharides and gluconic acid. Ultrason. Sonochem. 67, 104945, http://dx.doi.org/10.1016/j.ultsonch.2019.104945.

Rodrigues Mano, M.C., Paulino, B.N., Pastore, G.M., 2019. Whey permeate as the raw material in galacto-oligosaccharide synthesis using commercial enzymes. Food Res. Int. 124, 78–85, http://dx.doi.org/10.1016/j.foodres.2018.09.019.

Romano, N., Santos, M., Mobili, P., Vega, R., Gómez-Zavaglia, A., 2016. Effect of sucrose concentration on the composition of enzymatically synthesized short-chain fructo-oligosaccharides as determined by FTIR and multivariate analysis. Food Chem. 202, 467–475, http://dx.doi.org/10.1016/j.foodchem.2016.02.002.

Ruiz-Matute, A.I., Corzo-Martínez, M., Montilla, A., Olano, A., Copovi, P., Corzo, N., 2012. Presence of mono-, di- and galactooligosaccharides in commercial lactose-free UHT dairy products. J. Food Compos. Anal. 28, 164–169, http://dx.doi.org/10.1016/j.jfca.2012.06.003.

Sangwan, V., Tomar, S.K., Singh, R.R.B., Singh, A.K., Ali, B., 2011. Galactooligosaccharides: novel components of designer foods. J. Food Sci. 76, R103–R111, http://dx.doi.org/10.1111/j.1750-3841.2011.02131.x.

Schiele, S.A., Meinhardt, R., Eder, C., Briesen, H., 2020. ATR-FTIR spectroscopy for in-line anomer concentration measurements in solution: a case study of lactose. Food Control 110, 107024, http://dx.doi.org/10.1016/j.foodcont.2019.107024.

Warmerdam, A., Zisopoulos, F.K., Boom, R.M., Janssen, A.E.M., 2014. Kinetic characterization of galacto-oligosaccharide (GOS) synthesis by three commercially important β-galactosidases. Biotechnol. Prog. 30, 38–47, http://dx.doi.org/10.1002/btpr.1828.

Yin, H., Bultema, J.B., Dijkhuizen, L., van Leeuwen, S.S., 2017. Reaction kinetics and galactooligosaccharide product profiles of the β-galactosidases from *Bacillus circulans*, *Kluyveromyces lactis* and *Aspergillus oryzae*. Food Chem. 225, 230–238, http://dx.doi.org/10.1016/j.foodchem.2017.01.030.