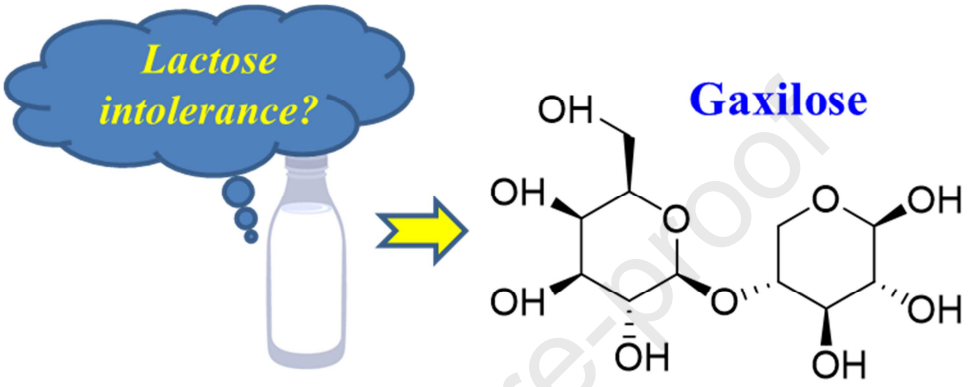


Characterization of crystalline forms of gaxilose, a diagnostic drug

Federica Catti[a,†] , Santos Hernández Gallego[a], Mónica Benito[b], Elies Molins[b,*], Francisco Marquillas Olóndriz[b,*]

[a] Interquim S.A., R&D Department, C/Joan Buscallà, 10, 08173 Sant Cugat del Vallès, Barcelona, Spain.

[b] Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus UAB, 08193 Bellaterra, Spain

Abstract

Lactose intolerance is a pathology caused by lactase enzyme deficiency, usually produced in the intestinal cells provoking symptoms as abdominal pain, bloating, diarrhoea, gas and nausea. Gaxilose, 4-*O*-β-D galactopyranosyl-D-xylose, is used as a diagnostic drug for a non-invasive method for hypolactasia diagnosis. To date, no definitive guide for identifying gaxilose and distinguishing between crystalline forms is available. Data have been collected from a number of different analytical techniques in order to provide a full characterization of the compound and a simple method to discriminate between two solid forms.

Keywords: Diagnostic drug, lactose intolerance, crystal structure, enzyme, NMR spectroscopy

[*]Corresponding authors

Tel.: +034 935 801 853, *E-mail address:* elies.molins@icmab.es (E. Molins),

Tel.: +034 935 044 200, *E-mail address:* fmarquillas@ferrer.com (F. Marquillas Olóndriz)

[†]Present address: Arkansas State University, Campus Queretaro, Carretera Estatal 100, km 17.5, Municipio Colón, 76270, Querétaro, Mexico

1. Introduction

Intestinal lactase is a glycoprotein [1] localized at the enterocyte's brush-border surface membrane and it is essential for the hydrolysis of the lactose component of milk in young mammals. Lactose is not absorbed as it, but it is essential firstly to be submitted to a hydrolysis into its D-Galactose and D-Glucose components and therefore those can be absorbed. [2] A loss of intestinal lactase could produce an uncompleted digestion or bad absorption and therefore lactose intolerance. [3,4] In fact, without this enzyme the lactose cannot be hydrolyzed in the small intestine and it gives rise to delays in gastric emptying, increases in osmotic pressure eventually bringing bacterial fermentation in the colon, and also gas production either hydrogen, methane and/or carbon dioxide. [5-7] Hence, digestion process is reduced and monosaccharides absorption decreases producing symptoms as abdominal pain, bloating, diarrhea, gas and nausea.

This is really a widespread affection: it is estimated that around 65% of the world population is affected with lactose intolerance [8]. In the case of adults, the treatment typically consists on decreasing the amount of lactose in the diet, taking lactase supplements, and treating the underlying disease; in the case of newborn children, this congenital deficiency prevents the correct assimilation of lactose causing severe disorders, such as severe diarrhea and dehydration, resulting both from decreased energy intake and intestinal accumulation of the non-hydrolyzed disaccharide [9]. Lactase deficiency also occurs as a secondary deficiency related to a significant number of intestinal pathologies involving deterioration in various degrees of intestinal mucosa, such as celiac disease, chronic inflammatory bowel disease (Crohn's disease and ulcerative colitis), intestinal resections, cystic fibrosis, premature infants, administration of chemotherapy, or as an additional disorder to the elderly, among others. The evaluation of the lactase activity is therefore of particular interest in gastroenterology, pediatrics and, in general, pathological processes in which it is necessary to evaluate the functional integrity of the intestinal mucosa or to make the differential diagnosis with the deficiency of this enzyme [10].

Gaxilose, the international non-proprietary name (INN) of 4-*O*-β-D galactopyranosyl-D-xylose [11], is a disaccharide formed by a D-Galactose and a D-Xylose, joined together through a β (1→4) bond (Figure 1) [12]. When gaxilose is orally administered, it can act as substrate for the intestinal lactase enzyme. In fact, human intestine does not absorb gaxilose if it is not previously hydrolyzed by an intestinal lactase enzyme. Once hydrolyzed, the monosaccharides are absorbed, and D-Galactose is metabolized, whilst D-Xylose is eliminated through the urine. The amount of D-Xylose secreted is correlated with the intestinal lactase enzyme levels and it can be evaluated by a simple colorimetric method. It is therefore used as a diagnostic drug for a non-invasive method for hypolactasia diagnosis.

4-*O*-β-D galactopyranosyl-D-xylose was firstly described in 1957 [11] as a syrup resulting from a controlled depolymerization of corn hull hemicellulose by hydrolysis using 0.01N HCl. A gelatinous mass submitted to slow crystallization from ethanol/water affords rectangular plate crystals having a melting point (m.p.) of 210 - 211°C with a sintering at 110 - 120 °C and an optical rotation of $[\alpha]_D$ -1→ 15° (H_2O). Elemental analysis suggested the presence of two molecules of water.

In a subsequent publication [13], the disaccharide was prepared following a Koenigs-Knorr reaction through condensation of tetra-*O*-acetyl-α-D-galactopyranosyl bromide and benzyl 2,3-anhydro β-D-ribopyranoside, followed by deacetylation, alkaline opening of epoxide and catalytic hydrogenation [14]. The syrup collected after column chromatography was reported to be pure, but no crystals were isolated. It was characterized as an amorphous product showing a m.p. of 210-211 °C and $[\alpha]_D$ +15° (H_2O).

In 1978, the same group published [15] the use 4-*O*-β-D galactopyranosyl-D-xylose as starting material in the synthesis of oligosaccharides fragments in order to study protein chain linkage in heparin and other proteoglycans. In this paper, the authors followed the same process described in the previous article, but the yield was increased due to the use of silver triflate and 2,3,4,6-tetra-*O*-benzoyl-α-D-galactopyranosyl bromide. The product could not be crystallized and only optical rotation was given ($[\alpha]_D$ +18° (H_2O)).

A Slovenian patent [16] was then approved in 1989 reporting the synthesis of the product and its characterization as $[\alpha]_D$ -4° (H_2O) and elemental analysis for a molecular mass of 312.27.

The first paper reporting the use of 4-*O*-β-D galactopyranosyl-D-xylose as diagnostic in urine for adult type alactasia was published in 1992 [12,17,18]. The disaccharide was prepared from benzyl 2,3-*O*-isopropylidene-β-D-xylopyranoside by glycosylation with 2,3,4,6-tetra-*O*-benzoyl-α-D-galactopyranosyl bromide and subsequent deprotection. The isolation of a syrup and its purification by column chromatography gave a pure fraction as a white solid. However, the characterization was pretty confusing. Although elemental analysis and optical rotation corresponded to the previously reported in the literature, the lower melting point described (160-168 ºC) suggests that either a different polymorphic form might be isolated or a highly unpure product.

The same group published several other patents introducing new synthetic methods for the preparation of 4-*O*-β-D galactopyranosyl-D-xylose and its use in the hypolactasia detection [19,20]. In those patents, the synthesis of 4-*O*-β-D galactopyranosyl-D-xylose and its isomers is described by the reaction between a β-D-galactopyranoside substrate with D-xylose in presence of a β-galactosidase enzyme. It was stated that a mixture of three possible isomers was obtained. However, no details about characterization data were available.

Finally, in a more recent patent from 2006 [21], several examples were reported using optimized enzymatic conditions for the preparation of 4-*O*-β-D galactopyranosyl-D-xylose. It is described the use of recrystallization methods for isolating a pure product with a melting point of 171-176 °C.

Consequently, we have found a big dispersion of the data regarding this compound. And, for instance, different melting points could suggest that the molecule may exhibit crystalline polymorphism.

Crystalline polymorphism is the ability of a chemical compound to crystallize into different structures where the atoms, ions or molecules are packed differently within the crystal lattice [22,23]. Moreover, crystalline solids can incorporate solvents in the lattice during crystallization leading to the formation of solvates, hydrates when the solvent is

water. This is commonly called as pseudopolymorphism [24]. It is critical to study the solid forms landscape (anhydrous crystalline forms, solvates and amorphous forms) of the substance especially in the pharmaceutical industry because distinct solid forms can have different chemical and physical properties, such as solubility or melting point.

To date, no definitive contribution for identifying gaxilose and its polymorphism using modern analytical techniques is available. A similar situation has been recognized for other small disaccharides [25-31]. Recently, a review on the well-known compound Lactose, which has been extensively studied in the past [32,33], was published to clarify its solid form landscape [34]. Moreover, as for many other carbohydrates, the existence of anomers as a consequence of the different orientation of the hydrogen and hydroxyl group in the anomeric carbon (C1') of the second monosaccharide unit makes the situation more complex (Figure 1 for gaxilose α and β anomers). Mutarotation can take place in solution affording different anomeric compositions and thus several solid forms, as have been demonstrated by Altamimi *et al*. during the analysis of several commercial lactose samples [35].

As a contribution to clarify the solid state of gaxilose, this paper aims to outline the key data from a number of different analytical techniques that allows the unequivocal recognition of two solid forms of this compound, Forms A and B.

2. Experimental

2.1. Materials

All reagents and solvents were used as received from commercial suppliers without further purification.

2.2. Synthesis

Gaxilose was prepared following example 9 in EP1408118B1 (Scheme 1). In summary, 4.12 g of o-nitrophenyl β-D-galactopyranoside (Gal-ONP) and 20.6 g of D-xylose were dissolved in 66 V of water buffered at pH = 7 (0.05 M KH_2PO_4/K_2HPO_4, 1 mM $MgCl_2$, 5 mM mercaptoethanol), 66 units of *E.coli* β-galactosidase enzyme were added and the solution was heated at 37 °C until complete consumption. The reaction was stopped by cooling to 0 °C and the o-nitrophenol was filtered as solid. 60 g of active carbon were added to the filtrate and stirred for 30 min. The mixture was filtered and active carbon

washed with water and different isopropanol solutions. The fractions containing 1,4 β–gaxilose were concentrated and a crystallization from acetone-water was attempted. In our hands an impure solid was obtained.

Two further successive recrystallizations from isopropanol-water were carried out giving rise to a pure and crystalline solid, named gaxilose Form A. **4-*O*-β-D galactopyranosyl-D-xylose Form A**, β anomer m.p. 95.4°C, 208.7 °C, lit. 210-211 °C [11], 171-176 °C [21]; $[\alpha]_D^{25}$ +14.9° (H_2O) [11]; R_f = 0.14; ATR-FTIR ν_{max} (cm^{-1}): 3579, 3313, 2919, 2886, 1093, 1067; H-NMR (*DMSO-d_6*, 400 MHz): δ 6.60 (d,1H, J=6.4 Hz, $OH_{1'\,\beta}$), 4.95 (d, 1H, J=8Hz, $OH_{2'\,\beta}$), 4.92 (d, 1H, J=3.2 Hz, $OH_{3'\,\beta}$); 4.81 (d, 1H, J=4.4 Hz, $OH_{2\,\beta}$), 4.78 (d, 1H, J=8 Hz, $OH_{3\,\beta}$), 4.59 (t, 1H, J=4 Hz, $OH_{6\,\beta}$), 4.39 (d, 1H, J=4 Hz, $OH_{4\,\beta}$), 4.20 (dd, 1H, J=6.8, 8 Hz, $H_{1'\,\beta}$), 4.16 (d, 1H, J=8 Hz, $H_{1\,\beta}$), 3.75 (dd, 1H, J= 5.2, 11.2 Hz, $H_{5'\,\beta}$), 3.57 (t, 1H. J= 4.4 Hz, $H_{4\,\beta}$), 3.51-3.42 (m, 3H, $H_{4'\,\beta}$, $H_{6\,\beta}$), 3.36-3.33 (m, 1H, $H_{5\,\beta}$), 3.31-3.29 (m, 1H, $H_{2\,\beta}$), 3.25 (dd, 1H, J= 3.2, 5.2 Hz, $H_{3\,\beta}$), 3.18 (dt, 1H, J= 3.6, 8.8Hz, $H_{3'\,\beta}$), 3.06 (t, 1H, 13.2 Hz, $H_{5'\,\beta}$), 2.92-2.86 (m, 1H, $H_{2'\,\beta}$); ^{13}C-NMR (100.62 MHz, *DMSO-d_6*): δ 102.6 ($C_{1\beta}$), 97.9 ($C_{1'\,\beta}$), 76.7 ($C_{4'\,\beta}$), 75.9 ($C_{5\,\beta}$), 75.1 ($C_{3'\,\beta}$), 74.9 ($C_{2'\,\beta}$), 73.5 ($C_{3\,\beta}$), 70.2 ($C_{2\,\beta}$), 68.5 ($C_{4\,\beta}$), 63.8 ($C_{5'\,\beta}$), 60.9 ($C_{6\,\beta}$); $[M+35]^-$ m/z = 347.1, 349.1 and 348.1.

Gaxilose Form B was prepared by heating Form A above 125 °C or preparing a slurry in methanol for three days. A solid was obtained. **4-*O*-β-D galactopyranosyl-D-xylose Form B**, α and β anomers, m.p. 216 °C, lit. 210-211 °C [11], 171-176 °C [21]; ATR-FTIR ν_{max} (cm^{-1}): 3331, 2896, 1044, 1000; ^{1}H-NMR (400 MHz, *DMSO-d_6*, δ in ppm): 6.62 (d, 0.66H, J=6.4 Hz, $OH_{1'\,\beta}$), 6.24 (d,0.34H, J=8 Hz, $OH_{1'\alpha}$), 4.98 (d, 0.66H, J= 8 Hz, $OH_{2'\beta}$), 4.95 (d, 0.66H, J=8 Hz, $OH_{3'\beta}$), 4.88-4.82 (m, 0.68H, $OH_{2'\alpha}$, $OH_{3'\alpha}$); 4.80 (d, 1H, J=5.6 Hz, $OH_{2\alpha,}$ $OH_{2\beta}$), 4.73 (d, 1H, J=8 Hz, $OH_{3\alpha}$, $OH_{3\beta}$), 4.63-4.55 (m, 1H, $OH_{6\alpha}$, $OH_{6\beta}$), 4.42 (d, 1H, J=7.2 Hz, $OH_{4\alpha}$,$OH_{4\beta}$), 4.28-4.15 (m, 2H, $H_{1\alpha}$,$H_{1\beta}$, $H_{1'\alpha}$, $H_{1'\beta}$),3.82-3.80 (dd, 0.66H, J= 5.2, 11.2 Hz, $H_{5'\beta}$), 3.60 (t, 1H. J= 4.4 Hz, $H_{4\alpha,}$ $H_{4\beta}$), 3.56-3.42 (m, 3H, $H_{2'\alpha}$,$H_{3'\alpha}$,$H_{4'\alpha}$, $H_{4'\beta,.}$ $H_{5\alpha'}$, $H_{6\alpha}$, $H_{6\beta}$), 3.42-3.36 (m, 1H, $H_{5\alpha.}$ $H_{5\beta}$), 3.32-3.23 (m, 2H, $H_{2\alpha}$, $H_{2\beta}$, $H_{3\alpha}$, $H_{3\beta}$); 3.21 (dt, 0.66H, J= 3.6, 8.8 Hz, $H_{3'\beta}$), 3.1 (t, 0.66H, 13.2 Hz, $H_{5'\beta}$), 2.98-2.86 (m, 0.66H, $H_{2'\beta}$); ^{13}C-NMR (100.62 MHz, *DMSO-d_6*, δ in ppm): δ 102.3 ($C_{1\beta}$, $C_{1\,\alpha}$), 97.5 ($C_{1'\beta}$), 92.3 ($C_{1'\alpha}$), 76.7 ($C_{4'\alpha}$), 76.3 ($C_{4'\beta}$), 75.6 ($C_{5\alpha,}$ $C_{5\beta}$), 74.8

($C_{3'\beta}$), 74.6 ($C_{2'\beta}$), 73.1 ($C_{3\alpha,}$ $C_{3\beta}$), 72.2 ($C_{3'\alpha}$), 71.4 ($C_{2'\alpha}$), 69.7 ($C_{2\alpha}$, $C_{2\beta}$), 68.2 ($C_{4\alpha}$,$C_{4\beta}$), 63.5 ($C_{5'\beta}$). 60.5 ($C_{6\alpha,}$ $C_{6\beta}$), 59.1 ($C_{5'\alpha}$).

2.3. Physical Characterization

2.3.1 Gas Chromatography (GC-FID)

The GC-FID analysis was carried out using an Agilent 7890A gas chromatograph with a DB-1701 [30m x 0.25mm x 0.25µm] chromatographic column or equivalent. Helium (2.4 mL/min, P = 33 psi at 170 °C) was used as gas carrier. The GC-FID parameters used are: Injection Split mode, ratio (80:1); Injection temperature 220 °C; FID Temperature at 260 °C. Main peak retention time: around 54.2 minutes, Injection volume: 3 µL; Oven initial temperature: 170 °C (kept for 11.82 min) and oven final temperature: 210 °C (kept for 59.10 min); Rate temperature: 15.23 °C / min.

2.3.2 Nuclear Magnetic Resonance Spectroscopy (NMR)

The ^{1}H and ^{13}C NMR spectra were recorded on an Agilent VNMRS-400 spectrophotometer using deuterated dimethyl sulfoxide (DMSO-d_6) (^{1}H at 400.10 MHz and ^{13}C at 100.62 MHz). Chemical shifts (δH or δC) are given in parts per million (ppm) relative to the given solvent. Where appropriate, coupling constants (J) are quoted in Hz and are recorded to the nearest 0.1 Hz. The multiplicity of each signal is indicated by: singlet (s), broad singlet (br s), doublet (d), broad doublet (br d), triplet (t), doublet of doublets (dd), doublet of doublet of doublets (ddd), doublet of triplets (dt), triplet of triplets (tt), triplet of doublets (td) or multiplet (m).

Signals were assigned by means of two-dimensional NMR spectroscopy: ^{1}H-^{1}H COSY, ^{1}H-^{13}C HSQC (Heteronuclear Single Quantum Coherence) and long-range ^{1}H-^{13}C HMBC (Heteronuclear Multiple Bond Correlation).

2.3.3 Liquid Chromatography-Mass Spectrometry (LC-MS)

Optical rotations were determined at room temperature with a Schmidt+Haensch Model Polartronic-D polarimeter.

2.3.4 Karl-Fischer (KF) method

The amount of water was determined by Karl-Fischer volumetric titration method using a Metrohm Titrando KF 841 equipment.

2.3.5 Liquid Chromatography-Mass Spectrometry (LC-MS)

Negative electrospray ionization from a sample of 4-*O*-β-D galactopyranosyl-D-xylose was performed on an Agilent 1260 MSD Spectrometer. Sample was injected (10 μL) in a chromatograph equipped with a mass spectrometer detector (6130 MS) using no chromatographic column, at a 0.2 mL/min flow. The MS spectrum was acquired using a API-ES (-).

2.3.6 Thin-Layer Chromatography (TLC)

TLC was performed on pre-coated Aldrich TLC plates Macherey-Nagel silica gel 60 UV_{254} (layer: 0.20 mm silica gel with fluorescent indicator UV_{254}). Developed plates were air dried and visualized after stained with a solution of p-anisaldehyde/H_2SO_4/EtOH and heating.

2.3.7 Powder X-ray Diffraction (PXRD)

X-ray measurements of powder samples were performed with a Siemens D5000 diffractometer using Cu radiation (CuKα = 1.54056 Å, 45 kV, 35 mA) in Bragg-Brentano geometry. The diffraction patterns were recorded in the range $2° < 2\theta < 50°$ in steps of 0.02° per second and 1 second per step.

2.3.8 Single Crystal X-ray Diffraction (SC-XRD)

Single crystals of gaxilose were obtained by slow diffusion of an isopropanol-water mixture. Crystallographic data were collected on a Bruker Smart APEX-II diffractometer, using Mo-Kα radiation (λ = 0.71073 Å) and a graphite crystal monochromator. Frames taken with a 0.3° separation afforded 67577 reflections up to a 2θ max of ca. 56°. Data integration was performed using SAINT V6.45A and SORTAV [36] in the diffractometer package. The structure was solved by direct methods, using the program SIR2014 [37]. The positional parameters and the anisotropic thermal parameters of the non-H atoms were refined using SHELXL-97 [38]. All hydrogen atoms were considered as ideal and geometrically placed, except those of water molecules. Selected crystal and data collection parameters are reported in the corresponding Table 1. The crystallographic plots were made with ORTEP-3 [39] and Mercury [40]. The calculations were made using WinGX [41] and PLATON [42]. Complete crystallographic data for the structural analysis have been deposited with the Cambridge Crystallographic Data Centre, CCDC nº 1846352. Copies of this information may be obtained free of charge from the Director, Cambridge

Crystallographic Data Centre, 12 Union Road, Cambridge, CB21EZ, UK. (fax: +44-1223-336033, e-mail: deposit@ccdc.cam.ac.uk or via: www.ccdc.cam.ac.uk).

2.3.9 Thermal analysis

Thermal analyses were carried out on a simultaneous thermogravimetric analysis (TGA) - differential scanning calorimetry/differential thermal analysis (heat flow DSC /DTA) system NETZSCH -STA 449 F1 Jupiter. Accurately weighted sample was placed in an alumina pan and measured at a scan speed of 10 °C min^{-1} from ambient temperature to 300 °C under N_2 atmosphere as protective and purge gas (their respective flow velocities were 20 and 40 ml/min).

2.3.10 Attenuated Total Reflection Fourier Transform Infrared spectroscopy (ATR-FT-IR)

Infrared spectra were recorded with a Thermo Scientific Nicolet Series IS5 spectrophotometer with attenuated total reflectance accessory. The scanning range was 4000 to 450 cm^{-1} and at a resolution of 4.0 cm^{-1}.

3. Results and discussion

The desired 4-*O*-β-D galactopyranosyl-D-xylose (gaxilose) product was prepared following the procedure described in Example 9 in EP1408118B1 (Scheme 1). Unfortunately, the isolation of the solid using the recrystallization solvents therein described did not afford a product with high purity. Therefore, we decided to carry out some crystallizations to improve the quality of the product.

From two successive recrystallizations using isopropanol-water a crystalline solid, called Form A, was isolated and fully characterized as follows.

By gas chromatography the resulting product showed a purity of 98.59% in area (Figure S1 in Supplementary Information, SI).

Interpretation of the NMR spectra confirms the chemical structure of gaxilose, β anomer (Figure 2a and Figures S2-S5 and Tables S1-S2 in SI) for the obtained solid. In fact, in the proton spectrum performed in DMSO-d_6 it was possible to identify small

signals corresponding of the α anomer. Relative areas of signals 6.60 (d, J=6.4 Hz, $OH_{1'\beta}$) and 6.21 (d, J=6.4 Hz, $OH_{1'\alpha}$), show an anomeric ratio of 3:97 (α:β) (Figure 2a). ^{1}H NMR assignments were confirmed by two-dimensional ^{1}H-^{1}H-COSY experiment (Figures S6-S7), and ^{13}C-NMR assignments were done by two-dimensional ^{1}H-^{13}C-HSQC (Figure S8) and ^{1}H-^{13}C-HMBC (Figures S9-S10).

For identification purpose, LC-MS and TLC were carried out. The low-fragmentation mass spectrum clearly shows a main peak that corresponds to the adduct of molecular ion of 4-*O*-β-D galactopyranosyl-D-xylose product $[M+35]^-$, with an m/z relationship = 347.1 accompanied by a lower-intensity peaks due to isotopic effects, which appears at a m/z relationship = 349.1 and 348.1 (Figure S11 in SI). Thin-Layer Chromatography (TLC) was carried out in iPrOH:H_2O:NH_4OH (0.7:0.2:0.05) and 4-*O*-β-D galactopyranosyl-D-xylose identified (R_f = 0.14).

Optical rotation was carried out, confirming the value previously found by Montgomery et at., $[\alpha]_D$ +14.9° (H_2O) [11].

Stability of Form A was tested by stirring a slurry of this solid in methanol for three days. Analysis of the starting solid and the resulting after treatment in methanol by means of powder X-ray diffraction (PXRD) allowed us to distinguish two distinct single phases, the initial one Form A (Figure 3) and a new form hereinafter known as Form B (Figure 3 and Figure S12 in Supplementary Information).

Form A has characteristic diffraction peaks at 2theta (2θ) values: 10.8, 15.2, 15.6, 16.0, 17.5, 19.2, 23.5° and Form B, although it is not as crystalline as Form A, at 15.2, 15.95 and 20.4°.

DSC curve of Form A shows first an endotherm at 95.4 °C (ΔH = 69.5 J/g), associated to a mass loss of 10.1 % in the TGA (Figure 4a). This event is in line with a water content (by Karl Fischer method) of about 10.6%, suggesting the presence of two molecules of water. A second endothermic peak at approximately T_{onset} 204.5 °C (with a melting point at 208.7 °C) shows the melting of the compound followed by decomposition according to the important loss of weight observed in the TGA trace. Therefore, the first isolated solid Form A corresponds to gaxilose dihydrate.

To confirm this, a sample of Form A was heated to 125 °C, cooled to room temperature and analyzed by PXRD which showed the characteristic pattern of Form B (Figure 3 and Figure S12 in SI). Form B hereby isolated was further characterized as follows. Interpretation of the NMR spectra in DMSO-d_6 confirms the chemical structure of gaxilose as a mixture of α and β anomers (Figure 2b and Figures S13-S14 and Tables S3 and S4 in SI) for the obtained solid. The ^{1}H-NMR and ^{13}C-NMR were carried out by superposition of Form B and Form A structure (Figure 2c) and confirmed by two-dimensional ^{1}H-^{13}C-HSQC (Figure S15) and ^{1}H-^{13}C-HMBC (Figure S16). The ratio of the α and β anomers was measured as 34:66 for the obtained sample and at the moment of analysis.

Figure 4b shows the DSC/TGA traces of Form B. In the DSC, a single endotherm with an onset temperature at 203.4 °C was observed for the melting of the compound. In the TGA a minimal mass loss of 0.25% is observed before degradation of the compound at approximately 216 °C. Karl Fischer analysis (0.2%) corroborated the water absence. The new phase Form B is assigned as anhydrous gaxilose (Figure 3).
Hygroscopicity of Form B was studied by treatment of a sample in a climatic cabinet set at 25±1 °C and 80±2% relative humidity, as described in European Pharmacopeia 8.5 (2015) obtaining a sample of gaxilose with a KF value of 10.6%. This value corresponded to a dihydrate gaxilose (Form A). To corroborate this, the sample was therefore analysed by powder X-ray diffraction (PXRD) and ATR-FT-IR which showed the characteristic patterns of Form A (Figure 3, after hygroscopicity and Figure S17c).

Further analysis by ATR-FTIR shows that the main differences between both forms were found in the region between 3600 and 1500 cm^{-1} (Figure S17). Form A shows bands at ν_{max} (cm^{-1}): 3579, 3313, 2919, 2886, 1093, 1067 while Form B at ν_{max} (cm^{-1}): 3331, 2896, 1044, 1000 (assignments in Table 2). Stretching signals at 3578, 1666 and 1625 cm^{-1} are diagnostic for the presence of water in the sample. In the FT-IR spectrum for Form B, those signals are missing, confirming again the absence of water [22,43].

Suitable single crystals of gaxilose for structure determination by X-ray diffraction were isolated by slow diffusion from an isopropanol-water mixture at room temperature. Crystals are tetragonal with $P4_12_12$ space group. An ORTEP drawing of this structure is shown in Figure 5a with its labeling scheme. It corresponds to the β anomer. The unit

cell contains eight gaxilose molecules and sixteen water molecules (Z = 8, Table 1). Figure S18 shows a perspective view of the unit cell of gaxilose β anomer dihydrate. The water molecules are involved in eight hydrogen bonds, of a total of fifteen. All the hydroxyl groups from both monosaccharide rings establish hydrogen bonds through O-H···O interactions. From one side, towards the water molecules such as: O(1W)-H···O(2), O(3)-H···O(1W), O(4)-H···O(1W) and O(1W)-H···O(12) for the first water molecule, and O(2W)-H···O(3), O(2W)-H···O(4), O(2W)-H···O(5) and O(13)-H···O(2W) for the second one (see Figure 5b). In all these contacts, the solvent molecules act as bridges among gaxiloses. On the other hand, other hydrogen bonds are formed among disaccharide molecules through the oxygen O6, which presents disorder over two positions with the same occupation factor, which are represented as O(6) and O(6A). The O(6) oxygen in both disordered positions show contacts to the hydroxyl groups O(11) and O(13) from xylose units through the following interactions: O(11)-H···O(6) or O(11)-H···O(6A) and, also, O(13)-H···O(6) or O(13)-H···O(6A), as shown in Figure 5c. This curious disorder can be described as a *double well swing:* an intermediate site for the oxygen is less stable than lying at the O(6) and O(6A) positions.

Further intermolecular hydrogen bonds among gaxiloses were established through the glycosidic oxygen O1 (O(2)-H···O(1)), among hydroxyl groups from two galactose units (O(2)-H···O(2) and O(3)-H···O(4)) or through the xyloses (O(12)-H···O(15) or C(14)-H···O(11)). All this complex network of contacts stabilizes the crystal packing, (Table S5 in SI with Hydrogen bonds). Finally, the simulated PXRD is in excellent agreement with the experimental measured PXRD for Form A (Figure 3).

4. Conclusions

To summarize, in this paper we have isolated pure 4-*O*-β-D galactopyranosyl-D-xylose (gaxilose). Moreover, two crystalline forms (Form A and B) have been identified and characterized by many different techniques (X-ray diffraction, thermal analysis or IR spectroscopy among others) showing that Form A is a hydrated form, while Form B is an anhydrous solid. Moreover, the anomeric composition has been determined by proton NMR in DMSO-d_6 showing a different α/β ratio for both forms. Single crystal resolution of gaxilose β anomer has allowed establishing its dihydrate state confirming what was described in very early-stage literature. A second form has been obtained by

dehydration, but until now unfruitful attempts for obtaining more crystalline samples of form B resulted. Its anhydrous nature has been confirmed by TGA, KF and FTIR analysis. Also, O(6) exhibits a curious orientational disorder successfully modelled as a symmetrical double well swing.

Finally, the techniques outlined within this article have helped to clarify the large dispersion of the data regarding this compound in the literature up to date.

Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

E. Molins and M. Benito thank to Severo Ochoa FunFuture project (MICINN, CEX2019-917S) and Generalitat de Catalunya (2017SGR1687).

Conflict of interests

There are no conflicts to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at….

References

[1] M. Amiri, L. Diekmann, M. von Köckritz-Blickwede, YH. Naim, *Nutrients* 2015, **7**, 7209–7230. https://doi.org/10.3390/nu7095332

[2] M. Kuokkanen, NS. Enattah, A. Oksanen, E. Savilahti, A. Orpana, I. Järvelä, *Gut* 2003, **52**, 647–652. http://dx.doi.org/10.1136/gut.52.5.647

[3] MS. Dallas, *Genetics* 2003, **37**, 197–219. https://doi.org/10.1146/annurev.genet.37.110801.143820

[4] V. Alexandre, P.C. Even, C. Larue-Achagiotis, J. M. Blouin, F. Blachier, R. Benamouzig, D. Tomé, A.-M. Davila, *Br. J. Nutr*. 2013, **110**, 625–631. https://doi.org/10.1017/S0007114512005557

[5] FL. Suarez, DA. Savaiano, MD. Levitt, *Aliment. Pharmacol. Ther.* 1995, **9**, 589–597. https://doi.org/10.1111/j.1365-2036.1995.tb00427.x
[6] T. He, MG. Priebe, HJ. Harmsen, F. Stellaard, X. Sun, GW- Welling, RJ. Vonk, *J. Nutr.* 2006, **136**, 58–63. https://doi.org/10.1093/jn/136.1.58
[7] B. Misselwitz, D. Pohl, H. Frühauf, M. Fried, SR. Vavricka, M. Fox, *United Eur. Gastroenterol. J.* 2013, **1**, 151–159. https://doi.org/10.1177%2F2050640613484463
[8] T. Sahi, *Scand. J. Gastroenterol.* 1994, **29**, 7–20. https://doi.org/10.3109/00365529409091740
[9] F. Lifshitz, *J. Pediatr.* 1966, **69**, 229–237. https://doi.org/10.1016/S0022-3476(66)80325-6
[10] H. Arola, *Scand. J. Gastroenterol. Suppl.* 1994, **202**, 26–35. https://doi.org/10.3109/00365529409091742
[11] R. Montgomery, F. Smith, HC. Srivastava, *J. Am. Chem. Soc*. 1957, **79**, 698–700. https://doi.org/10.1021/ja01560a053
[12] JJ. Aragón, A. Fernández-Mayoralas, J. Jiménez-Barbero, M. Martín-Lomas, A. Rivera-Sagredo, D. Villanueva, *Clin. Chim. Acta* 1992, **210**, 221–226. https://doi.org/10.1016/0009-8981(92)90207-7
[13] B. Lindberg, L. Róden, B. Silvander, *Carbohydr. Res.* 1966, **2**, 413–417. https://doi.org/10.1016/S0008-6215(00)80336-5
[14] GO. Aspinall, KM. Ross, *J. Chem. Soc*. 1961, 3674-3677. https://doi.org/10.1039/JR9610003674
[15] B. Erbing, B. Lindberg, T. Norberg, *Acta Chem. Scand. B* 1978, **32**, 308-310. 10.3891/acta.chem.scand.32b-0308
[16] E. Petrakova, I. Krupova, A. Ebringerova, CS patent 271,999B1. Filed August 18, 1989.
[17] JJ. Aragón, G. Corrales, A. Fernández Mayoralas, J. Jiménez Barbero, M. Martín Lomas, A. Rivera Sagredo, D. Villanueva ES patent 2,023,556. Filed June 18, 1990.
[18] A. Rivera-Sagredo, A. Fernandez-Mayoralas, J. Jiménez-Barbero, M. Martin-Lomas, *Carbohydr. Res.* 1992, **228**, 129-135. https://doi.org/10.1016/S0008-6215(00)90554-8
[19] JJ. Aragón, FJ. Canada Vicinay, A. Fernández-Mayoralas Alvarez, R. López Alvárez, M. Martín Lomas, D. Villanueva Torregroza, EP patent 0,819,764B1. Filed November 11, 1996.
[20] F. Cañada Vicinay, A. Fernández-Mayoralas Álvarez, M. Martín Lomas, E. Montero Prieto, ES patent 2,130,073B1. Filed May 28, 1997.
[21] FJ. Cañada Vicinay, G. Corrales Morales, A. Fernández-Mayoralas Álvarez, M. Martin Lomas, JJ. Aragón Reyes, EP patent 1,408,118B1. Filed June 14, 2002.
[22] JD. Dunitz, J. Bernstein, *Acc. Chem. Res.* 1995, **28**, 193–200. https://doi.org/10.1021/ar00052a005
[23] J. Bernstein, *Polymorphism in molecular crystals*, Oxford University Press, USA. 2002.
[24] A. Nangia, G. Desiraju, *Chem. Commun.* 1999, 605–606. https://doi.org/10.1039/A809755K
[25] F. Takusagawa, RA. Jacobson, *Acta Cryst. B.* 1978, **34**, 213 – 218. https://doi.org/10.1107/S0567740878002630
[26] C. Burden, W. Mackie, B. Sheldrick, *Acta Cryst. C*. 1986, **42**, 177–179. https://doi.org/10.1107/S0108270186096865

[27] GA. Jeffrey, R. Nanni, *Carbohydr. Res.* 1985, **137**, 21–30. https://doi.org/10.1016/0008-6215(85)85146-6
[28] H. Nagase, N. Ogawa, T. Endo, M. Shiro, H. Ueda, M. Sakurai, *J. Phys. Chem. B* 2008, **112**, 9105–9111.
[29] A. Nezzal, L. Aerts, M. Verspaille, G. Henderickx, A. Redl, *J. Cryst. Growth* 2009, **311**, 3863–3870. https://doi.org/10.1016/j.jcrysgro.2009.06.003
[30] S. Kamitori, A. Ueda, Y. Tahara, H. Yoshida, T. Ishii, J. Uenishi, *Carbohydr. Res.* 2011, **346**, 1182–1185. https://doi.org/10.1016/j.carres.2011.04.003
[31] N. Verhoeven, TL. Neoh, T. Furuta, C. Yamamoto, T. Ohashi, H. Yoshii, *Food Chem.* 2012, **132**, 1638–1643. https://doi.org/10.1016/j.foodchem.2011.06.010
[32] S. Garnier, S. Petit, G. Coquerel, *J. Therm. Anal. Calorim*. 2002, **68**, 489–502. https://doi.org/10.1023/A:1016087702409
[33] JH. Smith, SE. Dann, MRJ. Elsegood, SH. Dale, CG. Blatchford, *Acta Cryst. E*. 2005, **61**, o2499 – o2501. https://doi.org/10.1107/S1600536805021367
[34] JH. Kirk, SE. Dann, CG. Blatchford, *Int. J. Pharm.* 2007, **334**, 103–114. https://doi.org/10.1016/j.ijpharm.2006.10.026
[35] MJ. Altamimi, K. Wolff, A. Nokhodhci, GP. Martin, PG. Royall, *Int. J. Pharm*. 2019, **555**, 237-249. https://doi.org/10.1016/j.ijpharm.2018.10.061
[36] Blessing, R. H. *Acta Cryst*. A. 1995, **51**, 33-38. https://doi.org/10.1107/S0108767394005726
[37] MC. Burla, R. Caliandro, B. Carrozzini, GL. Cascarano, C. Cuocci, C. Giacovazzo, M. Mallamo, A. Mazzone, G. Polidori. *J. Appl. Crystallogr.* 2015, **48**, 306–309. https://doi.org/10.1107/S1600576715001132
[38] GM. Sheldrick, *Acta Cryst A.* 2008, **64**, 112–122. https://doi.org/10.1107/S0108767307043930
[39] LJ. Farrugia, *J. Appl. Crystallogr.* 1997, **30**, 565. https://doi.org/10.1107/S0021889897003117
[40] CF. Macrae, IJ. Bruno, JA. Chisholm, PR. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, PA. Wood, *J. Appl. Crystallogr*. 2008, **41**, 466–470. https://doi.org/10.1107/S0021889807067908
[41] LJ. Farrugia, *J. Appl. Crystallogr*. 1999, **32**, 837–838. https://doi.org/10.1107/S0021889899006020
[42] AL. Spek, *Acta Cryst. D*. 2009, **65**, 148–155. https://doi.org/10.1107/S090744490804362X
[43] E. Wiercigroch, E. Szafraniec, K. Czamara, MZ. Pacia, K. Majzner, K. Kochan, A. Kaczor, M. Baranska, K. Malek, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc*. 2017, **185**, 317–335. https://doi.org/10.1016/j.saa.2017.05.045

Legend of Figures

Scheme 1. Synthetic scheme for gaxilose, 4-O-β–D-galactopyranosyl-β–D-xylopyranose, and its preparation following Example 9 in EP1408118B1.

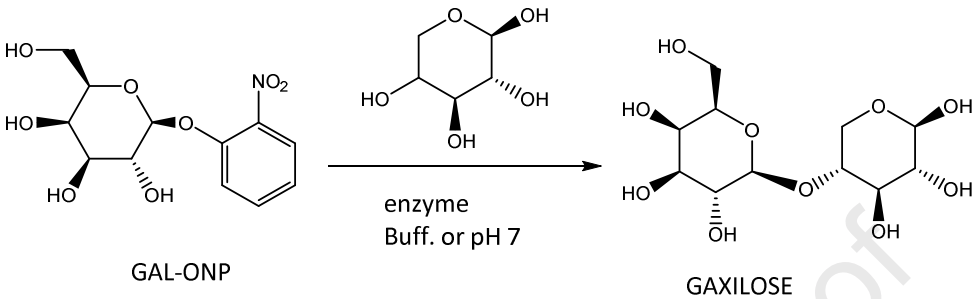
Figure 1. Molecular structure of gaxilose with numbering for NMR interpretation and its α and β anomers: 4-O-β-D-galactopyranosyl-α-D-xylopyranose (α anomer) and 4-O-β-D-galactopyranosyl-β-D-xylopyranose (β anomer).

Figure 2. ^{1}H-NMR spectra of gaxilose a) Form A, b) Form B and c) superposition of the two ^{1}H-NMR spectra showing anomeric compositions (insets) in DMSO-d_6.

Figure 3. PXRD patterns of simulated Form A, experimental Form A and Form B and Form A after hygroscopicity study of Form B.

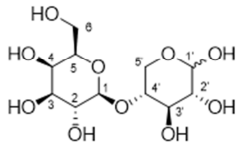
Figure 4. DSC and TGA curves of gaxilose Form A (a) and Form B (b).

Figure 5. a) ORTEP drawing of gaxilose β anomer (without the water molecules), b) hydrogen bonds established with the water molecules (O1W and O2W), c) hydrogen bonds formed with O(6) and O(6A), showing the *double well swing*.

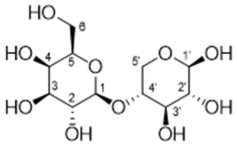


GAL-ONP

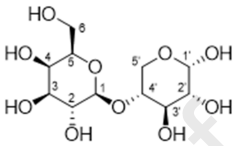
GAXILOSE



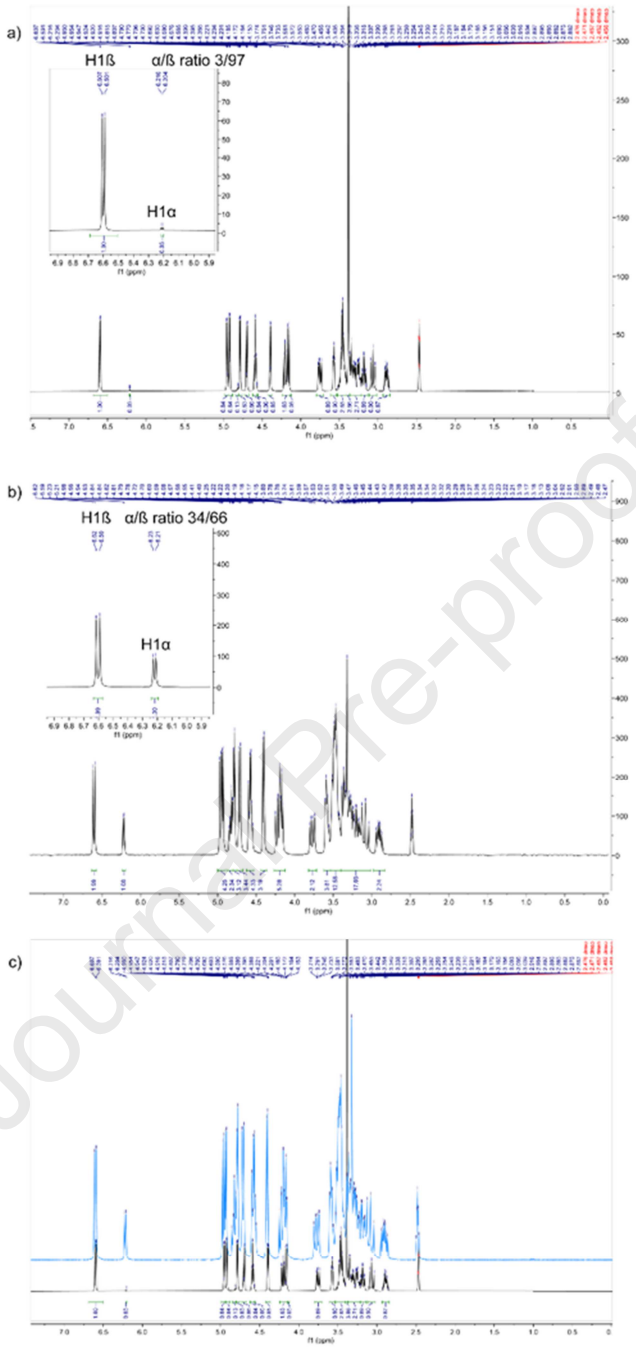
Gaxilose

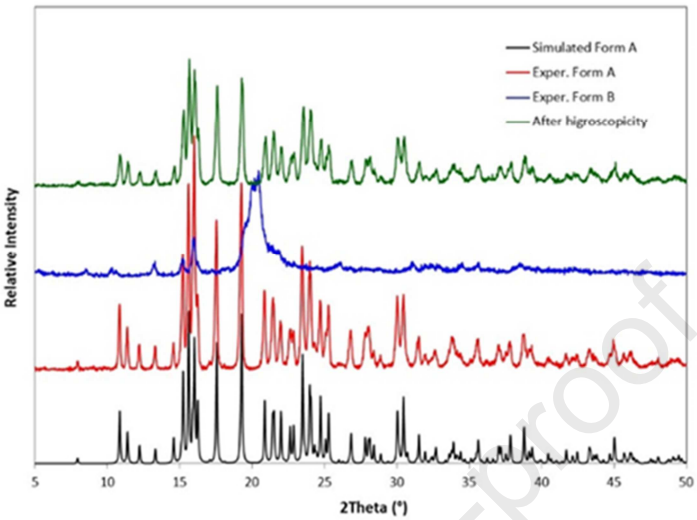


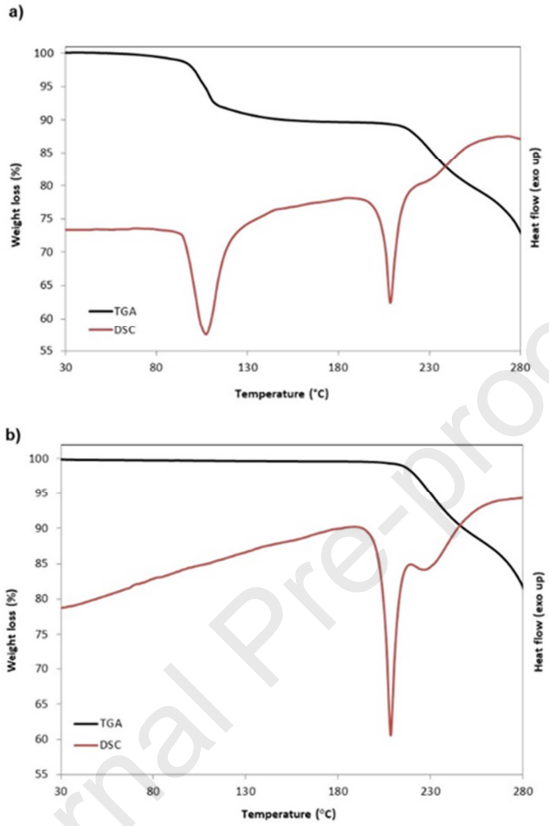
β-anomer



α-anomer







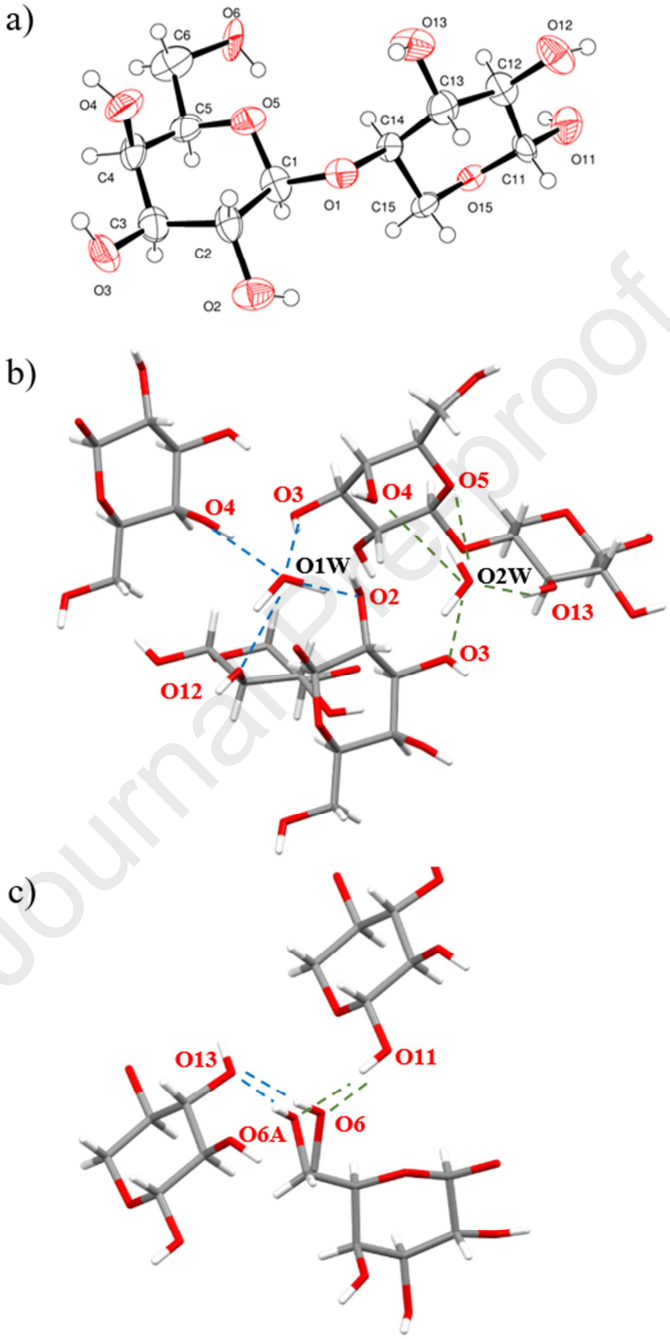


Table 1. Crystallographic data and structural refinement parameters.

Empirical formula	$C_{11}H_{24}O_{12}$
Formula weight	348.30
Temperature (K)	294(2)
Wavelength (Å)	0.71073
Crystal system	Tetragonal
Space group	P 4_1 2_1 2
a (Å)	8.2905(4)
b (Å)	8.2905(4)
c (Å)	44.583(2)
α (°)	90
β (°)	90
γ (°)	90
Volume ($Å^3$)	3064.3(3)
Z	8
ρ_{calc} (Mg/m^3)	1.510
R_1 ($I>2\sigma(I)$)	0.0710
wR_2	0.1392

Table 2. IR spectral assignments of gaxilose Forms A, Form B and Form A from rehydration of Form B.

Form A	**Form B**	**Form A from rehydration of Form B**	**Assignment**
3579	-	3578	Water, Stretching OH
3313	3331	3311	Alcohol, Stretching OH
2919, 2886	2896	2919, 2886	Aliphatic, Stretching C-H
1667, 1625	-	1669, 1625	Water, Bending OH
1093	1044	1093	Ether, asym Stretching C-O-C
1068	1000	1068	Alcohol, Stretching C-O

Highlights

- Synthesis of pure 4-*O*-β-D galactopyranosyl-D-xylose, known as gaxilose.
- Characterization by XRPD, thermal methods and FTIR of two solid forms of gaxilose.
- Single crystal resolution of the dihydrated gaxilose.

Supplementary Information

Characterization of crystalline forms of gaxilose, a diagnostic drug

Federica Catti[1,†], Santos Hernández[1], Mónica Benito[2], Elies Molins[2,*], Francisco Marquillas-Olóndriz[1,*]

[1]Interquim S.A., R&D Department, C/Joan Buscallà, 10, 08173 Sant Cugat del Vallès, Barcelona, Spain.

[2]Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus UAB, 08193 Bellaterra, Spain

[†] Present address: Arkansas State University, Campus Queretaro, Carretera Estatal 100, km 17.5, Municipio Colón, 76270, Querétaro, Mexico

[*]Corresponding authors

Tel.: +034 935 801 853, *E-mail address:* elies.molins@icmab.es (E. Molins),

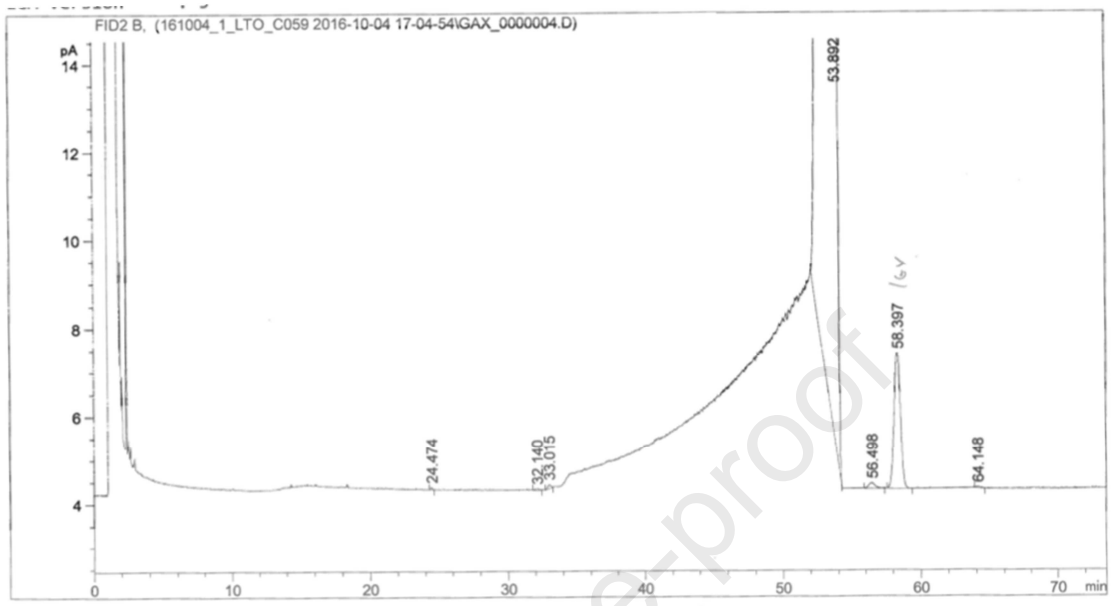
Tel.: +034 935 044 218, *E-mail address:* fmarquillas@ferrer.com (F. Marquillas Olóndriz)

Table of contents

Description
Figure S1. The GC-FID analysis of gaxilose Form A showed a purity of 98.59% in area.
Figure S2. Portion of ^{1}H-NMR spectrum of gaxilose Form A in DMSO-d_6 (range 6.8 – 4.0 ppm).
Figure S3. Portion of the ^{1}H-NMR spectrum of gaxilose Form A in DMSO-d_6 (range 4.0 - 2.70 ppm).
Table S1. ^{1}H NMR chemical shifts and assignments for gaxilose Form A in DMSO-d_6.
Figure S4. ^{13}C-NMR spectrum of gaxilose Form A in DMSO-d_6.
Figure S5. Portion ^{13}C-NMR spectrum of gaxilose Form A in DMSO-d_6.
Table S2. ^{13}C NMR chemical shifts and assignments for gaxilose Form A in DMSO-

d_6.
Figure S6. ^{1}H-^{1}H COSY NMR spectrum of gaxilose Form A in DMSO-d_6.
Figure S7. Portion of ^{1}H-^{1}H COSY NMR spectrum of gaxilose Form A in DMSO-d_6.
Figure S8. ^{1}H-^{13}C HSQC NMR spectrum of gaxilose Form A in DMSO-d_6.
Figure S9. ^{1}H-^{13}C HMBC NMR spectrum of gaxilose Form A in DMSO-d_6.
Figure S10. Portions of the ^{1}H-^{13}C HMBC NMR spectrum of gaxilose Form A in DMSO-d_6.
Figure S11. Negative electrospray ionization from a sample of gaxilose Form A. The low-fragmentation mass spectrum shows a main peak that corresponds to the adduct of molecular ion of gaxilose product $[M+35]^-$, with an m/z relationship = 347.1 accompanied by a lower-intensity peaks due to isotopic effects, which appears at a m/z relationship = 349.1 and 348.1.
Figure S12. Powder X-ray diffraction patterns (PXRD) of gaxilose Form B obtained from Form A by stirring in methanol for three days (batch 1) and by heating to 125 ºC and cooling to room temperature (batch 2)
Figure S13. Portions of the ^{1}H-NMR spectrum of gaxilose Form B in DMSO-d_6.
Table S3. ^{1}H NMR chemical shifts and assignments for gaxilose Form B in DMSO-d_6
Figure S14. ^{13}C-NMR spectrum of gaxilose Form B in DMSO-d_6.
Table S4. ^{13}C NMR chemical shifts and assignments for gaxilose Form B in DMSO-d_6.
Figure S15. ^{1}H-^{13}C HSQC NMR spectrum of gaxilose Form B in DMSO-d_6.
Figure S16. Portions of the ^{1}H-^{13}C HMBC NMR spectrum of gaxilose Form B in DMSO-d_6.
Figure S17. ATR-FTIR spectra of gaxilose Form A (a), Form B (b) and Form A from rehydration of Form B (c).
Figure S18. Perspective view of the crystalline unit cell of gaxilose form A.
Table S5. Hydrogen bonds for gaxilose dihydrate Form A [Å and °].

Figure S1. The GC-FID analysis of gaxilose Form A showed a purity of 98.59% in area.



Signal 1: FID2 B,

Peak #	RetTime [min]	Type	Width [min]	Area [pA*s]	Height [pA]	Area %
1	24.474	BB	0.1132	5.08307e-1	5.74534e-2	0.00670
2	32.140	BB	0.2113	5.90049e-1	3.32085e-2	0.00778
3	33.015	BV	0.1932	1.31858	8.21607e-2	0.01738
4	53.892	BB	0.6962	7481.36816	131.41833	98.59072
5	56.498	BB	0.4375	4.36690	1.26882e-1	0.05755
6	58.397	BB	0.4601	99.37482	3.08250	1.30958
7	64.148	BB	0.3462	7.81758e-1	2.69679e-2	0.01030
Totals :				7588.30858	134.82751	

Figure S2. Portion of ^{1}H-NMR spectrum of gaxilose Form A in DMSO-d_6 (range 6.8 – 4.0 ppm).

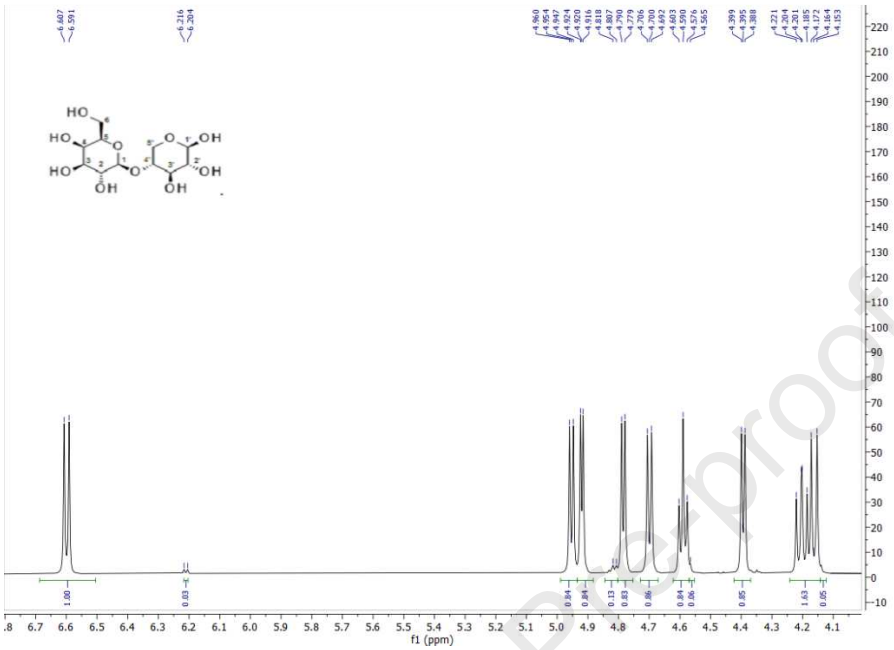


Figure S3. Portion of the ^{1}H-NMR spectrum of gaxilose Form A in DMSO-d_6 (range 4.0 - 2.70 ppm).

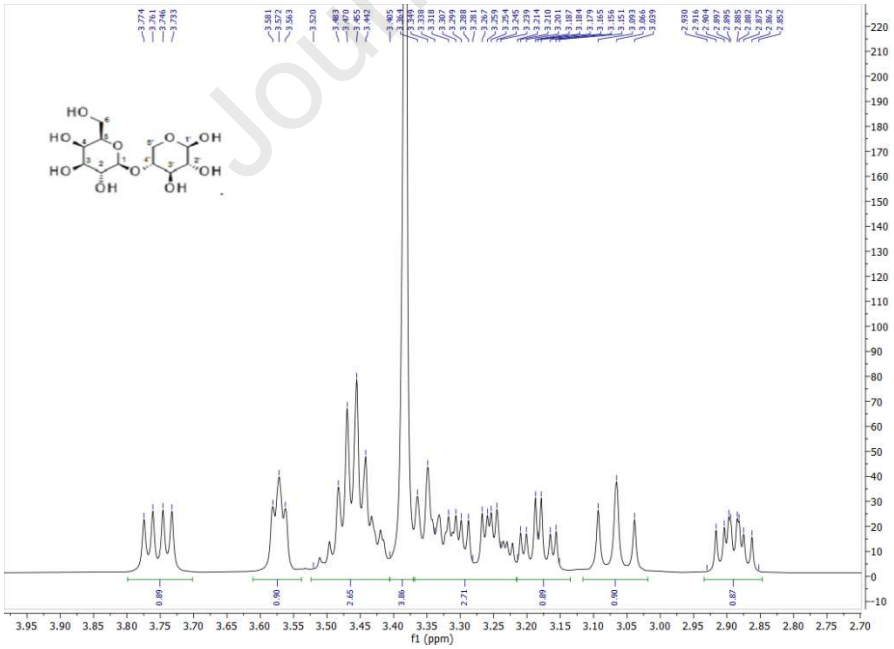


Table S1. ^{1}H NMR chemical shifts and assignments for gaxilose Form A in DMSO-d_6.

δ (ppm)	Assignments	δ (ppm)	Assignments
6.60	d, 1H, J=6.4 Hz, $OH_{1'}$	4.95	d, 1H, J=8 Hz, $OH_{2'}$
4.92	d, 1H, J=3.2 Hz, $OH_{3'}$	4.81	d, 1H, J=4.4 Hz, OH_2
4.78	d, 1H, J=8 Hz, OH_3	4.59	t, 1H, J=4 Hz, OH_6
4.39	d, 1H, J=4 Hz, OH_4	4.20	dd, 1H, J=6.8, 8 Hz, $H_{1'}$
4.16	d, 1H, J=8 Hz, H_1	3.75	dd, 1H, J= 5.2, 11.2 Hz, $H_{5'}$
3.57	t, 1H. J= 4.4 Hz, H_4	3.51-3.42	m, 3H, $H_{4'}$, H_6
3.36-3.33	m, 1H, H_5	3.31-3.29	m, 1H, H_2
3.25	dd, 1H, J= 3.2, 5.2 Hz, H_3	3.18	dt, 1H, J= 3.6, 8.8 Hz, $H_{3'}$
3.06	t, 1H, 13.2 Hz, $H_{5'}$	2.92-2.86	m, 1H, $H_{2'}$

Figure S4. ^{13}C-NMR spectrum of gaxilose Form A in DMSO-d_6.

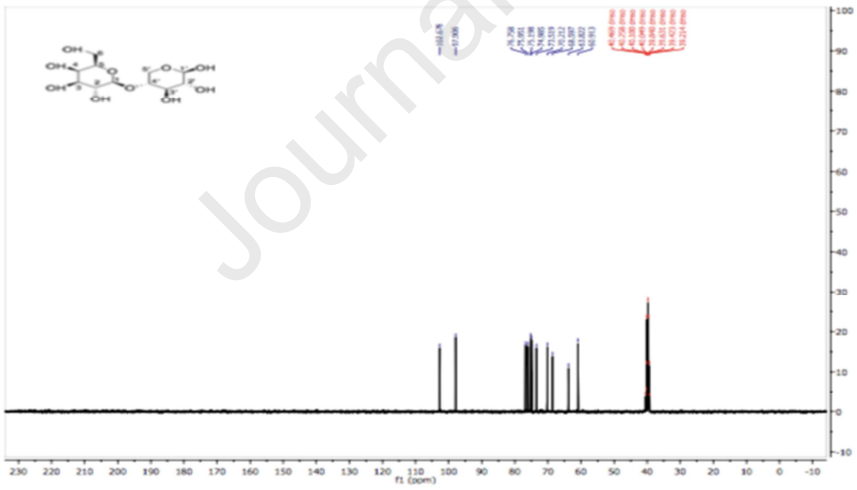


Figure S5. Portion ^{13}C-NMR spectrum of gaxilose Form A in DMSO-d_6.

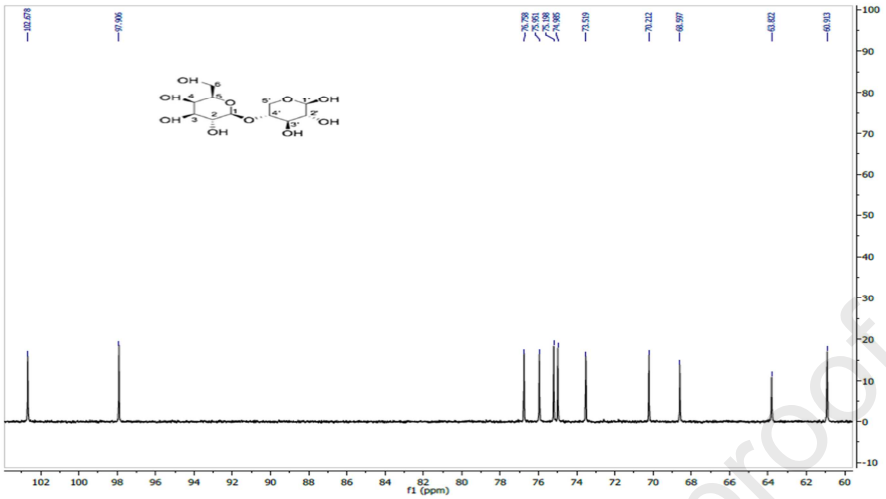


Table S2. ^{13}C NMR chemical shifts and assignments for gaxilose Form A in DMSO-d_6.

δ (ppm)	Assignments	δ (ppm)	Assignments
102.6	C_1	97.9	$C_{1'}$
76.7	$C_{4'}$	75.9	C_5
75.1	$C_{3'}$	74.9	$C_{2'}$
73.5	C_3	70.2	C_2
68.5	C_4	63.8	$C_{5'}$
60.9	C_6		

Figure S6. ^{1}H-^{1}H COSY NMR spectrum of gaxilose Form A in DMSO-d_6.

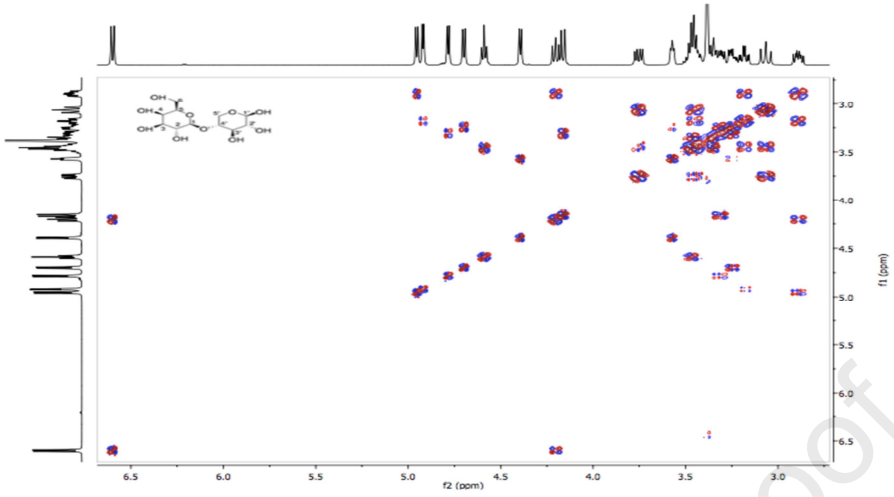


Figure S7. Portion of ^{1}H-^{1}H COSY NMR spectrum of gaxilose Form A in DMSO-d_6.

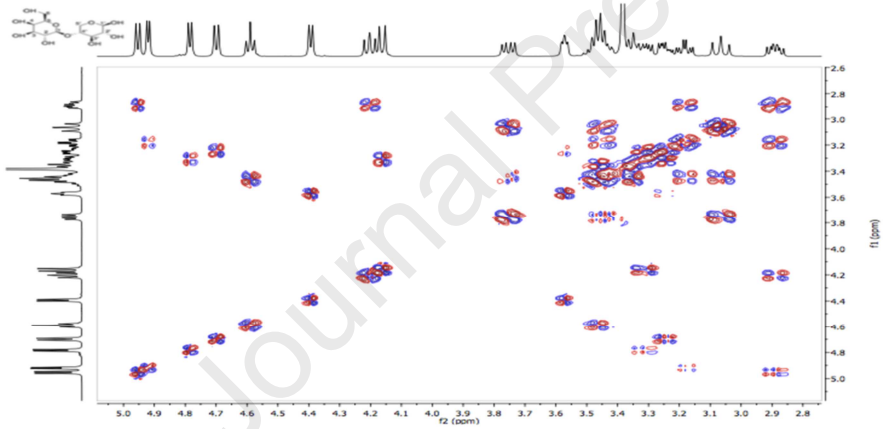


Figure S8. ^{1}H-^{13}C HSQC NMR spectrum of gaxilose Form A in DMSO-d_6.

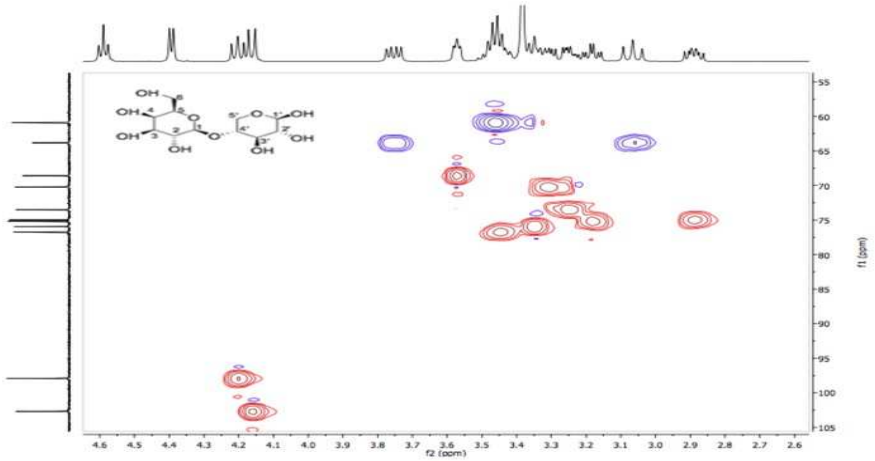


Figure S9. ^{1}H-^{13}C HMBC NMR spectrum of gaxilose Form A in DMSO-d_6.

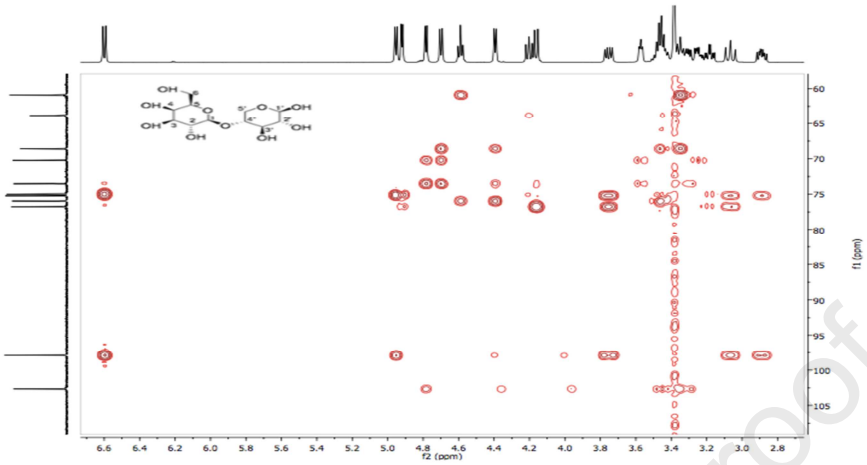


Figure S10. Portions of the ^{1}H-^{13}C HMBC NMR spectrum of gaxilose Form A in DMSO-d_6.

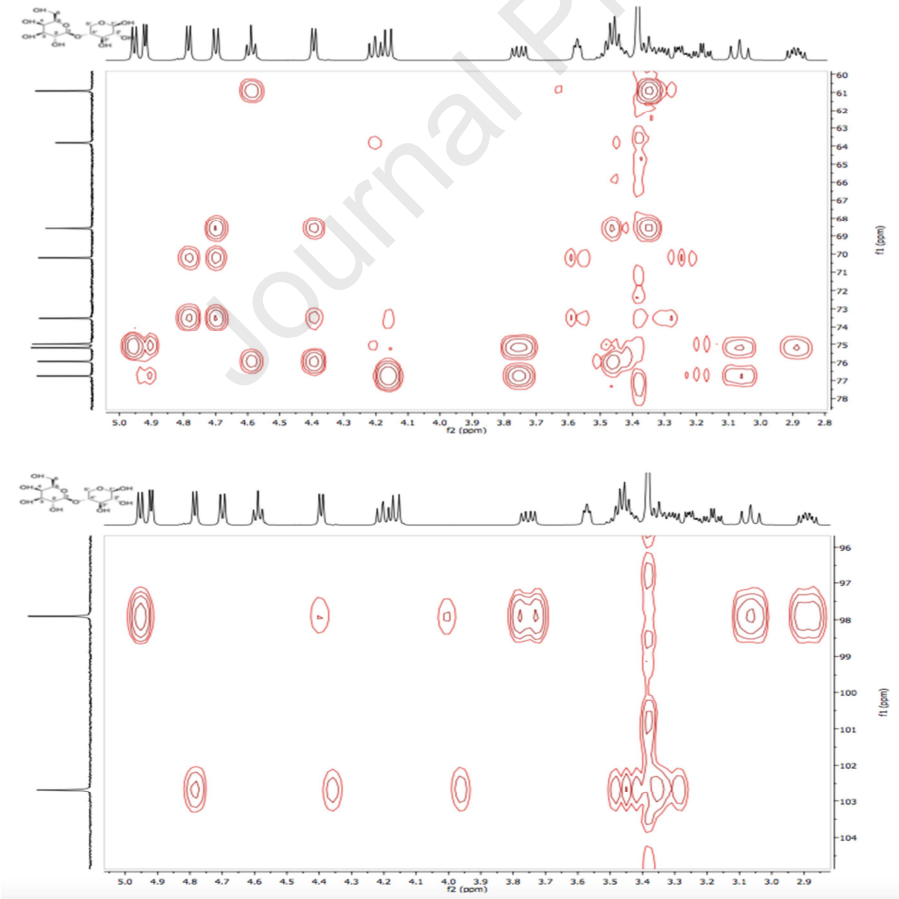


Figure S11. Negative electrospray ionization from a sample of gaxilose Form A. The low-fragmentation mass spectrum shows a main peak that corresponds to the adduct of molecular ion of gaxilose product $[M+35]^-$, with an m/z relationship = 347.1 accompanied by a lower-intensity peaks due to isotopic effects, which appears at a m/z relationship = 349.1 and 348.1.

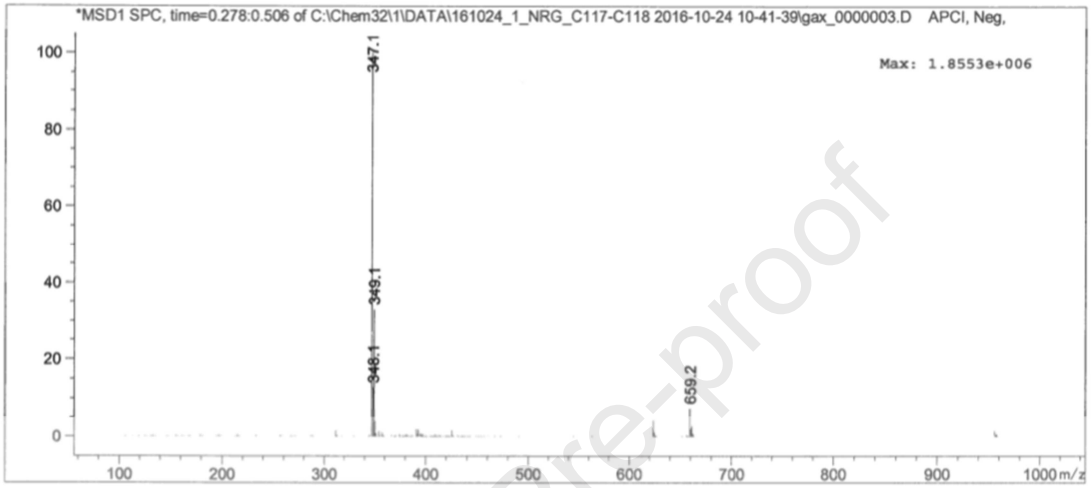


Figure S12. Powder X-ray diffraction patterns (PXRD) of gaxilose Form B obtained from Form A by stirring in methanol for three days (batch 1) and by heating to 125 ºC and cooling to room temperature (batch 2).

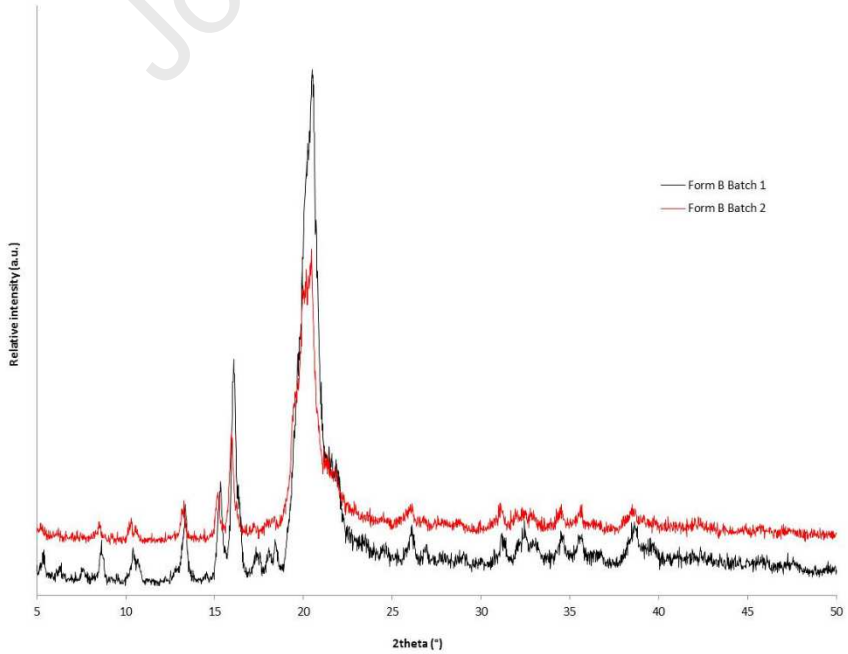


Figure S13. Portions of the ^{1}H-NMR spectrum of gaxilose Form B in DMSO-d_6.

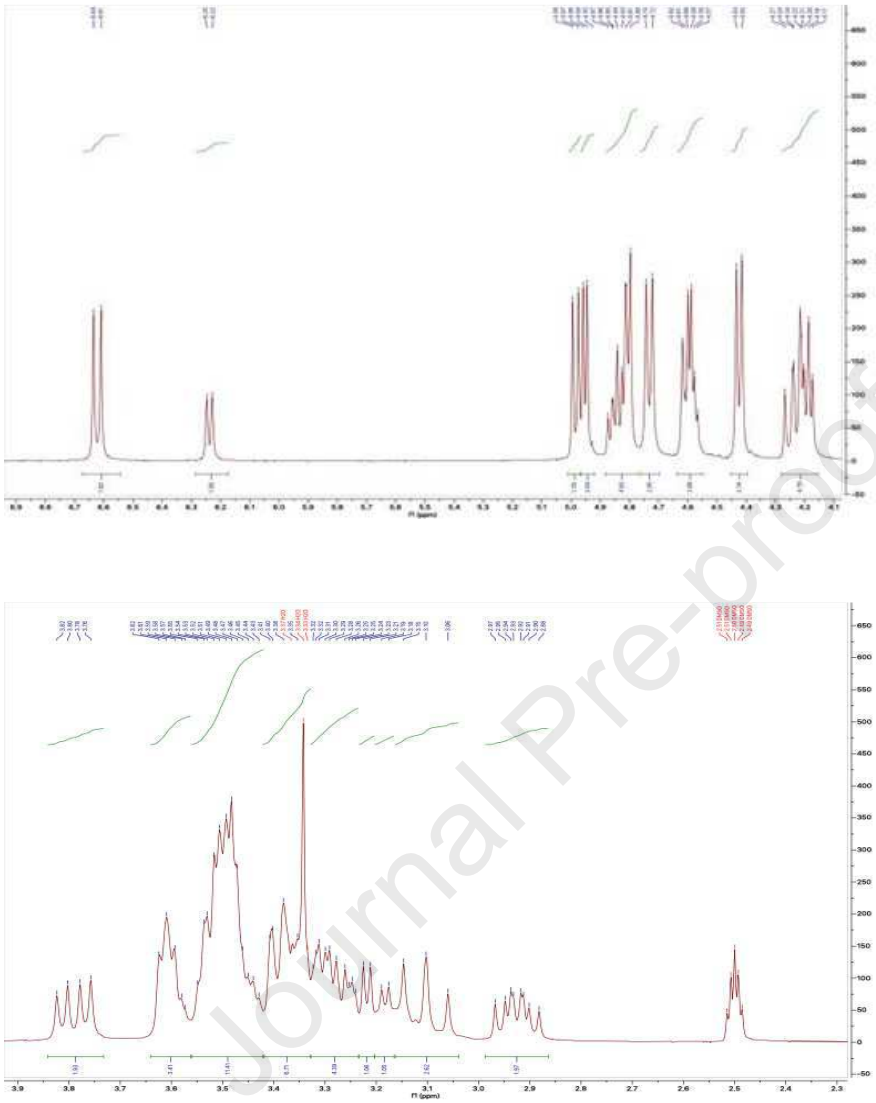


Table S3. ^{1}H NMR chemical shifts and assignments for gaxilose Form B in DMSO-d_6.

δ (ppm)	Assignments	δ (ppm)	Assignments
6.62	d, 1H, J=6.4 Hz, $OH_{1'\beta}$	6.24	d, 1H, J=8 Hz, $OH_{1'\alpha}$
4.98	d, 1H, J= 8 Hz, $H_{1'\alpha}$	4.95	d, 1H, J=8Hz, $OH_{3'\beta}$
4.88-4.82	m, 1H, $OH_{3'\alpha}$	4.80	d, 2H, J=5.6 Hz, $OH_{2\alpha}$, $OH_{2\beta}$
4.73	d, 2H, J=8 Hz, $OH_{3\alpha}$, $OH_{3\beta}$	4.63-4.55	m, 2H, $OH_{6\alpha}$, $OH_{6\beta}$
4.42	d, 2H, J=7.2 Hz, $OH_{4\alpha}$,$OH_{4\beta}$	4.28-4.15	m, 4H, $H_{1\alpha}$,$H_{1\beta}$, $H_{1'\alpha}$, $H_{1'\beta}$
3.82-3.80	dd, 1H, J= 5.2, 11.2 Hz, $H_{5'\beta}$	3.60	t, 2H. J= 4.4 Hz, $H_{4\alpha}$, $H_{4\beta}$
3.56-3.42	m, 7H, $H_{2'\alpha}$,$H_{3'\alpha}$,$H_{4'\alpha}$, $H_{4'\beta}$,. $H_{5\alpha'}$, $H_{6\alpha}$, $H_{6\beta}$	3.42-3.36	m, 2H, $H_{5\alpha}$. $H_{5\beta}$
3.32-3.23	m, 4H, $H_{2\alpha}$, $H_{2\beta}$, $H_{3\alpha}$, $H_{3\beta}$	3.21	dt, 1H, J= 3.6, 8.8Hz, $H_{3'\beta}$
3.1	t, 1H, 13.2 Hz, $H_{5'\beta}$	2.98-2.86	m, 1H, $H_{2'\beta}$

Figure S14. ^{13}C-NMR spectrum of gaxilose Form B in DMSO-d_6.

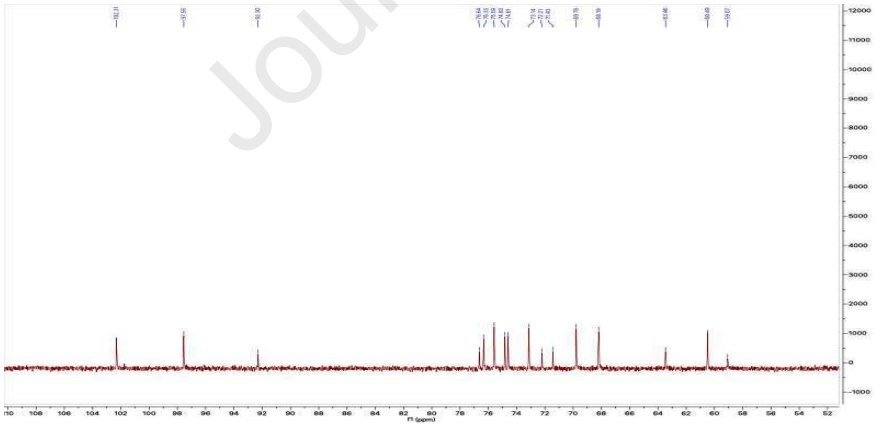


Table S4. ^{13}C NMR chemical shifts and assignments for gaxilose Form B in DMSO-d_6.

δ (ppm)	Assignments	δ (ppm)	Assignments
102.3	$C_{1\beta}$, $C_{1\,\alpha}$	97.5	$C_{1'\beta}$
92.3	$C_{1'\alpha}$	76.7	$C_{4'\alpha}$
76.3	$C_{4'\beta}$	75.6	$C_{5\alpha}$, $C_{5\beta}$
74.8	$C_{3'\beta}$	74.6	$C_{2'\beta}$
73.1	$C_{3\alpha}$, $C_{3\beta}$	72.2	$C_{3'\alpha}$
71.4	$C_{2'\alpha}$	69.7	$C_{2\alpha}$, $C_{2\beta}$
68.2	$C_{4\alpha}$,$C_{4\beta}$	63.5	$C_{5'\beta}$
60.5	$C_{6\alpha}$, $C_{6\beta}$	59.1	$C_{5'\alpha}$

Figure S15. ^{1}H-^{13}C HSQC NMR spectrum of gaxilose Form B in DMSO-d_6.

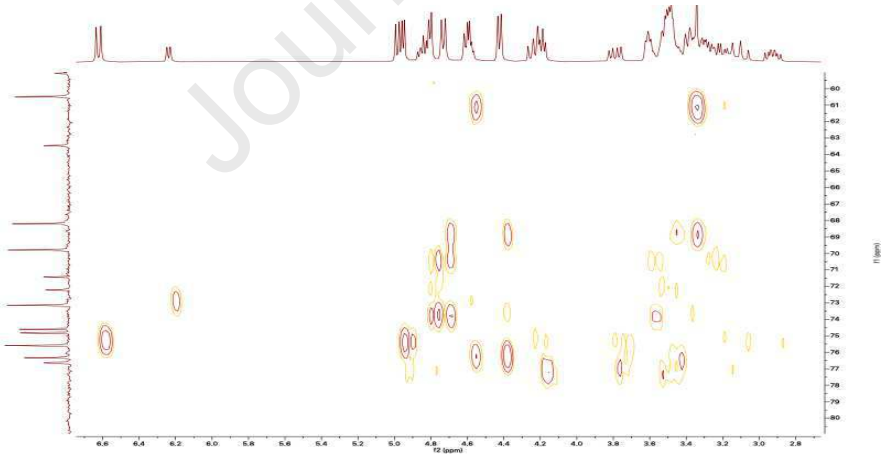


Figure S16. Portions of the ^{1}H-^{13}C HMBC NMR spectrum of gaxilose Form B in DMSO-d_6.

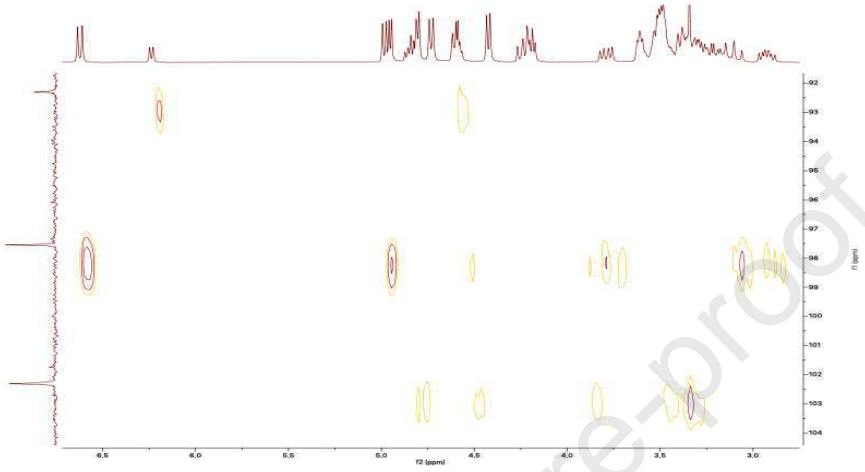


Figure S17. ATR-FTIR spectra of gaxilose Form A (a), Form B (b) and Form A from rehydration of Form B (c).

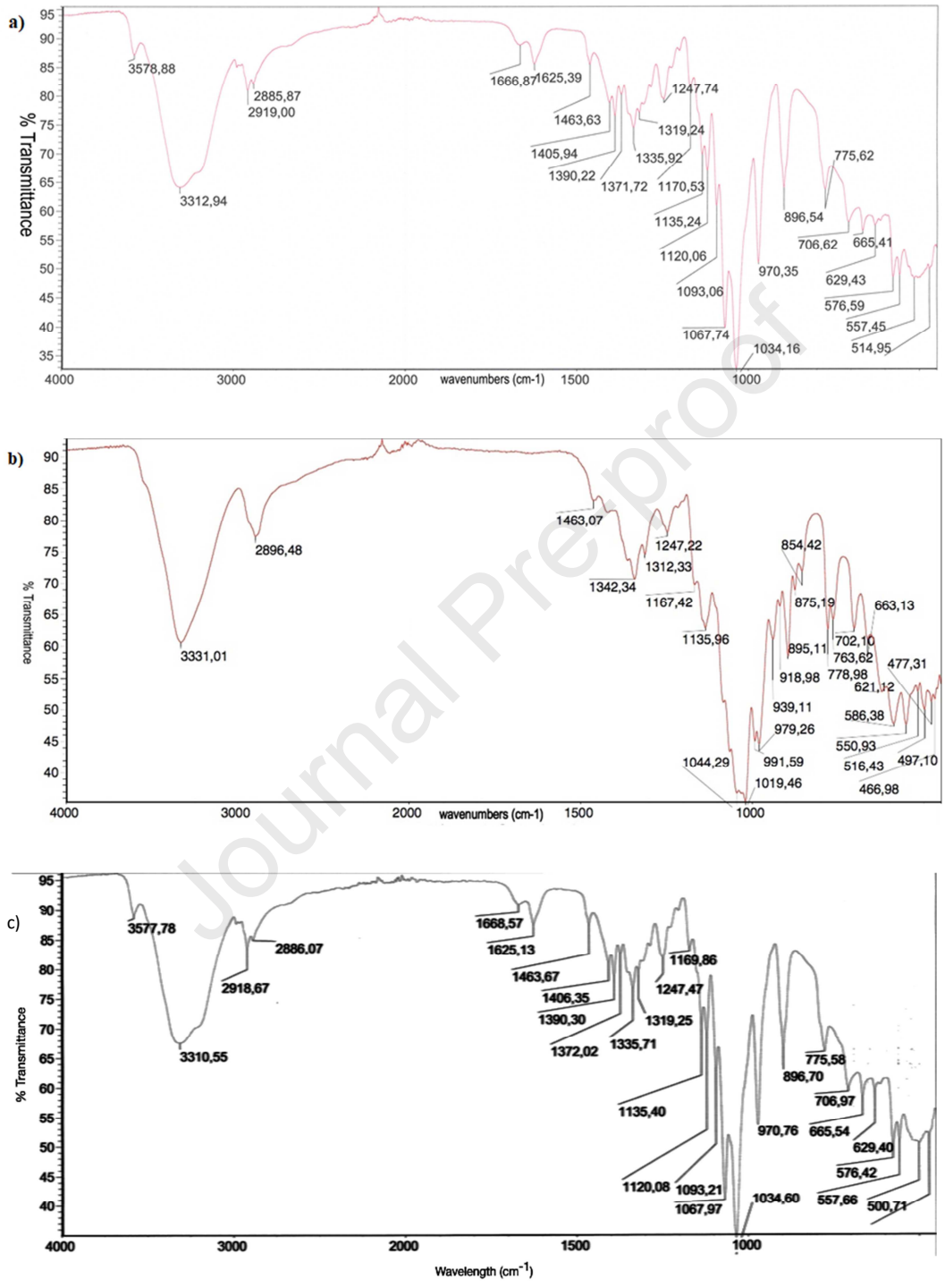


Figure S18. Perspective view of the crystalline unit cell of gaxilose form A.

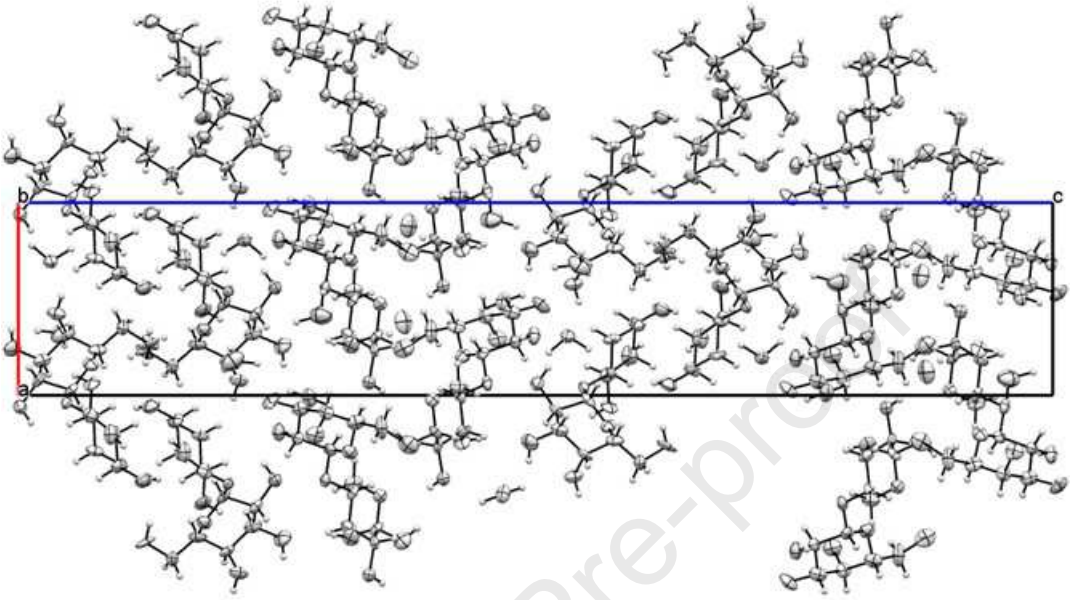


Table S5. Hydrogen bonds for gaxilose dihydrate Form A [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
O(2)-H(2)...O(1)#1	0.82	2.52	3.143(5)	134.2
O(2)-H(2)...O(2)#1	0.82	2.20	2.931(8)	149.3
O(3)-H(3)...O(4)	0.82	2.41	2.822(5)	112.1
O(4)-H(4)...O(1W)#2	0.82	1.99	2.786(6)	165.0
O(6)-H(6E)...O(13)#3	0.82	1.95	2.747(9)	164.4
O(6A)-H(6F)...O(13)#3	0.82	2.00	2.803(9)	167.6
O(11)-H(11)...O(6)#4	0.82	1.89	2.672(10)	159.5
O(11)-H(11)...O(6A)#4	0.82	2.12	2.929(12)	171.5
O(12)-H(12)...O(15)#4	0.82	2.09	2.874(5)	159.5
O(13)-H(13)...O(2W)	0.82	2.11	2.836(7)	148.1
C(14)-H(14)...O(11)#5	0.98	2.42	3.231(6)	140.3
O(1W)-H(1W)...O(12)#6	1.07(2)	1.95(7)	2.783(5)	132(6)
O(1W)-H(2W)...O(2)	1.09(3)	2.05(8)	2.791(5)	123(6)
O(2W)-H(4W)...O(4)	1.10(2)	2.45(6)	3.272(7)	130(4)
O(2W)-H(4W)...O(5)	1.10(2)	2.01(2)	3.033(6)	153(5)
O(3)-H(3)...O(1W)#7	0.82	2.07	2.848(6)	158.7
O(2W)-H(3W)...O(3)#7	1.05(2)	1.95(2)	2.729(6)	128(3)

Symmetry transformations used to generate equivalent atoms:

#1 y-1,x+1,-z #2 x-1,y,z #3 -x-1/2,y+1/2,-z+1/4

#4 -x+1/2,y-1/2,-z+1/4 #5 -x+1/2,y+1/2,-z+1/4

#6 x,y+1,z #7 y-2,x+1,-z

Highlights

- Synthesis of the disaccharide gaxilose, used as a diagnostic drug for a non-invasive method for hypolactasia diagnosis, and obtention of forms A and B.
- Characterization of gaxilose forms A and B by thermal analysis, spectroscopic, and X-ray diffraction methods.
- The crystal structure of the gaxilose form A, a dihydrate, has been obtained and discussed.
- The gaxilose form B is anhydrous.
- The network of hydrogen bonds in gaxilose form A has been discussed, showing a double well swing disorder at the O(6) position.