PAPER

[View Article Online](#)
[View Journal](#) | [View Issue](#)Highly efficient and regioselective enzymatic synthesis of β -(1 $\rightarrow$ 3) galactosides in biosolvents†Cite this: *RSC Advances*, 2013, 3, 12155Carlos Bayón,^a Álvaro Cortés,^b Antonio Aires-Trapote,^a Concepción Civera^c and María José Hernáiz^{*a}

A green synthesis for β -(1 $\rightarrow$ 3) galactosyldisaccharides that combines the use of a biodegradable biocatalyst, aqueous solutions, and solvent recycling (renewable and derived from biomass) has been developed. The use of biomass-derived solvents allows good catalytic activity in the synthesis of Gal- β -D-(1 $\rightarrow$ 3)GlcNAc and Gal- β -D-(1 $\rightarrow$ 3)GalNAc (99% and 95% yields respectively) with β -Gal-3-NTag β -galactosidase, preventing hydrolytic activity and with full regioselectivity. This represents a considerable improvement over the use of an aqueous buffer or conventional organic solvents. Furthermore, reaction scaling up and biosolvent recycling are feasible without losing catalytic action. In order to understand the role of these green solvents in the enzyme's synthetic behaviour, different structural studies were performed (fluorescence and molecular modelling) in the presence of some selected biosolvents to conclude that the presence of green biosolvents in the reaction media modifies the enzyme's tertiary structure allowing better substrate disposition in the active site, most probably due to solvation effects, explaining the behaviour observed.

Received 20th February 2013,
Accepted 3rd May 2013

DOI: 10.1039/c3ra40860d

www.rsc.org/advances

Introduction

There is a current need in the chemical and biochemical industry to develop green synthetic routes, involving renewable raw materials, to replace unfriendly ones that usually generate toxic residues and/or involve harsh reaction conditions.

Nowadays Gal- β -D-(1 $\rightarrow$ 3)GalNAc and Gal- β -D-(1 $\rightarrow$ 3)GlcNAc are very well known epitopes associated with carcinoma cells and other crucial intercellular recognition events.^{1–6} Both are also important constituents of mucin type or complex type glycoproteins⁷. Therefore, there is a great interest in the synthesis of these compounds in order to deepen our knowledge of their biological functions. Organic chemical methods for obtaining them have been developed^{8–13} but they involve several elaborate protection and deprotection procedures. On the other hand, enzymatic synthesis using glycosyl transferases has also been reported.^{14–16} Glycosyl transferases are widely used to perform regiospecific galactosylation and silylation on a preparative scale, but these enzymes belong to

the Leloir pathway and are difficult to obtain and have limited stability. Moreover, they require expensive cofactors as glycosyl donors. Glycosyl hydrolases (glycosidases) can also be used to synthesize oligosaccharides in a kinetically controlled reaction, in which a glycosyl donor is used to transfer its glycosyl residue to a sugar acceptor, present in the reaction medium.^{16,17} In spite of the increasing amount of work carried out with glycosyl hydrolases, their main drawbacks are low yields and lack of regioselectivity,¹⁸ which limits their use for synthetic purposes. In an important contribution to glycosidase-catalyzed oligosaccharides synthesis, Ito *et al.*¹⁹ reported for the first time the use of a β -galactosidase from *Bacillus circulans* to synthesise some β -D-(1 $\rightarrow$ 3) galactosyldisaccharides bearing a GlcNAc or a GalNAc residue at the reducing end. Some β -D-(1 $\rightarrow$ 6) linkages were formed as well, but to the best of our knowledge, it was the first example of a preparative scale β -D-galactosyl transfer that occurs preferentially at the O-3 position using a galactosyl hydrolase. There is considerable synthetic interest in the use of this galactosidase to perform this type of galactosyl transfer.

Recently, it has been shown that the use of green solvents,^{20,21} such as biomass derived solvents or ionic liquids, as co-solvents in the enzymatic reaction lead to a variation of the enzyme activity.^{22–29}

Biomass derived solvents show some new and attractive advantages, as there are made from a renewable source, they are (presumably) less toxic, exhibit low volatility, and possess tuneable physicochemical properties.^{20,21,30–32} We have previously described the use of some of these biomass derived

^aDepartment of Organic and Pharmaceutical Chemistry, Faculty of Pharmacy, Complutense University of Madrid, Campus de Moncloa, 28040 Madrid, Spain^bUnidad de Bioinformática. Centro de Biología Molecular "Severo Ochoa" (CBMSO), CSIC, Autónoma University of Madrid (UAM), C/Nicolás Cabrera 1, 28049, Madrid, Spain^cDepartment of Physics and Chemistry II, Faculty of Pharmacy, Complutense University of Madrid, Campus de Moncloa, 28040 Madrid, Spain.E-mail: mjhernai@ucm.es; Fax: (+) 34 9139 41822; Tel: (+) 34 91139 41821

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3ra40860d

solvents in the transglycosylation reaction catalyzed by Biolacta β -galactosidase, using *p*-nitrophenyl- β -D-galactopyranoside (*p*-NP- β -Gal) as donor and GlcNAc as acceptor, changing the regioselectivity of the enzyme from a mixture of β (1 $\rightarrow$ 4) and β (1 $\rightarrow$ 6) towards the synthesis of only β (1 $\rightarrow$ 6) oligosaccharides.²⁶

Thus, the aim of the present work was to produce a easily purified β -Gal-3 from *Bacillus circulans* ATCC 31382 and explore the possibility of combining the advantages of using of this enzyme, a biodegradable biocatalyst, and solvents from renewable sources to the efficient and regioselective synthesis of β -D-(1 $\rightarrow$ 3) galactooligosaccharides.

Results

Production of recombinant enzyme β -Gal-3 from *Bacillus circulans* ATCC 31382

Two recombinant β -Gal-3 enzymes were obtained, one cloned in pET28b+ carrying an N-terminal 6-histidine tag (β -Gal-3-NTag), and the other, cloned in pET22b+ carrying a C-terminal one (β -Gal-3-CTag). After purification (Fig. 1), the two enzymes were assayed to determine the influence of the HistTag on the hydrolytic activity. β -Gal-3-NTag showed an activity of 12.2 U mg^{-1} and β -Gal-3-CTag of 12.5 U mg^{-1} . Fujimoto *et al.* characterized the wild type enzyme, obtaining an activity of 5.13 U mg^{-1} .¹⁸ This lower activity indicates that the technology based on histidine tags achieves better enzyme purification than ion exchange and gel permeation techniques previously used. Although our recombinant enzymes showed equivalent activity, the different antibiotics added to the growing medium have an effect on the recombinant enzyme production. Larger amounts of enzyme are obtained in kanamycin medium than in ampicillin, 21 mg and 9 mg per liter of grown medium respectively. Because of that, the enzyme used in subsequent assays is β -Gal-3-NTag, cloned in pET28b+.

Fig. 1 DS-electrophoresis of the purification of β -Gal-3. Lane **A**, non-induced cell extract. Lane **B**, IPTG induced cell extract. Lane **C**, purified β -Gal-3-NTag. Lane **D**, β -Gal-3-CTag. Lane **E**, molecular weight standard.

Scheme 1 Synthesis of Gal- β -(1 $\rightarrow$ 3)-GlcNAc and Gal- β -(1 $\rightarrow$ 3)-GalNAc catalyzed by β -Gal-3-NTag.

Enzymatic synthesis in presence of green solvents with β -Gal-3 from *Bacillus circulans* ATCC 31382

The transglycosylation reaction of *p*-nitrophenyl- β -D-galactopyranoside (*p*-NP- β -Gal) as donor and *N*-acetyl glucosamine (GlcNAc) or *N*-acetyl galactosamine (GalNAc) as acceptors, catalyzed by β -Gal-3-NTag from *Bacillus circulans* ATCC 31382 will be used throughout this work (Scheme 1).

The regioselectivity showed by the enzyme was investigated when the reaction was carried out in 50 mM sodium phosphate buffer at pH 6 (Fig. 2 and 3). This reaction can afford Gal- β -(1 $\rightarrow$ 3)-GlcNAc as a major product (50.7% conversion) and 48.9% of hydrolysis product (Gal) when GlcNAc was used as acceptor, while with GalNAc the main product formed was Gal- β -(1 $\rightarrow$ 3)-GalNAc with similar yields (49.3%) and 48.4% of hydrolysis product.

Three different groups of solvents (Scheme 2) were screened in order to evaluate their effect over β -Gal-3-NTag activity during transglycosylation reactions: glycerol based solvents with cyclic structures (**S1**–**S3**), glycerol based solvents with open chain (**S4**–**S12**) and 3-*N,N*-dimethylamide based solvents (**S13**–**S15**). Main features of these solvents (density, Log *P* and solubility 2M) are summarized in Table 1.

Transglycosylation reactions catalyzed by the β -Gal-3-NTag were carried out following the general procedure described in the experimental section and monitored by HPLC. The concentration of green solvents was fixed to 2 M in the sodium phosphate buffered mixture. The results obtained are summarized in Fig. 2 (GlcNAc as acceptor) and Fig. 3 (GalNAc as acceptor).

The use of only buffer as reaction medium leads to a good conversion of *p*-NP- β -Gal but there is a considerable amount of hydrolytic product, galactose (48.9 and 48.4% respectively).

In presence of green solvents, the Gal- β -(1 $\rightarrow$ 3)-GlcNAc was formed as main product (Fig. 2), except for **S3**, **S5**, **S6** and **S14**. In fact, **S6** and **S14** give rise to a total loss of enzymatic activity when used as co-solvents in this reaction. However, solvents **S10** and **S13** lead to the best yields for the synthesis of Gal- β -

Fig. 2 Transglycosylation yields obtained with β -Gal-3-NTag using buffer 2 M solvents derived from biomass with *p*-NP-Gal as donor and GlcNAc as acceptor.

Fig. 3 Transglycosylation yields obtained with β -Gal-3-NTag using buffer-2 M solvents derived from biomass with *p*-NP-Gal as donor and GalNAc as acceptor.

(1 $\rightarrow$ 3)-GlcNAc (93% and 99%, respectively), with no appearance of hydrolysis product. This means not only an important increase of enzymatic activity in comparison to the enzyme's

natural behaviour in 50 mM phosphate buffer at pH 6, but also an improvement in the selectivity of the reaction (no hydrolysis product), with full regioselectivity (β -(1 $\rightarrow$ 3) isomer).

The synthetic activity of β -galactosidase was examined when GalNAc was used as acceptor. Results are shown in Fig. 3, and are similar of that obtained with GlcNAc. In the presence of solvents **S4**, **S9** and **S11**, Gal- β -(1 $\rightarrow$ 3)-GalNAc was formed as main product, but for **S3**, **S5**, **S6**, **S7** and **S14** no enzymatic activity was observed. Again, solvents **S10** and **S13** lead to the best yields for the synthesis of Gal- β -(1 $\rightarrow$ 3)-GlcNAc (86% and 95%, respectively), with no appearance of hydrolysis product.

Scheme 2 Structure of the different green solvents employed in transglycosylation reaction with β -Gal-3-NTag.

Table 1 Density, Log *P* and solubility 2 M of the different solvents

Solvent	Solvent density (g ml ⁻¹)	Log <i>P</i>	System composition
S1	1.24	-0.057	Monophasic
S2	1.41	-0.024	Monophasic
S3	1.06	0.030	Monophasic
S4	1.07	-0.60	Monophasic
S5	0.94	0.14	Monophasic
S6	0.91	0.27	Monophasic
S7	1.12	1.14	Biphasic
S8	1.36	1.42	Biphasic
S9	1.27	1.71	Biphasic
S10	0.90	1.93	Biphasic
S11	0.89	2.07	Biphasic
S12	0.89	2.48	Biphasic
S13	1.06	-0.69	Monophasic
S14	1.15	1.41	Monophasic
S15	0.91	1.42	Biphasic

Fig. 4 Transglycosylation yields (%) obtained with β -Gal-3-NTag at different **S13** concentrations (1M, 2M and 3M) with *p*-NP-Gal as donor and GlcNAc as acceptor in (a) and GalNAc in (b).

The use of **S13** showed yields over 95% in the synthesis of both disaccharides, therefore this solvent was selected for further assays.

There is not a clear relationship between solvent structures and the available physical parameters ($\log P$ and density) with the noticeable increase of synthetic activity (Table 1).

Effect of the co-solvent concentration on the transglycosylation reactions

In order to study the influence of co-solvent concentration on the yield boost, the reaction was carried out with different amounts of solvent **S13** (1, 2 and 3 M). Results obtained are shown in Fig. 4. As it can be seen, in a **S13** 2 M concentration, the best yields for (Gal- β -(1 $\rightarrow$ 3)-GlcNAc and Gal- β -(1 $\rightarrow$ 3)-GalNAc) were obtained (99% and 95% respectively). The same effect was previously observed in β -galactosidase from *Thermus thermophilus* and β -galactosidase from *Escherichia coli* when these biosolvents were used as co-solvents. It was attributed to a conformational change in the enzyme secondary and tertiary structures caused by the co-solvent.^{28,29}

Purification stage

The feasibility of the target molecule isolation from the reaction medium must be evaluated after the synthesis of an oligosaccharide. Fig. 5 shows a typical HPLC chromatogram of a sample that was withdrawn during an enzymatic synthesis of Gal- β -(1 $\rightarrow$ 3)-GlcNAc using **S13** as a cosolvent. After the enzymatic synthesis of Gal- β -(1 $\rightarrow$ 3)-GlcNAc in the presence of 2 M **S13**, a mixture of Gal- β -(1 $\rightarrow$ 3)-GlcNAc and GlcNAc (which is used overconcentrated) is obtained as observed in Fig. 5.

Solvent **S13** is water soluble and cannot be eliminated by lyophilisation, percolation, rotary evaporation or organic extraction. An alternative protocol was developed for green

Fig. 5 HPLC chromatogram of the reaction catalyzed by β -Gal-3-NTag using buffer-2 M **S13** with *p*-NP-Gal as donor and GlcNAc as acceptor. (a) Reaction debris; (b) GlcNAc; (c) Gal; (d) Gal- β -(1 $\rightarrow$ 3)-GlcNAc.

solvent elimination, based on direct hexane precipitation at -196°C without prior extraction. The pellets obtained (composed by the mixture of saccharides) were purified by filtration. **S13** was separated by rotary evaporation of the hexane, recovering the total amount of both solvents. After the purification process, **S13** can be reused in further reactions.

On the other hand, an advantage of solvents **S9** and **S10**, compared with other solvents used in this study, is that under these experimental conditions a biphasic mixture is created between these solvents and aqueous buffer (Table 1), then after reaction these solvents can be separated from reaction media by centrifugation. Moreover, carbohydrate compounds in the reaction media are not soluble in these solvents phase and remains in the aqueous phase. Centrifugation becomes a very useful tool for the isolation of these solvents from the reaction media, allowing its reuse in further reactions.

With these optimized separation procedures in hand, we proceeded to the systematic semi-preparative experiments in 80 ml of reaction media in presence of **S9**, **S10** and **S13**. In all cases the yield of pure disaccharide after purification through a carbon-celite column, compared with initial *p*-NP-Gal, was 85% (2,17 g of Gal- β -(1 $\rightarrow$ 3)-GlcNAc).

Solvents effects in tertiary structure by fluorescence study

To explore the effect of **S13** on the enzyme structure and its potential conformational changes, we measured the fluorescence variations at different **S13** concentrations. The enzyme was first dissolved in 10 mM sodium phosphate buffer (pH 6.00) to record a native emission spectrum, then some aliquots of solvent were added sequentially. As a result, a red shift was observed when the enzyme is placed in 2 M **S13** (concentration for optimal transglycosylation yield), the maximum fluorescence emission is shifted to 345.87 nm from 340.63 nm (Fig. 6). The spectra obtained at different solvent concentrations tested for transglycosylation reactions (Fig. 6) show a maximum peak at 344.76 nm for 1 M solvent concentration and a maximum peak on 346.59 nm for 3 M. These results suggest a modification of the tryptophan chemical environment which is the main fluorescence residue, which most likely involves a conformational modification. However, it is well known that red-shift of the emission wavelength is due to

Fig. 6 Normalized emission fluorescence spectra for β -Gal-3-Ntag at different concentrations of **S13**.

an increased exposure of tryptophan residues to a more polar environment, which is usually related with protein denaturation/enzyme deactivation.³³ Nevertheless, since CD spectrum of **S13** in the far UV range (200–250 nm) was too noisy to reveal any change of secondary structure of protein, we have studied the thermostability of the β -gal 1–3 in buffer and with **S13** 2 M by fluorescence spectroscopy (see Fig. S1 in ESI†). As it shows in this Figure the presence of **S13** at higher temperatures of 60 °C shows a sharper slope indicating that **S13** destabilizes the protein, however, in fact this change occurs at high temperature not at 37 °C, our running temperature. According to this result, the effect of **S13** on the enzyme's structure could be attributed to a more open conformation of the enzyme that would destabilize the protein.

Molecular modelling studies

In previous work, our group found that the change of regioselectivity cannot be easily related to the biomass-derived solvent's physical parameters (density, system composition, log *P*, etc.).²⁹ To get some insights on the origin of the activity boost observed in some of the solvents tested, a molecular dynamics simulation along with a fluorescence study were carried out, comparing pure water with the **S13**-water mixture, as solvating medium of the β -gal-3 from *Bacillus circulans* ATCC 31382.

Active site. Since no ligand or any other indication of the active site was available, we have performed multiple sequence alignment and structure superimposition with the β -galactosidase (lacZ) of *Escherichia coli*. Two glutamic residues appeared to be conserved across the compared sequences and corresponded roughly in the structure superimposition with the catalytic pair of glutamic acids in the lacZ. These two residues, Glu-157 and Glu-233 were located in the center of a well-defined cavity in the homology model and were selected as the catalytic pair. Glu-233 corresponds with the nucleophilic residue in lacZ while Glu-157 is aligned with the acid-base catalyst (Fig. 7).

Docking results. Most populated docking results for the second substrate in the glycosylated enzyme may confirm the role of the previously mentioned residues, Trp-235 and Tyr-450, in the active site substrate recognition and orientation to

Fig. 7 Modelling of the β -Gal-3-Ntag active site.

the catalytic glutamic acid residues pair (Fig. 7). The sugar core is located on top of the aromatic residue while the N-acetyl chain is involved in a hydrogen bond. These interactions allow the hydroxyl groups to be located within a distance of the glycosylated glutamic acid and the free one, and could allow the reaction to occur. Accommodation of the substrate in the active site may evolve very differently depending on the solvent and the flexibility conditions of the enzyme, being the reason to perform molecular dynamic simulations to obtain some insight into these.

Molecular dynamics. In previous studies, we have found some correlation between reactant groups distance and regioselectivity and/or reaction yields.²⁹ Measuring the distances between the glycosylated glutamic residue (Glu 233-galactose intermediate) and the hydroxyl group in the O3 position of GlcNAc in the presence of water as solvent and a mixture of **S13**-water, we found that shorter distances are translated in enhanced conversion (Fig. 8).

The solvent could play a key role in these changes modifying the interaction pattern between substrate and enzyme. Therefore, we measured the calculated non-bonded energy terms of substrate and protein to try to estimate the different contributions that could improve the binding energy and therefore, the overall conversion.

Fig. 8 Distances between the glycosylated glutamic residue and the hydroxyl group in the O3 position of GlcNAc in water (black) and a mixture of **S13**-water (blue).

Table 2 Average force field energy terms in molecular dynamics simulation

Solvent	Electrostatic term (kJ mol ⁻¹)	Van der Waals term (kJ mol ⁻¹)	Total MM energy (kJ mol ⁻¹)
Water	−13.27	−76.54	−89.81
S13-water	−14.64	−74.62	−89.27

As it can be seen in Table 2, the total calculated energy is almost equal in the mixture of **S13**-water. However, in electrostatic terms it is larger in mixed solvents than in water alone and by the Circe effect principle, which states that an increase of reaction speed in enzymatic environments can be mostly explained by a increase in the electrostatic interaction between enzyme and substrate, can provide a reasonable explanation of the enhanced conversion rate under mixed solvent conditions. Also, in the mixture case, the cost of desolvation is predicted to be less unfavourable since the substrate and the active site are both polar in nature and therefore interact strongly with the water molecules.

After this analysis we can conclude that the great regioselectivity changes during transglycosylation reactions observed in this study are caused by the interaction of the co-solvents with the enzyme, causing the modification of its secondary and tertiary structures (Fig. 6), also explaining the molecular modelling and docking results (Fig. 8).

Materials and methods

Materials and reagents

Solvents **S1**, **S2**, **S3**, **S4**, **S5**, **S6**, **S7**, **S8**, **S9**, **S10**, **S11** and **S12** were synthesized using the same procedures previously described.³⁴ **S13**, **S14** and **S15** were a gift from COGNIS IP Management GmbH (Germany).

Bovine serum albumin (BSA), *p*-nitrophenol (*p*-NP), *p*-nitrophenyl-β-D-galactopyranoside (*p*-NP-β-Gal), *N*-acetyl-D-glucosamine (GlcNAc), *N*-acetyl-D-galactosamine (GalNAc) and analytical standards of monosaccharides for HPLC were purchased from Sigma Aldrich. All other chemicals were of analytical grade.

Production of the enzyme and purification

Recombinant β-galactosidase from *Bacillus circulans* 31382 was cloned in *Escherichia coli* BL21 using pET28b+ and pET22b+ vectors (Novagen). *E. coli* cultures were grown aerobically at 37 °C in LB broth with kanamycin (30 mg L⁻¹) and induced with IPTG (isopropyl β-D-thiogalactopyranoside, 1 mM) at 37 °C for 5 h. Cells were disrupted by sonic disruption, unbroken cells and insoluble debris were eliminated by centrifugation (14.000 g for 15 min at 4 °C). The solution obtained was passed through a Ni²⁺-agarose column (3 mL) according to manufacturer's protocol (BioRad). Fractions were monitored by absorbance at 280 nm, pooled, and concentrated and desalted in an Amicon ultra centrifuge filter (Millipore). The purification process was followed by SDS-PAGE.³⁵ Protein quantifica-

tion was done by Bradford method,³⁶ using bovine serum albumin as standard.

Hydrolytic reactions

Hydrolytic activity was determined by spectrophotometrical quantification of *p*-NP liberated from the hydrolysis of *p*-NP-β-galactopyranoside 5 mM in sodium phosphate buffer 50 mM, pH 7 in a 300 μL cell by measuring the increase in absorbance at 410 nm during 3 min at 37 °C. Each experimental assay was run at least three times with standard deviation under 5% of the samples average. One enzyme unit (U) was defined as the amount (mg) of protein that hydrolyzes 1.0 μmol of substrate per minute.

Transglycosylation reactions

p-NP-β-Gal (85 mM) and GlcNAc or GalNAc (425 mM) were dissolved in 1.00 mL of buffer sodium phosphate 50 mM pH 6 and 2M of solvents from biomass as final concentration, and pre-warmed at reaction temperature (37 °C). Reaction started by addition of biocatalyst to the mixture: 5 U of recombinant enzyme. In order to check the effect of different concentrations of **S13**, reactions were run in 1.00 mL of buffer sodium phosphate 50 mM pH 6 and 1M, 2M and 3M of solvent as final concentration.

Aliquots (50 μL) were withdrawn from reaction media at different times. The reaction was stopped after 3 h by heating to 100 °C for 5 min and conserved immediately at −20 °C. Analytical determination of products were performed by HPLC using a NH2P50-4E amino column (Asahipak, Japan) using three detectors: ELSD (Evaporative Light Scattering), UV-Vis at 317 nm and CD (Circular Dichroism).

Purification of the reaction products in semi-preparative conditions

The scaled-up reaction mixture was composed by *p*-NP-β-Gal 85 mM and GlcNAc 425 mM dissolved in 80 mL of buffer sodium phosphate 50 mM pH 6 with **S13** 2M. The reaction was stopped after 3 h by heating to 100 °C for 5 min. After that, it was lyophilized with the aim of eliminating the water. The crude obtained (30 mL) was mixed with 30 mL of hexane and freeze at −196 °C using liquid nitrogen. The resulting pellet was recovered by filtration and dried. Solvent from biomass were separated from hexane by rotary evaporation, recovering the total amount of both solvents.

The pellet was directly loaded onto a carbon-celite column (50% m/m), eluted with a linear gradient from 0% to 15% (v/v) of ethanol in water. Solvents were eliminated and disaccharide stored in freeze. 30 mg were dissolved in D2O to be characterized by ¹H-NMR and ¹³C-NMR spectroscopy (D2O,

700 MHz–500 MHz). Spectra were consistent with previous references.^{15,17,26,37}

Gal-β(1→3)GalNAc (Galacto-*N*-Biose, GNB)¹⁵. ¹H-NMR (700 MHz, D₂O): 1.92 (s, Ac), 4.58 (d, J_{1β}, 2 = 8.47 Hz, H-1β), 5.11 (d, J_{1α}, 2 = 3.71 Hz, H-1α). ¹³C-NMR (700 MHz, D₂O): 21.96 (Me of Ac, α), 22.19 (Me of Ac, β), 48.92 (C-2α), 52.39 (C-2β), 60.89 (C-6α), 60.93 (C-6β), 61.11 (C-6'), 68.01 (C-4'), 68.51 (C-4α), 68.68 (C-4β), 70.13 (C-5α), 70.57 (C-2'), 72.48 (C-3'), 74.78 (C-5β), 74.90 (C-5'), 76.99 (C-3α), 80.01 (C-3β), 91.13 (C-1, α), 95.12 (C-1β), 104.64 (C-1'β), 104.81 (C-1'α), 174.61 (C=O of Ac, α), 174.91 (C=O of Ac, β).

Gal-β(1→3)GlcNAc (Lacto-*N*-Biose, LNB)^{15,17}. ¹H-NMR (500 MHz, D₂O): 1.96 (s, 3H, Ac), 5.11 (d, J_{1α}, 2 = 3.45 Hz, H-1α). ¹³C-NMR (500 MHz, D₂O): 22.39 (Me of Ac, α), 22.64 (Me of Ac, β), 53.28 (C-2α), 56.02 (C-2β), 60.98 (C-6), 61.39 (C-6'), 68.94 (C-4'), 69.10 (C-4), 71.12 (C-2'), 71.62 (C-5α), 72.95 (C-3'), 75.63 (C-5'), 75.85 (C-5β), 80.57 (C-3α), 83.01 (C-3β), 91.42 (C-1α), 95.11 (C-1β), 103.83 (C-1'β), 103.96 (C-1'α), 174.93 (C=O of Ac, α), 175.19 (C=O of Ac, β).

Gal-β(1→6)GlcNAc (*N*-Acetyl-allolactosamine)^{26,37}. ¹H-NMR (700 MHz, D₂O): 1.94 (s, Ac), 4.33 (d, H_{1'}, J_{1'}, 2' = 7.91 Hz), 4.61 (d, H_{1β}, J_{1β}, 2 = 8.47 Hz), 5.09 (d, J_{1α}, 2 = 3.58 Hz, H-1α). ¹³C-NMR: 21.81 (Me of Ac, α), 22.08 (Me of Ac, β), 53.95 (C-2 α), 56.52 (C-2β), 60.94 (C-6'), 68.49 (C-4'), 68.57 (C-6β), 68.59 (C-6α), 69.62 (C-3α), 69.77 (C-5α), 70.52 (C-4β), 70.56 (C-4α), 70.56 (C-2), 72.58 (C-3β), 72.61 (C-3'), 74.85 (C-5β), 75.11 (C-5'), 90.81 (C-1α), 94.93 (C-1β), 103.26 (C-1'β), 103.27 (C-1'α), 174.45 (C=O of Ac α), 174.70 (C=O of Ac β).

Solvents effects in tertiary structure by fluorescence study

Fluorescence emission spectra were recorded on a Perkin Elmer LS50B 295 nm was set as excitation wavelength because its high selectivity for tryptophan. Emission was acquired from 310 to 400 nm. The excitation and emission bandwidth were adjusted at 3 nm, measurements were done in 1 cm path length cell at 25 °C. First, the fluorescence spectrum of **S13** was recorded using a 2M solution of this solvent, then for β-Gal-3-NTag (0.1 mg mL⁻¹) in 10 mM sodium phosphate buffer at pH 6.00. The **S13** emission spectrum was negligible compared to the one of the enzyme in buffer solution. Then, several additions of **S13** were performed in the cuvette in order to evaluate the effect of this solvent in the tryptophan exposition of the enzyme. Thermal ramps of enzyme in buffer and at 2M **S13** were performed from 20 to 90 °C at a rate of and 60 °C h⁻¹. Spectra were acquired in a PTI modular spectrometer with thermoelectric temperature regulator TLC 50, Spectra analysis were done using Felix 32 software. (supplementary material)

Computational methods

Homology modelling

Due to the lack of a public structure with atomic resolution for the β-gal-3 β-galactosidase of *Bacillus circulans* ATCC 31382, we have generated a three dimensional homology model. With this aim, we have located the closest homologue structure deposited in the Protein Data Bank performing a Basic Local Alignment Search Tool (BLAST) search with the National

Center for Biotechnology Information (NCBI) web interface. The results (e-value: 1e-146) yielded a 42% of sequence identity with the β-galactosidase of *Bacterioides thetaiotaomicron* (PDB ID: 3D3A, resolution 2.15 Å) which was selected for performing the model. We have used two automatic modelling servers: SWISS-Model and CPHmodels-3.0 to build the model based on the alignment with the β-galactosidase of *Bacterioides thetaiotaomicron*. These two generated models were refined using the GROMOS 96 43a1 force field³⁸ to energetically minimize the structure using the 1000 steps of the Steepest descent algorithm. For validation and selection of the best model, we have employed an analysis of the Ramachandran plot, the ERRAT score and the Verify3D residue plots. After comparison, the model generated with CPH models was selected since it was the most compliant with the parameters found in naturally folded proteins for the validation criteria.

Active site search and docking

The active site of the enzyme was obtained from structure superimposition with *E. coli* β-galactosidase (PDB code 1DP0), sequence conservation analysis and blind docking, yielding a match of the catalytic pair Glu-157 and Glu-233 (nucleophilic residue). Bearing in mind the catalytic mechanism of the enzyme, and as we have done in other previous works,^{28,29} we followed a two steps modelling on the enzyme. First, we performed a docking simulation of *p*-nitro-phenol-β-galactopyranose in the active site using Autodock 4.2,³⁹ the default Lamarckian genetic algorithm parameters and a grid of 60 × 60 × 60 points (spacing 0.375 Å). The best pose according to both energetic (favourable docking energy) and geometrical criteria (close enough to Glu-233 to allow any reaction) was selected to build a covalent complex model using the galactopyranose moiety and the Glu-233 residue. Then, a second docking simulation, using the glycosylated enzyme, was carried out with the GlcNAc, selecting the most favourable pose according to its predicted docking energy. This complex was used as the starting point for molecular dynamics studies.

Solvent parameters

The solvent used in the simulations was already parameterized for the GROMOS 96 43a1 force field,⁴⁰ briefly: we first optimized the solvent molecules using the Density Functional Theory (DFT) at the B3LYP 6-31G** level, then the bond lengths and angles obtained from this calculation were used and adapted to the force field to match both experimental density and the enthalpy of vaporization.

Molecular dynamics

The final complex obtained by docking the two molecules implicated in the catalytic mechanism of the enzyme was used as the initial model for molecular dynamics simulation.

The protein was parameterized using the GROMOS 96 43a1 parameters, including galactose-glutamic parameters already used in other works.²⁸ GlcNAc parameters were generated by the Dundee PRODRG 2.5 Server, carefully checking the generated topology as pointed out by Lemkul *et al.*⁴¹

For both, water and mixed solvent systems, we first solvated the complex with the selected solvent and then we performed the following protocol: a energy minimization with 2000 steps

of SD algorithm and 7000 steps of Polak–Ribiere Conjugate gradients (CG) was performed; a equilibration of 300 ps at 300 K (NVT), using the Particle-Mesh-Ewald (PME) to deal with long range electrostatics, the v-rescale thermostat and restrained positions in protein and ligand heavy atoms, followed by 300 ps at 300 K and 1 atm (NPT) using the Parrinello–Rahman barostat with the same conditions; and finally, a 2 ns production simulation recording energy and coordinates each 1 ps and a step of 2 fs.⁴⁰ All simulations were performed using the GROMACS v4.0.7 suite.³⁸

Conclusions

An efficient strategy for the synthesis of Gal- β -D-(1 $\rightarrow$ 3)GlcNAc and Gal- β -D-(1 $\rightarrow$ 3)GalNAc, combining renewable raw materials and recycled solvents (derived from biomass), buffer solutions and a biodegradable biocatalyst has been explored. The use of biosolvents allows good yields in the enzymatic synthesis of β -D-(1 $\rightarrow$ 3) linkages, preventing hydrolysis and with full regioselectivity. Yields of up to 99% and 95% of Gal- β -D-(1 $\rightarrow$ 3)GlcNAc and Gal- β -D-(1 $\rightarrow$ 3)GalNAc were obtained with total substrate conversion. These reactions represent a considerable improvement over the use of aqueous buffer or conventional organic solvents. Optimization of the reaction conditions allows the reaction's scale up to 80 ml with further solvent recycling and reuse. This recovery strategy could, in principle, be extrapolated to other enzymatic reactions carried out in these kind of green solvents.

In order to understand the possible effect of these green solvents on the synthetic behaviour of the enzyme, we performed different structural studies (fluorescence and molecular modelling) in the presence of selected biosolvents. We are able to conclude that the presence of green co-solvents in the reaction media modifies the enzyme tertiary structure, permitting a better disposition of the substrates in the active site, due to solvation effects, that could explain the behaviour observed.

Notes and references

- 1 P. K. Mandal and W. B. Turnbull, *Carbohydr. Res.*, 2010, **346**, 2113–2120.
- 2 G. F. Springer, *Science*, 1984, **224**, 1198–1206.
- 3 A. Holeman and P. H. Seeberger, *Curr. Opin. Biotechnol.*, 2004, **15**, 615–622.
- 4 E. Rapoport, S. Khaidukov, O. Baidina, V. Bojenko, E. Moiseeva, G. Pasynina, U. Karsten, N. Nifant'ev, J. LePendu and N. Bovin, *Eur. J. Cell Biol.*, 2003, **82**, 295–302.
- 5 H. Shirato, S. Ogawa, H. Ito, T. Sato, A. Kameyama, H. Narimatsu, Z. Xiaofan, T. Miyamura, T. Wakita, K. Ishii and N. Takeda, *J. Virol.*, 2008, **82**, 10756–10767.
- 6 M. E. C. Caines, H. Zhu, M. Vuckovic, L. M. Willis, S. G. Withers, W. W. Wakarchuk and N. C. J. Strynadka, *J. Biol. Chem.*, 2008, **283**, 31279–31283.
- 7 A. Takada, K. Ohmori, T. Yoneda, K. Tsuyuoka, A. Hasegawa, M. Kiso and R. Kannagi, *Cancer Res.*, 1993, **53**, 354–361.
- 8 R. U. Lemieux and H. Driguez, *J. Am. Chem. Soc.*, 1975, **97**, 4063–4069.
- 9 D. P. Khare, O. Hindsgaul and R. U. Lemieux, *Carbohydr. Res.*, 1985, **136**, 285–308.
- 10 M. Wilstermann and G. Magnusson, *Carbohydr. Res.*, 1995, **272**, 1–7.
- 11 R. R. Schmidt and E. Rücker, *Tetrahedron Lett.*, 1980, **21**, 1421–1424.
- 12 R. R. Schmidt, *Angew. Chem., Int. Ed. Engl.*, 1986, **25**, 212–235.
- 13 H. Paulsen, *Angew. Chem., Int. Ed. Engl.*, 1982, **21**, 155–173.
- 14 G. Baisch, R. Ohrlein, M. Streiff and F. Kolbinger, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 751–754.
- 15 H. Yu, V. Thon, K. Lau, L. Cai, Y. Chen, S. Mu, Y. Li, P. G. Wang and X. Chen, *Chem. Commun.*, 2010, **46**, 7507–7509.
- 16 A. D'Almeida, M. Ionata, V. Tran, C. Tellier, M. Dion and C. Rabiller, *Tetrahedron: Asymmetry*, 2009, **20**, 1243–1246.
- 17 A. Vetere, M. Miletich, M. Bosco and S. Paoletti, *Eur. J. Biochem.*, 2000, **267**, 942–949.
- 18 H. Fujimoto, M. Miyasato, Y. Ito, T. Sasaki and K. Ajisaka, *Glycoconjugate J.*, 1998, **15**, 155–160.
- 19 Y. Ito and T. Sasaki, *Biosci., Biotechnol., Biochem.*, 1997, **61**, 1270–1276.
- 20 M. J. Hernaiz, A. R. Alcántara, J. I. García and J. V. Sinisterra, *Chem.–Eur. J.*, 2010, **16**, 9422–9437.
- 21 M. Perez, J. V. Sinisterra and M. J. Hernaiz, *Curr. Org. Chem.*, 2010, **14**, 2366–2383.
- 22 S. V. Kamat, E. J. Beckman and A. J. Russell, *J. Am. Chem. Soc.*, 1993, **115**, 8845–8846.
- 23 K. Ryu and J. S. Dordick, *Biochem.*, 1992, **31**.
- 24 S. Tawaki and A. M. Klibanov, *J. Am. Chem. Soc.*, 1992, **114**, 1882–1884.
- 25 C. R. Wescott and A. M. Klibanov, *J. Am. Chem. Soc.*, 1993, **115**, 1629–1631.
- 26 M. Pérez-Sánchez, M. Sandoval, A. Cortés-Cabrera, H. García-Marín, J. V. Sinisterra, J. I. García and M. J. Hernaiz, *Green Chem.*, 2011, **13**, 2810–2817.
- 27 M. Perez-Sanchez, M. Sandoval and M. J. Hernaiz, *Tetrahedron*, 2012, **68**, 2141–2145.
- 28 M. Sandoval, A. Cortes, C. Civera, J. Trevino, E. Ferreras, M. Vaultier, J. Berenguer, P. Lozano and M. J. Hernaiz, *RSC Adv.*, 2012, **2**, 6306–6314.
- 29 M. Pérez-Sánchez, Á. Cortés-Cabrera, H. García-Martín, J. V. Sinisterra, J. I. García and M. J. Hernaiz, *Tetrahedron*, 2011, **67**, 7708–7712.
- 30 A. Wolfson, C. Dlugy and Y. Shotland, *Environ. Chem. Lett.*, 2007, **5**, 67–71.
- 31 A. Wolfson, C. Dlugy, D. Tavor, J. Blumenfeld and Y. Shotland, *Tetrahedron: Asymmetry*, 2006, **17**, 2043–2045.
- 32 L. H. Andrade, L. Piován and M. D. Pasquini, *Tetrahedron: Asymmetry*, 2009, **20**, 1521–1525.
- 33 J. T. Vivian and P. R. Callis, *Biophys. J.*, 2001, **80**, 2093–2109.
- 34 J. I. García, H. García-Marín, J. A. Mayoral and P. Pérez, *Green Chem.*, 2010, **12**, 426–434.
- 35 U. K. Laemmli, *Nature*, 1970, **227**, 680–685.
- 36 M. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248–254.

- 37 N. Bridiau and T. Maugard, *Biotechnol. Prog.*, 2011, **27**, 386–394.
- 38 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
- 39 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.
- 40 A. Amadei, A. B. M. Linssen and H. J. C. Berendsen, *Proteins: Struct., Funct., Genet.*, 1993, **17**, 412–425.
- 41 J. A. Lemkul, W. J. Allen and D. R. Bevan, *J. Chem. Inf. Model.*, 2010, **50**, 2221–2235.