Tetrahedron 69 (2013) 4973–4978

Contents lists available at SciVerse ScienceDirect

Tetrahedron

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

Highly efficient enzymatic synthesis of Galβ-(1→3)-GalNAc and Galβ-(1→3)-GlcNAc in ionic liquids

Carlos Bayón [a], Álvaro Cortés [b], José Berenguer [b], María J. Hernáiz [a,*]

[a] Department of Organic and Pharmaceutical Chemistry, Faculty of Pharmacy, Complutense University of Madrid, Campus de Moncloa, 28040 Madrid, Spain

[b] Centro de Biología Molecular "Severo Ochoa" (CBMSO), CSIC, Autónoma University of Madrid (UAM), C/ Nicolás Cabrera 1, 28049 Madrid, Spain

ARTICLE INFO

Article history:
Received 7 February 2013
Received in revised form 24 March 2013
Accepted 4 April 2013
Available online 9 April 2013

Keywords:
Enzyme catalysis
Glycosylation
Ionic liquids
Oligosaccharides
Regioselectivity

ABSTRACT

Ionic liquids (ILs) have emerged as an alternative to conventional organic media due to their high thermal and chemical stability, negligible vapour pressure, non-flammability and easy recycling. In this context, this work assesses the catalytic activity of a β-galactosidase from *Bacillus circulans* ATCC 31382 (β-Gal-3-NTag) in the synthesis of β-(1→3)-galactosides using different ILs. A noticeably increase in activity, retaining total regioselectivity was found in the synthetic behaviour of *B. circulans* β-galactosidase in ILs as co-solvents and using a 1:5 molar ratio of donor (*p*NP-β-Gal):acceptor (GlcNAc or GalNac). Yields up to 97% of β-(1→3) with different ILs were found. These reactions take place without noticeable hydrolytic activity and with total regioselectivity, representing a considerable improvement over the use of aqueous buffer or conventional organic solvents. Furthermore, reaction scaling up and IL recovery and recycling are feasible without losing catalytic action. Molecular modelling studies performed predict a three-dimensional interaction at the active centre between the acceptor and the water–IL mixture, which could explain the results obtained.

1. Introduction

Oligosaccharides present on the cell surface are involved in many biological processes, such as cell–cell recognition and communication, growth regulation and antibody interaction, bacterial and viral infection and other crucial intercellular recognition events.[1–4] Specifically, β-(1→3) linked galactosides, such as galacto-*N*-biose (GNB, Gal-β-(1→3)-GalNAc) are key carbohydrate structures that integrated in the Tn antigen that is associated to carcinoma cells or lacto-*N*-biose (LNB, Gal-β-1→3-GlcNAc), the critical disaccharide of Lewis antigen (sLea), which has been reported to be a ligand in metastasis of cancer cells. GNB and LNB are also important constituents of mucin type or complex type glycoproteins.[5–11]

There is consequently much interest in the chemical or enzymatic synthesis of GNB and LNB. For that reason related oligosaccharides containing this linkage have been synthesized by chemical[12–17] and enzymatic approaches using β-glycosidases[18,19] or β-(1→3)-galactosyltransferases.[19,20] Chemical synthesis of oligosaccharides is time-consuming, requires numerous steps and use expensive and toxic reagents. The use of β-(1→3)-galactosyltransferases as catalysts has proven to be efficient, but its application is somehow difficult because of the limited access to a sufficient amount of enzymes and the expensive sugar nucleotide required. Recently, an efficient process has been developed using D-galactosyl-β-1→3-*N*-acetyl-D-hexosamine phosphorylase from *Bifidobacterium infantis* coupled with a galactokinase from *Escherichia coli*, but with some drawbacks, such as the use of two enzymes and long reaction times.[21] From a practical viewpoint, the use of glycosidases is attractive for oligosaccharide synthesis because it offers great control over chemo-, regio- and stereoselectivity, though yields are still low.[22–25] Until now only two β-galactosidases are known to transgalactosylate a galactose residue with β-1→3-linkage. One is obtained purifying bovine testes,[26–28] with the drawback that the source of the enzyme is an animal organ and therefore it is not appropriate for a large-scale synthesis. The other is a β-galactosidase obtained by recombinant technology using DNA from *Bacillus circulans* ATCC 31382.[29] This enzyme recognizes GlcNAc and GalNAc as acceptors and is able to use them to synthesize β-(1→3)-galactooligosaccharides with high regioselectivity but a low yield.[30]

At the same time, ionic liquids (ILs) have received increasing attention as new co-solvents for chemo- and biocatalytic organic synthesis in general and in carbohydrate chemistry in particular.[31,32] The use of them as co-solvents in enzymatic

* Corresponding author. Tel: +34 913941821; fax: +34 913841822; e-mail address: mjhernai@farm.ucm.es (M.J. Hernáiz).

0040-4020/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tet.2013.04.015

reactions, has already been applied to lipases,[33] proteases,[34] glycosidases,[35] oxidoreductases,[36] peroxidases[37] and whole cells.[38]

ILs are constitutively salts, composed entirely of ions and have remarkable special properties. They are liquids at room temperature,[39] thermally stable,[40] highly polar,[41] inert,[42] non-flammable,[43] present an almost inexistent vapour pressure,[44] and have good ionic conductivity,[45] low melting temperature[46] and wide electrochemical windows.[47]

In this work, we propose a highly efficient and regioselective enzymatic process for synthesizing β-(1→3)-galactosides using β-Gal-3-NTag as biocatalyst in a reaction system based on ILs.

2. Results and discussion

2.1. Hydrolytic activity of recombinant enzymes β-Gal-3 from *B. 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 (see figure in Supplementary data), the two enzymes were assayed in order 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 an hydrolytic activity of 12.5 U mg^{-1}. Fujimoto et al. have already characterized the wild type obtaining an activity of 5.13 U mg^{-1}.[30] This lower activity indicates that the technology based in histidine tags achieves a better enzyme purification than the ion exchange and gel permeation previously used. Although our recombinant enzymes showed an equivalent activity, the different antibiotics added to the medium affected to the amount of enzyme obtained. This, in kanamycin medium, is larger than in ampicillin, 21 mg and 9 mg per litre of grown medium, respectively. Therefore, the enzyme used in subsequent assays was β-Gal-3-NTag, cloned in pET28b+.

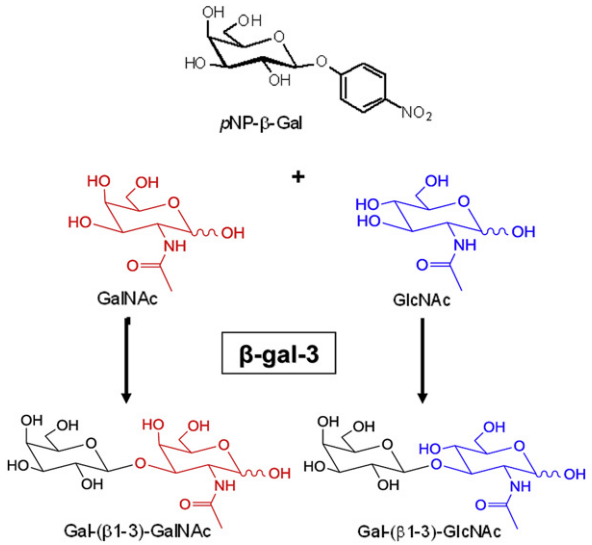
2.2. Enzymatic synthesis in the presence of ILs with β-Gal-3-NTag from *B. circulans* ATCC 31382

Glycosidases can 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.[22–25] The regioselectivity of enzymatic transgalactosylation depends on the source of the galactosidase used. In this respect, β-Gal-3 from *B. circulans* ATCC 31382 has proven to be a valuable biocatalyst for galactosyl transfer from suitable donor (*p*NP-gal) to an *N*-acetyl glucosamine (GlcNAc) or *N*-acetyl galactosamine (GalNAc) as acceptors. The synthesis of Gal-β-(1→3)-GlcNAc and Gal-β-(1→3)-GalNAc employing β-Gal-3 from *B. circulans* have been previously reported (Scheme 1).[30] This reaction can afford Gal-β-(1→3)-GlcNAc as a major product and Gal-β-(1→6)-GlcNAc as a minor product in some cases.

The transgalactosylation reaction used throughout this work was the 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 *B. circulans* ATCC 31382.

First, enzyme regioselectivity was investigated in this work if the reaction was carried out in 50 mM sodium phosphate buffer at pH 6 (Tables 1 and 2). This reaction could afford Gal-β-(1→3)-GlcNAc (LNB) as a major product (51% conversion) and 49% of hydrolytic product (Gal) when GlcNAc was used as acceptor, while with GalNac the main product formed was Gal-β-(1→3)-GalNAc (GNB) with similar yields, 49% and 51% of hydrolysis.

In a recent work we have shown that the use of ILs as co-solvents in the reaction medium allows to direct the regioselectivity of the transgalactosylation reactions in order to obtain



Scheme 1. Synthesis of Gal-β-(1→3)-GlcNAc and Gal-β-(1→3)-GalNAc catalyzed β-Gal-3-NTag.

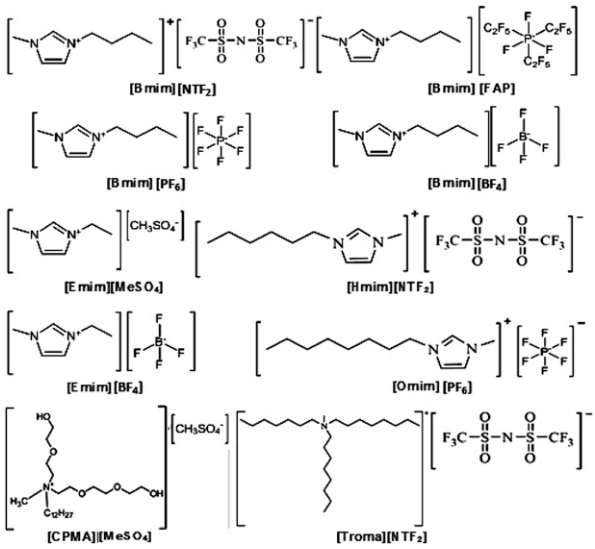
Table 1
Transglycosylation yields obtained with β-Gal-3-NTag using buffer-30% IL with *p*NP-Gal as donor and GlcNAc as acceptor

Solvent	Acceptor	Transgalactosylation		Hydrolysis
		β-(1-3) (%)	β-(1-6) (%)	(Galactose) (%)
Buffer	GlcNAc	51	0	49
[Bmim][Bf_4]	GlcNAc	40	14	30
[Bmim][FAP]	GlcNAc	4	0	0
[Bmim][NTf_2]	GlcNAc	51	27	22
[Bmim][PF_6]	GlcNAc	**97**	0	3
[CPMA][$MeSO_4$]	GlcNAc	53	33	15
[Emim][BF_4]	GlcNAc	37	26	6
[Emim][$MeSO_4$]	GlcNAc	20	0	80
[Hmim][NTf_2]	GlcNAc	37	19	44
[Omim][PF_6]	GlcNAc	43	16	41
[Troma][NTf_2]	GlcNAc	88	0	12

only the β(1→4) regioisomer when *Thermus thermophilus* β-galactosidase is employed.[35] In this work, we have extended the use of these ILs to the study of oligosaccharide synthesis catalyzed by the cloned β-Gal-3-NTag. Several ILs have been included in this study: [Bmim][BF_4], [Bmim][FAP], [Bmim][NTf_2], [Bmim][PF_6], [CPMA][$MeSO_4$], [Emim][BF_4], [Omim][PF_6], [Troma][NTf_2], [Hmim][NTf_2], [Emim][$MeSO_4$] (Scheme 2).

Table 2
Transglycosylation yields obtained with β-Gal-3-NTag using buffer-30% IL with *p*NP-Gal as donor and GalNAc as acceptor

Solvent	Acceptor	Transgalactosylation		Hydrolysis
		β-(1-3) (%)	β-(1-6) (%)	(Galactose) (%)
Buffer	GalNAc	49	0	51
[Bmim][Bf_4]	GalNAc	55	5	28
[Bmim][FAP]	GalNAc	5	0	0
[Bmim][NTf_2]	GalNAc	50	8	42
[Bmim][PF_6]	GalNAc	**98**	0	2
[CPMA][$MeSO_4$]	GalNAc	32	23	45
[Emim][BF_4]	GalNAc	59	10	7
[Emim][$MeSO_4$]	GalNAc	22	0	78
[Hmim][NTf_2]	GalNAc	61	6	33
[Omim][PF_6]	GalNAc	64	4	32
[Troma][NTf_2]	GalNAc	91	0	9



Scheme 2. Structure of the different ILs employed in transgalactosylation reaction with β-Gal-3-NTag.

Transgalactosylation reaction catalyzed by the β-Gal-3-NTag was carried out following the general procedure described in the Experimental section and monitored by HPLC. The percentage of IL was fixed to 30% in sodium phosphate buffer mixture, which is the percentage with which we obtained best yields in previous ILs work.[35] Results obtained for the transgalactosylation reaction with β-Gal-3-NTag in presence of different ILs are summarized in Table 1 (GlcNAc as acceptor) and Table 2 (GalNAc as acceptor).

The use of only buffer reaction medium leads to a good conversion of *p*NP-β-Gal but there is a considerable amount of hydrolytic product, galactose (49 and 51%, respectively).

In the presence of ILs, Gal-β-(1→3)-GlcNAc was formed as the main product and the best yields were achieved with [Bmim][PF_6] (97%) and [Troma][NTf_2] (88%), no other products were observed. [Bmim][Bf_4], [Bmim][NTf2], [CPMA][$MeSO_4$], [Emim][BF_4], [Hmim][NTf_2] and [Omim][PF_6] show a variation of the transgalactosylation regioselectivity, leading to the appearance of Gal-β-(1→6)-GlcNAc as minor product. [Emim][$MeSO_4$] causes the hydrolysis of the *p*NP-β-Gal against the transgalactosylation that shows a yield of 20%. Finally, the use of [Bmim][FAP] results in a loss of enzymatic activity. The result with [Bmim][PF_6] means an important increase of yield compared to the natural behaviour of this enzyme in 50 mM sodium phosphate buffer at pH 6, retaining β-(1→3) regioselectivity (Table 1).

The synthetic activity of the β-galactosidase was also examined when GalNAc was used as acceptor and the results, similar of that obtained with GlcNAc, are shown in Table 2. In the presence of ILs the Gal-β-(1→3)-GalNAc was formed as the main product, the best yields obtained once more in the presence of [Bmim][PF_6] (97%). Therefore, this solvent was selected for further assays.

In order to evaluate the influence of the co-solvent concentration on the yield enhancement, the reaction was carried out with different amounts of [Bmim][PF_6] (15, 30 and 45%) using *p*NP-Gal as donor and GlcNAc and GalNAc as acceptors. Percentages of IL over 45% resulted in media excessively dense to run reactions. Results obtained are shown in Table 3.

As can be seen, the best yield in disaccharide Gal-β-(1→3)-GlcNAc, was obtained with a 30% of IL (97% yield) with almost total substrate conversion (Table 3). The same behaviour was observed in the synthesis of Gal-β-(1→3)-GalNAc with a yield of 98%, indicating the need to work at the adequate co-solvent concentrations in order to obtain optimal effects over the reaction. Similar results over regioselectivity have been reported by our group in the enzymatic synthesis of Gal-β-(1→4)-GlcNAc using a β-galactosidase from *T. thermophilus* HB27.[35]

Table 3
Transgalactosylation yields (%) obtained with β-Gal-3-NTag at different [Bmim][PF6] concentrations (15, 30 and 45%)

Solvent	Acceptor	Transgalactosylation	Hydrolysis
		β-(1-3) (%)	(Gal) (%)
[Bmim][PF_6] 15%	GlcNAc	66	34
[Bmim][PF_6] 30%	GlcNAc	97	3
[Bmim][PF_6] 45%	GlcNAc	78	22
[Bmim][PF_6] 15%	GalNAc	68	32
[Bmim][PF_6] 30%	GalNAc	98	2
[Bmim][PF_6] 45%	GalNAc	80	20

2.3. Semi-preparative synthesis

An important point that must be evaluated after the synthesis of an important disaccharide is the feasibility of isolation of the target molecule from the reaction mixture. In this respect, an advantage of [Bmim][PF_6] compared with other solvents, such as solvents derived from biomass,[48,49] is that under stirring conditions an emulsion is created between IL and aqueous buffer, but after reaction this IL can be separated from reaction media by centrifugation. Moreover, carbohydrate compounds in the reaction media are not soluble in the IL phase and remains in the aqueous phase. Centrifugation becomes a very useful tool for the isolation of the IL from the reaction media, allowing its reuse in further reactions.

With this optimized separation procedure in hand, we proceeded to the systematic semi-preparative experiments in 80 ml of reaction media in presence of [Bmim][PF_6]. This scale up of the reaction used carried our using *p*NP-β-Gal as donor and GlcNAc and GalNAc as acceptors.

The yield of pure disaccharides after purification through a carbon–Celite column, compared with initial *p*NP-Gal, were; 88% for GlcNAc (2.25 g of Gal-β-(1→3)-GlcNAc) and 90% for GalNAc (2.30 g of Gal-β-(1→3)-GlalNAc).

2.4. Molecular modelling studies

In order to explain the increase in activity that might occur due to the interaction between the GlcNAc and the buffer–IL mixture at the active site of the enzyme, we have used molecular modelling study. With this aim, the active site of the enzyme was identified as the only cavity available in the structure with two glutamic acid residues (Glu-157 and Glu-233) (Fig. 1). Visual inspection of the cavity

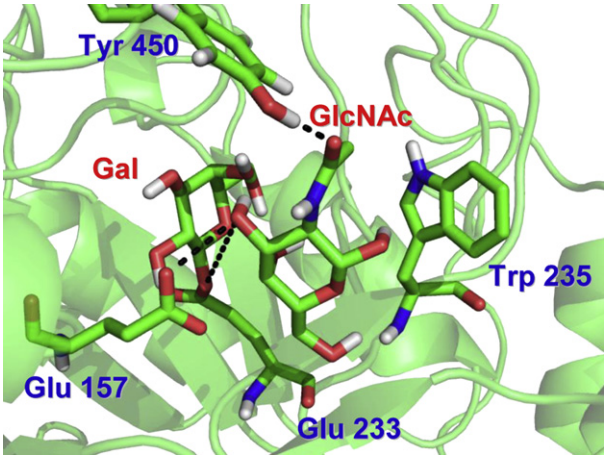


Fig. 1. Modelling of the β-Gal-3 active site.

indicated that other residues, commonly found in other β-galactosidases surrounding the catalytic pair, could be involved in the molecular recognition of the substrates. These are Trp-235, which is located under the candidate sugar-core binding site, and aromatic residues, such as Trp and Tyr that have been described in the literature to act on these enzymes as key residues to allow the binding of sugars probably due to π–π stacking interactions. Tyr-450 is also located near the catalytic pair and could be also involved in hydrogen bonds with the hydroxyl groups of the substrates.

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 the catalytic glutamic acid residues pair (Fig. 1). 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.

In previous studies we have found some correlation between reactant groups distance and regioselectivity and/or reaction yield.[48] Measuring the distances between the glycosylated glutamic residue (Glu 233-galactose intermediate) and the hydroxyl group in the O3 position of GlcNAc in presence of water as solvent and a mixture of [Bmim][PF_6]–water, we found that shorter distances are translated in enhanced conversion (Fig. 2).

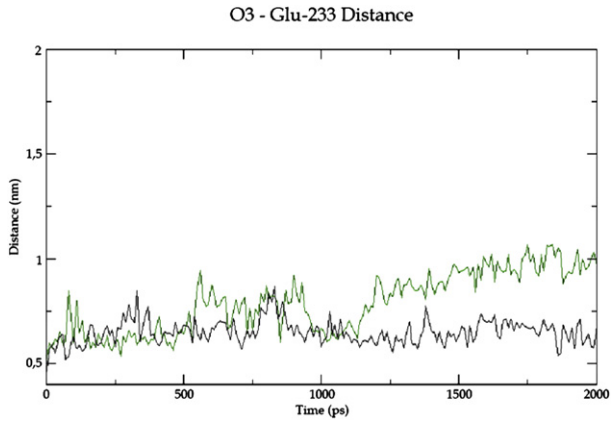


Fig. 2. Distances between the galactosylated glutamic residue and the hydroxyl group in the O3 position of glcNAc in water (green) and a mixture of [BMIM][PF_6]–water (black).

The solvent could play a key role in these changes modifying the interaction pattern between substrate and enzyme. Therefore, we have 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.

As can be seen in Table 4, the total calculated energy is larger for the case of [Bmim][PF_6]–water. Electrostatics terms are 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 an 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 case of mixtures, 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.

Table 4
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
[Bmim][PF_6]–water	−19.13	−74.83	−93.96

3. Conclusions

The application of ILs to the enzymatic synthesis of β-1→3-galactosides using a β-Gal-3-NTag was assessed. The results clearly showed an important increase in enzymatic activity retaining regioselectivity. Yields of up to 99% of Gal-β-(1→3)-GlcNac or Gal-β-(1→3)-GlcNac were obtained with full substrate conversion. These reactions take place without noticeable hydrolytic activity and keeping total regioselectivity, representing a considerable improvement over the use of aqueous buffers or conventional organic solvents. These results reinforce the concept of driving enzymatic syntheses to the desired product simply by adjusting the reaction medium with small amounts of ILs. By using a three dimensional structural simulation of the β-Gal-3 enzyme we have developed a model for the three-dimensional arrangement between GlcNAc and the water–IL mixture in the active centre that explains the enzyme's regioselectivity towards the synthesis of β(1→3) linkages as observed experimentally.

4. Experimental section

4.1. Materials and reagents

1-Butyl-3-methylimidazolium tetrafluoroborate ([Bmim][BF_4]), 1-butyl-3-methylimidazolium tris (pentafluoroethyl) trifluoro-phosphate ([Bmim][FAP]), 1-butyl-3-methylimidazolium bis-(trifluoromethyl sulfonyl)-imide ([Bmim][NTf_2]), 1-butyl-3-methylimidazolium hexafluorophosphate ([Bmim][PF_6]), 1-ethyl-3-methylimidazolium methylsulfate ([Emim][$MeSO_4$]) 1-hexyl-3-methylimidazolium bis(trifluoromethyl sulfonyl)-imide ([Hmim][NTf_2]), 99% purity were obtained from Merck. Cocosalkyl pentaethoxy methyl ammonium methylsulfate ([CPMA][$MeSO_4$]), 1-ethyl-3-methylimidazolium tetrafluoroborate ([Emim][BF_4]), 1-methyl-3-octylimidazolium hexafluorophosphate ([Omim][PF_6]) and methyltrioctylammonium bis(trifluoromethyl sulfonyl)-imide, ([Troma][NTf_2]), were obtained from Solvent Innovation 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.

4.2. Production of the enzyme and purification

Recombinant β-Gal-3-NTag was cloned in *E. coli* BL21 using pET28b+ and pET22b+ vectors (Novagen). *E. coli* cultures were grown aerobically at 37 °C in LB broth with antibiotics (kanamycin (30 mg l^{-1}) in the case of pET28b+ and ampicillin (50 mg l^{-1}) in the case of pET22b+) 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 Ni^{2+}–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.[50] Protein quantification was done by Bradford method,[51] using bovine serum albumin as standard.

4.3. Hydrolytic reactions

Hydrolytic activity was determined by spectrophotometrical quantification of *p*NP liberated by 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 determined at least three times with standard deviation under 5% of the average of samples. One enzyme unit (U), was defined as the amount (mg) of protein that hydrolyzes 1.0 μmol of substrate per minute.

4.4. Transgalactosylation 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 different percentages of ILs, and pre-warmed at reaction temperature (37 °C). Reaction started by addition of biocatalyst to the mixture: 5 U of recombinant enzyme. 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 an NH2P50-4E amino column (Asahipak, Japan) using three detectors: ELSD (Evaporative Light Scattering), UV−vis at 317 nm and CD (Circular Dichroism). Each experimental assay was determined at least three times with standard deviation under 5% of the average of samples.

4.5. 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 (or GalNAc 425 mM) dissolved in 80 ml buffer sodium phosphate 50 mM pH 6 with [Bmim][PF_6] (30%). The reaction was stopped after 3 h by heating to 100 °C for 5 min. The crude was centrifuged at 14,000 rpm with the aim of separating the aqueous phase (containing carbohydrate compounds) from the IL. After that, aqueous phase was lyophilized to eliminate the water. Powder was loaded on activated carbon and Celite® column (50% m/m), products were eluted with milliQ water in linear gradient of ethanol (from 0% v/v to 15% v/v). Disaccharide enriched fractions were collected in 10% or 15% ethanol; samples were pooled, solvent was removed and then lyophilized, purity of the solid powder was analyzed by HPLC. The yield of pure disaccharides after purification through the carbon−Celite column, compared with initial *p*NP-Gal, were; 88% for GlcNAc (2.25 g of Gal-β-(1→3)-GlcNAc) and 90% for GalNAc (2.25 g of Gal-β-(1→3)-GlalNAc).

Structural determination was done by ^{1}H NMR and ^{13}C NMR (D_2O, 500 MHz). Spectra were consistent with previous references.[18,21,49,52]

4.5.1. Gal-β(1→3)GalNAc (galacto-N-biose).[21] ^{1}H NMR (500 MHz, D_2O): δ ppm 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). ^{13}C NMR (700 MHz, D_2O): 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).

4.5.2. Gal-β(1→3)GlcNAc (lacto-N-biose).[18,21] ^{1}H NMR (500 MHz, D_2O): δ ppm 1.96 (s, 3H, Ac), 5.11 (d, *J*1α,2=3.45 Hz, H-1α). ^{13}C NMR (500 MHz, D_2O): 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).

4.5.3. Gal-β(1→6)GlcNAc (N-acetyl-allolactosamine).[49,52] ^{1}H NMR (500 MHz, D_2O): 1.94 (s, Ac), 4.33 (d, H1′, *J*1′,2′=7.91 Hz), 4.61 (d, H1β, *J*1β,2=8.47 Hz) 5.09 (d, *J*1α,2=3.58 Hz, H-1α). ^{13}C NMR (500 MHz, D_2O): 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).

4.6. Computational methods

4.6.1. Homology modelling. Due to the lack of a public structure with atomic resolution for the β-Gal-3-NTag, 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 thethaiotaomicron* (PDB ID: 3D3A, resolution 2.15 A), which was selected for performing the model. We have used two automatic modelling servers: SWISS-Model and CPH models-3.0, to build the model based on the alignment with the β-galactosidase of *B. thethaiotaomicron*. These two generated models were refined using the GROMOS 96 43a1 force field[53] 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.

4.6.2. 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,[35,48] 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,[54] 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.

4.6.3. Solvent parameters. The solvent used in the simulations was already parameterized for the GROMOS 96 43a1 force field,[53] 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.

4.6.4. 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.[35] GlcNac parameters were generated by the Dundee PRODRG 2.5 Server, carefully checking the generated topology as pointed out by Lemkul et al.[55] For both water and mixed solvent systems we first solvated the complex with the selected solvent and then we performed the following protocol: an energy minimization with 2000 steps of SD algorithm and 7000 steps of Polak-Ribiere Conjugate gradients (CG) was performed; an 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.[56] All simulations were performed using the GROMACS v4.0.7 suite.[53]

Acknowledgements

This work has been supported by the MICINN ('Ministerio de Ciencia e Innovación'), CTQ2009-11801 and CTQ2012-32042. Carlos Bayón is grateful for a Predoctoral fellowship granted by Complutense University of Madrid.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.tet.2013.04.015.

References and notes

1. Hakomori, S. *Glycoconjugate J.* **2004**, *21*, 125–137.
2. Yagi, H.; Yanagisawa, M.; Suzuki, Y.; Nakatani, Y.; Ariga, T.; Kato, K.; Yu, R. K. *J. Biol. Chem.* **2010**, *285*, 37293–37301.
3. Niederer, K. E.; Morrow, D. K.; Gettings, J. L.; Irick, M.; Krawiecki, A.; Brewster, J. L. *Cell. Signalling* **2005**, *17*, 177–186.
4. Yanagisawa, M.; Yu, R. K. *Biochem. Biophys. Res. Commun.* **2009**, *386*, 101–104.
5. Mandal, P. K.; Turnbull, W. B. *Carbohydr. Res.* **2010**, *346*, 2113–2120.
6. Springer, G. F. *Science* **1984**, *224*, 1198–1206.
7. Holeman, A.; Seeberger, P. H. *Curr. Opin. Biotechnol.* **2004**, *15*, 615–622.
8. Rapoport, E.; Khaidukov, S.; Baidina, O.; Bojenko, V.; Moiseeva, E.; Pasynina, G.; Karsten, U.; Nifant'ev, N.; LePendu, J.; Bovin, N. *Eur. J. Cell Biol.* **2003**, *82*, 295–302.
9. Shirato, H.; Ogawa, S.; Ito, H.; Sato, T.; Kameyama, A.; Narimatsu, H.; Xiaofan, Z.; Miyamura, T.; Wakita, T.; Ishii, K.; Takeda, N. *J. Virol.* **2008**, *82*, 10756–10767.
10. Caines, M. E. C.; Zhu, H.; Vuckovic, M.; Willis, L. M.; Withers, S. G.; Wakarchuk, W. W.; Strynadka, N. C. J. *J. Biol. Chem.* **2008**, *283*, 31279–31283.
11. Takada, A.; Ohmori, K.; Yoneda, T.; Tsuyuoka, K.; Hasegawa, A.; Kiso, M.; Kannagi, R. *Cancer Res.* **1993**, *53*, 354–361.
12. Lemieux, R. U.; Driguez, H. *J. Am. Chem. Soc.* **1975**, *97*, 4063–4069.
13. Khare, D. P.; Hindsgaul, O.; Lemieux, R. U. *Carbohydr. Res.* **1985**, *136*, 285–308.
14. Wilstermann, M.; Magnusson, G. *Carbohydr. Res.* **1995**, *272*, 1–7.
15. Schmidt, R. R.; Rücker, E. *Tetrahedron Lett.* **1980**, *21*, 1421–1424.
16. Schmidt, R. R. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 212–235.
17. Paulsen, H. *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 155–173.
18. Vetere, A.; Miletich, M.; Bosco, M.; Paoletti, S. *Eur. J. Biochem.* **2000**, *267*, 942–949.
19. D'Almeida, A.; Ionata, M.; Tran, V.; Tellier, C.; Dion, M.; Rabiller, C. *Tetrahedron: Asymmetry* **2009**, *20*, 1243–1246.
20. Baisch, G.; Ohrlein, R.; Streiff, M.; Kolbinger, F. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 751–754.
21. Yu, H.; Thon, V.; Lau, K.; Cai, L.; Chen, Y.; Mu, S.; Li, Y.; Wang, P. G.; Chen, X. *Chem. Commun.* **2010**, 7507–7509.
22. Anastas, P. T.; Warner, J. C. *Green Chemistry: Theory and Practice*; Oxford University Press: New York, NY, 1998.
23. Anastas, P. T. *Chem. Rev.* **2007**, *107*, 2167–2168.
24. Horvath, I. T.; Anastas, P. T. *Chem. Rev.* **2007**, *107*, 2169–2173.
25. Kates, R. W.; Clark, W. C.; Corell, R.; Hall, J. M.; Jaeger, C. C.; Lowe, I.; McCarthy, J. J.; Schellnhuber, H. J.; Bolin, B.; Dickson, N. M.; Faucheux, S.; Gallopin, G. C.; Grubler, A.; Huntley, B.; Jager, J.; Jodha, N. S.; Kasperson, R. E.; Mabogunje, A.; Matson, P.; Mooney, H.; Moore, B.; O'Riordan, T.; Svedin, U. *Science* **2001**, *292*, 641–642.
26. Hedbys, L.; Johansson, E.; Mosbach, K.; Larsson, P. O.; Gunnarsson, A.; Svensson, S.; Lonn, H. *Glycoconjugate J.* **1989**, *6*, 161–168.
27. Gambert, U.; Lio, R. G.; Farkas, E.; Thiem, J.; Bencomo, V. V.; Liptak, A. *Bioorg. Med. Chem.* **1997**, *5*, 1285–1291.
28. Schroder, S.; Schmidt, U.; Thiem, J.; Kowalczyk, J.; Kunz, M.; Vogel, M. *Tetrahedron* **2004**, *60*, 2601–2608.
29. Ito, Y.; Sasaki, T. *Biosci. Biotechnol. Biochem.* **1997**, *61*, 1270–1276.
30. Fujimoto, H.; Miyasato, M.; Ito, Y.; Sasaki, T.; Ajisaka, K. *Glycoconjugate J.* **1998**, *15*, 155–160.
31. Hernaiz, M. J.; Alcántara, A. R.; García, J. I.; Sinisterra, J. V. *Chem.—Eur. J.* **2010**, *16*, 9422–9437.
32. Perez, M.; Sinisterra, J. V.; Hernaiz, M. J. *Curr. Org. Chem.* **2010**, *14*, 2366–2383.
33. De Diego, T.; Lozano, P.; Abad, M. A.; Steffensky, K.; Vaultier, M.; Iborra, J. L. *J. Biotechnol.* **2009**, *140*, 234–241.
34. Sangeetha, K.; Morris, V. B.; Abraham, T. E. *Appl. Catal., A* **2008**, *341*, 168–173.
35. Sandoval, M.; Cortes, A.; Civera, C.; Trevino, J.; Ferreras, E.; Vaultier, M.; Berenguer, J.; Lozano, P.; Hernaiz, M. J. *RSC Adv.* **2012**, *2*, 6306–6314.
36. Hussain, W.; Pollard, D. J.; Truppo, M.; Lye, G. J. *J. Mol. Catal. B: Enzym.* **2008**, *55*, 19–29.
37. Eker, B.; Zagorevski, D.; Zhu, G.; Linhardt, R. J.; Dordick, J. S. *J. Mol. Catal. B: Enzym.* **2009**, *59*, 177–184.
38. Musa, M. M.; Fjeld, K. I. Z.; Vieille, C.; Phillips, R. S. *Org. Biomol. Chem.* **2008**, *6*, 887–892.
39. Galínski, M.; Lewandowski, A.; Stepniak, I. *Electrochim. Acta* **2006**, *51*, 5567–5580.
40. Kosmulski, M.; Gustafsson, J.; Rosenholm, J. B. *Thermochim. Acta* **2004**, *412*, 47–53.
41. Dzyuba, S. V.; Bartsch, R. A. *Tetrahedron Lett.* **2002**, *43*, 4657–4659.
42. Chiappe, C.; Pieraccini, D. *J. Phys. Org. Chem.* **2005**, *18*, 275–297.
43. Fox, D. M.; Awad, W. H.; Gilman, J. W.; Maupin, P. H.; De Long, H. C.; Trulove, P. C. *Green Chem.* **2003**, *5*.
44. Earle, M. J.; Esperanca, J. M. S. S.; Gilea, M. A.; Canongia Lopes, J. N.; Rebelo, L. P. N.; Magee, J. W.; Seddon, K. R.; Widegren, J. A. *Nature* **2006**, *439*, 831–834.
45. Tsuzuki, S.; Tokuda, H.; Hayamizu, K.; Watanabe, M. *J. Phys. Chem. B* **2005**, *109*, 16474–16481.
46. Wasserscheid, P.; Keim, W. *Angew. Chem., Int. Ed.* **2000**, *39*, 3772–3789.
47. Hagiwara, R.; Ito, Y. *J. Fluorine Chem.* **2000**, *105*, 221–227.
48. Pérez-Sánchez, M.; Cortés-Cabrera, Á.; García-Martín, H.; Sinisterra, J. V.; García, J. I.; Hernaíz, M. J. *Tetrahedron* **2011**, *67*, 7708–7712.
49. Pérez-Sánchez, M.; Sandoval, M.; Cortés-Cabrera, A.; García-Marín, H.; Sinisterra, J. V.; García, J. I.; Hernaiz, M. J. *Green Chem.* **2011**, *13*, 2810–2817.
50. Laemmli, U. K. *Nature* **1970**, *227*, 680–685.
51. Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248–254.
52. Bridiau, N.; Maugard, T. *Biotechnol. Prog.* **2011**, *27*, 386–394.
53. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
54. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
55. Lemkul, J. A.; Allen, W. J.; Bevan, D. R. *J. Chem. Informat. Model.* **2010**, *50*, 2221–2235.
56. Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins: Struct. Funct. Bioinf.* **1993**, *17*, 412–425.