



An expeditious and efficient synthesis of β -D-galactopyranosyl-(1 \rightarrow 3)-D-N-acetylglucosamine (lacto-N-biose) using a glycosynthase from *Thermus thermophilus* as a catalyst

Ayité D'Almeida, Marina Ionata, Vinh Tran, Charles Tellier, Michel Dion, Claude Rabiller*

Biotechnologie, Biocatalyse, Biorégulation (UMR CNRS 6204), Nantes University, 2, rue de la Houssinière, BP 92208, F-44322 Nantes, France

ARTICLE INFO

Article history:

Received 16 March 2009

Accepted 8 May 2009

Available online 6 June 2009

ABSTRACT

Mutant glycosynthases or transglycosidases obtained from a *Thermus thermophilus* β -D-glycosidase (TbGly) efficiently catalyzed the synthesis of β -(1 \rightarrow 3)-disaccharides. Unfortunately, this regioselectivity was changed to the β -(1 \rightarrow 4) one when N-acetylglucosamine derivatives were used as acceptors, thus precluding the possibility of synthesizing D-Galp- β -(1 \rightarrow 3)-D-GlcPNAc (lacto-N-biose) or D-GlcP- β -(1 \rightarrow 3)-D-GlcPNAc, which are useful synthons for the synthesis of antigen determinants. In contrast, we show in this work that, in the presence of phenyl 2-amino-1-thio- β -D-glucofuranoside, the 'normal' β -(1 \rightarrow 3) regioselectivity of E338G TbGly glycosynthase takes place. Thus, transglycosylations using α -galactosyl or α -glucosyl fluorides gave the corresponding phenyl β -D-glucofuranosyl-(1 \rightarrow 3)-2-amino-2-deoxy-1-thio- β -D-glucofuranosides in high yields (88–97%). Subsequent selective N-acylation followed by NBS/water deprotection of the thiophenyl group afforded lacto-N-biose in high overall yields.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Carbohydrate ligands play an important role in a number of biological mechanisms and there remains a need for green and low cost methods for their preparation despite the considerable development of efficient synthetic methods in this field. Over the last decades, enzyme-catalyzed reactions have proved to be very efficient in the construction of the glycosidic bond due to their stereoselectivity. For this purpose, retaining glycoside hydrolases is particularly useful, but their use was limited by the competition between hydrolysis and transglycosylation, both reactions catalyzed by these enzymes. However, site-directed mutagenesis, a strategy developed at the same time by Withers^{1–7} and Planas^{8–12} on retaining β -glycosidases, afforded new enzymes called glycosynthases. This approach was based on the mutation of the active nucleophile site, an aspartate or a glutamate, by a neutral amino acid (glycine, serine or alanine) and the use of donors with α anomeric configurations (such as α -glycosyl fluorides). These glycosynthases provided very high transglycosylation yields because they were unable to hydrolyze the transglycosylation product. In our laboratory, we have cloned and characterized a thermophilic β -glycosidase from *Thermus thermophilus*. This enzyme has been proven to accept β -D-fucosides (main activity), β -D-glucosides and β -D-galactosides as donors while the β -1,3 regioselectivity was observed in the hydrolytic and in the transglycosylation reac-

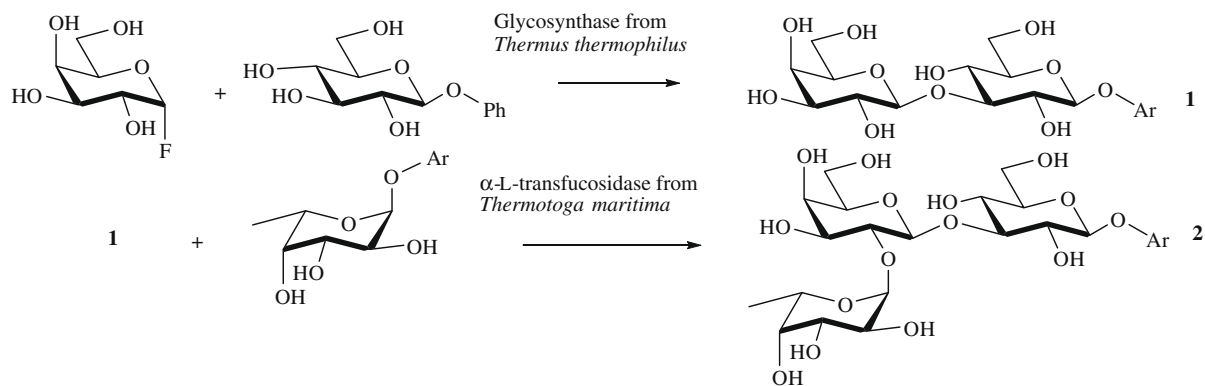
tions.^{13,14} We have validated the Withers and Planas strategy with this β -D-glycosidase by producing very efficient mutant glycosynthases able to catalyze the one-step synthesis of β -(1 \rightarrow 3) disaccharides.^{15,16} In addition, we have developed a new strategy based on directed evolution techniques which provided clones completely devoid of their hydrolytic activity while keeping a high transfer capability.^{17,18} These enzymes were called transglycosidases. Combining the two strategies, we have described the two-step synthesis of 4-nitrophenyl L-Fucp- α -(1 \rightarrow 2)-D-Galp- β -(1 \rightarrow 3)-D-GlcP **2**, an analogue of antigen H1, as shown in Scheme 1 (overall yield 35%).¹⁹ Unfortunately, when aryl 2-(acylamino)-2-deoxy- β -D-glucofuranosides [compounds necessary for the synthesis of antigen HI: L-Fucp- α -(1 \rightarrow 2)-D-Galp β -(1 \rightarrow 3)-D-GlcPNAc] were used as acceptors, the *T. thermophilus* glycosynthases, or the evolved *T. thermophilus* transglycosidases obtained by directed evolution, both catalyzed the synthesis of D-Galp β -(1 \rightarrow 4)-D-GlcPNAc disaccharides and not that of the β -(1 \rightarrow 3) regioisomer.²⁰

Thus, an alternative would be to replace the bulky acetamido group of the acceptor by the amino group, assuming that the latter would not induce any significant conformational changes in the active site of the enzyme.

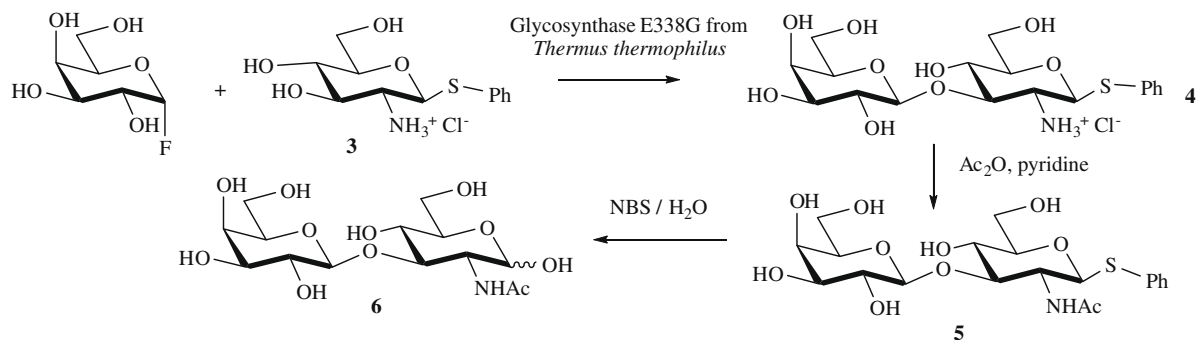
2. Results and discussion

Herein we report the successful synthesis of phenyl β -D-galactopyranosyl-(1 \rightarrow 3)-2-amino-2-deoxy-1-thio- β -D-glucofuranoside **4** using phenyl 2-amino-1-thio- β -D-glucofuranoside **3** as an acceptor substrate. Disaccharide **4** easily affords the 2-acetamido analogue **5**

* Corresponding author. Tel.: +33 (0)2 51 12 57 32; fax: +33 51 12 56 37.
E-mail address: claude.rabiller@univ-nantes.fr (C. Rabiller).



Scheme 1. Two-step enzymatic synthesis of L-Fucp- α -(1 \rightarrow 2)-D-Galp- β -(1 \rightarrow 3)-D-Glcp **2**.¹⁷



Scheme 2. Synthesis of D-Galp β -(1 \rightarrow 3)-D-GlcpNAc (lacto-*N*-biose) **6**.

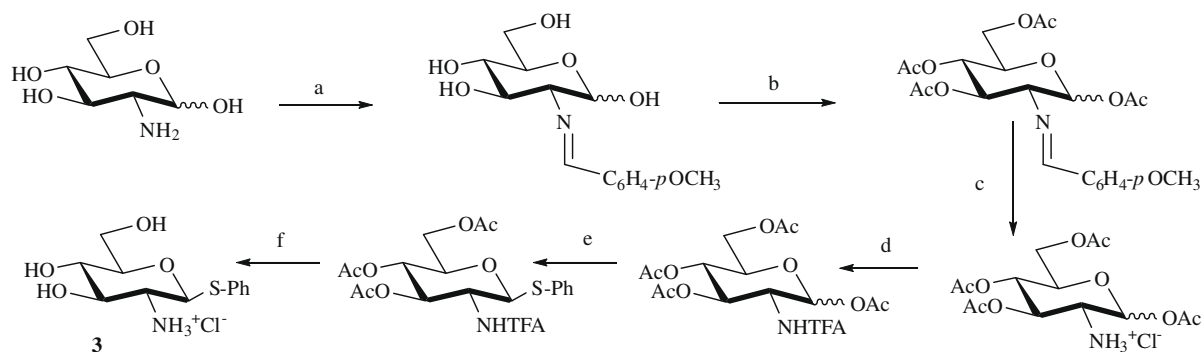
via selective acylation of nitrogen. Next, NBS/H₂O deprotection of the thiophenyl group gives β -D-galactopyranosyl-(1 \rightarrow 3)-2-acetamido-2-deoxy- β -D-glucopyranose **6** (Scheme 2).

The choice of the S-Ph group at the anomeric position for the acceptor was made since we have already shown that the presence of a β -aryl anomeric group is necessary for either the catalytic activity of wild type glycosidase from *T. thermophilus* or for the derived glycosynthases to ensure a correct positioning of the acceptor in the active site.¹⁵ Thus, the synthesis of **3** was achieved according to Scheme 3.²¹

Alternatively, **3** can be prepared via a slightly shorter strategy starting from 2-(acetylamino)-2-deoxy-2,3,4,6-tetra-*O*-acetyl- β -D-glucopyranoside.²²

Herein, the reaction shown in Scheme 2 was catalyzed by the glycosynthase E358G from *T. thermophilus*, using the readily available α -galactopyranosyl fluoride (GalF)¹² as a donor. Only two

reactions were observed in the presence of GalF: transglycosylation and the spontaneous hydrolysis of the fluoride. When operating at 37 °C and with a sufficient amount of enzymatic activity, the latter reaction slowly resulted in mixtures containing low amounts of galactose. In previous work, we had shown that the glycosynthases from *T. thermophilus* did not catalyze the hydrolysis of GalF or its self-condensation, which led to only one transglycosylation product.¹⁵ Our experimental results confirmed our predictions: acceptor **3** was a very good substrate for the enzyme and disaccharide **4** (structure established by means of NMR spectroscopy, see Section 4) could be selectively obtained in high yields. Moreover, we have already reported that our glycosynthase also accepted α -glucopyranosyl fluoride (GlcF) as a donor.¹⁵ Thus, in order to compare the reactivities of GalF and GlcF in the presence of **3**, the results obtained with both fluorides under different experimental conditions are shown in Table 1. As expected, phenyl



Scheme 3. Synthesis of phenyl 1-amino-deoxy-1-thio- β -D-glucopyranoside **3**. Reagents and conditions: (a) *p*-CH₃O-C₆H₄-CHO, NaOH 1 mol/L (95%); (b) (Ac)₂O, pyridine, DMAP (97%); (c) HCl 5 mol/L, acetone (98%); (d) TFA, pyridine, CH₂Cl₂ 95%; (e) PhSH, BF₃O(Et)₂ (75%); (f) (i) NaOH/MeOH 5 mol/L, (ii) HCl 2 mol/L, MeOH (95%).

Table 1

Yields (*calculated from the acceptor) obtained for the synthesis of phenyl β -D-glycopyranosyl-(1 \rightarrow 3)-2-amino-2-deoxy-1-thio- β -D-glycopyranosides **4** and **7**

Donor	Fluoride (mmol/L)	Acceptor 3 (mmol/L)	Yield* for 4 or 7
GalF	100	100	82
GalF	140	100	88
GlcF	100	100	92
GlcF	120	100	97

β -D-glycopyranosyl-(1 \rightarrow 3)-2-amino-2-deoxy-1-thio- β -D-glycopyranoside **7** (structure established by means of NMR spectroscopy, see Section 4) was obtained in slightly higher yields than **4**.

Moreover, as expected, the chemoselective N-acylation (vs O-acylation) and the deprotection of the thiophenyl aglycon group afforded nearly quantitatively lacto-*N*-biose **5** (80% overall yield from **3**, see Section 4). A drawback of our strategy, inherent in the *T. thermophilus* enzyme specificity, is the required presence of an *S*- or *O*-phenyl anomeric group on the acceptor. Thus, these results should be compared to other approaches using glycosidases able to accept free *N*-acetylglucosamine as a substrate. For example, *Xanthomonas manihotis* β -D-galactosidase allowed the preparation of lacto-*N*-biose from *p*-nitrophenyl β -D-galactopyranoside and 2-acetamido-2-deoxy-D-glycopyranose in 50% yield;²³ however this result was obtained with more than a 10-fold excess of the acceptor. Scaling-up this reaction under such conditions would probably lead to difficulties in the high yielding separation of lacto-*N*-biose. Furthermore, the free amino disaccharides **4** and **7** obtained by our method offer the opportunity to synthesize different useful amido derivatives which could be incorporated into oligosaccharides.

3. Conclusion

In conclusion, we have shown the possibility of circumventing the lack of regioselectivity of *T. thermophilus* glycosynthase towards *N*-acetylglucosamine derivatives. Thus, phenyl 2-amino-1-thio- β -D-glycopyranoside, an acceptor well recognized by these enzymes, allowed the synthesis of phenyl β -D-glycopyranosyl-(1 \rightarrow 3)-2-amino-2-deoxy-1-thio- β -D-glycopyranosides in high yields, the latter being easily converted to the corresponding *N*-acetyl disaccharides.

4. Experimental

4.1. General procedures

Glycosynthase E338G solution was prepared according to a procedure already described by us.²⁴ Chemicals were supplied by Aldrich and used without further purification. The course of the reactions was followed by means of TLC (precoated Silica Gel 60 sheets Merck F254) and proton NMR spectroscopy. The components of the reaction mixtures were separated by silica gel chromatography. Analysis of the ¹H and ¹³C NMR resonances and subsequent structure assignments were carried out using standard 2D sequences (COSY, HMBC, HMQC, TOCSY correlations). The structures of the β -(1,3)-regioisomers were established on the basis of the identification of the C-3 carbon (shifted at lower fields). The spectra were recorded with a Bruker DRX500 spectrometer operating at 500 MHz for ¹H and 126 MHz for ¹³C. Due to their solubility in water, the spectra of the saccharides were recorded in D₂O and the chemical shifts (in ppm) are quoted from the resonance of methyl protons of sodium 3-(trimethylsilyl)-propanesulfonate (DSS) used as an internal reference.

4.1.1. Phenyl β -D-glycopyranosyl-(1 \rightarrow 3)-2-amino-2-deoxy-1-thio- β -D-glycopyranosides **4** and **7**

GalF (182 g/mol, 255 mg, 1.4 mmol) or GlcF (182 g/mol, 218 mg, 1.2 mol) and phenyl 2-amino-1-thio- β -D-glycopyranoside (HCl salt, 307.8 g/mol, 308 mg, 1 mmol), were dissolved in a phosphate buffer (150 mmol/L, pH 7, 10 mL). Then, 2 mL of glycosynthase E338G sol. prepared as previously described²⁴ was added and the reaction was allowed to proceed at 37 °C until complete consumption of the fluoride (about 20 h). After removal of the solvent under reduced pressure, purification by silica gel chromatography (8:4:1 CHCl₃-MeOH-NH₄OH, *R*_f of **4**, 0.24 and *R*_f of **7**, 0.25) afforded 381 mg of pure **4** (free amine, white solid, yield 88%, mp 159–161 °C, [α]_D²² = –23.7 (c 0.35 MeOH), Calcd for C₁₈H₂₇NO₉S: C, 49.88; H, 6.24; N, 3.23; S, 7.39. Found: C, 49.52; H, 6.32; N, 3.52; S, 7.21) and 407 mg of pure **7** (white solid, yield 94%, mp 165–167 °C, [α]_D²² = –29.6 (c 0.45 MeOH), Calcd for C₁₈H₂₇NO₉S: C, 49.88; H, 6.24; N, 3.23; S, 7.39. Found: C, 49.59; H, 6.22; N, 3.41; S, 7.26).

Compound 4: ¹H NMR [500 MHz, D₂O, reference: (CH₃)₃Si-CH₂-CH₂-CH₂-SO₃Na]: δ 7.65 (2H, Ph), 7.49 (3H, Ph), 4.98 (1H, d, *J* 10.5 Hz, H-1), 4.60 (1H, d, *J* 7.5 Hz, H-1'), 3.95 (1H, dd, *J* 10.3, 8.7 Hz, H-3), 3.93 (1H, dd, *J* –12.4, 2.3 Hz, H-6a), 3.93 (1H, dd, *J* 3.3 Hz, H-4'), 3.78 (1H, dd, *J* –12.4, 5.5 Hz, H-6b), 3.73 (1H, ddd, *J* not determined, H-5'), 3.69 (1H, dd, *J* 9.9, 3.3 Hz, H-3'), 3.68 (1H, dd, *J* 9.9, 8.7 Hz, H-4), 3.63 (1H, dd, *J* 9.9, 7.5 Hz, H-2'), 3.57 (1H, ddd, *J* 9.9, 5.5, 2.3 Hz, H-5), 3.34 (1H, dd, *J* 10.5, 10.3 Hz, H-2). ¹³C NMR (125 MHz, D₂O, reference: (CH₃)₃Si-CH₂-CH₂-CH₂-SO₃Na): δ 135.1 (2 \times CH Ph), 133.0 (C, Ph), 132.2 (2 \times CH Ph), 131.7 (1 \times CH Ph), 105.9 (C-1'), 86.7 (C-1), 85.9 (C-3), 82.6 (C-5), 78.2 (C-5'), 75.2 (C-4), 73.6 (C-2'), 71.1 (C-4'), 70.4 (C-3'), 63.5 (C-6'), 63.2 (C-6), 56.3 (C-2).

Compound 7: ¹H NMR [500 MHz, D₂O, reference: (CH₃)₃Si-CH₂-CH₂-CH₂-SO₃Na]: δ 7.65 (2H, Ph), 7.49 (3H, Ph), 4.82 (1H, d, *J* 10.4 Hz, H-1), 4.61 (1H, d, *J* 7.9 Hz, H-1'), 3.79 (1H, dd, *J* –12.3, 2.3 Hz, H-6'a), 3.78 (1H, dd, *J* 10.3, 9.5 Hz, H-3), 3.75 (1H, dd, *J* –12.2, 2.2 Hz, H-6a), 3.64 (1H, dd, *J* –12.3, 5.9 Hz, H-6'b), 3.63 (1H, dd, *J* –12.2, 5.5 Hz, H-6b), 3.51 (1H, dd, *J* 9.2, 9.5 Hz, H-4), 3.45 (1H, dd, *J* 8.1, 7.9 Hz, H-4'), 3.40 (1H, ddd, *J* 9.2, 5.5, 2.2 Hz, H-5), 3.40 (1H, ddd, *J* 7.9, 5.9, 2.3 Hz, H-5'), 3.34 (1H, dd, *J* 9.4, 7.9 Hz, H-2'), 3.33 (1H, dd, *J* 9.4, 8.1 Hz, H-3'), 3.18 (1H, dd, *J* 10.4, 10.3 Hz, H-2). ¹³C NMR (125 MHz, D₂O, reference: (CH₃)₃Si-CH₂-CH₂-CH₂-SO₃Na): δ 135.1 (2 \times CH Ph), 133.0 (C, Ph), 132.2 (2 \times CH Ph), 131.7 (1 \times CH Ph), 106.1 (C-1'), 89.3 (C-1), 89.3 (C-3), 83.2 (C-5'), 82.9 (C-5), 78.7 (C-3'), 72.5 (C-4), 71.2 (C-4'), 63.9 (C-6), 63.7 (C-6'), 57.3 (C-2).

4.1.2. Phenyl β -D-galactopyranosyl-(1 \rightarrow 3)-2-acetamido-2-deoxy-1-thio- β -D-glycopyranoside **5**

Disaccharide **4** (free amino group, 433 g/mol, 43 mg, 0.1 mmol) was dissolved in 3 mL of MeOH. Then, sodium bicarbonate (84 mg, 1 mmol) and acetic anhydride (102 g/mol, 100 μ L, 1 mmol) were added. The mixture was maintained at rt for 45 min while stirring. After this time, TLC plates indicated the total consumption of **5**. Next, MeOH was distilled under reduced pressure, purification by silica gel chromatography (8:4:1 CHCl₃-MeOH-NH₄OH, *R*_f of **5**, 0.36) afforded 45 mg of pure **5** (white solid, yield 94%, mp 162–164 °C, Calcd for C₂₀H₂₉NO₁₀S: C, 50.53; H, 6.10; N, 2.95; S, 6.74. Found: C, 50.27; H, 6.22; N, 3.12; S, 6.59). ¹H NMR [500 MHz, D₂O, reference: (CH₃)₃Si-CH₂-CH₂-CH₂-SO₃Na]: δ 7.65 (2H, Ph), 7.49 (3H, Ph), 4.93 (1H, d, *J* 10.6 Hz, H-1), 4.42 (1H, d, *J* 7.7 Hz, H-1'), 3.92 (1H, dd, *J* 10.6, 10.3 Hz, H-2), 3.91 (1H, dd, *J* –12.5, 2.1 Hz, H-6b), 3.90 (1H, dd, *J* 3.3, nd Hz, H-4'), 3.82 (1H, dd, *J* 10.3, 8.2 Hz, H-3), 3.76 (1H, dd, *J* –12.5, 5.2 Hz, H-6a), 3.69 (1H, ddd, *J* nd, H-5'), 3.63 (1H, dd, *J* 10.0, 3.3 Hz, H-3'), 3.57 (1H, dd, *J* 10.0, 8.2 Hz, H-4), 3.52 (1H, ddd, *J* 10.0, 5.2, 2.1 Hz, H-5), 3.50 (1H, dd, *J* 10.0, 7.7 Hz, H-2'). ¹³C NMR (125 MHz, D₂O, reference:

(CH₃)₃Si-CH₂-CH₂-CH₂-SO₃Na]: δ 183.5 (CO), 135.5 (2 \times CH Ph), 134.5 (C Ph), 132.2 (2 \times CH Ph), 130.9 (1 \times CH Ph), 106.2 (C-1'), 88.8 (C-1), 86.3 (C-3), 82.3 (C-5), 78.0 (C-5'), 75.2 (C-3'), 73.4 (C-2'), 71.2 (C-4'), 71.2 (C-4), 63.7 (C-6), 63.5 (C-6'), 25.6 (CH₃).

4.1.3. β -D-Galactopyranosyl-(1 \rightarrow 3)-2-acetamido-2-deoxy-D-glucopyranoside **6**

Phenyl β -D-glycopyranosyl-(1 \rightarrow 3)-2-acetamido-2-deoxy-1-thio- β -D-glucopyranoside **5** (475 g/mol, 48 mg, 0.1 mmol) was dissolved in 2 mL of distilled water. Then, *N*-bromosuccinimide (178 g/mol, 35.6 mg, 0.2 mmol) was added and the mixture was stirred for 2 h at rt. At this time, the disaccharide was completely consumed as revealed by TLC. Purification of **6** by silica gel chromatography (2:1:1 *n*BuOH-AcOH-water, *R_f* of **6**, 0.28), afforded 37 mg of **6** (white solid, yield 97%, mp 175–177 °C, $[\alpha]_D^{22} = +3.1$ (equilibrium value α , β anomers) (c 0.65 MeOH), Calcd for C₁₄H₂₅NO₁₁: C, 43.86; H, 6.53; N, 3.66. Found: C, 43.62; H, 6.38; N, 3.52). ¹H and ¹³C NMR spectra were identical to those already described.²³

Acknowledgements

Thanks are due to the French Ministry of Education and Research, to the Centre National de la Recherche Scientifique and to the Region des Pays de la Loire (VANAM Program) for financially supporting this work.

References

- Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. *J. Am. Chem. Soc.* **1998**, *120*, 5583–5584.
- Jahn, M.; Stoll, D.; Warren, R. A. J.; Szabo, L.; Singh, P.; Gilbert, H. J.; Ducros, V. M. A.; Davies, G. J.; Withers, S. G. *Chem. Commun.* **2003**, 1327–1329.
- Jahn, M.; Chen, H.; Muellegger, J.; Marles, J.; Warren, R. A. J.; Withers, S. G. *Chem. Commun.* **2004**, 274–275.
- Mayer, C.; Zechel, D. L.; Reid, S. P.; Warren, R. A.; Withers, S. G. *FEBS Lett.* **2000**, *466*, 40–44.
- Tolborg, J. F.; Petersen, L.; Jensen, K. J.; Mayer, C.; Jakeman, D. L.; Warren, R. A. J.; Withers, S. G. *J. Org. Chem.* **2002**, *67*, 4143–4149.
- Viladot, J.-L.; Canals, F.; Batllori, X.; Planas, A. *Biochem. J.* **2001**, *355*, 79–86.
- Nashiru, O.; Zechel, D. L.; Stoll, D.; Mohammadzadeh, T.; Warren, R. A. J.; Withers, S. G. *Angew. Chem., Int. Ed.* **2001**, *40*, 417–420.
- Faijes, M.; Fairweather, J. K.; Driguez, H.; Planas, A. *Chem. Eur. J.* **2001**, *7*, 4651–4655.
- Faijes, M.; Perez, X.; Perez, O.; Planas, A. *Biochemistry* **2003**, *42*, 13304–13318.
- Faijes, M.; Imai, T.; Bulone, V.; Planas, A. *Biochem. J.* **2004**, *380*, 635–641.
- Fairweather, J. K.; Faijes, M.; Driguez, H.; Planas, A. *ChemBioChem* **2002**, *3*, 866–873.
- Malet, C.; Planas, A. *FEBS Lett.* **1998**, *440*, 208–212.
- Fourage, L.; Dion, M.; Colas, B. *Glycoconjugate J.* **2001**, *17*, 377–383.
- Chiffolleau-Giraud, V.; Spangenberg, P.; Dion, M.; Rabiller, C. *Eur. J. Org. Chem.* **1999**, 757–763.
- Drone, J.; Feng, H.-y.; Tellier, C.; Hoffmann, L.; Tran, V.; Rabiller, C.; Dion, M. *Eur. J. Org. Chem.* **2005**, 1977–1983.
- Marton, Z.; Tran, V.; Tellier, C.; Dion, M.; Drone, J.; Rabiller, C. *Carbohydr Res.* **2008**, *343*, 2939–2946.
- Osanjo, G.; Dion, M.; Drone, J.; Solleux, C.; Tran, V.; Rabiller, C.; Tellier, C. *Biochemistry* **2007**, *46*, 1022–1033.
- Feng, H.-Y.; Drone, J.; Hoffmann, L.; Tran, V.; Tellier, C.; Rabiller, C.; Dion, M. *J. Biol. Chem.* **2005**, *280*, 37088–37097.
- Osanjo, G., Thesis Nantes, France, 2007.
- Hoffmann, L.; Tran, V.; Rabiller, C.; Tellier, C.; Dion, M. *Protein Eng. Des. Sel.*, submitted for publication.
- Silva, D. J.; Wang, H.; Allanson, N. M.; Jain, R. K.; Sofia, M. J. *J. Org. Chem.* **1999**, *64*, 5926–5929.
- Benakli, K.; Zha, C.; Kerns, R. J. *J. Am. Chem. Soc.* **2001**, *123*, 9461–9462.
- Vetere, A.; Miletich, M.; Bosco, M.; Paoletti, S. *Eur. J. Biochem.* **2000**, *267*, 942–949.
- Dion, M.; Fourage, L.; Hallet, J.-N.; Colas, B. *Glycoconjugate J.* **1999**, *16*, 27–37.