ChemComm www.rsc.org/chemcomm

Efficient synthesis of globoside and isogloboside tetrasaccharides by using β(1→3) *N*-acetylgalactosaminyltransferase/ UDP-*N*-acetylglucosamine C4 epimerase fusion protein†

Jun Shao, Jianbo Zhang, Przemyslaw Kowal, Yuquan Lu and Peng George Wang
Department of Chemistry, Wayne State University, Detroit, MI 48202, USA. E-mail: pwang@chem.wayne.edu; Fax: 313-577-2554; Tel: 313-993-6759

Received (in Cambridge, MA) 19th February 2003, Accepted 7th April 2003
First published as an Advance Article on the web 15th May 2003

The β(1→3) *N*-acetylgalactosaminyltransferase/UDP-*N*-acetylglucosamine C4 epimerase fusion protein was constructed and used in coupled enzymatic reactions to synthesize a variety of globotetraose and isoglobotetraose derivatives from the corresponding lactoside acceptors.

Globoside and isogloboside are glycosphingolipids expressed in human red blood cells and other tissues. Evidence has shown that these glycolipids are receptors in the adhesion of uropathogenic bacteria to urinary tract epithelial cells.[1] The binding of bacterial toxins to the oligosaccharide segments of these glycans is a critical step in causing serious clinical manifestations of human infections, such as hemorrhagic colitis and hemolytic uremic syndrome.[2] In addition, human parvovirus B19 binds the erythroid progenitor cells *via* interaction with three globotetraose molecules.[3] Therefore, efficient synthesis of the oligosaccharide derivatives as potential inhibitors is of growing value in disease diagnosis, treatment and related biochemical studies.

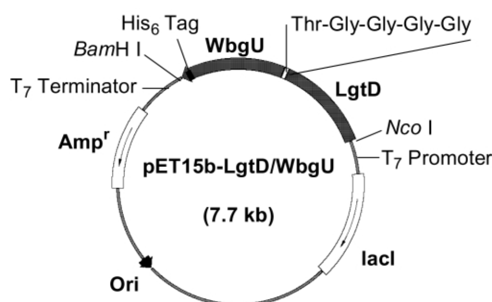


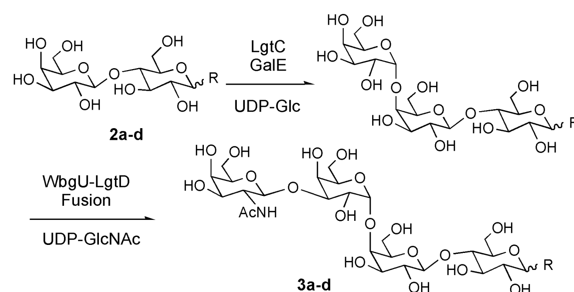
Fig. 1 The plasmid map of LgtD-WbgU fusion protein.

Table 1 Acceptor specificity of LgtD-WbgU fusion enzyme

Entry	Acceptor compounds	Relative activity (%)[a]
1a	Galα1→4Galβ1→4Glc	100
1b	Galα1→4Galβ1→4Glc–OBn	121
1c	Galα1→4Galβ1→4Glc–OMe	166
1d	Galα1→4Galβ1→4Glc–N_3	92
1e	Galα1→4Galβ1→4Glc–SPh	82
1f	Galα1→4Galβ1→4GlcNAc	116
1g	Galα1→3Galβ1→4Glc	64
1h	Galα1→3Galβ1→4Glc–OBn	77
1i	Galα1→3Galβ1→4Glc–OMe	86
1j	Galα1→3Galβ1→4Glc–NHAc	68
1k	Galα1→3Galβ1→4Glc–SPh	54
1l	Galα1→3Galβ1→4GlcNAc	76

[a] The activity of β(1→3) *N*-acetylgalactosaminyltransferase was assayed according to a previously published protocol[7] except that the acceptor was varied.

† Electronic supplementary information (ESI) available: further experimental details. See http://www.rsc.org/suppdata/cc/b3/b300831b/

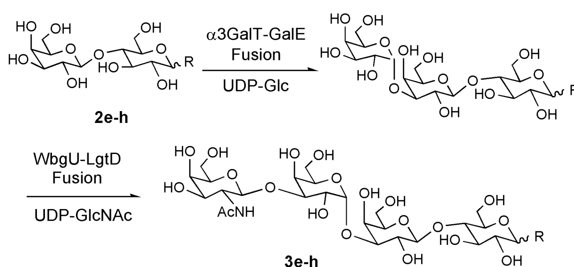


Entry	Compound	R	Yield (%)
1	3a	OH	75
2	3b	OBn	78
3	3c	OMe	85
4	3d	N_3	69

Scheme 1 Coupled enzymatic synthesis of globotetraose derivatives.

Due to multiple protection and purification steps, chemical synthesis of (iso)globotetraose and some derivatives[4] have suffered relatively low yields. Enzymatic approaches, on the other hand, are highly stereo- and regioselective for the glycosidic linkage formation, thereby allowing the straightforward synthesis of carbohydrates and glycoconjugates.[5] Recently, several fusion enzymes involving glycosyltransferases have been developed such as the bovine α(1→3) galactosyltransferase/*E. coli* UDP-glucose C4 epimerase, the *N. meningitidis* β(1→4) galactosyltransferase/*S. thermophilus* UDP-glucose C4 epimerase, and the *N. meningitidis* α(2→3) sialyltransferase/CMP-*N*-acetylneuraminic acid synthetase fusion proteins.[6] It has been proved that these recombinant enzymes are highly effective tools in large-scale synthesis of oligosaccharides.

β(1→3) *N*-Acetylgalactosaminyltransferase (LgtD) from *H. influenzae* has previously been expressed in *E. coli*.[7] We report

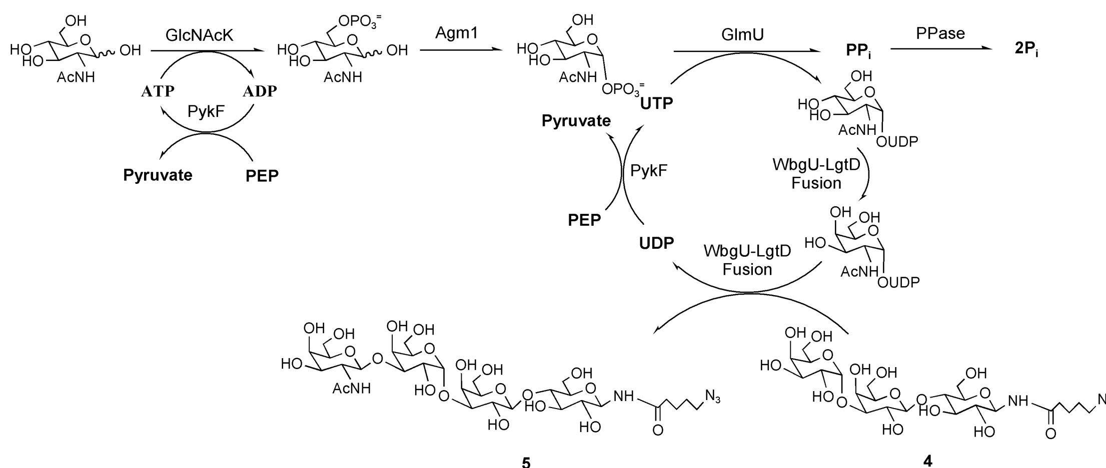


Entry	Compound	R	Yield (%)
1	3e	OH	65
2	3f	OBn	67
3	3g	OMe	72
4	3h	NHAc	60

Scheme 2 Coupled enzymatic synthesis of isoglobotetraose derivatives.

DOI: 10.1039/b300831b

This journal is © The Royal Society of Chemistry 2003



Scheme 3 Synthesis of GalNAcβ1–3Galα1–3Galβ1–4Glcβ-5-azidopentamide **5** with *in situ* UDP-GlcNAc regeneration.

here the construction of LgtD/WbgU (UDP-GlcNAc C4 epimerase from *P. shigelloides*)[8] fusion protein. WbgU catalyzes the interconversion of UDP-GlcNAc to UDP-GalNAc; meanwhile, LgtD catalyzes the introduction of GalNAc residue to the acceptor substrates. The overall function of the recombinant protein is the transfer of the GlcNAc residue from UDP-GlcNAc to the acceptors. Thus, the construction alters the donor substrate requirement of the reaction from UDP-GalNAc to UDP-GlcNAc This change is cost effective because the commercial price of UDP-GlcNAc is 20 times less than that of UDP-GalNAc. Moreover, it avoids the need to express and purify two recombinant enzymes for reactions

The bifunctional LgtD-WbgU fusion enzyme was prepared by in-frame linking of the *P. shigelloides wbgU* gene downstream to the *H. influenzae lgtD* gene through a five-residue peptide linker (Thr-Gly-Gly-Gly-Gly). The C-terminus of the fusion protein includes a 6 × His tag for convenient purification by nickel–nitrilotriacetic acid affinity chromatography (Fig. 1). The enzyme was expressed in *E. coli* BL21(DE3) at a relatively high level (22 U L^{-1}). The one-step purified enzyme has an estimated molecular weight of 70 kDa and a specific activity of 1.68 U mg^{-1}.

To investigate its acceptor specificity, the GalNAc transferase activity was tested with 12 carbohydrate compounds including globotriose, Galα1,3Lac and some derivatives **1a–l**. As shown in Table 1, all these oligosaccharides are good substrates though compounds with a terminal α(1→4)-linked Gal-Gal structure are better acceptors than those with an α(1→3) linkage (compare **1a** with **1g**, and **1f** with **1l**). Apparently, the fusion protein has very broad acceptor substrate specificity to the anomeric aglycon.

The fusion protein was then employed in coupled enzymatic glycosylation reactions (100 mg scale). Globotetraose **3a** and its β-benzyl **3b**, β-methyl **3c** and β-azido **3d** derivatives were synthesized from the corresponding lactosides **2a–d** in reactions catalyzed sequentially by the α(1→4) galactosyltransferase (LgtC) from *N. meningitidis* and the fusion enzyme (Scheme 1). The UDP-Glc C4 epimerase (GalE) from *E. coli* K12 was also used so that UDP-Glc and UDP-GlcNAc, instead of expensive UDP-Gal and UDP-GalNAc, were supplied as donor substrates. Compound **3d** has a purposely-introduced azido group, which makes it flexible in the solid-phase synthesis of glycopeptides and glycopolymers. Meanwhile, isoglobotetraose **3e** and its β-benzyl **3f**, β-methyl **3g** and β-acetylamide **3h** derivatives were synthesized from the corresponding lactosides **2e–h** in one-pot reactions catalyzed by GalE from *E. coli* K12, α (1→3) galactosyltransferase (α1,3GalT) from bovine and the LgtD-WbgU fusion protein (Scheme 2). It should be noted that these compounds also possess the α-Gal (Galα1,3Gal) structure, which is believed to have potential applications in xenotransplantation. After two steps of purification, overall yields from 60% to 85% were achieved in the reactions. It is a significant improvement as compared to the chemical synthesis of these tetrasaccharides, for which the total yields from lactosides are normally below 40% due to multiple protection and purification steps.[4]

The synthetic efficiency of the fusion protein was also tested by coupling with a multiple-enzyme UDP-GlcNAc regeneration cycle which consists of UDP-GlcNAc pyrophosphorylase (GlmU), pyruvate kinase (PykF), and inorganic pyrophosphatase (PPase) from *E. coli* K12; GlcNAc phosphate mutase (Agm1) from *S. cerevisiae*; and GlcNAc kinase (GlcNAcK) from *C. albicans* (Scheme 3). This system further reduces the cost of glycosylation reactions because it avoids the use of expensive sugar nucleotides. An overall yield of 70% was achieved in the reaction. We have previously used the accepter **4** in the synthesis of α-Gal conjugated polymers that significantly inhibit the binding of human anti-Gal antibody to mouse laminin glycoprotein and mammalian PK15 cells.[9] Similarly, the oligosaccharide product **5** can be easily transformed into other functional glycoconjugates that may serve as tools in bacterial toxin inhibition, xenotransplantation and other pharmaceutical studies.

This work was supported by National Institutes of Health Grant AI 44040.

Notes and references

1 M. J. Kuehn, J. Heuser, S. Normark and S. J. Hultgren, *Nature*, 1992, **356**, 252–255.
2 M. A. Karmali, B. T. Steele, M. Petric and C. Lim, *Lancet*, 1983, **1**, 619–620.
3 P. R. Chipman, M. Agbandje-McKenna, S. Kajigaya, K. E. Brown, N. S. Young, T. S. Baker and M. G. Rossman, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 7502–7506.
4 (*a*) S. Bhattacharyya, B. G. Magnusson, U. Wellmar and U. J. Nilsson, *J. Chem. Soc., Perkin Trans. 1*, 2001, 886–890; (*b*) U. Nilsson, A. K. Ray and G. Magnusson, *Carbohydr. Res.*, 1994, **252**, 117–136; (*c*) H. Ishida, R. Miyawaki, M. Kiso and A. Hasegawa, *Carbohydr. Res.*, 1996, **284**, 179–190.
5 K. M. Koeller and C. H. Wong, *Chem. Rev.*, 2000, **100**, 4465–4494.
6 (*a*) J. Fang, X. Chen, W. Zhang, A. Janczuk and P. G. Wang, *Carbohydr. Res.*, 2000, **329**, 873–878; (*b*) O. Blixt, J. Brown, M. J. Schur, W. Wakarchuk and J. C. Paulson, *J. Org. Chem.*, 2001, **66**, 2442–2448; (*c*) M. Gilbert, R. Bayer, A. M. Cunningham, S. DeFrees, Y. Gao, D. C. Watson, M. N. Young and W. W. Wakarchuk, *Nat. Biotechnol.*, 1998, **16**, 769–772.
7 J. Shao, J. Zhang, P. Kowal, Y. Lu and P. G. Wang, *Biochem. Biophys. Res. Commun.*, 2002, **295**, 1–8.
8 P. Kowal and P. G. Wang, *Biochemistry*, 2002, **41**, 5410–4.
9 J.-Q. Wang, X. Chen, W. Zhang, S. Zacharek, Y. Chen and P. G. Wang, *J. Am. Chem. Soc.*, 1999, **121**, 8174–8181.