COMMUNICATION

Enzymatic synthesis of *N*-acetyllactosamine from lactose enabled by a reversible galactosyltransferase

Received 00th January 20xx,
Accepted 00th January 20xx

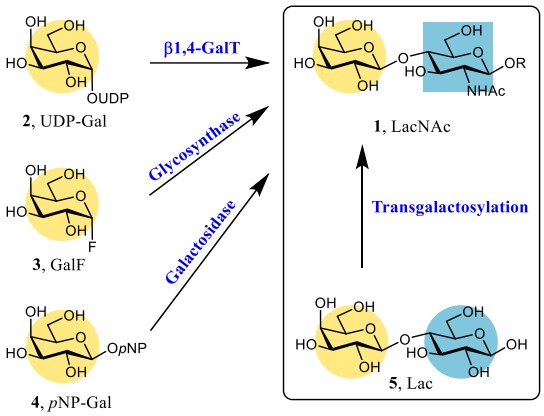
DOI: 10.1039/x0xx00000x

Kun Huang,[a] Fabio Parmeggiani,[a] Helene Ledru,[b] Kristian Hollingsworth,[c] Jordi Mas Pons, [b] Andrea Marchesi,[a] Peter Both,[a] Ashley P. Mattey,[a] Edward Pallister,[a] Gregory S. Bulmer, [a] Jolanda M. van Munster, [a] W. Bruce Turnbull, [c] M. Carmen Galan,*[b] Sabine L. Flitsch*[a]

Utilising a fast and sensitive screening method based on imidazolium tagged probes, we report unprecedented reversible activity of bacterial β1,4-galactosyltransferases to catalyse the transgalactosylation from lactose to *N*-acetylglucosamine to form *N*-acetyllactosamine in the presence of UDP. The process is demonstrated by the preparative scale synthesis of *p*NP-β-LacNAc from lactose using β1,4-galactosyltransferase NmLgtB-B as the only biocatalyst.

Galactosides are among the most abundant glycans in the mammalian glycome and generally biosynthesised by Leloir galactosyltransferases. In particular, *N*-acetyllactosamine (LacNAc) is a common core glycan motif (Type 2 glycan) in free oligosaccharides, glycoproteins and glycolipids. Galactosides, including LacNAc, are important constituents of human milk oligosaccharides, which have great health benefits for infants.[1,2] Sialylated and fucosylated LacNAc such as sialyl Lewis^X have been described as ligands of various lectins, such as selectins.[3] The demand of galactosides for biological investigation and commercialisation as additives to formula milk has increased considerably over the last decade and synthetic methods employing biocatalysts are very attractive compared to multi-step chemical strategies.[4–10]

The central role of galactosides in these bioactive oligosaccharides, in particular LacNAc (**1**), has prompted the development of several enzymatic synthetic strategies (Scheme 1). Key to all is the activation of the galactose anomeric centre, since direct glycosidic bond formation from free reducing sugars is unfavourable. In biosynthesis, UDP-galactose (**2**) is commonly utilised as the activated substrate by a wide range of galactosyltransferases, but the cost of this substrate can be prohibitive in large scale synthesis.

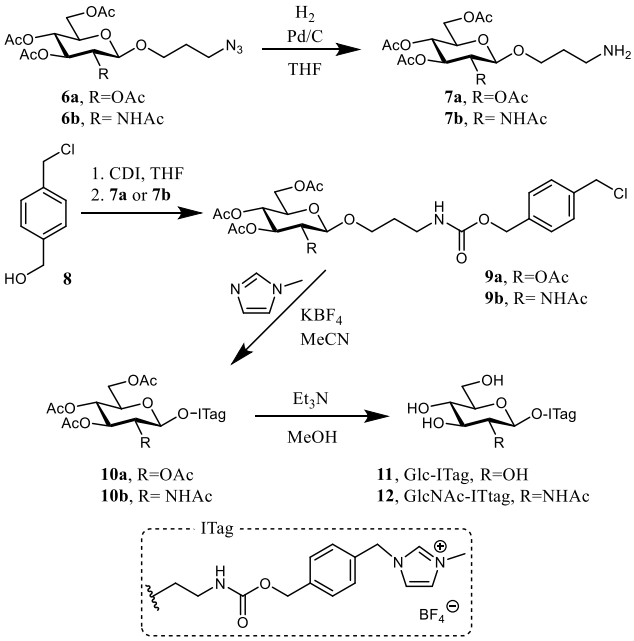


Scheme 1. Enzymatic approaches for LacNAc derivative (**1**) synthesis using GlcNAc-R as acceptor and **2-5** as donor substrates.

Some elegant alternative synthetic approaches have been developed using activated substrates such as galactosyl fluoride (**3**) [11] and *p*NP-galactose (**4**).[12] UDP-galactose (**2**) can also be regenerated using a multienzyme system either from galactose or sucrose, which adds additional steps and potential side-products.[13–15] It has been recognised that the most cost-efficient galactosyl donor would be lactose (**5**), a waste product of the cheese industry that is produced at 6 million ton scale every year. So far, the use of lactose (**5**) as substate has been limited to galactosidases, which have some inherent transgalactosylation activity, but show generally poor regioselectivity and low yield. [16,17]

[a] *Manchester Institute of Biotechnology (MIB) & School of Chemistry, The University of Manchester, 131 Princess Street, M1 7DN, Manchester, United Kingdom. E-mail: Sabine.Flitsch@manchester.ac.uk*
[b] *School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom. E-mail: m.c.galan@bristol.ac.uk*
[c] *School of Chemistry and Asbury Centre for Structural Molecular Biology, The University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, United Kingdom.*
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

These findings prompted us to look for biocatalysts that would be able to generate the donor UDP-galactose (**2**) directly *in situ* from lactose (**5**) and UDP without any further cofactor. In general, Leloir glycosyltransferases are perceived to catalyse unidirectional reactions.[18] However, the reversible catalytic activity of natural product glycosyltransferases has been described,[19] and subsequently employed to synthesise a wealth of different nucleotide sugars in the presence of nucleotides.[20] In addition, sucrose synthase, a type of Leloir glycosyltransferase, was widely used to produce nucleotide diphosphate glucose according to its reversible activity.[21] Furthermore, mammalian sialyltransferase ST3Gal-II was able to synthesise CMP-Neu5Ac from α2,3 sialylated glycans and glycoconjugates in the presence of CMP, [22]which was utilised to label sialic acid containing glycoproteins and gangliosides.[23] Recently, two bacterial α2,6-sialyltransferases were exploited as specific α2,6-sialidases based on their reverse activity.[24,25] Furthermore, glycosyl transfer catalysed by β1,4-*N*-acetylglucosaminyltransferase III (GnT-III) was also reversible.[26] However, to the best of our knowledge there is



Scheme 2. Chemical synthesis of Glc-ITag (**11**) and GlcNAc-ITag (**12**).

no report on galactosyltransferases exhibiting reverse transfer from cheaply available lactose.

For the initial screening experiments of transgalactosylation activity, it was important to develop a fast and robust assay that would detect even weak galactose transfer activity in the presence of high excess of lactose **5.** Our previous work had shown that sugar acceptors tagged with imidazolium-based probes (I-Tags) allow for the monitoring of glycosylation reactions by mass spectrometry (MALDI-ToF),[27–30] even in complex mixtures.[31] The cationic I-Tag generates a strong MS signal that dominates the ionisation of the analytes[30] and allows for a reasonable estimate of overall yields by measurement of starting material and product peaks.

For the purpose of the present study, a new class of benzyl carbamate-containing I-Tagged glucosides of Glc (**11**) and GlcNAc (**12**) were chemically synthesised from 1-azidopropyl (**6a**)[32] and (**6b**)[29] in 4 steps, giving 18% and 21% overall yield respectively (Scheme 2, See supporting information for further details). As expected, MALDI-ToF spectra of Glc-ITag (**11**) and GlcNAc-ITag (**12**) provided strong peaks with expected mass in aqueous and buffer solutions (Figure 1). Furthermore, when both I-Tagged substrates were treated with a galactosyltransferase (NmLgtB-A) and UDP-Gal (**2**) the glycosylation products could be clearly observed by MALDI-ToF without any further purification (Figure 1), providing an excellent basis for further screening studies.

The I-Tag methodology was next used to screen galactosyltransferase activity from lactose (**5**) instead of UDP-Gal (**2**). Three bacterial β1,4-galactosyltransferases were cloned from *Neisseria meningitidis* serogroup A strain

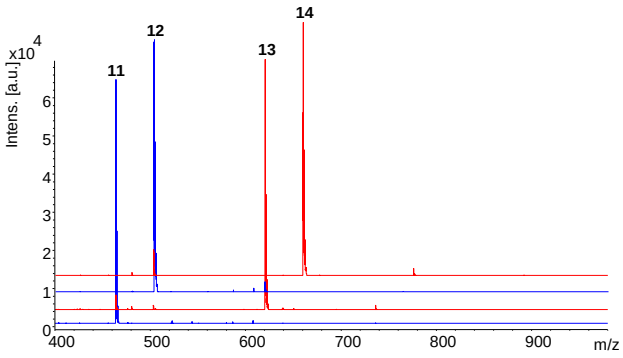
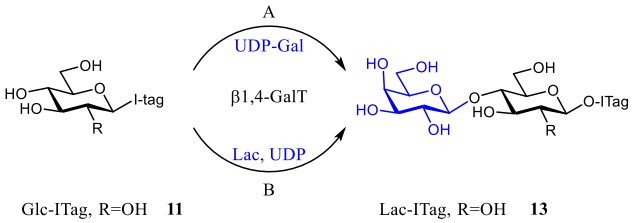


Figure 1. MALDI-ToF mass spectra of ITag substrates (**11** and **12**, blue traces) and their biotransformation products (**13** and **14**, red traces). Glc-ITag **11** $[M]^+$ = 466, GlcNAc-ITag **12** $[M]^+$ = 507, Lac-ITag **13** $[M]^+$ = 628 , and LacNAc-ITag **14** $[M]^+$ = 669.

Z2491(NmLgtB-A), *Neisseria meningitidis* serogroup B strain MC58 (NmLgtB-B)[8] and *Neisseria meningitidis* (NmLgtH). NmLgtB-A and NmLgtB-B are homologous proteins from different strains (92% identity), while NmLgtH shows 71% and 72% identity to NmLgtB-A and NmLgtB-B, respectively, as shown by the amino acid sequence alignment (Supporting information, Figure S2). All three proteins were produced

recombinantly in *E. coli* BL21 (DE3) with an N-terminal His_6-tag and purified by affinity chromatography. In the first instance, the activity of three β1,4-galactosyltransferases was confirmed by using Glc-ITag (**11**) and GlcNAc-ITag (**12**) as substrates (Scheme 3), although LgtH displayed only low activity against GlcNAc-ITag (**12**) (data not shown).

Having established a sensitive MS-based assay on I-Tagged substrates, lactose (**5**) was tested as galactose donor in the presence of UDP (Scheme 3). Given that the equilibrium between lactose and UDP-Gal would be expected to be unfavourable towards the latter, UDP-Gal generation was monitored by coupling the reactions with subsequent galactose transfer to the I-Tagged acceptor substrates **11** and **12** (Scheme 3), which would result in transfer of galactose from lactose to substrate *via in situ* formation of UDP-Gal (**2**). Rewardingly, formation of Lac-ITag (**13**) and LacNAc-ITag (**14**) from **11** and **12** (Scheme 3) could be detected by MALDI-ToF spectrometry in the NmLgtB-B catalysed reactions, while NmLgtB-A afforded only a very low conversion and NmLgtH no conversion at all (Supporting information, from Figure S3 to S8). Therefore, NmLgtB-B was used in subsequent experiments.



Glc-ITag, R=OH **11**
GlcNAc-ITag, R=NHAc **12**
Lac-ITag, R=OH **13**
LacNAc-ITag, R=NHAc **14**

Scheme 3. Investigation of galactosyltransferase activity with I-Tag acceptor substrates **11** and **12** using A: UDP-Gal (**2**) or B: lactose (**5**) as sugar donor substrates.

The scope of galactose donor substrates beyond lactose was also tested using *p*NP-β-Lac, *p*NP-α-Gal, *p*NP-β-Gal and LacNAc (**1**) as galactose donors in the presence of UDP. Galactose transfer was also observed in reactions containing *p*NP-β-Lac and LacNAc (Supporting information, Figure S9 and S10). However, since lactose (**5**) is very inexpensive and easily available, it was subsequently used as a preferable galactose donor.

The optimal reaction conditions for the transgalactosylation activity of NmLgtB-B were explored by studying the effects of pH, UDP and lactose (5) concentrations on reaction yields (Figure 2), using *p*NP-β-GlcNAc (**15**) as acceptor. Interestingly, with NmLgtB-B higher conversions could be achieved in acidic environment, with a maximum around pH 5.0 (Figure 2A). In the absence of UDP no conversion was observed, suggesting that UDP is required for the transfer reaction and providing the first mechanistic clue that the reactions would proceed *via* UDP-galactose (**2**). Product formation increased gradually when the concentration of UDP was increased, with a plateau reached at 2 mM (Figure 2B). The yield of reaction was measured at different lactose (**5**) concentrations (Figure 2C), establishing that substrate concentrations above 20mM did not afford significant increases in the conversion.

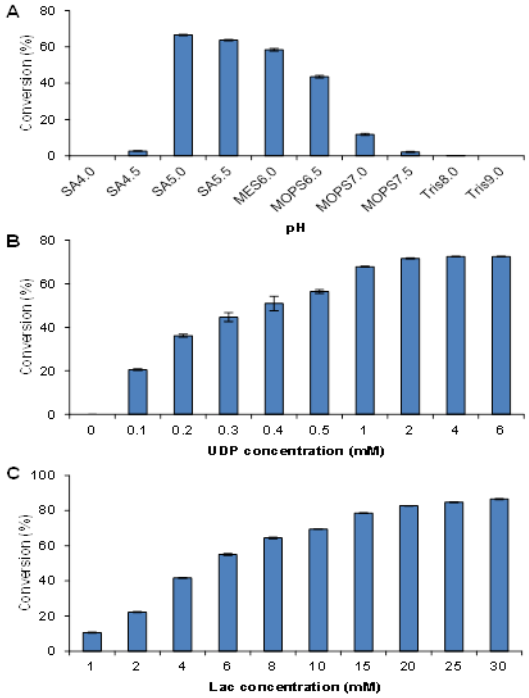
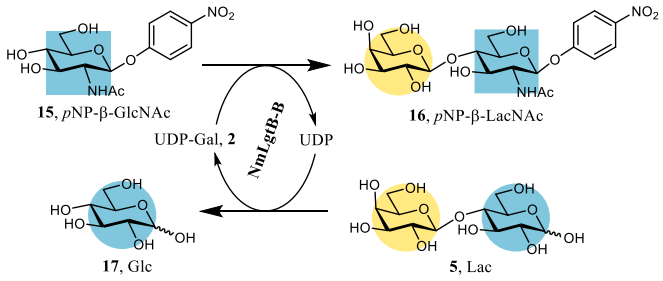


Figure 2. Transgalactosylation reaction condition optimisation. A: pH dependence, using 1mM *p*NP-β-GlcNAc, 5mM UDP and 10mM Lac; B: UDP dependence, using 1mM *p*NP-β-GlcNAc, 10mM Lac and sodium acetate buffer pH 5; C: lactose (**5**) dependence, using 1mM *p*NP-β-GlcNAc, 2mM UDP and sodium acetate buffer pH 5.

The formation of UDP-Gal (**2**) as an intermediate was also verified directly by high resolution mass spectrometry of the crude incubation mixture. In a reaction mixture containing lactose (**5**), UDP and NmLgtB-B in acetate buffer pH5.0 incubated at 37°C overnight it was possible to detect the presence of UDP-Gal (**2**), while a negative control without enzyme did not show the diagnostic peaks (Supporting information, Figure S11).

The practical application of galactosylation from lactose (**5**) was demonstrated by the preparative scale synthesis of *p*NP-β-LacNAc (**16**) (Scheme 4). Initially, a reaction time course was carried out to monitor the reaction process. No significant product increase was observed after 7 hours incubation at 37°C (Supporting information, Figure S12). Hence, preparative scale reaction mixture containing *p*NP -β-GlcNAc (**15**; 51µmole), lactose (**5**; 1.0mmole) and UDP (112

µmole) were incubated with NmLgtB-B at 37°C for 14 hours, which gave 90% conversion as measured by HPLC (Supporting information, Figure S13). The unreacted *p*NP-β-GlcNAc (**15**) was removed by hydrolysis with β-*N*-acetylhexosaminidase, which allowed for the product to be isolated by preparative reverse phase HPLC, yielding *p*NP-β-LacNAc (**16**) in 50% overall yield.



Scheme 4. Preparative synthesis of *p*NP-β-LacNAc (**16**).

Conclusions

The reversibility of the catalytic activity of Leloir β1,4-galactosyltransferases, in particular NmLgtB-B, was demonstrated for the first time by utilising novel imidazolium tagged substrates. This reversibility allowed for the *in situ* formation of UDP-Gal (**2**) from inexpensive lactose (**5**) in the presence of UDP. By adding a further acceptor substrate the overall transfer of galactose from lactose (**5**) to acceptor substrates **11**, **12** and **15** could be observed. The practical applicability of this simple transgalactosylation method was demonstrated by preparative scale synthesis of *p*NP-β-LacNAc (**16**) from lactose using NmLgtB-B as the only biocatalyst. These results are particularly interesting because lactose **5** is a waste product in the milk industry and the current transglycosylation strategy provides opportunities for converting bio-waste into high value products.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This study was funded by the BBSRC, EPSRC and InnovateUK: IBCatalyst programme [BB/M029034/1, BB/M028836/1 (SLF), BB/M028976/1 (MCG), BB/P011462/1 (JvM), BB/M028747/1 (WBT)], Marie Skłodowska-Curie Innovative Training Network [H2020-MSCA-ITN-2014-ETN-642870] and the European Research Council (ERC) COG: 648239 (to MCG) and ProgrES-DVL-788231 (to SLF). We are grateful for the help provided by Dr. Reynard Spiess and Dr. Matt Cliff for HRMS and NMR analyses, respectively.

Notes and references

1 B. Petschacher and B. Nidetzky, *J. Biotechnol.*, 2016, **235**, 61–83.
2 Y. Vandenplas, B. Berger, V. P. Carnielli, J. Ksiazyk, H. Lagström, M. S. Luna, N. Migacheva, J. M. Mosselmans, J. C. Picaud, M. Possner, A. Singhal and M. Wabitsch, *Nutrients*, , DOI:10.3390/nu10091161.
3 L. A. Lasky, *Annu. Rev. Biochem.*, 1995, **64**, 113–140.
4 R. Šardzík, A. P. Green, N. Laurent, P. Both, C. Fontana, J. Voglmeir, M. J. Weissenborn, R. Haddoub, P. Grassi, S. M. Haslam, G. Widmalm and S. L. Flitsch, *J. Am. Chem. Soc.*, 2012, **134**, 4521–4524.
5 Z. Wang, Z. S. Chinoy, S. G. Ambre, W. Peng, R. Mcbride, R. P. De Vries, J. Glushka, J. C. Paulson and G. Boons, *Science (80-.).*, 2013, **341**, 379–384.
6 S. S. Shivatare, S. H. Chang, T. I. Tsai, S. Y. Tseng, V. S. Shivatare, Y. S. Lin, Y. Y. Cheng, C. T. Ren, C. C. D. Lee, S. Pawar, C. S. Tsai, H. W. Shih, Y. F. Zeng, C. H. Liang, P. D. Kwong, D. R. Burton, C. Y. Wu and C. H. Wong, *Nat. Chem.*, 2016, **8**, 338–346.
7 Z. Wu, Y. Liu, C. Ma, L. Li, J. Bai, L. Byrd-Leotis, Y. Lasanajak, Y. Guo, L. Wen, H. Zhu, J. Song, Y. Li, D. A. Steinhauer, D. F. Smith, B. Zhao, X. Chen, W. Guan and P. G. Wang, *Org. Biomol. Chem.*, 2016, **14**, 11106–11116.
8 K. Lau, V. Thon, H. Yu, L. Ding, Y. Chen, M. M. Muthana, D. Wong, R. Huang and X. Chen, *Chem. Commun.*, 2010, **46**, 6066–6068.
9 H. Kunz, J. C. Paulson and C. Unverzagt, *J. Am. Chem. Soc.*, 1990, **112**, 9308–9309.
10 A. Cabanettes, L. Perkams, C. Spies, C. Unverzagt and A. Varrot, *Angew. Chem. Int. Ed. Angew. Chem*, 2018, **10**, 10178–10181.
11 C. Mayer, D. L. Zechel, S. P. Reid, R. A. J. Warren and S. G. Withers, *FEBS Lett.*, 2000, **466**, 40–44.
12 M. J. Hernaiz and D. H. G. Crout, *J. Mol. Catal. - B Enzym.*, 2000, **10**, 403–408.
13 C. H. Wong, R. Wang and Y. Ichikawa, *J. Org. Chem.*, 1992, **57**, 4343–4344.
14 A. Zervosen and L. Elling, *J. Am. Chem. Soc.*, 1996, **118**, 1836–1840.
15 M. M. Muthana, J. Qu, Y. Li, L. Zhang, H. Yu, L. Ding, H. Malekan and X. Chen, *Chem. Commun.*, 2012, **48**, 2728–2730.
16 Y. Ishido, H. Ohi and T. Usui, *J. Carbohydr. Chem.*, 1992, **11**, 553–565.
17 N. Kaftzik, P. Wasserscheid and U. Kragl, *Org. Process Res. Dev.*, 2002, **6**, 553–557.
18 K. M. Koeller and C. H. Wong, *Glycobiology*, 2000, **10**, 1157–1169.
19 G. Reactions, C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I. Lee, L. Li and J. S. Thorson, 2006, **313**, 1291–1295.
20 R. W. Gantt, P. Peltier-Pain, S. Singh, M. Zhou and J. S. Thorson, *Proc. Natl. Acad. Sci.*, 2013, **110**, 7648–7653.
21 A. Gutmann and B. Nidetzky, *Adv. Synth. Catal.*, 2016, **358**, 3600–3609.

22 E. V. Chandrasekaran, J. Xue, J. Xia, R. D. Locke, K. L. Matta and S. Neelamegham, *Biochemistry*, 2008, **47**, 320–330.
23 E. V. Chandrasekaran, J. Xue, J. Xia, R. D. Locke, S. A. Patil, S. Neelamegham and K. L. Matta, *Biochemistry*, 2011, **50**, 9475–9487.
24 J. B. McArthur, H. Yu, N. Tasnima, C. M. Lee, A. J. Fisher and X. Chen, *ACS Chem. Biol.*, 2018, **13**, 1228–1234.
25 P. Both, M. Riese, C. J. Gray, K. Huang, E. G. Pallister, I. Kosov, L. P. Conway, J. Voglmeir and S. L. Flitsch, *Glycobiology*, 2018, **28**, 261–268.
26 T. Okada, H. Ihara, R. Ito, N. Taniguchi and Y. Ikeda, *Glycobiology*, 2009, **19**, 368–374.
27 M. C. Galan, A. T. Tran, K. Bromfield, S. Rabbani and B. Ernst, *Org. Biomol. Chem.*, 2012, **10**, 7091–7097.
28 I. Sittel and M. C. Galan, *Bioorganic Med. Chem. Lett.*, 2015, **25**, 4329–4332.
29 M. C. Galan, A. T. Tran and C. Bernard, *Chem. Commun.*, 2010, **46**, 8968–8970.
30 M. C. Galan, R. A. Jones and A. T. Tran, *Carbohydr. Res.*, 2013, **375**, 35–46.
31 I. Sittel and M. C. Galan, *Org. Biomol. Chem.*, 2017, **15**, 3575–3579.
32 C. He, S. Wang, M. Liu, C. Zhao, S. Xiang and Y. Zeng, *Org. Biomol. Chem.*, 2016, **14**, 1611–1622.