Glycobiology Advance Access published December 23, 2015

Biochemical characterization of an α1,2-colitosyltransferase from *Escherichia coli* O55:H7

Zhigang Wu[†,§], Guohui Zhao[†,§], Tiehai Li[†], Jingyao Qu[†], Wanyi Guan[±], Jiajia Wang[‡], Cheng Ma[†], Xu Li[†], Wei Zhao[‡], Peng G. Wang[†,‡,*], and Lei Li[†,*]

[†]Department of Chemistry and Center for Diagnostics & Therapeutics, Georgia State University, Atlanta, GA 30303

[‡]State Key Laboratory of Medicinal Chemical Biology and College of Pharmacy, Nankai University, Tianjin 30071, China

[±]College of Life Science, Hebei Normal University, Shijiazhuang, Hebei 050024, China

[*]To whom correspondence should be addressed: Phone: +1-4044135607, Email: lli22@gsu.edu (L. Li), or Phone: +1-4044135607, Email: pwang11@gsu.edu (P.G. Wang)

[§]Contributed Equally

The following **Supplementary data** is included:

1. DNA sequence of codon optimized WbgN gene
2. MS spectra of FKP and ColT catalyzed reactions
3. NMR spectra of GDP-colitose
4. NMR spectra of synthesized trisaccharide Colα1,2-Galα1,3-GlcNAc
5. NMR spectra of synthesized Type 1 H-antigen

Abstract

Colitose, also known as 3,6-dideoxy-L-galactose or 3-deoxy-L-fucose, is one of the only five naturally occurred 3,6-dideoxyhexoses. Colitose was found in lipopolysaccharide (LPS) of a number of infectious bacteria, including *Escherichia coli* O55 & O111, *Vibrio cholera* O22 & O139, etc. To date, no colitosyltransferase has been characterized, probably due to the inaccessibility of the sugar donor, GDP-colitose. In this study, starting with chemically prepared colitose, 94.6 mg of GDP-colitose was prepared via a facile and efficient one-pot two-enzyme system involving an L-fucokinase/GDP-L-Fuc pyrophosphorylase (FKP) and an inorganic pyrophosphatase (EcPpA). WbgN, a putative colitosyltransferase from *E. coli* O55:H5 was then cloned, overexpressed, purified and biochemically characterized by using GDP-colitose as a sugar donor. Activity assay and structural identification of the synthetic product clearly demonstrated that *wbgN* encodes an α1,2-colitosyltransferase (α1,2-ColT). Biophysical study showed that WbgN does not require metal ion, and is highly active at pH 7.5 to 9.0. In addition, acceptor specificity study indicated that WbgN exclusively recognize lacto-*N*-biose (Galβ1,3-GlcNAc). Most interestingly, it was found that WbgN exhibits similar activity toward GDP-L-Fuc (k_{cat}/K_m = 9.2 min^{-1} mM^{-1}) as that toward GDP-colitose (k_{cat}/K_m = 12 min^{-1} mM^{-1}). Finally, taking advantage of this, type 1 H-antigen was successfully synthesized in preparative scale.

Keywords

Colitosyltransferase; GDP-colitose; *Escherichia coli* O55; H-antigen; Synthesis

Introduction

Deoxysugars are widely found in natural products of plant and fungi, and secondary metabolites of bacteria, which are common sources for antibiotics or anticancer agents, e.g., anthracyclines, avermectins, enediynes, angucyclines, etc. A great number of bacterial lipopolysaccharide (LPS) and capsular polysaccharides also contains diverse deoxysugars, for example, L-rhamnose (6-deoxy-L-mannose, found in nearly one third of *Escherichia coli* LPS), 6-deoxy-L-talose, L-fucose (6-deoxy-L-galactose), D-FucNAc (6-deoxy-GalNAc), etc. Deoxysugars are not only structural components of these bioactive molecules, but in many cases are critical for the biological recognition of the mother compounds, serving as ligands for cell-cell interaction or as targets for toxins, antibodies, and microorganisms (Weymouth-Wilson, A.C. 1997).

Among the large number of monosaccharide residues found as components of O-antigen of bacteria LPS, 3,6-dideoxyhexoses have drawn special attentions due to their highly immunogenic characteristics. Of the eight possible stereoisomers of 3,6-dideoxyhexoses, only five exist naturally, including ascarylose, abequose, paratose, tyvelose and colitose. In particular, colitose was found in a number of pathogenic bacteria, for example, *E. coli* O55 & O111, *Vibrio cholera* O22 & O139, *Salmonella enterica* O35 & O50, *Yersinia pseudotuberculosis* O6, O7 & O10, etc (Kenyon, J.J., De Castro, C., et al. 2011, Knirel, Y.A., Senchenkova, S.N., et al. 1998, Knirel, Y.A., Senchenkova, S.N., et al. 1996, Ovchinnikova, O.G., Kocharova, N.A., et al. 2007, Senchenkova, S.N., Shashkov, A.S., et al. 1997). In these cases, colitose locates at the termini of O-antigens, serving as a special antigenic component of LPS (Luderitz, O., Staub, A.M., et al. 1958, Luderitz, O., Westphal, O., et al. 1960, Westphal, O., Luderitz, O., et al. 1959). The presence of colitose in pathogenic *E. coli* may enhance the lipophilic character of endotoxin LPS (Medearis, D.N., Jr., Camitta, B.M., et al. 1968). Colitose was also considered an essential component of specific ligands of certain LPS-binding lectins, e.g., houseshoe crab tachylectin-4 (specifically binds to *E. coli* O111 LPS) (Saito, T., Hatada, M., et al. 1997). In addition, aerobic marine heterotrophic bacteria of

the genera *Alteromonas*, *Echinicola* and *Pseudoalteromonas* also contain colitose in their O-antigens (Gorshkova, N.M., Gorshkova, R.P., et al. 2001, Muldoon, J., Perepelov, A.V., et al. 2001, Silipo, A., Molinaro, A., et al. 2005, Tomshich, S.V., Kokoulin, M.S., et al. 2015). Given the emerging importance, chemical synthesis had been performed to prepare colitose and colitose-containing oligosaccharides for vaccine development (Calin, O., Eller, S., et al. 2013, Calin, O., Pragani, R., et al. 2012, Ruttens, B. and Kovac, P. 2004, Senchenkova, S.N., Shashkov, A.S., et al. 1997, Turek, D., Sundgren, A., et al. 2006). In contrast, there are limited studies on the biosynthesis of colitose-containing O-antigen, especially on characterization of colitosyltransferases that are responsible for the transfer of colitose. Lack of efficient approaches for the access of sufficient amounts of the sugar donor, GDP-colitose, has been a roadblock.

Different from the other four 3,6-dideoxyhexoses (ascarylose, abequose, paratose and tyvelose) that are derived from CDP-glucose, the biosynthetic pathway of colitose (also known as 3,6-dideoxy-L-galactose or 3-deoxy-L-fucose) begins from mannose-1-phosphate, and share a few enzyme-catalyzed steps with that of GDP-L-galactose and GDP-L-Fuc. This is understandable when concerning structural similarities between colitose and L-galactose/L-fucose (L-Fuc) (**Fig. 1**). Functional assignment and biochemical characterization (Alam, J., Beyer, N., et al. 2004, Beyer, N., Alam, J., et al. 2003) of putative enzymes in GDP-colitose biosynthesis gene clusters revealed a similar pathway as that of other 3,6-dideoxyhexoses. As depicted in **Figure 1**, the biosynthesis of GDP-colitose begins with GDP-mannose pyrophosphorylase (GMP, or ColE)-catalyzed pyrophosphorylation of mannose-1-phosphate to form GDP-mannose, followed by GDP-mannose 4,6-dehydratase (GMD, or ColB)-catalyzed intramolecular oxidoreduction of GDP-mannose to afford GDP-4-keto-6-deoxy-mannose (GKDM), an key intermediate for biosynthesis of both GDP-colitose and GDP-L-Fuc. Subsequently, GKDM was converted into GDP-4-keto-3,6-dideoxy-mannose (GKDDM) by GDP-4-keto-6-deoxy-mannose-3-dehydrase (ColD)-catalyzed C-3 deoxygenation in the presence of L-glutamate and cofactor pyridoxal-5'-phosphate (PLP) (Alam, J., Beyer, N., et al. 2004). Finally, GDP-colitose is generated by GDP-colitose synthase (ColC), a bifunctional enzyme catalyzing the C-5 epimerization of GKDDM and the subsequent C-4 keto reduction

(Alam, J., Beyer, N., et al. 2004). The last step requires equal moles of NAD(P)H as that of GKDDM. Given the involvement of multiple steps and cofactors, and the structural similarities between final product and intermediates, one can imagine it must be complicated to produce and purify milligrams of GDP-colitose *via* such a biosynthetic pathway (Heath, E.C. and Elbein, A.D. 1962, Isshiki, Y., Kondo, S., et al. 1996). A facile and efficient approach is thus in demand for large scale preparation of GDP-colitose.

E. coli O55:H7 belongs to enteropathogenic *E. coli* (EPEC), a group of *E. coli* isolates among the major causes of acute and prolonged infantile diarrhea in developing countries (Ochoa, T.J. and Contreras, C.A. 2011). Genomic analysis revealed the *E. coli* O55:H7 is closely related to the most notorious *E. coli* isolate O157:H7, which accounts for majority of *E. coli* infections in North America every year, leading to severe acute haemorrhagic diarrhea, complication hemolytic uremic syndrome (HUS), even death (Sodha, S., Heiman, K., et al. 2015). The O-antigen repeating unit of *E. coli* O55:H7 contains five monosaccharide residues, with a structure of Colα1,2-Galβ1,3-GlcNAcβ1,3-Galα1,3GalNAc (**Fig. 2**) (Lindberg, B., Lindh, F., et al. 1981). The O-antigen biosynthesis gene cluster has been sequenced (Wang, L., Huskic, S., et al. 2002), unveiling four putative glycosyltransferase genes (*wbgM*, *wbgN*, *wbgO* and *wbgP*) for the assembly of the repeating unit (**Fig. 2**). Among these four genes, we have biochemically demonstrated that *wbgO* encodes an β1,3-galactosyltransferase (Liu, X.W., Xia, C., et al. 2009). WbgN shares 23% sequence identity and 47% sequence similarity with human FUT2 (α1,2-fucosyltransferase), and was thought to the colitosyltransferases (ColT). In this study, we cloned the *wbgN* gene, overexpressed and purified the recombinant protein, and performed biochemical characterization using chemoenzymatically prepared GDP-colitose as sugar donor.

Results and Discussion

Chemoenzymatic synthesis of GDP-colitose

De novo biosynthetic pathways of GDP-L-Fuc and GDP-colitose share the same initial steps including generation of intermediates mannose-1-phosphate, GDP-mannose and GKDM (**Fig. 1**). The complexity of the pathways and lack of cost-effective approaches to access intermediates greatly restricted large scale preparation of these valuable sugar donors. In efforts to develop facile and cost-effective approaches for large scale synthesis of GDP-L-Fuc and derivatives, we and others exploited L-fucokinase/GDP-L-Fuc pyrophosphorylase (FKP) (Wang, W., Hu, T., et al. 2009, Yi, W., Liu, X., et al. 2009, Zhao, G., Guan, W., et al. 2010), a bifunctional enzyme found in the fucosylation salvage pathway of a human intestinal bacterium, *Bacteroides fragilis* 9343 (Coyne, M.J., Reinap, B., et al. 2005). As depicted in **Figure 3**, the salvage pathway involves 1) L-fucokinase activity of FKP catalyzed formation of L-Fuc-1-phosphate (L-Fuc-1-P) from L-Fuc in presence of ATP, and 2) GDP-L-Fuc pyrophosphorylase activity of FKP catalyzed generation of GDP-L-Fuc and byproduct pyrophosphate in presence of GTP. Using a one-pot two-enzyme system which contains FKP and an inorganic pyrophosphatase (drive reaction forward by digesting byproduct pyrophosphate), GDP-L-Fuc and a number of C-5 modified derivatives including GDP-L-galactose were successfully prepared in large scale (Wang, W., Hu, T., et al. 2009, Yi, W., Liu, X., et al. 2009). In another example, a C-2 modified derivative, GDP-2-deoxy-2-fluoro-L-fucose (GDP-Fuc2F), was also efficiently synthesized (Rillahan, C.D., Antonopoulos, A., et al. 2012). These reports indicated that FKP is very promiscuous towards L-Fuc and derivatives.

To explore whether FKP could tolerate colitose (3-deoxy-L-fucose) as a substrate, we chemically prepared colitose starting with L-Fuc as previously reported (Ruttens, B. and Kovac, P. 2004). Activity assays of FKP towards colitose were carried out in presence of ATP (for fucokinase activity) or both ATP and GTP (for pyrophosphorylase activity), with positive controls using L-Fuc as substrate instead of colitose. As shown in **Figure 4**, after incubation with FKP, both spots corresponding to colitose and ATP

weakened, whereas two new spots appeared (**Fig. 4**, lane 5). Mass spectroscopy (MS) analysis of the reaction mixture (**Supplementary Data 2**) showed two peaks at m/z 227.0290 and 426.0151, corresponding to colitose-1-phoshate and ADP. In addition, when GTP was added to above reaction mixture, the spot corresponding to colitose-1-phoshate disappeared, yielding a new spot (**Fig. 4**, lane 6) of GDP-colitose, which is later confirmed by MS analysis with a m/z peak of 572.0705 $[M-H]^-$. These results clearly indicate that FKP can well accept colitose and efficiently catalyze the formation of GDP-colitose. To confirm the structure, large scale synthesis of GDP-colitose was carried out using the one-pot two-enzyme system as illustrated in **Figure 3**. After P2 purification, 94.6 mg of GDP-colitose was obtained, with a total yield of 83%. The chemical structure was confirmed by high resolution MS and NMR (see **Supplementary Data 3** for spectra). Such a one-pot two-enzyme system provides a facile, efficient and cost-affordable approach for the access of large scales of GDP-colitose.

wbgN encodes an α1,2-colitosyltransferase

According to sequence similarities, WbgN was classified into glycosyltransferase family 11 (www.cazy.org), a group of enzymes with known activities of α1,2-L-fucosyltransferase (α1,2-FucT) and α1,3-L-fucosyltransferase (α1,3-FucT). In addition, amino acid sequence of WbgN shares 37% identity and 57% similarity with a recently identified α1,2-FucT, WbgL, which is involved in the biosynthesis of *E. coli* O126 O-antigen (Engels, L. and Elling, L. 2014). Considering the structure similarities between O-antigen of *E. coli* O55:H5 (contains Colα1,2-Galβ1,3-GlcNAc motif) and O126 (contains L-Fucα1,2-Galβ1,3-GlcNAc motif), it is assumed that *wbgN* encodes the α1,2-colitosyltransferase that is responsible for transfer of colitose from GDP-colitose to the Gal residue of Galβ1,3-GlcNAc via an α1,2-linkage. To test this hypothesis, the *wbgN* gene was codon optimized (host *E. coli*), synthesized and cloned into pQE80 L for expression. Even though induction was performed under a low temperature (16 °C) with minimum amounts of inducers (0.1 mM IPTG), majority of WbgN proteins tend to precipitate (**Fig. 5**, lane 4), further optimization to increase soluble form proteins failed. Such phenomena were found

common among O-antigen synthesis enzymes (especially fucosyltransferases) (Engels, L. and Elling, L. 2014, Li, M., Liu, X.W., et al. 2008, Li, M., Shen, J., et al. 2008), which are thought membrane associated and with low solubility (Raetz, C.R. and Whitfield, C. 2002). Nevertheless, 1.2 mg of soluble WbgN was purified from 1 L of cell culture. As shown in **Figure 5** (Lane 7), the N-terminal His-tagged WbgN has an apparent MW of 30 kDa on SDS-PAGE, slightly lower than the calculated value (35 kDa). This is normal for membrane associated hydrophobic proteins (Rath, A., Glibowicka, M., et al. 2009). Western blot using an anti-His antibody as a primary antibody confirmed the result. Alternatively, fusion protein strategy has been extensively used to increase the solubility of α1,2-FucTs (Engels, L. and Elling, L. 2014, Li, M., Liu, X.W., et al. 2008, Li, M., Shen, J., et al. 2008), which usually resulted in greatly decreased activity. Therefore, it was not applied in this study.

To reveal the function of WbgN, a panel of sugar nucleotides (GDP-colitose, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc) were tested using Lacto-*N*-biose (LNB, Galβ1,3-GlcNAc) as an acceptor. MS analysis of the reaction with GDP-colitose as donor found a new peak at m/z of 536.1933 (**Supplementary Data 2**), corresponding to the trisaccharide product Colα1,2-Galβ1,3-GlcNAc [M+Na]. No new peaks were found when using other sugar nucleotides as donors. Thus, the donor specificity study indicated that *wbgN* encodes colitosyltransferase, consistent with bioinformatics analysis. To investigate whether WbgN catalyzes the formation of the α1,2-linkage, large scale preparation of the trisaccharide were performed. Briefly, a 5 mL reaction system was set up containing 30 mg of LNB (16 mM), 1.25 equiv. of GDP-colitose, WbgN (20 μg) and 50 mM Tris-HCl (pH 8.0) buffer. After LNB was totally converted (monitored by TLC), the trisaccharide product was isolated by gel filtration chromatography with a total yield of 92% (37 mg), The structure of trisaccharide was confirmed by NMR spectroscopy including ^{1}H, ^{1}H-^{1}H-COSY, ^{13}C, HSQC and HMBC (**Supplementary Data 4**). The α-linkage of Col with Gal was identified by the coupling constant (J_{H1-H2} = 4.8 Hz) of colitose. Moreover, HMBC revealed a good correlation of the anomeric carbon (C-1) of colitose with H-2 of Gal, which indicated 1,2-linked colitose with Gal. Collectively, these results provide solid evidences that *wbgN* encodes an α1,2-ColT

that transfers colitose onto LNB via a α1,2-linkage. To the best of our knowledge, WbgN is the first biochemically characterized colitosyltransferase.

Biophysical Properties of WbgN

Study on pH optimum with LNB as an acceptor revealed that WbgN is highly active with in a broad pH spectrum ranging from pH 7.5 to 9.0, and exhibits around 60% activity at a higher basic condition (pH 10.2) (**Fig. 6A**). Sequence analysis showed that WbgN doesn't contain a DXD motif, a conserved amino acid sequence exists in all glycosyltransferase superfamily A (GT-A) members (Breton, C., Šnajdrová, L., et al. 2006). Such DXD motifs play a role in coordinating the binding of sugar nucleotides in the active site with the requirement of certain divalent metal cation. This is consistent with activity assay results showing metal ion is not necessary for WbgN. To further investigate the metal ion dependence, a panel of divalent metal cations and EDTA were tested with positive (no additive) and negative (no enzyme) controls. As shown in **Figure 6B**, WbgN exhibited highest activities in the presence or the absence of EDTA, indicating that a metal ion is not required for the activity of WbgN. In contrast, the addition of metal ions dramatically decreased its activity to 0- 40%. It is thus assumed that instead of GT-A, WbgN belongs to GT-B superfamily, which is in accordance with similar fucosyltransferases WbgL (Engels, L. and Elling, L. 2014), WbsJ (Li, M., Liu, X.W., et al. 2008), and WbiQ (Pettit, N., Styslinger, T., et al. 2010).

WbgN can efficiently use GDP-Fuc as a sugar donor

Concerning the sequence similarities between WbgN and a number of α1,2-FucT (e.g., WbgL from *E. coli* O126, WbsJ from *E. coli* O128 and WbwK from *E. coli* O86), as well as the structural similarities between sugar donors GDP-colitose and GDP-L-Fuc, it is of great interest to reveal whether WbgN could recognize GDP-Fuc. To this end, reaction was conducted using GDP-Fuc as the sugar donor and LNB as the acceptor. Briefly, a 50 μL reaction was set up containing 50 mM Tris-HCl (pH 8.0), 5 mM acceptor

Galβ1,3-GlcNAc, 5 mM of acceptor GDP-L-Fuc, and 50 μg of recombinant WbgN. The reaction was allowed to proceed at 37 °C for 2 h before subjecting to MS analysis. Two new peaks at m/z of 530.2071 and 552.1888 were found on MS profile, corresponding to the trisaccharide product L-Fucα1,2-Galβ1,3-GlcNAc $[M+H]^+$, and $[M+Na]^+$. In addition, MS peak corresponding to the acceptor LNB disappeared (**Supplementary Data 2**). This result clearly indicated that WbgN can employ GDP-L-Fuc as a sugar donor efficiently. To further investigate the donor preference of WbgN, kinetic characterization were performed. As listed in **Table 1**, WbgN has a lower K_m value toward GDP-colitose (0.76 ± 0.18 mM) than that toward GDP-L-Fuc (1.3 ± 0.42 mM), implying that GDP-colitose has a higher enzyme affinity. In contrast, a higher turnover rate was found toward GDP-L-Fuc (k_{cat} = 12 ± 1.5 min^{-1}) than that toward GDP-colitose (k_{cat} = 8.8 ± 0.8 min^{-1}), indicating that WbgN-catalyzed reaction is more favorable towards the trisaccharide product containing L-Fuc. Collectively, WbgN exhibits comparable activities toward both sugar nucleotide donors, with a k_{cat}/K_m value of 12 min^{-1} mM^{-1} toward GDP-colitose, and that of 9.2 min^{-1} mM^{-1} toward GDP-L-Fuc.

Among all colitose-containing bacterial polysaccharide identified so far (http://csdb.glycoscience.ru/bacterial/), none contains L-Fuc residue(s). This could be caused by 1) lack of sugar donor GDP-L-Fuc, or/and 2) lack of corresponding FucT(s). Genetic analysis revealed that GDP-L-Fuc synthase (GFS) encoding gene (*fcl*) is always missing or inactivated in gene clusters responsible for the biosynthesis of these bacterial polysaccharides (Bastin, D.A. and Reeves, P.R. 1995, Cunneen, M.M., Pacinelli, E., et al. 2011, Kenyon, J.J., De Castro, C., et al. 2011, Wang, L., Huskic, S., et al. 2002, Wang, L. and Reeves, P.R. 2000, Yamasaki, S., Shimizu, T., et al. 1999), supporting the first assumption. On the other hand, our results clearly showed that, in *E. coli* O55:H7, ColT can also function as FucT (or possibly is intrinsic FucT), excluding the second assumption. This phenomenon could be common among other colitose-containing bacteria, as putative ColTs usually share significant sequence similarities with FucTs (Bastin, D.A. and Reeves, P.R. 1995, Cunneen, M.M., Pacinelli, E., et al. 2011, Kenyon, J.J., De

Castro, C., et al. 2011, Wang, L., Huskic, S., et al. 2002, Wang, L. and Reeves, P.R. 2000, Yamasaki, S., Shimizu, T., et al. 1999). Biochemical characterizations are needed to confirm such hypothesis.

WbgN is specific for Lacto-N-biose

The purified WbgN was used to investigate its substrate specificity towards a group of mono- and disaccharide acceptors. As shown in **Table 2**, WbgN is highly specific for LNB in the presence of either GDP-colitose or GDP-L-Fuc, and cannot tolerate any other acceptors tested, including galacto-*N*-biose (GNB, Galβ1,3-GalNAc), LacNAc (Galβ1,4-GlcNAc) and lactose. These results indicated that not only the terminal Gal, but also the adjacent sugar residue and the way they coupled together are involved in the enzyme-acceptor binding. Such a strict acceptor preference can be applied specifically for the preparation of type 1 H-antigen (L-Fucα1,2-Galβ1,3-GlcNAc), or complex glycans that contain the antigen. In a test trial, 40 mg (95% yields) of type 1 H-antigen was synthesized in a WbgN-catalyzed reaction, starting with LNB and GDP-L-Fuc (see **Supplementary Data 5** for NMR characterization). Taken together, the acceptor pattern shown in the assay is in agreement with the proposed function of WbgN in the biosynthesis of the *E. coli* O55:H7 O-antigen.

Conclusion

In summary, taken advantage of relaxed substrate specificity of FKP, we have developed a facile and cost-efficient system for enzymatic synthesis of rare sugar nucleotide GDP-colitose in preparative scale. Using the synthesized GDP-colitose, we have biochemically characterized the first ColT (WbgN) from *E. coli* O55:H7. WbgN was found high active in a pH range of 7.5-9.0, and a metal ion is not required for its activity. Interestingly, WbgN exhibited similar activity towards both GDP-colitose and GDP-L-Fuc, taking together with the fact that ColTs and FucTs share significant sequence similarities, it is possible that many ColTs may have FucT activities. Acceptor substrate specificity study revealed that WbgN is

specific for Lacto-*N*-biose, which may found great use in the preparation of type 1 H-antigen-containing complex glycans.

Materials and Methods

Materials

E. coli DH5α competent cells were purchased from Life Technologies (Carlsbad, CA), *E. coli* BL21(DE3) competent cells and Nitrocellulose membrane were purchased from New England BioLabs (Ipswich, MA). Vector pQE80L and Ni-NTA agarose were from Qiagen (Valencia, CA). Anti-His monoclonal antibody (from mouse) was from Sigma (St. Louis, MO). HRP-linked anti-mouse IgG, and ECL plus Western Blot detection reagents were purchased from GE healthcare (Piscataway, NJ). Fast Digest restriction enzymes were purchased from Fermentas (Glen Burnie, MD). Shrimp alkaline phosphatase was from New England Biolabs (Ipswich, MA). GDP-Fuc, Galβ1,3-GlcNAc, Galβ1,3-GalNAc, Galβ1,4-GlcNAc (LacNAc), FKP and EcPpA were prepare previously in in our group (Li, L., Liu, Y., et al. 2015, Li, L., Liu, Y., et al. 2013, Zhao, G., Guan, W., et al. 2010). All other chemicals and solvents were purchased from Sigma-Aldrich.

Plasmid construction, protein expression and enzyme purification

Full length putative α1,2-L-colitosyltransferase gene (*wbgN*) from *E. coli* O55:H7 (GeneBank: 18266395) was codon optimized (target host, *E. coli*) and synthesized (Qinglan Bio, China) together with 5'-terminal *BamH* I restriction sites and 3'-terminal *Hind* III restriction sites (see **Supplementary Data 1** for optimized gene sequence). The gene was then digested with *BamH*I and *Hind*III, and inserted into pQE80L pre-treated with same restriction enzymes, yielding an expression plasmid pQE-WbgN. After DNA sequencing, the plasmid was transformed into *E. coli* BL21(DE3) for overexpression.

E. coli BL21 (DE3) harboring the pQE-WbgN was grown at 37 °C in 1 L of LB medium (Lennox) with 100 μg/mL Ampicillin until OD_{600} reached 0.6-0.8. After cooling the culture on ice for 20 min,

isopropyl 1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 0.1 mM. Expression was allowed to proceed at 16 °C for 20 h. Cells were harvested by brief centrifugation and stored at -20 °C until use.

To purify soluble recombinant enzymes, cells were resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 0.3 M NaCl, 5 mM imidazole, 0.2% Triton X-100, 10% glycerol), and disrupted by a Microfluidics M-110P Homogenizer. The lysate was cleared by centrifugation (15,000 × *g*, 30 min, 4 °C) and loaded onto a 3 mL Ni-NTA gravity column preequilibrated with buffer B (20 mM Tris-HCl, pH 8.0, 0.3 M NaCl, 5 mM imidazole, 10% glycerol). The column was washed with 200 mL of buffer B and 200 mL of buffer C (20 mM Tris-HCl, pH 8.0, 0.3 M NaCl, 25 mM imidazole, 10% glycerol). The protein was finally eluted with buffer D (20 mM Tris-HCl, pH 8.0, 0.3 M NaCl, 250 mM imidazole, 10% glycerol) and desalted against buffer E (50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 20% glycerol) for long term storage at -20 °C.

Chemoenzymatic synthesis of GDP-colitose and purification

To test the fucokinase activity of FKP towards colitose, reactions were carried out in a 50 μL system, containing 100 mM of Tris-HCl (pH 8.0), 20 mM of colitose or L-Fuc (as positive control), 20 mM of ATP, 20 mM of $MgCl_2$, and 10 μg of purified FKP. To test the GDP-L-Fuc pyrophosphorylase activity, 20 mM of GTP and 2 μg of EcPpA were added to the above reaction mixtures. Reactions were incubated at 37 °C for 1 h, and analyzed by MS and TLC (developing solvent: isopropanol/ammonium acetate (1 M)/Acetic acid = 7:3:0.1 (v/v/v)). The compounds on TLC plates were stained with anisaldehyde/acetic acid/H_2SO_4/H_2O = 7:3:10:27 (v/v/v/v), and visualized by brief heating. For large scale GDP-colitose synthesis, a one-pot two-enzyme system similar as described before was used (Zhao, G., Guan, W., et al. 2010). Briefly, a reaction was carried out at 37 °C for 5 h in a total volume of 10 mL in 100 mM Tris-HCl (pH 8.0) buffer, containing 20 mM colitose, 22 mM ATP, 22 mM GTP, 20 mM $MgCl_2$, 20 μg/mL FKP and 2 μg/mL EcPpA. The reaction was monitored by TLC. To purify GDP-colitose, the reaction was

quenched by the addition of equal volumes of cold ethanol. The mixture was placed on ice for 30 min, and centrifuged to remove debris. The supernatant was then concentrated and loaded onto bio-gel P2 for separation. The fractions contain pure GDP-colitose (detected by TLC) were then concentrated and lyophilized for MS and NMR analysis. The fractions contain impure GDP-colitose were also collected and treated with shrimp alkaline phosphatase (New England Biolabs) to digest ADP byproduct following instructions (100 U, 30 °C for 5 h), and then subject to bio-gel P2 for further separation.

GDP-L-Colitose: ^{1}H NMR (400MHz, D_2O) δ: 8.00 (s, 1 H), 5.83 (d, J = 6.0 Hz,1 H), 4.82 (t, J = 8.0 Hz, 1 H), 4.43 (t, J = 4.0 Hz, 1 H), 4.25 (s,1 H), 4.10 (d, J = 4.4 Hz,1 H), 3.56 - 3.76 (m, 4 H), 2.07- 2.11 (m, 1 H), 1.56-1.63 (m, 1 H), 1.07 (d, J = 6.4 Hz, 3 H); ^{13}C NMR (100MHz, D_2O) δ:172.89, 158.97, 153.89, 140.84, 116.23, 100.34,100.28, 86.73, 83.83, 83.74, 74.70, 73.45, 70.42, 67.93, 36.24, 15.57; ESI HRMS: $[M-H]^-$ calcd for 572.0946, found: 572.0941**.**

Activity assay of WbgN

Briefly, a 50 μL reaction was set up containing 50 mM Tris-HCl (pH 8.0), 5 mM acceptor Galβ1,3-GlcNAc, 5 mM of acceptor GDP-colitose, and 20 μg of recombinant WbgN. The reaction was performed at 37 °C for 1 h, and quenched by adding equal volume of ice cold ethanol. After brief centrifugation, the supernatant was analyzed by mass spectrometry.

Kinetics study

To study the kinetics of WbgN towards GDP-colitose and GDP-L-Fuc, reactions were performed at 37 °C for 10 min in a reaction buffer containing 50 mM Tris-HCl (pH 8.0), 5 mM Galβ1,3-GlcNAc, 5.25 μM recombinant WbgN and varied concentrations of GDP-colitose or GDP-L-Fuc (0.25, 0.5, 0.75, 1.0, 2.0, and 4.0 mM). To study the kinetics of WbgN towards an acceptor Galβ1,3-GlcNAc, reactions were performed at 37 °C for 10 min in a reaction buffer containing 20 mM Tris-HCl (pH 8.0), 5 mM GDP-colitose, 18 μg recombinant WbgN and varied concentrations of Galβ1,3-GlcNAc (0.25, 0.5, 0.75, 1.0,

2.0, and 4.0 mM). Reactions were quenched immediately by addition of equal volume of ice-cold ethanol and subject to capillary electrophoresis (CE) analysis using Beckman Coulter P/ACE MDQ Capillary electrophoresis to detect substrate GDP-colitose or GDP-Fuc and by-product GDP. CE conditions were as follows: 75 μm i.d. capillary, 25 KV/170 μA, 5 s vacuum injection, with monitoring at 254 nm and 50 mM sodium tetraborate running buffer, pH 9.4. The parameters K_m and V_{max} were obtained by Lineweaver-Burk plots of substrate concentration-initial velocity.

Biophysical property study

To study pH dependence of WbgN, reactions were carried out in a total volume of 50 μL, containing 20 μg of recombinant WbgN, 5 mM GDP-colitose, 5 mM Galβ1,3-GlcNAc, and 50 mM of different buffer (PBS, pH 6.5; Tris-HCl, pH 7.5; Tris-HCl, pH 7.5; Glycine-NaOH, pH 9.0; Glycine-NaOH, 10.2). To study metal ion dependence, reactions were carried out in a total volume of 50 μL, containing 50 mM Tris-HCl, pH 8.0, 10 μg of recombinant WbgN, 5 mM GDP-colitose, 5 mM Galβ1,3-GlcNAc, and 10 mM of different additions (EDTA, None, $MgCl_2$, $MnCl_2$, $CoCl_2$, $CaCl_2$, $ZnCl_2$ or $CuCl_2$). A negative control was also carried out containing EDTA but no acceptor Galβ1,3-GalNAc. All reactions were performed in triplicate. Reactions were incubated at 37 °C for 60 min, quenched and analyzed by CE as described above.

Acceptor specificity study

Reactions were carried out in a total volume of 50 μL, containing 50 mM Tris-HCl, pH 8.0, 10 μg of recombinant WbgN, 5 mM GDP-colitose, and 2.5 mM of different acceptors, including galactose (Gal), Galβ1,3GlcNAc, Galβ1,4GlcNAc (LacNAc), Galβ1,3GalNAc, and Galβ1,4Glc (Lac). All reactions were performed in triplicate. Reactions were incubated at 37 °C for 60 min, quenched and analyzed by CE as described above.

Enzymatic synthesis of Colα1,2-Galβ1,3-GlcNAc and Type 1 H-antigen

In a 15 mL centrifuge tube, reactions were carried out in a total volume of 5 mL, containing 50 mM Tris-HCl, pH 8.0, 300 μg of recombinant WbgN, 20 mM GDP-colitose (or GDP-L-Fuc for the synthesis of type 1 H-antigen), and 16 mM of Galβ1,3-GlcNAc. The reactions were allowed to proceed at 37 °C and monitored by TLC (developing solvent: isopropanol: NH_4OH: H_2O = 7: 3: 2, v/v/v). After Galβ1,3-GlcNAc was totally converted to products, reactions were quenched by the addition of equal volumes of ice-cold ethanol, followed by centrifugation at 4 °C for 10 min (10,000 × *g*). The supernatant was ten concentrated and subject for separation by P2 gel filtration. The fractions contain pure products were pooled, lyophilized, and stored at -20 °C until chemical characterization. The fractions contain impure products were also collected and treated with shrimp alkaline phosphatase (New England Biolabs) to digest GDP byproduct following instructions (50 U, 30 °C for 5 h), and then subject to bio-gel P2 for further separation. NMR and ESI HRMS data of trisaccharides (see **Supplementary Data 4**, **5** for spectra):

Colα1,2-Galβ1,3-GlcNAc: ^{1}H NMR (D_2O, 400 MHz):*δ* 1.16 (t, *J* = 5.6 Hz 3 H, *CH*$_3$), 1.76-1.87 (m, 1 H, H″-3), 1.93-1.96 (m, 1 H, H″-3), 2.07 (s, 3 H, NHCO*CH*$_3$), 3.52-3.57 (m, 1 H, H′-2), 3.63 (dd, *J* = 4.8 Hz, *J* = 6.8 Hz,1 H, H″-2), 3.70-3.71(m, 1 H, H-2), 3.73-3.76 (m, 2 H), 3.79-3.83 (m, 4 H), 3.90-3.93 (m, 2 H), 3.99-4.03 (m, 1 H), 4.15-4.28 (m, 2 H), 4.63-4.71 (m, 2.2 H, H″-5, H′-1, H-1-β), 5.12 (m, 1.8 H, H-1-α, H″-1); ^{13}C NMR (D_2O, 100 MHz): *δ* 15.25 (*CH*$_3$), 21.79 (NHCO*CH*$_3$), 32.74 (C″-3), 53.63, 60.43, 60.93, 63.03, 66.05, 68.29, 68.59, 69.01, 71.14 (C″-2), 73.48 (C-2), 74.86, 74.98, 76.41 (C′-2), 90.68 (C-1-α), 95.46 (C-1-β), 98.73 (C″-1), 100.12 (C′-1), 173.55 (NH*CO*CH_3); ESI HRMS: m/z calcd for $[C_{20}H_{35}NO_{14}+Na]^+$, 536.1955, found: 536.1933

Fucα1,2-Galβ1,3-GlcNAc: ^{1}H NMR (D_2O, 400 MHz): *δ* 1.23 (t, *J* = 5.6 Hz, 3 H, *CH*$_3$), 2.08 (s, 3 H, NHCO*CH*$_3$), 3.52-3.65 (m, 2 H, H′-2, H″-2), 3.68-3.90 (m, 9 H), 4.14-4.33 (m, 2 H), 4.63-4.70 (m, 2.3 H, H″-5, H′-1, H-1-β), 5.12 (d, *J* = 2.8 Hz, 0.7 H, H-1-α), 5.20 (d, *J* = 4.0 Hz, 1 H, H″-1); ^{13}C NMR (D_2O, 100 MHz): *δ* 14.41 (*CH*$_3$), 21.06 (NHCO*CH*$_3$), 52.87, 59.69, 60.18, 60.23, 65.60, 67.22, 67.85, 68.27,

68.66, 70.41, 70.97, 72.68, 74.14, 75.76, 89.94 (C-1-α), 98.58 (C″-1), 99.25 (C-1-β), 99.34 (C′-1), 172.79 (NH*CO*CH_3); ESI HRMS: m/z calcd for $[C_{20}H_{35}NO_{15} + Na]^+$ 552.1904, found: 552.1888.

Supplementary Data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

This work was supported by National Institute of General Medical Sciences (U01GM116263 to P.G. Wang & L. Li, R01GM085267 to P.G. Wang) and National Institute of Allergy and Infectious Diseases (R01AI083754 to P.G. Wang).

Abbreviations

ADP, adenine 5'-diphosphate; ATP, adenine 5'-triphosphate; CDP, cytosine 5'-diphosphate; Gal, galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcNAc, *N*-acetylglucosamine; Col, colitose; GDP; guanosine 5'-diphosphate; GDP-L-Fuc, GDP-L-fucose; GNB, Galacto-*N*-biose; GTP, guanosine 5'-triphosphate; L-Fuc, L-fucose; LNB, Lacto-*N*-biose.

Conflict of interest

None declared.

References

Alam J, Beyer N, Liu HW. 2004. Biosynthesis of colitose: expression, purification, and mechanistic characterization of GDP-4-keto-6-deoxy-D-mannose-3-dehydrase (ColD) and GDP-L-colitose synthase (ColC). *Biochemistry*, 43:16450-16460.

Bastin DA, Reeves PR. 1995. Sequence and analysis of the O antigen gene (rfb) cluster of Escherichia coli O111. *Gene*, 164:17-23.

Beyer N, Alam J, Hallis TM, Guo Z, Liu HW. 2003. The biosynthesis of GDP-L-colitose: C-3 deoxygenation is catalyzed by a unique coenzyme B6-dependent enzyme. *Journal of the American Chemical Society*, 125:5584-5585.

Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A. 2006. Structures and mechanisms of glycosyltransferases. *Glycobiology*, 16:29R-37R.

Calin O, Eller S, Hahm HS, Seeberger PH. 2013. Total synthesis of the Escherichia coli O111 O-specific polysaccharide repeating unit. *Chemistry*, 19:3995-4002.

Calin O, Pragani R, Seeberger PH. 2012. De novo synthesis of L-colitose and L-rhodinose building blocks. *The Journal of organic chemistry*, 77:870-877.

Coyne MJ, Reinap B, Lee MM, Comstock LE. 2005. Human symbionts use a host-like pathway for surface fucosylation. *Science*, 307:1778-1781.

Cunneen MM, Pacinelli E, Song WC, Reeves PR. 2011. Genetic analysis of the O-antigen gene clusters of Yersinia pseudotuberculosis O:6 and O:7. *Glycobiology*, 21:1140-1146.

Engels L, Elling L. 2014. WbgL: a novel bacterial alpha1,2-fucosyltransferase for the synthesis of 2'-fucosyllactose. *Glycobiology*, 24:170-178.

Gorshkova NM, Gorshkova RP, Ivanova EP, Nazarenko EL, Zubkov VA. 2001. [Diversity in the monosaccharide composition of antigenic polysaccharides from proteobacteria Pseudoalteromonas and Marinomonas genera]. *Mikrobiologiia*, 70:651-655.

Heath EC, Elbein AD. 1962. The enzymatic synthesis of guanosine diphosphate colitose by a mutant strain of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 48:1209-1216.

Isshiki Y, Kondo S, Iguchi T, Sano Y, Shimada T, Hisatsune K. 1996. An immunochemical study of serological cross-reaction between lipopolysaccharides from Vibrio cholerae O22 and O139. *Microbiology*, 142 (Pt 6):1499-1504.

Kenyon JJ, De Castro C, Cunneen MM, Reeves PR, Molinaro A, Holst O, Skurnik M. 2011. The genetics and structure of the O-specific polysaccharide of Yersinia pseudotuberculosis serotype O:10 and its relationship with Escherichia coli O111 and Salmonella enterica O35. *Glycobiology*, 21:1131-1139.

Knirel YA, Senchenkova SN, Jansson PE, Weintraub A. 1998. More on the structure of Vibrio cholerae O22 lipopolysaccharide. *Carbohydrate research*, 310:117-119.

Knirel YA, Senchenkova SN, Jansson PE, Weintraub A, Ansaruzzaman M, Albert MJ. 1996. Structure of the O-specific polysaccharide of an Aeromonas trota strain cross-reactive with Vibrio cholerae O139 Bengal. *European journal of biochemistry / FEBS*, 238:160-165.

Li L, Liu Y, Li T, Wang W, Yu Z, Ma C, Qu J, Zhao W, Chen X, Wang PG. 2015. Efficient chemoenzymatic synthesis of novel galacto-N-biose derivatives and their sialylated forms. *Chemical communications*, 51:10310-10313.

Li L, Liu Y, Wan Y, Li Y, Chen X, Zhao W, Wang PG. 2013. Efficient enzymatic synthesis of guanosine 5'-diphosphate-sugars and derivatives. *Organic letters*, 15:5528-5530.

Li M, Liu XW, Shao J, Shen J, Jia Q, Yi W, Song JK, Woodward R, Chow CS, Wang PG. 2008. Characterization of a novel alpha1,2-fucosyltransferase of Escherichia coli O128:b12 and functional investigation of its common motif. *Biochemistry*, 47:378-387.

Li M, Shen J, Liu X, Shao J, Yi W, Chow CS, Wang PG. 2008. Identification of a new alpha1,2-fucosyltransferase involved in O-antigen biosynthesis of Escherichia coli O86:B7 and formation of H-type 3 blood group antigen. *Biochemistry*, 47:11590-11597.

Lindberg B, Lindh F, Lonngren J. 1981. Structural studies of the O-specific side-chain of the lipopolysaccharide from Escherichia coli O 55. *Carbohydrate research*, 97:105-112.

Liu XW, Xia C, Li L, Guan WY, Pettit N, Zhang HC, Chen M, Wang PG. 2009. Characterization and synthetic application of a novel beta1,3-galactosyltransferase from Escherichia coli O55:H7. *Bioorganic & medicinal chemistry*, 17:4910-4915.

Luderitz O, Staub AM, Stirm S, Westphal O. 1958. [Colitose, 3-desoxy-L-fucose, a new sugar building block with immunospecific functions in the endotoxins, lipopolysaccharides, of some gram-negative bacteria]. *Biochemische Zeitschrift*, 330:193-197.

Luderitz O, Westphal O, Staub AM, Le Minor L. 1960. Preparation and immunological properties of an artificial antigen with colitose (3-deoxy-1-fucose) as the determinant group. *Nature*, 188:556-558.

Medearis DN, Jr., Camitta BM, Heath EC. 1968. Cell wall composition and virulence in Escherichia coli. *The Journal of experimental medicine*, 128:399-414.

Muldoon J, Perepelov AV, Shashkov AS, Gorshkova RP, Nazarenko EL, Zubkov VA, Ivanova EP, Knirel YA, Savage AV. 2001. Structure of a colitose-containing O-specific polysaccharide of the marine bacterium Pseudoalteromonas tetraodonis IAM 14160(T). *Carbohydrate research*, 333:41-46.

Ochoa TJ, Contreras CA. 2011. Enteropathogenic E. coli (EPEC) infection in children. *Current opinion in infectious diseases*, 24:478.

Ovchinnikova OG, Kocharova NA, Wykrota M, Shashkov AS, Knirel YA, Rozalski A. 2007. Structure of a colitose-containing O-polysaccharide from the lipopolysaccharide of Providencia alcalifaciens O6. *Carbohydrate research*, 342:2144-2148.

Pettit N, Styslinger T, Mei Z, Han W, Zhao G, Wang PG. 2010. Characterization of WbiQ: An alpha1,2-fucosyltransferase from Escherichia coli O127:K63(B8), and synthesis of H-type 3 blood group antigen. *Biochemical and biophysical research communications*, 402:190-195.

Raetz CR, Whitfield C. 2002. Lipopolysaccharide endotoxins. *Annual review of biochemistry*, 71:635-700.

Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 106:1760-1765.

Rillahan CD, Antonopoulos A, Lefort CT, Sonon R, Azadi P, Ley K, Dell A, Haslam SM, Paulson JC. 2012. Global metabolic inhibitors of sialyl- and fucosyltransferases remodel the glycome. *Nature chemical biology*, 8:661-668.

Ruttens B, Kovac P. 2004. A facile synthesis of armed and disarmed colitose thioglycosides. *Synthesis*:2505-2508.

Saito T, Hatada M, Iwanaga S, Kawabata S. 1997. A newly identified horseshoe crab lectin with binding specificity to O-antigen of bacterial lipopolysaccharides. *The Journal of biological chemistry*, 272:30703-30708.

Senchenkova SN, Shashkov AS, Knirel YA, Schwarzmuller E, Mayer H. 1997. Structure of the O-specific polysaccharide of Salmonella enterica ssp. arizonae O50 (Arizona 9a,9b). *Carbohydrate research*, 301:61-67.

Silipo A, Molinaro A, Nazarenko EL, Gorshkova RP, Ivanova EP, Lanzetta R, Parrilli M. 2005. The O-chain structure from the LPS of marine halophilic bacterium Pseudoalteromonas carrageenovora-type strain IAM 12662T. *Carbohydrate research*, 340:2693-2697.

Sodha S, Heiman K, Gould L, Bishop R, Iwamoto M, Swerdlow D, Griffin P. 2015. National patterns of Escherichia coli O157 infections, USA, 1996–2011. *Epidemiology and infection*, 143:267-273.

Tomshich SV, Kokoulin MS, Kalinovsky AI, Komandrova NA, Nedashkovskaya OI. 2015. Structure of the O-specific polysaccharide from a marine bacterium Echinicola vietnamensis Kcapital EM, Cyrilliccapital EM, Cyrillic 6221capital TE, Cyrillic. *Carbohydrate research*, 402:1-5.

Turek D, Sundgren A, Lahmann M, Oscarson S. 2006. Synthesis of oligosaccharides corresponding to Vibrio cholerae O139 polysaccharide structures containing dideoxy sugars and a cyclic phosphate. *Organic & biomolecular chemistry*, 4:1236-1241.

Wang L, Huskic S, Cisterne A, Rothemund D, Reeves PR. 2002. The O-antigen gene cluster of Escherichia coli O55:H7 and identification of a new UDP-GlcNAc C4 epimerase gene. *Journal of bacteriology*, 184:2620-2625.

Wang L, Reeves PR. 2000. The Escherichia coli O111 and Salmonella enterica O35 gene clusters: gene clusters encoding the same colitose-containing O antigen are highly conserved. *Journal of bacteriology*, 182:5256-5261.

Wang W, Hu T, Frantom PA, Zheng T, Gerwe B, Del Amo DS, Garret S, Seidel RD, 3rd, Wu P. 2009. Chemoenzymatic synthesis of GDP-L-fucose and the Lewis X glycan derivatives. *Proceedings of the National Academy of Sciences of the United States of America*, 106:16096-16101.

Westphal O, Luderitz O, Staub AM, Tinelli R. 1959. [Demonstration & immunological differentiation of 2 optical antipodes-abequose & colitose (3-desoxy-D- & -L-fucose) as determinant endgroups in Salmonella O antigens]. *Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene. 1. Abt. Medizinisch-hygienische Bakteriologie, Virusforschung und Parasitologie. Originale*, 174:307-314.

Weymouth-Wilson AC. 1997. The role of carbohydrates in biologically active natural products. *Natural product reports*, 14:99-110.

Yamasaki S, Shimizu T, Hoshino K, Ho ST, Shimada T, Nair GB, Takeda Y. 1999. The genes responsible for O-antigen synthesis of vibrio cholerae O139 are closely related to those of vibrio cholerae O22. *Gene*, 237:321-332.

Yi W, Liu X, Li Y, Li J, Xia C, Zhou G, Zhang W, Zhao W, Chen X, Wang PG. 2009. Remodeling bacterial polysaccharides by metabolic pathway engineering. *Proceedings of the National Academy of Sciences of the United States of America*, 106:4207-4212.

Zhao G, Guan W, Cai L, Wang PG. 2010. Enzymatic route to preparative-scale synthesis of UDP-GlcNAc/GalNAc, their analogues and GDP-fucose. *Nature protocols*, 5:636-646.

Legends to Figures

Figure 1. *De novo* biosynthetic pathways of GDP-L-galactose, GDP-L-Fuc and GDP-colitose. GMP, GDP-mannose pyrophosphorylase; GME, GDP-mannose 3,5-epimerase; GMD, GDP-mannose 4,6-dehydratase; GFS, GDP-L-Fuc synthase; ColD, GDP-4-keto-6-deoxy-mannose-3-dehydratase; ColC, GDP-colitose synthase; GKDM, GDP-4-keto-6-deoxy-mannose. GKDDM, GDP-4-keto-3,6-dideoxy-mannose; PLP, pyridoxal-5'-phosphate.

Figure 2. Structure of *E. coli* O55:H7 O-antigen repeating unit. WbgN, putative α1,2-colitosyltransferases. The repeating units are polymerized between the GalNAc and GlcNAc residue to form O-antigen.

Figure 3. The salavage biosynthetic pathway of GDP-L-Fuc and chemoenzymatic synthesis of GDP-colitose. FKP, L-fucokinase/GDP-L-Fuc pyrophosphorylase from *Bacteroides fragilis* 9343; EcPPA, inorganic pyrophosphatase from *E. coli* O157.

Figure 4. TLC analysis of FKP catalyzed reactions. 1, ATP; 2, GTP; 3, ADP; 4, colitose; 5, FKP reaction containing colitose, ATP and $MgCl_2$; 6, 5 + GTP and EcPpA; 7, L-Fuc; 8, FKP reaction containing L-Fuc, ATP and $MgCl_2$; 9, 8 + GTP and EcPpA; 10, GDP-L-Fuc.

Figure 5. SDS-PAGE analysis of gene expression and purification (lanes 1-7) and western blot analysis of purified WbgN (lane 8). 1, before IPTG induction; 2, after IPTG induction; 3, supernatant; 4, precipitate; 5, wash; 6, flow through; 7, eluted enzyme; 8, western blot of purified WbgN using anti-His tag antibody as first antibody.

Figure 6. Biophysical studies of recombinant WbgN (buffers were used in a concentration of 50 mM). (A) pH dependent study, and (B) metal ion dependent study (metal ion were used in a concentration of 10 mM).

Tables

Table 1. Kinetic parameters of recombinant WbgN. LNB was used as an acceptor to study kinetics towards a sugar donor, and GDP-colitose was used as a sugar donor to study kinetics towards the acceptor LNB.

Substrate	K_m (mM)	V_{max} ($mM\,mg^{-1}$)	k_{cat} (min^{-1})	k_{cat}/K_m ($\min^{-1}mM^{-1}$)
GDP-colitose	0.76 ± 0.18	0.046 ± 0.004	8.8 ± 0.8	12[a]
GDP-L-Fuc	1.3 ± 0.42	0.062 ± 0.008	12 ± 1.5	9.2[a]
Galβ1,3-GlcNAc	1.4 ± 0.34	0.027 ± 0.003	5.2 ± 0.6	3.7[a]

[a] k_{cat}/K_m values were calculated using the average values of K_m and k_{cat}.

Table 2. Relative acceptor specificity of WbgN. Activity of WbgN towards LNB in presence of GDP-colitose was set as 100%.

Acceptor	**Sugar Nucleotide Donor**	
	GDP-colitose	GDP-L-Fuc
Gal	<1%	<1%
GlcNAc	<1%	<1%
Galβ1,3-GlcNAc (LNB)	100%	98%
Galβ1,3-GalNAc (GNB)	~1%	~1%
Galβ1,4-GlcNAc (LacNAc)	<1%	<1%
Galβ1,4-Gal (Lac)	<1%	<1%