Efficient Synthesis of the Pentasaccharide Repeating Unit of the O-Antigenic Polysaccharide of *Escherichia coli* O166 Strain

Anup Kumar Misra*

Bose Institute, Division of Molecular Medicine, P-1/12, C.I.T.
Scheme VII-M, Kolkata-700054, India
akmisra69@gmail.com

Received: 20.06.2014
Accepted after revision: 06.10.2014
Published online: 10.11.2014

**Abstract** An efficient strategy has been developed for the synthesis of the pentasaccharide repeating unit of the cell-wall polysaccharide of *Escherichia coli* O166 strain through sequential stereoselective glycosylations of monosaccharide intermediates. All the glycosylation steps were high-yielding with high stereoselectivities.

**Key words** carbohydrates, glycosylations, glycosides, stereoselectivity, oligosaccharides

Diarrheal outbreaks and gastrointestinal complications are important health problems in developing countries. In general, intake of contaminated food and water and a lack of adequate sanitation are the leading causes of enteric disorders. Recently, gastrointestinal infections have also become significant health hazards in developed countries. Among the several enteropathogenic microbes that are responsible for the diarrheal infections, pathogenic strains of *Escherichia coli* merit particular attention. These are associated with several gastrointestinal infections, particularly ‘travelers’ diarrhea’ and they can be classified into several pathotypes, such as enteropathogenic, enterohemorrhagic, enterotoxigenic, enteroinvasive, enteroaggregative, or diffusely adherent. The O166 strain of *E. coli* is generally classed as belonging to the enteraggregative pathotype and causes diarrhea in human by producing a heat-stable enterotoxin. *E. coli* O166 has also been isolated from the environment and from cattle, and has also been classified as an enterohemorrhagic strain. In 1996, *E. coli* O166 was identified as the cause of an outbreak of diarrhea in Japan.

Cell-wall polysaccharides of virulent strains of bacteria play crucial roles in the initial stages of bacterial infections in hosts. As result, researchers have focused attention on the characterization of cell-wall O-antigenic polysaccharides from several bacterial strains. Recently, the structure of the pentasaccharide repeating unit of the O-antigenic polysaccharide of *E. coli* O166 was reported by Ali et al. This pentasaccharide contains D-glucose, D-galactose, and N-acetyl-D-galactosamine moieties. As the result of the acceleration in the failure of antibiotics to act on multidrug-resistant strains of bacteria, the development of alternative approaches to the control of bacterial infections is currently a major area in drug-discovery research. It is therefore relevant to develop therapeutics based on glycoconjugate derivatives related to cell-wall polysaccharide O-antigens, because of their involvement in the process of bacterial infection. Sufficient quantities of oligosaccharides free of biological impurities are required for biological studies, and these cannot readily be isolated from natural sources. Therefore, the development of efficient strategies for chemical synthesis of these oligosaccharides would be extremely useful in providing access to significant quantities of pure oligosaccharides with appropriate structures. In this context, an efficient synthesis of the pentasaccharide repeating unit of the O-antigenic polysaccharide of *E. coli* O166 has been developed. The target pentasaccharide was synthesized as its 4-methoxypyphenyl glycoside by a series of stereoselective sequential glycosylation reactions of suitably functionalized monosaccharide intermediates. For this purpose, the monosaccharide intermediates 2, 3, 4, 5, and 6 were prepared by following the methods previously reported (Figure 1).
Treatment of the known ethyl 4,6-O-benzylidene-2-deoxy-2-(N-phthalimido)-1-thio-β-D-galactopyranoside (7)\textsuperscript{14} with triethylsilane in the presence of molecular iodine,\textsuperscript{15} followed by acetylation with acetic anhydride and pyridine\textsuperscript{16} gave ethyl 3,4-di-O-acetyl-6-O-benzyl-2-deoxy-2-(N-phthalimido)-1-thio-β-D-galactopyranoside (2) in 75% overall yield (Scheme 1).

Stereoselective 1,2-cis glycosylation of the D-galactosyl donor 3\textsuperscript{10} with the D-galactosyl acceptor 4\textsuperscript{11} in the presence of N-iodosuccinimide and trimethylsilyl trifluoromethanesulfonate\textsuperscript{17,18} in dichloromethane–diethyl ether gave the disaccharide derivative 8 in 72% yield, together with a minor quantity (~8%) of another isomer. Disaccharide 8 was purified by column chromatography and its stereoselective formation was confirmed by spectroscopic analysis. Removal of the allyl ether group from disaccharide 8 by using palladium(II) chloride\textsuperscript{19} gave the partially deprotected disaccharide 9 in 75% yield. Stereoselective 1,2-trans glycosylation of disaccharide 9 with the D-galactosamine donor 2 in the presence of N-iodosuccinimide and trimethylsilyl trifluoromethanesulfonate\textsuperscript{17,18} gave the trisaccharide derivative 10 in 74% yield. The exclusive formation of compound 10 was confirmed by NMR spectroscopy. De-O-acetylation of compound 10 with sodium methoxide\textsuperscript{20} gave the trisaccharide diol derivative 11, which was selectively 4-O-acetylated through the formation of an ortho ester\textsuperscript{21} and subsequent acidic hydrolysis to give trisaccharides 12 in 86% overall yield. Stereoselective 1,2-cis glycosylation of trisaccharide 12 with the D-galactosamine derivative 5\textsuperscript{12} as a glycosyl donor in the presence of N-iodosuccinimide and trimethylsilyl trifluoromethanesulfonate\textsuperscript{17,18} in dichloromethane–diethyl ether gave the tetrasaccharide derivative 13 in 68% yield, together with a minor quantity of the trans-glycosylation product (~10%), which was separated by column chromatography. Stereoselective formation of the tetrasaccharide derivative 13 was confirmed by spectroscopic analysis. Selective removal\textsuperscript{22} of the 3-O-acetyl group from compound 13 by using sodium methoxide left the internally located 4-O-acetyl group unaffected and gave the tetrasaccharide acceptor 14 in 95% yield. N-Iodosuccinimide–trimethylsilyl trifluoromethanesulfonate-mediated 1,2-trans-glycosylation of tetrasaccharide 14 with the D-glucose thioglycoside derivative 6\textsuperscript{13} in dichloromethane gave the pentasaccharide derivative 15 in 72% yield. NMR spectroscopic analysis of compound 15 confirmed that it was formed exclusively. Finally, pentasaccharide 15 was subjected to a series of reactions to remove the protecting groups completely. These reactions included (a) removal of the N-phthaloyl group by treatment with hydrazine monohydrate,\textsuperscript{23} followed by acetylation of the resulting amine using acetic anhydride and pyridine; (b) removal of the benzyl ethers and benzylidene acetics and reduction of the azido group by hydrogenolysis over palladium(II) hydroxide/carbon,\textsuperscript{24} followed by acetylation of the resulting amine using acetic anhydride and methanol; and (c) removal of the acetyl and benzoyl groups by treatment with sodium methoxide to give the target penta-

---

**Figure 1** Structures of the synthesized pentasaccharide repeating unit of the O-antigenic polysaccharide of *E. coli* O166 (1) and various monosaccharide intermediates (2–6).
Scheme 1  Reagents and conditions: (a) Et3SiH, MeCN, 0–5 °C; (b) Ac2O, py, r.t., 2 h, 75% (two steps); (c) NIS, TMSOTf, CH2Cl2–Et2O (1:3), 0 °C, MS-4Å, 45 min, 72% for compound 8, 68% for compound 13; (d) PdCl2, MeOH, r.t., 3 h, 75%; (e) NIS, TMSOTf, CH2Cl2, MS-4Å, –20 °C, 30 min, 74% for compound 10, 72% for compound 15; (f) 0.1 M NaOMe, MeOH, r.t., 1.5 h, quant for compound 11, 95% for compound 14; (g) (EtO)3CMe, TsOH, DMF, r.t., 2 h then H2O, r.t., 30 min, 86%; (h) NH2NH2·H2O, EtOH, 80 °C, 10 h; (i) Ac2O, py, r.t., 1 h; (j) H2, 20% Pd(OH)2/C, MeOH, r.t., 15 h; (k) Ac2O, MeOH, r.t., 1 h; (l) 0.1 M NaOMe, MeOH, r.t., 1.5 h, 54% (overall).

In summary, a straightforward strategy has been developed for the synthesis of the pentasaccharide repeating unit of the O-antigen of *Escherichia coli* O166 by a series of sequential stereoselective glycosylations of monosaccharide intermediates. The glycosylation steps were high-yielding and gave an excellent stereochemical outcome. Similar reaction conditions were used in each of the glycosylation reactions.

All reactions were monitored by TLC on silica gel coated plates. TLC spots were visualized by spraying the plates with ceric sulfate [2% Ce(SO4)2 in 2 N H2SO4] and warming the sprayed TLC plates on a hot-plate. Silica gel (230–400 mesh) was used for column chromatography. NMR spectra were recorded on Bruker Avance 500 MHz spectrometer with CDCl3 as solvent and TMS as the internal reference, unless stated otherwise. Chemical shifts (δ) are expressed in ppm. Complete assignment of the 1H and 13C NMR spectra was carried out by means of a standard set of NMR experiments, e.g. 1H NMR, 13C NMR, 1H–1H COSY, 2D HSQC. MALDI mass spectra were recorded on a Bruker mass spectrometer. Optical rotations were recorded in a JASCO P-2000 polarimeter. Commercially available grades of organic solvents of adequate purity were used in all reactions.

Ethyl 3,4-di-O-acetyl-6-O-benzyl-2-deoxy-2-(N-phthalimido)-1-thio-β-D-galactopyranoside (2)

Et3SiH (1.8 mL, 11.27 mmol) and I2 (250 mg, 0.98 mmol) were added sequentially to a solution of monosaccharide derivative 7 (2 g, 4.53 mmol) in MeCN (10 mL) at 0–5 °C, and the mixture was stirred at 0–5 °C for 40 min. The mixture was then diluted with CH2Cl2 (100 mL) and washed successively with 5% aq Na2S2O3 (50 mL) and H2O (100 mL), then dried (Na2SO4) and concentrated. A solution of the crude product in Ac2O (5 mL) and pyridine (5 mL) was kept at r.t. for 2 h. The reagents were removed under reduced pressure and the crude product was purified by chromatography [silica gel, hexane–EtOAc (3:1)] to give a yellow oil; yield: 1.8 g (75%); [α]21° +43 (c 1.0, CHCl3).

IR (neat): 3024, 1736, 1515, 1372, 1216, 1096, 766 cm–1.

1H NMR (500 MHz, CDCl3): δ = 7.83–7.21 (m, 9 H, Ar-H), 5.79 (dd, J = 11.0, 3.0 Hz, 1 H, H-3), 5.55 (d, J = 2.0 Hz, 1 H, H-4), 5.44 (d, J = 10.5 Hz, 1 H, H-1), 4.57 (d, J = 12.0 Hz, 1 H, PhCH2), 4.53 (t, J = 10.5 Hz, 1 H, H-2), 4.43 (d, J = 12.0 Hz, 1 H, PhCH2), 4.07–4.04 (m, 1 H, H-5), 3.59–3.56 (m, 1 H, H-6a), 3.50–3.46 (m, 1 H, H-6b), 2.71–2.63 (m, 2 H, SCH2CH2), 2.08–1.82 (2 s, 6 H, 2 COCH3), 1.26 (t, J = 7.4 Hz, 3 H, SCH2CH2).

13C NMR (125 MHz, CDCl3): δ = 170.0, 169.5 (2 COCH3), 167.8, 167.3 (PhthCO), 137.6–123.6 (m, Ar-C), 81.5 (C-1), 75.9 (C-3), 73.5 (PhCH2), 69.0 (C-4), 67.6 (C-6), 67.4 (C-5), 50.3 (C-2), 24.4 (SCH2CH2), 20.7, 20.5 (2 COCH3), 14.9 (SCH2CH2).


4-Methoxyphenyl [3,4-Di-O-acetyl-6-0-benzyl-2-deoxy-2-(N-<br>phthalimido)-β-D-galactopyranosyl]-(1→3)-(2,4,6-tri-O-<br>benzyl-α-D-galactopyranosyl)-(1→6)-2,3,4-tri-O-benzyl-α-D-galactopyranoside (10)

MS-4A (1 g) was added to a solution of compound 9 (1.2 g, 1.21 mmol) and compound 2 (960 mg, 1.82 mmol) in anhyd CHCl₃ (10 mL) and the mixture was cooled to −20 °C under argon. The cooled mixture was stirred at −20 °C for 30 min. The mixture was then diluted with CH₂Cl₂ (100 mL) and washed successively with 5% aq Na₂SO₄ (50 mL), sat. aq NaHCO₃ (100 mL), and H₂O (100 mL). The organic phase was dried (Na₂SO₄) and concentrated under reduced pressure to give a crude product that was purified by chromatography [silica gel, hexane-EtOAc (4:1)] to give a colorless oil; yield: 1.3 g (74%); [α]D₂+33 +35 (c 1.0, CHCl₃).

IR (KBr): 3020, 2362, 1719, 1610, 1216, 1098, 761 cm⁻¹.

1H NMR (500 MHz, CDCl₃): δ = 7.46–6.75 (m, 43 H, Ar-H), 5.84 (d, J = 8.5 Hz, 1 H, H-1c), 5.44 (d, J = 3.0 Hz, 1 H, H-4c), 5.27 (d, J = 2.0 Hz, 1 H, H-1c), 4.93–4.49 (10 d, J = 11.5 Hz each, 10 H, PhCH₃), 4.41 (d, J = 2.5 Hz, 1 H, H-1c), 4.36–4.32 (m, 4 H, 2 H-2c, PhCH₂), 4.22 (d, J = 11.5 Hz, 1 H, PhCH₂), 4.09–4.01 (m, 4 H, 2 H-2c, H-3c, H-5c), 3.98 (d, J = 10.5, 3.0 Hz, 1 H, H-3c), 3.96–3.87 (m, 3 H, 3 H, H-3c, H-6c, H-4c), 3.82–
3.80 (m, 1 H, H-5c), 3.72 (s, 3 H, OCH3), 3.67 (dd, J = 10.5, 3.0 Hz, 1 H, H-2a), 3.58–3.51 (m, 3 H, H-6a-D, H-6a-D), 3.42–3.35 (m, 1 H, H-6a-D), 3.34–3.26 (m, 2 H, H-6a-D, H-6a-D), 2.05 (s, 3 H, COCH3).

\[ {^{13}C} \text{ NMR (125 MHz, CDCl}_3): \delta = 171.5 (\text{COCH}_3), 167.7, 167.4 (\text{PhCH}O), 155.0–114.5 (m, Ar-C), 100.6 (\text{PhCH}), 99.8 (C-1C), 97.8 (2 C, C-1C, 1C), 97.0 (C-1C), 78.8 (2 C, C-5C, C-6C), 76.4 (C-3D), 76.3 (C-2D), 75.5 (C-2C), 75.2 (C-3C, 74.5 (2 C, PhCH2), 73.6 (PhCH2), 73.5 (PhCH3), 73.3 (C-1C), 73.0 (C-1C), 72.9 (PhCH2), 72.6 (C-4C), 69.8 (C-3C), 69.7 (C-3C), 69.2 (C-4C), 69.1 (C-6C), 68.3 (2 C, C-6C, C-6C), 67.0 (C-6C), 66.4 (C-6C), 62.9 (C-5C), 57.2 (C-2D), 55.5 (OCH3), 53.4 (C-2C), 20.9, 20.7 (2 COCH3).

MALDI-MS: 1751.6 [M + Na]+.

Anal. Calcd for C66H64NO34 (1728.67): C, 68.74; H, 5.83. Found: C, 68.60; H, 6.00.

4-Methoxypyrenyl (2-Azido-4,6-benzyldiene-2-deoxy-\(\beta\)-galactopropyl)osyl(1–3)-[4-O-acetyl-6-benzyl-2-deoxy-\(\beta\)-D-galactopyranosyl]-[1–3]-[2,4,6-tri-o-benzyl-\(\alpha\)-galactopyranosyl](1–6)-2,3,4-tri-o-benzyl-\(\alpha\)-galactopropyranoside (14)

A solution of compound 10 (700 mg, 0.40 mmol) in 0.1 M methanolic NaOMe (20 mL) was stirred at rt for 1.5 h. The mixture was then neutralized with Dowex 50W X8 (H+), filtered, and concentrated under reduced pressure. The crude product was passed through a short pad of silica gel with elution by hexane–EtOAc (2:1) to give a white solid; yield: 650 mg (95%); mp 95–96 °C (EtOH); [\(\alpha\)]\text{D}21 +61 (c 1.0, CHCl3).

IR (KBr): 2927, 1746, 1508, 1456, 1369, 1233, 1109, 1055, 987, 826, 739 cm\(^{-1}\).

1H NMR (500 MHz, CDCl3): \(\delta = 7.45–6.76\) (m, 48 H, Ar-H), 5.66 (d, J = 3.0 Hz, 1 H, H-4C), 5.47 (d, J = 8.5 Hz, 1 H, H-1C), 5.27 (s, 1 H, PhCH), 5.26 (d, J = 2.5 Hz, 1 H, H-1C), 5.09 (d, J = 3.0 Hz, 1 H, H-1C), 4.95–4.69 (6 d, J = 11.5 Hz each, 6 H, PhCH6), 4.65–4.58 (m, 4 H, H-2C, PhCH2), 4.45 (d, J = 11.5 Hz, 1 H, PhCH), 4.40 (d, J = 11.5 Hz, 1 H, PhCH), 4.38 (d, J = 3.0 Hz, 1 H, H-1C), 4.35–4.22 (3 d, J = 11.5 Hz each, 3 H, PhCH3), 4.07–4.00 (m, 4 H, H-2C, H-2C, H-3C, H-4C), 3.92–3.84 (m, 2 H, H-3C, H-3C), 3.82–3.75 (m, 3 H, H-4C, H-4C, H-5C), 3.73 (s, 3 H, OCH3), 3.67 (d, J = 10.0, 3.0 Hz, 1 H, 2C), 3.62–3.60 (m, 1 H, H-6C), 3.58–3.48 (m, 4 H, H-2C, H-2C, H-6C, H-6C), 3.43–3.35 (m, 2 H, H-6C, H-6C), 3.32–3.25 (m, 2 H, H-6C, H-6C), 3.12–3.11 (m, 1 H, H-5C), 2.10 (s, 3 H, COCH3).

13C NMR (125 MHz, CDCl3): \(\delta = 171.2\) (COCH3), 167.7, 167.4 (PhCHO), 155.2–114.5 (m, Ar-C), 101.1 (PhCH), 99.8 (C-1C), 97.7 (2 C, C-1C, C-1C), 96.6 (C-1C), 78.8 (2 C, C-5C, C-6C), 76.7 (C-3C), 76.3 (C-3C), 75.5 (C-2C), 75.1 (C-2C), 74.9 (C-3C), 74.5 (2 C, PhCH2), 73.6 (PhCH2), 73.3 (PhCH3), 73.1 (2 C, PhCH2), 72.5 (PhCH2), 72.4 (C-4C), 72.3 (C-4C), 69.8 (C-4C), 69.1 (C-3C), 68.0 (C-6C), 68.5 (C-6C), 68.3 (C-6C), 67.6 (C-3C), 67.0 (C-6C), 66.0 (C-4C), 63.1 (C-5C), 60.7 (C-2C), 55.9 (OCH3), 53.3 (C-2C), 20.7 (COCH3).

MALDI-MS: 1709.6 [M + Na]+.

Anal. Calcd for C$_{796}$H$_{863}$N$_{4}$O$_{23}$ (1866.68): C, 69.03; H, 5.85. Found: C, 68.84; H, 6.00.

4-Methoxypyrenyl (2,3,4-Benzylidene-4,6-benzylidene-\(\beta\)-D-glucopyranosyl)-(1–3)-(2,4-benzylidene-2-deoxy-\(\beta\)-D-galactopyranosyl)-(1–3)-(2,4,6-tri-o-benzyl-\(\alpha\)-galactopyranosyl)-(1–6)-2,3,4-tri-o-benzyl-\(\alpha\)-galactopropyranoside (15)

MS-4A (500 mg) were added to a solution of compound 14 (600 mg, 0.36 mmol) and compound 6 (225 mg, 0.43 mmol) in anhyd CHCl$_3$ (5 mL), and the mixture was cooled to –20 °C under argon. NIS (100 mg, 0.44 mmol) and TMSOTf (3 μL) were added and the mixture was stirred at –20 °C for 30 min. The mixture was then diluted with CHCl$_3$ (50 mL) and washed successively with 5% aq Na$_2$SO$_4$ (25 mL), sat. aq NaHCO$_3$ (50 mL), and H$_2$O (50 mL). The organic phase was dried (Na$_2$SO$_4$) and concentrated under reduced pressure to give a crude product that was purified by chromatography (silica gel, hexane–EtOAc (4:1)) to give a white solid; yield: 550 mg (72%); mp 162–163 °C (EtOH); [\(\alpha\)]\text{D}23 +72 (c 1.0, CHCl3).

IR (KBr): 2932, 1776, 1748, 1720, 1509, 1388, 1229, 1107, 1081, 1052, 1029, 998, 827, 738, 722 cm\(^{-1}\).

1H NMR (500 MHz, CDCl3): \(\delta = 7.89–6.70\) (m, 63 H, Ar-H), 5.64 (t, J = 8.5 Hz, 1 H, H-1C), 5.56 (d, J = 3.0 Hz, 1 H, H-4C), 5.44 (d, J = 8.5 Hz, 1 H, H-1C), 5.30 (t, J = 9.0 Hz, 1 H, H-1C), 5.37 (s, 1 H, PhCH), 5.21 (d, J = 2.5 Hz, 1 H, H-1C), 5.19 (s, 1 H, PhCH), 4.97 (d, J = 3.0 Hz, 1 H, H-1C), 4.95
(\(J = 8.5 \text{ Hz}, 1 \text{ H}, \text{H-1}a\)), 4.91–4.52 (8 d, \(J = 11.5 \text{ Hz each}, 8 \text{ H}, \text{PhCH}_2\)), 4.51 (t, \(J = 8.5 \text{ Hz}, 1 \text{ H}, \text{H-2}a\)), 4.41 (d, \(J = 11.5 \text{ Hz}, 1 \text{ H}, \text{PhCH}_2\)), 4.34 (d, \(J = 3.0 \text{ Hz}, 1 \text{ H}, \text{H-1}b\)), 4.32–4.27 (m, 4 \text{ H}, \text{PhCH}_2\)), 4.26–4.21 (m, 1 \text{ H}, \text{H-6}a\)), 4.17 (t, \(J = 11.5 \text{ Hz}, 1 \text{ H}, \text{PhCH}_2\)), 4.02–3.96 (36 \text{ H}, 5 \text{ H}, \text{H-2}b, 3 \text{ H}, \text{H-3}a, 5 \text{ H}, \text{H-5}a, 3.92–3.89 (m, 3 \text{ H}, \text{H-3}b, 3 \text{ H}, \text{H-4}c)), 3.85–3.80 (m, 2 \text{ H}, \text{H-3}b)), 3.78–3.62 (m, 6 \text{ H}, 6 \text{ H}, \text{H-2}b, 6 \text{ H}, \text{H-4}b, 6 \text{ H}, \text{H-6}a)), 3.72 (s, 3 \text{ H}, \text{OCH}_3)), 3.55–3.46 (m, 4 \text{ H}, \text{H-6}a\), \text{H-6}b\), \text{H-6}c\), 3.43–3.40 (m, 1 \text{ H}, \text{H-6}a\), \text{H-6}b\), \text{H-6}c\), 3.38–3.32 (m, 2 \text{ H}, \text{H-6}a\), \text{H-6}b\), \text{H-6}c\), 3.06–3.05 (m, 1 \text{ H}, \text{H-5}b\), 1.95 (s, 3 \text{ H}, \text{H-5}c).

**Acknowledgments**

A.S. thanks CSIR, India, for providing a Junior Research Fellowship. This work was supported by CSIR, New Delhi, [Project no. 02(0038)/11/EMR-II] and the Bose Institute, Kolkata.

**Supporting Information**

Supporting information for this article is available online at http://dx.doi.org/10.1055/s-0034-1378911.

**References**