PNAS
www.pnas.org

Supplementary Information for

Chlorovirus PBCV-1 Protein A064R has three of the transferase activities necessary to synthesize its capsid protein N-linked glycans

Immacolata Speciale[a,b,1], Maria Elena Laugieri[c,1], Eric Noel[d,e,1], Sicheng Lin[f], Todd L. Lowary[f,g], Antonio Molinaro[a], Garry A. Duncan[d], Irina V. Agarkova[d,h], Domenico Garozzo[i], Michela G. Tonetti[c,2], James L. Van Etten[d,h,2], Cristina De Castro[b,2]

[a] Department of Chemical Sciences, University of Napoli Federico II, Via Cintia 4, 80126 Napoli, Italy

[b] Department of Agricultural Sciences, University of Napoli Federico II, Via Università 100, 80055 Portici NA, Italy

[c] Department of Experimental Medicine and Center of Excellence for Biomedical Research, University of Genova, Viale Benedetto XV/1, 16132, Genova, Italy.

[d] Nebraska Center for Virology, University of Nebraska, Lincoln, NE 68583-0900

[e] School of Biological Sciences, University of Nebraska, Lincoln, NE 68588-0118, USA

[f] Department of Chemistry, University of Alberta, Gunning-Lemieux Chemistry Centre, Edmonton, Alberta T6G 2G2, Canada

[g] Institute of Biological Chemistry, Academia Sinica,128, Section 2, Academia Road, Nangang, Taipei, 11529, Taiwan

[h] Department of Plant Pathology, University of Nebraska, Lincoln, NE 68583-0722 USA.

[i] CNR, Institute for Polymers, Composites and Biomaterials, Via P. Gaifami 18, 95126 Catania, Italy

[1] contributed equally

[2] Corresponding author: C. De Castro, James L. Van Etten, Michela Tonetti

E-mail: decastro@unina.it; jvanetten1@unl.edu; tonetti@unige.it.

www.pnas.org/cgi/doi/10.1073/pnas.2016626117

This PDF includes **Pag.**

NMR characterization.
Tetrasaccharide **1** S-3
Pentasaccharide **2** S-4
Hexasaccharide **3** S-4
Rha-Rha disaccharide S-4
Trisaccharide **7** S-5
Hexasaccharide **4** S-5
Hexasaccharide **5** S-6

Material and methods
Cloning and expression of A061L S-7
Purification of A061L enzyme S-7
Synthesis of tetrasaccharide 1 S-7
Synthesis of Xyloside **6** S-17

Figures
Figure S1 S-19
Figure S2 S-20
Figure S3 S-21
Figure S4 S-22
Figure S5 S-23
Figure S6 S-24
Figure S7 S-25
Figure S8 S-26
Figure S9 S-27
Figure S10 S-28
Figure S11 S-29
Figure S12 S-30
Figure S13 S-31
Figure S14 S-32
Figure S15 S-33
Figure S16 S-34
Figure S17 S-35
Figure S18 S36
Figure S19 S37
Figure S20 S38
Figure S21 S39

Tables
Table S1 S-40
Table S2 S-41
Table S3 S-42
Table S4 S-43
Table S5 S-44
Table S6 S-45
Table S7 S-46
Table S8 S-47

SI References S-48

NMR characterization

Tetrasaccharide 1. [Table S1; Fig. S2 and S3] Structure of the synthetic product, here defined as **1** (Fig. 2), was determined by analyzing the complete set of ^{1}H-^{1}H homonuclear (COSY, TOCSY, T-ROESY) and ^{1}H-^{13}C heteronuclear (HSQC, HMBC, HSQC-TOCSY) 2D NMR spectra recorded at 310 K to reduce the overlap of the anomeric signals (attributions in Table S1).

Inspection of the low field region (5.4 – 4.5 ppm) of the HSQC spectrum (Fig. S2) displayed four anomeric carbon atoms labelled with capital letters from $\mathbf{A_1}$ to $\mathbf{D_1}$ (Table S1), as expected, and NMR analysis started from the anomeric proton of **C** residue at 5.16 ppm, identified as α-L-Rha. The analysis of the total correlation spectroscopy (TOCSY) spectrum established the *manno* stereochemistry of this residue. In fact, proton H-1 showed only one cross peak (Fig. S3A) relative to H-1/H-2 correlation, whereas the connections up to the methyl group H-6, which resonated at 1.27 ppm, were visible from H-2, due to favourable proton-proton coupling constant. This information, combined with those of the HSQC spectrum (Fig. S2), identified this residue as a terminal L-Rha unit, α-configured at the anomeric center based on the anomeric proton shape and its H-5/C-5 chemical shifts at 3.96/70.4 ppm, respectively (1).

Using a similar approach, **B** was identified as a Gal unit: starting from anomeric proton, at 5.10 ppm, its TOCSY pattern displayed three correlation with H-2, H-3 and H-4 protons (two of which are overlapped; Fig. S3A), beyond which both COSY and TOCSY experiments failed in the detection of the H-5, typical features of the *galacto* configured residue. In detail, C-2 and C-3 chemical shifts were identified analyzing the HSQC and the HSQC-TOCSY spectra; identification of the H-5 occurred through the analysis of the TOCSY correlation between H-4 and H-5. The latter was in turn correlated with a hydroxymethylene group (at 62.4 ppm), as indicated by the HMBC and TOCSY (Fig. S2 and S3A) spectra, recognizing this residue as a terminal Gal α-configured at the anomeric center due to the carbon chemical shift values (1).

Residue **A** showed the same TOCSY pattern, with the difference that its H-5 proton had a correlation with a methyl group at 1.31 ppm. Similar to the previous case, COSY and TOCSY were useful to establish the first four proton and carbon chemical shifts, instead the H-5 proton was achieved with the T-ROESY experiment. Hence, **A** was an α-Fuc residue ($^{3}J_{H1H2}$ 3.6 Hz), and the displacement at low field values of the carbons C-2, C-3 and C-4 (at 76.5, 74.7 and 81.6 ppm, respectively) with respect to the standard value, indicated that it was fully substituted. In agreement, the T-ROESY experiment (Fig. S3b) showed the key NOE contacts between proton H-1 of **C** with the H-3 of **A**, H-1 of **B** with H-2 of **A** and H-1 of **D'** with H-4 of **A**, data also supported by the HMBC spectrum (Fig. S2).

Regarding **D'** unit, its H-1 displayed intense correlations in the TOCSY spectrum (Fig. S3a) with five other protons, with two of them (4.01 and 3.2. ppm) located on the same carbon atom. Indeed this monosaccharide was identified as a Xyl in the pyranose form (C-5 at 66.2 ppm), β-configured as disclosed by the anomeric signal that appeared as a doublet with a high coupling constant value ($^{3}J_{H1H2}$ 7.7 Hz), and not further substituted as deduced by the chemical shift values of the other carbon atoms.

The remaining densities were ascribed to the octyl tail, and both T-ROESY (Fig. S3b) and HMBC (Fig. S2) spectra displayed that the proton of the residue **A** correlated with the C-1 of the octyl, as expected.

Pentasaccharide 2. [Table S2; Fig. S4 and S5] This product was obtained from the enzymatic reaction between the A064R-D1 enzyme and **1** in the presence of UDP-β-L-Rha and the manganese cation.
Comparing the proton spectrum of **2** with that of the starting substrate **1** (Fig. 2), the presence of a new anomeric signal, at 4.79 ppm confirmed the elongation of the acceptor with one sugar unit. The comparison of the HSQC spectrum (Fig. S2) of **1** with that of **2** (Fig. S4) identified a set of equivalent residues, namely **A**, **B** and **C**, whereas some variations were recognized for residue **D**. Indeed, C-4 of **D** was displaced at a lower field (79.4 ppm) because it was glycosylated (Table S2).
As for the new residue **E**, the TOCSY pattern of its anomeric proton at 4.79 ppm (Fig. S5A) was similar to that of **C**, and it showed only one correlation with H-2, from which all others were visible, suggesting that it was a Rha unit. However, in this case the H-5 proton resonated at 3.39 ppm, a value typical of a terminal β-configured Rha; also confirmed by the T-ROESY spectrum (Fig. S5B), in which the anomeric proton H-1 correlated with both H-3 and H-5, because of the cis diaxal orientation of these three protons. Moreover, the HMBC spectrum (Fig. S4) showed a long-range correlation between H-1 of **E** and C-4 of **D**, confirmed the analogue H-1 of **E**/H-4 of **D** correlation in the T-ROESY spectrum (Fig. S5b). Hence, **2** turned outs to be **1** elongate with a β-Rha unit linked at position 4 of the Xyl monosaccharide (Structure in Fig. 2).

Hexasaccharide 3. [Table S3; Fig. S8] It was obtained from the enzymatic reaction between the A064R-D2 and **2** in the presence of UDP-β-L-Rha. Its structure was achieved using the same strategy described above.

Most of the densities of its HSQC spectrum (Fig. S8A) overlapped with those of **2** (namely **A**, **B**, **C** and **D**) with some variations in residue **E**, due to its glycosylation. Consistently, the density relative to the carbon C-2 of the residue **E'** was shifted at low field ($^{1}H/^{13}C$ 4.12/78.1 ppm), with respect to the terminal β-Rha value ($^{1}H/^{13}C$ 4.06/71.8 ppm), causing a minor variation in the H-3 and C-3 values (see Table S3).

The additional residue, labelled with **F**, was ascribed to a terminal α-Rha as suggested by the TOCSY pattern, which showed only one correlation H-1/H-2 (Fig. S8B), while starting from H-2 all the other correlations were seen. The shape of the anomeric signal along with the C-5 chemical shift value at 69.7 ppm proved the α-configuration of this residue (24).

Rha-Rha disaccharide. [Table S4; Fig. S9] The A064R-D2 enzyme reacts with the Rha monosaccharide obtaining a Rha-Rha disaccharide as final product.

The HSQC spectrum (Fig. S9B) showed two densities into the anomeric region, at $^{1}H/^{13}C$ 5.22/93.7 ppm and 4.96/103.3 ppm, labelled with **R** and **T**, respectively. Both anomeric protons have only one correlation in the TOCSY spectrum (Fig. S9C) with proton H-2, from which all other correlations, until the methyl group H-6 which resonated at 1.28 ppm, can be found, suggesting that both residues are Rha units. Combining the data from the TOCSY and HSQC spectra, it was possible to assign all proton and carbon chemical shift values (Table S4), disclosing the presence of two α-Rha units. More in detail, the **T** residue is a terminal Rha unit, instead the **R** residue is a rhamnose unit with the free reducing end, and 2-linked based on the C-2 value that resonated at 80.3 ppm rather than at 71.1 ppm, as occurred for the terminal

Rha. A more detailed inspection of the proton NMR of the final product (Fig. S9) revealed traces of the α-Rha-(1→2)-β-Rha disaccharide; indeed, two anomeric signals appear at 5.03 and 4.89 ppm relative to the α- and β-rhamnose units of the terminal and reducing ends, respectively.

Trisaccharide 7. [Table S5; Fig. S11] It was obtained from the reaction between the A064R-D1D2 enzyme with **6** in the presence of the UDP-β-L-Rha and the manganese cation.

The low field region of the HSQC spectrum (Fig. S11A) displayed three anomeric carbon atoms at 4.36, 4.84 and 5.03 ppm, labelled with **D**, **E'** and **F**, respectively. Starting from anomeric proton H-1 of the **D** residue, its TOCSY pattern (Fig. S11B) showed correlations with the other five protons ring, identifying it as a Xyl, β-configurated at the anomeric centre due to its shape, that appeared as a doublet with a high coupling constant value, $^{3}J_{H1H2}$ 8.0 Hz. All proton and carbon chemical shift values were accounted for by combining information from the TOCSY and the HSQC spectra (Table S5) revealing that the carbon density C-4 of the **D** residue resonated at 79.4 ppm, because O-glycosylated.

Regarding the other two residues, they had the same TOCSY pattern (Fig. S11B), in which the anomeric proton showed only one main correlation with proton H-2, from which all other correlations are visible, meaning that they are both Rha units, that differ in some features. The HSQC experiment (Fig. S11A) revealed that the C-2 of the **E'** residue resonated at $^{1}H/^{13}C$ 4.12/78.1 ppm, instead the same density of the **F** residue resonated at $^{1}H/^{13}C$ 4.09/71.3 ppm, explaining that the **E'** residue was glycosylated at position 2, instead the **F** residue is a terminal Rha. In addition, the anomeric signal proton and carbon chemical shift values together with that of the C-5 (Table S5) allows us to identify the **E**' residue as β-Rha and the **F** residue as α-Rha.

Hexasaccharide 4. [Table S7; Fig. S14] It was obtained when the full length A064R enzyme reacted with **1** in presence of two equivalents of UDP-β-L-Rha and the methyltransferases precursor, SAM.

Analysis of the low field region of the HSQC spectrum (Fig. S14B) displayed six anomeric carbon atoms, similarly to **3** with some differences: densities of the residues from **A** to **D** were completely overlapped with those of the previous compound; **E''** differed only by a slight shift of the anomeric proton at 4.85 ppm (with respect 4.83 ppm in **3**); instead the **F** residue seems to be missing and a new anomeric signal (named **F'**) appeared at $^{1}H/^{13}C$ 5.17/99.2 ppm (Table S7). Starting from it, TOCSY correlation from H-1 stopped at H-2, but, from it, all other ring protons correlations up to H-6, which resonated at 1.28 ppm, were seen, identifying **F'** as another Rha unit α-configured at the anomeric center (C-5 at 69.5 ppm; 24). Moreover, its C-2 value (at 81.2 ppm) indicated that the hydroxyl function was capped with a methyl group ($^{1}H/^{13}C$ 3.45/59.6 ppm), also confirmed by a NOE contact in the T-ROESY spectrum (Fig. S14C), that was also used to established the linkage of the **F'** residue onto the position 2 of the residue **E''**.

A more detailed inspection revealed the presence of minor traces of **3** (in a ratio of 1:5 in **2**), proven by the anomeric signals at 5.04 and 4.83 ppm relative to the terminal α-Rha (**F**) and the 2-α-Rha (**E'**), respectively.

Hexasaccharide 5. [Table S8; Fig. S16] It was obtained when A061L enzyme reacted with **4** and SAM.

Comparison of the HSQC spectrum of **5** (Fig. S16A) with that of **4** identified a set of equivalent residues (**A**, **B**, **C**, **D** and **E''**) leaving only a few densities out, all ascribed to the residue **F''**. The anomeric signal at 5.22 ppm (Table S8) displayed the TOCSY pattern typical of a *manno* configured unit. In detail, the presence of a cross peak between the H-2 proton with a methyl group at 1.26 ppm, identified this unit as a Rha, α-configured at the anomeric center (C-5 at 69.5 ppm; Bock 1983). Furthermore, C-2 and C-3 carbon density values (at 76.9 and 80.1 ppm, respectively) suggested that both were substituted with two methyl groups at O-2 and O-3, in agreement with the presence of the NOE contacts in T-ROESY spectrum (Fig. S19B). Again, the T-ROESY spectrum showed a clear correlation between the anomeric proton of residue **F''** with the H-2 of the β-Rha (**E''**).

In conclusion, **5** was **1** implemented with two Rha units: a β-Rha linked at the fourth position of Xyl, in turn glycosylated at position 2 by a α-Rha with two methyl groups at positions 2 and 3, similarly to the N-glycan structure produced by PBCV-1 (2).

Material and methods

Cloning and expression of A061L. PBCV-1 *a061l* was cloned from PCR-amplified viral DNA using oligonucleotide primers with EcoI–NotI restriction site. PCR fragments of the expected size were digested and inserted into the restriction sites of the pGEX-5X-1 expression vector (GE Healthcare). This process produced a GST-tag at the N-terminus of the target protein. The resulting plasmid was transformed into *E. coli* strain One Shot TOP10 competent cells (Invitrogen) for maintenance. The *E. coli* cells containing positive cloned plasmids were selected with 100 µg/ml ampicillin. The structure of the vector was sequence verified. Plasmids were isolated with a QIAprep Spin Miniprep kit (Qiagen) according to the manufacturer's instructions and transformed into BL21-Gold (DE3)pLysS for expression. Viral gene was expressed by growing cells over-night at 37 °C in 10 ml of LB medium containing 100 µg/ml ampicillin. Then, 5 ml of the over-night culture was subcultured into 200 ml LB medium containing 100 µg/ml ampicillin. The batch culture was grown to an OD_{600} of 0.6 at 37 °C and then induced with 0.1 mM IPTG and incubated at 16 °C overnight. The cells were harvested by centrifugation at 5000 rpm for 5 min at 4°C and resuspended in 35 ml of PBS with 2 mM PMSF. After incubation on ice for 30 min, cells were disrupted by sonication for 3 min using a Tekmar sonic disruptor at 30% amplitude, in 5-s pulses. Samples were centrifuged at 10,000 rpm for 15 min to separate soluble and insoluble fractions.

Purification of recombinant A061L enzyme. Glutathione Sepharose 4B slurry (GE Healthcare) was loaded on a 5 ml self-packing column with a 45- to 90-µm-pore-size polyethylene filter (frit) (Life Science Products), and the resin was allowed to settle. The column was equilibrated with 5 column volumes of cold wash buffer (50 mM NaH_2PO_4, 150 mM NaCl, 1 mM DTT, 1 mM EDTA, pH 7.2). The soluble bacterial sonicate was applied to the column and allowed to drain. The column was washed again with 5 column volumes of cold wash buffer.

The recombinant protein was eluted with the GST moiety with 500 µl of cold elution buffer (50 mM Tris-HCl, 100 mM NaCl, 0.1% v/v Triton X-100, 50 mM reduced glutathione, 1 mM DTT, pH 8.0) and visualized by SDS-PAGE to access purity (Fig. S1C). Protein concentration was determined by NanoDrop spectrophotometer (NanoDrop Technologies).

Synthesis of tetrasaccharide 1 (Figs. S19-S20)

Octyl α-L-fucopyranoside (**S2**)**:** To a stirred solution of octanol (6.04 mL, 38.5 mmol) and 1,2,3,4-tetra-*O*-acetyl-L-fucopyranose **S1** (3) (5.12 g, 15.4 mmol) in dry CH_2Cl_2 (40 mL) was added tin (IV) chloride at 0 °C. The resulting solution was stirred for 1 h at room temperature under an Ar atmosphere. The crude residue was diluted with CH_2Cl_2 (200 mL), then washed with saturated $NaHCO_3$ (aq.). The organic layer was dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (4:1 hexane–EtOAc) to afford octyl 2,3,4-tri-*O*-acetyl-α-L-fucopyranoside which was contaminated with octanol. To a stirred mixture in CH_3OH (60 mL) was added a solution of $NaOCH_3$ in CH_3OH (6.0 mL, 0.5 M). The reaction mixture was stirred for 3 h at room temperature then neutralized by addition of Amberlite® IR-120 (H^+) cation exchange resin, filtered and the filtrate was concentrated to afford

S2 (1.70 g, 41%) as a white amorphous solid. R_f 0.41 (9:1 CH_2Cl_2–CH_3OH); $[\alpha]_D$ –106.4 (*c* 0.6, CH_3OH); ^1H NMR (400 MHz; CD_3OD): δ 4.72 (d, 1H, *J* = 2.7 Hz, H-1), 3.93 (qd, 1H, *J* = 6.6, 0.7 Hz, H-5), 3.71–3.70 (m, 2H, H-2, H-3), 3.67–3.61 (m, 2H, H-4, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.43 (dt, 1H, *J* = 9.7, 6.4 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 1.66–1.56 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.39–1.27 (m, 10H, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 1.19 (d, 3H, *J* = 6.6 Hz, H-6), 0.89 (t, 1H, *J* = 6.9 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; CD_3OD): δ 100.5 (C-1), 73.7 (C-4), 71.8 (C-2 or C-3), 70.1 (C-2 or C-3), 69.3 ($O\underline{C}H_2(CH_2)_6CH_3$), 67.5 (C-5), 33.0 ($OCH_2(\underline{C}H_2)_6CH_3$), 30.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 30.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 30.4 ($OCH_2(\underline{C}H_2)_6CH_3$), 27.3 ($OCH_2(\underline{C}H_2)_6CH_3$), 23.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 16.6 (C-6), 14.4 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{14}H_{28}NaO_5$: 299.1829; Found 299.1833.

Octyl 3,4-*O*-isopropylidene-α-L-fucopyranoside (S3): To a stirred solution of **S2** (783.1 mg, 2.83 mmol) in dry CH_3CN (24 mL) were added 2,2-dimethoxypropane (0.54 mL, 4.25 mmol) and *p*-toluenesulfonic acid monohydrate (53.3 mg, 0.28 mmol) successively, the reaction mixture was stirred at room temperature for 2 h. Triethylamine was added, the mixture was filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (4:1 hexane–EtOAc) to afford **S3** (873.2 mg, 97%) as a viscous oil. R_f 0.69 (1:1 hexane–EtOAc); $[\alpha]_D$ –105.1 (*c* 1.2, $CHCl_3$); ^1H NMR (400 MHz; $CDCl_3$): δ 4.82 (d, 1H, *J* = 4.0 Hz, H-1), 4.21 (t, 1H, *J* = 6.2 Hz, H-3), 4.13 (qd, 1H, *J* = 6.7, 2.2 Hz, H-5), 4.07 (dd, 1H, *J* = 6.0, 2.3 Hz, H-4), 3.82–3.73 (m, 2H, H-2, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.48 (dt, 1H, *J* = 9.7, 6.6 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 2.26 (d, 1H, *J* = 7.2 Hz, 2-OH), 1.65–1.59 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.53 (s, 3H, $C(C\underline{H}_3)_2$), 1.37-1.29 (m, 16H, $C(C\underline{H}_3)_2$, C-6, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 0.90 (t, 3H, *J* = 6.9 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 109.2 ($\underline{C}(CH_3)_2$), 97.3 (C-1), 76.3 (C-3), 75.7 (C-4), 69.4 (C-2), 68.3 ($O\underline{C}H_2(CH_2)_6CH_3$), 63.8 (C-5), 31.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.4 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.2 ($OCH_2(\underline{C}H_2)_6CH_3$), 27.8 ($C(\underline{C}H_3)_2$), 26.1 ($OCH_2(\underline{C}H_2)_6CH_3$), 25.9 ($C(\underline{C}H_3)_2$), 22.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 16.3 (C-6), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{17}H_{32}NaO_5$: 339.2142; Found 339.2147.

Octyl 3,4-*O*-isopropylidene-2-*O*-levulinoyl-α-L-fucopyranoside (S4): To a stirred solution of **S3** (873.2 mg, 2.76 mmol) in dry CH_2Cl_2 (20 mL) was added EDC·HCl (1.06 g, 5.52 mmol), levulinic acid (0.64 g, 5.52 mmol) and 4-(dimethylamino)pyridine (34.2 mg, 0.28 mmol) successively. The reaction mixture was stirred overnight at room temperature. Then poured into the saturated $NaHCO_3$ (aq.) and extracted with CH_2Cl_2 (60 mL × 3). The combined organic layers were dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (4:1 hexane–EtOAc) to afford **S4** (1.04 g, 91%) as a viscous oil. R_f 0.23 (4:1 hexane–EtOAc); $[\alpha]_D$ –133.6 (*c* 0.6, $CHCl_3$); ^1H NMR (400 MHz; $CDCl_3$): δ 4.90–4.87 (m, 2H, H-1, H-1), 4.34–4.30 (m, 1H, H-3), 4.14 (qd, 1H, *J* = 6.7, 2.5 Hz, H-5), 4.09 (dd, 1H, *J* = 5.4, 2.6 Hz, H-4), 3.65 (dt, 1H, *J* = 9.8, 6.8 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.42 (dt, 1H, *J* = 9.8, 6.6 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 2.86–2.73 (m, 2H, $COCH_2C\underline{H}_2COCH_3$), 2.72–2.62 (m, 2H, $COC\underline{H}_2CH_2COCH_3$), 2.20 (s, 3H, $COC\underline{H}_3$), 1.60–1.55 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.54 (s, 3H, $C(C\underline{H}_3)_2$), 1.38-1.27 (m, 16H, $C(C\underline{H}_3)_2$, C-6, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 0.90 (t, 3H, *J* = 6.9 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 206.2 (C=O), 172.3 (C=O), 109.3 ($\underline{C}(CH_3)_2$), 96.0 (C-1), 76.2 (C-4), 73.5 (C-3), 72.3 (C-2), 68.5 ($O\underline{C}H_2(CH_2)_6CH_3$), 63.1 (C-5), 38.0 ($COCH_2\underline{C}H_2COCH_3$), 31.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.8 ($CO\underline{C}H_3$), 29.4 (2 ×

$OCH_2(\underline{C}H_2)_6CH_3$), 29.3 ($OCH_2(\underline{C}H_2)_6CH_3$), 28.0 ($C(\underline{C}H_3)_2$, $CO\underline{C}H_2CH_2COCH_3$), 26.4 ($C(\underline{C}H_3)_2$), 26.1 ($OCH_2(\underline{C}H_2)_6CH_3$), 22.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 16.3 (C-6), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{22}H_{38}NaO_7$: 437.2510; Found 437.2509.

Octyl 2-*O*-levulinoyl-α-L-fucopyranoside (S5): To a stirred solution of **S4** (1.12 g, 2.70 mmol) in CH_3CN–CH_3OH (22 mL, 10:1) was added *p*-toluenesulfonic acid monohydrate (1.54 g, 8.10 mmol) at room temperature. The reaction mixture was stirred for 3 h at room temperature. Triethylamine was added, the mixture was filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (1:2 hexane–EtOAc) to afford **S5** (808.9 mg, 80%) as a viscous oil. R_f 0.17 (1:2 hexane–EtOAc); $[\alpha]_D$ –119.8 (*c* 0.9, $CHCl_3$); ^{1}H NMR (500 MHz; $CDCl_3$): δ 5.03 (dd, 1H, *J* = 10.1, 3.8 Hz, H-2), 4.93 (d, 1H, *J* = 3.8 Hz, H-1), 4.08–4.01 (m, 2H, H-3, H-5), 3.85–3.84 (m, 1H, H-4), 3.66 (dt, 1H, *J* = 9.8, 6.8 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.43 (dt, 1H, *J* = 9.8, 6.6 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 2.99 (d, 1H, *J* = 5.2 Hz, 3-OH), 2.88–2.76 (m, 2H, $COCH_2C\underline{H}_2COCH_3$), 2.71–2.58 (m, 2H, $COC\underline{H}_2CH_2COCH_3$), 2.44 (br s, 1H, 4-OH), 2.21 (s, 3H, $COC\underline{H}_3$), 1.63–1.57 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.37–1.26 (m, 13H, C-6, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 0.90 (t, 3H, *J* = 7.0 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 207.2 (C=O), 173.1 (C=O), 96.2 (C-1), 72.2 (C-4), 72.0 (C-2), 68.7 (C-3), 68.4 ($O\underline{C}H_2(CH_2)_6CH_3$), 65.4 (C-5), 38.3 ($COCH_2\underline{C}H_2COCH_3$), 31.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.8 ($CO\underline{C}H_3$), 29.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.4 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.3 ($OCH_2(\underline{C}H_2)_6CH_3$), 28.2 ($CO\underline{C}H_2CH_2COCH_3$), 26.1 ($OCH_2(\underline{C}H_2)_6CH_3$), 22.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 16.1 (C-6), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{19}H_{34}NaO_7$: 397.2197; Found 397.2195.

Octyl 4-*O*-acetyl-2-*O*-levulinoyl-α-L-fucopyranoside (S6): To a stirred solution of **S5** (689.8 mg, 1.84 mmol) in dry CH_3CN (18 mL) was added trimethyl orthoacetate (0.34 mL, 2.76 mmol) and *p*-toluenesulfonic acid monohydrate (35.0 mg, 0.18 mmol) at 0 °C. The reaction mixture was stirred for 1 h under an Ar atmosphere. Triethylamine was added, the mixture was filtered and the filtrate was concentrated. The reaction was then dissolved in CH_2Cl_2 (50 mL) and 1 N HCl was added (50 mL). The reaction mixture was stirred for 1 h at room temperature, then the aqueous layer was extracted with CH_2Cl_2 (50 mL × 3) and combined organic layers were dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (1:1 hexane–EtOAc) to afford **S6** (748.4 mg, 98%) as a white solid. R_f 0.23 (1:1 hexane–EtOAc); $[\alpha]_D$ –122.2 (*c* 0.6, $CHCl_3$); ^{1}H NMR (400 MHz; $CDCl_3$): δ 5.27 (dd, 1H, *J* = 3.6, 1.3 Hz, H-4), 5.01 (dd, 1H, *J* = 10.1, 3.7 Hz, H-2), 4.97 (d, 1H, *J* = 3.8 Hz, H-1), 4.24 (dt, 1H, *J* = 9.9, 4.1 Hz, 1H), 4.10 (qd, 1H, *J* = 6.6, 0.9 Hz, H-5), 3.66 (dt, 1H, *J* = 9.8, 6.7 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.44 (dt, 1H, *J* = 9.8, 6.6 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 2.81–2.78 (m, 2H, $COCH_2C\underline{H}_2COCH_3$), 2.67–2.63 (m, 2H, $COC\underline{H}_2CH_2COCH_3$), 2.34 (d, 1H, *J* = 5.0 Hz, 3-OH), 2.20 (s, 3H, $COC\underline{H}_3$), 2.20 (s, 3H, $COC\underline{H}_3$), 1.62–1.56 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.37–1.27 (m, 10H, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 1.16 (d, 3H, *J* = 6.6 Hz, H-6), 0.90 (t, 3H, *J* = 6.9 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 206.7 (C=O), 172.9 (C=O), 171.3 (C=O), 96.3 (C-1), 73.5 (C-4), 71.8 (C-2), 68.6 ($O\underline{C}H_2(CH_2)_6CH_3$), 67.1 (C-3), 64.7 (C-5), 38.1 ($COCH_2\underline{C}H_2COCH_3$), 31.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.8 ($CO\underline{C}H_3$), 29.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.4 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.3 ($OCH_2(\underline{C}H_2)_6CH_3$), 28.1 ($CO\underline{C}H_2CH_2COCH_3$), 26.1 ($OCH_2(\underline{C}H_2)_6CH_3$), 22.7

($OCH_2(\underline{C}H_2)_6CH_3$), 20.8 ($CO\underline{C}H_3$), 16.2 (C-6), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{21}H_{36}NaO_8$: 439.2302; Found 439.2304.

***p*-Tolyl 6-*O*-(*p*-toluenesulfonyl)-1-thio-α-D-mannopyranoside (S8):** To a stirred solution of 1-thio-α-D-mannopyranoside **S7** (4) (5.15 g, 18.0 mmol) in dry pyridine (60 mL) was added *p*-toluenesulfonyl chloride (5.15 g, 27.0 mmol) at 0 °C. The reaction mixture was stirred for 18 h at 0 °C. The solvent was evaporated and the crude residue was purified by flash chromatography (1:4 hexane–EtOAc) to afford **S8** (4.76 g, 60%) as a white foam. R_f 0.53 (9:1 CH_2Cl_2–CH_3OH); $[\alpha]_D$ +141.7 (*c* 0.6, $CHCl_3$); ^{1}H NMR (400 MHz; $CDCl_3$): *δ* 7.80–7.77 (m, 2H, Ar), 7.32–7.29 (m, 4H, Ar), 7.10–7.08 (m, 2H, Ar), 5.42 (d, 1H, *J* = 1.1 Hz, H-1), 4.43 (dd, 1H, *J* = 11.0, 4.4 Hz, H-6a), 4.30 (ddd, 1H, *J* = 9.4, 4.4, 1.9 Hz, H-5), 4.25 (dd, 1H, *J* = 11.0, 2.0 Hz, H-6b), 4.21 (br s, 1H, H-2), 3.89 (t, 1H, *J* = 9.4 Hz, H-4), 3.83 (dd, 1H, *J* = 9.4, 2.7 Hz, H-3), 3.33 (br s, 2H, 3-OH, 4-OH), 3.16 (br s, 1H, 2-OH), 2.44 (s, 3H, Ar-$C\underline{H}_3$), 2.33 (s, 3H, Ar-$C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): *δ* 145.0 (Ar), 137.9 (Ar), 132.6 (Ar), 132.2 (Ar), 129.9 (2 × Ar), 129.6 (Ar), 128.0 (Ar), 88.4 (C-1), 72.1 (C-3), 72.0 (C-2), 70.9 (C-5), 68.9 (C-6), 67.5 (C-4), 21.7 (Ar-$\underline{C}H_3$), 21.1 (Ar-$\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{20}H_{24}NaO_7S_2$: 463.0856; Found 463.0865.

***p*-Tolyl 2,3-*O*-isopropylidene-6-*O*-(*p*-toluenesulfonyl)-1-thio-α-D-mannopyranoside (S9):** To a stirred solution of **S8** (3.75 g, 8.51 mmol) in dry CH_3CN (25 mL) were added 2,2-dimethoxypropane (1.56 mL, 12.75 mmol) and *p*-toluenesulfonic acid monohydrate (323.0 mg, 1.70 mmol) successively, the reaction mixture was stirred at room temperature for 2 h. Triethylamine was added, the mixture was filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (3:1 hexane–EtOAc) to afford **S9** (3.45 g, 84%) as a white foam. R_f 0.37 (2:1 hexane–EtOAc); $[\alpha]_D$ +98.3 (*c* 0.8, $CHCl_3$); ^{1}H NMR (500 MHz; $CDCl_3$): *δ* 7.75–7.73 (m, 2H, Ar), 7.36–7.34 (m, 2H, Ar), 7.32–7.30 (m, 2H, Ar), 7.13–7.12 (m, 2H, Ar), 5.65 (s, 1H, H-1), 4.35–4.31 (m, 2H, H-2, H-6a), 4.2–-4.18 (m, 2H, H-5, H-6b), 4.15 (dd, 1H, *J* = 7.5, 5.6 Hz, H-3), 3.72 (ddd, 1H, *J* = 10.3, 7.6, 4.2 Hz, H-4), 2.60 (d, 1H, *J* = 4.4 Hz, 4-OH), 2.46 (s, 3H, Ar-$C\underline{H}_3$), 2.35 (s, 3H, Ar-$C\underline{H}_3$), 1.52 (s, 3H, $C(C\underline{H}_3)_2$), 1.38 (s, 3H, $C(C\underline{H}_3)_2$); ^{13}C NMR (125 MHz; $CDCl_3$): *δ* 144.9 (Ar), 138.2 (Ar), 132.8 (Ar), 132.7 (Ar), 129.9 (Ar), 129.8 (Ar), 128.8 (Ar), 128.0 (Ar), 110.0 ($\underline{C}(CH_3)_2$), 84.4 (C-1), 78.0 (C-3), 76.1 (C-2), 69.2 (C-4), 68.9 (C-5), 68.5 (C-6), 28.1 ($C(\underline{C}H_3)_2$), 26.3 ($C(\underline{C}H_3)_2$), 21.7 (Ar-$\underline{C}H_3$), 21.2 (Ar-$\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{23}H_{28}NaO_7S_2$: 503.1169; Found 503.1173.

***p*-Tolyl 4-*O*-benzyl-2,3-*O*-isopropylidene-6-*O*-(*p*-toluenesulfonyl)-1-thio-α-D-mannopyranoside (S10):** To a stirred solution of **S9** (1.22 g, 2.55 mmol) in dry DMF (5.0 mL) was added sodium hydride (204.0 mg, 5.10 mmol, 60% dispersion in mineral oil) in one portion at 0 °C, After stirring for 30 min, benzyl bromide (0.36 mL, 3.06 mmol) was added dropwise and the reaction mixture was warmed to room temperature and stirred for 1 h under an Ar atmosphere. Then, CH_3OH was added at 0 °C and the solution was concentrated. The crude residue was diluted with CH_2Cl_2 (50 mL), washed with saturated $NaHCO_3$ (aq.), dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (6:1 hexane–EtOAc) to afford **S10** (1.39 g, 96%) as a syrup. R_f 0.59 (3:1 hexane–EtOAc); $[\alpha]_D$ +141.1 (*c* 1.0, $CHCl_3$); ^{1}H NMR (500 MHz; $CDCl_3$): *δ* 7.70–7.69 (m, 2H, Ar), 7.37–7.29 (m, 7H, Ar), 7.26–7.25 (m, 2H,

Ar), 7.12–7.11 (m, 2H, Ar), 5.60 (s, 1H, H-1), 4.89 (d, 1H, J = 11.3 Hz, PhC$\underline{H}_2$), 4.53 (d, 1H, J = 11.3 Hz, PhC$\underline{H}_2$), 4.34–4.28 (m, 3H, H-2, H-3, H-5), 4.23 (dd, 1H, J = 10.5, 1.9 Hz, H-6a), 4.18 (dd, 1H, J = 10.6, 5.3 Hz, H-6b), 3.51 (dd, 1H, J = 10.2, 6.5 Hz, H-4), 2.43 (s, 3H, Ar-C$\underline{H}_3$), 2.35 (s, 3H, Ar-C$\underline{H}_3$), 1.51 (s, 3H, C(C$\underline{H}_3$)$_2$), 1.39 (s, 3H, C(C$\underline{H}_3$)$_2$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 144.6 (Ar), 138.1 (Ar), 137.7 (Ar), 132.9 (Ar), 132.7 (Ar), 129.9 (Ar), 129.7 (Ar), 129.0 (Ar), 128.4 (Ar), 128.0 (3 × Ar), 127.8 (Ar), 109.7 (C$\underline{C}$(CH$_3$)$_2$), 84.4 (C-1), 78.3 (C-3), 76.2 (C-2), 75.4 (C-4), 72.9 (Ph$\underline{C}$H$_2$), 68.9 (C-6), 67.9 (C-5), 28.0 (C(C$\underline{H}_3$)$_2$), 26.4 (C(C$\underline{H}_3$)$_2$), 21.6 (Ar-C$\underline{H}_3$), 21.2 (Ar-C$\underline{H}_3$); HRMS (ESI) Calc. for [M + Na]$^+$ $C_{30}H_{34}NaO_7S_2$: 593.1638; Found 593.1641.

***p*-Tolyl 4-*O*-benzyl-2,3-*O*-isopropylidene-1-thio-α-D-rhamnopyranoside (S11):** To a stirred solution of **S10** (1.29 g, 2.26 mmol) in dry THF (20 mL) was added lithium aluminum hydride (214.0 mg, 5.65 mmol) at room temperature. The reaction mixture was stirred for 2.5 h at 65 °C, then cooled to 0 °C and ethyl acetate was added. The crude residue was then washed with 1 N HCl and saturated $NaHCO_3$ (aq.), the aqueous layers were extracted with EtOAc (60 mL), and the combined organic layers were dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (15:1 hexane–EtOAc) to afford **S11** (843.4 mg, 93%) as a white solid. R_f 0.54 (9:1 hexane–EtOAc); $[\alpha]_D$ +199.9 (*c* 2.4, $CHCl_3$); ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.41–7.29 (m, 7H, Ar), 7.15–7.13 (m, 2H, Ar), 5.68 (d, 1H, J = 0.3 Hz, H-1), 4.94 (d, 1H, J = 11.6 Hz, PhC$\underline{H}_2$), 4.66 (d, 1H, J = 11.6 Hz, PhC$\underline{H}_2$), 4.38–4.32 (m, 2H, H-2, H-3), 4.18 (dq, 1H, J = 9.8, 6.1 Hz, H-5), 3.32 (dd, 1H, J = 9.8, 6.9 Hz, H-4), 2.35 (s, 3H, Ar-C$\underline{H}_3$), 1.54 (s, 3H, C(C$\underline{H}_3$)$_2$), 1.41 (s, 3H, C(C$\underline{H}_3$)$_2$), 1.26 (d, 3H, J = 6.2 Hz, H-6); ^{13}C NMR (125 MHz; $CDCl_3$): δ 138.3 (Ar), 137.8 (Ar), 132.5 (Ar), 129.8 (Ar), 129.7 (Ar), 128.3 (Ar), 128.0 (Ar), 127.7 (Ar), 109.4 (C$\underline{C}$(CH$_3$)$_2$), 84.2 (C-1), 81.5 (C-4), 78.5 (C-3), 76.7 (C-2), 73.1 (Ph$\underline{C}$H$_2$), 66.2 (C-5), 28.1 (C(C$\underline{H}_3$)$_2$), 26.5 (C(C$\underline{H}_3$)$_2$), 21.1 (Ar-C$\underline{H}_3$), 17.7 (C-6); HRMS (ESI) Calc. for [M + Na]$^+$ $C_{23}H_{28}NaO_4S$: 423.1601; Found 423.1610.

Octyl 4-*O*-benzyl-2,3-*O*-isopropylidene-α-D-rhamnopyranosyl-(1→3)-4-*O*-acetyl-2-*O*-levulinoyl-α-L-fucopyranoside (S12): To a stirred solution of acceptor **S6** (748.4 mg, 1.80 mmol) and donor **S11** (826.2 mg, 2.06 μmol) in dry CH_2Cl_2 (20 mL) was added molecular sieves (2.0 g, 4 Å, powder). After stirring for 30 min at room temperature, the reaction mixture was cooled to –25 °C, and then *N*-iodosuccinimide (566.9 mg, 2.52 mmol) and silver trifluoromethanesulfonate (92.5 mg, 0.36 mmol) were added successively. The resulting solution was stirred for 1 h at –25 °C under an Ar atmosphere. Triethylamine was added, the mixture was filtered and the filtrate was washed with saturated $Na_2S_2O_3$ (aq.) and saturated $NaHCO_3$ (aq.). The aqueous layer was extracted with CH_2Cl_2 (30 mL × 3), dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (4:1 hexane–EtOAc) to afford **S12** (1.06 g, 85%) as a syrup. R_f 0.51 (2:1 hexane–EtOAc); $[\alpha]_D$ –50.1 (*c* 1.3, $CHCl_3$); ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.39–7.26 (m, 5H, Ar), 5.24 (dd, 1H, J = 3.6, 1.1 Hz, Fuc-H-4), 5.14 (s, 1H, Rha-H-1), 5.09 (dd, 1H, J = 10.5, 3.7 Hz, Fuc-H-2), 4.99 (d, 1H, J = 3.7 Hz, Fuc-H-1), 4.88 (d, 1H, J = 11.8 Hz, PhC$\underline{H}_2$), 4.65 (d, 1H, J = 11.8 Hz, PhC$\underline{H}_2$), 4.26 (dd, 1H, J = 10.5, 3.5 Hz, Fuc-H-3), 4.20–4.15 (m, 2H, Rha-H-3, Rha-H-2), 4.10 (qd, 1H, J = 6.8, 0.8 Hz, Fuc-H-5), 3.73 (dq, 1H, J = 9.8, 6.1 Hz, Rha-H-5), 3.65 (dt, 1H, J = 9.8, 6.6 Hz, OC$\underline{H}_2$CH$_2$(CH$_2$)$_5$CH$_3$), 3.42 (dt, 1H, J = 9.8, 6.6 Hz, OC$\underline{H}_2$CH$_2$(CH$_2$)$_5$CH$_3$), 3.18 (dd, 1H, J = 9.8, 6.7 Hz,

Rha-H-4), 2.84–2.70 (m, 2H, $COCH_2C\underline{H}_2COCH_3$), 2.68–2.55 (m, 2H, $COC\underline{H}_2CH_2COCH_3$), 2.19 (s, 3H, $COC\underline{H}_3$), 2.16 (s, 3H, $COC\underline{H}_3$), 1.60–1.55 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.49 (s, 3H, $C(C\underline{H}_3)_2$), 1.36–1.27 (m, 16H, $C(C\underline{H}_3)_2$, Rha-H-6, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 1.12 (d, 3H, J = 6.6 Hz, Fuc-H-6), 0.90 (t, 3H, J = 6.9 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 205.8 (C=O), 172.2 (C=O), 170.5 (C=O), 138.5 (Ar), 128.2 (Ar), 128.0 (Ar), 127.5 (Ar), 108.9 ($\underline{C}(CH_3)_2$), 99.7 (Rha-C-1, $^1J_{C\text{-}H}$ = 172.2 Hz), 96.2 (Fuc-C-1), 80.7 (Rha-C-4), 78.4 (Rha-C-3), 76.2 (Rha-C-2), 73.3 (Fuc-C-4), 72.9 ($Ph\underline{C}H_2$), 72.8 (Fuc-C-3), 71.2 (Fuc-C-2), 68.6 ($O\underline{C}H_2(CH_2)_6CH_3$), 65.4 (Rha-C-5), 65.0 (Fuc-C-5), 37.8 ($COCH_2\underline{C}H_2COCH_3$), 31.9 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.8 ($CO\underline{C}H_3$), 29.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.4 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.3 ($OCH_2(\underline{C}H_2)_6CH_3$), 28.0 ($CO\underline{C}H_2CH_2COCH_3$, $C(\underline{C}H_3)_2$), 26.3 ($C(\underline{C}H_3)_2$), 26.1 ($OCH_2(\underline{C}H_2)_6CH_3$), 22.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 20.8 ($CO\underline{C}H_3$), 17.7 (Rha-C-6), 16.1 (Fuc-C-6), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{37}H_{56}NaO_{12}$: 715.3664; Found 715.3661.

Octyl 4-*O*-benzyl-2,3-*O*-isopropylidene-α-D-rhamnopyranosyl-(1→3)-4-*O*-acetyl-α-L-fucopyranoside (S13): To a stirred solution of **S12** (989.4 mg, 1.43 mmol) in CH_2Cl_2–CH_3OH (20 mL, 9:1) was added hydrazine acetate (263.0 mg, 2.86 mmol) at room temperature. The reaction mixture was stirred for 2 h at room temperature. The reaction mixture was then concentrated and the crude residue was purified by flash chromatography (3:1 hexane–EtOAc) to afford **S13** (814.3 mg, 96%) as a syrup. R_f 0.13 (4:1 hexane–EtOAc); $[\alpha]_D$ –53.2 (*c* 1.2, $CHCl_3$); ^{1}H NMR (700 MHz; $CDCl_3$): δ 7.35–7.30 (m, 4H, Ar), 7.27–7.24 (m, 1H, Ar), 5.28 (s, 1H, Rha-H-1), 5.15 (d, 1H, J = 2.9 Hz, Fuc-H-4), 4.88–4.85 (m, 2H, $PhC\underline{H}_2$, Fuc-H-1), 4.63 (d, 1H, J = 11.8 Hz, $PhC\underline{H}_2$), 4.21–4.17 (m, 2H, Rha-H-2, Rha-H-3), 4.02 (q, 1H, J = 6.6 Hz, Fuc-H-5), 3.92 (dd, 1H, J = 10.0, 3.3 Hz, Fuc-H-3), 3.88 (td, 1H, J = 10.1, 3.9 Hz, Fuc-H-2), 3.74–3.68 (m, 2H, Rha-H-5, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.44 (dt, 1H, J = 9.7, 6.6 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.17 (dd, 1H, J = 9.8, 7.1 Hz, Rha-H-4), 2.12 (s, 3H, $COC\underline{H}_3$), 1.94 (d, 1H, J = 10.3 Hz, 2-OH), 1.62–1.58 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.47 (s, 3H, $C(C\underline{H}_3)_2$), 1.35–1.24 (m, 16H, $C(C\underline{H}_3)_2$, Rha-H-6, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 1.10 (d, 3H, J = 6.6 Hz, Fuc-H-6), 0.88 (t, 3H, J = 7.1 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (175 MHz; $CDCl_3$): δ 170.4 (C=O), 138.5 (Ar), 128.2 (Ar), 127.9 (Ar), 127.5 (Ar), 108.9 ($\underline{C}(CH_3)_2$), 99.3 (Rha-C-1), 98.6 (Fuc-C-1), 80.8 (Rha-C-4), 78.4 (Rha-C-3), 76.1 (Rha-C-2), 75.4 (Fuc-C-3), 73.1 (Fuc-C-4), 72.9 ($Ph\underline{C}H_2$), 69.3 (Fuc-C-2), 68.5 ($O\underline{C}H_2(CH_2)_6CH_3$), 65.4 (Fuc-C-5), 65.3 (Rha-C-5), 31.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.3 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.2 ($OCH_2(\underline{C}H_2)_6CH_3$), 28.0 ($C(\underline{C}H_3)_2$), 26.4 ($C(\underline{C}H_3)_2$), 26.1 ($OCH_2(\underline{C}H_2)_6CH_3$), 22.6 ($OCH_2(\underline{C}H_2)_6CH_3$), 20.8 ($CO\underline{C}H_3$), 17.7 (Rha-C-6), 16.2 (Fuc-C-6), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{32}H_{50}NaO_{10}$: 617.3296; Found 617.3294.

Octyl 4-*O*-benzyl-2,3-*O*-isopropylidene-α-D-rhamnopyranosyl-(1→3)-[2,3-di-*O*-benzyl-4,6-*O*-di-*tert*-butylsilylene-α-D-galactopyranosyl-(1→2)]-4-*O*-acetyl-α-L-fucopyranoside (S15): To a stirred solution of acceptor **S13** (258.8 mg, 435.2 μmol) and 2,3-di-*O*-benzyl-4,6-*O*-di-*tert*-butylsilylene-α-D-galactopyranoside **S14** (5) (316.9 mg, 522.2 μmol) in dry CH_2Cl_2 (10 mL) was added molecular sieves (1.0 g, 4 Å, powder). After stirring for 30 min at room temperature, the reaction mixture was cooled to –10 °C, and then *N*-iodosuccinimide (146.9 mg, 652.8 μmol) and silver trifluoromethanesulfonate (22.4 mg, 87.0

μmol) were added successively. The resulting solution was stirred for 1 h at 0 °C under an Ar atmosphere. Triethylamine was added, the mixture was filtered and the filtrate was washed with saturated $Na_2S_2O_3$ (aq.) and saturated $NaHCO_3$ (aq.), the aqueous layer was extracted with CH_2Cl_2 (30 mL × 3), dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (8:1 hexane–EtOAc) to afford **S15** (438.6 mg, 94%) as a syrup. R_f 0.56 (4:1 hexane–EtOAc); $[\alpha]_D$ –3.4 (*c* 1.3, $CHCl_3$); ^{1}H NMR (700 MHz; $CDCl_3$): δ 7.43–7.42 (m, 2H, Ar), 7.36–7.24 (m, 12H, Ar), 5.48 (s, 1H, Rha-H-1), 5.17 (d, 1H, *J* = 2.6 Hz, Fuc-H-4), 4.89 (d, 1H, *J* = 3.6 Hz, Fuc-H-1), 4.87 (d, 1H, *J* = 11.9 Hz, PhC$\underline{H}_2$), 4.81 (d, 1H, *J* = 12.4 Hz, PhC$\underline{H}_2$), 4.73–4.70 (m, 3H, Gal-H-1, PhC$\underline{H}_2$), 4.68 (d, 1H, *J* = 12.4 Hz, PhC$\underline{H}_2$), 4.63 (d, 1H, *J* = 11.9 Hz, PhC$\underline{H}_2$), 4.38 (d, 1H, *J* = 2.6 Hz, Gal-H-4), 4.24 (d, 1H, *J* = 5.6 Hz, Rha-H-2), 4.21 (dd, 1H, *J* = 10.2, 3.5 Hz, Fuc-H-3), 4.14 (dd, 1H, *J* = 12.3, 2.0 Hz, Gal-H-6a), 4.08 (dd, 1H, *J* = 7.3, 5.7 Hz, Rha-H-3), 4.00–3.96 (m, 2H, Fuc-H-5, Gal-H-6b), 3.88 (dd, 1H, *J* = 10.1, 3.5 Hz, Gal-H-2), 3.83–3.78 (m, 2H, Fuc-H-2, Rha-H-5), 3.77–3.74 (m, 2H, Gal-H-5, Gal-H-3), 3.57 (dt, 1H, *J* = 9.5, 7.0 Hz, OC$\underline{H}_2$CH$_2$(CH$_2$)$_5$CH$_3$), 3.19–3.14 (m, 2H, Rha-H-4, OC$\underline{H}_2$CH$_2$(CH$_2$)$_5$CH$_3$), 2.20 (s, 3H, COC$\underline{H}_3$), 1.53–1.45 (m, 2H, OCH$_2$C$\underline{H}_2$(CH$_2$)$_5$CH$_3$), 1.41 (s, 3H, C(C$\underline{H}_3$)$_2$), 1.28–1.22 (m, 13H, Rha-H-6, OCH$_2$CH$_2$(C$\underline{H}_2$)$_5$CH$_3$), 1.11 (s, 3H, C(C$\underline{H}_3$)$_2$), 1.09 (d, 3H, *J* = 6.6 Hz, Fuc-H-6), 1.02 (s, 9H, C(C$\underline{H}_3$)$_3$), 0.91 (s, 9H, C(C$\underline{H}_3$)$_3$), 0.87 (t, 3H, *J* = 7.1 Hz, OCH$_2$CH$_2$(CH$_2$)$_5$C$\underline{H}_3$); ^{13}C NMR (175 MHz; $CDCl_3$): δ 170.4 (C=O), 139.4 (Ar), 138.9 (Ar), 138.7 (Ar), 128.3 (Ar), 128.2 (2 × Ar), 128.1 (Ar), 127.9 (Ar), 127.7 (Ar), 127.4 (2 × Ar) , 127.3 (Ar), 108.7 (C(CH$_3$)$_2$), 101.2 (Gal-C-1, $^1J_{C\text{-}H}$ = 169.5 Hz), 98.7 (Rha-C-1), 98.3 (Fuc-C-1), 81.0 (Rha-C-4), 79.3 (Fuc-C-2), 78.3 (Rha-C-3), 77.6 (Gal-C-3), 76.0 (Rha-C-2), 73.9 (Gal-C-2), 73.5 (Fuc-C-4), 73.2 (PhCH$_2$), 72.7 (PhCH$_2$), 71.6 (Gal-C-4), 71.3 (PhCH$_2$), 71.0 (Fuc-C-3), 68.1 (OCH$_2$(CH$_2$)$_6$CH$_3$), 68.0 (Gal-C-5), 67.3 (Gal-C-6), 64.9 (Rha-C-5), 64.6 (Fuc-C-5), 31.8 (OCH$_2$(CH$_2$)$_6$CH$_3$), 29.7 (OCH$_2$(CH$_2$)$_6$CH$_3$), 29.4 (OCH$_2$(CH$_2$)$_6$CH$_3$), 29.3 (OCH$_2$(CH$_2$)$_6$CH$_3$), 28.1 (C(CH$_3$)$_2$), 27.6 (C(CH$_3$)$_3$), 27.2 (C(CH$_3$)$_3$), 26.4 (C(CH$_3$)$_2$), 26.3 (OCH$_2$(CH$_2$)$_6$CH$_3$), 23.4 (C(CH$_3$)$_3$), 22.6 (OCH$_2$(CH$_2$)$_6$CH$_3$), 20.8 (COCH$_3$), 20.6 (C(CH$_3$)$_3$), 17.8 (Rha-C-6), 16.1 (Fuc-C-6), 14.1 (OCH$_2$(CH$_2$)$_6$CH$_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{60}H_{88}NaO_{15}Si$: 1099.5785; Found 1099.5790.

Octyl 4-*O*-benzyl-2,3-*O*-isopropylidene-α-D-rhamnopyranosyl-(1→3)-[2,3-di-*O*-benzyl-4,6-*O*-di-*tert*-butylsilylene-α-D-galactopyranosyl-(1→2)]-α-L-fucopyranoside (S16): To a stirred solution of **S15** (438.6 mg, 407.1 μmol) in CH_3OH (8.0 mL) was added a solution of $NaOCH_3$ in CH_3OH (1.0 mL, 0.5 M). The reaction mixture was stirred for 7 h at room temperature, then neutralized by addition of Amberlite® IR-120 (H^+) cation exchange resin, filtered and the filtrate was concentrated to afford **S16** (421.4 mg, quant.) as a syrup. R_f 0.37 (4:1 hexane–EtOAc); $[\alpha]_D$ +6.4 (*c* 0.9, $CHCl_3$); ^{1}H NMR (700 MHz; $CDCl_3$): δ 7.44–7.43 (m, 2H, Ar), 7.36–7.23 (m, 13H, Ar), 5.61 (s, 1H, Rha-H-1), 4.90 (d, 1H, *J* = 11.6 Hz, PhC$\underline{H}_2$), 4.86 (d, 1H, *J* = 3.6 Hz, Fuc-H-1), 4.81 (d, 1H, *J* = 12.5 Hz, PhC$\underline{H}_2$), 4.72–4.70 (m, 4H, 3 × PhC$\underline{H}_2$, Gal-H-1), 4.61 (d, 1H, *J* = 11.6 Hz, PhC$\underline{H}_2$), 4.40 (d, 1H, *J* = 2.6 Hz, Gal-H-4), 4.32 (d, 1H, *J* = 5.7 Hz, Rha-H-2), 4.21 (t, 1H, *J* = 6.3 Hz, Rha-H-3), 4.13 (dd, 1H, *J* = 12.3, 1.9 Hz, Gal-H-6a), 4.10 (dd, 1H, *J* = 10.0, 3.3 Hz, Fuc-H-3), 3.97 (dd, 1H, *J* = 12.3, 1.3 Hz, Gal-6-Hb), 3.91–3.84 (m, 3H, Fuc-H-5, Gal-H-2, Fuc-H-2), 3.79–3.70 (m, 4H, Gal-H-3, Gal-H-5, Rha-H-5, Fuc-H-4), 3.58 (dt, 1H, *J* = 9.5, 7.0 Hz, OC$\underline{H}_2$CH$_2$(CH$_2$)$_5$CH$_3$), 3.24 (dd, 1H, *J* = 9.6, 7.1 Hz, Rha-H-4), 3.14 (dt, 1H, *J* = 9.4, 6.8 Hz, OC$\underline{H}_2$CH$_2$(CH$_2$)$_5$CH$_3$), 2.23 (s, 1H, Fuc-4-OH), 1.52–

1.45 (m, 5H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$, $C(C\underline{H}_3)_2$), 1.29–1.24 (m, 16H, Rha-H-6, Fuc-H-6, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 1.14 (s, 3H, $C(C\underline{H}_3)_2$), 1.01 (s, 9H, $C(C\underline{H}_3)_3$), 0.88–0.86 (m, 12H, $C(C\underline{H}_3)_3$, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 139.4 (Ar), 138.8 (Ar), 138.3 (Ar), 128.3 (3 × Ar), 128.17 (Ar), 128.0 (Ar), 127.7 (Ar), 127.6 (Ar), 127.4 (Ar), 127.3 (Ar), 109.0 ($\underline{C}(CH_3)_2$), 101.4 (Gal-C-1), 98.5 (Rha-C-1), 98.2 (Fuc-C-1), 81.2 (Rha-C-4), 78.8 (Fuc-C-2), 78.3 (Rha-C-3), 77.7 (Gal-C-3), 75.8 (Gal-C-4), 73.4 (Gal-C-2), 73.3 (Fuc-C-3), 73.0 (2 × $Ph\underline{C}H_2$), 72.3 (Fuc-C-4), 71.5 (Gal-C-4), 71.1 ($Ph\underline{C}H_2$), 68.0 (Gal-C-5, $O\underline{C}H_2(CH_2)_6CH_3$), 67.3 (Gal-C-6), 65.0 (Rha-C-5, Fuc-C-5), 31.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.4 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.3 ($OCH_2(\underline{C}H_2)_6CH_3$), 28.1 ($C(\underline{C}H_3)_2$), 27.6 ($C(\underline{C}H_3)_3$), 27.2 ($C(\underline{C}H_3)_3$), 26.4 ($C(\underline{C}H_3)_2$, $OCH_2(\underline{C}H_2)_6CH_3$), 23.4 ($\underline{C}(CH_3)_3$), 22.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 20.6 ($\underline{C}(CH_3)_3$), 18.0 (Rha-C-6), 16.0 (Fuc-C-6), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{58}H_{86}NaO_{14}Si$: 1057.5679; Found 1057.5681.

Octyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→4)-[4-*O*-benzyl-2,3-*O*-isopropylidene-α-D-rhamnopyranosyl-(1→3)]-[2,3-di-*O*-benzyl-4,6-*O*-di-*tert*-butylsilylene-α-D-galactopyranosyl-(1→2)]-α-L-fucopyranoside (S18): To a stirred solution of acceptor **S16** (383.9 mg, 370.8 μmol) and *p*-tolyl 2,3,4-tri-*O*-acetyl-1-thio-β-D-xylopyranoside **S17** (6) (351.8 mg, 919.9 μmol) in dry CH_2Cl_2 (8.0 mL) was added molecular sieves (800 mg, 4 Å, powder). After stirring for 30 min at room temperature, the reaction mixture was cooled to 0 °C, and then *N*-iodosuccinimide (32.2 mg, 143.0 μmol) and silver trifluoromethanesulfonate (4.9 mg, 19.1 μmol) were added successively. The resulting solution was stirred for 1 h at room temperature under an Ar atmosphere. Triethylamine was added, the mixture was filtered and the filtrate was washed with saturated $Na_2S_2O_3$ (aq.) and saturated $NaHCO_3$ (aq.), the aqueous layer was extracted with CH_2Cl_2 (25 mL × 3), dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (4:1 hexane–EtOAc) to afford **S18** (273.5 mg, 57%) as a syrup. R_f 0.32 (3:1 hexane–EtOAc); $[\alpha]_D$ –21.2 (*c* 0.9, $CHCl_3$); ^{1}H NMR (700 MHz; $CDCl_3$): δ 7.43–7.42 (m, 2H, Ar), 7.35–7.23 (m, 13H, Ar), 5.51 (s, 1H, Rha-H-1), 5.11 (t, 1H, *J* = 7.4 Hz, Xyl-H-3), 5.05 (dd, 1H, *J* = 7.4, 5.7 Hz, Xyl-H-2), 4.98 (td, 1H, *J* = 7.0, 4.6 Hz, Xyl-H-4), 4.89 (d, 1H, *J* = 11.7 Hz, $PhC\underline{H}_2$), 4.82 (d, 1H, *J* = 3.6 Hz, Fuc-H-1), 4.81 (d, 1H, *J* = 12.4 Hz, $PhC\underline{H}_2$), 4.74–4.69 (m, 4H, Gal-H-1, 3 × $PhC\underline{H}_2$), 4.59–4.58 (m, 2H, $PhC\underline{H}_2$, Xyl-H-1), 4.37 (d, 1H, *J* = 2.7 Hz, Gal-H-4), 4.34–4.31 (m, 2H, Xyl-H-5a, Rha-H-2), 4.15–4.10 (m, 3H, Rha-H-3, Gal-H-6a, Fuc-H-3), 3.96–3.91 (m, 2H, Gal-H-6b, Rha-H-5), 3.88–3.82 (m, 3H, Gal-H-2, Fuc-H-5, Fuc-H-2), 3.76 (br s, 1H, Gal-H-5), 3.74 (dd, 1H, *J* = 10.1, 2.8 Hz, Gal-H-3), 3.63 (d, 1H, *J* = 2.9 Hz, Fuc-H-4), 3.5 (dt, 1H, *J* = 9.5, 7.1 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.36 (dd, 1H, *J* = 12.2, 7.1 Hz, Xyl-H-5b), 3.20 (dd, 1H, *J* = 9.7, 7.1 Hz, Rha-H-4), 3.12 (dt, 1H, *J* = 9.3, 6.9 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 2.06 (s, 3H, $COC\underline{H}_3$), 2.03 (s, 3H, $COC\underline{H}_3$), 1.99 (s, 3H, $COC\underline{H}_3$), 1.49–1.41 (m, 5H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$, $C(C\underline{H}_3)_2$), 1.29–1.24 (m, 13H, Rha-H-6, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 1.18 (d, 3H, *J* = 6.6 Hz, Fuc-H-6), 1.11 (s, 3H, $C(C\underline{H}_3)_2$), 1.01 (s, 9H, $C(C\underline{H}_3)_3$), 0.89 (s, 9H, $C(C\underline{H}_3)_3$), 0.86 (t, 3H, *J* = 7.0 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 170.0 (C=O), 169.7 (C=O), 169.4 (C=O), 139.5 (Ar), 138.8 (2 × Ar), 128.4 (Ar), 128.3 (Ar), 128.1 (Ar), 127.9 (Ar), 127.8 (Ar), 127.4 (2 × Ar), 127.2 (Ar), 108.9 ($\underline{C}(CH_3)_2$), 101.2 (Gal-C-1), 100.9 (Xyl-C-1, $^1J_{C\text{-}H}$ = 165.6 Hz), 98.4 (Rha-C-1), 98.3 (Fuc-C-1), 81.5 (Rha-C-4), 80.4 (Fuc-C-4), 79.2 (Fuc-C-2), 78.5 (Rha-C-3), 77.9 (Gal-C-3), 75.8 (Rha-C-2), 73.7 (Gal-C-2), 73.1 ($Ph\underline{C}H_2$), 72.7 ($Ph\underline{C}H_2$), 71.7 (Gal-C-4), 71.4

(Ph<u>C</u>H$_2$), 71.3 (Fuc-C-3), 70.6 (Xyl-C-3), 70.3 (Xyl-C-2), 68.6 (Xyl-C-4), 68.0 (Gal-C-5), 67.9 (O<u>C</u>H$_2$(CH$_2$)$_6$CH$_3$), 67.4 (Gal-C-6), 65.7 (Fuc-C-5), 64.5 (Rha-C-5), 61.2 (Xyl-C-5), 31.8 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 29.7 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 29.4 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 29.3 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 28.2 (C(<u>C</u>H$_3$)$_2$), 27.6 (C(<u>C</u>H$_3$)$_3$), 27.2 (C(<u>C</u>H$_3$)$_3$), 26.5 (C(<u>C</u>H$_3$)$_2$), 26.4 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 23.4 (<u>C</u>(CH$_3$)$_3$), 22.7 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 20.80 (3 × CO<u>C</u>H$_3$), 20.6 (<u>C</u>(CH$_3$)$_3$), 18.0 (Rha-C-6), 16.3 (Fuc-C-6), 14.1 (OCH$_2$(CH$_2$)$_6$<u>C</u>H$_3$); HRMS (ESI) Calc. for [M + Na]$^+$ $C_{69}H_{100}NaO_{21}Si$: 1315.6419; Found 1315.6430.

Octyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→4)-[4-*O*-benzyl-2,3-*O*-isopropylidene-α-D-rhamnopyranosyl-(1→3)]-[2,3-di-*O*-benzyl-α-D-galactopyranosyl-(1→2)]-α-L-fucopyranoside (S19): To a stirred solution of **S18** (225.5 mg 174.3 μmol) in THF–pyridine (6.0 mL, 1:1) was added HF·pyridine (0.60 mL, pyridine ~30%, hydrogen fluoride ~70%) at 0 °C under an Ar atmosphere. The reaction mixture was stirred for 1 h at 0 °C, before being poured into saturated $NaHCO_3$ (aq.). The aqueous layer was extracted with EtOAc (25 mL × 3), dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (2:3 hexane–EtOAc) to afford **S19** (180.8 mg, 92%) as a syrup. R_f 0.21 (1:1 hexane–EtOAc); $[\alpha]_D$ –25.1 (*c* 1.3, $CHCl_3$); ^{1}H NMR (700 MHz; $CDCl_3$): δ 7.36–7.23 (m, 15H, Ar), 5.52 (s, 1H, Rha-H-1), 5.12 (t, 1H, *J* = 7.3 Hz, Xyl-H-3), 5.05 (dd, 1H, *J* = 7.5, 5.6 Hz, Xyl-H-2), 4.98 (td, 1H, *J* = 7.0, 4.6 Hz, Xyl-H-4), 4.93–4.91 (m, 2H, Gal-H-1, Fuc-H-1), 4.88 (d, 1H, *J* = 11.8 Hz, PhC<u>H</u>$_2$), 4.82 (d, 1H, *J* = 11.5 Hz, PhC<u>H</u>$_2$), 4.76 (d, 1H, *J* = 12.5 Hz, PhC<u>H</u>$_2$), 4.71 (d, 1H, *J* = 12.5 Hz, PhC<u>H</u>$_2$), 4.66 (d, 1H, *J* = 11.5 Hz, PhC<u>H</u>$_2$), 4.60 (d, 1H, *J* = 5.5 Hz, Xyl-H-1), 4.58 (d, 1H, *J* = 11.8 Hz, PhC<u>H</u>$_2$), 4.34 (dd, 1H, *J* = 12.2, 4.5 Hz, Xyl-H-5a), 4.28 (d, 1H, *J* = 5.7 Hz, Rha-H-2), 4.18 (dd, 1H, *J* = 10.3, 3.1 Hz, Fuc-H-3), 4.13 (t, 1H, *J* = 6.3 Hz, Rha-H-3), 4.03 (d, 1H, *J* = 2.0 Hz, Gal-H-4), 3.99 (t, 1H, *J* = 4.2 Hz, Gal-H-5), 3.95–3.84 (m, 5H, Rha-H-5, Fuc-H-5, Fuc-H-2, Gal-H-3, Gal-H-6a), 3.77 (dd, 1H, *J* = 9.9, 3.4 Hz, Gal-H-2), 3.71 (dd, 1H, *J* = 11.8, 4.4 Hz, Gal-H-6b), 3.65 (d, 1H, *J* = 3.0 Hz, Fuc-H-4), 3.57 (dt, 1H, *J* = 9.4, 7.0 Hz, OC<u>H</u>$_2$CH$_2$(CH$_2$)$_5$CH$_3$), 3.37 (dd, 1H, *J* = 12.2, 7.1 Hz, Xyl-H-5b), 3.31 (dt, 1H, *J* = 9.4, 6.8 Hz, OC<u>H</u>$_2$CH$_2$(CH$_2$)$_5$CH$_3$), 3.19 (dd, 1H, *J* = 9.8, 7.1 Hz, Rha-H-4), 2.07 (s, 3H, COC<u>H</u>$_3$), 2.03 (s, 3H, COC<u>H</u>$_3$), 2.00 (s, 3H, COC<u>H</u>$_3$), 1.57–1.51 (m, 2H, OCH$_2$C<u>H</u>$_2$(CH$_2$)$_5$CH$_3$), 1.38 (s, 3H, C(C<u>H</u>$_3$)$_2$), 1.30–1.25 (m, 13H, Rha-H-6, OCH$_2$CH$_2$(C<u>H</u>$_2$)$_5$CH$_3$), 1.20 (d, 3H, *J* = 6.6 Hz, Fuc-H-6), 1.01 (s, 3H, C(C<u>H</u>$_3$)$_2$), 0.87 (t, 3H, *J* = 7.0 Hz, OCH$_2$CH$_2$(CH$_2$)$_5$C<u>H</u>$_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 169.9 (C=O), 169.7 (C=O), 169.4 (C=O), 138.7 (Ar), 138.5 (Ar), 138.4 (Ar), 128.5 (2 × Ar), 128.1 (Ar), 128.0 (Ar), 127.9 (2 × Ar), 127.8 (Ar), 127.7 (2 × Ar), 108.9 (<u>C</u>(CH$_3$)$_2$), 100.9 (Xyl-C-1), 100.7 (Gal-C-1), 98.4 (Rha-C-1), 98.1 (Fuc-C-1), 81.4 (Rha-C-4), 80.6 (Fuc-C-4), 79.5 (Fuc-C-2), 78.5 (Rha-C-3), 75.8 (Rha-C-2), 75.3 (Gal-C-2), 73.2 (Ph<u>C</u>H$_2$), 72.9 (Ph<u>C</u>H$_2$), 72.7 (Ph<u>C</u>H$_2$), 71.3 (Fuc-C-3), 70.6 (Xyl-C-3), 70.3 (Xyl-C-2), 70.2 (Gal-C-4), 69.3 (Gal-C-5), 68.6 (Xyl-C-4), 67.8 (O<u>C</u>H$_2$(CH$_2$)$_6$CH$_3$), 65.7 (Fuc-C-5), 64.6 (Rha-C-5), 63.5 (Gal-C-6), 61.2 (Xyl-C-5), 31.8 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 29.8 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 29.5 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 29.3 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 28.1 (C(<u>C</u>H$_3$)$_2$), 26.43 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 26.4 (C(<u>C</u>H$_3$)$_2$), 22.7 (OCH$_2$(<u>C</u>H$_2$)$_6$CH$_3$), 20.8 (3 × CO<u>C</u>H$_3$), 18.0 (Rha-C-6), 16.3 (Fuc-C-6), 14.1 (OCH$_2$(CH$_2$)$_6$<u>C</u>H$_3$); HRMS (ESI) Calc. for [M + NH$_4$]$^+$ $C_{61}H_{88}NO_{21}$: 1170.5843; Found 1170.5862.

Octyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→4)-[4-*O*-benzyl-α-D-rhamnopyranosyl-(1→3)]-[2,3-di-*O*-benzyl-α-D-galactopyranosyl-(1→2)]-α-L-fucopyranoside (S20): To a stirred solution of **S19** (180.8 mg, 156.8 μmol) in CH_3CN–CH_3OH (18 mL, 10:1) was added *p*-toluenesulfonic acid monohydrate (89.5 mg, 470.3 μmol) at room temperature. The reaction mixture was stirred for 3.5 h at room temperature. Triethylamine was added, the mixture was filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (1:2 hexane–EtOAc) to afford **S20** (150.2 mg, 86%) as a white amorphous solid. R_f 0.31 (1:2 hexane–EtOAc); $[\alpha]_D$ –10.9 (*c* 0.6, $CHCl_3$); ^{1}H NMR (700 MHz; $CDCl_3$): δ 7.39–7.23 (m, 15H), 5.11–5.08 (m, 2H, Rha-H-1), 5.04 (dd, 1H, *J* = 8.3, 6.5 Hz), 4.99–4.95 (m, 2H, Gal-H-1), 4.92 (d, 1H, *J* = 3.5 Hz, Fuc-H-1), 4.81–4.75 (m, 3H), 4.70–4.67 (m, 2H), 4.58 (d, 1H, *J* = 11.4 Hz), 4.52 (d, 1H, *J* = 6.4 Hz, Xyl-H-1), 4.21 (dd, 1H, *J* = 12.0, 4.9 Hz), 4.11 (d, 1H, *J* = 1.6 Hz), 4.06–4.03 (m, 2H), 3.92–3.80 (m, 7H), 3.75–3.72 (m, 2H), 3.67 (d, 1H, *J* = 3.0 Hz), 3.56 (dt, 1H, *J* = 9.4, 6.9 Hz), 3.31 (dt, 1H, *J* = 9.4, 6.8 Hz), 3.23 (dd, 1H, *J* = 12.0, 8.4 Hz), 3.18 (t, 1H, *J* = 8.7 Hz), 2.71 (s, 1H), 2.41 (dd, 1H, *J* = 6.4, 0.4 Hz), 2.05 (s, 3H), 2.02 (s, 3H), 1.96 (s, 3H), 1.89 (d, 1H, *J* = 3.7 Hz), 1.87 (d, 1H, *J* = 6.1 Hz), 1.55-1.51 (m, 2H), 1.30–1.24 (m, 13H), 1.18 (d, 3H, *J* = 6.6 Hz), 0.87 (t, 3H, *J* = 7.0 Hz); ^{13}C NMR (125 MHz; $CDCl_3$): δ 170.1, 169.6, 169.5, 138.8, 138.2, 138.0, 128.7, 128.6, 128.3 (3 × C), 128.0, 127.8, 127.7, 127.5, 101.3 (Xyl-C-1), 101.3 (Rha-C-1), 100.9 (Gal-C-1), 98.1 (Fuc-C-1), 81.9, 80.5, 79.7, 77.6, 76.2, 74.5, 74.1, 72.8, 71.9, 71.5, 71.2, 71.0, 70.6, 69.5, 69.2, 68.7, 67.7, 67.4, 65.5, 63.4, 61.7, 31.9, 29.8, 29.5, 29.3, 26.5, 22.7, 20.8 (2 × C), 20.7, 18.3, 16.2, 14.1; HRMS (ESI) Calc. for $[M + Na]^+$ $C_{58}H_{80}NaO_{21}$: 1135.5084; Found 1135.5090.

Octyl β-D-xylopyranosyl-(1→4)-[4-*O*-benzyl-α-D-rhamnopyranosyl-(1→3)]-[2,3-di-*O*-benzyl-α-D-galactopyranosyl-(1→2)]-α-L-fucopyranoside (S21): To a stirred solution of **S20** (103.5 mg, 93.0 μmol) in CH_3OH (4.0 mL) was added a solution of $NaOCH_3$ in CH_3OH (0.4 mL, 0.5 M). The reaction mixture was stirred for 4 h at room temperature, then neutralized by addition of Amberlite® IR-120 (H^+) cation exchange resin, filtered and the filtrate was concentrated to afford **S21** (91.8 mg, quant.) as a white amorphous solid. R_f 0.32 (9:1 CH_2Cl_2–CH_3OH); $[\alpha]_D$ +2.7 (*c* 1.5, CH_2Cl_2–CH_3OH); ^{1}H NMR (700 MHz; CD_3OD): δ 7.46–7.44 (m, 4H), 7.38–7.23 (m, 11H), 5.30 (d, 1H, *J* = 1.1 Hz, Rha-H-1), 5.02 (d, 1H, *J* = 3.6 Hz, Gal-H-1), 4.95 (d, 1H, *J* = 11.4 Hz), 4.87 (d, 1H, *J* = 12.8 Hz), 4.83–4.81 (m, 2H, Fuc-H-1), 4.77 (d, 1H, *J* = 11.6 Hz), 4.71 (d, 1H, *J* = 11.6 Hz), 4.63 (d, 1H, *J* = 11.4 Hz), 4.29 (d, 1H, *J* = 7.4 Hz, Xyl-H-1), 4.13–4.08 (m, 3H), 4.06–4.05 (m, 2H), 4.00–3.94 (m, 3H), 3.87–3.82 (m, 4H), 3.65–3.61 (m, 3H), 3.53 (ddd, 1H, *J* = 10.0, 8.8, 5.4 Hz), 3.47 (dt, 1H, *J* = 9.4, 6.5 Hz), 3.39–3.34 (m, 2H), 3.32–3.30 (m, 1H), 3.09 (t, 1H, *J* = 11.0 Hz), 1.60 (quintet, 2H, *J* = 6.9 Hz), 1.39–1.27 (m, 13H), 1.24 (d, 3H, *J* = 6.3 Hz), 0.90 (t, 3H, *J* = 6.9 Hz); ^{13}C NMR (125 MHz; CD_3OD): δ 140.5, 140.3, 140.2, 129.6, 129.5, 129.3 (2 × C), 128.9 (2 × C), 128.7, 128.5, 128.4, 106.3 (Xyl-C-1), 102.6 (Rha-C-1), 102.5 (Fuc-C-1), 99.8 (Gal-C-1), 83.0, 81.0, 80.3, 79.2, 78.0, 76.1, 75.7, 75.5, 74.0, 73.0, 72.8, 72.6, 72.5, 72.3, 71.3, 69.3, 68.7 (2 × C), 67.8, 66.8, 62.5, 33.1, 30.8, 30.6 (2 × C), 27.6, 23.8, 18.4, 16.5, 14.5; HRMS (ESI) Calc. for $[M + Na]^+$ $C_{52}H_{74}NaO_{18}$: 1009.4767; Found 1009.4772.

Octyl β-D-xylopyranosyl-(1→4)-[α-D-rhamnopyranosyl-(1→3)]-[α-D-galactopyranosyl-(1→2)]-α-L-fucopyranoside (1): To a stirred solution of **S21** (91.8 mg, 93.0 μmol) in dry THF (5.0 mL) was added 20%

palladium hydroxide on carbon (10.0 mg). After stirring for 24 h under an H_2 atmosphere (1 atm), the reaction mixture was filtered through Celite and concentrated. The residue was dissolved in water and then lyophilized to afford **1** (66.7 mg, quant.) as a white solid. R_f 0.28 (1:2:1 CH_2Cl_2–CH_3OH–H_2O); $[\alpha]_D$ +2.9 (*c* 0.4, CH_2Cl_2–CH_3OH); ^{1}H NMR (700 MHz; CD_3OD): δ 5.14 (d, 1H, *J* = 1.2 Hz, Rha-H-1), 5.04 (d, 1H, *J* = 3.7 Hz, Fuc-H-1), 4.97 (d, 1H, *J* = 3.8 Hz, Gal-H-1), 4.28 (d, 1H, *J* = 7.1 Hz, Xyl-H-1), 4.08 (dd, 1H, *J* = 11.6, 5.4 Hz), 4.04 (dd, 1H, *J* = 10.4, 3.1 Hz), 4.00 (q, 1H, *J* = 6.7 Hz), 3.98 (dd, 1H, *J* = 3.3, 1.6 Hz), 3.96–3.92 (m, 3H), 3.88 (d, 1H, *J* = 2.6 Hz), 3.84 (d, 1H, *J* = 3.1 Hz), 3.81–3.78 (m, 2H), 3.75 (dd, 1H, *J* = 10.2, 3.2 Hz), 3.73–3.65 (m, 3H), 3.53–3.48 (m, 2H), 3.40 (t, 1H, *J* = 9.4 Hz), 3.33–3.28 (m, 2H), 3.10 (t, 1H, *J* = 11.0 Hz), 1.62 (dt, 2H, *J* = 14.0, 7.0 Hz), 1.42–1.40 (m, 2H), 1.36–1.27 (m, 11H), 1.24 (d, 3H, *J* = 6.3 Hz), 0.91 (t, 3H, *J* = 7.0 Hz); ^{13}C NMR (125 MHz; CD_3OD): δ 106.3 (Xyl-C-1), 103.4 (Rha-C-1), 103.2 (Gal-C-1), 99.9 (Fuc-C-1), 81.2, 78.4, 77.9, 75.5, 74.3, 74.1, 72.7, 71.9, 71.3, 71.2, 71.1, 70.1, 68.9, 68.0, 66.7, 62.8, 33.1, 30.8, 30.6, 30.5, 27.6, 23.7, 18.0, 16.6, 14.4; HRMS (ESI) Calc. for $[M + Na]^+$ $C_{31}H_{56}NaO_{18}$: 739.3359; Found 739.3364.

Synthesis of Xyloside 6 (Fig. S21)

Octyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranoside (S22): To a stirred solution of octanol (48.0 μL, 305.9 μmol) and *p*-tolyl 2,3,4-tri-*O*-acetyl-1-thio-β-D-xylopyranoside **S17** (6) (130.1 mg, 367.1 μmol) in dry CH_2Cl_2 (4.0 mL) added molecular sieves (400 mg, 4 Å, powder). After stirring for 30 min at room temperature, the reaction mixture was cooled to 0 °C, and then *N*-iodosuccinimide (103.3 mg, 459.2 μmol) and silver trifluoromethanesulfonate (15.7 mg, 61.2 μmol) were added successively. The resulting solution was stirred for 1 h at room temperature under an Ar atmosphere. Triethylamine was added, the mixture was filtered and the filtrate was washed with saturated $Na_2S_2O_3$ (aq.) and saturated $NaHCO_3$ (aq.), the aqueous layer was extracted with CH_2Cl_2 (25 mL × 3), dried over Na_2SO_4, filtered and the filtrate was concentrated. The crude residue was purified by flash chromatography (3:1 hexane–EtOAc) to afford **S22** (59.4 mg, 50%) as a white amorphous solid. R_f 0.51 (2:1 hexane–EtOAc); $[\alpha]_D$ –48.5 (*c* 0.7, $CHCl_3$); ^{1}H NMR (600 MHz; $CDCl_3$): δ 5.17 (t, 1H, *J* = 8.6 Hz, H-3), 4.96 (td, 1H, *J* = 8.7, 5.1 Hz, H-4), 4.93 (dd, 1H, *J* = 8.7, 6.8 Hz, H-2), 4.48 (d, 1H, *J* = 6.8 Hz, H-1), 4.13 (dd, 1H, *J* = 11.8, 5.1 Hz, H-5a), 3.82 (dt, 1H, *J* = 9.6, 6.5 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.47 (dt, 1H, *J* = 9.6, 6.7 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.37 (dd, 1H, *J* = 11.8, 8.8 Hz, H-5b), 2.07 (s, 3H, $COC\underline{H}_3$), 2.06 (s, 3H, $COC\underline{H}_3$), 2.05 (s, 3H, $COC\underline{H}_3$), 1.60–1.55 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.34–1.28 (m, 10H, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 0.89 (t, 3H, *J* = 7.0 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; $CDCl_3$): δ 170.1 (C=O), 169.8 (C=O), 169.4 (C=O), 100.7 (C-1), 71.5 (C-3), 70.9 (C-2), 69.7 ($O\underline{C}H_2(CH_2)_6CH_3$), 69.0 (C-4), 62.0 (C-5), 31.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 29.3 (2 × $OCH_2(\underline{C}H_2)_6CH_3$), 25.9 ($OCH_2(\underline{C}H_2)_6CH_3$), 22.6 ($OCH_2(\underline{C}H_2)_6CH_3$), 20.8 ($CO\underline{C}H_3$), 20.7 (2 × $CO\underline{C}H$), 14.1 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{19}H_{32}NaO_8$: 411.1989; Found 411.1984.

Octyl β-D-xylopyranoside (6): To a stirred solution of **S22** (25.9 mg, 66.7 μmol) in CH_3OH (4.0 mL) was added a solution of $NaOCH_3$ in CH_3OH (0.4 mL, 0.5 M). The reaction mixture was stirred for 4 h at room

temperature, then neutralized by addition of Amberlite® IR-120 (H^+) cation exchange resin, filtered and the filtrate was concentrated. The residue was dissolved in water and then lyophilized to afford **6** (17.5 mg, quant.) as a white solid. R_f 0.49 (9:1 CH_2Cl_2–CH_3OH); $[\alpha]_D$ –38.6 (*c* 0.2, CH_2Cl_2–CH_3OH); ^1H NMR (600 MHz; CD_3OD): δ 4.17 (d, 1H, *J* = 7.6 Hz, H-1), 3.83 (dd, 1H, *J* = 11.5, 5.4 Hz, H-5a), 3.79 (dt, 1H, *J* = 9.5, 6.8 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.51 (dt, 1H, *J* = 9.6, 6.7 Hz, $OC\underline{H}_2CH_2(CH_2)_5CH_3$), 3.46 (ddd, 1H, *J* = 10.2, 8.8, 5.4 Hz, H-4), 3.28 (t, 1H, *J* = 8.9 Hz, H-3), 3.17 (dd, 1H, *J* = 11.2, 10.1 Hz, H-5b), 3.14 (dd, 1H, *J* = 9.0, 7.5 Hz, H-2), 1.62–1.57 (m, 2H, $OCH_2C\underline{H}_2(CH_2)_5CH_3$), 1.38–1.27 (m, 10H, $OCH_2CH_2(C\underline{H}_2)_5CH_3$), 0.89 (t, 3H, *J* = 7.0 Hz, $OCH_2CH_2(CH_2)_5C\underline{H}_3$); ^{13}C NMR (125 MHz; CD_3OD): δ 105.1 (C-1), 77.9 (C-3), 74.9 (C-2), 71.2 (C-4), 70.9 ($O\underline{C}H_2(CH_2)_6CH_3$), 66.9 (C-5), 33.0 ($OCH_2(\underline{C}H_2)_6CH_3$), 30.8 ($OCH_2(\underline{C}H_2)_6CH_3$), 30.5 ($OCH_2(\underline{C}H_2)_6CH_3$), 30.4 ($OCH_2(\underline{C}H_2)_6CH_3$), 27.1 ($OCH_2(\underline{C}H_2)_6CH_3$), 23.7 ($OCH_2(\underline{C}H_2)_6CH_3$), 14.4 ($OCH_2(CH_2)_6\underline{C}H_3$); HRMS (ESI) Calc. for $[M + Na]^+$ $C_{13}H_{26}NaO_5$: 285.1672; Found 285.1670.

A

A064R

```
  1 MTTPCITILS GHFPKETIYA RKTKELVEEY CSIHGYNFYY EESEPLETEE HALHFRRSWI
 61 IQQAAEKFPS TEWFLWLDSD VYVNPKNKNK PITSFIDLSD PNILYHTFHE APWGSYPINT
121 GVKFVHKDAL EIEKIVWSLR NEAPWNTFPY EQKTVYEYVF PRIPGRYIVH DPYTLNCIVK
181 AYPEHVKDAL FVHMCGTSRA ERDEHMEMVA TGEIQLPKKT IGYIHVCQKD GWERSFGMIL
241 NEVKRSGLYD KTDEIRCVIV NDIDRPYYNE MFEDTKLNLL YAGESQKYER PALLHMKQQS
301 ETEYCNYWYV HTKGLRWFNT PKENNVVDWI NLLIYWNVTK WEEAISNLKT HDVYGCNFTD
361 TPTPHFSGNF WWATSEYVRK LPAFISSGYN DPEFWLFLGN PVYKNIFSSG LEGMGHYDNR
421 FERNKYAPEA DFASTEETPL VKSGTLASLV DNTKTDKNTT HSYLNLYEEI LKEKKYSAKN
481 IIEIGIGDFG EKNGGSIKMW RDYFPNATIY GVDILPKDRV MDELLEDERV AIFTETDGYT
541 TDFINREINN IKFDFAIDDG PHTLESMVQF VRLYHTLLTD DGVLIVEDIP HLEWLEILAN
601 EIPQDMLHCV QGYDLRSVKN RFDDIVFVFD KRCVNTQE
```

A061L

```
  1 MRHKPRVLSR SMAWRNAVNG VLQQLTGYQL RRVTVPAART APAESLAAST PAPAPPKPEA
 61 EEGQAKKPET KKSPAKKPSP AFPEDYDDEA RDIIRAVKPY SMTSPERLNA FILATRYIAR
121 HNIPGDIVEC GVWRGGSMQA CARTLLSVGE TERELYLFDT YEGMTEPTAE DLRRDGRPAQ
181 ELLDAQGKDR PIWAVASLDD VKAGFENVPY PKERVHYVQG RVEDTVPGQA PEQISILRLD
241 TDWYASTKHE LEHLYSRLVS GGVLLIDDYG YWQGSRQAVD EFLDKTGERL LLLRMDEGRI
301 AVKP
```

B

GST-fusion proteins

GST-*SDLIEGRGIP*-**A061L**

GST-*LEVLFQ*|*GPLGS*-**A064R-D1**
GST-*LEVLFQ*|*GPLGS*-**A064R-D2L**
GST-*LEVLFQ*|*GPLGS*-**A064R-D1D2L**

Prescission protease cleavage site

Full length A064R – C-terminal 6XHis tag

A064R-*LEHHHHHH*

C

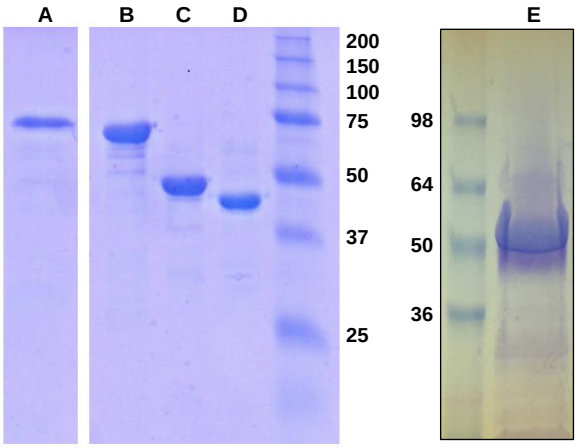


Figure S1. A) Amino acid sequence of chlorovirus PBCV-1 encoded proteins A064R and A061L. The A064R protein has 3 domains: domain 1 (red, from 1 to 212 aa) and domain 2 (black) each harbor a glycosyl transferase activity; domain 3 (blue, from 438 to 638 aa) is an O-methyltransferase. Several sequences have been explored involving the exact range of the second domain: the **D2** sequence encompasses aa 191–405; **D2L** sequence range is 191–438 aa; **D2L$_2$** range is 213–438 aa. Highlighted in yellow are aa straddling the first two domains and highlighted in blue are aa added to the C-terminus of **D2L** and **D2L$_2$**. The A061L protein is the O-methyltransferase whose function relies on A064R-D3 methylation of the terminal Rha first. B) Organization of the expressed proteins used to test the enzymatic activities. **D1**, **D2L,** **D1D2L**, and A061L are N-terminal GST-fusion proteins; the connector between GST and the proteins is indicated in italic; the protease cleavage site is indicated for the A064R domains by the blue dotted line. **D2L$_2$** (same as D2L) was also produced as GST-fusion protein, lacking the first N-terminal residues of **D2L**, as indicated in panel A. The full length A064R was produced with a C-terminal 6xHis tag. C) SDS-PAGE analysis of the expressed proteins as in panel B. Lane A: full length A064R-6xHis; lane B: GST-D1D2L; lane C: GST-D1; lane D: GST-D2L; lane E: GST-A061L.

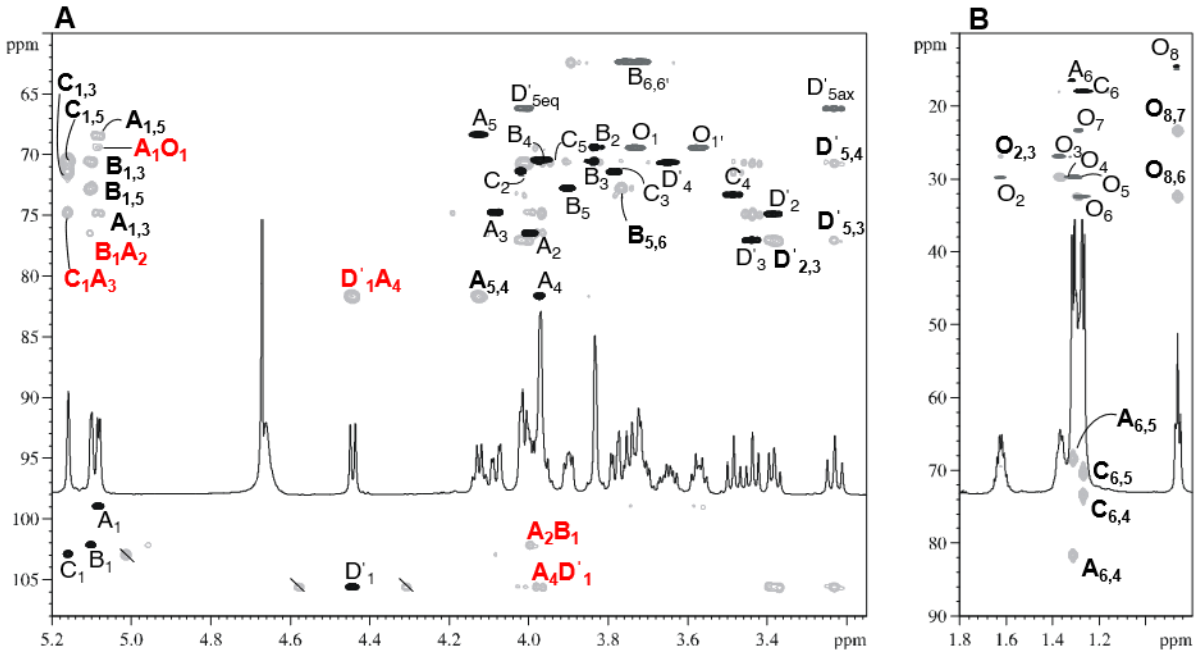


Figure S2. (600 MHz, 310 K, D_2O) Superimposition of the HMBC (light gray) and HSQC (black/dark gray) NMR spectra of tetrasaccharide **1**. Densities appearing in dark gray in the HSQC spectrum have the opposite sign with respect to the other and represent carbons bearing two hydrogen atoms. A) and B) panels are two different zooms of the same HSQC/HMBC spectra overlap. HMBC densities are in labeled in bold and interresidual correlations are drawn in red. Labelling follows the letter system of Table S1. The peak at ca. 4.7 ppm is the residual solvent signal.

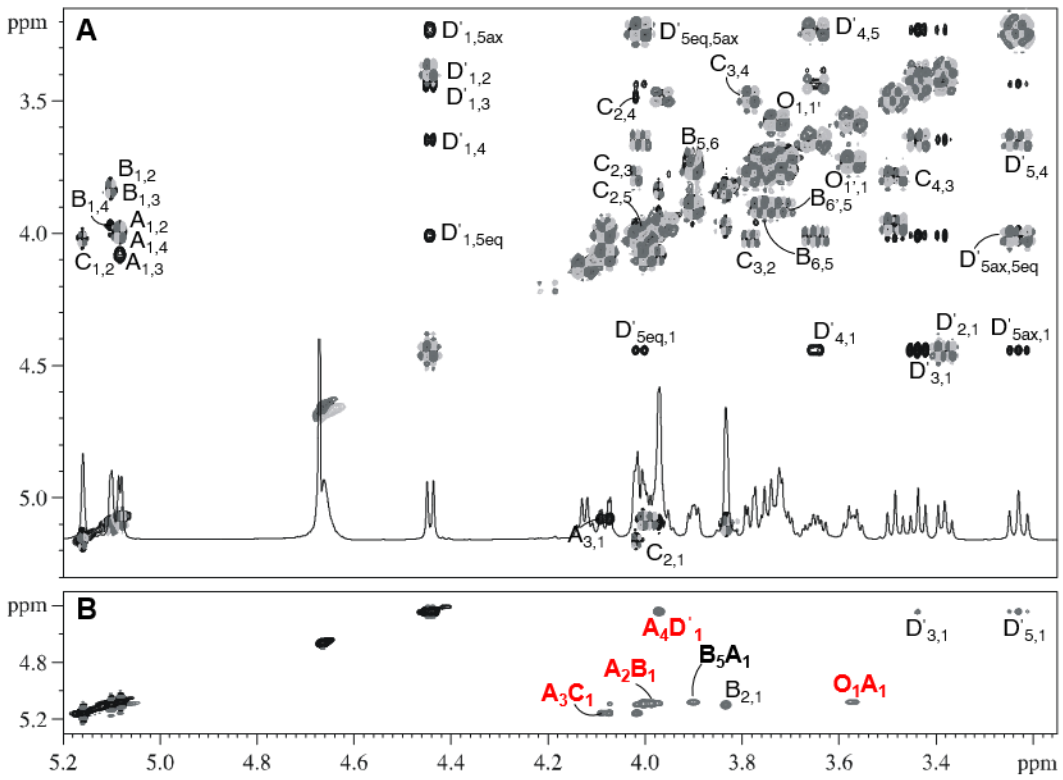


Figure S3. (600 MHz, 310 K, D_2O) NMR spectra of oligosaccharide **1**. A) Zoom of the superimposition of the TOCSY (black) and COSY (light and dark gray) spectra. B) Zoom of the T-ROESY spectrum. In red are reported key NOE contacts. Labelling follows the letter system of Table S1. The peak at ca. 4.7 ppm is the residual solvent signal.

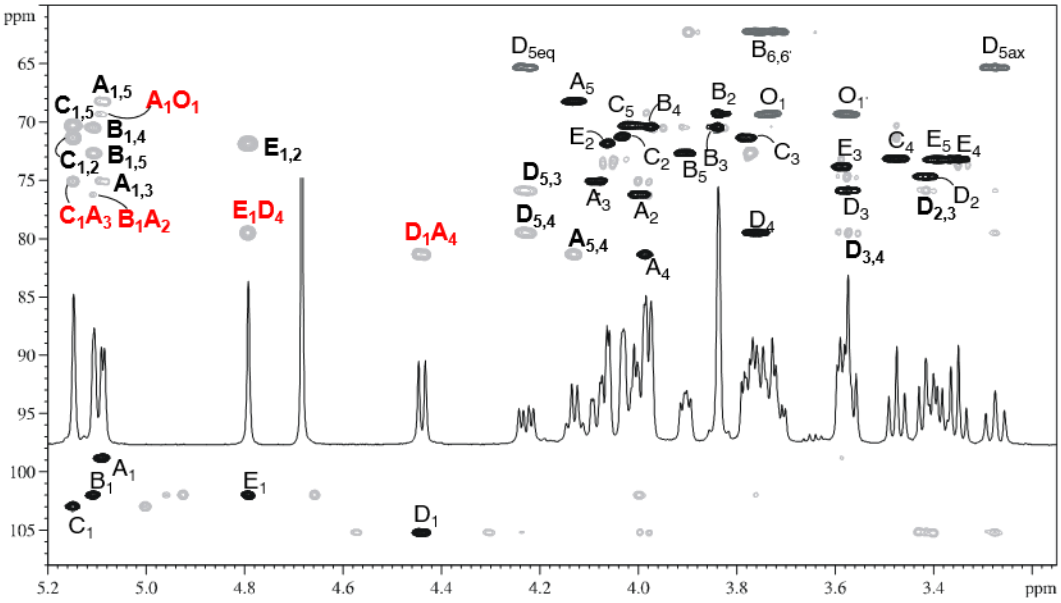


Figure S4. (600 MHz, 310 K, D_2O) NMR spectra of oligosaccharide **2**. Superimposition of the HMBC (light gray) and HSQC (black/dark gray) spectra. HMBC densities are in labeled in bold and interresidual correlations are drawn in red. Labelling follows the letter system of Table S2. The peak at ca. 4.7 ppm is the residual solvent signal.

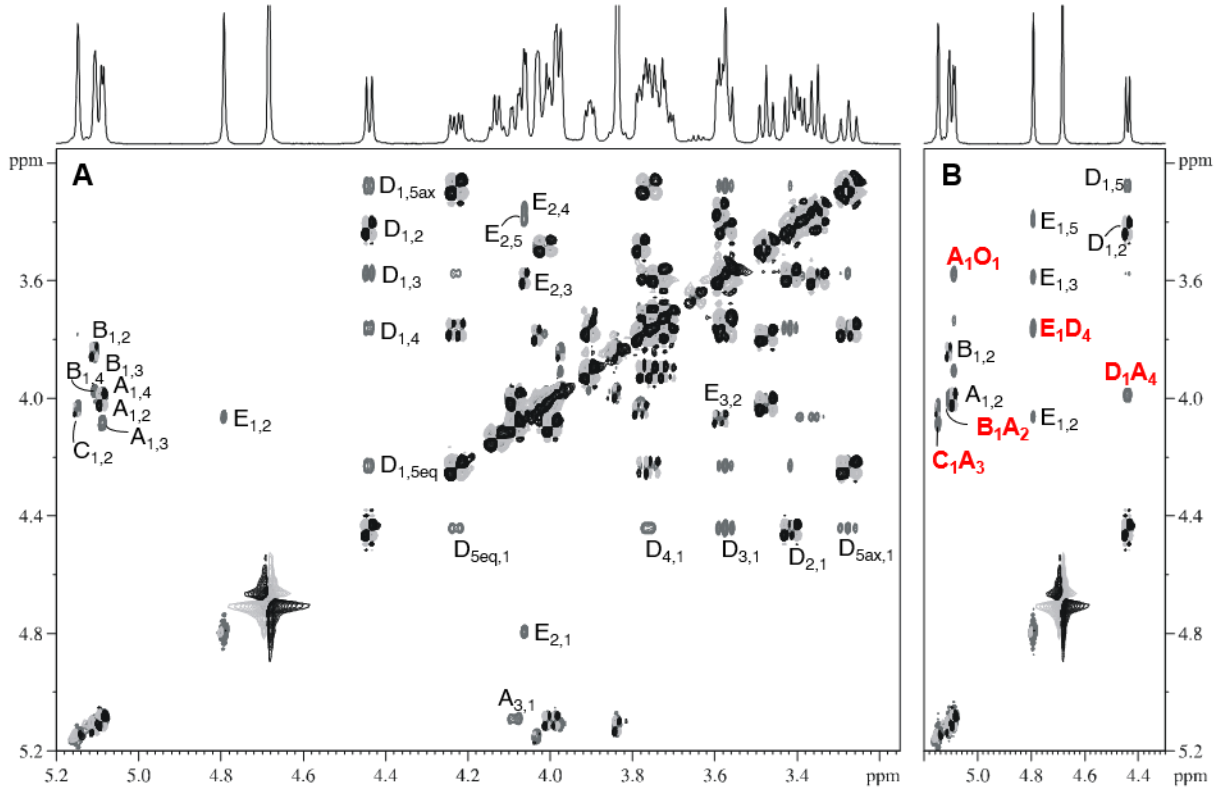


Figure S5. (600 MHz, 310 K, D_2O) NMR spectra of oligosaccharide **2**. A) Zoom of the overlapping of the TOCSY (black) and COSY (light and dark gray) spectra. B) Zoom of the T-ROESY (dark gray) and COSY (light gray/black) spectra. In red are reported the key NOE contacts. Proton NMR is reported at the top of each panel. Labelling follows the letter system of Table S2. The peak at ca. 4.7 ppm is the residual solvent signal.

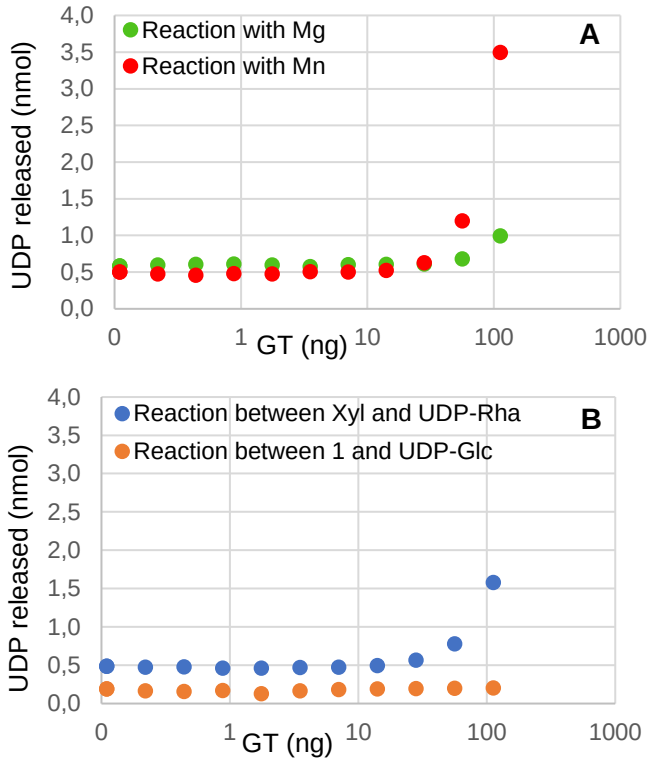


Figure S6. Graphs obtained from the UDP-Glo bioluminescent assay reporting the amount of released UDP molecules vs. the concentration of A064R-D1. A) Graph obtained from the reaction of A064R-D1 in the presence of manganese (red) and magnesium (green) cations. B) Reaction curves obtained by reacting (blue curve): UDP-β-L-Rha with D-Xyl (which replaced the acceptor substrate **1)**, and (orange curve): UDP-α-D-Glc (rather than the donor UDP-β-L-Rha) with the acceptor substrate **1**.

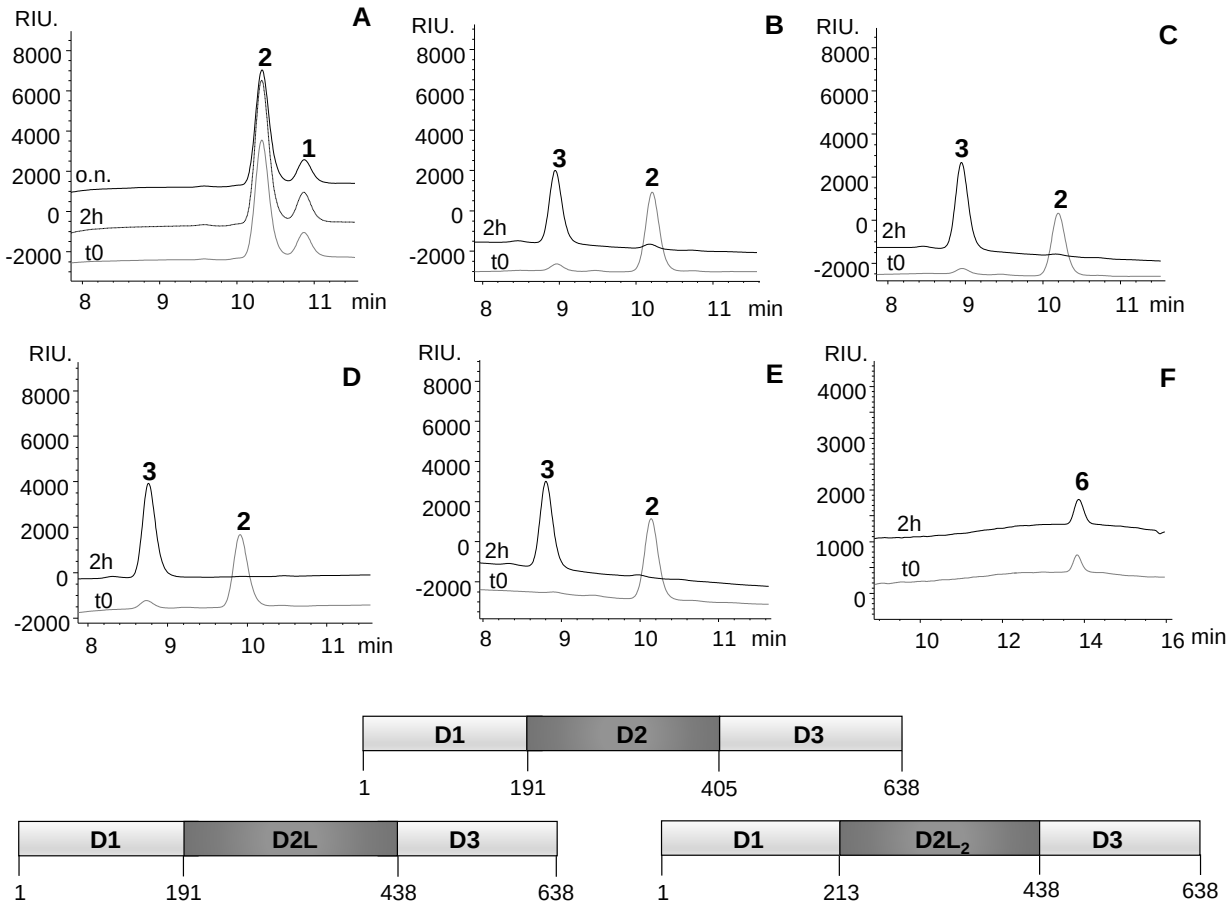


Figure S7. HPLC chromatographic profiles obtained from the reactions of different A064R-D2 enzyme constructs (schematic representations at bottom). A) Reaction between the **D2** enzyme with **2** as acceptor and UDP-Rha as donor, in the presence of Mn^{2+} and Mg^{2+} cations. B and C) Reactions of the **D2L** and **D2L2**, respectively, between **2** and UDP-Rha substrates without cations. D) Same reaction as in panel B repeated 5 days later. E and F) **D2L** tested in the presence of EDTA and with **6** as the acceptor, respectively. Structures are reported in Figure 2.

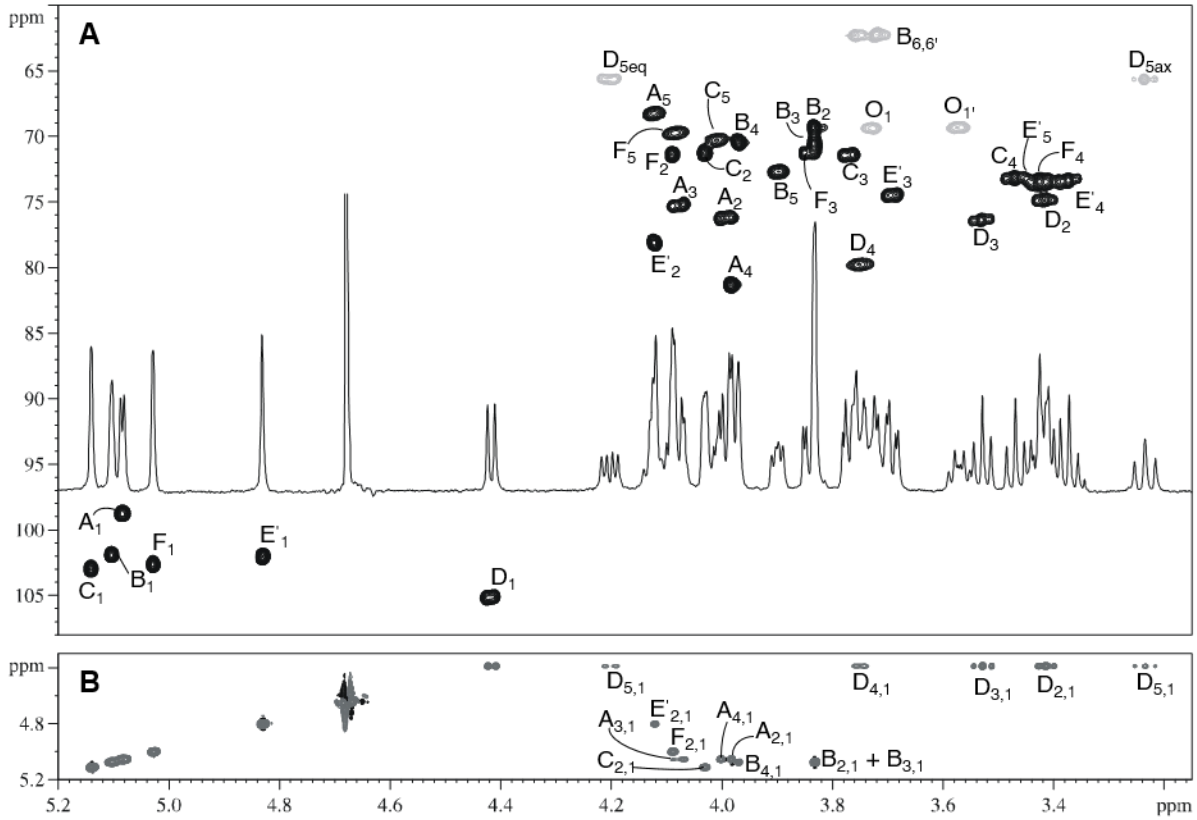


Figure S8. (600 MHz, 310 K, D_2O) NMR spectra of oligosaccharide **3**. A) Expansions of HSQC (black/light gray) spectrum and B) of the TOCSY spectrum. Labelling follows the letter system of Table S3. The peak at ca. 4.7 ppm is the residual solvent signal.

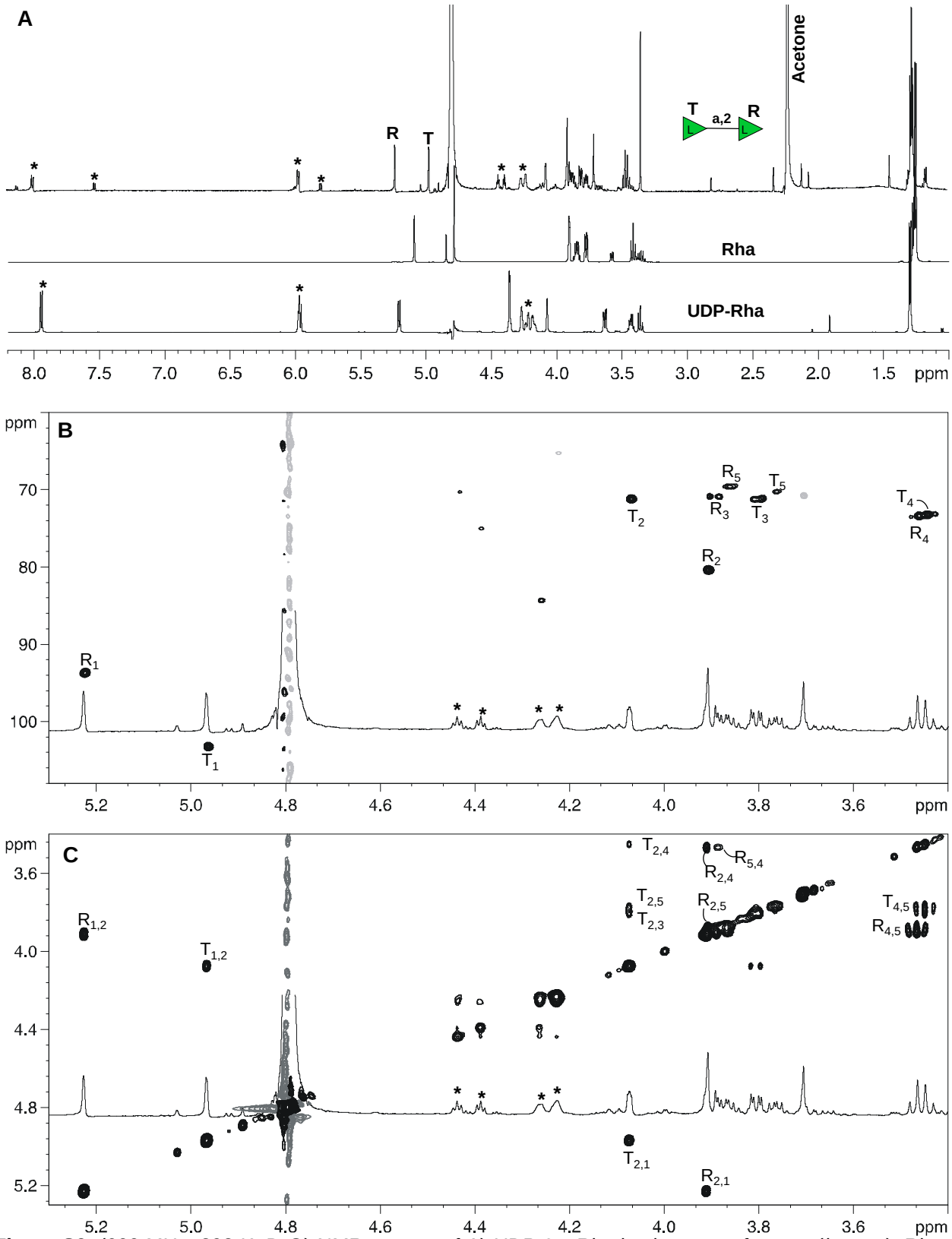


Figure S9. (600 MHz, 298 K, D_2O) NMR spectra of A) UDP-β-L-Rha is given as reference (*bottom*), Rha monosaccharide (*middle*), and of α-L-Rha-(1→2)-L-Rha (*top*); signals starred arise from the nucleoside portion or from the ribose moiety of UDP-Rha. B and C) (HSQC) and (TOCSY) spectra of the disaccharide. Labelling follows the letter system of Table S4. Green triangles indicate the L-Rha monosaccharide. The peak at ca. 4.7 ppm is the residual solvent signal.

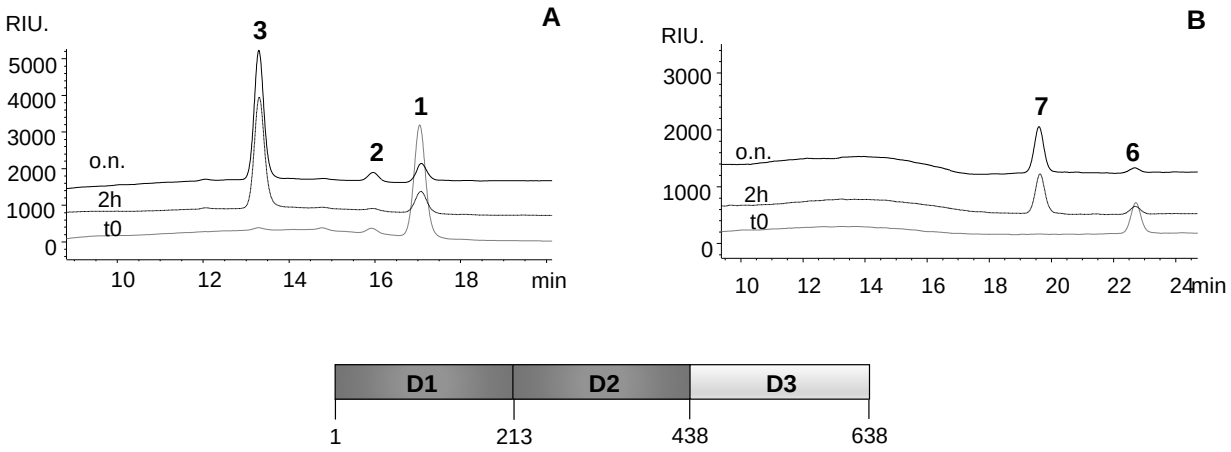


Figure S10. HPLC chromatographic profiles obtained from the reaction of the A064R-D1D2 enzyme (schematic representation at bottom) with two different acceptor substrates: **1** in panel A, and **6** in panel B. The compound to which each peak refers is shown in figure 2.

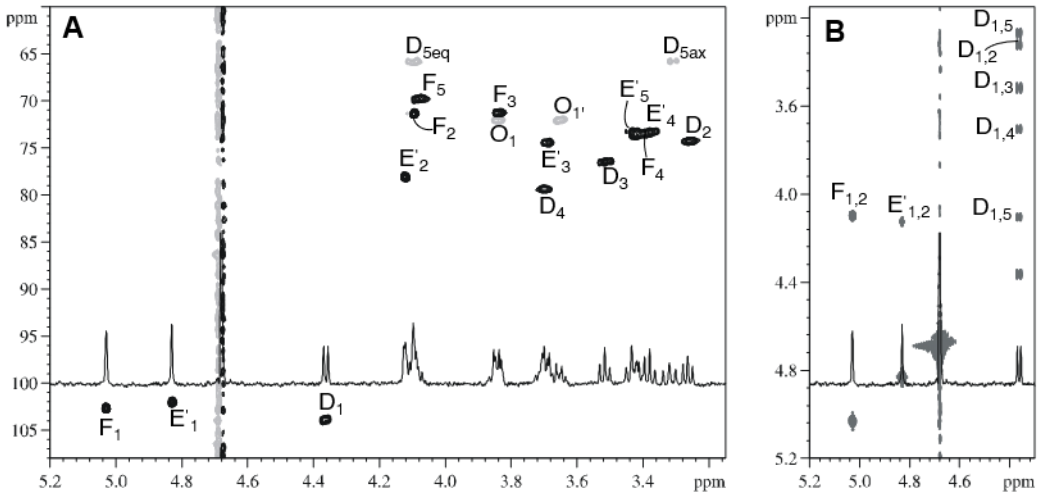


Figure S11. (600 MHz, 310 K, D_2O) NMR spectra of trisaccharide **7**: A) HSQC and B) TOCSY spectra. Labelling follows the letter system of Table S5. The peak at ca. 4.7 ppm is the residual solvent signal.

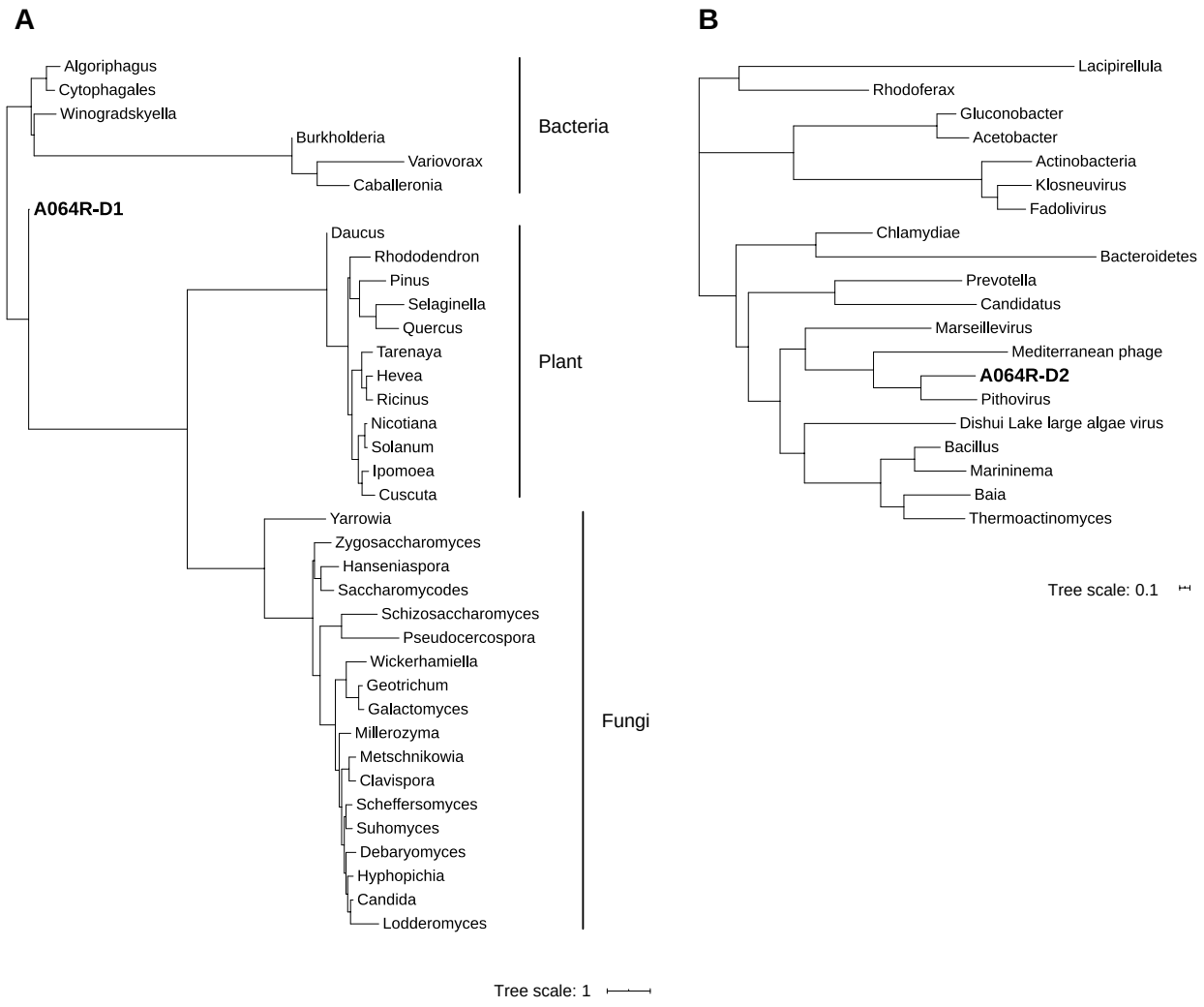


Figure S12. Maximum-likelihood phylogenetic trees of A) domain 1 and B) domain 2 of PBCV-1 encoded protein A064R involved in glycan synthesis. Unlike its partner domains, A064R-D1 shares an ancestral node of origin with eukaryotic homologs. Organism class denoted by vertical lines. Phylogenetic tree members were derived from a protein BLAST query search for A064R-D1 and A064R-D2 using the NCBI web database.

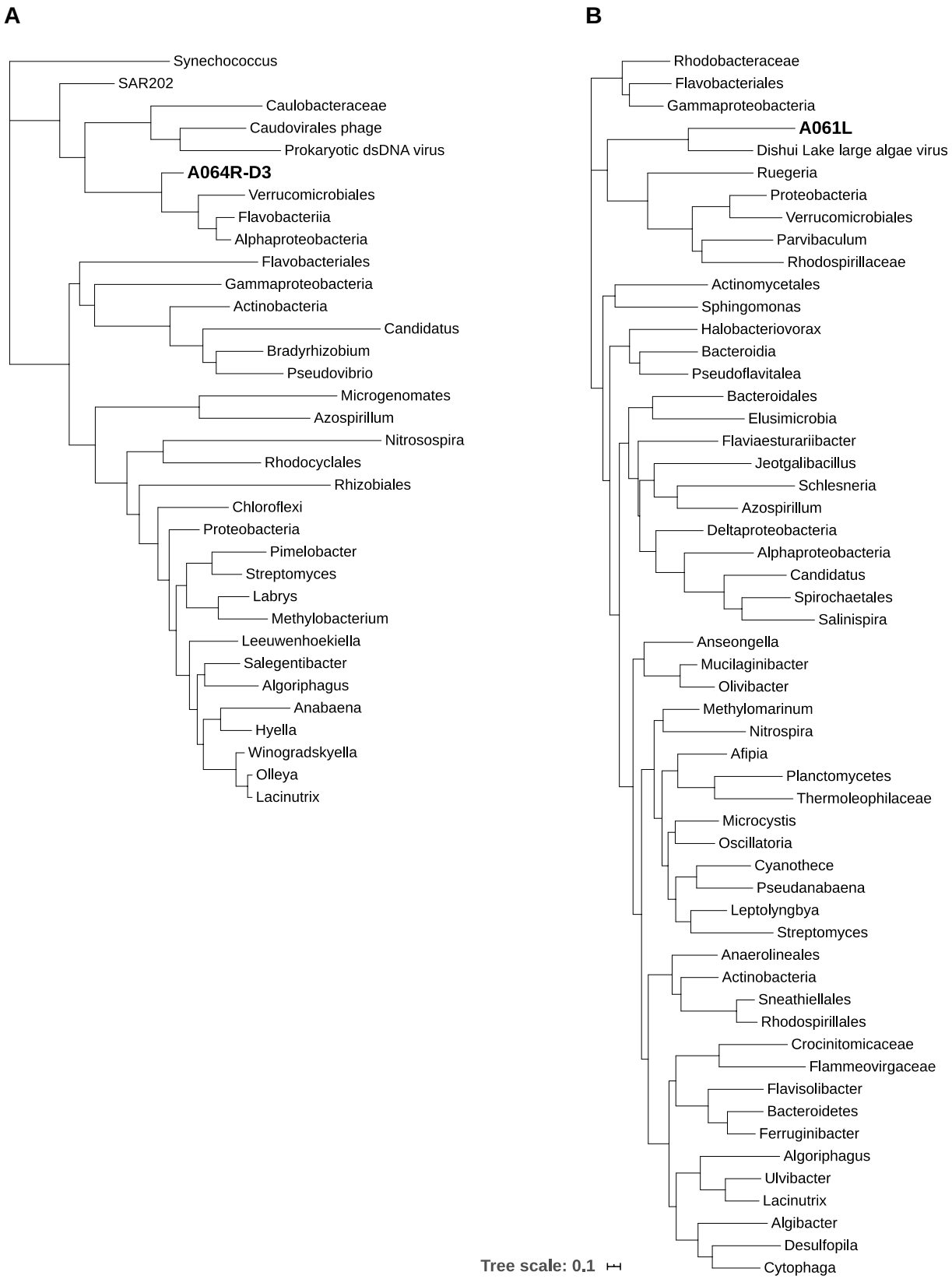


Figure S13. Maximum-likelihood phylogenetic trees of A) A064R-D3 and B) A061L involved in glycan methylation. Phylogenetic tree members were derived from a protein BLAST query search for virus-encoded proteins using the NCBI web database

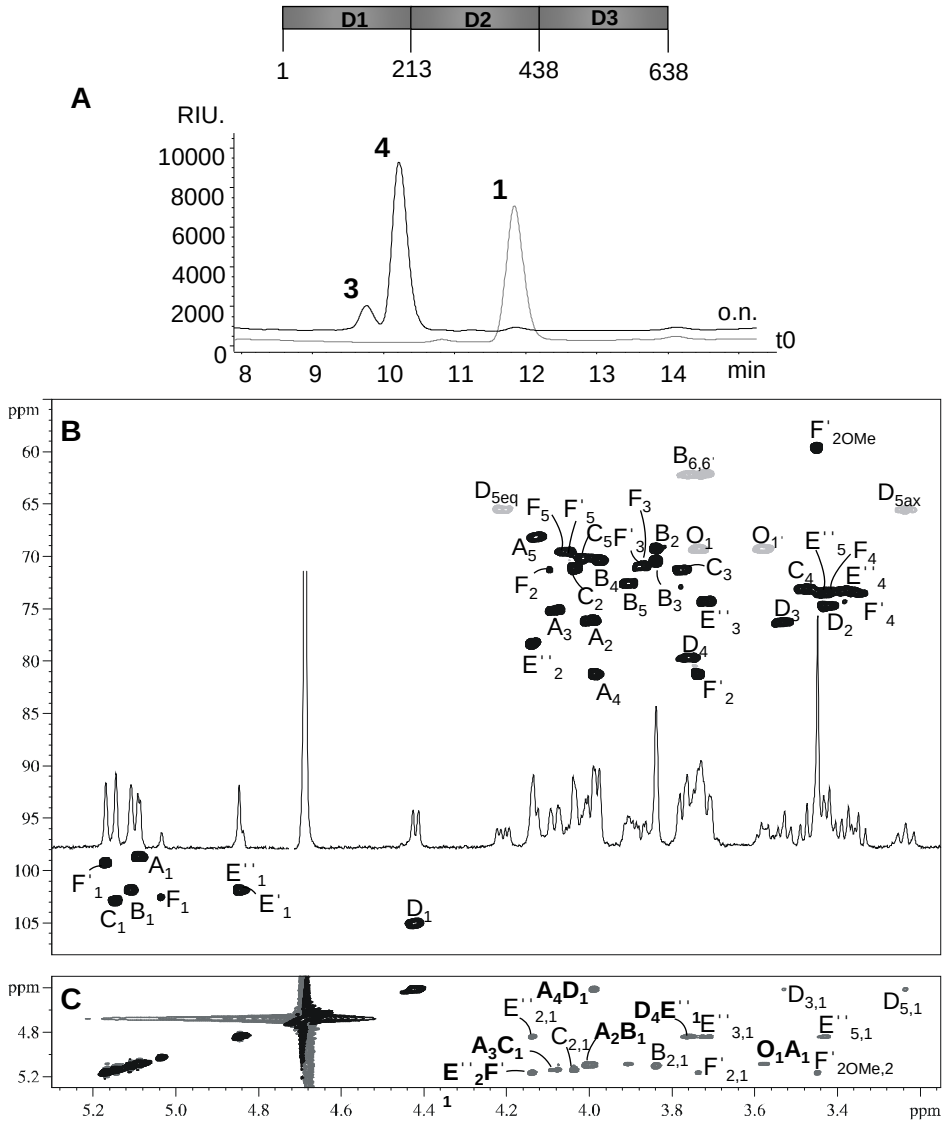


Figure S14. Data results of oligosaccharide **4** (Structure in Fig. 2). A) HPLC chromatographic profiles obtained from the reaction of the entire A064R protein (schematic representation on the right) with **1** as acceptor substrate, in the presence of UDP-Rha and SAM (methyltransferase precursor).B) (600 MHz, 310 K, D_2O) HSQC and C) TOCSY spectra of the reaction product **4** with traces of **3**. Labelling follows the letter system of Table S7. The peak at ca. 4.7 ppm is the residual solvent signal.

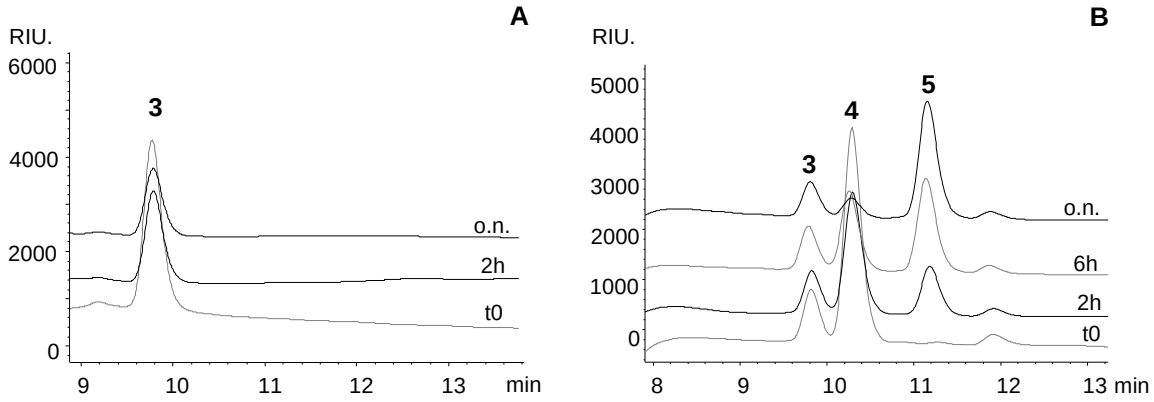


Figure S15. HPLC chromatographic profiles obtained from the reaction of the A061L enzyme with A) **3**; and B) **4**. The two substrates differ by one methyl group, located on the O-2 of the terminal Rha in **4**, but absent in **3**. The compound to which each peak refers to is shown in Figure 2. Note that **4** had minor amounts of **3** (peak at 9.8 min), which remained during the reaction confirming the requirement of A061L for a substrate previously methylated.

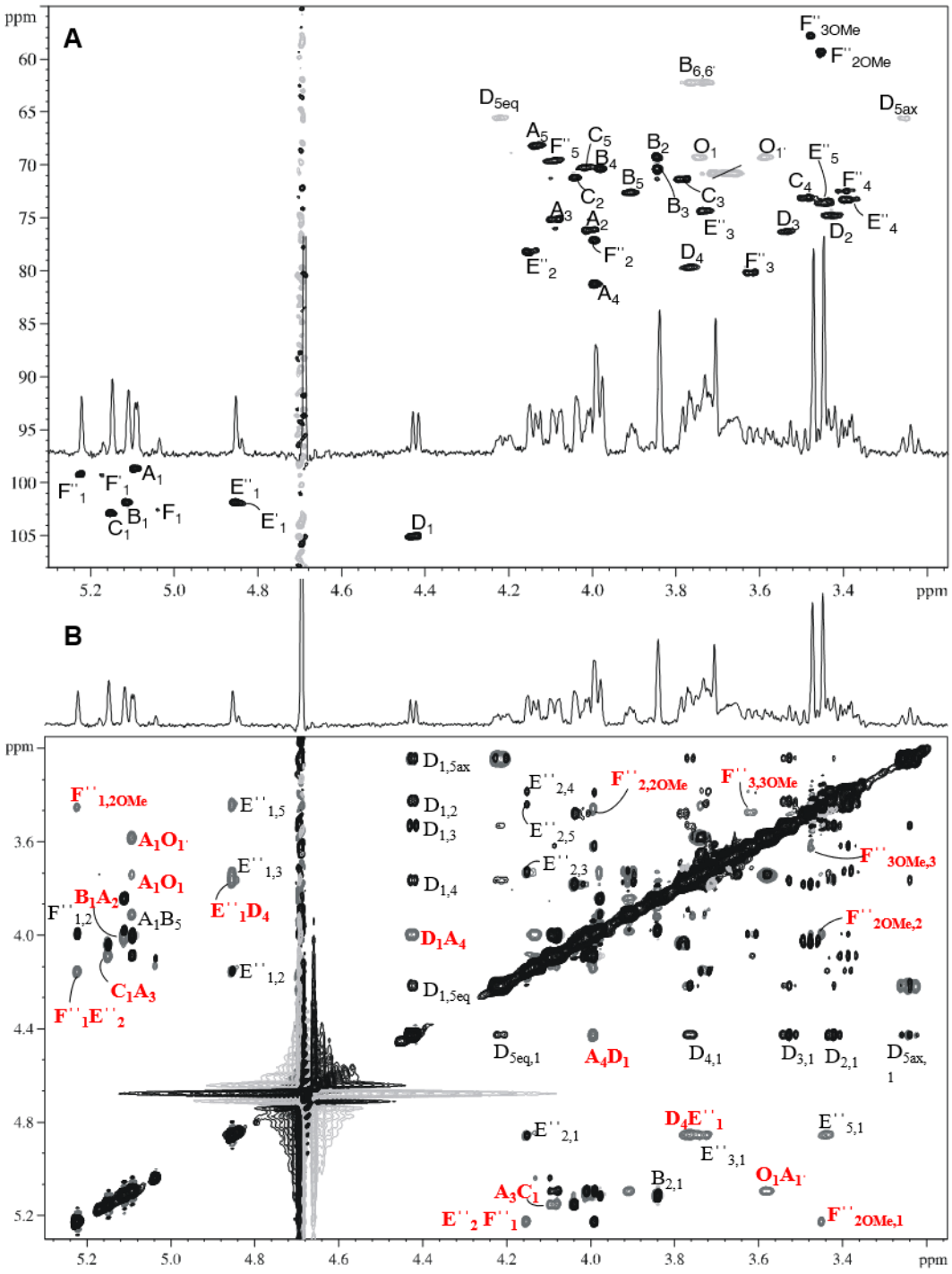


Figure S16. (600 MHz, 310 K, D_2O) NMR spectra of hexasaccharide **5**. A) Zoom of the HSQC spectrum, densities in gray represent carbons bearing two hydrogen atoms. Signals crossed are impurities. B) Superimposition of T-ROESY (gray) and TOCSY (black) spectra. NOE contacts are portrayed in red. Labelling follows the letter system of Table S8. The peak at ca. 4.7 ppm is the residual solvent signal.

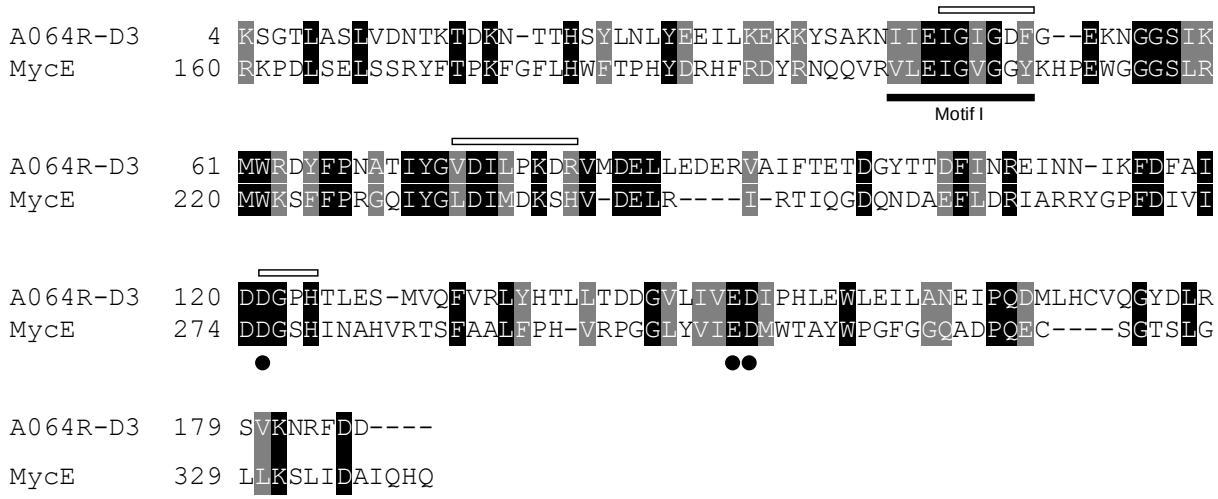


Figure S17. Protein sequence alignment of methyltransferases A064R-D3 and MycE. Identical and similar residues in the multiple alignment are represented by black and gray, respectively. Conserved methyltransferase motif elements are labeled. Predicted A064R-D3 SAM-binding residues positioned within 5 Å of the SAM ligand are organized into three major regions: 46-51, 74-81 and 121-124 and denoted by white bars. Residues involved in metal ion coordination are marked with black circles. Multiple-alignment was performed by T-coffee (7) using structural information and homology extension. File output was compiled by BOXSHADE.

Figure S18. Protein sequence alignment of methyltransferases A061L and NovP. Identical and similar residues in the multiple alignment are represented by black and gray, respectively. Conserved methyltransferase motif elements are labeled. Motif I lies between β1 and α4 of A061L and forms the expected interactions with the amino acyl portion of SAM via Glu21 and Gly23, equivalent to residues Glu 92 and Gly94 in NovP denoted by white bars. The active center of NovP contains a strictly conserved metal-binding site marked with black circles. Multiple-alignment was performed by T-coffee (7) using structural information and homology extension. File output was compiled by BOXSHADE.

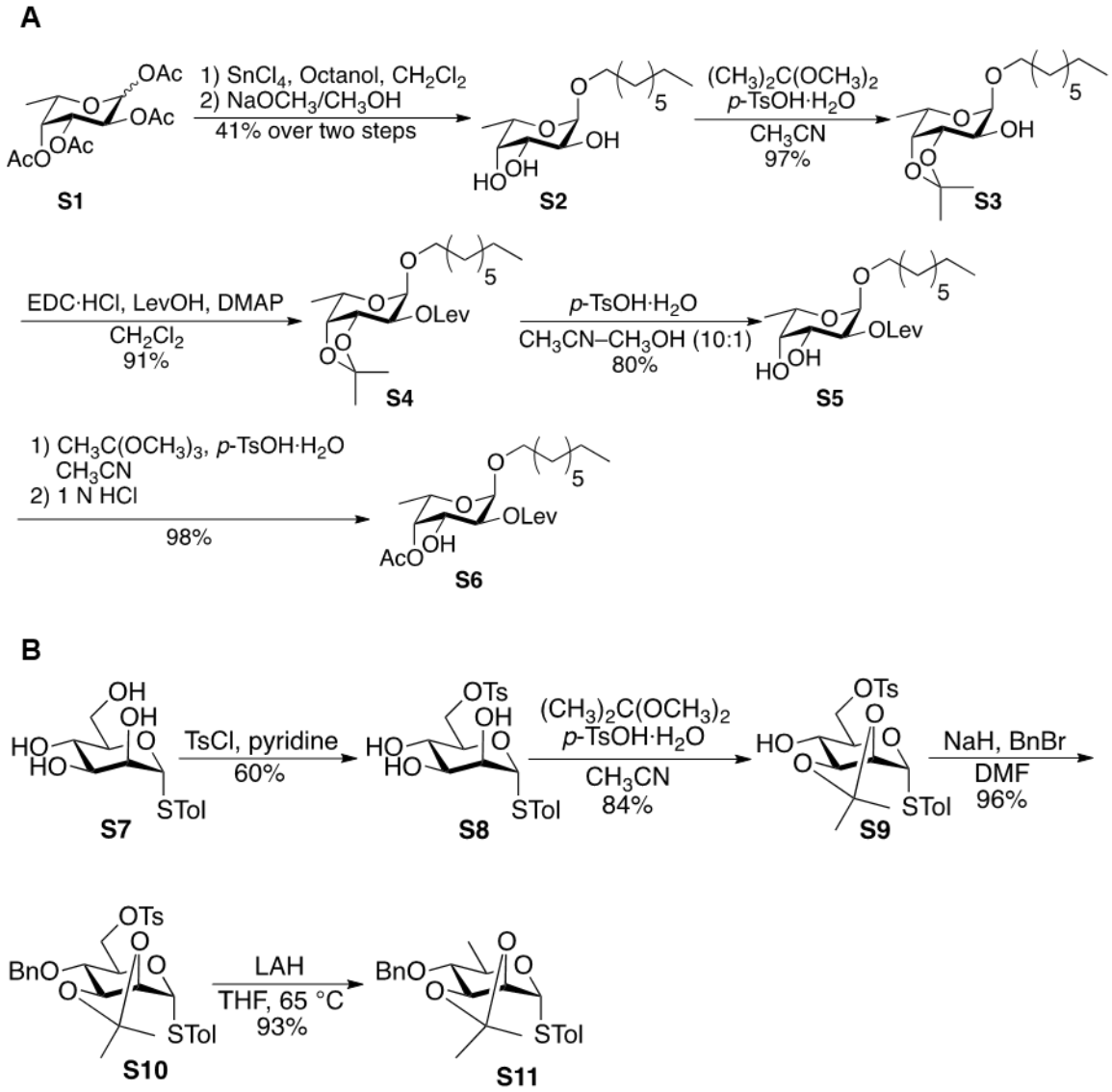


Figure S19. Synthesis scheme of the two precursors of tetrasaccharide **1**. **A.** Synthesis of octyl 4-*O*-acetyl-2-*O*-levulinoyl-α-L-fucopyranoside (**S6**). **B.** Synthesis of *p*-Tolyl 4-*O*-benzyl-2,3-*O*-isopropylidene-1-thio-α-D-rhamnopyranoside (**S11**). Reagents and conditions are showed in figure.

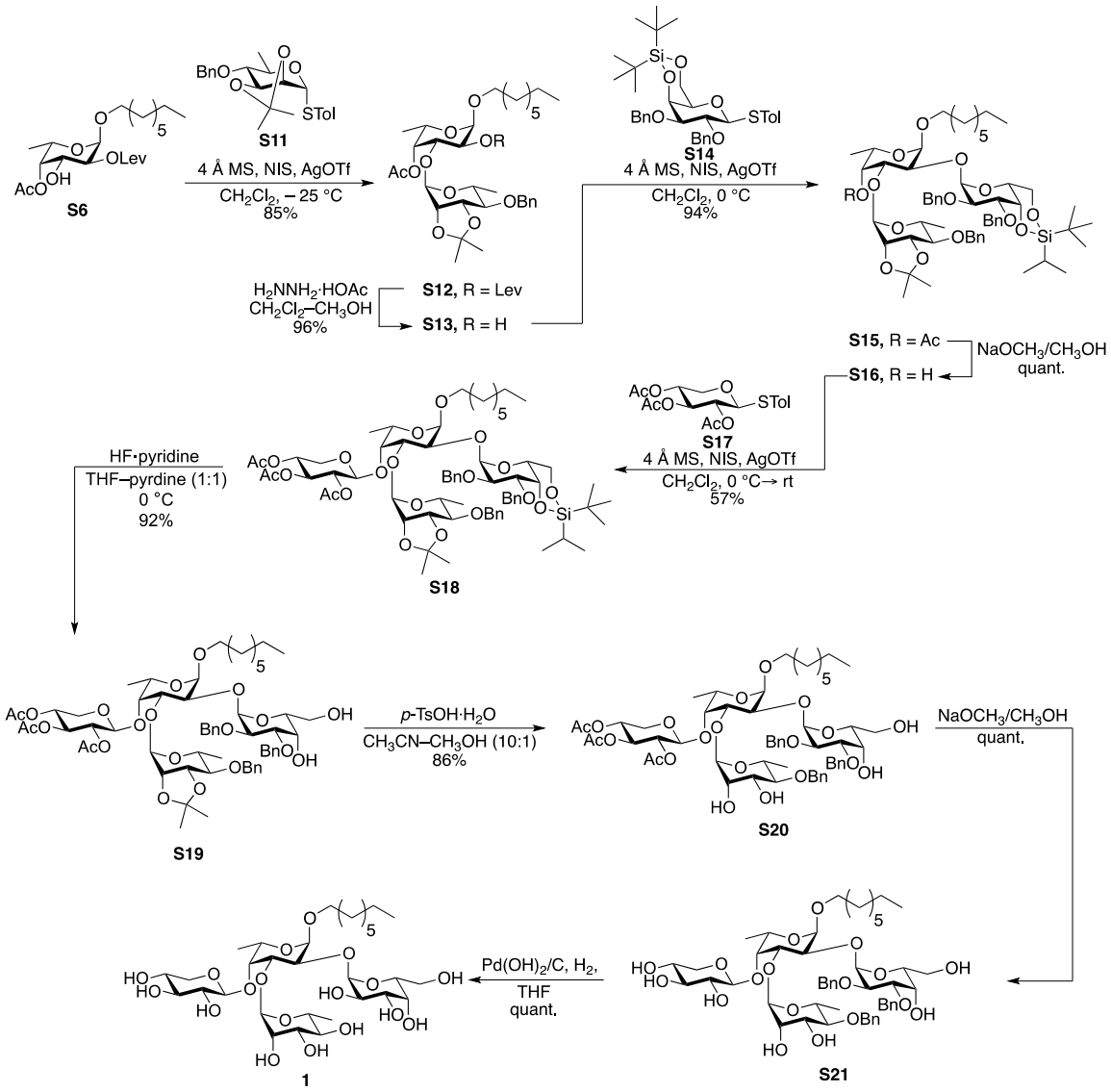
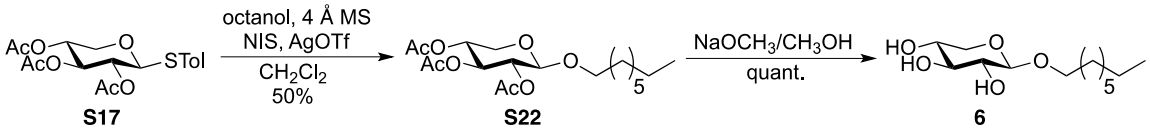


Figure S20. Synthesis of tetrasaccharide **1** using **S6** and **S11** as reagents (Fig. S19). Reagents and conditions are showed in the figure.



S17 S22 6

Figure S21. Synthesis scheme of **6.** Reagents and conditions are showed in the figure.

Table S1. (600 MHz, 310 K, D_2O) Proton and carbon chemical shifts of the synthetic substrate **1**. For structure, refer to Figure 2.

		1,1'	**2**	**3**	**4**	**5_{eq},5_{ax}**	**6,6'**	**7**	**8**
A	^{1}H	5,08	3,99	4,08	3,97	4,13	1,31		
2,3,4-α-Fuc	*^{13}C*	*99,0*	*76,5*	*74,7*	*81,6*	*68,3*	*16,4*		
B	^{1}H	5,10	3,83	3,84	3,97	3,90	3,75 - 3,72		
t-α-Gal	*^{13}C*	*102,2*	*69,4*	*70,6*	*70,5*	*72,7*	*62,4*		
C	^{1}H	5,16	4,02	3,78	3,48	3,96	1,27		
t-α-Rha	*^{13}C*	*102,9*	*71,4*	*71,4*	*73,3*	*70,4*	*17,9*		
D'	^{1}H	4,44	3,38	3,44	3,65	4,01 - 3,23	-		
t-β-Xyl	*^{13}C*	*105,6*	*74,9*	*77,0*	*70,7*	*66,2*	-		
O	^{1}H	3,73-3,57	1,62	1,37	1,31	1,30	1,28	1,29	0,86
Octyl	*^{13}C*	*69,4*	*29,8*	*26,9*	*29,7*	*29,7*	*32,4*	*23,3*	*14,6*

Table S2. (600 MHz, 310 K, D_2O) ^{1}H and ^{13}C chemical shift values of pentasaccharide **2** obtained from the reaction between the A064R-D1 enzyme and **1**. For structure, refer to Figure 2. Chemical shift of the octyl group are the same reported in Table S1.

		1	**2**	**3**	**4**	**5_{eq},5_{ax}**	**6,6'**
A	^{1}H	5,09	3,99	4,08	3,98	4,13	1,31
2,3,4-α-Fuc	*^{13}C*	*98,8*	*76,2*	*75,0*	*81,4*	*68,3*	*16,4*
B	^{1}H	5,11	3,83	3,84	3,97	3,90	3,76 - 3,73
t-α-Gal	*^{13}C*	*101,9*	*69,3*	*70,4*	*70,3*	*72,6*	*62,2*
C	^{1}H	5,15	4,03	3,78	3,48	4,01	1,27
t-α-Rha	*^{13}C*	*102,9*	*71,2*	*71,3*	*73,1*	*70,3*	*17,9*
D	^{1}H	4,44	3,42	3,57	3,76	4,23 - 3,28	-
4-β-Xyl	*^{13}C*	*105,2*	*74,7*	*75,9*	*79,4*	*65,3*	-
E	^{1}H	4,79	4,06	3,58	3,34	3,39	1,30
t-β-Rha	*^{13}C*	*102,0*	*71,8*	*73,8*	*73,2*	*73,2*	*18,1*

Table S3. (600 MHz, 310 K, D_2O) Proton and carbon chemical shifts of hexasaccharide **3** obtained from the reaction of A064R-D2 with oligosaccharide **2**. For structure, refer to Figure 2. Chemical shift of the octyl group are the same reported in Table S1.

		1	**2**	**3**	**4**	**5_{eq},5_{ax}**	**6,6'**
A	^{1}H	5,09	3,99	4,08	3,98	4,12	1,30
2,3,4-α-Fuc	*^{13}C*	*98,7*	*76,2*	*75,3*	*81,4*	*68,3*	*16,3*
B	^{1}H	5,10	3,83	3,83	3,97	3,90	3,75 - 3,72
t-α-Gal	*^{13}C*	*101,9*	*69,3*	*70,4*	*70,4*	*72,6*	*62,3*
C	^{1}H	5,14	4,03	3,77	3,47	4,01	1,27
t-α-Rha	*^{13}C*	*103,0*	*71,2*	*71,4*	*73,2*	*70,2*	*18,1*
D	^{1}H	4,42	3,41	3,53	3,75	4,20 - 3,24	-
4-β-Xyl	*^{13}C*	*105,1*	*74,8*	*76,3*	*79,7*	*65,6*	-
E'	^{1}H	4,83	4,12	3,69	3,44	3,38	1,30
2-β-Rha	*^{13}C*	*102,0*	*78,1*	*74,4*	*73,4*	*73,4*	*18,1*
F	^{1}H	5,03	4,09	3,84	3,42	4,08	1,30
t-α-Rha	*^{13}C*	*102,6*	*71,3*	*71,2*	*73,4*	*69,7*	*18,1*

Table S4. (600 MHz, 298 K, D_2O) Proton and carbon chemical shifts of the Rha disaccharide obtained from the reaction of the A064R-D2 with the Rha monosaccharide as an acceptor substrate.

		1	**2**	**3**	**4**	**5**	**6,6'**
R	^{1}H	5,22	3,91	3,89	3,46	3,86	1,28
2-α-Rha	*^{13}C*	*93,7*	*80,3*	*70,9*	*73,4*	*69,5*	*17.0.*
T	^{1}H	4,96	4,07	3,80	3,44	3,76	1,28
t-α-Rha	*^{13}C*	*103,3*	*71,1*	*71,2*	*73,2*	*70,2*	*17.0.*

Table S5. (600 MHz, 310 K, D_2O) Proton and carbon chemical shifts of the trisaccharide **7** obtained from the reaction of the A064R-D1D2 with **6** as acceptor substrate. For structure, refer to Figure 2. Chemical shift of the octyl group are the same reported in Table S1.

		1	**2**	**3**	**4**	**5_{eq},5_{ax}**	**6,6'**
D	^{1}H	4,36	3,26	3,52	3,70	4,10 - 3,32	-
4-β-Xyl	*^{13}C*	*103,9*	*74,3*	*76,4*	*79,4*	*65,7*	-
E'	^{1}H	4,84	4,12	3,69	3,38	3,43	1,31
2-β-Rha	*^{13}C*	*102,0*	*78,1*	*74,4*	*73,3*	*73,3*	*n.d.*
F	^{1}H	5,03	4,09	3,84	3,42	4,08	1,28
t-α-Rha	*^{13}C*	*102,6*	*71,3*	*71,1*	*73,4*	*69,7*	*n.d.*

Table S6. Organism identification by UniProtKB gene accession number or GenBank ID.

Organism	ID
Afipia	WP_156946979.1
Aggregatibacter	Q9JRS1
Anseongella	WP_132130041.1
Azospirillum	WP_119832210.1
Bacteroidetes	A0A317GZE4
Burkholderia	H8WGC9
Caulobacteraceae	NBW23558.1
Dishui Lake large algae virus 1	QIG60142.1
Enterococcus	Q8GC54
Franconibacter	A0A172WYM6
Gammaproteobacteria	A0A2E6AA56
Hyella	WP_144864280.1
Klebsiella	A0A193SDA5
Lacinutrix	F6GHX4
Lactobacillus	Q58Z06
Leeuwenhoekiella	WP_072983890.1
Lobelia	A0A140KFJ1
Morus	A0A0M4KE42
Mucilaginibacter	WP_100340013.1
Mycobacterium	O69000
Olivibacter	WP_093328425.1
Olleya	WP_036599855.1
PBCV-1	Q89399, Q89396
Prokaryotic dsDNA virus	A0A516L7Y9
Pseudomonas	Q51560
SAR202 cluster bacterium	A0A5N9AEU9
Schlesneria	HGT38096.1
Shigella	A0A510B1Y5
Solanum	Q2I6N7
Sphingomonas	WP_116840279.1
Streptococcus	F8WRN6
Streptococcus	E5F5C9
Streptomyces	WP_051767384.1
Uncultured archaeon	A0A482D791

Table S7. (600 MHz, 310 K, D_2O) Proton and carbon chemical shifts of the hexasaccharide **4** obtained from the reaction of the A064R with **1**. ^{1}H and ^{13}C chemical shifts of the methyl group are 3, 45 and *59,6* ppm, respectively. Symbol * means that residues are in traces. For structure, refer to Figure 2. Chemical shift of the octyl group are the same reported in Table S1.

		1	**2**	**3**	**4**	**5_{eq},5_{ax}**	**6,6'**
A	^{1}H	5,09	3,99	4,08	3,99	4,13	1,31
2,3,4-α-Fuc	*^{13}C*	*98,7*	*76,1*	*75,1*	*81,2*	*68,2*	*16,2*
B	^{1}H	5,10	3,84	3,84	3,98	3,90	3,76 - 3,72
t-α-Gal	*^{13}C*	*101,8*	*69,2*	*70,4*	*70,4*	*72,6*	*62,1*
C	^{1}H	5,14	4,04	3,78	3,48	4,01	1,26
t-α-Rha	*^{13}C*	*102,9*	*71,1*	*71,2*	*73,0*	*70,1*	*17,9*
D	^{1}H	4,42	3,42	3,53	3,75	4,21 - 3,24	-
4-β-Xyl	*^{13}C*	*105,0*	*74,7*	*76,3*	*79,7*	*65,5*	-
***E'**	^{1}H	4,83	4,13	3,69	3,38	3,44	1,30
2-β-Rha	*^{13}C*	*101,9*	*78,1*	*74,3*	*73,4*	*73,4*	*17,9*
E''	^{1}H	4,85	4,14	3,72	3,37	3,44	1,30
2-β-Rha	*^{13}C*	*101,9*	*78,3*	*74,3*	*73,4*	*73,4*	*17,9*
***F**	^{1}H	5,04	4,09	3,86	3,42	4,07	1,30
t-α-Rha	*^{13}C*	*102,5*	*71,3*	*70,9*	*73,4*	*69,5*	*17,9*
F'	^{1}H	5,17	3,74	3,87	3,35	4,05	1,28
t-2OMe-α-Rha	*^{13}C*	*99,2*	*81,2*	*70,9*	*73,5*	*69,5*	*17,9*

Table S8. (600 MHz, 310 K, D_2O) Proton and carbon chemical shifts of the hexasaccharide **5** obtained from the reaction of A061L with **4**. The **F''** residue was substituted at both O-2 and O-3 with a methyl group ($^1H/^{13}C$ 3,45/*59,3* and 3,47/*57,7* ppm, respectively). n.d. means not detected. For structure, refer to Figure 2. Chemical shift of the octyl group are the same reported in Table S1.

		1	**2**	**3**	**4**	**5_{eq},5_{ax}**	**6,6'**
A	^1H	5,09	4,00	4,08	3,99	4,13	1,31
2,3,4-α-Fuc	*^{13}C*	*98,6*	*76,0*	*75,0*	*81,2*	*68,0*	*16,1*
B	^1H	5,11	3,84	3,84	3,98	3,91	3,76 - 3,72
t-α-Gal	*^{13}C*	*101,8*	*69,2*	*70,4*	*70,3*	*72,5*	*62,1*
C	^1H	5,15	4,04	3,78	3,48	4,01	1,26
t-α-Rha	*^{13}C*	*102,8*	*71,1*	*71,2*	*73,0*	*70,1*	*17,8*
D	^1H	4,42	3,42	3,53	3,75	4,21 - 3,24	-
4-β-Xyl	*^{13}C*	*104,9*	*74,7*	*76,2*	*79,6*	*65,5*	-
E''	^1H	4,85	4,15	3,72	3,38	3,44	1,31
2-β-Rha	*^{13}C*	*101,8*	*78,1*	*74,2*	*73,2*	*73,4*	*n.d.*
F''	^1H	5,22	3,99	3,62	3,39	4,08	1,26
t-2,3OMe-α-Rha	*^{13}C*	*99,1*	*76,9*	*80,1*	*72,4*	*69,5*	*17,8*

References

1. K. Bock, S. Pedersen, Carbon-13 nuclear magnetic resonance spectroscopy of monosaccharides. *Adv Carbohydr Chem Biochem* **41**, 27–65 (1983).
2. C. De Castro et al., Structure of the N-linked oligosaccharides attached to virus PBCV-1 major capsid protein: an unusual class of complex N-glycans. *Proc. Natl. Acad. Sci. USA* **110**, 13956-13960 (2013).
3. C.-A. Tai, S. S Kulkarni, S. -C. Hung, Facile $Cu(OTf)_2$-Catalyzed Preparation of Per-*O*-acetylated Hexopyranoses with Stoichiometric Acetic Anhydride and Sequential One-Pot Anomeric Substitution to Thioglycosides under Solvent-Free Conditions *Org. Chem*. **68**, 8719–8722 (2003).
4. S. Meng et al., Efficient assembly of oligomannosides using the hydrophobically assisted switching phase method *Org. Biomol. Chem*. **13**, 6711–6722 (2015).
5. M. Shiozaki et al., Synthesis and biological activity of hydroxylated analogues of KRN7000 (a-galactosylceramide) *Carbohydr. Res.* **370**, 46–66 (2013).
6. B. Yang et al., Chemical Synthesis of a Heparan Sulfate Glycopeptide: Syndecan-1. *Angew. Chem. Int. Ed.* **51**, 10185–10189 (2012).
7. P. Di Tommaso, S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J. M. Chang, J. F. Taly, C. Notredame. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res*. **39**, W13-7 (2011).