Supporting information

Designing interactions by control of protein-ligand complex conformation: Tuning arginine-arene interaction geometry for enhanced electrostatic protein-ligand interactions

A.-L. Noresson,[a] O. Aurelius,[b] C. T. Öberg,[a] O. Engström,[a] A. P. Sundin,[a] M. Håkansson,[c] O. Stenström,[d] M. Akke,[d] D. T. Logan,[b,c] H. Leffler,[e] and U. J. Nilsson[a]

[a]Centre for Analysis and Synthesis, Department of Chemistry, Lund University, Box 124, SE-221 00 Lund, Sweden

[b]Section for Biochemistry and Structural Biology, Center for Molecular Protein Science, Department of Chemistry, Lund University, Box 124, SE-221 00 Lund, Sweden

[c]SARomics Biostructures AB, Scheelevägen 2, SE-223 63 Lund, Sweden

[d]Biophysical Chemistry, Center for Molecular Protein Science, Department of Chemistry, Lund University, Box 124, SE-221 00 Lund, Sweden

[e]Department of Laboratory Medicine, Section MIG, Lund University, Sölvegatan 23, SE-223 62, Lund, Sweden

*Corresponding author: Tel.: +46 46 2228218. Fax: +46 46 2228209. Email: ulf.nilsson@chem.lu.se.

Contents

Synthesis experimental procedures and compound physical data

General synthetic methods

NMR spectra were recorded with Bruker DRX 400 MHz and Avance II MHz spectrometers at ambient temperature. ^{1}H-NMR spectra were assigned using 2D methods (COSY). Chemical shifts are given in ppm downfield from the signal for Me_4Si, with reference to residual $CHCl_3$, DMSO, HDO or CD_2HOD. Fast atom bombardment ionization high-resolution mass spectrometry (FAB-HRMS) was recorded on a Jeol SX-120 mass spectrometer. Electrospray ionization mass spectrometry (ESI-MS) and electrospray ionization high-resolution mass spectrometry (ESI-HRMS) were recorded on a Micromass Q-TOF micro spectrometer. Reactions were monitored by TLC using aluminum-backed silica gel plates (Merck $60F_{254}$) and visualized using UV light and by charring with ethanolic H_2SO_4 (7%). Preparative chromatography was performed using silica gel (Amicon Matrex 35-70 μm, 60 Å) columns. Preparative TLC was performed using glass-backed silica gel plates (200*200*1 mm, $60F_{254}$). Preparative HPLC was performed on a Waters 600 Series HPLC, RP-C18, with a mobile phase H_2O-MeCN gradient containing 0.1% TFA. Analytical purities for compounds **7**, **19**, and **20** were determined on a Waters Acquity UPLC system with a Waters Acquity CSH C18 column, 0.5 mL/min H_2O-MeCN gradient 5-95%, 13 min (0.1% formic acid), and detection at 254 nm. Analytical purity for compound **16** was determined on an Agilent 1260 HPLC system, column Agilent S8-C18, 0.5 mL/min H_2O-MeCN gradient 5-95%, 8 min (0.1% formic acid), and detection at 254 nm. DMF was distilled and stored over 4Å molecular sieves. Other solvents were dried with activated 4Å molecular sieves. Reagents were supplied by Sigma-Aldrich and Acros and used as-is. The anionic compounds were typically stored as their triethylammonium-salts, prepared by concentrating a solution of the anion with excess TEA.

Methyl 2-*O*-acetyl-3-(4-hydroxy-2,3,5,6-tetrafluoro-benzamido)-4,6-*O*-benzylidene-3-deoxy-ß-D-galactopyranoside 4

Compound **3** (81 mg, 0.23 mmol) was dissolved in MeOH and HOAc and palladium on carbon was added. The reaction mixture was degassed under reduced pressure and stirred under hydrogen gas for three hours, filtered and concentrated. The residue was chromatographed (DCM/MeOH: 100:5⇒100:10), evaporated and then solved in DCM (10 mL). BOP-chloride (242 mg, 0.95 mmol) and triethylamine (184 μL, 1.33 mmol) was added and the reaction was stirred overnight, concentrated and chromatographed (Toluene/MeOH, 100:5⇒100:10 gradient) to give **4** (23.4 mg) in 21% yield. ^{1}H NMR (400 MHz, MeOD) δ 8.09 (d, *J* = 9.2 Hz, 1H, *NH*), 7.60-7.49 (m, 2H, ArH), 7.41-7.36 (m, 3H, ArH), 5.62 (s, 1H, benzyl), 5.02 (dd, *J*=11.0, 7.9 Hz, 1H, H2), 4.56 (d, *J*=7.9 Hz, 1H, H1), 4.46-4.35 (m, 1H, H3), 4.27-4.12 (m, 3H, H6+H4), 3.71 (d, J=1.2 Hz, 1H, H5), 3.53 (s, 3H, OCH_3), 2.03 (s, 3H, Ac).

Methyl 2-*O*-acetyl-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-4,6-*O*-benzylidene-3-deoxy-ß-D-galactopyranoside 5

Compound **4** (75 mg, 0.146 mmol) was solved in acetone (35 mL) and Me_2SO_4 (138 μL, 1.46 mmol) and K_2CO_3 (201 mg, 1.46 mmol) were added. The reaction mixture was stirred under nitrogen gas overnight and then evaporated. The residue was solved in EtOAc and washed with HCl (5%, aq, 2*50 mL) and brine (50 mL), dried with $MgSO_4$ and evaporated. Chromatography (DCM/MeOH, 100:1⇒100:10 gradient) gave **5** (52.3 mg) in 68% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56-7.46 (m, 2H, ArH), 7.43-7.34 (m, 3H, ArH), 6.44 (d, *J* = 8.7 Hz, 1H, NH), 5.56 (s, 1H, benzyl), 5.15 (dd, *J*=10.8, 7.9 Hz, 1H, H2), 4,71-4.44 (m, 2H, H3+H1), 4.38 (dd, *J*=12.5, 1.3 Hz, 1H, H6), 4.29 (d, J=3.2 Hz, 1H, H4), 4.12 (d, *J*=1.7 Hz, 1H, H6), 4.09 (d, *J*=1.5 Hz, 3H, OCH_3), 3.63 (s, 1H, H5), 3.56 (s, 3H, OCH_3), 2.11 (s, 3H, Ac).

Methyl 3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-ß-D-galactopyranoside 7

Compound **5** (15.5 mg, 0.029 mmol) was dissolved in MeOH (18 mL) and DCM (18 mL) and NaOMe (500 μL, 1M) were added. The reaction was stirred for three hours, neutralized with Amberlite IR120 (H^+), filtrated and concentrated. The residue was chromatographed (DCM/MeOH: 100:2) and then concentrated and solved in HOAc (4 mL, 80%, aq). The reaction was stirred at 60°C for 5 h, evaporated and chromatographed (DCM/MeOH, 100:2⇒100:8 gradient) to give **7** (6.4 mg) in 56% yield and 99.3% analytical hplc-purity. ^{1}H NMR (400 MHz, MeOD) δ 4.27 (d, *J* = 7.7 Hz, 1H, *H*1), 4.13 (t, *J* = 1.5 Hz, 3H, OC*H*$_3$), 4.10 (dd, *J* = 10.9, 3.2 Hz, 1H, *H*3), 4.00 (d, *J* = 2.6 Hz, 1H, *H*4), 3.80 – 3.69 (m, 2H, H6), 3.68 – 3.58 (m, 2H, H5+H3), 3.57 (s, 3H, OC*H*$_3$); FAB-HRMS m/z calcd for $[C_{15}H_{17}F_4NO_7Na]^+$, 422.0839; found 422.0822.

Methyl 3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-2-*O*-sulfo-ß-D-galactopyranoside 9

Compound **5** (9 mg, 0.018 mmol)) was solved in DMF (4 mL) and SO_3-NMe_3 (26 mg, 0.19 mmol) was added. The reaction was stirred under nitrogen athmosphere at 50°C overnight, concentrated, chromatographed (DCM/MeOH: 100:8) and solved in HOAc (5 mL, 80%, aq). The reaction was stirred at 60°C for two hours, evaporated and chromatographed (Toluene/MeOH, 100:20⇒100:40 gradient) to give **9** (7.2 mg) in 83% yield. Compound **9** was highly unstable in analytical hplc analysis conditions and purity was instead estimated to be >95% according to nmr

analyses (see nmr spectra in SI). ^{1}H NMR (400 MHz, MeOD) δ 4.63 (s, 1H), 4.50 – 4.43 (m, 2H, *H*2+*H*1), 4.32 (d, *J* = 2.2 Hz, 1H, *H*4), 4.13 (t, *J* = 1.5 Hz, 3H, OC*H*$_3$), 4.11 – 4.04 (m, 1H, *H*3), 3.99 (d, *J* = 1.5 Hz, 1H), 3.82 – 3.63 (m, 3H, H5+H6), 3.56 (s, 3H, OC*H*$_3$), 3.22 (q, *J* = 7.3 Hz, 8H, TEA), 1.33 (t, *J* = 7.3 Hz, 14H, TEA); FAB-HRMS m/z calcd for $[C_{15}H_{16}F_4NO_{10}SNa_2]^+$, 524.0232; found, 524.0239.

1,2,4,6-Tetra-O-acetyl-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-ß-D-galactopyranose 13

Compound **12** (1.6 g, 4.29 mmol) was solved in ethanol (90 mL) and cyclohexene (180 mL). Palladium hydroxide (200 mg, 1.42 mmol) was added and the reaction was refluxed at 90°C for 20 min. The reaction was filtered, concentrated and solved in DCM (200 mL) and pyridine (1040 µl, 12.9 mmol) and 4-methoxy-2,3,5,6-tetrafluoro-benzoylchloride (1.25 mg, 5.14 mmol) was added. The reaction was stirred for four hours under nitrogen atmosphere and then washed with $NaHCO_3$ (2*150 mL) and dried with $MgSO_4$. The residue was chromatographed (heptane/EtOAc, 1:1) to give **13** (934 mg) in 40% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.26 (d, *J* = 7.3 Hz, 1H, NH), 5.81 (dd, *J* = 8.2, 3.1 Hz, 1H, H1), 5.64 – 5.41 (m, 1H, H4), 5.16 (ddd, *J* = 11.2, 8.2, 3.0 Hz, 1H, H2), 4.54 (td, *J* = 7.8, 4.5 Hz, 1H, H3), 4.23 – 3.96 (m, 6H, OCH_3+H5+H6), 2.28 – 1.93 (m, 12H, Ac). ESI-HRMS m/z calcd for $[C_{22}H_{23}F_4NO_{11}Na]^+$, 576.1105; found 576.1104.

2,4,6-Tri-O-acetyl-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-α-D-galactopyranosyl bromide 14

Compound **13** (22 mg, 0.040 mmol) was solved in DCM (4 mL) and Ac_2O (10 µL, 0.11 mmol) was added. The reaction mixture was stirred under nitrogen atmosphere and cooled to 0 ˚C. HBr/AcOH (88 µL, 0.043 mmol) was added and the reaction was stirred for 2 hours, washed with $NaHCO_3$ (1*50 mL), brine (1*50 mL), dried with $MgSO_4$ and concentrated to give **14** (20 mg, 0.035 mmol, 89%). The unstable crude bromide was not further purified, but used immediately in the synthesis of **16.** ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.59 (d, *J* = 3.8 Hz, 1H, H1), 6.12 (d, *J* = 7.9 Hz, 1H, NH), 5.66 (dd, *J* = 3.1, 1.1 Hz, 1H, H4), 5.13 (dd, *J* = 11.3, 3.8 Hz, 1H, H2), 4.93 (ddd, *J* = 11.2, 8.0, 3.1 Hz, 1H, H3), 4.55 (t, *J* = 6.4 Hz, 1H, H5), 4.19 (dd, *J* = 11.6, 6.0 Hz, 1H, H6), 4.13 (t, *J* = 1.7 Hz, 3H, OCH_3), 4.06 (dd, *J* = 11.6, 6.9 Hz, 1H, H6), 2.16 (s, 3H, Ac), 2.14 (s, 3H, Ac), 2.07 (s, 3H, Ac).

1,1'-Sulfanediyl-bis-[3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-β-D-galactopyranoside] 16

Compound **14** (20 mg, 0.035 mmol) was solved in Acetonitrile (5 mL). Na_2S (6.7 mg, 0.086 mmol) and molecular sieves (4Å, 2 mg) was heated under vacuum and then added to the solution of **14**. The reaction was stirred overnight and then filtered and concentrated under reduced pressure. The residue was solved in MeOH (10 mL) and NaOMe (160 µL, 1M) was added. The reaction was stirred for 4 hours and then neutralized with Amberlite IR120 (H^+), filtrated and concentrated under reduced pressure. The residue was chromatographed (DCM/MeOH, 6:1) to give **16** (5.9 mg) in 44% yield and 99.8% analytical hplc-purity. ^{1}H NMR (400 MHz, MeOD) δ 4.84 (d, *J* = 9.7 Hz, 2H, H1), 4.18 (dd, *J* = 10.3, 3.1 Hz, 2H, H4), 4.14 (t, *J* = 1.4 Hz, 6H, OCH_3), 4.03 (d, *J* = 3.1 Hz, 2H, H3), 3.84 – 3.65 (m, 8H, H2+H5+H6). FAB-HRMS m/z calcd for $[C_{28}H_{28}F_8N_2O_{12}SNa]^+$, 791.1133; found 791.1152.

1,1'-Sulfanediyl-bis-[4,6-*O*-benzylidene-3-deoxy-3-(3-methoxy-benzamido)-β-D-galactopyranoside] 17

Compound **15** (55 mg, 80 µmol) and catalytic pTsOH-H_2O (1.7 mg, 8.8 µmol) was stirred in CH_3CN (5 mL) and α,α-dimetoxytoluene (132 µL, 882 µmol) at 70°C under an open condenser. After 3.5 h, the reaction was concentrated and purified by column chromatography (10 g silica, DCM/MeOH, 100:5) to give **17** in 48% yield. ^{1}H-NMR (DMSO-d6, 400 MHz) δ 8.43 (d, J = 7.7 Hz, 1 H, NH), 7.52-7.32 (m, 8H, ArH), 7.07 (ddd, J = 8.2, 2.6, 0.9 Hz, 1 H, ArH), 5.59 (s, 1 H, CH), 4.90 (d, J = 9.7 Hz, 1 H, H1), 4.30 (app d, J = 3.4 Hz, 1 H, H4), 4.20 (ddd, J = 11.0, 7.5, 3.5 Hz, 1 H, H3), 4.14 (br d, J = 11.5 Hz, 1 H, H6), 4.08 (br d, J = 12.6 Hz, 1 H, H6), 3.85 (app t, J = 9.8 Hz, 1 H, H2), 3.79 (s, 3 H, OCH_3), 3.70 (br s, 1 H, H5).

1,1'-Sulfanediyl-bis-[4,6-*O*-benzylidene-3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-β-D-galactopyranoside] 18

Compound **16** (24 mg, 0.039 mmol) was solved in acetonitrile (90 mL) and stirred under nitrogen gas. $VO(OTF)_2$ (15 mg, 0.0389 mmol) and benzaldehyde (79 µL, 0.78 mmol) was dissolved in acetonitrile (10 mL), stirred under nitrogen atmosphere for 10 min and then transferred to the solution of **16**. The reaction was stirred under nitrogen atmosphere for 48 h, neutralized with triethylamine (40 µL, 0.29 mmol) and then washed in $NaHCO_3$ (1*100 mL). The water phase was extracted in EtOAC (4*100 mL) and then dried with $MgSO_4$. The residue was chromatographed (DCM/MeOH, 100:1—>100:10 gradient) to give **18** (14.5 mg, 39%). NMR spectra could not be obtained due to very low solubility. ESI-MZ m/z calcd for $[C_{42}H_{36}F_8N_2O_{12}SH]^+$, 944.2; found 944.6.

3,3'-Dideoxy-3,3'-[3-methoxy-benzamido]-2-*O*-sulfo-1,1'-sulfanediyl-di-β-D-galactopyranoside 19 and 1,1'-sulfanediyl-bis-[3-deoxy-3-(3-methoxy-benzamido)-2-sulfo-β-D-galactopyranoside] 21

To a stirred solution of **17** (15 mg, 18.7 μmol) in DMF (0.5 mL) at 70°C was added SO_3-NMe_3 (total amount ca 31 mg, 220 μmol) in portions over 7 days during continuous monitoring. The reaction was concentrated to ca half volume before being purified by preparative TLC (DCM:MeOH 100:10) to give a mix of mono- and di-sulfated products (1.0:0.3 ratio). The crude mix was stirred in HOAc (1.8 mL) at rt for 13 h and 60°C for 1 h before being concentrated and purified by preparative TLC (DCM:MeOH:H_2O 130:35:5) followed by preparative hplc to give **19** (2 mg) and **21** (1.4 mg) in 11% and 7% yield respectively. **19** had analytical hplc-purity 98.5% and data: ^{1}H-NMR (CD_3OD, 400 MHz) δ 8.47 (d, J = 5.3 Hz, 0.6 H, NH), 7.58-7.45 (m, 4 H, ArH), 7.39-7.32 (m, 2 H, ArH), 7.12-7.05 (m, 2 H, ArH), 5.16 (d, J = 9.7 Hz, 1 H, H1), 4.80 (d, J = 9.6 Hz, 1 H, H1'), 4.71 (app t, J = 10.3 Hz, 1 H, H2), 4.40 (br d, J = 2.6 Hz, 1 H, H4), 4.17-4.11 (dd, J = 10.4, 3.0 Hz, 2 H, H3+H3'), 4.05 (br d, J = 2.9 Hz, 1 H, H4') 3.91 (app t, J = 9.9 Hz, 1 H, H2'), 3.86 (s, 3 H, OCH_3), 3.86 (s, 3 H, OCH_3'), 3.85-3.65 (m, 6 H, H5+H5'+2*H6+2*H6'), 3.14 (q, J =7.3 Hz, 18 H, TEA), 1.29 (t, J = 7.3 Hz, 30 H, TEA); ESI-HRMS m/z calcd for $[C_{28}H_{35}N_2O_{15}S_2]^-$, 703.1479 found, 703.1489. **21** had data: ^{1}H-NMR (CD_3OD, 400 MHz) δ 8.58-8.53 (m, 0.3 H, NH), 7.56-7.49 (m, 2 H, ArH), 7.35 (app t, J = 8.0 Hz, 1 H, ArH), 7.07 (ddd, J = 8.2, ~2.6, ~1 Hz, 1 H, ArH), 5.22 (d, J = 9.7 Hz, 1 H, H1), 4.73 (app t, J =10.1 Hz, 1 H, H2), 4.41 (br d, J = 2.2 Hz, 1 H, H4), 4.13 (dd, J = 10.5, 2.7 Hz, 1 H, H3), 3.88-3.81 (m, 4 H, H5+OCH_3), 3.75 (dd, J = 11.4, 4.0 Hz, 1 H, H6), 3.70 (dd, J = 10.8, 4.2 Hz, 1 H, H6), 3.17 (q, J = 7.3 Hz, 2 H, TEA), 1.30 (t, J = 7.3 Hz, 4 H, TEA); ESI-HRMS m/z calcd for $[C_{28}H_{35}N_2O_{18}S_3]^-$, 783.1047; found, 783.1022.

3,3'-Dideoxy-3,3'-[4-methoxy-2,3,5,6-tetrafluoro-benzamido]-2-*O*-sulfo-1,1'-sulfanediyl-di-β-D-galactopyranoside 20

Compound **18** (12.3 mg, 0.013 mmol) and SO_3-NMe_3 (18.1 mg, 0.13 mmol) was dissolved in DMF (4 mL). The reaction mixture was stirred under nitrogen atmosphere at 80 ˚C overnight and then concentrated under reduced pressure. The residue was chromatographed (DCM/MeOH, 8:1), concentrated and solved in HOAc (5 mL, 80%, aq). The reaction mixture was stirred for 4 hours at 60 ˚C and concentrated under reduced pressure. The residue was purified by flash chromatography (DCM/MeOH, 4:1) and preparative HPhplcLC to give **20** (3.4 mg, 36% over two steps and 99.4% analytical hplc-purity). ^{1}H NMR (400 MHz, MeOD) δ 5.11 (d, *J* = 9.7 Hz, 1H, H1), 4.79 (d, *J* = 9.7 Hz, 1H, *H*1'), 4.67 – 4.54 (m, 1H, H2), 4.34 (d, *J* = 2.6 Hz, 1H, H4), 4.24 – 4.09 (m, 8H, OCH_3+OCH_3'+H3+H3'), 4.03 (d, *J* = 3.1 Hz, 1H, H4'), 3.87 – 3.65 (m, 7H, H2'+H5+H5'+2*H6+2*H6'); FAB-HRMS m/z calcd for for $[C_{28}H_{27}F_8N_2O_{15}S_2]^-$, 847.0731; found, 847.0725.

Competitive fluorescence polarization experiments determining galectin affinities

Human galectin-3[28] and mutants R144S,[18] R186S,[18] R144K,[29] and R186K[29] were expressed and purified as earlier described. Fluorescence polarization experiments were performed on a POLARStar plate reader with software FLUOstar Galaxy software or a PheraStarFS plate reader with software PHERAstar Mars version 2.10 R3 (BMG, Offenburg, Germany) and fluorescence anisotropy of fluorescein tagged probes measured with excitation at 485 nm and emission at 520 nm. K_d values were determined in PBS as previously described[15, 16, 30] with specific conditions for each galectin as described below. Compounds were dissolved in neat DMSO at 100 mM and diluted in PBS to 3-6 different concentrations to be tested in duplicates. K_d average and SEM were calculated from 4 to 25 single point measurements showing between 30-70% inhibition. Experiments were done at 20°C.

Galectin-3 affinities. Experiments were done with galectin-3 at 0.20 μM and the fluorescent probe 3,3'-dideoxy-3-[4-(fluorescein-5-yl-carbonylaminomethyl)-1*H*-1,2,3-triazol-1-yl]-3'-(3,5-dimethoxybenzamido)-1,1'-sulfanediyl-di-β-D-galactopyranoside[31] at 0.02 μM.

Galectin-3 R144S affinities. Experiments were done with galectin-3 R144S at 0.30 μM and the fluorescent probe 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1–2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside[15, 16] at 0.02 μM.

Galectin-3 R144K affinities. Experiments were done with galectin-3 R144K at 0.40 μM and the fluorescent probe 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1–2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside at 0.02 μM.

Galectin-3 R186S affinities. Experiments were done with galectin-3 R186S at 3.50 μM and the fluorescent probe 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1–2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside at 0.1 μM.

Galectin-3 R186K affinities. Experiments were done with galectin-3 R186K at 0.90 μM and the fluorescent probe β-D-galactopyranosyl(1—4)-2-acetamido-2-deoxy-β-D-glucopyranosyl(1—3)-β-D-galactopyranosyl(1—4)-(*N*1-fluorescein-5-yl-carbonylaminomethylcarbonyl)-β-D-glucopyranosylamine[32] at 0.1 μM.

Crystallization, data collection and refinement

All diffraction data, except for compound **16**, were collected at MAX-lab beamlines I911-2[33] on a 165 mm marMosaic CCD detector. Data for compound **16** were collected at station I911-3[34] on a 225 mm marMosaic detector. Generally, a low dose, low-resolution dataset was collected first, followed by a higher dose, high-resolution dataset, to avoid saturation of the detector. Data were integrated and scaled using XDS and XSCALE.[35] All structures were solved by

direct refinement of existing PDB structures against the new data. Five percent of the reflections were used for calculation of the free R-factors. Manual model building was carried out in Coot.[36] Water molecules were added to all structures based on three criteria: positive peaks (> 3σ) in the difference electron density maps (m|Fo|-d|Fc|); also present at 1σ in the 2m|Fo|-d|Fc| maps; reasonable hydrogen bonds. Riding hydrogen atoms were added to all models. The quality of all structures was checked using the MolProbity server.[37] The sections below describe the unique features of the structure determination of each complex. Molecular images were generated using PyMOL (Schrodinger LLC).

Co-crystallization of galectin-3 C-terminal domain with compound 15

A galectin-3C[4] (C-terminal domain) solution (9.5 µL, 19 mg/ml in 10 mM phosphate pH 7.5, 100 mM NaCl, 100 mM Lactose, 10 mM ß-mercaptoethanol and 0.02 % NaN_3) was mixed with compound **15** (0.5 µL, 100 mM in 100% DMSO) and incubated on ice for 1.5 hour. Crystallization drops of 2+2 µL were set up over 1 mL reservoir solution (29% PEG 4000, 0.1 M Tris/HCl pH 7.5, 0.1 M $MgCl_2$, 0.4 M NaSCN, 8 mM ß-mercaptoethanol). Immediately after set-up the drops were seeded by adding 0.3 µL of crushed co-crystals with lactose. Crystals were crushed using Hampton Research seed bead and the stabilizing solution contained 1 mM lactose, 29 % PEG 4000, 0.1 M Tris/HCl pH 7.5, 0.1 M $MgCl_2$, 0.4 M NaSCN and 8 mM ß-mercaptoethanol. The galectin-3 C-terminal lactose co-crystals spontaneously nucleated at the same condition as above. Co-crystals of compound **15** and galectin-3C formed within a few days. A crystal measuring 0.1 x 0.05 x 0.05 mm was flash-frozen in cryo solution (15 % glycerol, 25.5 w/v % PEG 4000, 0.25 M NaSCN, 85 mM Tris/HCl pH 7.5, 85 mM $MgCl_2$, 4 mM compound **15**). The structure was determined using PDB entry 1A3K[38] as starting model using SHELXL.[39] Initial R-factors after rigid body refinement to 2.5 Å were 0.389 (0.380). Then the model was subjected to conjugate gradient least squares refinement, gradually increasing the resolution to 1.2 Å with isotropic B-factors. Water molecules were added using SHELXWAT,[39] then anisotropic B-factors were introduced. Restraints for compound **15** were generated using the CCP4[40] monomer sketcher. Bond and angle restraints were used for all atoms and planar restraints for the phenyl rings. Alternative side chain conformations were built for 25 amino acids and refined with an occupancy of 0.5 each. Twenty-four partially occupied water molecules (occupancy 0.5) were added in areas where these molecules are bound to side chains with occupancy of 0.5. In a few areas the main chain was also built with double conformations. Final R-factors were 0.123 (0.177).

Co-crystallization of galectin-3 C-terminal domain with compound 16

A galectin-3C:lactose co-crystal was made by mixing equal volumes of 9 mg/ml galectin-3C in buffer A with 10 mM lactose and a reservoir consisting of 0.1 M Tris-HCl pH 7.8, 0.3 M NaSCN, 0.1 M $MgCl_2$, 32% [w/v] PEG6000 and 20 mM ß-mercaptoethanol in a hanging drop vapor diffusion setup, was soaked in the same reservoir supplemented with 19% (v/v) PEG400 and a 1:27 dilution of an approximately 100 mM **16** stock in 100% DMSO. Rigid body refinement with the galectin-3C:lactose complex to 0.86 Å (PDB code 3ZSJ[20]) was done at 3 Å with Refmac5.[41] Further work with Refmac5 involved restrained refinement using the full resolution range up to 1.12 Å. For the final refinement cycles phenix.refine[42] was used with ADP and stereochemistry weight optimization and anisotropic B-factors for all atoms except hydrogens. Restraints for **16** were generated with eLBOW[43] in the Phenix package. Final R-factors were 0.130 (0.156).

Co-crystallization of galectin-3 C-terminal domain with compound 19

Soaking and complex formation was done as described for **15**. A crystal measuring 0.2 x 0.2 x 0.2 mm was flash-frozen in cryo solution (15 % glycerol, 25.5 w/v % PEG 4000, 0.25 M NaSCN, 85 mM Tris/HCl pH 7.5, 85 mM $MgCl_2$, 4 mM compound **19**). The structure was determined as for compound 15. Initial R-factors after isotropic B-factor refinement to 1.08 Å were 0.283 (0.290). This was followed by anisotropic B-factor refinement to 1.08 Å with R-factors 0.256 (0.265). The ligand was fitted in the electron density and refined, leading to a drop in R-factors of 0.02. About 200 water molecules were automatically added using the ARP/wARP[44] procedure in Refmac5, leading to R-factors of 0.150 (0.166). Subsequently the structure was refined in SHELXL, using bond and angle restraints for all atoms and planar restraints for the phenyl rings and immediate connecting atoms. Subsequent anisotropic refinement gave improved R-factors: 0.126 (0.158). Several refinement and rebuilding cycles were done. The loop region 185–188 could be traced with double conformations, leading to R-factor 0.106 (0.142). Double conformations have been built for 34 amino acids. Twenty-one water molecules were refined with partial occupancy.

Co-crystallization of galectin-3 C-terminal domain with compound 20

A galectin-3C solution (9.5 µL, 19 mg/mL in 10 mM phosphate pH 7.5, 100 mM NaCl, 10 mM ß-mercaptoethanol and 0.02 % NaN_3) was mixed with 0.5 µL of 50 mM compound **20** in 100% DMSO and incubated on ice for 1.5 hours. The solution was centrifuged at 7000 rpm, 4°C for 10 min. A 2+2 µL hanging drop was set up over 1 mL reservoir

solution (29% PEG 4000, 0.1 M Tris/HCl pH 7.5, 0.1 M $MgCl_2$, 0.4 M NaSCN, 8 mM ß-mercaptoethanol). A crystal measuring 0.3 x 0.3 x 0.8 mm was flash-frozen in cryo solution (15 % glycerol, 25.5 w/v % PEG 4000, 0.25 M NaSCN, 85 mM Tris/HCl pH 7.5, 85 mM $MgCl_2$, 2.5 mM compound **20** and 8 mM ß-mercaptoethanol). The structure was determined by using the apo structure of galectin-3C ; PDB ID 3ZSL[20]) determined to 1.08 Å resolution as starting model. All refinement was done using SHELXL program. Gradually increasing the resolution from 2.5 to 1.2 Å resulted in R-factors 0.150 (0.200). A model for **20** was generated using **19** as a starting point, exchanging the hydrogen atoms on the phenyl rings for fluorine atoms. Adding the ligand and refining anisotropic B-factors to 0.96 Å resulted in R-factors 0.149 (0.179). Final R-values after addition of H atoms and alternate conformations for 28 amino acids were 0.128 (0.178).

Procedures for isothermal titration calorimetry

Isothermal titration calorimetry (ITC): ITC measurements were performed using a VP-ITC microcalorimeter from Microcal, Inc. (Northampton, MA, USA). Galectin-3 and ligands **15**, **16**, **19**, and **20** were dissolved in the same PBS buffer (10 mM sodium phosphate buffer, 0.15 M sodium chloride, 0.02% sodium azide and 4% DMSO at pH 7.2) to minimize heat of dilution effects. The pH of the solutions of galectin-3, **15**, **16**, **19**, and **20** was 7.16, 7.15, 7.14, 7.16, and 7.15, respectively. Individual injections (10 μL) of the ligand (0.100–0.050 mM) were added to the galectin-3 solution (10 μM, cell volume 1.4288 mL) from a stirring microsyringe (300 rpm) at intervals of 300 s for **15** and **19,** and 150 s for **16** and **20**, for a total of 30 injections. Experiments were carried in duplicate for ligand **15** and triplicate for all other ligands. The temperature was set to 25°C for ligands **15** and **19** and 28°C for ligands **16** and **20**. The parameters ΔH, K_a (= $1/K_d$), and n were fitted to the ITC data using in-house Matlab routines that enable global fitting of multiple datasets. All datasets were fitted using a single-site binding model:

$$q_i = V_0 \Delta H \left([ML]_i - [ML]_{i-1} \left[1 - \frac{\upsilon_i}{V_0} \right] \right)$$

where q_i is the heat of binding upon injection i, V_0 is the active cell volume, ΔH is the enthalpy of binding, υ_i is the volume of injection i, $[ML]_i$ is the concentration of protein–ligand complex after injection i:

$$[ML]_i = \frac{1 + n[M]_{T,i}K_a + K_a[L]_{T,i} - \sqrt{\left(1 + n[M]_{T,i}K_a + K_a[L]_{T,i}\right)^2 - 4n[M]_{T,i}K_a^2[L]_{T,i}}}{2K_a}$$

in which $[M]_{T,i}$ is the total protein concentration after injection i, $[L]_{T,i}$ is the total ligand concentration in the cell after injection i, K_a is the association constant and n is the effective number of binding sites on the protein, which accounts for the possibility that a fraction of the protein is inactive and cannot bind ligand, or that the concentrations of ligand or protein are not perfectly accurate.

Table S1. Data processing and refinement statistics for the X-ray crystal structures.

Values in parentheses are for the highest resolution shell, unless noted otherwise.

compound	**15**	**16**	**19**	**20**
PDB code	4BLJ	5IUQ	4BLI	4BM8
station	I911-2	I911-3	I911-2	I911-2
wavelength [Å]	1.0379	1.0000	1.0379	1.0408
unit cell (Å)	a =36.2	a= 36.1	a = 36.0	a = 35.8
	b = 57.9	b= 57.6	b = 58.1	b = 58.2
	c = 62.6	c= 62.2	c = 62.4	c = 62.8
resolution range [Å]	30.00–1.20	28.8–1.12	30.0–1.08	30.0–0.96
	(1.23–1.20)	(1.20–1.12)	(1.11–1.08)	(0.98–0.96)
completeness [%]	97.3 (94.9)	99.4 (98.0)	99.2 (97.5)	98.0 (92.3)
unique reflections	40 084	50 259	56 420	79 320
multiplicity	5.1 (4.0)	4.8 (3.5)	5.0 (3.9)	5.2 (2.7)
R_{merge} [%]	5.3 (54.0)	5.0 (53.9)	4.0 (11.7)	3.2 (56.5)
mean I/σ(I)	15.8 (2.7)	16.4 (2.3)	22.2 (10.6)	20.9 (2.0)
Wilson B-factor [Å^2]	8.9	10.9	8.7	12.3
refinement program	SHELXL-97	phenix.refine	SHELXL-97	SHELXL-97
R_{model} (F) [%]	12.3	13.0 (22.4)	10.6	12.8
R_{free} (F) [%]	17.7	15.6 (22.9)	14.2	15.5
reflections used in refinement	38795	50 258	53 600	75 348
(for R_{free})	(2019)	(2 579)	(4 260)	(3 965)
average B-factors [Å^2]	protein: 13.6	protein: 13.5	protein:12.7	protein: 15.8
	ligand: 25.4	ligand: 22.5	ligand: 28.7	ligand: 27.4
	solvent: 34.0	solvent: 29.4	solvent: 31.9	solvent: 32.8
Ramachandran outliers [%]	0.0	0.0	0.0	0.0
rotamer outliers [no. and %]	2 (1.3%)	1 (0.7%)	4 (2.6%)	3 (2.0%)
MolProbity clash score	5.26	7.80	4.29	5.50
bond length rmsd from ideal [Å]	0.012	0.025	0.014	0.006
bond angle rmsd from ideal [Å or °]	0.030	2.29*	0.032	0.023

*rms deviations from ideal angles are measured in degrees in phenix.refine and as distances in SHELXL.

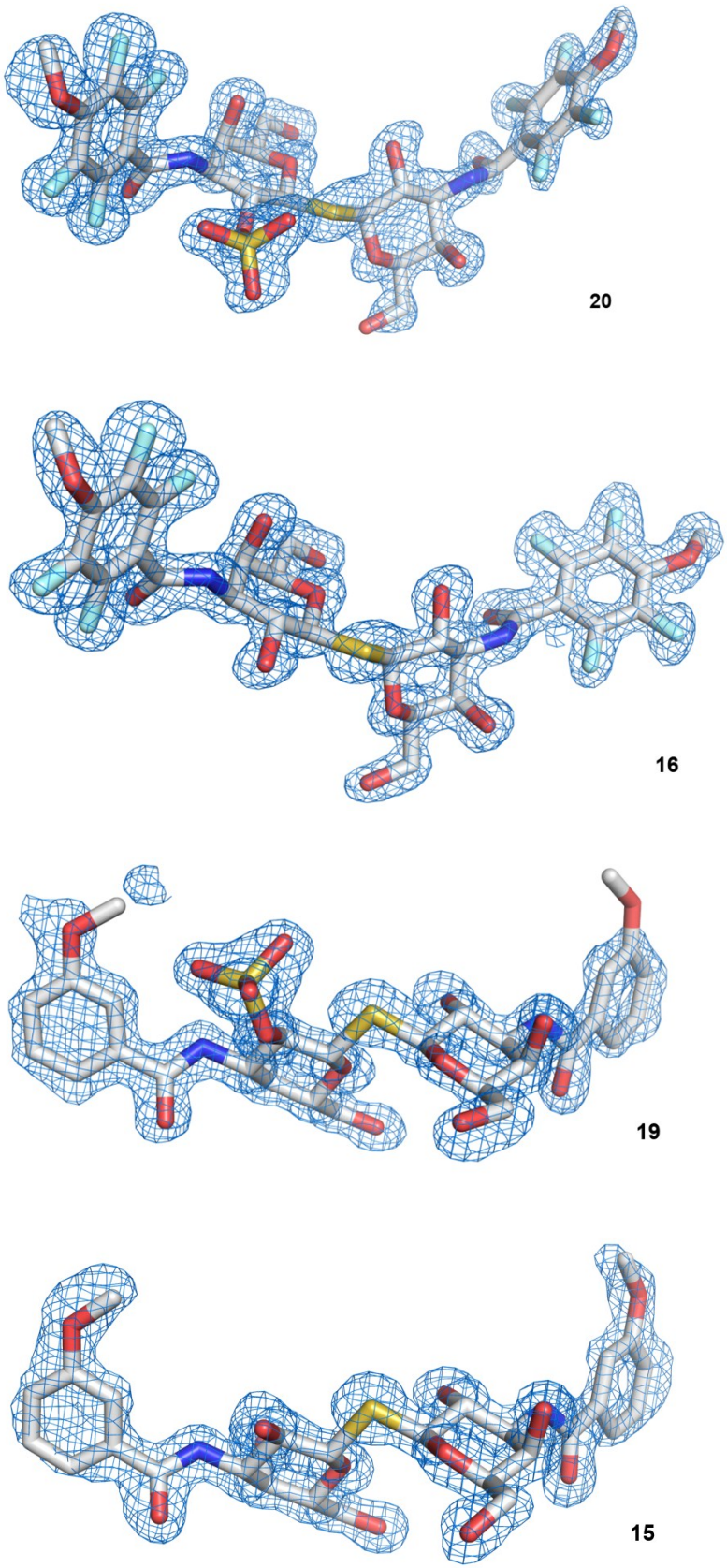


Figure S1. $2F_o$-F_c electron density maps contoured at 1.1 σ for ligands **15**, **16**, **19**, and **20** when in complex with galectin-3C.

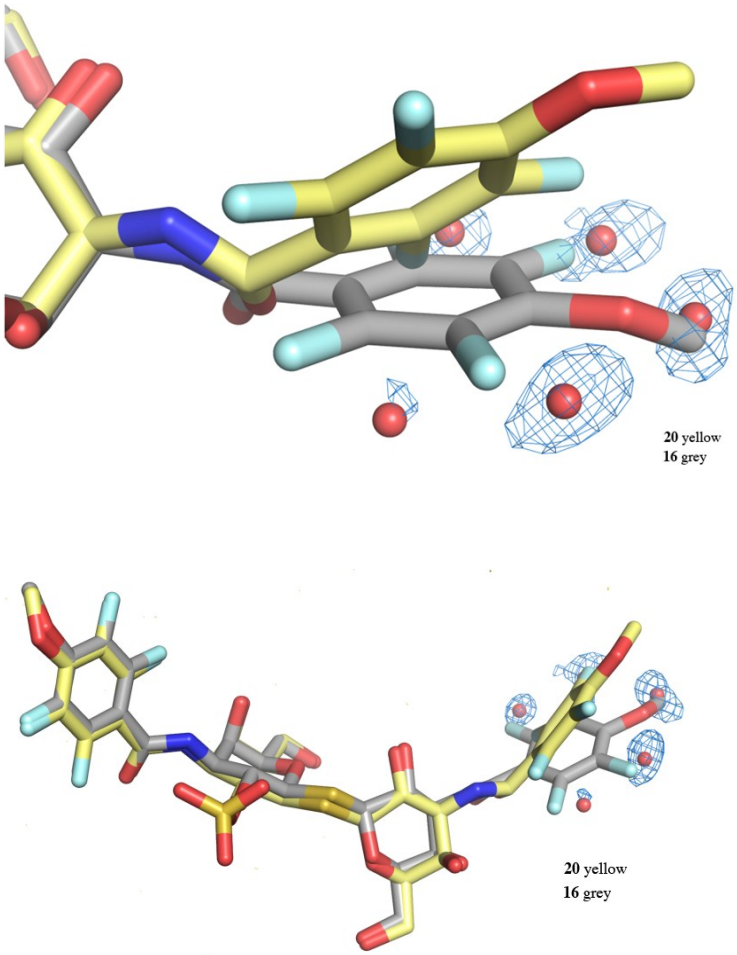


Figure S2. Electron densities (blue grid with red spheres) for the possible alternate positions of **20** (yellow) and its position compared to **16** (grey).

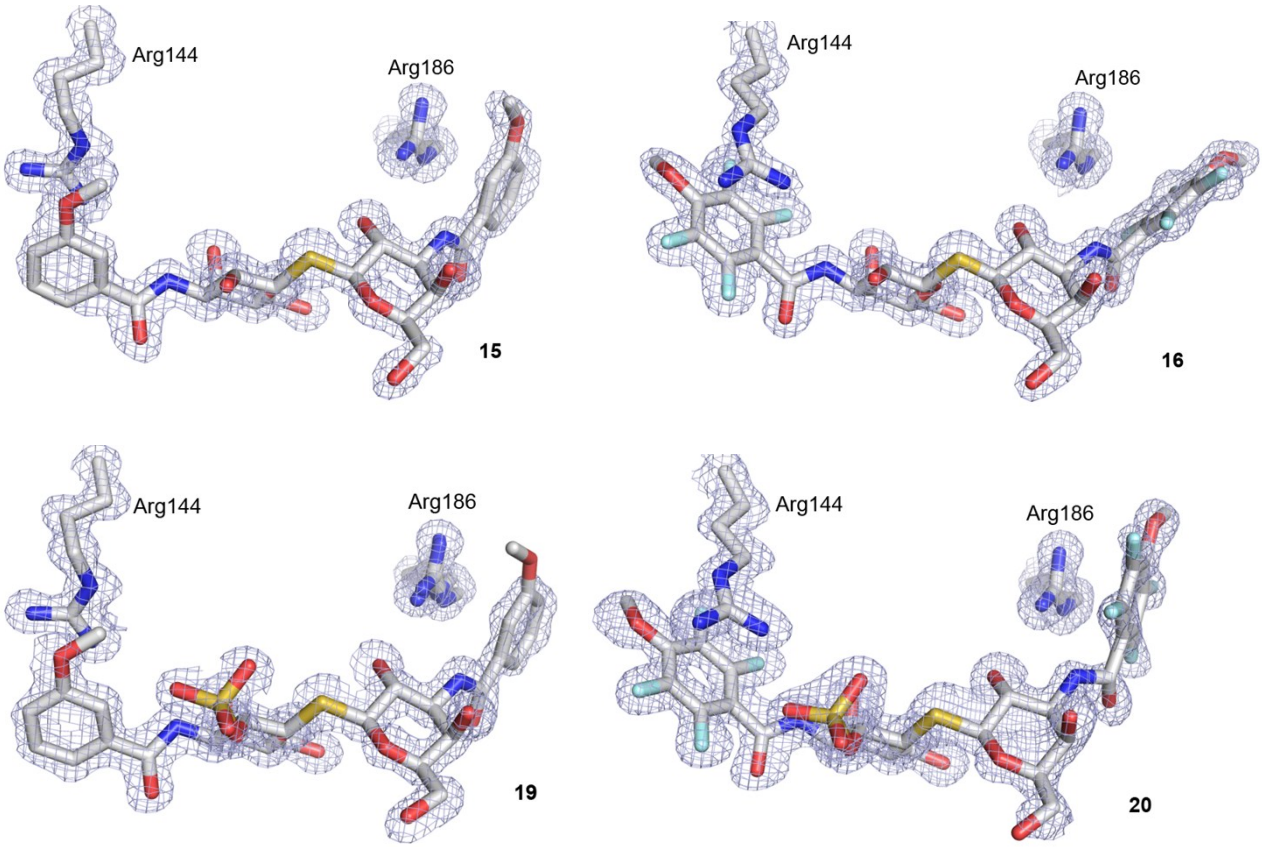
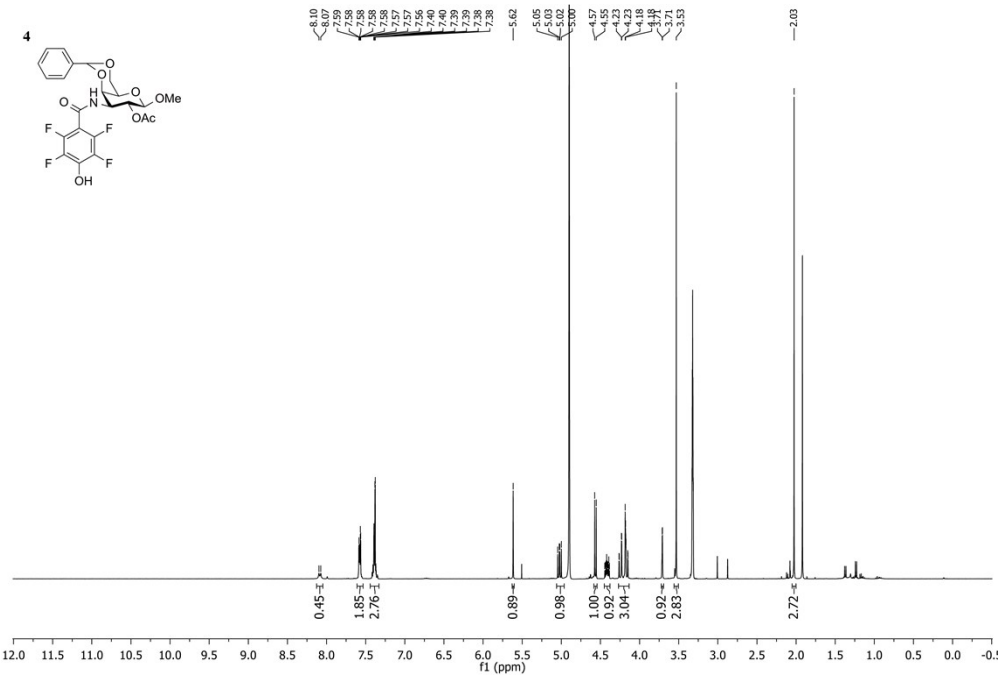


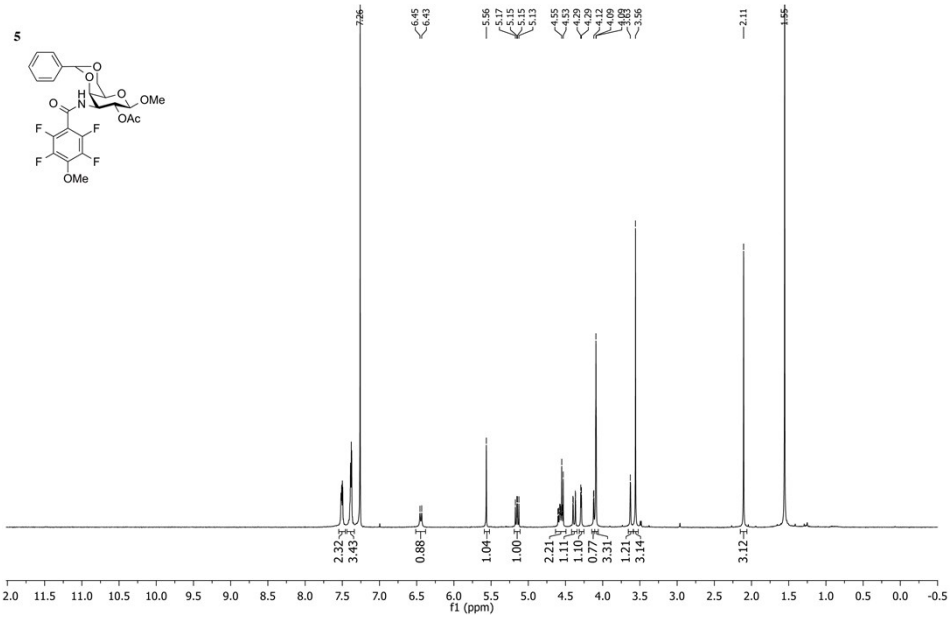
Figure S3. $2F_o$-F_c electron density maps contoured at 1.1 σ showing the galectin-3 complex and R144 and R186 interactions with **15**, **16**, **19** and **20**.

^{1}H nmr spectra

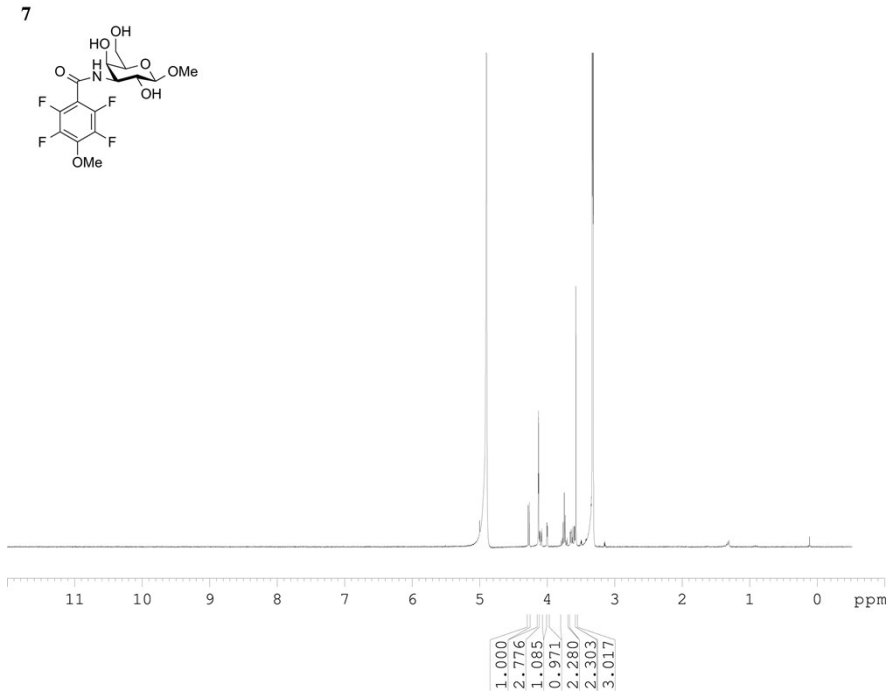
Methyl 2-*O*-acetyl-3-(4-hydroxy-2,3,5,6-tetrafluoro-benzamido)-4,6-*O*-benzylidene-3-deoxy-β-D-galactopyranoside 4.



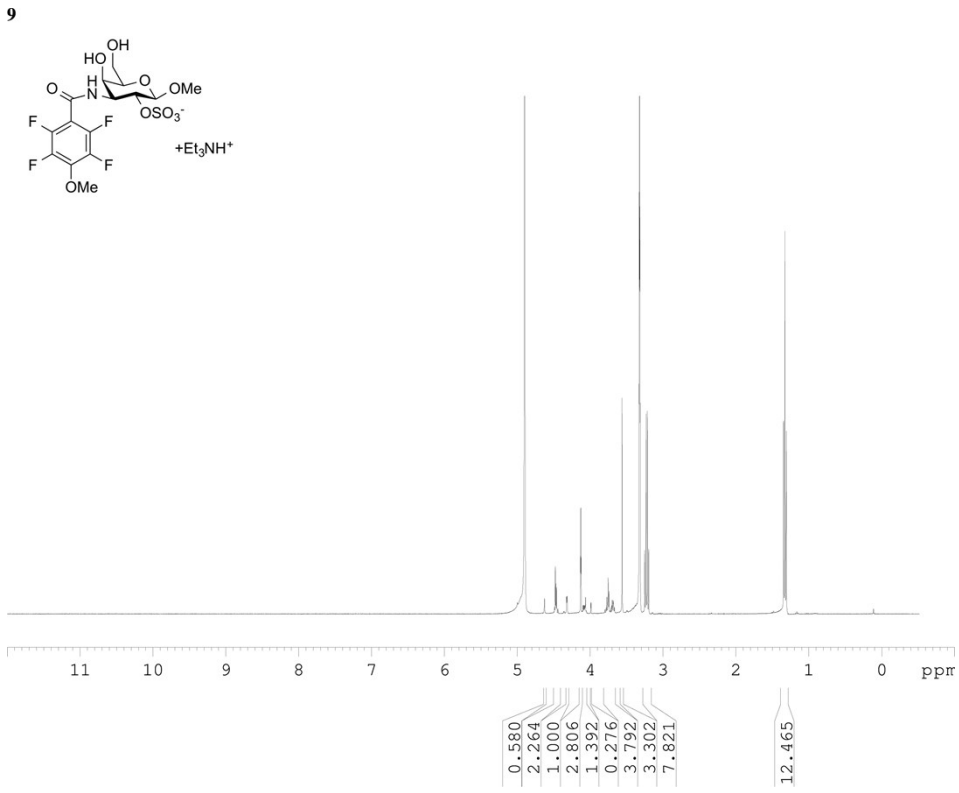
Methyl 2-*O*-acetyl-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-4,6-*O*-benzylidene-3-deoxy-β-D-galactopyranoside 5.



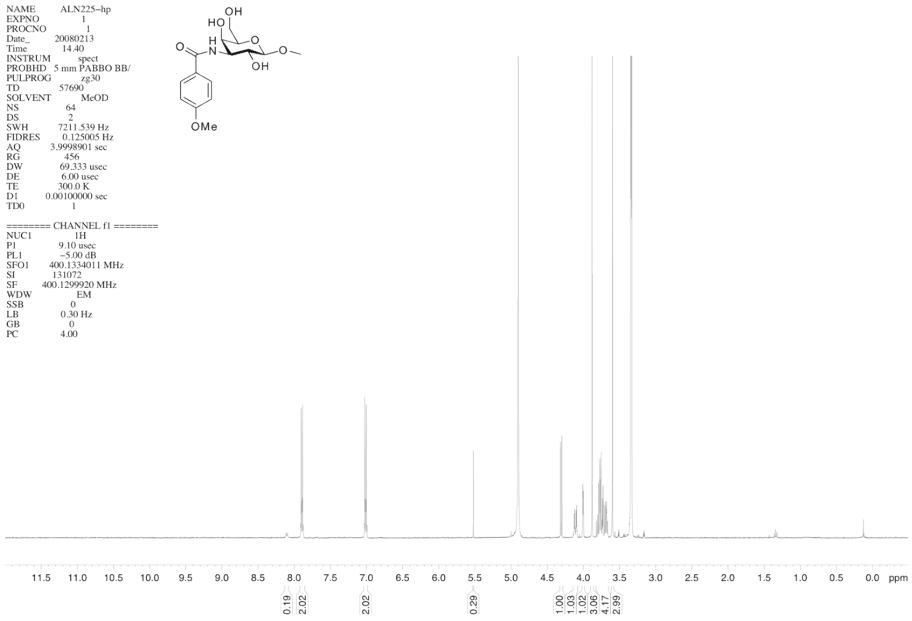
Methyl 3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-β-D-galactopyranoside 7.



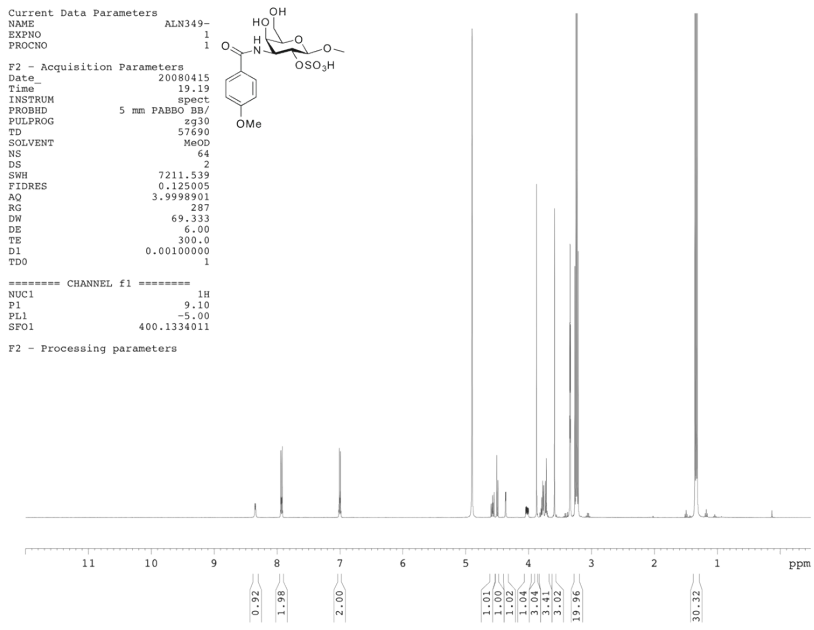
Methyl 3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-2-*O*-sulfo-β-D-galactopyranoside 9.



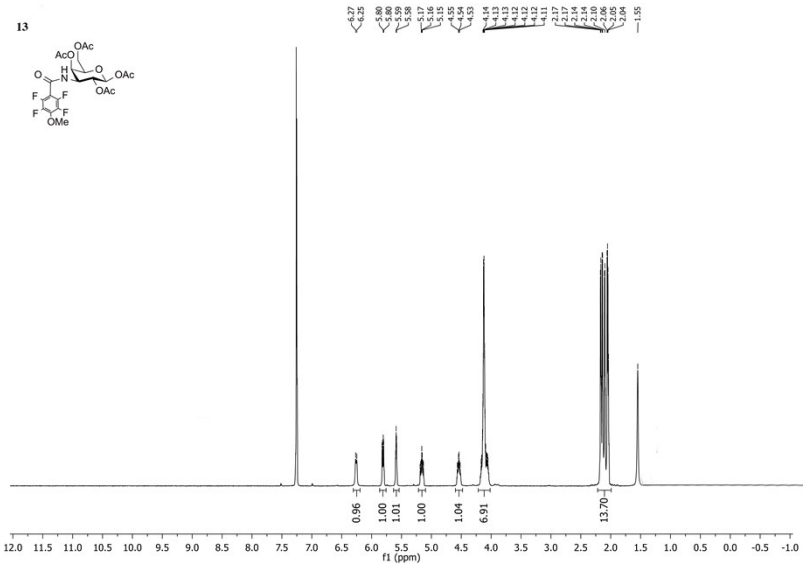
Methyl 3-deoxy-3-(4-methoxybenzamido)-β-D-galactopyranoside 10.



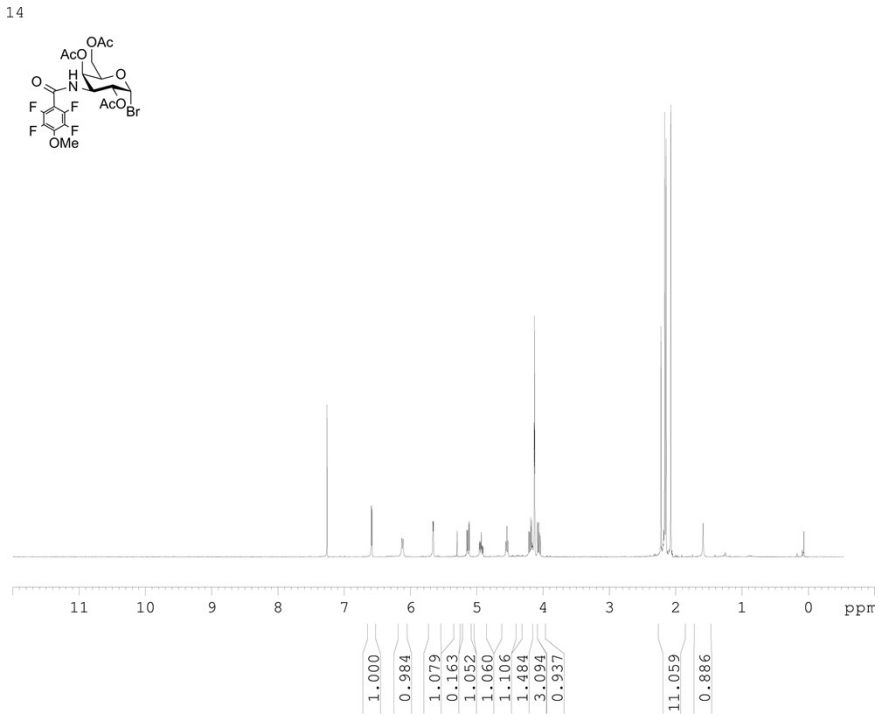
Methyl 3-deoxy-3-(4-methoxybenzamido)-2-*O*-sulfo-β-D-galactopyranoside 11.



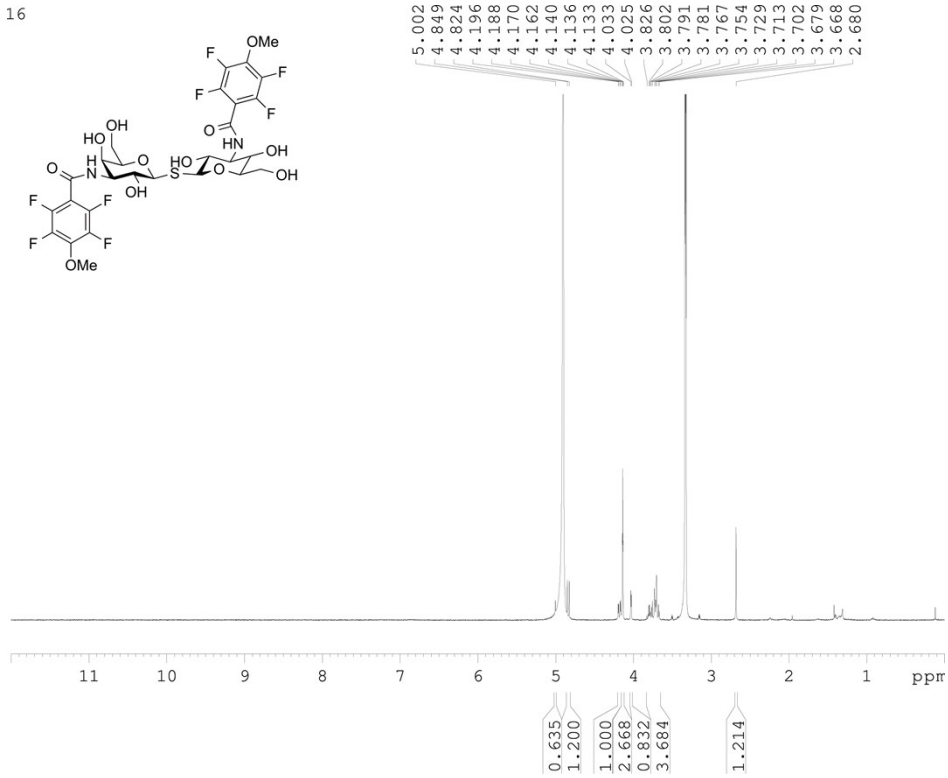
1,2,4,6-Tetra-O-acetyl-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-β-D-galactopyranose 13.



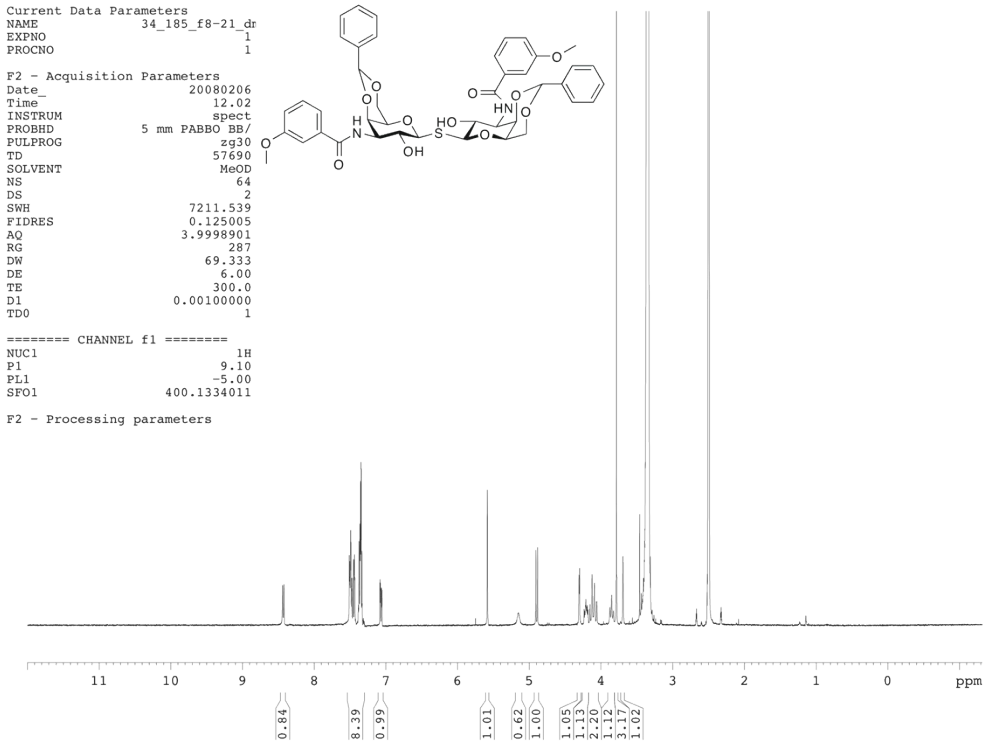
2,4,6-Tri-O-acetyl-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-β-D-galactopyranosyl bromide 14.



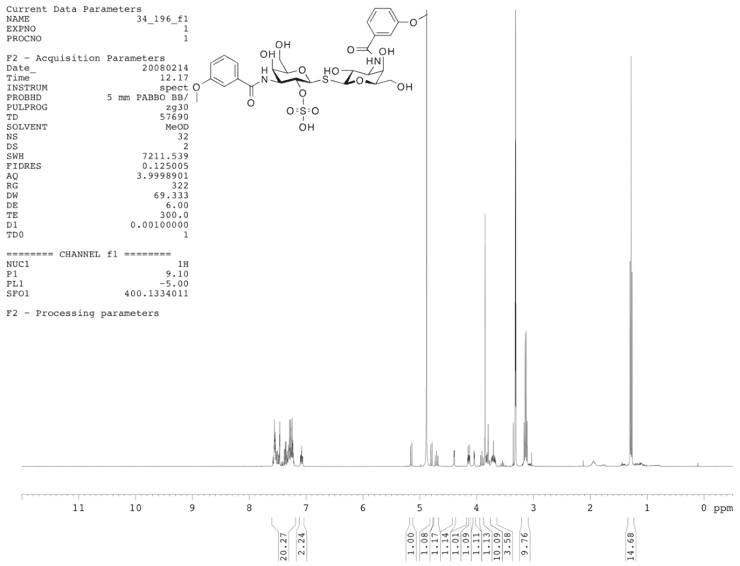
Di-(3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-β-D-galactopyranosyl)sulfane 16.



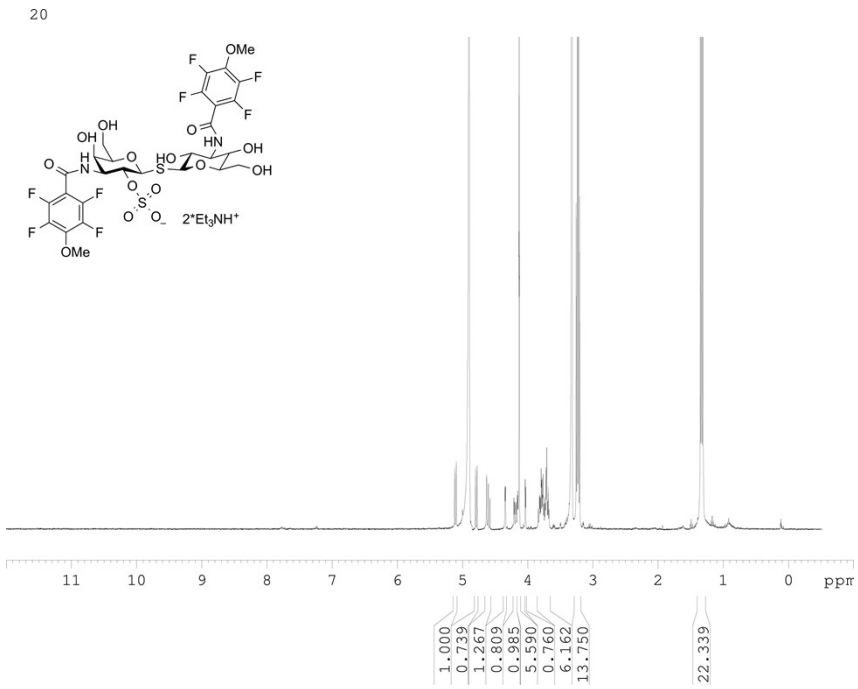
Di-[4,6-*O*-benzylidene -3-deoxy-3-(3-methoxybenzamido)-β-D-galactopyranosyl]sulfane 17.



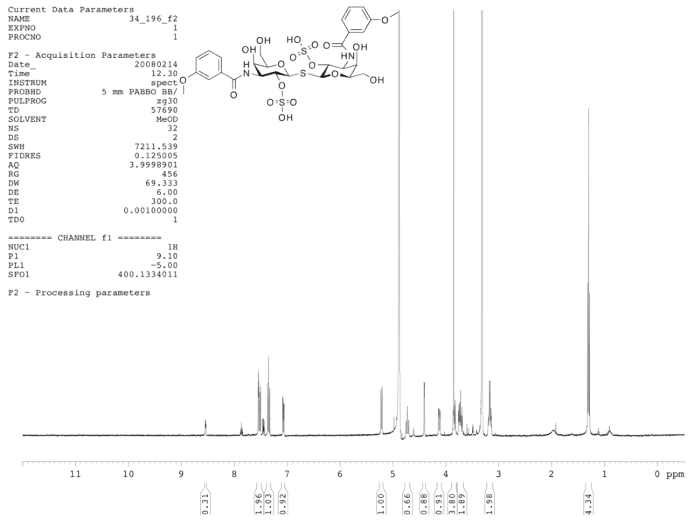
(3-Deoxy-3-methoxybenzamido-β-D-galactopyranosyl)-(3-deoxy-3-methoxybenzamido-2-sulfo-β-D-galactopyranosyl)sulfane 19.



(3-Deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro-benzamido)-β-D-galactopyranosyl)-(3-deoxy-3-(4-methoxy-2,3,5,6-tetrafluoro -benzamido)-2-sulfo-β-D-galactopyranosyl)sulfane 20.



Di-(3-deoxy-3-methoxybenzamido-2-sulfo-β-D-galactopyranosyl)sulfane 21.



Methyl β-D-galactopyranosyl-(1→4)-2-*O*-(4-methoxy-2,3,5,6-tetrafluorobenzoyl)-β-D-glucopyranoside 22.