

Expanding the Thioglycoligase Strategy to the Synthesis of α -linked Thioglycosides Allows Structural Investigation of the Parent Enzyme/Substrate Complex

Young-Wan Kim,[†] Andrew L. Lovering,[‡] Hongming Chen,[†] Terrence Kantner,^{†,‡}
Lawrence P. McIntosh,[‡] Natalie C. J. Strynadka,[‡] Stephen G. Withers^{*,†}

[†] Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia V6T 1Z1, Canada

[‡] Department of Biochemistry and Molecular Biology, University of British Columbia, 2146 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3, Canada

Contents

- 1. Experimental Procedures and Data**
- 2. Additional Figures**
- 3. Complete References (7) and (8)**

1. Experimental procedures and Data

Materials.

All chemicals and buffer salts were obtained from the Sigma Chemical Co. unless otherwise specified. *Pwo* polymerase was purchased from Roche and restriction enzymes from Fermentas (Germany). α -D-Xylopyranosyl fluoride (α XylF), α -D-glucopyranosyl fluoride (α GlcF), 4-nitrophenyl 4-deoxy-4-thio- β -D-glucopyranoside (pNP β Glc4S), and 4-nitrophenyl 4-deoxy-4-thio- β -D-xylopyranoside (pNP β Xyl4S) were synthesized according to literature procedures ((a) Hayashi, M.; Hashimoto, S.; Noyori, R. *Chem. Lett.* **1984**, *10*, 1747-1750. (b) Jahn, M.; Marles, J.; Warren, R. A. J.; Withers, S. G. *Angew. Chem. Int. Ed.* **2003**, *42*, 352-354). The synthetic details of 4-nitrophenyl 6-deoxy-6-thio- β -D-glucopyranoside (pNP β Glc6S) will be published elsewhere.

General analysis.

All ^1H and ^{13}C nuclear magnetic resonance (NMR) spectra were recorded at 400 MHz using a Bruker AV-400 spectrometer. Mass spectrometry for small molecules was recorded using a PE-Sciex API 300 triple quadrupole mass spectrometer (Sciex, Thornhill, ON, Canada) equipped with an electrospray ionization ion source. Thin layer chromatography (TLC) was performed on aluminum-backed sheets of silica gel 60F₂₅₄ (Merck) of thickness 0.2 mm. The plates were visualized using UV light (254 nm) and/or by exposure to 10 % sulfuric acid in methanol followed by charring. The SEP-PAK Cartridge for solid phase extraction of C18 was purchased from Waters Corporation (Milford, MA.). Silica flash column chromatography was carried out using Silica Gel 60 (230 \pm 400 mesh).

Construction of YicI D482A and MalA D416A mutants.

The alanine mutations of the acid/base residues of YicI and MalA were introduced using a mega primer PCR method. Plasmid pET29EcoXyl31A¹ was used as template for the mutation of YicI. The gene (*malA*) encoding for α -glucosidase from *S. solfataricus* (MalA) was amplified using the genomic DNA of *S. solfataricus* as template and two primers (MalA-TOP-Nde, 5'-TGTATACGGTGATAACATATGCAGACAATA-3' and

MalA-END-Xho, 5'-ACTACTCGAGGTTAATCTTTCCCCTAAT-3') designed based on the *S. solfataricus* genome sequence (GenBank NC_002754). The amplified PCR fragment was digested with NdeI and XhoI, followed by subcloning into pET29b (Novagen). The resulting plasmid was designated pET29MalA(His)₆ and used for the further mutagenesis of MalA. First the rear parts of the gene fragment for YicI D482A and MalA D416A were amplified with T7 terminator primer and the respective mutation primers (YicI-D482A-fw (5'-GTACACTGGGGTGGCGCGTGTTACGCTAACTAC-3') and MalA-D416-fw (5'-GCATTCATCTGGACTGGTGCGAATACCCCT-3'), respectively). The PCR products were purified using a QIAquick Gel Extraction Kit (Qiagen) and then used as megaprimers to obtain full size genes with the T7 promoter primer (5'-TAATACGACTCACTATAGGG-3'). The full-size PCR products were digested with NdeI and NotI, then extracted from agarose gel using a QIAquick Gel Extraction Kit and ligated with pTKNd vector (Kim, Y. W.; Lee, S. S.; Warren, R. A. J.; Withers, S. G. *J. Biol. Chem.* **2004**, *279*, 31033-31040) that had been digested with the corresponding restriction enzymes. The resulting plasmids, designated pBLYicI-D482A(His)₆ and pBLMalA-D416A(His)₆, were transformed into *E. coli*, and cultured in Luria-Bertani media containing 20 µg/mL of kanamycin, followed by purification using nickel-affinity chromatography as described previously.¹ In the case of MalA D416A the protein purified using nickel-affinity chromatography was heat-treated at 70 °C for 20 min to eliminate heat-labile *E. coli* proteins.

Inhibition Kinetics for YicI.

Inhibition studies were performed by measuring enzyme activity in the presence of various concentrations of each thioglycoside, using pNP α-D-xylopyranoside (pNPαX) as substrate at 30 °C in pH 7.0, 50 mM phosphate buffer. Fifty microliter of wild type YicI, preincubated at 30 °C, was added to 50 µL of buffer solution containing a fixed concentration of pNPαX and varying amounts of inhibitors, also preincubated at the corresponding temperature. The release of 4-nitrophenol was monitored at 400 nm using a plate reader (SPECTRAMax plus, Molecular Devices Corporation). The experiments were repeated at different concentrations of pNPαX. A Dixon plot of 1/rate versus inhibitor concentration for each substrate concentration intersects a line given by $1/V_{\max}$ at an

inhibitor concentration equal to $-K_i$.

Transglycosylation and isolation.

The thioglycoligase and transglycosylation reactions were carried out at room temperature in 100 mM sodium phosphate buffer, pH 7.0. Reactions were monitored by thin layer chromatography (TLC). Upon completion, the reaction mixtures were subjected to a C18 SEP PAK cartridge (Waters) to remove non-aryl sugars, enzyme, and salts. The cartridge was washed with 6 mL of water and then the pNP-sugars were eluted with 4 mL of 50 % (v/v) methanol.

4-Nitrophenyl α -D-xylopyranosyl-(1,4)-S- β -D-glucopyranoside (1) and its per-O-acetylated derivative. A mixture of the α XylF (15.5 mg 101.9 μ mol) and pNP β Glc4S (32.1 mg, 101 μ mol) in phosphate buffer (5 mL of 0.2 M, pH 7.0) was treated with YicI D482A (12 mg) and the mixture then incubated (30 h, 25 °C). The aryl glycoside products were purified using a C18 SEP PAK cartridge, then the solvent was evaporated under reduced pressure. Flash chromatography (EtOAc/MeOH/H₂O = 17.5: 2: 0.5) of the residue gave the disaccharide **1** (36 mg, 79 %). ¹H NMR (CD₃OD, 400 MHz): δ 8.21 (m, 2H, Ar-H), 7.23 (m, 2H, Ar-H), 5.57, (d, 1 H, $J_{1',2'}$ 4.8 Hz, H-1'), 5.09 (d, 1 H, $J_{1,2}$ 7.8 Hz, H-1), 4.06, (dd, 1 H, $J_{5,6a}$ 2.0 Hz, $J_{6a,6b}$ 12.1 Hz H-6a), 3.86 (d, 1 H, $J_{5,6b}$ 5.4 Hz, $J_{6a,6b}$ 12.1 Hz, H-6b), 3.69 (4 H, H-3, -2', -5, -5'a), 3.80 (m, 1 H, H-4'), 3.50 (3 H, H-2, -3', -5'b), 2.83 (t, 1 H, $J_{3,4} = J_{4,5}$ 10.7 Hz, H-4). ¹³C NMR (CD₃OD, 100 MHz): 163.99, 144.03, 126.75 (2C), 117.87 (2C) (Ar-C); 101.47, 88.68, 78.34, 77.90, 75.71, 75.06, 73.63, 71.19, 65.43, 63.36. EISMS: calc. for C₁₇H₂₃NO₁₁S + Na⁺ = 472.1; Found 472.1.

Pure **1** (26 mg, 58 μ mol) was acetylated in pyridine (6 mL) and Ac₂O (4 mL) overnight at RT. This mixture was quenched by the addition of MeOH (6 mL) at 0 °C and concentrated under reduced pressure. The residue was chromatographed on Silica Gel using (EtOAc/hexanes = 1: 9 \rightarrow 6: 4) to give the per-O-acetylated derivative of **1** (35.5 mg, 87 %). ¹H NMR (CDCl₃, 400 MHz): δ 8.21 (m, 2-H, Ar-H), 7.06 (m, 2 H, Ar-H), 5.82 (d, 1 H, $J_{1',2'}$ 5.7 Hz, H-1'), 5.34 (dd, 1 H, $J_{2,3}$ 8.9 Hz, $J_{3,4}$ 10.7 Hz, H-3), 5.24 (t, 1 H, J 10.0 Hz, H-3'), 5.16 (d, 1 H, $J_{1,2}$ 7.3 Hz, H-1), 5.13 (obs., H-2), 4.93 (m, 1 H, H-4'), 4.90 (dd, 1 H, $J_{1',2'}$ 5.7 Hz, $J_{2',3'}$ 10.0 Hz, H-2'), 4.56 (dd, 1 H, $J_{4',5'a}$ 2.1 Hz, $J_{5'a,5'b}$ 12.1 Hz, H-5'a), 4.35

(dd, 1 H, $J_{4',5'} 5.5$ Hz, $J_{5',6'} 12.1$ Hz, H-5'b), 3.85 (ddd, 1 H, J 2.1, 5.5, 11.2 Hz, H-5), 3.81 (s, 1 H, H-6a), 3.79 (s, 1 H, H-6b), 3.16 (t, 1 H, J 10.9 Hz, H-4), 2.11, 2.08, 2.05 (6 x CH_3CO). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 170.58, 170.30, 170.27, 170.04, 169.89, 169.70 (CH_3CO); 161.41, 143.42, 125.97 (2C), 116.83 (2C) (Ar-C); 98.00 (C-1), 84.5 (C-1'), 75.38 (C-3), 73.59 (C-5), 72.48 (C-2), 70.81 (C-2'), 69.28 (C-3'), 68.78 (C-4'), 64.0 (C-5'), 60.0 (C-6), 44.80 (C-4); 21.00, 20.93, 20.91, 20.89, 20.81, 20.73, (CH_3CO). EISMS: calc. for $C_{29}H_{35}NO_{17}S + Na^+ = 724.1$; Found 724.1.

4-Nitrophenyl α -D-glucopyranosyl-(1, 4)-S-4-deoxy-4-thio- β -D-glucopyranoside (2) and its per-O-acetylated derivative. A mixture of α GlcF (17 mg, 93.4 μ mol) and pNP β Glc4S (28 mg, 91.5 μ mol) in phosphate buffer (15 mL of 0.1 M) was added MalA D416A (8 mg) and the mixture then incubated (72 h, RT). After C18 purification as for **1** above, flash chromatography ($CH_2Cl_2/MeOH = 4:1 \rightarrow 1:1$) of the residue gave, the disaccharide **2** (21 mg, 48 %). 1H NMR (CD_3OD , 400 MHz): δ 8.15 (m, 2 H, Ar-H), 7.16 (m, 2 H, Ar-H), 5.65 (d, 1 H, $J_{1',2'}$ 5.6 Hz, H-1'), 5.03 (d, 1 H, $J_{1,2}$ 8.0 Hz, H-1), 4.00 ~ 3.10 (m, 11 H, H-2~3, H-5 ~ 6 & H-2' ~ 6'), 2.88 (t, 1 H, $J_{3,4} = J_{4,5}$ 10.6 Hz, H-4). ^{13}C NMR (CD_3OD , 100 MHz): δ 163.99, 144.01, 126.75 (2 C), 117.86 (2 C) (Ar-C); 101.46, 88.13, 78.38, 77.58, 75.79, 75.67, 75.30, 73.43, 71.98, 63.07, 62.89, 48.45. EISMS: Calcd for $C_{18}H_{25}NO_{12}S + Na^+ = 502.1$; Found: 502.1.

After acetylation of the pure compound **2** (10 mg, 21 μ mol) as for **1** above, the solution was concentrated under diminished pressure, coevaporated with toluene, and the residue was dissolved in CH_2Cl_2 (20 mL). The organic phase was successively washed with 1 N HCl (2 \times 20 mL), water (20 mL) and brine (20 mL), and then dried. The solvent was evaporated and crystals of the per-O-acetylated derivative of **2** (9 mg, 43 %) were obtained from methanol-diethyl ether. 1H NMR ($CDCl_3$, 400 MHz): δ 8.21 (m, 2 H, Ar-H), 7.06 (m, 2 H, Ar-H), 5.95 (d, 1 H, $J_{1',2'}$ 5.6 Hz, H-1'), 5.36 (dd, 1 H, $J_{2,3}$ 10.4 Hz, $J_{3,4}$ 9.2 Hz, H-3), 5.28 (t, 1 H, $J_{3',4'}$ 10.0 Hz, H-3'), 5.17 (d, 1 H, $J_{1,2}$ 8.0 Hz, H-1), 5.13 (dd, 1 H, H-2), 5.07 (t, 1 H, $J_{4',5'}$ 9.6 Hz, H-4'), 4.98 (dd, 1 H, $J_{2',3'}$ 10.4 Hz, H-2'), 4.69 (dd, 1 H, $J_{5,6a}$ 2.0 Hz, $J_{6a,6b}$ 11.6 Hz, H-6a), 4.34 (dd, 1 H, $J_{5',6'a}$ 4.4 Hz, $J_{6'a,6'b}$ 12.4 Hz, H-6'a), 4.26 (m, 1 H, H-5'), 4.23 (dd, 1 H, $J_{5,6b}$ 7.2 Hz, H-6b), 4.12 (dd, 1 H, $J_{5,6'b}$ 2.0 Hz, H-6'b), 3.88 (m, 1 H, H-5), 3.10 (t, 1 H, $J_{4,5}$ 10.8 Hz, H-4), 2.11 (s, 3 H, CH_3CO), 2.108 (s, 3 H, CH_3CO), 2.10 (s, 3 H,

CH_3CO), 2.05 (s, 9 H, 3 x CH_3CO), 2.03 (s, 3 H, CH_3CO). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 170.70, 170.48, 170.35, 170.24, 170.05, 169.67 (2 C) (CH_3CO); 161.42, 143.46, 125.95 (2 C), 116.86 (2 C) (Ar-C); 98.00 (C-1), 82.72 (C-1'), 75.64 (C-3), 73.39 (C-5), 72.51 (C-2), 70.66 (C-2'), 69.92 (C-3'), 68.97 (C-5'), 68.20 (C-4'), 63.83 (C-6), 61.84 (C-6'), 43.63 (C-4); 21.02, 20.98, 20.91, 20.82, 20.80, 20.78, 20.69 (CH_3CO). EISMS: Calcd for $C_{32}H_{39}NO_{19}S + Na^+ = 796.2$; Found: 796.2.

4-Nitrophenyl α -D-xylopyranosyl-(1,4)-S- β -D-xylopyranoside (3) and its per-O-acetylated derivative. To a mixture of α XylIF (15 mg, 0.1 mmol) and pNP β Xyl4S (22 mg, 0.076 mmol) in phosphate buffer (5 mL of 0.2 M, pH 7.0) was added YicI D482A (20 mg) and the mixture then incubated (30 h, RT). Aryl-glycosides were purified using a C18 SEP PAK cartridge (Waters) and then evaporated under reduced pressure. Flash chromatography (EtOAc/MeOH/H₂O = 17.5: 2: 0.5) of the residue gave the disaccharide **3** (13 mg, 41 %). 1H NMR (CD_3OD , 400 MHz): δ 8.21 (m, 2 H, Ar-H), 7.19 (m, 2 H, Ar-H), 5.45 (d, 1 H, $J_{1',2'}$ 4.8 Hz, H-1'), 5.02 (d, 1 H, $J_{1,2}$ 6.0 Hz, H-1), 4.07 (dd, 1 H, $J_{4,5a}$ 5.2 Hz, $J_{5a,5b}$ 11.9 Hz, H-5a), 3.83 (m, 1 H, H-4'), 3.66 (H-5'), 3.64 (H-2'), 3.57 (t, 1H, J 11.9 Hz, H-5b), 3.49 (4 H, H-2, -3, -3', -5'), 2.89 (m, H-4). ^{13}C NMR (CD_3OD , 100 MHz): 163.73, 143.93, 126.62 (2C), 117.63 (2C) (Ar-C); 101.99, 89.02, 76.74, 75.95, 74.92, 73.36, 71.13, 67.79, 64.54. EISMS: calc. for $C_{26}H_{31}NO_{15}S + Na^+ = 652.1$; Found 652.1.

After acetylation of the pure compound **3** (10 mg, 23.8 μ mol) as for **1** above, the solution was concentrated under diminished pressure, coevaporated with toluene, and the residue was dissolved in CH_2Cl_2 (20 mL). The organic phase was successively washed with 1 N HCl (2 \times 20 mL), water (20 mL) and brine (20 mL), and then dried. The solvent was evaporated and flash chromatography (EtOAc/hexanes = 1: 9 \rightarrow 6: 4) of the residue gave the per-O-acetylated derivative of **3** (12.4 mg, 83 %). 1H NMR ($CDCl_3$, 400 MHz): δ 8.21 (m, 2 H, Ar-H), 7.06 (m, 2 H, Ar-H), 5.75 (d, 1 H, $J_{1',2'}$ 5.4 Hz, H-1'), 5.27 (t, 1 H, J 9.3 Hz H-3'), 5.17 (obsc, H-3), 5.15 (obsc, H-1), 5.13 (obsc, H-2), 4.93 (m, 1 H, H-4'), 4.89 (dd, 1 H, $J_{1',2'}$ 5.4 Hz, $J_{2',3'}$ 9.7 Hz, H-2'), 4.19 (dd, 1 H, $J_{4,5a}$ 5.1 Hz, $J_{5a,5b}$ 12.1 Hz, H-5a), 3.94 (t, 1 H, J 11.1 Hz, H-5a'), 3.85 (dd, 1 H, $J_{4',5'b}$ 5.5 Hz, $J_{5'a,5'b}$ 11.1 Hz, H-5'b), 3.55 (t, 1 H, J 11.5 Hz, H-5b), 3.16 (m, 1 H, H-4), 2.1, 2.06 (5 x CH_3CO). ^{13}C NMR ($CDCl_3$, 100 MHz): δ 169.96, 169.88, 169.83, 169.65, 169.44, (CH₃CO); 161.16, 143.1, 125.81 (2C), 116.54 (2C)

(Ar-C); 98.45 (C-1), 83.46 (C-1'), 73.14 (C-2), 71.85 (C-3), 70.69 (C-2'), 68.98 (C-3'), 68.65 (C-4'), 65.90 (C-5), 59.54 (C-5'), 43.38 (C-4); 20.70 (2C), 20.63, 20.59 (2C), (CH₃CO). EISMS: calc. for C₂₆H₃₁NO₁₅S + Na⁺ = 652.1; Found 652.1.

4-Nitrophenyl α -D-xylopyranosyl-(1,6)-S-6-deoxy-6-thio- β -D-glucopyranoside (4) and its per-O-acetylated derivative. A mixture of α XylF (22.8 mg, 150 μ mol) and pNP β Glc6S (47.5 mg, 150 μ mol) in phosphate buffer (7.5 mL of 0.1 M) was treated with YicI D482A (12 mg) and the mixture then incubated (18 h, RT). After acetylation as for **1** above, flash chromatography (EtOAc/petrol = 3: 4) of the residue gave the acetylated derivative of **4** (90 mg, 87%). ¹H NMR (CDCl₃, 400 MHz): δ 8.22 (m, 2 H, Ar-H), 7.10 (m, 2 H, Ar-H), 5.53 (d, 1 H, J_{1',2'} 5.4 Hz, H-1'), 5.30 (t, 1 H, J_{3',4'} 9.4 Hz, H-3'), 5.25 (m, 2 H, H-2 & H-3), 5.14 (d, 1 H, J_{1,2} 7.5 Hz, H-1), 5.04 (t, 1 H, J_{3,4} = J_{4,5} 9.4 Hz, H-4), 4.91 (dd, 1 H, J_{2',3'} 9.7 Hz, H-2'), 4.86 (ddd, 1 H, J_{4',5'b} 5.6 Hz, H-4'), 3.85 (t, 1 H, J_{4',5'a} 10.4 Hz, H-5'a), 3.78 (m, 1 H, H-5), 3.65 (dd, 1 H, J_{5'a,5'b} 11.3 Hz, H-5'b), 2.81 (dd, 1 H, J_{6a,6b} 14.2 Hz & J_{5,6a} 2.6 Hz, H-6a), 2.62 (dd, 1 H, J_{5,6b} 7.8 Hz, H-6b); 2.09 (s, 3 H, CH₃CO), 2.04 (s, 3 H, CH₃CO), 2.03 (s, 3 H, CH₃CO), 2.013 (s, 3 H, CH₃CO), 2.01 (s, 3 H, CH₃CO), 1.94 (s, 3 H, CH₃CO). ¹³C NMR (CDCl₃, 100 MHz): δ 170.08, 169.85, 169.82, 169.59, 169.47, 169.14 (CH₃CO); 161.17, 143.31, 125.85 (2 C), 116.71 (2 C) (Ar-C); 98.22 (C-1), 81.96 (C-1'), 73.82 (C-5), 72.27 (C-2 or C-3), 70.94 (C-2 or C-3), 70.67 (2C, C-4 & C-2'), 69.22 (C-3'), 68.79 (C-4'), 59.41 (C-5'), 30.24 (C-6); 20.72, 20.65, 20.61, 20.55 (2C), 20.53 (CH₃CO). EISMS: Calcd for C₂₉H₃₅NO₁₇S + Na⁺ = 704.2; Found 704.2.

The per-O-acetylated **4** (65 mg, 92.5 μ mol) was dissolved in MeOH (20 mL) and the solution was cooled to 0 °C. NH₃ was bubbled through the solution for 10 min and the reaction mixture was stirred at room temperature for 1 h. The solvent was then evaporated and the residue was purified on Silica Gel using EtOAc/MeOH/water (12: 2: 1) as eluent. The eluate was concentrated and taken up in H₂O (3 mL). Freeze-drying of the eluate gave **4** (40 mg, 93 %) as a white solid. ¹H NMR (CD₃OD, 400 MHz): δ 8.22 (m, 2H, Ar-H), 7.30 (m, 2H, Ar-H), 5.29 (d, 1 H, J_{1',2'} 4.8 Hz, H-1'), 5.01 (d, 1 H, J_{1,2} 7.6 Hz, H-1), 3.87~3.28 (m, 9 H, H-2~5, H-2'~5'), 3.14 (dd, 1H, J_{6a,6b} 14.4 Hz & J_{6a,5} 2.0 Hz, H-6a), 2.62 (dd, 1H, J_{6b,5} 8.4 Hz, H-6b). ¹³C NMR (CD₃OD, 100 MHz): δ 163.82, 143.97, 126.65 (2 C), 118.04 (2 C) (Ar-C); 101.84 (C-1), 87.08 (C-1'), 77.75, 77.13, 75.00, 74.85, 74.06,

73.24, 71.33, 64.16, 31.99 (C-6). EISMS: Calcd for $C_{17}H_{23}NO_{11}S + Na^+ = 472.1$; Found 472.5.

Crystallization, Data Collection, Structure Determination and Refinement.

Crystals of YicI were obtained by following the previously reported protocol for the orthorhombic crystal form.¹ Crystals were cryoprotected and derivatised by transfer into a drop containing 100 mM MOPS pH 6.5, 2 M ammonium sulfate, 16 % (w/v) glycerol and 1 mM compound **4**, and left to stabilize for a period of 45 minutes. Cryoprotected crystals were then flash frozen at 100K and used for data collection at beamline 8.2.2. of the Advanced Light Source (ALS). Data collection and refinement statistics are provided in Table 1. The YicI:**4** ligand complex was solved using the isomorphous 1XSI co-ordinates,¹ and performing several cycles of rigid-body refinement to allow for small changes in unit cell dimensions. Map calculation showed clear density for compound **4** in all six active sites, and one further molecule located at a crystal packing interface. Topology and parameter files for **4** were obtained using the PRODRG program (Schuttelkopf, A. W.; van Aalten, D. M. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 1355-1363.). The model was refined using REFMAC (Murshudov, G. N. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53*, 240-255), and all other data manipulations were carried out using the CCP4 suite of programs. (Collaborative Computational Project *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50*, 760-763) The data presented here represent the highest resolution structure of YicI to date.

Table 1. Data collection and refinement statistics.

Crystal Parameters	
Spacegroup	P2 ₁ 2 ₁ 2 ₁
Unit Cell (Å)	162.3 x 175.8 x 210.7
Data Collection Statistics ^a	
Resolution (Å)	1.95 (2.06-1.95)
Wavelength (Å)	0.95516
No. reflections	1286463
No. unique reflections	375731
Redundancy	3.4 (3.0)
Completeness (%)	86.5 (88.9)
I/σI	8.2 (3.4)
R _{sym} ^b	9.7 (43.3)
Refinement Statistics	
ASU Contents	6x 772 AA, 7x compound 4, 14x glycerol, 9x SO ₄ , 5x MOPS, 2003x H ₂ O
Rwork (%)	18.0
Rfree ^c (%)	22.5
RMS bonds (Å)	0.018
RMS angles (°)	1.75
Average B / Wilson B (Å ²)	18.3 / 16.6

a Values in parentheses are for the outer shell.

b $R_{sym} = \frac{\sum(|I_{hkl} - \langle I \rangle|)}{\sum(I_{hkl})}$.

c 5% of reflections excluded from refinement.

2. Additional Figures

Figure legends

Fig S1. Comparison sugar rings between YicI:4 Michaelis complex (A) and glycosyl enzyme intermediate (B) including the nucleophiles Asp416 and the acid/base Asp482. The two structures are represented in bond-style format and colored as follows: Michaelis complex carbon atoms, yellow; intermediate carbon atoms, purple; compound 4 carbon atoms, beige; 5-fluoro xylose (eq5FX) carbon atoms, olive green; oxygen atoms, red; and fluorine atoms, grey.

Fig S2. Michaelis complex of YicI:4 including seventh molecule of 4, bound at N-terminal end of $\alpha_8\beta_8$ barrel / C-terminal domain interface of monomer F. Compound 4 shown in bond-style form, and colored according to atom type: C, purple (compound 4 in the active sites) or beige (the seventh molecule); O, red; N, blue; S, yellow. Each monomer of YicI is represented as stick-style colored in a different color.

Fig S3. Superposition of the Michaelis complex (2F2H) and the glycosyl enzyme intermediate (1XSK). The averaged root-mean-square deviation for backbone atoms between these structures is 0.37 \AA^2 , with larger shifts for three active site loops: 1.35 \AA^2 for amino acids 275-282, 1.44 \AA^2 for 513-520 and 1.32 \AA^2 for 539-547. Compound 4 shown in bond-style form, and colored according to atom type: C, beige; O, red; N, blue; S, yellow. C α -trace of the glycosyl enzyme intermediate (red) and the Michaelis complex (blue), with mobile loops 275-282, 513-520 and 539-547 shown in non-transparent form for emphasis.

Fig S1.

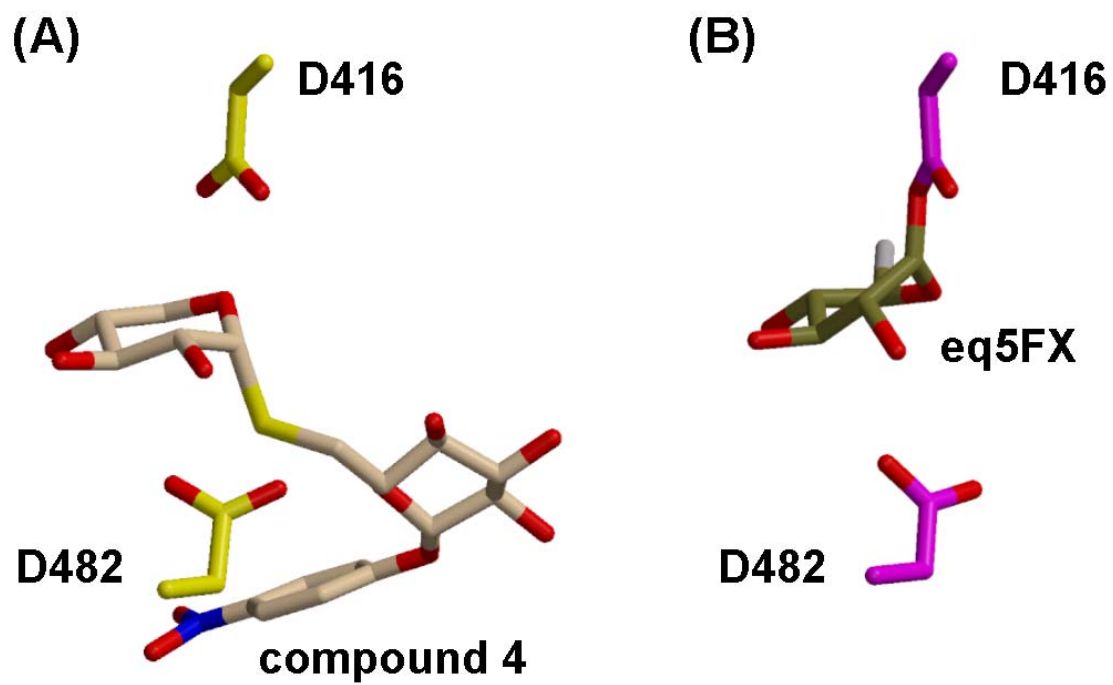


Fig S2.

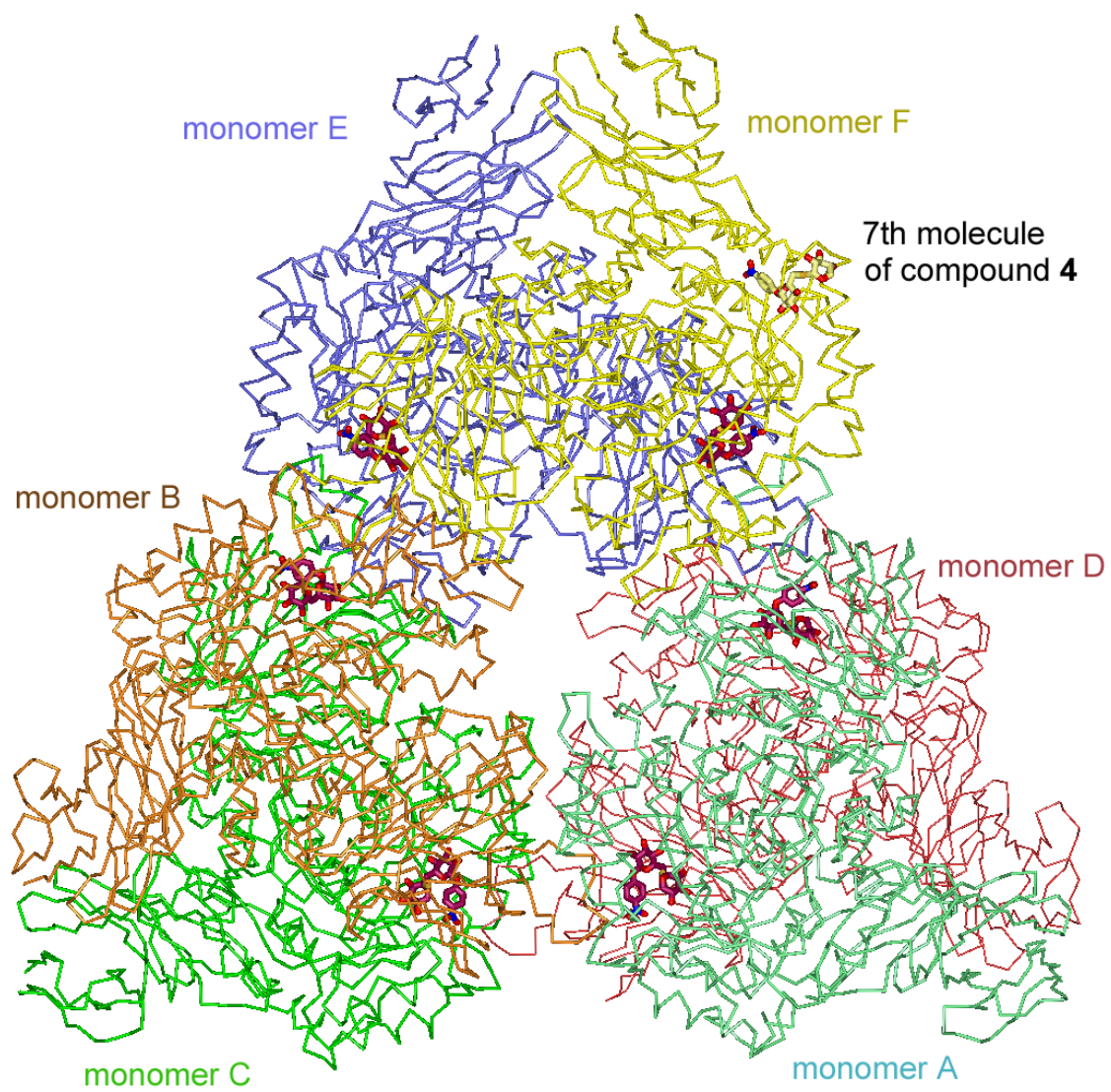
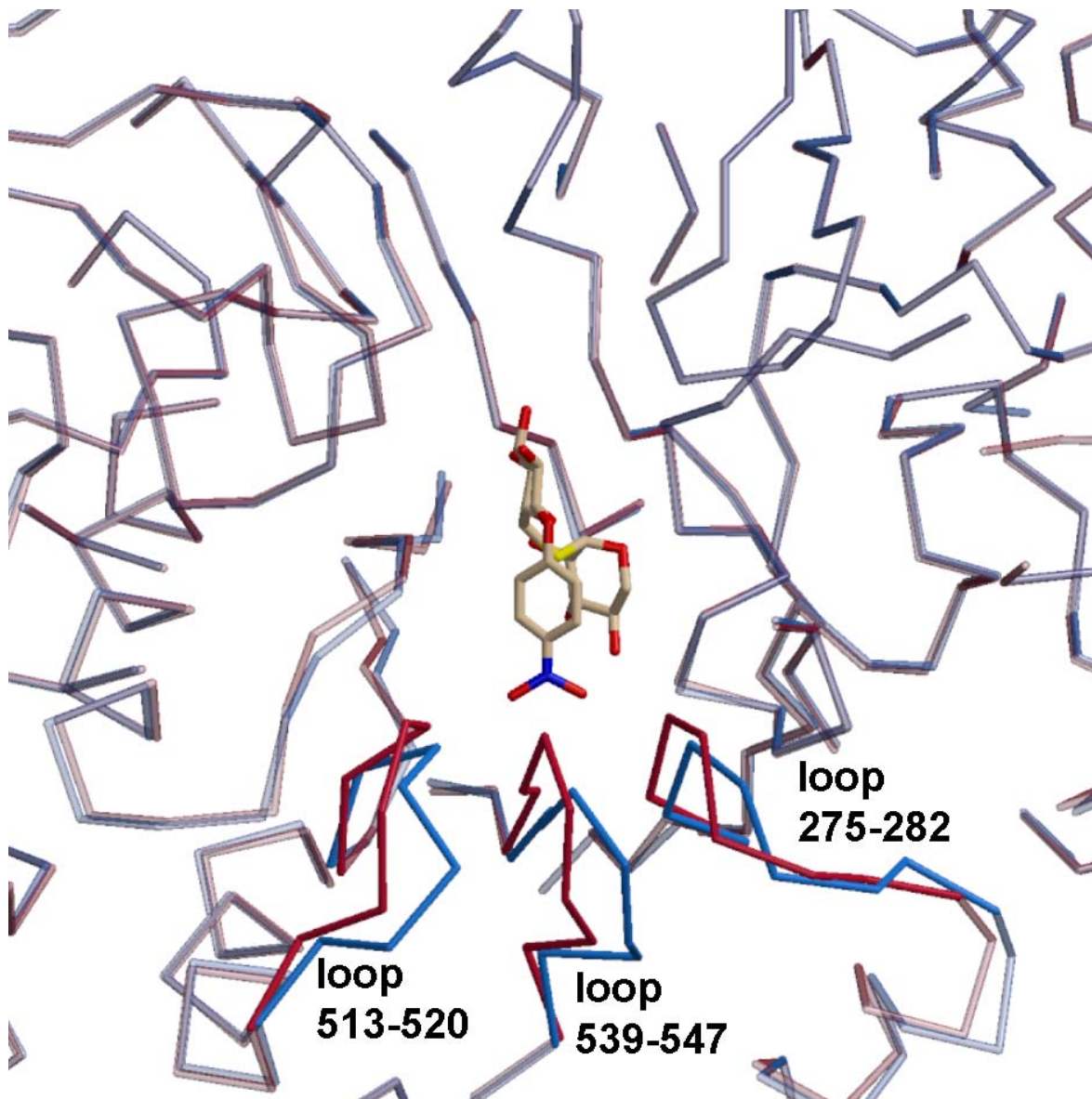


Fig S3.



3. Complete Reference (8)

(7) Moracci, M.; Cobucci-Ponzano, B.; Trincone, A.; Fusco, S.; De Rosa, M.; van Der Oost, J.; Sensen, C. W.; Charlebois, R. L.; Rossi, M. *J. Biol. Chem.* **2000**, *275*, 22082-22089.

(8) (a) Huie, M. L., Tsujino, S., Sklower-Brooks, S., Engel, A., Elias, E., Bonthron, D. T., Bessley, C., Shanske, S., DiMauro, S., Goto, Y. I., Hirschhorn, R. *Biochem. Biophys. Res. Commun.* **1998**, *244*, 921-927. (b) Hermans, M. M., van Leenen, D., Kroos, M. A., Beesley, C. E., Van Der Ploeg, A. T., Sakuraba, H., Wevers, R., Kleijer, W., Michelakakis, H., Kirk, E. P., Fletcher, J., Bosshard, N., Basel-Vanagaite, L., Besley, G., Reuser A. *J. Hum. Mutat.* **2004**, *23*, 47-56.