Supporting Information

NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity

Romain Irague,[†,‡,Ŧ] Stéphane Massou, [†,‡,Ŧ] Claire Moulis, [†,‡,Ŧ] Olivier Saurel,[⊥] Alain Milon, [⊥] Pierre Monsan, [†,‡,Ŧ] Magali Remaud-Siméon, [†,‡,Ŧ] Jean-Charles Portais[†,‡,Ŧ], Gabrielle Potocki-Véronèse*[†,‡,Ŧ]*

Université de Toulouse, INSA, UPS, INP, LISBP, 135 avenue de Rangueil, F-31077 Toulouse, France, UMR5504, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, CNRS, INRA, F-31400 Toulouse, France, CNRS, UPS, Institut de Pharmacologie et de Biologie Structurale (IPBS) ; 205 rte de Narbonne, Toulouse, France

[†]Université de Toulouse

[‡]CNRS UMR 5504

[Ŧ]INRA UMR 792

[⊥]Institut de Pharmacologie et de Biologie Structurale

* Corresponding authors: veronese@insa-toulouse.fr; jean-charles.portais@insa-toulouse.fr

Table of contents

Materials and Methods section

DNA manipulations. A variant library was constructed by using the ISOR method[1] from the parental gene encoding the dextransucrase DSR-S vardel Δ4N. An *AatII* restriction site was first incorporated in the pBAD *Thio-dsrs* vardel Δ4N-*His* plasmid[2] by silent mutation using inverse PCR at the position 1881 of the gene dsr-s vardel Δ4N, to generate the pBAD *Thio-dsrs* vardel Δ4N_*aatII*1881-*His* plasmid. Then, a 1130 bp cassette, corresponding to the positions 829 to 1958 of the gene dsr-s vardel Δ4N was PCR amplified from the pBAD *Thio-dsrs* vardel Δ4N_*aatII*1881-*His* plasmid. Five µg of purified PCR products were digested with 1U of DNaseI in the supplied buffer at 20°C in a final volume of 50 µL. The reaction was stopped after 3 min by adding 15 µL of 0.5 M EDTA and heating 10 min at 75°C. Fragments were separated on a 2% agarose gel and those with a size between 50-100 bp were extracted using QIAquick Gel Extraction Kit (Qiagen). Gene reassembly was carried out with 100 ng of purified fragments in combination with 2 µM of degenerated oligonucleotides mix, to mutate positions D306, F353, N404, W440, D460, H463, T464 and S512 (DSR-S vardel Δ4N numbering). The reaction mixture (30 µL), containing 1U Phusion® High-Fidelity DNA Polymerase (Finnzyme) in the appropriate buffer and 0.4 mM of each dNTPs, was thermocycled according to the following program: 1 denaturation step at 98°C for 30 s; 40 cycles composed of a denaturation step at 98°C for 10 s, 6 successive hybridization steps separated by 4°C each, from 65°C to 41°C for 10 s each and an elongation step at 72°C for 20s; and finally a 2 min step at 72°C. The fully recombined cassettes were isolated from the reassembly products in a last amplification by nested PCR using the primers forK7A3nted (5'-CCACAGTGGAATGAAACTAGTGAAGATATG-3') and revK7A8nted (5'-ATGGCATCTTTACCATAGCGAACACTT-3'). The purified nested PCR products were digested with *SpeI* and *AatII* restriction enzymes and ligated into the pBAD *Thio-dsrs* vardel Δ4N_*aatII*1881-*His* plasmid to substitute the parental cassette. The ligation products were precipitated by adding 5 volumes of absolute ethanol and the DNA pellet was rinsed 2 times with 70% ethanol. The resuspended plasmids were transformed into *E. coli* TOP 10 electrocompetent cells and plated on LB agar supplemented with

ampicillin (100 µg/ml). The transformants were growth overnight (ON) at 37°C and the colonies were scraped from the plates for plasmid extraction, constituting the glucansucrase DNA libraries.

Selection of glucansucrase active clones. The glucansucrase DNA libraries were transformed into chemiocompetent *E.coli* BL21 AI cells (Invitrogen, Carlsbad, CA) and plated onto LB agar supplemented with 100 µg/ml ampicillin. After overnight growth at 37°C, the cells were scraped, resuspended with physiological water and diluted to an OD_{600nm} of 5.10^{-5}. The clones were subjected to selection pressure by plating them on 22x22 cm plates containing solid M9 mineral medium (42 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 8.5 mM NaCl, $2.5.10^{-2}$ mM $CaCl_2$, 1 mM $MgSO_4$) supplemented with 100 µg/ml ampicillin, 0.02 % arabinose (wt/vol) and 146 mM sucrose as the sole carbon source. The plates were incubated 7 days at 20°C allowing enough growth and expression of the recombinant proteins. Glucansucrase positive variants were picked and transferred to 96 well microplates (Nunc™ Brands Products, Roskilde, Denmark) filled with 250 µL LB per well supplemented with 100 µg/ml ampicillin using a Biomek2000 pipettor (Beckman Coulter, Brea, CA). After overnight growth at 30°C under horizontal shaking at 250 rpm, each starter cultures was duplicated into new microplates containing 250 µL LB supplemented with 100 µg/ml ampicillin and 9 % (wt/vol) glycerol. The resulting libraries were stocked at -80°C.

Screening of the active glucansucrase library

Small-scale expression and oligosaccharides synthesis. Starter cultures of active variants were inoculated from storage microplates using the QpixII automate (Genetix, Hampshire, UK). After growth for 24 h at 30°C under agitation, 50 µL of each starter culture were used to inoculate 96-deepwell plates (ABgene, Epsom, UK) containing 500 µL auto-inducing media ZYM-5052[3] supplemented with 100 µg/ml ampicillin and 0.1 % arabinose (wt/vol). Growth and gene expression were conducted during 48 h at 20°C in an incubator-shaker (INFORS HT, Bottmingen, Switzerland). Plates were centrifuged (5 min, 3700 g, 4°C) and the supernatants removed. Bacterial cell pellets were resuspended in 200 µL of

0.5 mg/ml lysozyme solution, incubated 20 min at 37°C and frozen at -80°C for 12 h. After thawing at room temperature, 800 µL of reaction mixture containing sucrose and iso-maltooligosaccharides (1 kDa, Pharmacosmos, Denmark) diluted in buffered deuterium water (50 mM NaAc, 0.05 g/L $CaCl_2$, pH 5.2) at a final concentration of 292 mM and 50 mM respectively, were added to each well. Enzymatic reactions were incubated 48 h at 25°C under agitation. After that, 50 µL of TSP-d4 (trimethylsilyl)propionic acid-2,2,3,3-d4 acid) in D_2O were added to each well at a final concentration of 37 mM. The deepwell plates were then loaded onto the flow injection NMR system rack holders and subjected to 1D ^1H NMR analysis.

***NMR spectroscopy*.** Measurements were performed by using the BEST system combining a Bruker Avance 600 MHz spectrometer with a Gilson Liquid Handler for sample train preparation and injection. A 5 mm TCi (Triple Cryocooler inverse) probe with a 120 µL active volume cryoFIT (CryoProbe Flow Conversion System) was used. Samples were transferred to the probe with a flow rate of 2.6 ml/min and eliminated after analysis. All NMR data were acquired at 298 K and processed with IconNMR software, requiring 3 min per sample. A ZGPR pulse sequence, for presaturation of the residual water signal, was applied and experiments were carried out using the parameters described in Supplementary tables 1 and 2. The ^1H-signal from D_2O was used for automatic lock and a gradient shimming was performed on each sample. Before Fourier transformation, the FIDs were multiplied by an exponential function with a line broadening of 0.3 Hz. Spectra were processed with a 64 k zero filling, baseline correction and referenced using the TSP-d4 signal at 0 ppm.

Analysis of the products synthesised by variants S512C, F353W and H463R/T464D/S512T

***Enzyme production*.** Variants were growth in flask containing 50 ml of ZYM-5052 medium supplemented with 100 µg/ml ampicillin and 0.1 % (wt/vol) arabinose. Growth and gene expression were conducted during 24 h at 20°C in an incubator-shaker. Bacterial cells were collected by centrifugation (15 min, 4500g, 4°C) and resuspended in 3 ml of sodium acetate buffer (50 mM NaAc,

0.05 g.L^{-1} $CaCl_2$, pH 5.2). Cells were sonicated and precipitated by centrifugation (3500 g, 30 min, 4°C), supernatants were used as enzymatic extracts.

Oligosaccharide and glucan synthesis. Reactions were carried out at 25°C, using 1 U.ml^{-1} of enzymes, in sodium acetate buffer (50 mM NaAc, 0.05 g.L^{-1} $CaCl_2$, pH 5.2), and 292 mM sucrose for glucan synthesis or 292 mM sucrose and 50 mM iso-maltooligosaccharides for oligosaccharide synthesis, until sucrose depletion. Reactions were stopped by 5 min incubation at 95°C.

HPLC analysis. Oligosaccharide analysis was performed by HPAEC-PAD using a Dionex Carbo-pack PA100 column (250 x 4 mm, Dionex). A gradient of sodium acetate in 150 mM NaoH was applied at 1 ml/min flow rate as following: from 0 to 300 mM in 30 min, 300 to 450 mM in 1 s, 450 to 0 mM in 5 min and 0 mM during 10 min. Detection was performed using an ED40 module with a gold electrode.

NMR spectroscopy. 1D ^{1}H NMR spectra were recorded on a Bruker Avance 500mHz spectrometer using a 5 mm z-gradient BBI probe at 298 K, an acquisition frequency of 500.13 MHz and a spectral width of 8012.82 Hz. Oligosaccharide samples were lyophilised and dissolved in 650 µL of D_2O. Glucan samples were precipitated with one volume of absolute ethanol, recovered by centrifugation, washed with distilled water, lyophilised and dissolved in 650 µL of D_2O. Chemical shifts are given in ppm by reference to TSP-d4 (δ 0 ppm).

Supporting information references

(1) Herman, A.; Tawfik, D. S. *Protein Eng. Des. Sel.* **2007,** *20*, 219-26.

(2) Moulis, C.; Arcache, A.; Escalier, P. C.; Rinaudo, M.; Monsan, P.; Remaud-Simeon, M.; Potocki-Veronese, G. *FEMS Microbiol. Lett.* **2006,** *261*, 203-10.

(3) Studier, F. W. *Protein Expr. Purif.* **2005,** *41*, 207-34.

Figures section

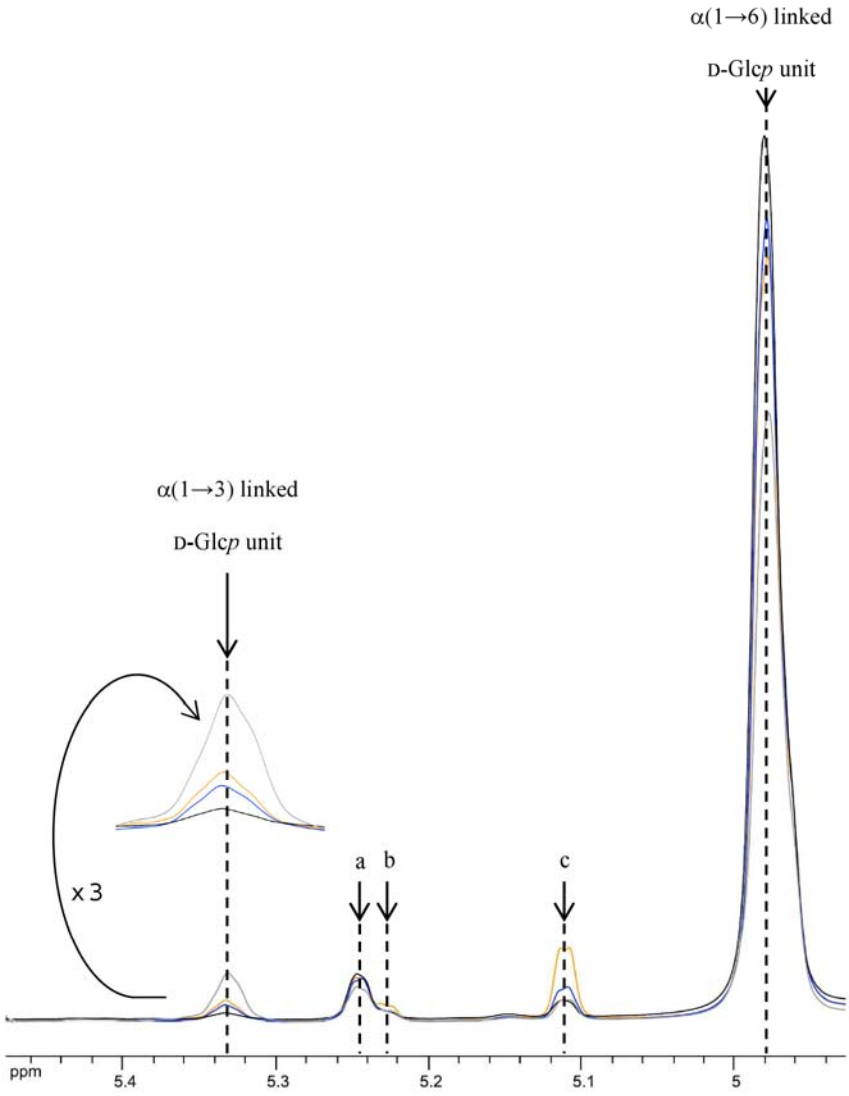

Figure S-1. Superimposition of 1D ^{1}H NMR spectra of the oligosaccharides synthesized by DSR-S vardel Δ4N (black) and the three mutants S512C (blue), F353W (orange) and H463R/T464D/S512T (gray). These mutants show an alteration in their capacity to synthesize α(1→3) glucosidic linkages, visible as disturbance of the relative intensities of α(1→3) linked D-Glc*p* units anomeric proton signals at 5.33 ppm. Spectra were calibrated according to the internal standard TSP-d4 present in each sample at the same concentration (^{1}H = 0 ppm, 37 mM). (a) α-D-Glc*p* reducing unit, (b) free D-Glc*p* unit and (c) α-D-Glc*p* (1→5)-D-fru*p* (leucrose).

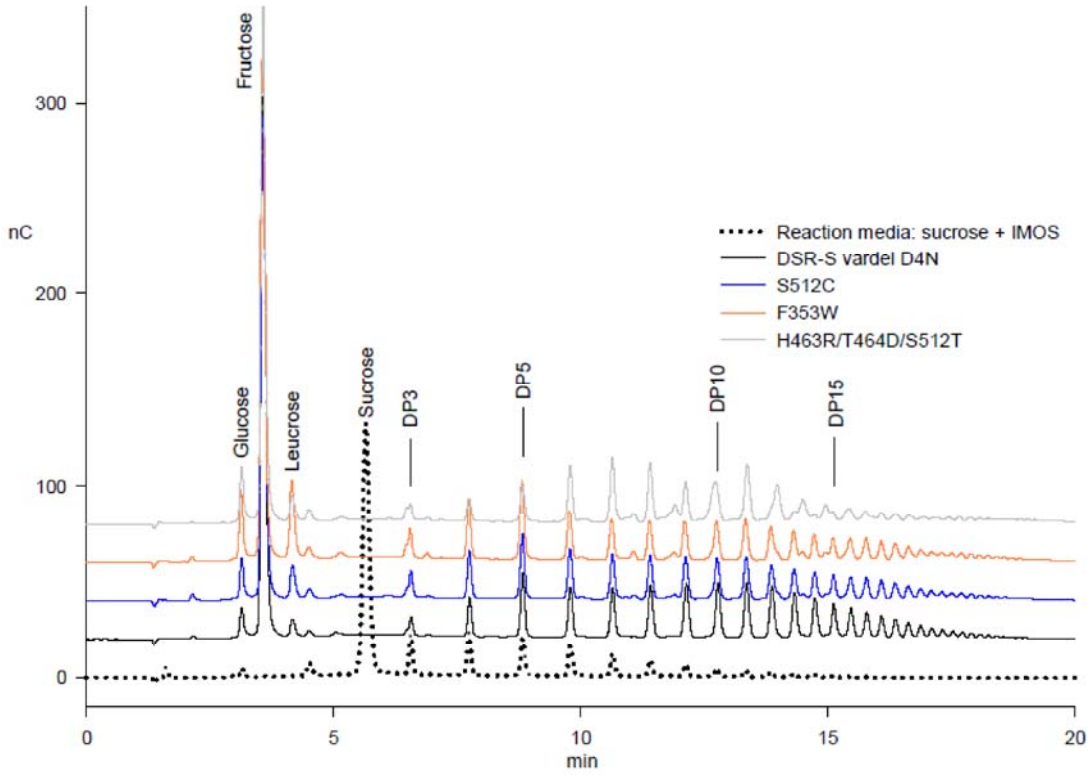

Figure S-2. HPAEC-PAD profiles of the reaction products obtained from sucrose and iso-maltooligosaccharides by using parental DSR-S vardel Δ4N and variants S512C, F353W and H463R/T464D/S512T. IMOS: Isomalto-oligosaccharides; DP: Polymerisation degree (referenced to IMOS used as acceptors)

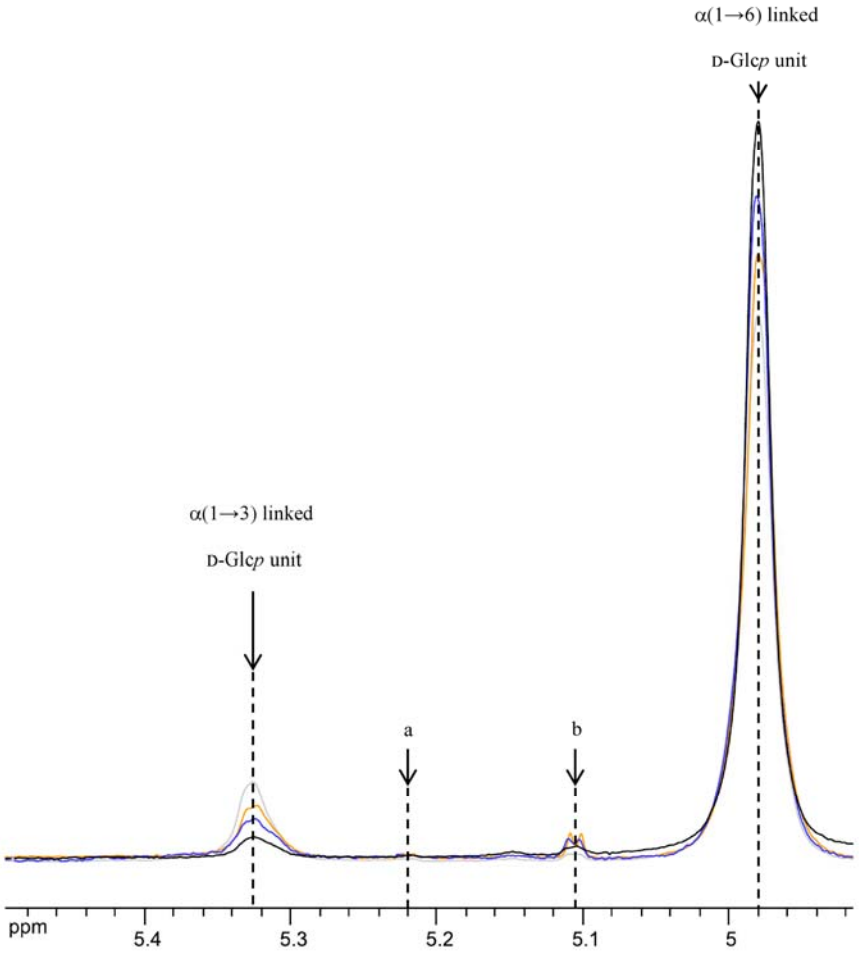

Figure S-3. Superimposition of 1D ^{1}H NMR spectra of the α-glucans synthesized by DSR-S vardel Δ4N (black) and the mutants S512C (blue), F353W (orange) and H463R/T464D/S512T (gray) that show an altered specificity. Spectra were calibrated according to the internal standard TSP-d4 present in each sample at the same concentration (^{1}H = 0 ppm, 37 mM). (a) α-D-Glc*p* reducing unit, (b) α-D-Glc*p* (1→5)-D-fru*p*.

Tables section

Table S-1. Parameters for sample train preparation and automated direct injection of crude reaction media to the NMR spectrometer.

Sample train		Gap composition	Volume (μL)
	1	Push solvent (H_2O)	
Trailing gaps	2	Gas	10
	3	H_2O	35
	4	Gas	10
	5	Sample	500
Leading gaps	6	Gas	10
	7	Sample	35
	8	Gas	10
	9	D_2O	50
	10	Gas	10

Table S-2. NMR measurement parameters.

Temperature	298 K
Time domain	32 k
Number of scans	16
Number of dummy scans	4
Relaxation delay	1 s
Acquisition time	1.95 s
Sweep width	8389 Hz
90° pulse lengh	8.2 µs
Presaturation radio frequency field	12.4 Hz

The experiments have been performed on a Bruker Avance 600 MHz spectrometer equipped with a triple resonance TCi cryoProbe. A flow through BEST system coupled with a cryoProbe flow conversion system insert was used to acquire NMR spectra in a continuous mode and to link the probe to a 96 well plates injection automate.

Table S-3. ^{1}H NMR chemical shifts (in ppm) of the synthesized products.

	Negative control	DSR-S vardel Δ4N	DSR-E	ASR C-del
β-D-Fru*f*-(2↔**1**)-α-D-Glc*p* **H1** (sucrose)	5,413			
α(1→3) linked D-Glc*p* unit		5,339 (**1**)	5,34 (**1**)	5,336 (**10**)
α-D-Glc*p* reducing unit		5,249	5,249	5,249
free D-Glc*p* unit		5,232	5,231	5,229
α(1→6) linked 2,6-di-*O*-substituted D-Glc*p* unit			5,187	
α-D-Glc*p* (1→5)-D-fru*p* (leucrose)		5,11	5,112	5,113
α(1→2) linked D-Glc*p* unit			5,09 (**5**)	
α(1→6) linked D-Glc*p* unit	4,976	4,976 (**99**)	4,977 (**94**)	4,976 (**90**)

Anomeric protons were referenced according to internal TSP-d4 (1H = 0 ppm). Numbers into brackets indicates the ratio of α(1→x) linked D-Glc*p* units involved in the synthesized oligosaccharide primary structures, calculated as follows:

Ratio of α(1→x) linked D-Glc*p* units = $\frac{I_x}{I_t} \times 100$, where I_x and I_t correspond to the relative intensities of the anomeric proton signals of α(1→x) linked D-Glc*p* units and the total α(1→) linked D-Glc*p* units respectively involved in the synthesized oligosaccharide primary structures.