Supplemental Information

Title: Elucidating the Enzymatic Framework for the Initiation of Pectic Galactan.

Authors: Charles J. Corulli[1,2], Pradeep K. Prabhakar[1,2], Santanu Jana[1], Liang Zhang[1], Lubana Shahin[1,2], Digantkumar Chapla[1], Tasleem Javaid[1,2], Dylon Jacob Quiros[1], Lan Na[1], Stephanann M. Costello[1], Kelley W. Moremen[1,2], Breeanna R. Urbanowicz[1,2,] *

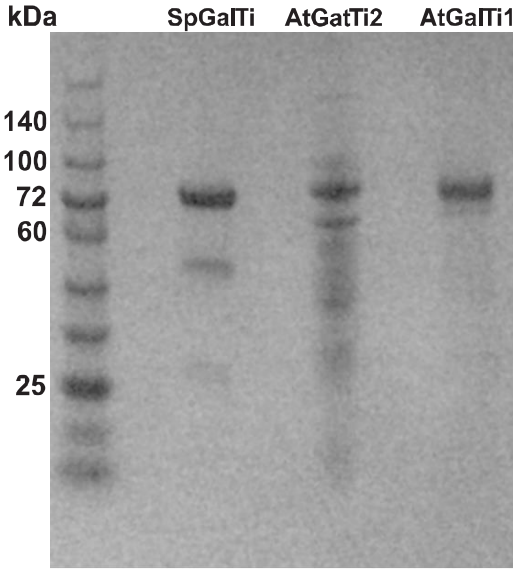

Figure S1: SDS-polyacrylamide gel electrophoretic analysis (SDS-PAGE) of SpGalTi, AtGalTi2, and AtGalTi1.

SDS PAGE gel of SpGalTi (left), AtGalTi2 (middle), and AtGalTi1 (right) all show a similar size of ~80 kDa, as expected from sequence data.

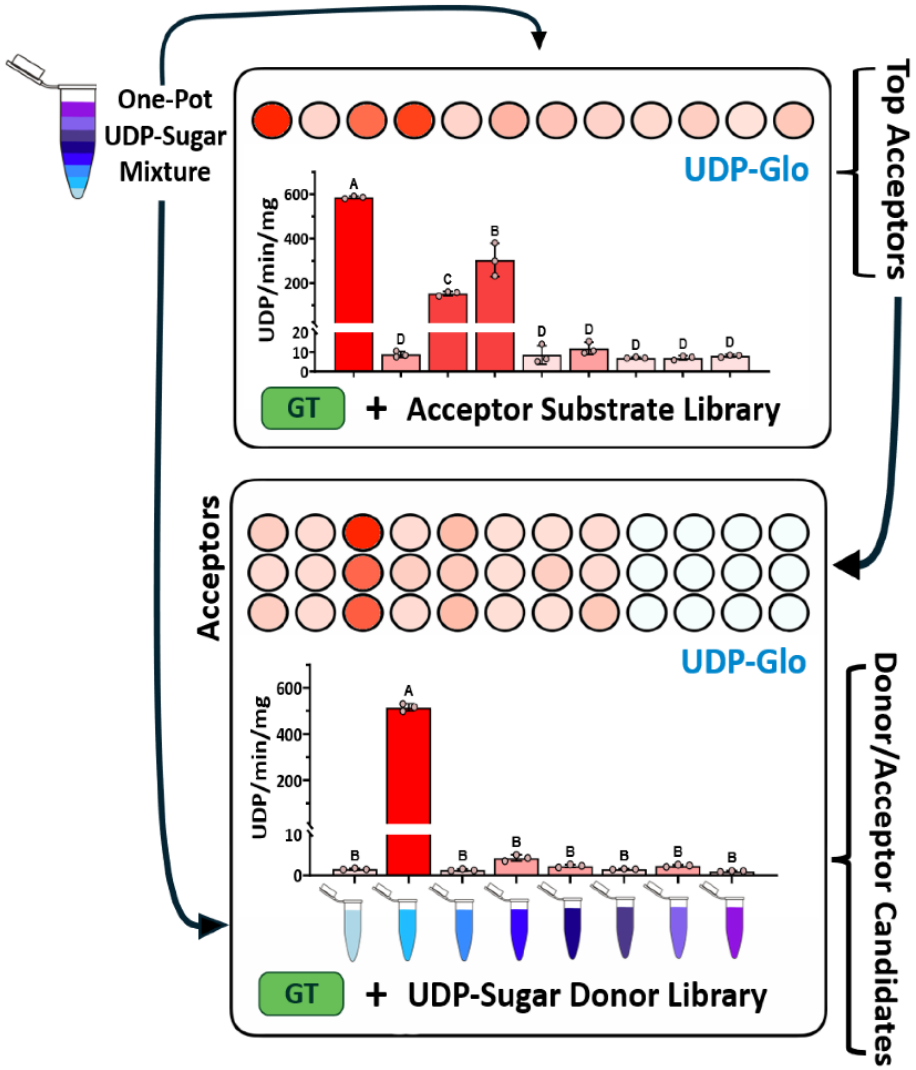

Figure S2: Workflow schematic used for screening GTs against acceptor and donor substrates for elucidating transferase activity.

GTs are screened against a library of plant-specific acceptor substrates while using a NDP-sugar donor mixture to act as an all-encompassing donor substrate. Top acceptor candidates are selected based on GT transferase activity quantified through UDP formation. Top acceptors are then used to screen the GT against a library of individual NDP-sugar donors from the NDP-sugar mixture. Initial screening concludes with the identification of potential donor-acceptor substrate candidates for expressed GTs.

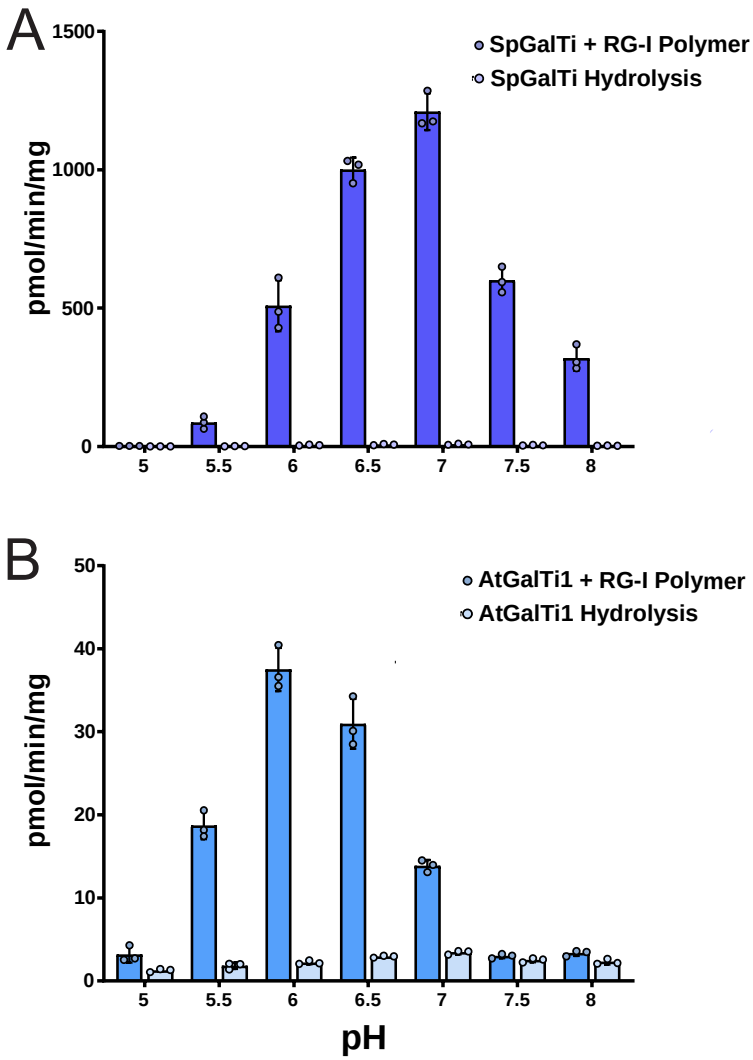

Figure S3: Screening SpGalTi and AtGalTi1 for transferase and hydrolytic activity in different pH conditions.

The transferase and hydrolytic potential of SpGalTi (A) and AtGalTi1 (B) to transfer UDP-Gal to either RG-I polymer or water was measured at various pH conditions (from pH 5 – 8). An equimolar mixture of 250 mM Tris/MOPS/MES was used as a buffer in all reactions to achieve the desired pH ranges. SpGalTi appears to exhibit optimal transferase activity at pH 7 while AtGalTi1 appears to show optimal activity at pH 6. Error bars indicate Mean ± SD with an n of 3 biological replicates.

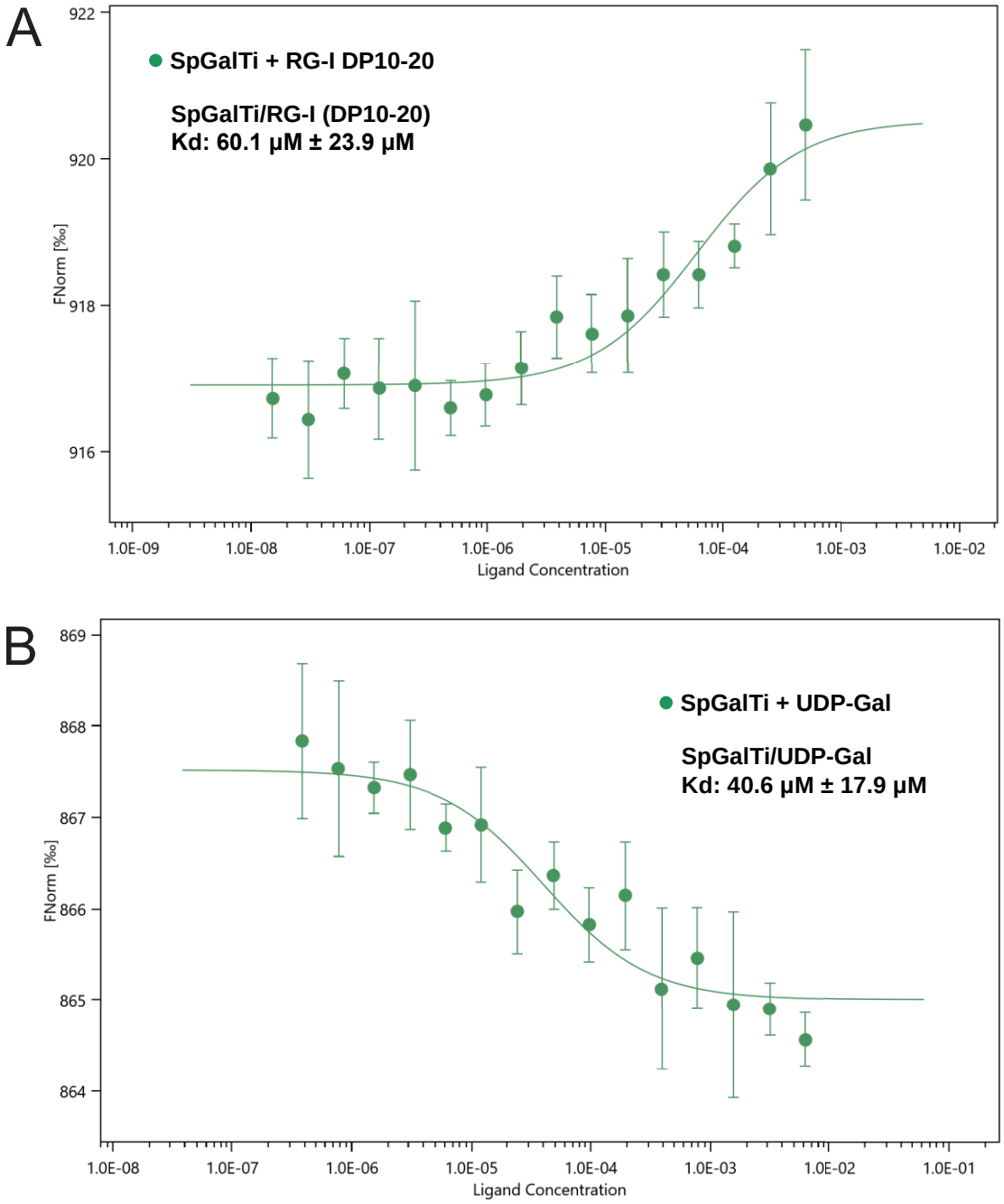

Figure S4: Dose–response curve for the binding interaction between SpGalTi and acceptor/donor substrates.

Dose–response curve for the binding interaction between SpGalTi and the acceptor (RG-I DP 10-20) (A) or the donor (UDP-Gal) (B). Dissociation constant (Kd) of SpGalTi was obtained by Microscale Thermophoresis (MST). The data shown is the Kd obtained after using the Kd fit model in MO.Affinity Analysis software (NanoTemper). Confidence (s.d.) values define the range where the K_D falls with 95% certainty. n = 6 (A) and n=7 (B) respectively.

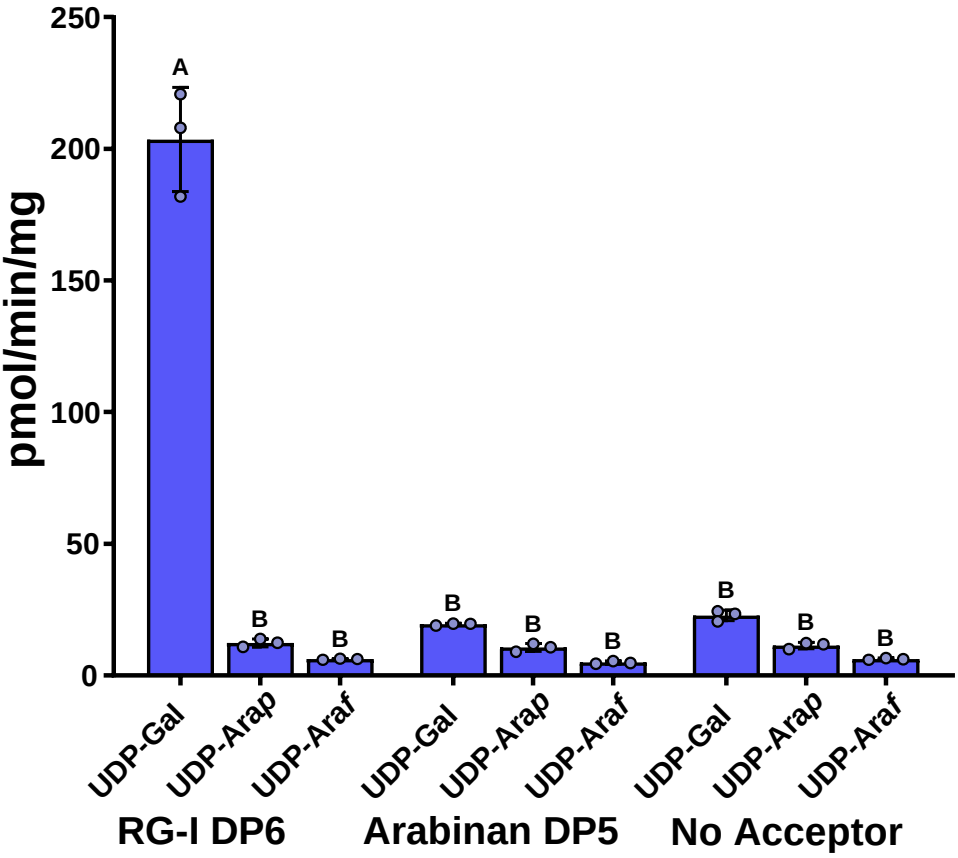

Figure S5: Comparing SpGalTi transferase activity with a RG-I DP 6 or arabinan DP 5 acceptor using UDP-Gal, UDP-Ara*p*, and UDP-Ara*f* donors.

The transferase activity of SpGalTi was compared between an equimolar concentration of two acceptors (RG-I DP 6 and arabinan DP 5) to assess the potential utilization of arabinan as an acceptor substrate. The donors UDP-Gal, UDP-Ara*p*, and UDP-Ara*f* were examined with both acceptor substrates to account for potential bifunctional activity. Error bars indicate Mean ± SD with an n of 3 biological replicates. Letters above indicate statistical significance accomplished by a two-way ANOVA followed by a Tukey's multiple comparison test ($P \leq 0.01$).

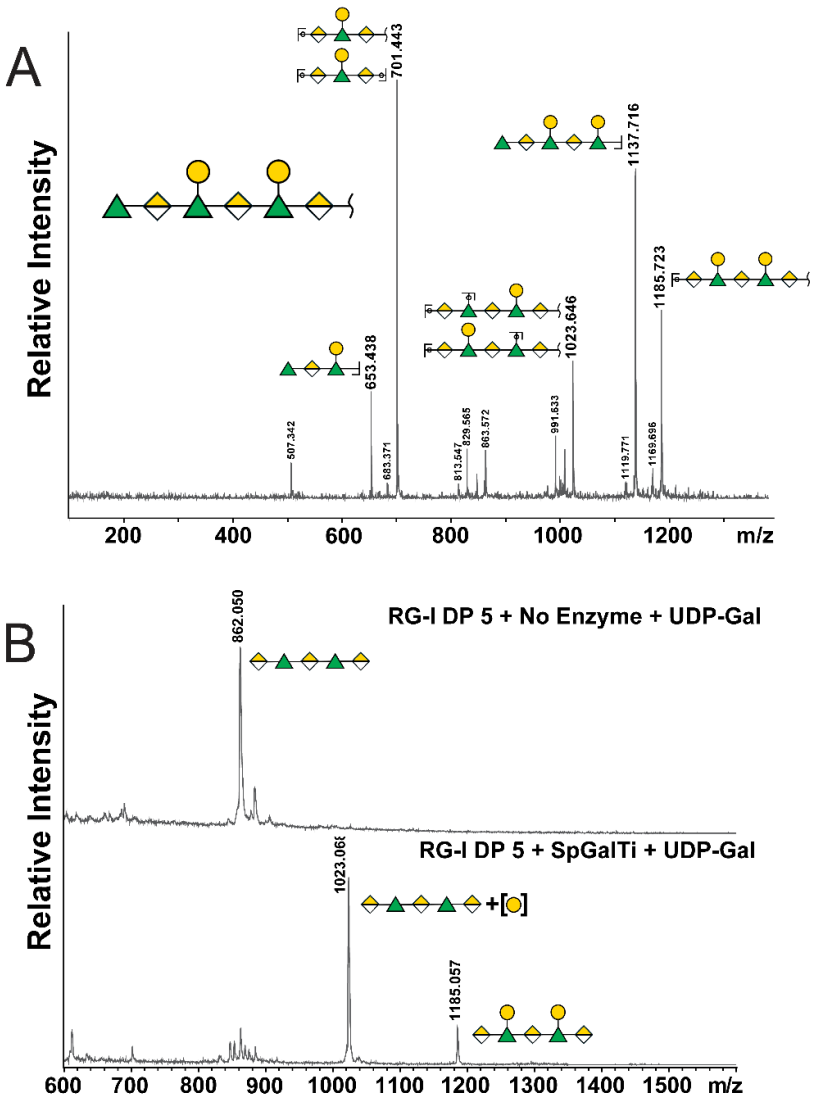

Figure S6: Analysis of SpGalTi reaction products reveal no galactosylation at free-end rhamnose.

(A) MS/MS of SpGalTi RG-I DP 6 reaction products with two Gal additions indicate a structure galactosylated at the 2nd and 3rd backbone Rha residues from the non-reducing end of the substrate. The fragment with a mass of 1185.723 corresponds to this particular patterning of Gal on the backbone. (B) MALDI-TOF MS analysis of SpGalTi reaction products with RG-I DP 5 oligosaccharides in the presence of a UDP-Gal donor. Masses corresponding to an increase of 162 Da represent an addition of one hexose (Gal). Symbol Nomenclature for Glycans (SNFG) was used for the graphical representation of hypothetical glycan structures.

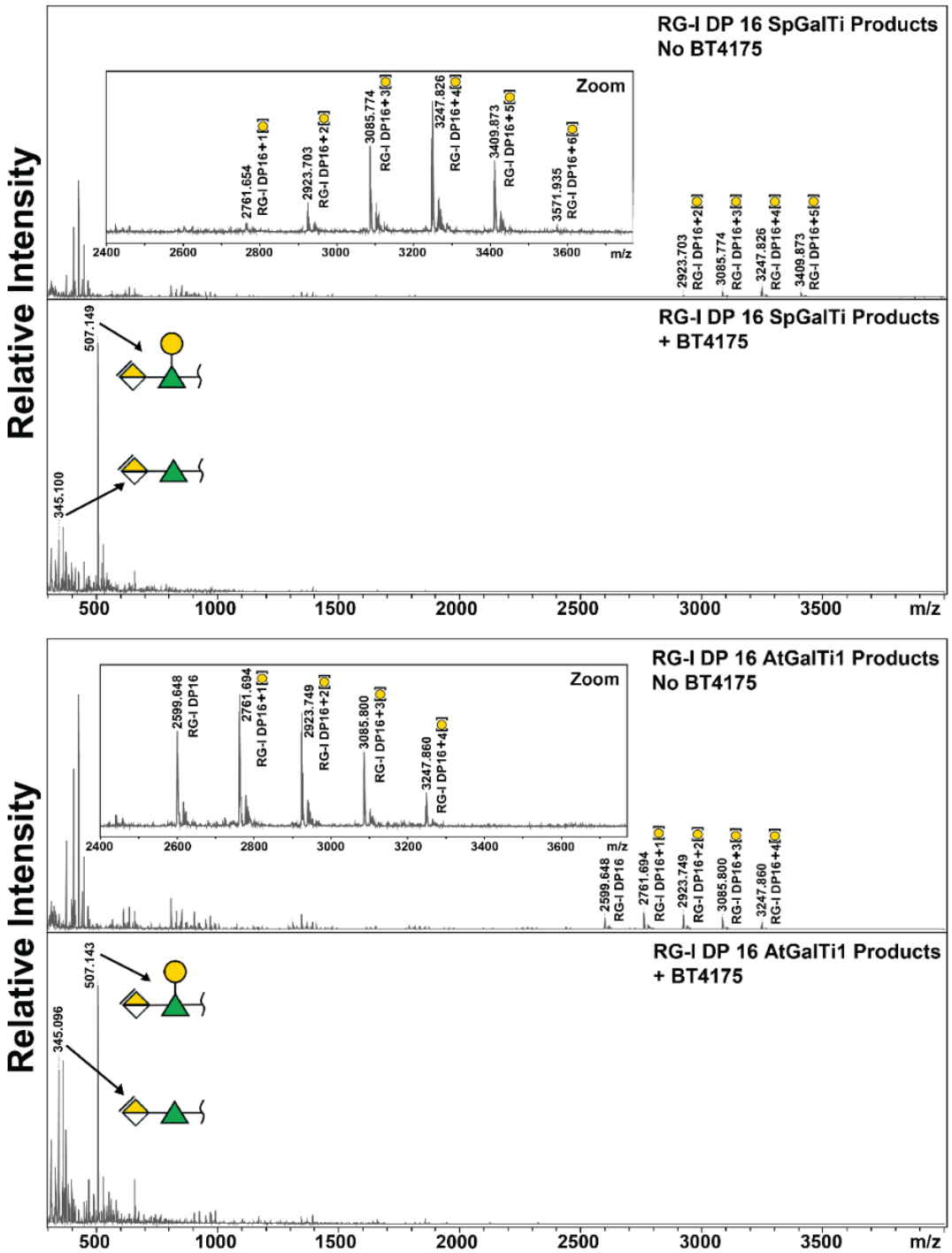

Figure S7: MALDI-TOF MS analysis of lyase digested SpGalTi (top) and AtGalTi1 (bottom) galactosylated RG-I DP 16 oligosaccharides.

After digestion, subsequent RG-I disaccharide products were found to have up to 1 Gal substitution, implying that only single Gal additions were being catalyzed by SpGalTi and AtGalTi1 per RG-I backbone Rha residue. Masses corresponding to an increase of 162 Da represent an addition of one hexose (Gal). Symbol Nomenclature for Glycans (SNFG) was used for the graphical representation of hypothetical glycan structures.

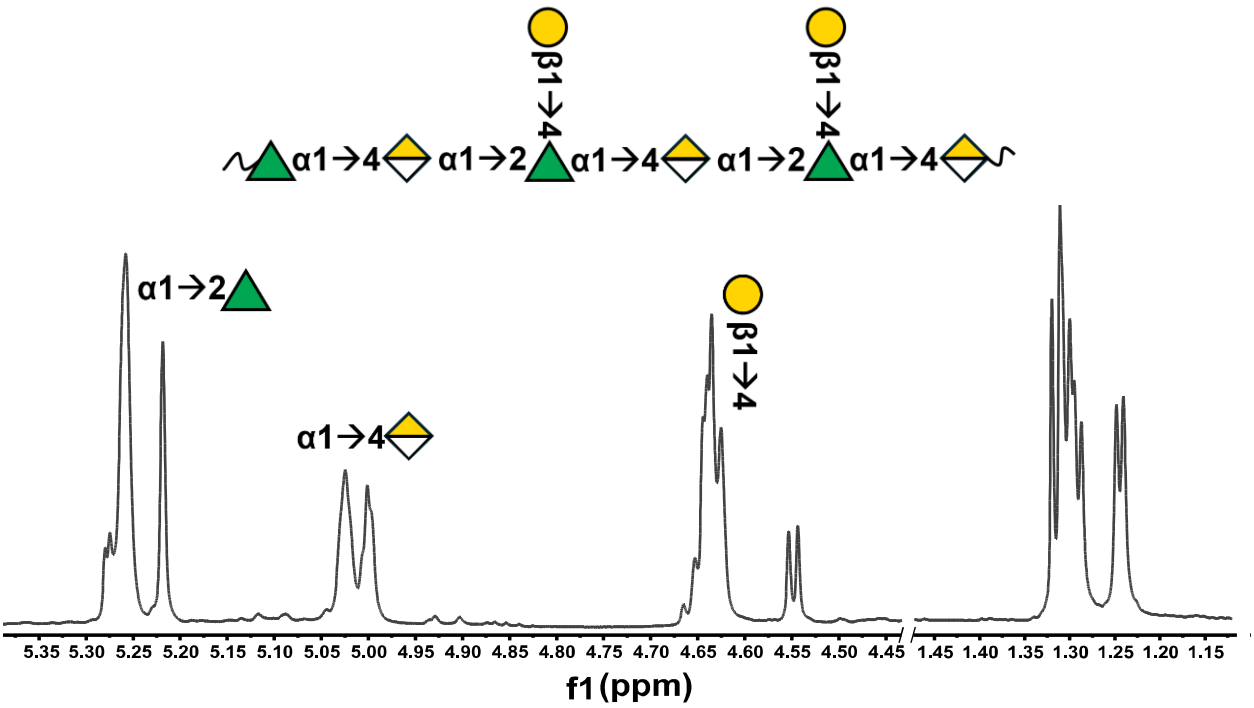

Figure S8: Simplified ^{1}H NMR analysis of the SpGalTi galactosylated RG-I oligosaccharides with product structural components assigned to specific peaks.

Symbol Nomenclature for Glycans (SNFG) was used for the graphical representation of hypothetical glycan structures.

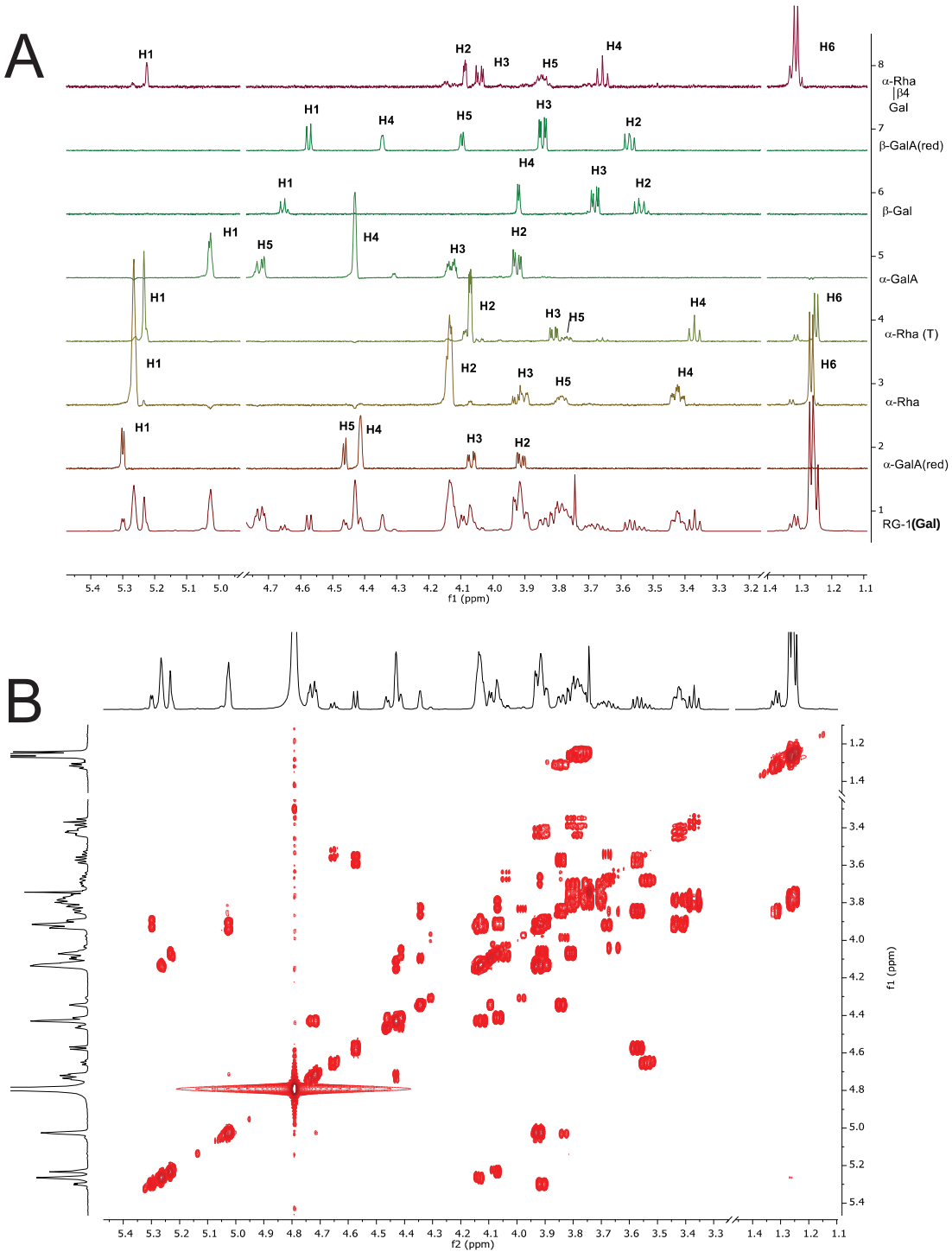

Figure S9: ^{1}H NMR and 2D ^{1}H,^{1}H-COSY spectra of SpGalTi glycosylated RG-I oligosaccharides.

(A) NMR analysis of the galactose-glycosylated product. The full ^{1}H NMR spectrum (bottom, 1) is shown. Above it (top, 2-8), a 1D selective gradient TOCSY spectrum is displayed, which was used to identify the spin system of an individual sugar residue. (B) 2D ^{1}H,^{1}H-COSY spectrum of the glycosylated product. This experiment was used to establish intra-residue proton correlations, identifying adjacent protons (e.g., H1-H2, H2-H3) from the cross-peaks.

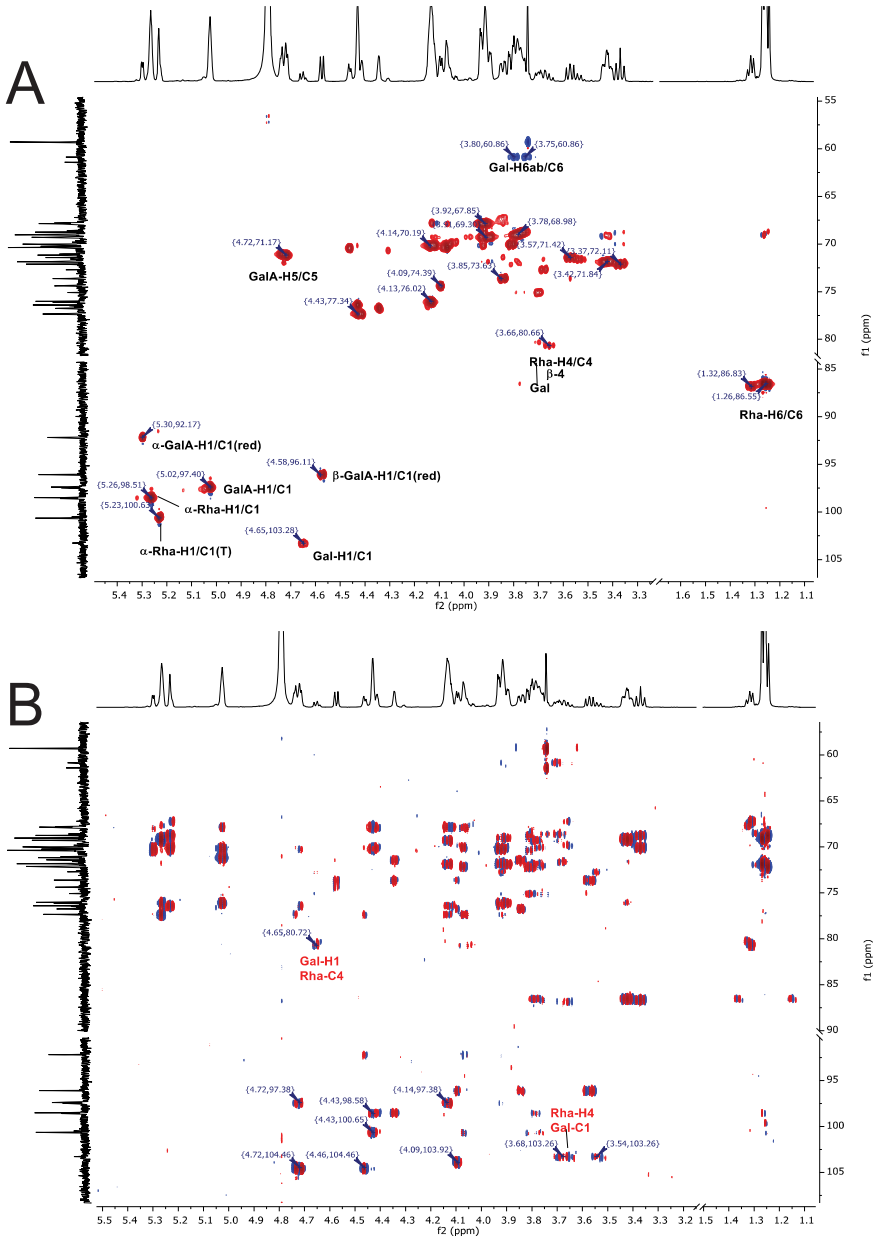

Figure S10: 2D $^{1}H,^{13}C$-HSQC and 2D $^{1}H,^{13}C$-HMBC spectra of SpGalTi glycosylated RG-I oligosaccharides.

(A) 2D $^{1}H,^{13}C$-HSQC spectrum of the glycosylated product. This experiment provides one-bond correlation peaks between each proton and the carbon atom it is directly attached to, which is essential for assigning the carbon skeleton. The spectral width (SW) and transmitter frequency offset (O1P) for the ^{13}C dimension (F2) were set to 70 ppm and 83 ppm, respectively. Peaks originating from outside this selected region (e.g., methyl groups) are folded into the 2D spectrum. (B) 2D $^{1}H,^{13}C$-HMBC spectrum of the glycosylated product. This experiment reveals long-range correlations (typically over 2-3 bonds) between protons and carbon atoms. These correlations are essential for establishing connectivity across glycosidic bonds (e.g., from an anomeric proton to the aglycone carbon) and for confirming the overall sequence. The spectral width (SW) and transmitter frequency offset (O1P) for the ^{13}C dimension (F2) were set to 70 ppm and 83 ppm, respectively. Peaks originating from outside this selected region (e.g., carbonyl and methyl groups) are folded into the 2D spectrum.

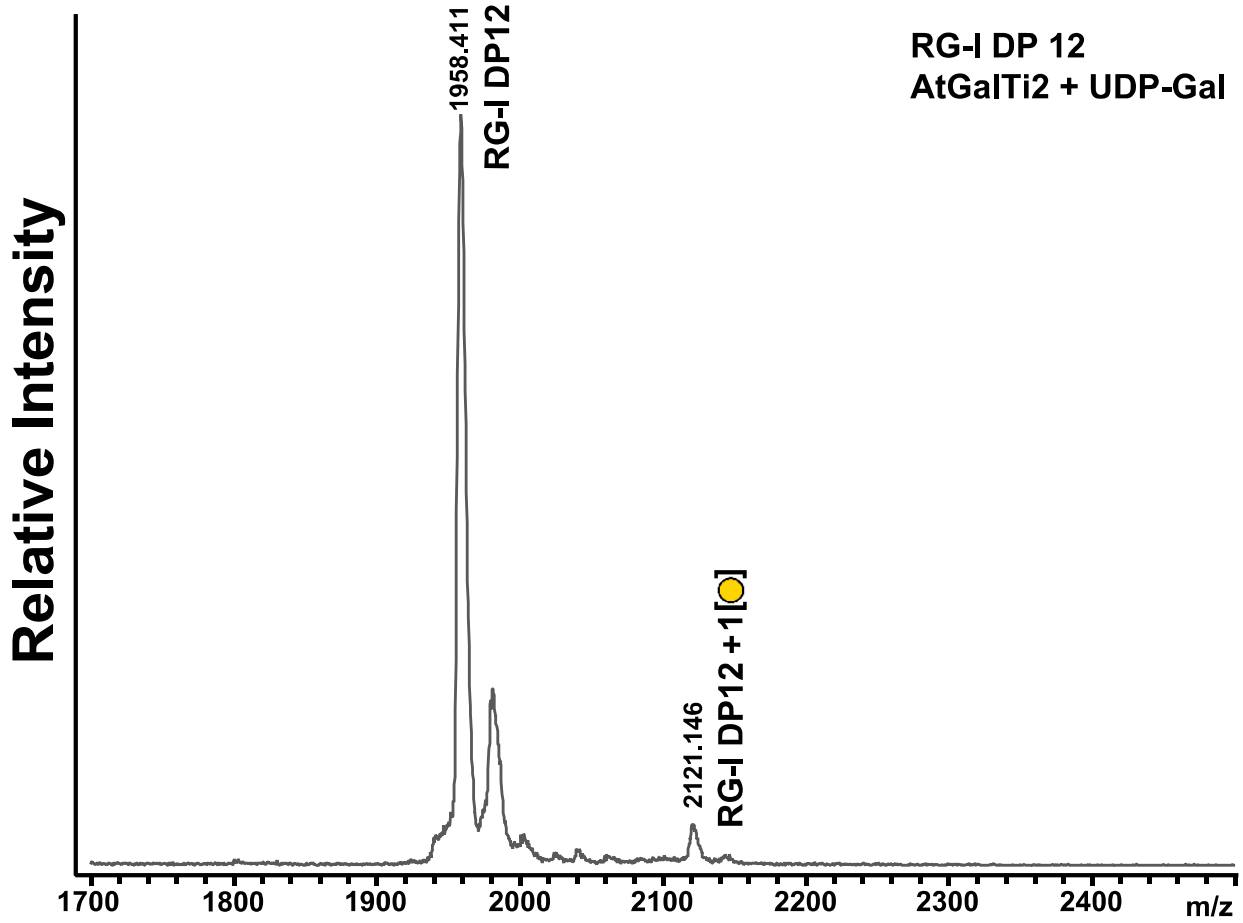

Figure S11: MALDI-TOF MS analysis of AtGalT2 reaction products.

MALDI-TOF MS analysis of AtGalTi2 reaction products with RG-I DP 12 oligosaccharides in the presence of a UDP-Gal donor. Masses corresponding to an increase of 162 Da represent an addition of one hexose (Gal). Symbol Nomenclature for Glycans (SNFG) was used for the graphical representation of hypothetical glycan structures.

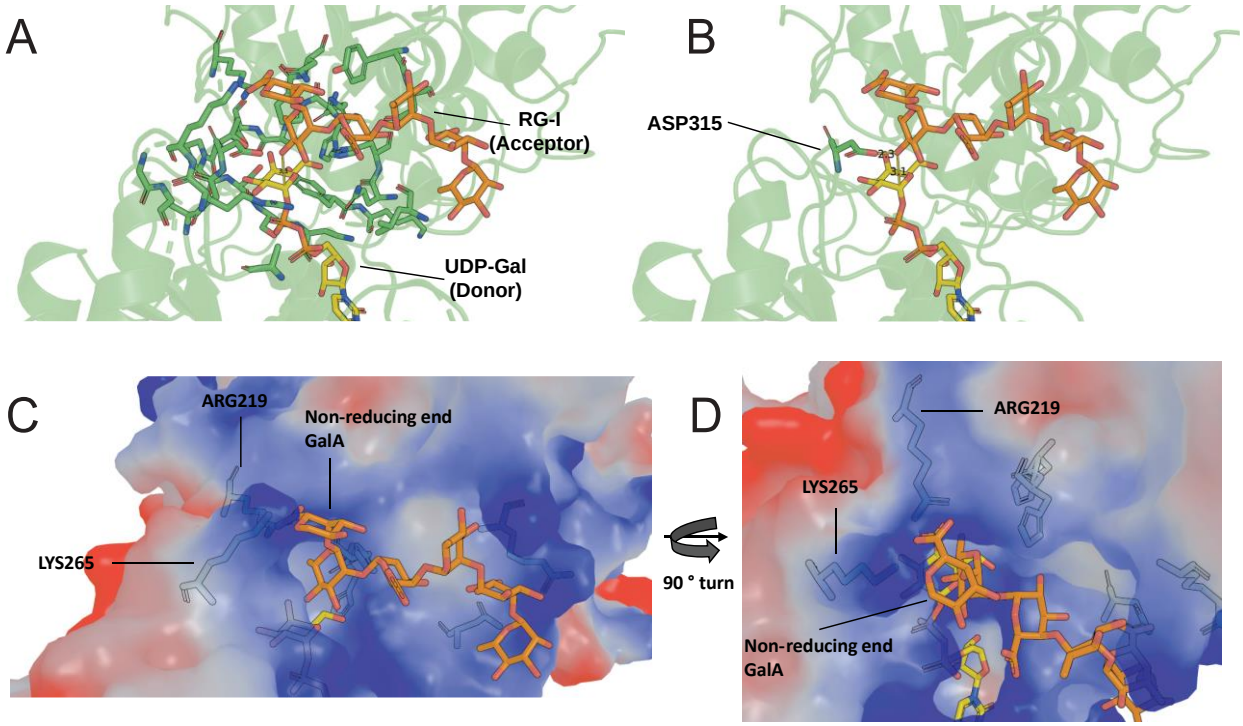

Figure S12: SpGalTi AlphaFold3 predicted structures and electrostatic basis of acceptor recognition.

(A) AF3-predicted ternary structure of SpGalTi with UDP-Gal (yellow) and a modeled hexasaccharide RG-I acceptor with an -2)-α-L-Rhap-(1-4)-α-D-GalpA- repeating disaccharide backbone (orange). potential interacting residues are shown in green. (B) close-up view showing Asp316 positioned on the β-face adjacent to the 4′-OH of the second rhamnose, located 3.1 Å from the anomeric carbon of UDP-Gal. This geometry supports an SN2-type inverting mechanism typical of GT-B glycosyltransferases. (C-D) Electrostatic surface of SpGalTi (red, negative; blue, positive) bound to a hexasaccharide acceptor (orange) and UDP-Gal (yellow). (C) the positively charged pocket explains specificity for the negatively charged acceptor saccharides including GalA. (D) LYS265 and ARG219 act as gateway residues interacting with the reducing end GalA, positioning the adjacent Rha in proximity to UDP-Gal for catalysis.

Supplemental Methods

Microscale Thermophoresis

The binding affinity between SpGalTi and the acceptor (RG-I DP 10-20) and donor (UDP-Gal) was determined using Microscale Thermophoresis (MST). Before the binding experiments, the assay buffer was systematically optimized to ensure protein stability and prevent surface adsorption in the capillaries. An initial buffer of 200 mM HEPES (pH 7.0) with 1% (v/v) Tween 20 resulted in significant adsorption. The final optimized assay buffer consisted of 200 mM HEPES (pH 7.0), 300 mM NaCl, and 0.1% (v/v) Triton X-100, which provided excellent sample homogeneity and stability. For MST measurements, SpGalTi was diluted in the optimized assay buffer to a final concentration of 20 nM, yielding a fluorescence signal between 300 and 800 arbitrary units. A 16-point, two-fold serial dilution of the RG-I ligand was prepared, with concentrations ranging from 0.5mM to 15.3 nM and with UDP-Gal concentrations ranging from 6.25 mM to 0.000381 mM. The protein solution was then mixed at a 1:1 ratio with the ligand dilution, resulting in a final SpGalTi concentration of 10 nM across all samples. The mixtures were incubated at room temperature for 15 min to reach binding equilibrium and subsequently centrifuged at 10,000 rpm for 5 min to remove any aggregates. All experiments were performed on a Monolith NT.115 Pico instrument (NanoTemper Technologies, Munich, Germany) using standard-treated capillaries. Measurements were conducted at 25 °C with instrument settings of 36% and 40% excitation power and 40% MST power. Each binding experiment was performed in triplicate to ensure reproducibility. The resultant data were analyzed using MO. Affinity Analysis software (v2.3, NanoTemper Technologies). The normalized change in the MST signal (Fnorm) was plotted against the logarithm of the ligand concentration to generate a dose-response curve. The dissociation constant (Kd) was determined by fitting the data to a model of single-site binding equilibrium, which calculates the ligand concentration at which SpGalTi was bound.

lyase Digestion of Galactosylated RG-I Substrates

10 µL of the galactosylation reaction was mixed with 7.5 µL 20 mM HEPES-Na pH 7. RG-I lyase BT4175 and $MnCl_2$ were added to achieve final concentrations of 30 µM and 10 mM, respectively. The RG-I lyase digestion was carried out at 37 °C overnight. 1 µL of the digest was diluted in 9 µL Milli-Q water. The diluted digest was then analyzed using a rapifleX spectrometer (Bruker).

Alpha Fold 3

The official AlphaFold 3 inference pipeline (v3.0.1) was executed on the University of Georgia HPC system (Sapelo2) using a single JSON input specifying the protein sequence, ligand CCD identifiers, and model seeds. Ten independent seeds were generated per complex. Structural visualization and figure preparation were performed in PyMOL.