

## Supplementary Information for

A pair of esterases from a commensal gut bacterium remove acetylations from all positions on complex  $\beta$ -mannans.

### **Authors:**

Leszek Michalak<sup>1</sup>, Sabina Leanti La Rosa<sup>1</sup>, Shaun Leivers<sup>1</sup>, Lars Jordhøy Lindstad<sup>1</sup>, Åsmund Kjendseth Røhr<sup>1</sup>, Finn Lillelund Aachmann<sup>2</sup>, Bjørge Westereng<sup>1\*</sup>

<sup>1</sup> Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås

<sup>2</sup> The Norwegian Biopolymer Laboratory, Department of Biotechnology and Food Science, Norwegian University of Science and Technology, Sem Sælands vei 6-8, N-7491 Trondheim, Norway

### **This PDF file includes:**

Supplementary text  
Supplementary materials and methods  
Figs. S1 to S10  
Table S1 to S3  
References for SI

## Supplementary Information Text

### Protein thermal shift assay

*RiCE17* showed great stability and negligible loss of activity after long term storage at 4°C, pH 8.0, while *RiCE2* was relatively unstable and required storage at pH 5.9, -80°C. To investigate the optimum storage pH and thermal stability of the esterases, we examined the denaturation of *RiCE17* and *RiCE2* in 50mM sodium phosphate buffer at pH 5.0-8.0 using the Protein thermal shift (PTS) assay (Thermo Scientific, USA). *RiCE17* was stable up to 65°C, with the highest observed melting temperature at pH 7.5. *RiCE2* has a markedly lower stability when exposed to increased temperatures, with a melting point of 39°C at pH 8.0 and a highest melting point of just 47°C at pH 6.0. The CBM35 domain does not fold properly on its own (Fig. S2C), while the SGNH domain appears to fold into a slightly less stable structure than the complete enzyme (denaturing at 52°C, pH 8.0, as compared with 65°C for the complete *RiCE17*). This indicates that both domains may be necessary for correct folding.

### HMM building

In order to build Hidden Markow Models (HMMs), the phmmer tool (1) on the European Bioinformatics Institute website ([www.ebi.ac.uk](http://www.ebi.ac.uk)) was used to search for homologous sequences in the UniProt's reference proteomes database. Homologous sequences were used to generate an HMM which was then used to search the UniProtKB (2) database for more distant relatives of the *RiCE17*.

Searching the UniProtKB (2) using the database with the CBM35 domain of *RiCE17* using a protein vs protein database tool phmmer (with significance threshold of  $e^{-20}$ ) resulted in 154 sequences of which 135 came from Firmicutes and were found in the same two domain arrangement with a Lipase\_GDSL\_2 domain. These results indicate that the CBM35 structure might be specific to the phylum.

A phmmer search in the UniProtKB reference proteomes using the whole sequence of *RiCE17* with an E-value for significant hit at  $1e^{-99}$  resulted in 32 sequences, all of which had the same two domain architecture and originated from Firmicutes. These were used for HMM building. A search in the UniProtKB database with an E-value  $<1e^{-50}$  using this HMM returned 482 bacterial sequences with a two domain GDSL2\_Lipase + CBM35 architecture. The majority of these sequences came from Firmicutes (435 sequences). 16 hits came from Bacteroidetes. The sequences of those 482 hits were aligned to identify conserved amino acids. All three catalytic residues (serine 41, aspartic acid 190, histidine 193) and the tryptophan 326 of the clamping domain that orients the substrate in the active site by aromatic stacking were present in the 90% consensus sequence (Fig. S3). The presence of these residues in consensus sequences indicates that they are conserved, and could be considered characteristic for the esterase family.

## Esterase activity

Substrate specificity of the two esterases was tested on a wide range of relevant substrates. Activity was only observed on mannose-based oligosaccharides, and the patterns of activity were always similar: a partial deacetylation when either esterase was used, and a near complete deacetylation when both enzymes were used (Fig. S4 A-C, and Fig. 4). None of the esterases were active on cellulose monoacetate, acetylated xylan and chitin oligosaccharides (Fig. S4 D-F). Transesterification reactions demonstrated a clear preference for acetylations rather than propylation/butyrylation (Fig. S5) and that galactosylations did not hinder *RiCE17* catalysis (Fig. S6).

## NMR analysis: assignment of chemical shifts

The chemical shift for acetylated mannose recorded at 25 °C were assigned (Table S1) by starting at the anomeric signal or primary alcohol and then following the proton-proton connectivity using DQF-COSY, <sup>13</sup>C H2BC and <sup>13</sup>C HSQC-[<sup>1</sup>H,<sup>1</sup>H] TOCSY spectra. <sup>13</sup>C-HSQC is used for assigning the carbon chemical shifts. The <sup>13</sup>C HMBC spectrum provides information of the position of the acetyl group in the monosaccharide residues. The following designations are used: Man – mannose, NRE – non reducing end, SRE – second before the reducing end, Ac – *O*-acetylation, M2 – 2-*O*-acetylated Man, M3 – 3-*O*-acetylated Man, M4 – 4-*O*-acetylated Man, M6 – 6-*O*-acetylated Man, α - the α anomer of Man, β - the β anomer of Man, H1-6; C1-6 refers to the ring carbon/proton number for the monosaccharides. The assignment was verified with previous values reported in the literature (3).

## Acetyl migration

The experiments with *RiCE17* transacetylated mannotriose have shown that temperatures of just 60 °C and pD 7.4 induce acetyl migration (Fig. S8). On the 2-*O*-acetylated mannotriose, the apparent direction of migration was the ‘clockwise’ 2-*O*- → 3-*O*-. To test if acetyl migration is a directional process, we devised a set of experiments complementary to those described in the NMR analysis described in Fig. S8. A sample of GGM was treated overnight with excess *RiCE17* (1 μM final concentration in sample) to remove all 2-*O*-acetylations, leaving a substantial amount of acetylation and exposing non-acetylated 2-*O*- hydroxyls as destination for 3-*O*- → 2-*O*- migrating acetylations. The release of acetate forced the pH of the solution to 5.6, ensuring that no unwanted migration was occurring during the following handling of samples. *RiCE17* was then removed from the solution by filtration, and the solution was split into two samples: one for pH induced migration, and one for temperature induced migration. Analogously to the pD in the NMR experiments (Fig. S8), pH induced migration in (Fig. S8 and 9) was accomplished by adjusting the sample pH to 7.4 by slow addition of a sodium phosphate buffer at pH 7.75, and incubation at 30°C overnight. MS analysis was carried out at every step to ensure that deacetylation was the result of 2-*O*- acetylations being enzymatically removed rather than chemical deacetylation. Temperature induced migration was

accomplished by adjusting the pH to 5.9, and incubation at 60°C for one hour. Conditions in both experiments were selected to match those in the 2-*O*- → 3-*O*- migration experiment (Fig. S8) described above. After the migration was induced, samples were treated with 1 μM final concentration of *RiCE17* again, saving a portion of the original sample as a control, before analysis by MALDI-ToF MS. The cycle of enzymatic treatment, enzyme removal, pH adjustment, migration induction, enzymatic treatment and analysis was completed twice for the pH induced migration and three times for the temperature induced migration before the acetylations were completely removed from the mannan.

Both pH (Fig. S9) and heating (Fig. S10) induce a migration process which replaced the 2-*O*- acetylations, providing fresh substrates for deacetylation by *RiCE17* at each round, until a complete deacetylation was achieved. Reaction rates for the acetyl migration between each pair of adjacent acetylations were presented in literature (4). This study reported the rate of 2-*O*- → 3-*O*- migration as well as the reverse rate as the highest (0.566 h<sup>-1</sup> and 0.395 h<sup>-1</sup> respectively at pD 8.0, 25°C) of all migration steps (4). Inducing migration on a mixture of heterogeneous manno-oligosaccharides with 2-*O*- acetylations selectively removed, allowed us to demonstrate that the migration occurs in all directions depending on the distribution of remaining acetylations. The elevated temperature and pH or pD do not drive this reaction in a particular direction, but rather create a permissive environment where the acetylations can redistribute to their most energetically favourable conformation. The apparent direction of acetyl migration depends on the distribution of acetylations prior to exposure to migration permissive conditions. Migration experiments presented here show that this process occurs both in model substrates and heterogeneous mixtures of hemicellulose obtained from a typical industrial hydrothermal extraction.

These results show that acetylations readily redistribute to an equilibrium when exposed to conditions that facilitate migration, such as heating or pH > 6.0. This finding is especially important for hemicellulose biorefining with enzymatic deacetylation steps – since conditions throughout the process can quickly change the acetyl distribution. As acetylations affect the solubility and viscosity of mannans in solution, selective enzymatic deacetylation and easily achieved redistribution could be used for fine-tuning the physicochemical properties of mannans for hydrocolloid applications. This observation has a special significance for biorefining of mannan rich feedstocks such as softwood mannans. Migration induced by a short exposure to just 60°C, at pH 5.9 as well as previously published data on migration caused by heating (5) imply that the distribution of acetylations present in hemicellulose produced by steam explosion and other common hydrothermal extraction methods may not represent the distribution of acetylations present in the hemicellulose *in vivo*.

## SI Materials and Methods:

### 1. Cloning and expression of *RiCE17* truncated versions.

*RiCE17* gene fragments were amplified from *R. intestinalis* genomic DNA using the primer pairs CE\_up/CE17catD\_rev1 and CE17cbm\_up/CE17\_down, respectively (Table S2). Fragments were cloned into the pNIC-CH expression vector with a C-terminal hexa-

histidine tag by ligation-independent cloning (LIC) (6), giving the constructs pNIC-*RiCE17*<sub>CATD</sub> and pNIC-*RiCE17*<sub>CBM35</sub>. Transformants were verified by sequencing. The two proteins were expressed in *Escherichia coli* BL21(DE3) cells harboring the appropriate recombinant plasmids. The recombinant *E. coli* strains were pre-cultured overnight in Luria Bertani (LB) broth supplemented with 50 µg/mL kanamycin (Sigma-Aldrich, Germany) and then used to inoculate 500 mL of medium consisting of 450 mL LB, 50 µg/mL kanamycin and 50 mL of potassium phosphate buffer (0.17 M KH<sub>2</sub>PO<sub>4</sub>, 0.72 M K<sub>2</sub>HPO<sub>4</sub>). Protein expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM for 16 h at 23°C. Cells were harvested by centrifugation (6000 g for 10 minutes) and resuspended in 30 mL lysis buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM imidazole). A cell-free extract was prepared by pulsed sonication and centrifugation at 15000 g for 15 minutes. The supernatant containing the soluble proteins was collected and filtered with 0.22-µm syringe filters. Recombinant *RiCE2* and *RiCE17* were purified by immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC). Protein purity was verified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Protein concentration was determined using the Bradford assay (Bio-Rad) with bovine serum albumin as a standard.

## 2. Crystallography.

For crystallographic studies, seleno-L-methionine substituted *RiCE17* was produced by first transforming pNIC-*RiCE17* (7) into *E. coli* B834(DE3) cells via the heat-shock method and plating the bacteria onto LB agar supplemented with 50 µg/mL kanamycin. Recombinant cells were grown in Medium A according to the EMBL protocol ([https://www.embl.de/pepcore/pepcore\\_services/protein\\_expression/ecoli/seleno/](https://www.embl.de/pepcore/pepcore_services/protein_expression/ecoli/seleno/)), supplemented with 50 µg/mL methionine (Sigma-Aldrich, Germany) and 50 µg/mL kanamycin at 37°C until an OD<sub>600</sub> of 0.8 was reached. At this point, cells were harvested by centrifugation at 6000 g for 15 minutes and resuspended in an equal amount of kanamycin-supplemented Medium A. Temperature was adjusted at 23°C and, following a 2 h starvation period, the flask was supplemented with 50 µg/mL seleno-L-methionine (Sigma-Aldrich, Germany). After 30 min of further incubation, seleno-L-methionine substituted *RiCE17* expression was induced with 0.5 mM IPTG and cultures were allowed to grow for an additional 48 h before being collected. Cells were harvested by centrifugation at 5000 g for 10 minutes, cell pellets were resuspended in 30 mL of 50 mM Tris buffer containing 10 mM Imidazole and 500 mM NaCl and lysed by sonication. The cell lysate was centrifuged at 12000 g for 15 minutes and the supernatant containing the proteins was collected, sterile filtered and purified as described above.

*RiCE17* and *RiCE2* crystallization conditions were screened using several commercial high throughput 96 condition sitting drop screens, using 20 mg/mL, 10 mg/mL and 5 mg/mL solutions of protein in a 1:1 ratio with ready-mixed mother liquors (200 nL of each). Screening plates were set up using a mosquito HTS liquid handling robot (ttp labtech, UK). Crystals were observed in spot G8 of the INDEX HT screen (Hampton Research, USA), containing 0.2 M ammonium acetate, HEPES pH 7.5 and 25% w/v PEG 3350. Hanging drop optimization grids were manually set up with 0.2 M

ammonium acetate, HEPES pH 7.0, 7.5 and 8.0, 20%, 25% and 30% w/v PEG 3350. Crystallization liquor was mixed 2  $\mu$ L:2  $\mu$ L and 1  $\mu$ L:1  $\mu$ L in the hanging drops, with additional 2  $\mu$ L or 1  $\mu$ L of 5 mg/mL solution of mannotriose, mannotetraose and mannopentaose (Megazyme, Ireland) for co-crystallization.

Crystals were transferred to a cryo-solution containing mother-liquor with 35 % glucose before flash freezing in liquid nitrogen. Diffraction data were collected at beamlines ID23-1 and ID-29 at the European synchrotron Research Facility in Grenoble, France.

The initial structure was solved by single-wavelength anomalous diffraction (SAD) using selenomethionine to obtain an anomalous signal. Data was processed by XDS (8) and scaled by AIMLESS (9). The Phenix software package (10) was used to phase (AUTOSOL) (11) and build (AUTOBUILD) (12) the first structure. Subsequent structures were solved by molecular replacement (Phaser) (13) and refined using REFMAC (14). Model manipulations were carried out using Coot (15) and molecular graphics were generated using Pymol2 (Schrödinger). Crystal data, data collection, and refinement statistics is presented in Table S3.

### 3. Preparative HPLC

Preparative chromatography was conducted using an Agilent 1260 Infinity preparative chromatography system with an XBridge BEH prep OBD 5  $\mu$ m particle size 30x250mm column. Analytical HILIC method was scaled up to 17.5 mL/min flow with 3.85 mL injections of 1-5 mg/mL carbohydrate concentrations. Elution started with a step of 0 – 3.57 min 75 % acetonitrile, followed by a linear gradient of 3.58 – 14.28 min 75 % - 50 % acetonitrile, a step of 14.29 - 21.42 min 50 % acetonitrile, and a final step of 21.43 – 33 minutes at 75 % acetonitrile. Fractions were collected as one minute time slices, acetonitrile evaporated in a fume hood and liquid fractions freeze-dried.

### 4. Quantification of acetate release and determination of degree of acetylation

Bound acetate was released by treating samples overnight (at 4°C) in 100 mM KOH or NaOH to determine the degree of acetylation of the substrates. Turnover rates and specific activities were calculated based on the release of acetate from a 100 mg/mL solution of Norway spruce GGM in 50mM sodium phosphate pH 7.0 treated with either 50 nM final concentration of either esterase, or 25 nM of each esterase when both were used in combination. Samples were drawn at 5 minute intervals, the reactions quenched by mixing the reaction solution 1:1 with a 1:1 mixture of acetonitrile and isopropyl alcohol, and heating for one minute in a heating block at 100 °C. Acetate content was analyzed on an RSLC Ultimate 3000 HPLC (Dionex, USA) using a REZEX ROA-Organic Acid H+ 300x7.8mm ion exclusion column (Phenomenex, USA) at 65°C, 10  $\mu$ L injection volume, with isocratic elution using 0.6 mL/min of 5 mM H<sub>2</sub>SO<sub>4</sub> as mobile phase and a UV detector set to 210 nm. Constituent monosaccharide of residuals and extracts were quantified either by GC via alditol acetates after acid hydrolysis for the Norway spruce GGM (16) or via sulphuric acid hydrolysis and HPAEC analysis according to (17), whereas for the Konjac mannan the figures on monosaccharide

composition provided in the datasheet where used (Megazyme). The acetate levels were quantified the same way as in (17). The degree of acetylation was determined as mole acetate/mole mannose.

## 5. Transesterification

Transesterification of oligosaccharides was conducted using vinyl acetate, vinyl butyrate and vinyl propionate (Thermo scientific, USA) as acyl donors. Enzymes were added to oligosaccharide solutions with concentrations from 1 – 10 mg/mL, and a volume of vinyl esters equal to 20 – 50 % sample volume was added. The samples were incubated in a thermomixer (Eppendorf, Norway) shaking at 600 rpm overnight, then moved to a freezer at -20°C. The vinyl acetate/butyrate/propionate, which remained liquid on top of the frozen aqueous phase was removed from the samples, and 96% ethanol was added on top of the frozen sample until the final concentration exceeded 80% in order to deactivate the enzymes. Samples were then thawed by vortexing and filtered through a pre-washed 1mL Amicon Ultracel 3kDa ultrafiltration device (Merck KGaA, Germany) to remove the enzymes completely and minimize the risk of deacetylation/depropylation/debutylation. Oligos were then dried in an Eppendorf Concentrator plus (Eppendorf, Norway) at room temperature.

For the purpose of preparative chromatography, samples were frozen, liquid layer of vinyl acetate was removed and samples were then diluted with acetonitrile to a 75 % concentration, foregoing the filtration step.

## 6. Computational details of molecular dynamics simulations

To compare binding of pNP-acetate to the active sites of *RiCE17* (PDBid 6HFZ) and *CjCE2B* (PDBid 2W9X), models of these enzymes were prepared for molecular dynamics. The most likely protonation state of each titratable amino acid side chain in models was assessed using the H++ server (18) and the histidine residue of the catalytic triad (His193 in *RiCE17*) and the dyad (His in *CjCE2B*) was assigned the HID state. This enables the -OH hydrogen of the catalytic serine to form a hydrogen bond to the N $\epsilon$  of this His residue. The disulfide bridge forming residues Cys165 and Cys333 in the *CjCE2B* model were renamed to CYX.

The pNP-acetate molecule was built in Avogadro (19) and geometry optimized using ORCA4 (20) using the hybrid functional B3LYP (21, 22) and the basis set def2-SVP (23, 24) while applying the Grimme D3 dispersion correction(25). Further, the pNP-acetate force field parameters were derived from the GAFF-force field using the programs antechamber and tLeap that are distributed with the AmberTools package (26). Sodium and chloride ions were added to a final concentration of 150 mM in ratios that ensured system neutrality. The models were solvated in a truncated octahedron of OPC water (27) with a distance of 12 Å to the periodic boundary, resulting in models consisting of about 90000 and 75000 atoms for *RiCE17* and *CjCE2B*, respectively.

Both models were subjected to 5000 steps of energy minimization. Then the models were heated to 300 K in the NVT ensemble over a period of 40 ps. Density equilibrations

were performed at 300 K for 0.5 ns at a constant pressure of 1 atm using the Berendsen barostat with a pressure relaxation time of 1 ps. In the following 0.5 ns equilibration step, performed in the NVT ensemble using the weak coupling algorithm and a time constant of 10 ps to regulate the temperature, a set of two distance restraints were added to both models. The first harmonic restraint ( $d = 2.0 \text{ \AA}$ ,  $k = 0.1 \text{ kcal.mol}^{-1} \cdot \text{\AA}^{-2}$ ) was applied to the distance between the catalytic Ser -OH hydrogen atom and the N $\epsilon$  of the catalytic triad/dyad His residue. The second harmonic restraint ( $d = 1.6 \text{ \AA}$ ,  $k = 0.1 \text{ kcal.mol}^{-1} \cdot \text{\AA}^{-2}$ ) was applied to the distance between the catalytic Ser -OH oxygen atom and the electrophilic carbon of the pNP-acetate molecule. The purpose of these restraints was to mimic scenarios where pNP-acetate is bound in the active sites of *RiCE17* and *CjCE2B*. Finally, data was collected from 100 ns production runs with both models (ensembles of 10000 structures), applying the same conditions and restrains as used in the last equilibration step.

In all simulations, periodic boundary conditions with a 10  $\text{\AA}$  cutoff for nonbonded interactions, PME treatment of long-range electrostatics (28) and a time step of 2 fs were applied. Hydrogen atoms were constrained by the SHAKE algorithm (29). Simulations were performed using the CUDA version of PEMEMD included in AMBER18 (30). Analysis of production trajectories was performed using the *cpptraj* module included in AmberTools19 (31) and in-house Python scripts.

## 7. NMR

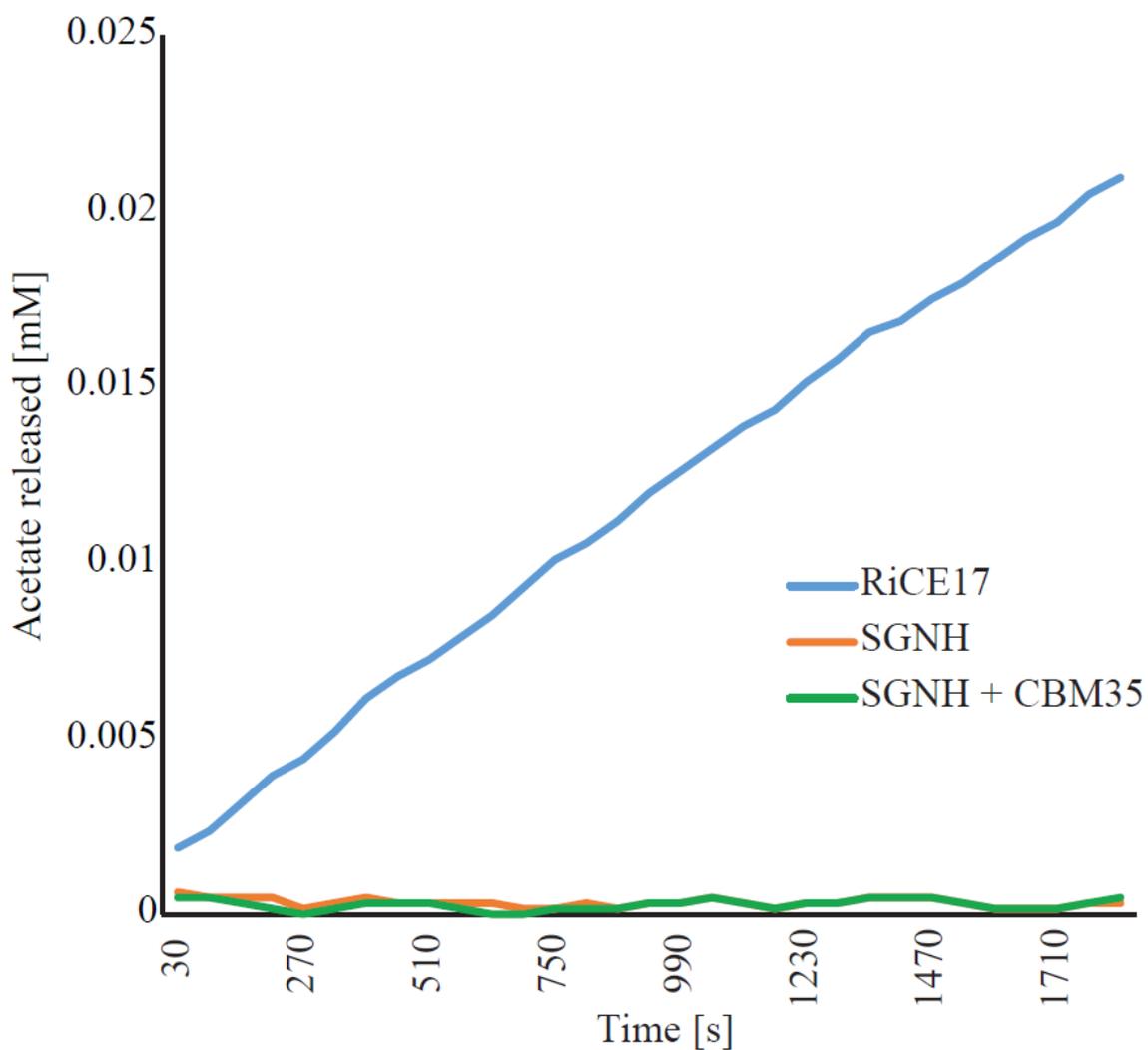
To reduce the interference of the water signal the substrate, *RiGH26* treated spruce galactoglucomannan was dissolved in 99.9% D<sub>2</sub>O (Sigma-Aldrich, Germany) and lyophilized. Similarly, 10 mL 40 mM phosphate buffer pH 5.9 and 250 mM phosphate buffer pH 8.0 were lyophilized and the powders were dissolved in 10 mL 99.9% D<sub>2</sub>O.

For the time-resolved NMR recordings: 4-5 mg of Norway spruce galactoglucomannan hydrolyzed with the *RiGH26* mannanase or *RiCE17* transacetylated mannotriose were dissolved in 500  $\mu\text{L}$  40 mM phosphate buffer pD 5.9 (99.9% D<sub>2</sub>O) and transferred to a 5 mm NMR tube. The sample was preheated in the NMR spectrometer for  $\sim 10$  min. Hereafter all recording parameters were set prior to the time-resolved NMR experiment. 2 or 5  $\mu\text{L}$  of enzyme solution (to a final concentration of 1  $\mu\text{M}$  *RiCE17* or 10  $\mu\text{M}$  *RiCE2*) was added to the preheated substrate and mixed by inverting the sample three times. The sample was then immediately inserted into the preheated NMR spectrometer and the experiment was started (time from adding the enzyme to the first spectra has been recorded was between 3-4 minutes in total). The recorded spectrum is a pseudo-2D type experiment recording a 1D proton NMR spectrum with weak water suppression (Based on Bruker 1D proton setup for metabolomics noesygppr1d) every 5 min with in total 200 time points. The recorded 1D proton spectrum contains 32K data points and has a spectral width of 10 ppm, 24 scans, and pre-saturation during relaxation delay and 10 ms mixing time with spoil gradient and relaxation delay of 1 s (total recording time of 89s).

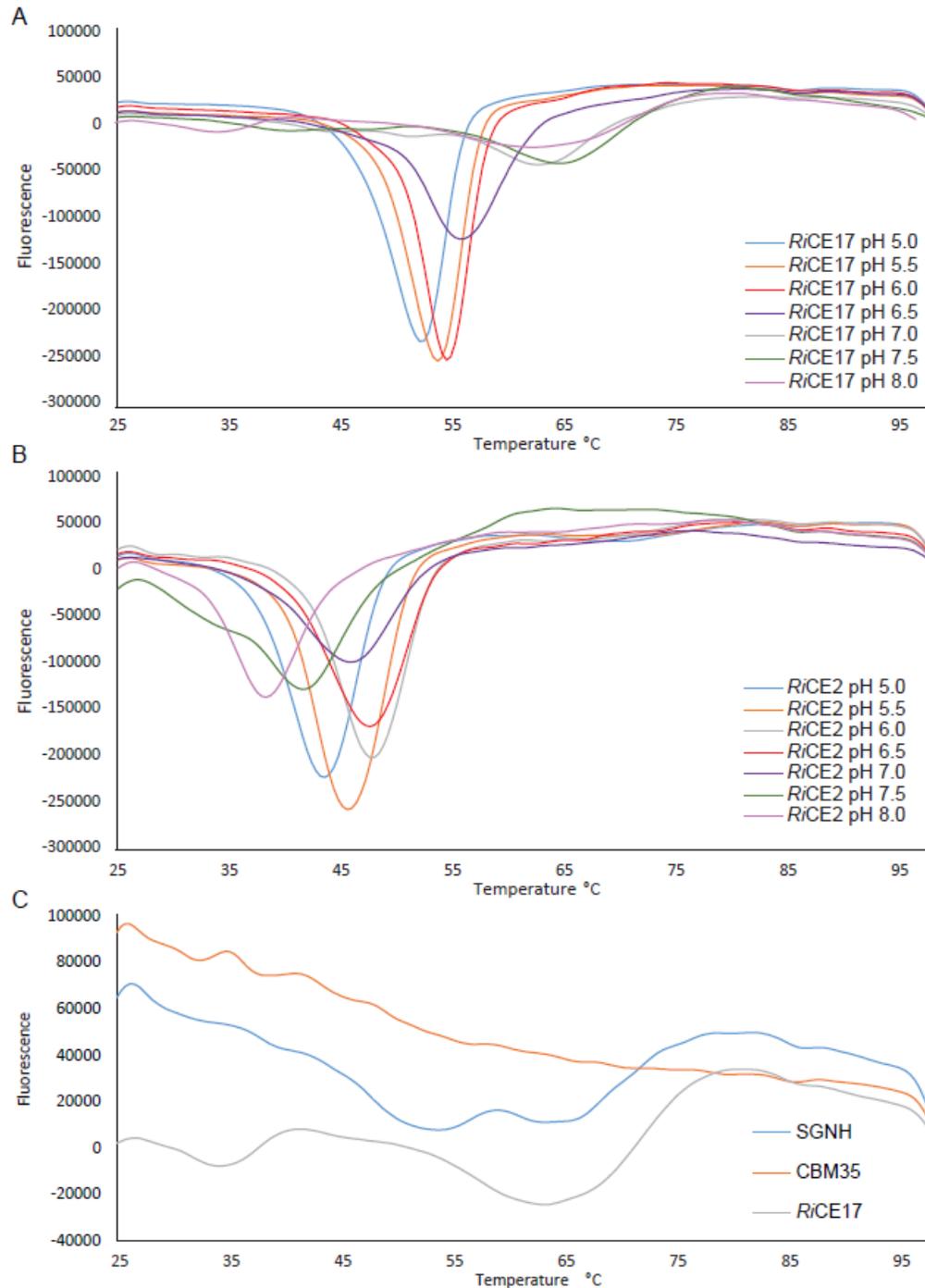
To monitor the effect of temperature and pH on acetyl migration, 2 mg transacetylated mannotriose was dissolved in 500  $\mu\text{L}$  40 mM phosphate buffer pD 5.9 (99.9% D<sub>2</sub>O) and it served as a control sample where 1D proton and 2D <sup>13</sup>C

heteronuclear single quantum coherence (HSQC) with multiplicity editing spectra were recorded. The sample was split into 3 samples of 160  $\mu\text{L}$  each and transferred into 3mm NMR tubes. Hereafter, the first sample was heated to 60°C for 60 min. In the second sample pD was adjusted pD 7.4 by adding 20  $\mu\text{L}$  of 250 mM phosphate buffer pD 8.0 and in the third sample the pD was adjusted pD 7.4 by adding 20  $\mu\text{L}$  of 250 mM phosphate buffer pD 8.0 and heated to 60 °C for 60 min. A 1D proton and 2D  $^{13}\text{C}$  HSQC spectra were recorded at 25 °C for each of the samples.

All homo- and heteronuclear NMR experiments were recorded on a BRUKER AVIIIHD 800 MHz (Bruker BioSpin AG, Switzerland) equipped with a 5 mm cryogenic CP-TCI. All NMR recordings were performed at 37°C. For chemical shift assignment of *RiCE17* transacetylated mannotriose, the following spectra were recorded: 1D proton, 1D proton with presaturation during relaxation delay and 10 ms mixing time with spoil gradient, 2D double quantum filtered correlation spectroscopy (DQF-COSY), 2D total correlation spectroscopy (TOCSY) with 70 ms mixing time, 2D  $^{13}\text{C}$  HSQC, 2D  $^{13}\text{C}$  Heteronuclear 2 Bond Correlation (H2BC), 2D  $^{13}\text{C}$  HSQC- $^1\text{H}$ ,  $^1\text{H}$ ]TOCSY with 70 ms mixing time on protons and 2D heteronuclear multiple bond correlation (HMBC) with BIRD filter to suppress first order correlations. The water signal at 4.75 ppm (at 25°C, pH 5.5 (32)) was used as chemical shift reference for protons, while  $^{13}\text{C}$  chemical shifts were referenced indirectly, based on the absolute frequency ratios (33). The spectra were recorded, processed and analyzed using TopSpin 3.5 software (Bruker BioSpin AG, Switzerland).



**Fig. S1** Activities of the SGNH domain (orange trace) and a cocktail of SGNH domain with the CBM35 (green trace) as two separate peptides were tested on *p*NP acetate, in comparison with a complete *RiCE17* (blue trace). The conditions (pH 7.0 and 30°C) were selected based on the pH optimum recorded previously. Despite using a much higher protein concentration in the single-domain samples (403 nM of the SGNH domain, 606 nM of the CBM35, as compared to 100 nM of *RiCE17*), no activity was observed.



**Fig. S2.** Melt plots of derivative data from protein thermal shift assays. Lowest point in each plot signifies the temperature point at which the protein is denatured. (A) The highest melting temperature (65°C) for *RiCE17* was observed at pH 7.5. (B) The highest melting temperature (47°C) for *RiCE2* was observed at pH range 6.0- 6.5. (C) The SGNH domain appears to fold into a stable complex, with its melting temperature at 52°C, while the CBM35 appears to hold no particular structure (melting temperature measured at pH 8.0).

```

cov pid 1 [ ..... 1 100
consensus/100% .....
consensus/90% .....s..hh.thht+h.t..tp...l.sls.lGGsITG...u.....t.pYs...hthhtp...h...
consensus/80% .....uhhp.Gs..Rltthhp+spt..Gcp...l.slualGGsITpG...u.....s.....ptsYu.hshphapp...tF...s
consensus/70% .....htpulshGshhRltphhc+App..Gcc...l.olualGGsITpG...hu.....s.....ppsYAhshshphac...pF...s

cov pid 101 ..... 2 200
consensus/100% .....
consensus/90% .....t.....hphhpGluuTsS.hGhhr.pcl...l.....p...PDhlh1-FuVND.....ttp.....sa-ul
consensus/80% .....p...s...p...hphlpAGlGuTsS.hGhhrhppDl...L...t.....p...PDhlh1EFuVND.....tss...hhtpsYEul
consensus/70% .....p...s...p...lpalNAGlGuTsS.hGhhrhpcDl...L...p...h.....p...PDlVh1EFuVND.....tss..p..hhtpsYEul

cov pid 201 ..... 3 300
consensus/100% .....
consensus/90% .....lhphht...tsAlhhl...ps.h..sh..p...lu.hYt1Ph1ShTssl.....h.....t...sp.....h...tp.h...
consensus/80% .....lRphlt..tpAlh1l.shh..ps.h..sh..Qttht.lu.hVp1Ph1Shpssl...h.th....t...t...sp.....h.ph.pp.h...
consensus/70% .....lR+lLpttspPAVlllsh1hh.pssh..sh..Qphht.lGpYs1Ph1Shpssl...h.ph....t...t...Gp.....h.sh.pp.h...

cov pid 301 ..... 4 400
consensus/100% .....
consensus/90% .....D.hHPss.GHthhup.l..hhp.h.....t.....
consensus/80% .....D.hHPss.GHtlhAphl..hhpph.....p..t.....
consensus/70% .....D.hHPss.GHplhAphlt.hhcphh.....p..p.....

cov pid 401 ..... 5 500
consensus/100% .....
consensus/90% .....s..hh...t...t...a.t..hh..p.....
consensus/80% .....t...t...t...t...s..hh...t...s...ta.pshhph.pp.....t.....
consensus/70% .....t...p...t...p.....tP...hh...s...s...tatssphhptps.....hp...

cov pid 501 ..... 6 600
consensus/100% .....
consensus/90% .....s.....t.....t.....a.s.....W.
consensus/80% .....t.G.h.....s.....t.....h.p.....F.s.....W.
consensus/70% .....s.h...h...t.G.a.....t...s.t...p.....s.....h.s.....F.ts.....W.

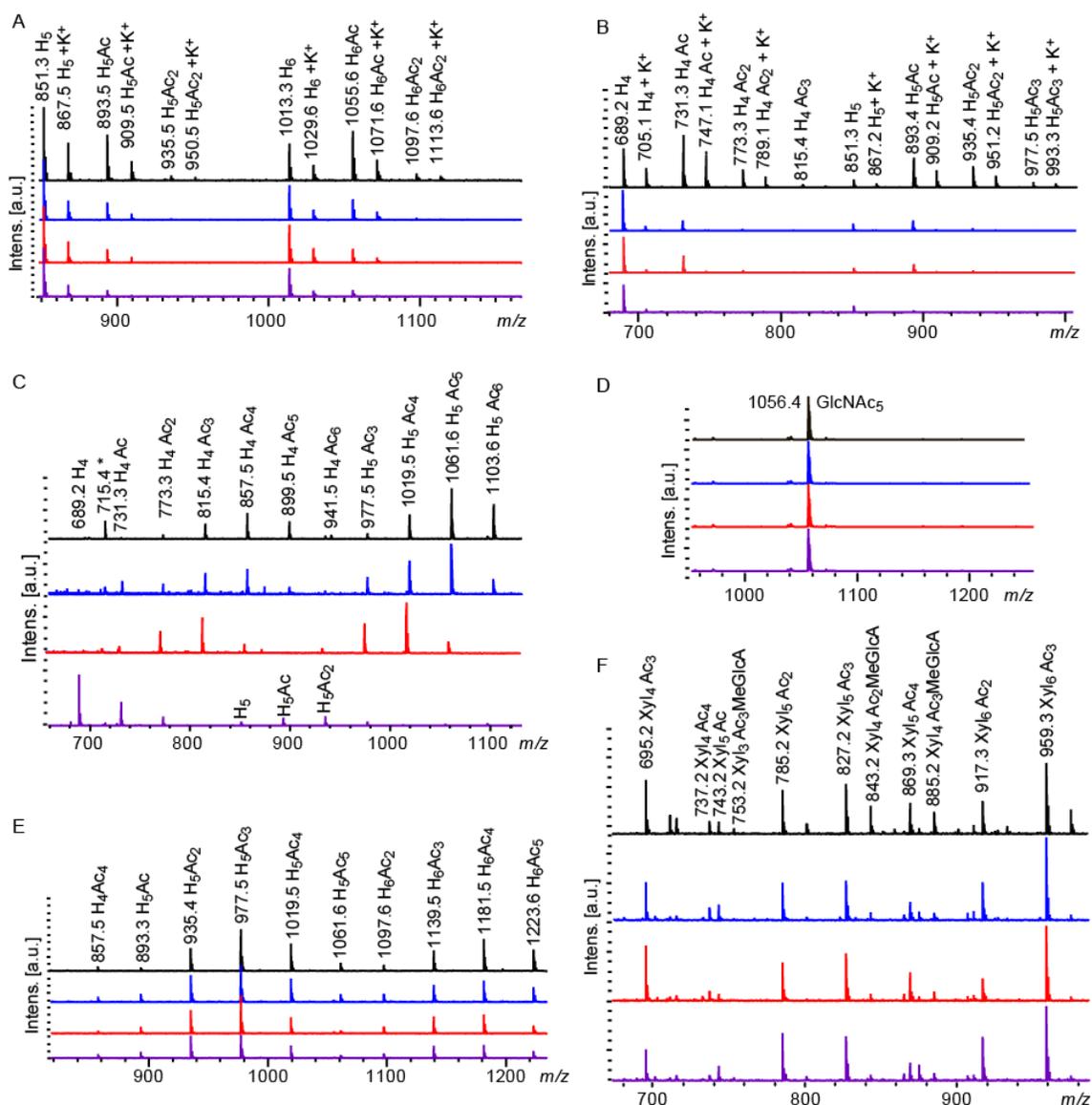
cov pid 601 ..... 7 700
consensus/100% .....
consensus/90% .....h...t.....t.....t.....h.....h..p..h...p.sp...h.hh..hh..t...p.....
consensus/80% .....h...h...p...t...t...t...s.....s.....h.....h..p..h...p.sp.t..l.hh..at..c..sh..p.....
consensus/70% .....h...t...p...p...p...p...s.....s.....p...h...p..h..c..l...p.sp.s..l.h.l..a...c..sh..pt..s..

cov pid 701 ..... 8 800
consensus/100% .....
consensus/90% .....h.hhl..Dst.....ss...s.W.tp.....lh.....t..a.lpl.h..t.....
consensus/80% .....hs...h...sthhl..Dsp...h...hss..h...ts.W.sph..h...lhpt..t..t.pHhlc1ph.ttppt...t.....
consensus/70% .....hs...h...Aplhl..Dgc...hhh...h-s..hp...ps.W.sph..h...lhpp.tpt.t.pHplc1phhptpt..p.....

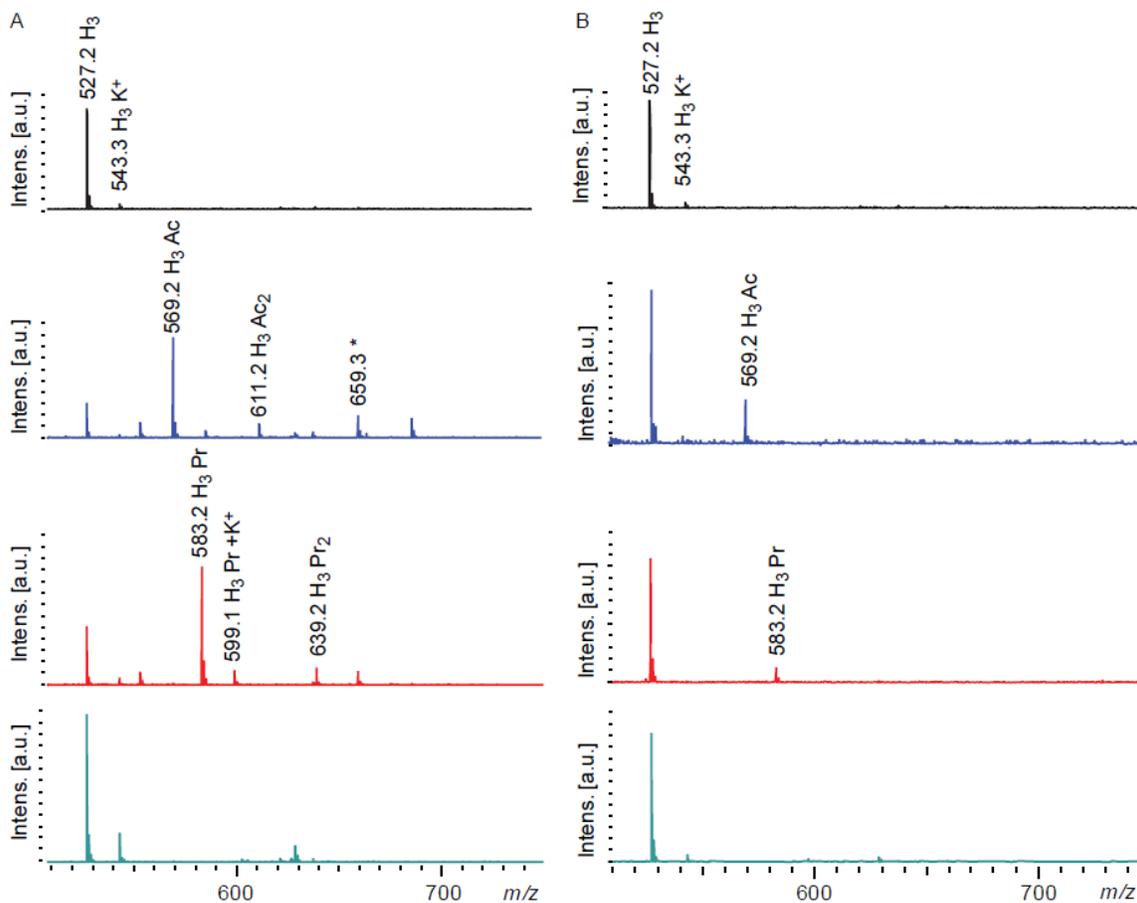
cov pid 801 ] 803
consensus/100% ...
consensus/90% ...
consensus/80% ...
consensus/70% ...

```

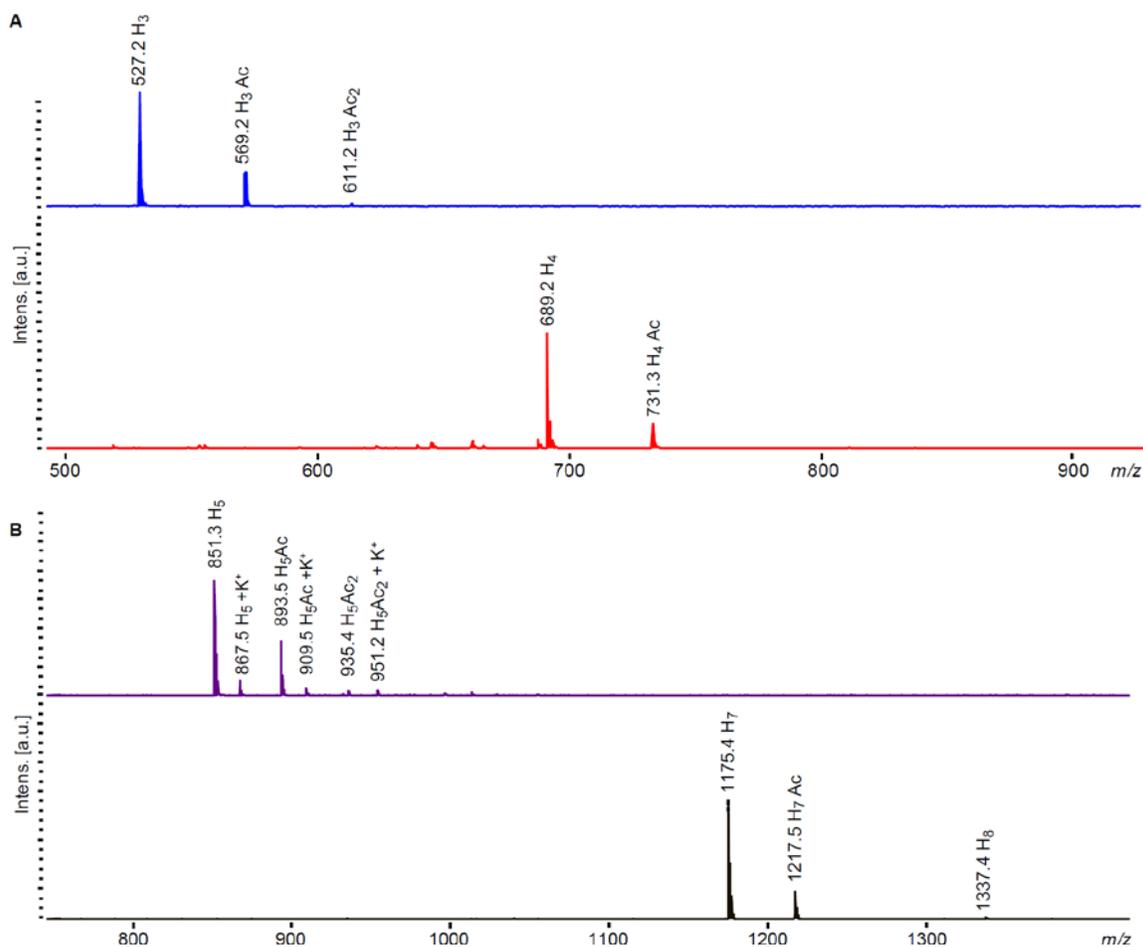
**Fig. S3.** MView generated consensus sequences of the 482 sequences found in UniprotKb database using an HMM of the *RiCE17*. Homologues of all three catalytic residues of *RiCE17* (Ser41 in the red box, Asp190 and His193 in the green box) as well as the Trp326 of the CBM35 domain (blue box) are present in the 90% consensus sequences. Capital letters signify conserved amino acids, small letters indicate conserved amino acid characters as described by Taylor et al. (34): l – aliphatic, a – aromatic, c – charged, h – hydrophobic, - – negative, p – polar, + – positive, s – small, u – tiny, t – turnlike.



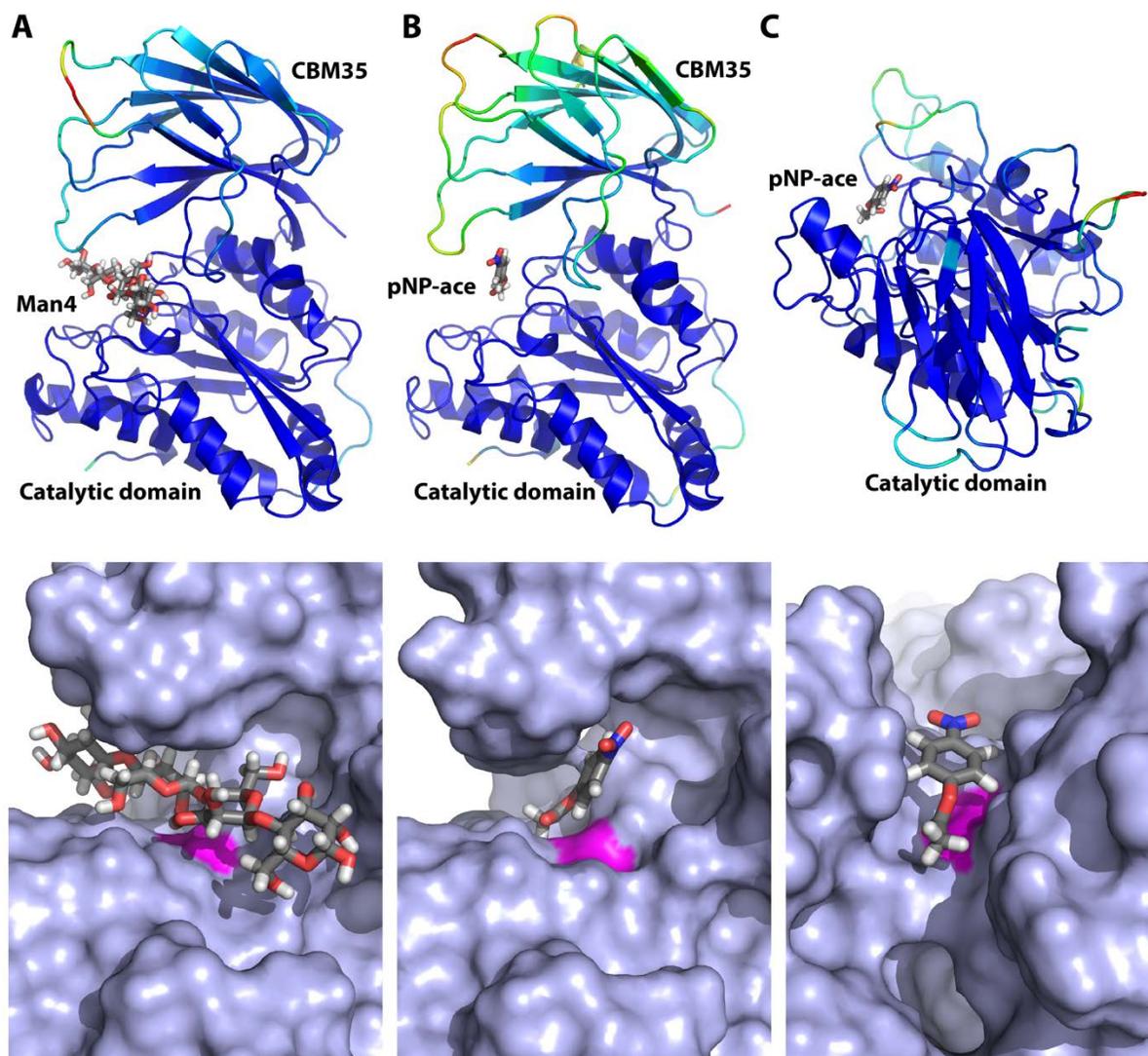
**Fig. S4.** MALDI-ToF MS traces of enzyme reactions on various substrates. (A) Konjac glucomannan deacetylated with *RiCE2* (in blue), *RiCE17* (in red), and both enzymes combined (in purple). (B) Chemically acetylated Konjac glucomannan deacetylated with *RiCE2* (in blue), *RiCE17* (in red) and both enzymes combined (in purple). (C) *Aloe vera* mannan treated with *RiCE2* (in blue), *RiCE17* (in red) and the two enzymes in combination (purple). (D) Chitopentaose (penta-*N*-acetylchitopentaose) (Megazyme, Ireland) treated with *RiCE2* (in blue), *RiCE17* (in red) and both enzymes combined (in purple), showing no signs of activity. (E) Acetylated cellulose oligosaccharides treated with *RiCE2* (in blue), *RiCE17* (in red) and both enzymes combined (in purple), the spectra show no signs of enzymatic activity. (F) Birch xylan oligosaccharides treated with *RiCE2* (in blue), *RiCE17* (in red) and both enzymes combined (in purple) showing no apparent activity on the xylooligosaccharides H-hexose, X-xylose, Ac-acetylations, Me-methylations, GlcA-glucuronic acid, GlcNAc<sub>5</sub> - Chitopentaose. All masses represent sodium adducts unless marked with K<sup>+</sup>, unlabeled peaks represent background signals from the sample matrix. Peak labelled 715.4\* in panel C is a persistent contaminant.



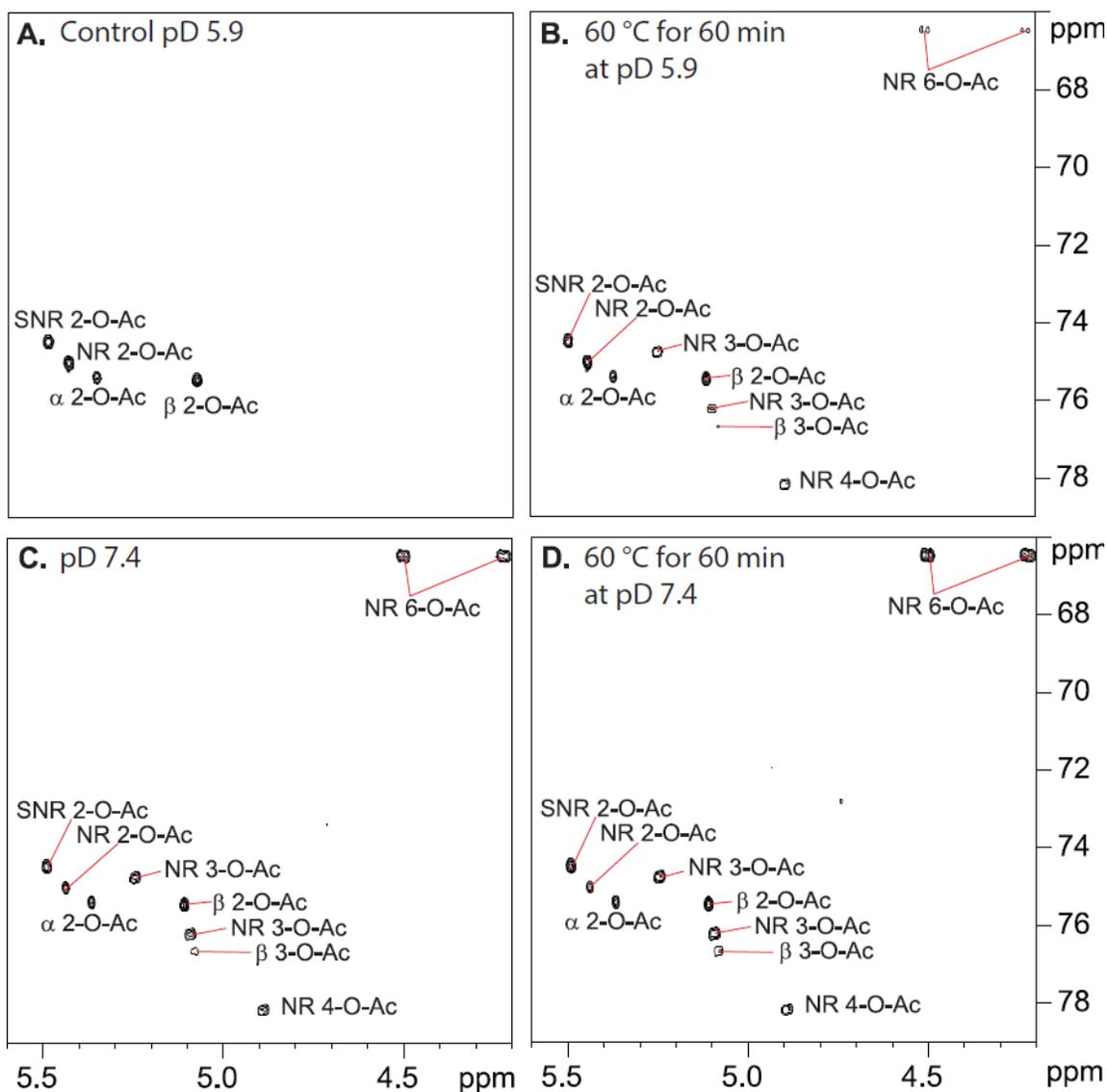
**Fig. S5.** MALDI-ToF MS spectra of transesterification reactions with mannotriose and vinyl esters. (A) *RiCE2* was able to transacetylate (in blue) and transpropylate (in red) mannotriose (in black), but not transbutyrylate (in green). (B) Similarly, *RiCE17* was able to transacetylate (in blue) and transpropylate (in red) mannotriose (in black), but was not able to transbutyrylate it (in green). Abbreviations: H- hexose, Ac- acetylation ( $m/z$  42), Pr- propylation ( $m/z$  56), K<sup>+</sup> signifies peaks of potassium adducts, all other  $m/z$  are of sodium adducts. Peak labelled \* in (A) ( $m/z$  659.3) is a persistent contaminant.



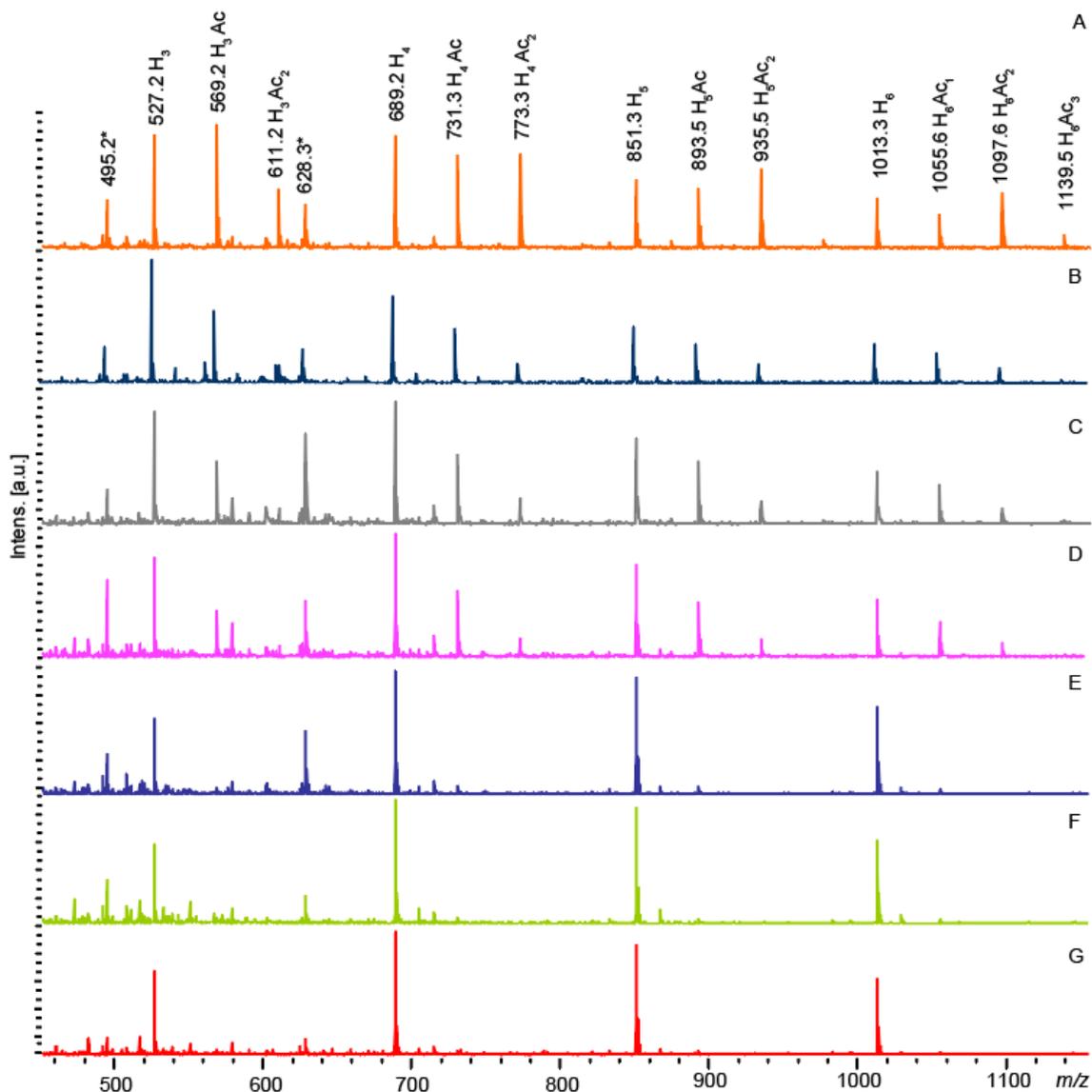
**Fig S6.** MALDI-ToF MS spectra of transacetylation reactions with *RiCE17* and vinyl acetate as acetyl donor. (A) *RiCE17* was able to transacetylate mannatriose producing single and double acetylated oligosaccharides (in blue). Transacetylation of 6<sup>1</sup>-α-D-galactosyl-mannatriose (in red) produced only single acetylated oligosaccharides. (B) Similarly, when transacetylating mannopentaose (in purple), *RiCE17* produced single and double acetylated oligosaccharides, while in a reaction with 6<sup>3</sup>, 6<sup>4</sup>-α-D-galactosyl-mannopentaose, only single acetylated oligosaccharides were observed (in black). The results indicate that *RiCE17* is active on oligosaccharides carrying galactosylations. This, together with the observation that *RiCE17* cooperatively with *RiCE2* completely deacetylates various galactosylated mannans, indicates that galactosylations (in all other subsites than 0) do not obstruct catalysis.



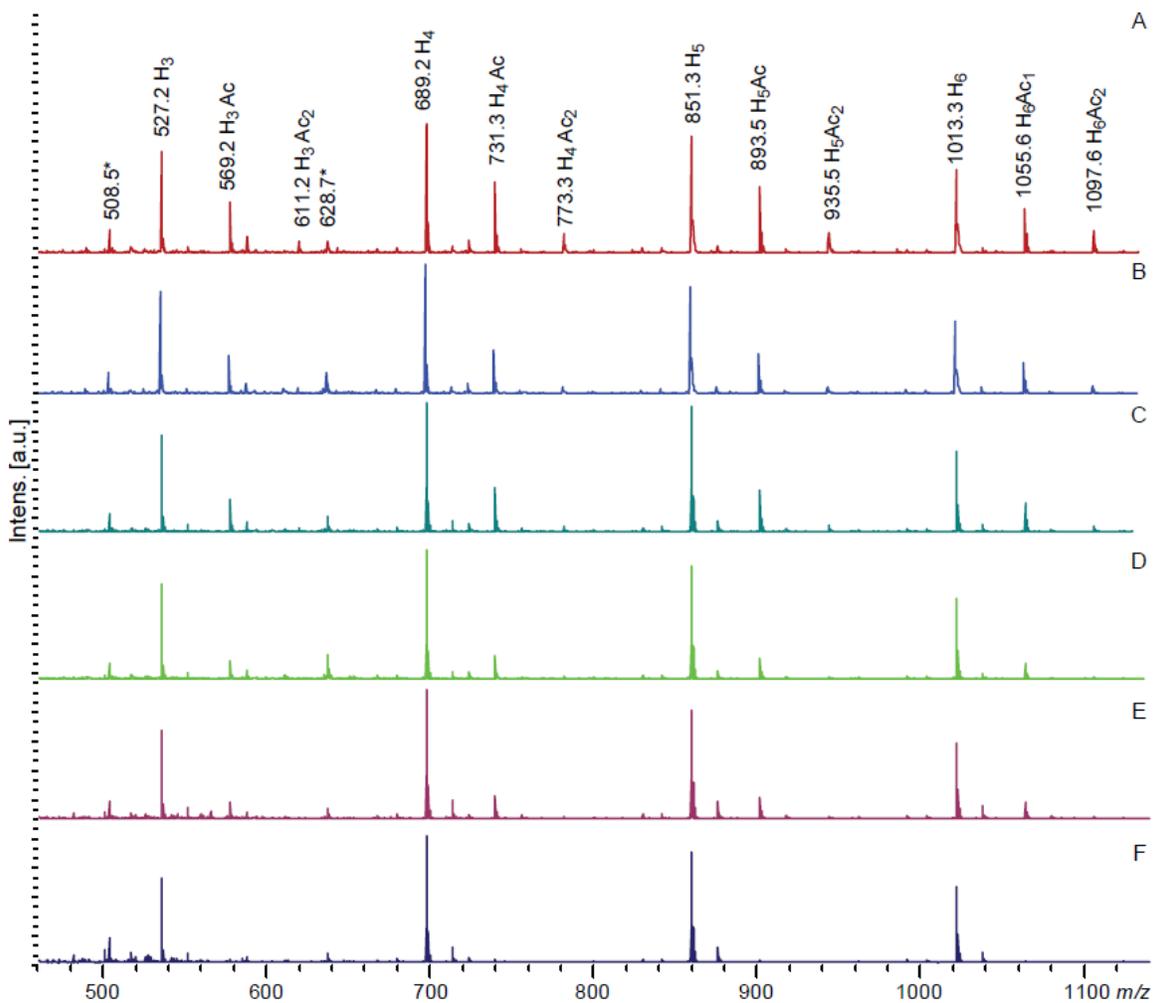
**Figure S7.** The effect of substrate binding on enzyme active site disorder. The models of *Ri*CE17-Man<sub>4</sub> (A), *Ri*CE17-pNP-acetate (B) and *Cj*CE2B-pNP-acetate (C) are displayed in cartoon representation. Close-up views of the enzyme active sites where the catalytic Ser residue is colored in magenta are displayed below the cartoon representation of the respective enzyme-substrate complexes. The cartoon representations are colored by B-factors that were obtained from the molecular dynamics simulations. Blue color shadings indicate low B-factors (small atomic positional fluctuations) and green to red color shadings indicate high B-factors (large atomic positional fluctuations).



**Fig. S8.** 2D  $^{13}\text{C}$  HSQC NMR spectra for area of interest of acetyl migration on *RiCE17* transacetylated mannotriose in 20 mM sodium phosphate. (A) *RiCE17* at pD 5.9 and 37°C acetylated mannotriose exclusively on the hydroxyl groups on carbon C2. Shift values after different treatment conditions; (B) after exposure to 60°C for one hour at pD 5.9, (C) after pD was increased to 7.4 by adding sodium phosphate and (D) after exposure to both high temperature and pD 7.4. In all conditions, the acetyl groups, initially present only as 2-*O*- acetylations migrated to other sites on the same sugar units. NR - nonreducing end mannose, SNR – Mannose internal in the mannotriose,  $\alpha$  and  $\beta$  – anomeric configuration of the reducing end mannose.



**Fig. S9.** MALDI-ToF MS spectra of Norway spruce GGM at each stage of pH induced migration. (A) The original sample had a prevalence of multiple acetylated oligosaccharides. (B) Overnight treatment with *RiCE17* removed a significant portion of the acetylations. (C) After *RiCE17* was removed by filtration, pH of the sample was adjusted to 7.4, but there was no apparent deacetylation in the process. (D) The sample retained the acetylations after incubation at 30°C overnight. (E) Treating the overnight incubated sample with *RiCE17* resulted in removal of a further portion of acetylations, with minor peaks for single acetylated oligosaccharides present. (F) Repeating the treatment from (C) did not affect the deacetylation. (G): Treating the sample as in (F) with *RiCE17* for the third time resulted in complete deacetylation.



**Fig. S10.** MALDI-ToF MS spectra of Norway spruce GGM at each step of temperature induced migration. (A) After the initial deacetylation and filtration, the pH of the sample adjusted to 5.9, and the sample was incubated at 60°C for one hour without a loss of acetylations in the process. (B) *Ri*CE17 treatment of the heat-treated sample removed more acetylations. (C) After *Ri*CE17 was removed by filtration, the sample pH was adjusted to 5.9, and the sample was incubated at 60°C for the second time, with no loss of acetylations. (D) After the second heat treatment, *Ri*CE17 removed a further proportion of acetylations, with peaks for double acetylated oligosaccharides disappearing completely. (E) The control sample after the third round of heat-induced migration (as in (C)). (F) After three rounds of heat-induced migration, the *Ri*CE17 treatment removed all acetylations.

Table S1: Assignment of chemical shifts.

Structural unit	Assignment							
	H-1; C-1	H-2; C-2	H-3; C-3	H-4; C-4	H-5; C-5	H-6; C-6	C=O	Ac-H; C
NRE-M2	4.91; 102.0	5.45; 75.0	3.84; 74.1	n.d	n.d	n.d	176.5	2.17; 23.2
NRE-M3	4.90; 96.8	4.18; 71.5	5.25; 74.7	4.06; 76.5	n.d	n.d	176.1	2.14; 23.3
NRE-M4	n.d	n.d	3.98;-	4.89; 78.14	n.d	n.d	176.1	2.14; 23.3
NRE-M6	n.d	n.d	n.d	3.62; 76.7	3.64; 69.4	4.50, 4.22; 66.4	177.0	2.12; 23.2
SRE-M2	4.93; 102.04	5.50; 74.4	4.00; 73.0	n.d	n.d	n.d	176.3	2.17; 23.2
SRE-M3	4.88; 96.6	4.17; 71.4	5.10; 76.2	4.00; 76.1	n.d	n.d	176.0	2.14; 23.3
$\alpha$ -M2	5.19; 94.2	5.37; 75.4	3.94; 73.2	n.d	n.d	n.d	176.7	2.15; 23.3
$\beta$ -M2	5.06; 95.3	5.12; 75.5	4.159; 70.4	n.d	n.d	n.d	176.2	2.18; 23.3

n.d: not determined

Table S2: Primers used in this study. 5' extension sequences used for molecular cloning are underlined.

Gene	Primer (5' -3')
CE17_up	F: <u>TTAAGAAGGAGATATACTATGGAATATCAAATTAATACGAAAACGGC</u>
CE17catD_rev1	R: <u>AATGGTGGTGATGATGGTGCGCCTGTCCGTTGCATCTTCTG</u>
CE17cbm_up	F: <u>TTAAGAAGGAGATATACTATGGATTATCCGGCACCTCTCAC</u>
CE17_down	R: <u>AATGGTGGTGATGATGGTGCGCTTTTTCAGAGGAACCAATGACAGAC</u>

Table S3. Crystal data, data collection, and refinement statistics.

<i>Crystal data</i>		
	<i>RiCE17</i>	<i>RiCE17-mannopentaose</i>
Space group	P 1 2 <sub>1</sub> 1	
Crystal parameters	a = 75.12, b = 135.52, c = 85.13 $\alpha = 90, \beta = 115.09, \gamma = 90$	a = 75.48, b = 136.69, c = 85.41 $\alpha = 90, \beta = 113.86, \gamma = 90$
<i>Data collection</i>		
X-ray source	ESRF, ID23-1	ESRF, ID23-1
Resolution (Å) <sup>a</sup>	48.2-1.75 (1.78-1.75)	48.6-2.40 (2.53-2.40)
Wavelength (Å)	0.97531	0.97625
Temperature (K)	100	100
Number of unique reflections	155949 (7703)	61043 (8942)
Completeness <sup>a</sup>	100 (99.9)	98.7 (99.2)
Redundancy <sup>a</sup>	6.7 (6.4)	3.5 (3.5)
CC half <sup>a</sup>	0.998 (0.815)	0.995 (0.713)
I/ $\sigma(I)$ <sup>a</sup>	11.5 (2.2)	8.2 (2.0)
R <sub>sym</sub> <sup>b</sup>	0.079 (0.631)	0.097 (0.593)
<i>Refinement statistics</i>		
R <sub>cryst</sub> <sup>c</sup>	0.190	0.181
R <sub>free</sub> <sup>d</sup>	0.232	0.253
Wilson B-factor (Å <sup>2</sup> )	25.4	42.1
Ramachandran plot, in most favored/other allowed regions (%)	96/4	95/5
Standard Uncertainty (Maximum Likelihood):	0.093	0.224
Added waters	808	487
<b>PDB code</b>	<b>6HFZ</b>	<b>6HH9</b>

<sup>a</sup> Values for outer shell in parenthesis,

<sup>b</sup>  $R_{sym} = \sum |I - \langle I \rangle| / \sum I$ .

<sup>c</sup>  $R_{cryst} = \sum (|F_{obs}| - |F_{calc}|) / \sum |F_{obs}|$

<sup>d</sup>  $R_{free}$  is the  $R_{cryst}$  value calculated on the 5 % reflections excluded for refinement.

## References:

1. Potter SC, *et al.* (2018) HMMER web server: 2018 update. *Nucleic Acids Res* 46(W1):W200-W204.
2. The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45(D1):D158-D169.
3. Hannuksela T & Herve du Penhoat C (2004) NMR structural determination of dissolved O-acetylated galactoglucomannan isolated from spruce thermomechanical pulp. *Carbohydrate research* 339(2):301-312.
4. Lassfolk R, *et al.* (2019) Acetyl Group Migration across the Saccharide Units in Oligomannoside Model Compound. *Journal of the American Chemical Society* 141(4):1646-1654.
5. Xu C, *et al.* (2010) Acetylation and characterization of spruce (*Picea abies*) galactoglucomannans. *Carbohydrate research* 345(6):810-816.
6. Aslanidis C & de Jong PJ (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* 18(20):6069-6074.
7. La Rosa SL, *et al.* (2019) The human gut Firmicute *Roseburia intestinalis* is a primary degrader of dietary beta-mannans. *Nature Communications* 10.
8. Kabsch W (2010) Xds. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):125-132.
9. Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallographica Section D* 69(7):1204-1214.
10. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* 66(2):213-221.
11. Terwilliger TC, *et al.* (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallographica Section D* 65(6):582-601.
12. Terwilliger TC, *et al.* (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica Section D* 64(1):61-69.
13. McCoy AJ, *et al.* (2007) Phaser crystallographic software. *Journal of Applied Crystallography* 40(4):658-674.
14. Vagin AA, *et al.* (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D* 60(12 Part 1):2184-2195.
15. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallographica Section D* 66(4):486-501.
16. Vestby LK, Moretro T, Ballance S, Langsrud S, & Nesse LL (2009) Survival potential of wild type cellulose deficient *Salmonella* from the feed industry. *Bmc Vet Res* 5.
17. La Rosa SL, *et al.* (2019) Wood-Derived Dietary Fibers Promote Beneficial Human Gut Microbiota. *mSphere* 4(1):e00554-00518.
18. Anandkrishnan R, Aguilar B, & Onufriev AV (2012) H++3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic acids research* 40(W1):W537-W541.
19. Hanwell MD, *et al.* (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminformatics* 4.
20. Neese F (2018) Software update: the ORCA program system, version 4.0. *Wires Comput Mol Sci* 8(1).
21. Becke AD (1993) Density-Functional Thermochemistry .3. The Role of Exact Exchange. *J Chem Phys* 98(7):5648-5652.

22. Stephens PJ, Devlin FJ, Chabalowski CF, & Frisch MJ (1994) Ab-Initio Calculation of Vibrational Absorption and Circular-Dichroism Spectra Using Density-Functional Force-Fields. *J Phys Chem-Us* 98(45):11623-11627.
23. Weigend F & Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* 7(18):3297-3305.
24. Weigend F, Furche F, & Ahlrichs R (2003) Gaussian basis sets of quadruple zeta valence quality for atoms H-Kr. *J Chem Phys* 119(24):12753-12762.
25. Grimme S, Ehrlich S, & Goerigk L (2011) Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J Comput Chem* 32(7):1456-1465.
26. D.A. Case IYB-S, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman (2018) AMBER 2018. *University of California, San Francisco*. .
27. Izadi S, Anandkrishnan R, & Onufriev AV (2014) Building Water Models: A Different Approach. *J Phys Chem Lett* 5(21):3863-3871.
28. Darden T, York D, & Pedersen L (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* 98(12):10089-10092.
29. Ryckaert JP, Ciccotti G, & Berendsen HJC (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys* 23(3):327-341.
30. Salomon-Ferrer R, Gotz AW, Poole D, Le Grand S, & Walker RC (2013) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J Chem Theory Comput* 9(9):3878-3888.
31. Roe DR & Cheatham TE (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* 9(7):3084-3095.
32. Cavanagh J, Fairbrother WJ, Palmer AG, & Skelton NJ (2007) *Protein NMR Spectroscopy (Second Edition)* (Academic Press, Burlington) pp vii-x.
33. Zhang H, Neal S, & Wishart DS (2003) RefDB: a database of uniformly referenced protein chemical shifts. *J Biomol NMR* 25(3):173-195.
34. Taylor WR (1986) The classification of amino acid conservation. *Journal of theoretical biology* 119(2):205-218.