SI Appendix

SI Materials and Methods

Materials

Polygalacturonic acid (PGA) and rhamnogalacturonan I (RGI) were purchased from Megazyme (Bray, Ireland) and two different pectins, PECF (pectin, esterified from citrus fruit; P9561) and pectin (P7536), chondroitin sulfate A (C9819), as well as sodium metaperiodate, biocytin, and Tween-20 were purchased from Sigma-Aldrich (St. Louis, USA). Porcine gastric mucin (PGM) type III and Dulbecco's phosphate buffered saline (PBS) were purchased from Sigma Aldrich (Dorset, UK). PGM was further purified (pPGM) as previously described (1). Polyclonal antiserum against immobilized metal affinity chromatography (IMAC)-purified His_6-$SRRP_{53608}$-BR was raised in rabbits by BioGenes GmbH (Berlin, Germany) and provided at a titre of >1:200000. Blocking reagent used in cell and tissue binding assays was from Perkin Elmer (Boston, MA, USA). Mouse monoclonal antibody to MUC5AC and Texas Red-conjugated goat polyclonal antibody to mouse IgG were from Abcam (Cambridge, UK), Alexa Fluor 488-conjugated goat anti-Rabbit IgG secondary antibody from Thermo Fischer Scientific (Eugene/OR, US). Fluorescein- and rhodamine-labelled Wheat Germ Agglutinin lectins (WGA-FITC and WGA-Rh) and Vectashield were from Vector laboratories (Peterborough, UK). DAPI was from Life Technologies, O.C.T. Compound from VWR and Hydromount from National Diagnostics (Atlanta/GA, USA).

Bioinformatics analyses

SignalP 4.1, with the D-score cutoff set for SignalP 3.0 sensitivity, was used to predict the signal peptide cleavage site in the relatively long N-terminal leader sequence of up to 100 aa in SRRPs (http://www.cbs.dtu.dk/services/SignalP/) (2). Domains within each SRRP were assigned based on sequence comparisons and previously published analyses for particular proteins. Defined N2 (BR) domain protein sequences were aligned as a single FASTA file using Clustal Omega (1.2.4) multiple sequence alignment (3) and an aa % identity matrix was

subsequently generated by Clustal 2.1 (http://www.ebi.ac.uk/Tools/msa/clustalo/). A Neighbour-Joining phylogram of the Clustal Omega BR alignment was visualised using EvolView (http://www.evolgenius.info/evolview/). In a few cases, especially for some SRRP fragments encoded by pseudogenes, different reading frames to those annotated in a genome were translated to give the appropriate SRRP domains and in certain genomes, the mis-annotation of some SRRP pseudogenes was corrected by translating an extended CDS to give a full-length SRRP.

Cloning of $SRRP_{53608}$-BR and $SRRP_{100\text{-}23}$-BR

L. reuteri ATCC 53608 and 100-23C were grown without shaking in 10 ml MRS broth at 37°C for 16-18 h. Cells were harvested by centrifugation, washed with sterile ultrapure water, resuspended to a cell density of 1.0 A_{600nm} and frozen prior their use as templates in PCRs. DNA sequences encoding the two BR domains were PCR-amplified using 10 µL thawed, washed cells with primers 1105N-for (5'-*AAGTTCTGTTTCAGGGCCCG*GCGAGTACAAGTCATTTAGAAGAAAATGG-3') and 1105N-rev (5'-*ATGGTCTAGAAAGCTTTA*TTCATTACTACTGATAAGTTCTG-3') for $SRRP_{53608}$-BR and primers 70902N-for (5'-*AAGTTCTGTTTCAGGGCCCG*AGTACCTTATCTGGACTCGATAATAATG-3') and 70902N-rev (5'-*ATGGTCTAGAAAGCTTTA*GGAAATCACCGTGCTACCTTGATATACAACAG-3') for $SRRP_{100\text{-}23}$-BR in a 50 µL reaction at an annealing temperature of 54 °C with HotStarTaq DNA polymerase (Qiagen) and a 72 °C extension time of 2 min for 30 cycles. Primer 5' sequences shown in italics are pOPINF sequences used in subsequent cloning (4). The amplified regions were aa positions 219-676 of $SRRP_{53608}$ ($SRRP_{53608}$-BR defined as 223-668 aa) and aa positions 198-757 of $SRRP_{100\text{-}23}$ ($SRRP_{100\text{-}23}$-BR defined as 202-687 aa). The purified DNA amplicon fragments (Qiaquick PCR purification kit) were cloned using In-Fusion technology (Clontech), following the manufacturer's instructions, into *Hin*dIII/*Kpn*I double-digested vector pOPINF which contains a N-terminal His_6-tag (4) and transformed into *E. coli* cloning strain

DH5α. Plasmids were DNA sequenced to confirm the integrity of the BR inserts and recombinant vectors transformed into expression strain *E. coli* BL21(DE3) (Novagen).

Mutagenesis of $SRRP_{53608}$-BR

A modified method by Liu and Naismith (2008) (5) was used for the creation of the single mutants K377A (K377A-for 5'-GTGTGGAGTGCAACTGGTGCTGGTGTTAAAACGTTGAACTTAGTG-3'; K377A-rev 5'-CAGTTGCACTCCACACTAATGTCCCACCGTAAGCC-3') and R512A (R512A-for 5'-GATAATGTAGCAATCATACCACAACTTGAAAATATCTTTACACGAGGAAATATTG-3'; R512A-rev 5'- GTGGTATGATTGCTACATTATCAACCTTAGCTGACATTTTAACAGCC-3'), and the deletion mutant ΔF411-T422 (Del-for 5'-GAGGCCGCCTTATATGTTTCTAATGCCATTAATATTGCAGAAAATGCTAATG-3'; Del-rev 5'- CATATAAGGCGGCCTCTGTACCACACGACTGTCCATC-3'). Primers were supplied by Eurogentec (Liège, Belgium). 20 µL PCR reactions consisted of 500 µM of each primer, 200 µM of each dNTP, 2-10 ng DNA template and 0.2 µL Q5 Hot-Start High-Fidelity DNA Polymerase, with its corresponding buffer (NEB, Hitchin, UK). PCR was performed in the T100^{TM} Thermal Cycler (Bio-Rad, Hemel Hempstead, UK) with the following steps: initial template denaturation at 98 °C for 30 s; 14 amplification cycles, each with denaturation at 98 °C for 10 s, the first annealing cycle (at 72 °C for K377A, 71 °C for Del, and 70 °C for R512A) for 20 s and extension at 72 °C for 7 min; the cycles were ended by a second annealing step (at 61 °C for K377A, 60 °C for Del and 59 °C for R512A) for 15 s and a final extension at 72 °C for 10 min. The PCR products were then treated with 5 units of Fermentas *Dp*nI (Thermo Scientific, Loughborough, UK) at 37 °C for 1 h, and then used for transformation into *E. coli* BL21 (DE3) cells.

Purification of recombinant *Lr*SRRP-BR

Cultures expressing wild-type and mutant $SRRP_{53608}$-BR, and $SRRP_{100-23}$-BR recombinant proteins were grown in 1 L flasks. Cells were grown with shaking in LB broth containing 1% (w/v) glucose and 100 µg.mL^{-1} carbenicillin or ampicillin at 37 °C until they reached an OD_{600nm}

of 0.8, then induced with the addition of 1 mM IPTG for a further 3 h at 37 °C, or 18 h at 20 °C. Harvested cell pellets were re-suspended at room temperature in 2 v/w of lysis buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 10% glycerol, 10 mM imidazole) and lysed twice using a cell disruptor (Constant Systems Ltd., Daventry, UK), which was pre-chilled to 5 °C. The first cycle was performed at 30 kpsi and the second at 20 kpsi, and the lysate was collected on ice. The clarified cell extracts were applied onto an Ni-NTA agarose IMAC resin (Qiagen, Hilden, Germany) packed in a gravity flow column, at 2-3 $mL.min^{-1}$; 1 mL of IMAC resin per L of cell culture was used. His-tag bound proteins were washed with 10 times bed volume of wash buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 10% glycerol, 30 mM imidazole), and eluted in 1.5 mL fractions with high imidazole elution buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 10% glycerol, 250 mM imidazole). IMAC-purified and dialysed His-tagged SRRP-BRs were purified by size exclusion gel filtration (SEGF) chromatography through Superdex 200 (16/600) in 20 mM Tris-HCl (pH 8.0), 100 mM NaCl running buffer at 1 $mL.min^{-1}$, 1.5 column volume, to separate monomeric and oligomeric forms which were confirmed by native gel electrophoresis. Fractions with the highest concentration of monomeric proteins were collected for buffer exchange into a protease cleavage buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 10% glycerol), using a 26/10 HiPrep Desalting column (GE Healthcare, Little Chalfont, UK) on an ÄKTApurifier UPC 100 chromatography system (GE Healthcare, Uppsala, Sweden) for overnight His-tag removal by 3C-GST protease (1 mg of protease per 50 mg of recombinant protein). The protease treated protein solutions were passed through an IMAC resin to capture the cleaved His-tag from the tag-free recombinant protein in the column flow through, which was concentrated to 2-4 mL for SEGF chromatography on a HiLoad™ 16/600 Superdex™ 200 pg prep grade column (using buffer: 20 mM Tris, pH 7.9, 0.15 M NaCl, 2.5 mM DTT). Proteins were eluted at between 80 and 90 mL retention volumes and concentrated to 10 $mg.mL^{-1}$ for subsequent structural and biochemical studies. The secondary structure of the recombinant $SRRP_{53608}$-BR was investigated by circular dichroism (CD) at the two pH's used for the various binding assays performed in this study, pH 4.0 and pH 7.4.

Production of SeMet $SRRP_{53608}$-BR

Overnight starter culture of *E. coli* BL21(DE3) expressing $SRRP_{53608}$-BR was grown in LB broth containing 100 µg.mL^{-1} ampicillin at 37 ℃, pelleted and washed three times with PBS to remove traces of the organic media prior to inoculation into minimal media. An amount equivalent to 5% (v/v) of the original overnight culture was inoculated into a litre of minimal essential media (12.8 g $Na_2HPO_4.7H_2O$, 3.0 g KH_2PO_4, 1.0 g NH_4Cl, 2% w/v glucose, 0.03% w/v $MgSO_4$, 0.001% w/v $Fe_2(SO_4)_3$ and 0.001% w/v thiamine), supplemented with 1.1 g.L^{-1} Glucose-free SelenoMethionine Nutrient Mix (Molecular Dimensions, UK) and 100 µg.mL^{-1} ampicillin, and incubated at 37 ℃ with shaking at 200 rpm until an OD_{600} of about 0.3 was obtained. Amino acid supplement (0.01% w/v each of lysine, phenylalanine and threonine; 0.005% each of w/v leucine, isoleucine and valine) was then added to the cultures, with further incubation for 30 min, after which SeMet was added (0.1 g.L^{-1}). The cultures were grown to an OD_{600} of 0.5-0.6, where after protein production was induced with 1 mM IPTG at 20 ℃ for 14 h. Cell harvesting, preparation of clarified extracts, and purification of SeMet labelled $SRRP_{53608}$-BR was performed as described for native $SRRP_{53608}$-BR.

Biotinylation of $SRRP_{53608}$-BR

Biotinylation of SRRP. For binding measurements by biolayer interferometry, $SRRP_{53608}$-BR was biotinylated using EZ-linkNHS-LC-LC-biotin (Thermo Fisher Scientific, Loughborough, UK) according to the 1:1 biotin:protein labelling technical note available from Pall BioForte. Briefly, a 1 mg.ml^{-1} solution of $SRRP_{53608}$-BR in PBS (pH 7.4) was incubated with an equimolar amount of NHS-LC-LC-biotin for 30 min at room temperature, with occasional agitation. Free biotin was then removed through the use of a PD-10 buffer exchange column (GE Healthcare, Buckinghamshire, UK). The resulting solution was diluted to 100 nM in PBS and aliquoted for later use in Octet biolayer interferometry binding assays.

Protein Crystallization, Data Collection, and Structure Determination

Crystallisation trials were dispensed in 96 well crystallisation trays (Molecular Dimensions, Newmarket, UK) for sitting drop vapour diffusion, using the Gryphon crystallisation robot (Art

Robbins Instruments, California, USA). For crystallisation by *in-situ* limited proteolysis, $SRRP_{53608}$-BR (7 $mg.mL^{-1}$) and SeMet $SRRP_{53608}$-BR (15 $mg.mL^{-1}$) were mixed with 1:500 (w/w) α-chymotrypsin, and $SRRP_{100-23}$-BR (9 $mg.mL^{-1}$) was mixed with 1:700 (w/w) thermolysin, and incubated at room temperature for 1.5 h prior to crystallisation screening with the PEG/Ion 1 and 2 screens (Hampton Research, California, USA), at 1:1 v/v ratio of protein sample to crystallisation buffer, with a total drop size of 2 µL. The plates were incubated at 22 ℃ and examined regularly for crystal growth. Crystals of $SRRP_{53608}$-BR were obtained in 0.1 M sodium malonate/pH 6.0, 12% PEG 3350, within 3 weeks. SeMet $SRRP_{53608}$-BR crystals were grown in 0.03 M Citric acid, 0.07 M Bis-Tris propane/pH 7.6, 20% PEG 3350, within 3 days. $SRRP_{100-23}$-BR crystals grew in 0.2 M Sodium tartrate dibasic dehydrate/pH 7.3, 20% PEG 3350, in under 3 days. Crystals were soaked in cryo-protectant (20% glycerol supplemented to the respective crystallisation mother liquors) and frozen in liquid nitrogen for data collection, at beamlines I03 and I04 of the Diamond Light Source (Harwell, UK). An Se edge scan was performed on a crystal of SeMet $SRRP_{53608}$-BR, giving f": 5.52 and f': -8.53, which indicated the presence of Se atoms in the crystal lattice. 3600 X-ray diffraction images were then collected at wavelength 0.9792 Å, exposure at 0.100 s and 30% transmission which were integrated and scaled by XIA2 (6) to a maximum resolution of 2.73 Å. 11 Se atomic positions and the handedness of SeMet $SRRP_{53608}$-BR was determined by SHELXD and SHELXE, respectively. Buccaneer (7) was then used to build a partial model of SeMet $SRRP_{53608}$-BR with R_{factor} 0.2334 and R_{free} 0.2887, which was used as the search model for molecular replacement of the high resolution native $SRRP_{53608}$-BR data. 1800 X-ray diffraction images for $SRRP_{53608}$-BR and $SRRP_{100-23}$-BR crystals were collected at a wavelength of 0.9795 Å, exposure of 0.100 s and 100% transmission. Diffraction images were indexed and integrated manually with MOSFLM (8), with merging and scaling performed by SCALA (9) in the CCP4 suite (10). Space groups were verified on POINTLESS (104). The phase of $SRRP_{100-23}$-BR was determined using Phaser (11), via molecular replacement with the native $SRRP_{53608}$-BR structure as a search model. All models were subsequently built automatically with ARP/wARP (12) and manually with Coot (13) and each building cycle was subjected to

structural refinements (rigid body and restrained) with REFMAC5 (14) and the final refinements were performed using Phenix (15). Structures were validated with MolProbity (16) and the complete data collection and refinement statistics are provided in **Table 1**.

Molecular Dynamics (MD) simulations

Protonation of Crystal Coordinates. Hydrogens were added to the coordinates of the $SRRP_{53608}$-BR crystal structure using the PROPKA (17) utility of Schrödinger's Maestro software suite (Maestro, Schrödinger, LLC, New York, NY, 2017) to predict the protonation state of amino acid sidechains at pH 4.0 and pH 7.4 respectively. The Protein Preparation Wizard was then used to optimise intramolecular hydrogen bonding and a restrained minimisation was performed using the OPLS3 forcefield, converging heavy atoms to a RMSD of 0.3 Å.

System setup. Both sets of coordinates were parameterised with the AMBER ff14SB forcefield (18). They were then separately solvated in a truncated octahedron of TIP4PEW water (19) scaled such that no protein atom was less than 10 Å from the boundary. Either 12 Cl^- or seven Na^+ ions were added to the system for simulation at pH 4.0 and pH 7.4 respectively, in order to neutralise overall charge. Ion coordinates were automatically assigned in *tleap* in order to minimise electrostatic potential and steric clashes.

Minimisation. Each system was first subjected to two rounds of conjugate gradient minimisation, the first with all protein atoms restrained with a 100 $kcal.mol^{-1}.Å^{-2}$ restraint and then with the whole system unrestrained. In both cases minimisation was considered complete when the root mean square of the Cartesian elements of the gradient was less than 10^{-4} $kcal.mol^{-1}.Å^{-1}$. A non-bonded cutoff of 8 Å was used.

Molecular dynamics. Each system was slowly heated at constant volume to 300 K over a period of 200 ps using the weak-coupling algorithm with a coupling constant of 1 ps. Each system was then equilibrated for a further 200 ps using the isobaric-isothermal ensemble. Pressure and temperature were maintained at 1 atm using the Berendsen algorithm with a relaxation time of 1 ps and 300 K using Langevin dynamics with a collision frequency of 2 ps^-

[1] respectively. Production dynamics were then run for a total of 1 μs using the same settings as during the equilibration period. All stages used a non-bonded cutoff of 8 Å, employed periodic boundary conditions and constrained bonded interactions involving hydrogen atoms using the SHAKE algorithm, allowing a timestep of 2 fs. Frames were saved at 100 ps intervals and analysed in *cpptraj*.

Bio-Layer Interferometry (BLI)

The binding affinity assay of $SRRP_{53608}$-BR to PGA, RGI, and PECF was performed using the Octet RED96 system (ForteBio, Freemont, California, USA). Assays were performed in black 96-well plates (Nunc™ F96 MicroWell™ plate, Thermo Scientific) at 25 °C using either 40 mM sodium acetate buffer (150 mM NaCl, pH 4.0) or PBS (pH 7.4) as running buffer and containing 3 mM EDTA and 0.005% Tween 20 (v/v) to reduce non-specific binding. Streptavidin-coated SA biosensor tips (ForteBio) were pre-hydrated in 200 μL PBS running buffer (without EDTA or Tween-20) for 10 min followed by equilibration in PBS for 60 s. Biosensor tips were non-covalently loaded with a 100 nM solution of biotinylated $SRRP_{53608}$-BR in running buffer for 600 s (threshold for binding set to 1 nm) followed by a wash of 60 s in the same buffer. $SRRP_{53608}$-BR mutants were immobilised to separate sets of sensors in the same manner. All biosensors, including reference sensors (without ligand), were blocked with 100 $\mu g.mL^{-1}$ biocytin (Sigma-Aldrich, Poole, Dorset, UK) for 60 s, to prevent non-specific interactions of protein to the sensor surface, followed by a further wash for 60 s. Association of biotinylated $SRRP_{53608}$-BR with PGA (1.56-50 μM in the appropriate running buffer) was performed for 300 s before dissociation of binding was performed using running buffer for 600 s. Experiments with RGI, PECF, pectin, and chondroitin sulfate A were performed in a similar manner. All experiments were performed in triplicate. Data were processed to calculate kinetic and affinity parameters using the ForteBio software.

Generation of DNA fragment for binding assays

Lactobacillus spp.-specific primers were used to generate a 341-bp DNA fragment for use in AFM analysis of $SRRP_{53608}$-BR:DNA binding. To facilitate the immobilisation of the DNA

fragment, a 5' C6-amino modification was incorporated into the forward primer. The primer set used to generate the fragment were: F: 5'-AC6-AGCAGTAGGGAATCTTCCA-3'; R: 5'-CACCGCTACACATGGAG-3'. Reaction volumes of 200 µL were used to generate sufficient product: 20 µL 10x reaction buffer, 4 µL 10 mM dNTP mix, 4 µL each 10 µM forward and reverse primer, 4 µL *L. reuteri* genomic DNA (ATCC 53608), 1 µL Taq DNA polymerase (New England Biolabs; Herts, UK), and 163 µL H_2O. Template genomic DNA was obtained using the Gene Jet kit. PCR conditions were as follows: denaturation: 95 ℃ for 30 s; 35 elongation cycles: 95 ℃ for 30 s, 50 ℃ for 45 s, 68 ℃ for 1 min; final extension: 65 ℃ for 5 min. DNA product purity was confirmed by 1.5% agarose gel electrophoresis and the DNA further cleaned up and concentrated by isopropanol precipitation followed by resuspension in PBS.

Atomic force microscopy

The atomic force microscope (AFM) used in this study was MFP-3D BIO (Oxford Instruments Company, Asylum Research, Santa Barbara, CA. USA). The AFM tips used were silicon nitride contact mode cantilevers (PNP-TR, Nanoworld AG, Neuchâtel, Switzerland). The interactions were examined by covalently attaching the purified $SRRP_{53608}$-BR molecules to AFM tips whereas purified porcine gastric mucin (pPGM), polygalacturonic acid (PGA) or the amino modified DNA fragments were immobilised onto glass slides to enable binding interactions to be measured in a specific manner (20). Functionalization of AFM tips followed a four step procedure (carried out at 21℃): the first step involved incubation of the tips in a 2% solution of 3-mercaptopropyltrimethoxy silane (MTS, Sigma-Aldrich, Poole, Dorset, UK) in toluene (dried over a 4 Å molecular sieve) for 2 h, followed by washing with toluene and then chloroform. In the second step, silanised tips were incubated for 1 h in a 1 $mg.mL^{-1}$ solution of a heterobifunctional linker: MAL-PEG-SCM, 2 kD (Creative PEGWorks, NC, USA) in chloroform. The tips were rinsed with chloroform and then dried with argon. The third step involved covalent attachment by incubation of the tips in 1 $mg.mL^{-1}$ protein solution in PBS (pH 7.4) for 1 h at 21 ℃, followed by a PBS washing step. The fourth step involved incubation of the protein functionalized cantilevers in a 10 $mg.mL^{-1}$ solution of glycine in PBS to 'amine'-

cap any unreacted succinimide groups, followed by washing in PBS. Glass slides were also functionalised in a four step manner. The glass was first silanised with MTS for 2 h and then incubated for 1 h in a specific heterobifunctional linker chemical for covalently attaching proteins or carbohydrates. Respectively, 5 mM N-γ-maleimidobutyryloxy-succinimide ester (GMBS) in ethanol (Thermo Fisher Scientific, Massachusetts, USA) was used for attaching mucin or amined DNA and 5 mM (4-(4-N-maleimidophenyl)butyric acid hydrazide) (MPBH) in ethanol for attaching PGA. In the third step, the molecules being covalently attached to the slides were incubated on the slides for 1 h at concentrations of 1 $mg.mL^{-1}$ in PBS and then rinsed with PBS. The fourth step was carried out to cap any unreacted agents of the slides. For the mucin and DNA slides, 10 $mg.mL^{-1}$ glycine was added to 'amine'-cap any unreacted succinimide groups and for PGA slides, 10 $mg.mL^{-1}$ glucose to 'sugar'-cap any unreacted hydrazide groups, followed by washing with PBS.

Force spectroscopy

Binding measurements were carried out in two saline buffers (137 mM NaCl) of different pH, PBS pH 7.4, and citrate buffered saline, pH 4.0, (Sigma-Aldrich, Poole, Dorset, UK). In addition one of the DNA interaction trials was run in pH 7.4 PBS buffer with varying salt concentration (137 mM NaCl and 1M NaCl). Experimental data were captured in 'force-volume' (FV) mode at a rate of 2 $\mu m.s^{-1}$ in the Z direction and at a scan rate of 1 Hz and a maximum load force of 300 pN (pixel density of 32 x 32 which collects 1024 force-distance curves). The spring constant, k, of the cantilevers was determined by fitting the thermal noise spectra (21), yielding typical values in the range 0.03-0.06 $N.m^{-1}$. Adhesion in the force spectra was quantified using a bespoke Excel macro (22) which fits a line to the baseline of the retract portion of the force-distance data. This method ensures that any non-specific tip-sample interactions (which appear at the tip-glass detachment-point) are eliminated from the measurements.

Competition experiments with PGA were carried out in the pH 4.0 buffer and DNA nucleotide competition experiments were all carried out in PBS (pH 7.4) 137 mM NaCl buffer. Force measurements were repeated on the same region of the samples after addition of 2 $mg.mL^{-1}$

PGA in the pH 4.0 buffer to the AFM's liquid cell. The concentration in the liquid cell was 0.56 mg.mL^{-1} PGA. Following addition of the PGA to the liquid cell a minimum incubation period of 15 min to allow sufficient mixing and settling of the AFM cantilever was included in the protocol. The same protocol was used for the addition of competitive molecules to the DNA or mucin slides.

Tissue adhesion assays

To assess binding of *Lr*SRRP-BR to mouse GI tissue, stomach and colon of wild-type C57 mouse were washed with PBS, fixed in methacarn (60% dry methanol, 30% chloroform and 10% acetic acid), embedded in optimal cutting temperature compound (OCT) and cut into 8 µm sections. Tissue sections were washed twice in PBS containing 0.05% BSA and blocked for 30 min at room temperature with TNB buffer (prepared by dissolving 0.5% w/v Blocking reagent (Perkin Elmer, Boston, MA, USA) at 60 °C in 100 mM Tris-HCl pH 7.5, 150 mM NaCl) supplemented with 5% goat serum. The slides were then washed once in PBS 0.05% BSA, followed by 2 h incubation of 150 µg.mL^{-1} $SRRP_{53608}$-BR in PBS at 37 °C. The negative control tissue sections were incubated with PBS only. After three PBS 0.05% BSA washes, the custom-made rabbit anti-$SRRP_{53608}$-BR antiserum (BioGenes GmbH, Berlin, Germany; diluted 1:100 in TNB buffer) was added to the tissue sections and incubated at room temperature for 1 h. The slides were then washed three times and Alexa Fluor 488-conjugated goat anti-rabbit IgG (Life Technologies A11034, 1:200 diluted in PBS) incubated at room temperature in the dark for 1 h on its own or together with WGA-Rh (at 4 µg.mL^{-1}). WGA-Rh binding at 4 µg.mL^{-1} was also performed on its own. The sections were then washed three times in the dark, counterstained with 25 µg.mL^{-1} 4'6-diamino-2 phenylindole (DAPI, Life Technologies) for 10 min at room temperature in the dark, washed three times and mounted in Hydromount mounting medium. The slides were imaged using a Zeiss Axio Imager 2 microscope.

For sodium periodate treatment of colonic tissue sections, methacarn-fixed OCT embedded C57BL/6 mouse colonic tissue sections (8 µm) were first washed in 0.1 M acetate buffer, pH 4.5, twice for 5 min, followed by an incubation in periodate buffer (10 mM periodate in 0.1 M

acetate buffer) for 2 h in the dark. Tissue was washed in 0.1 M acetate buffer once for 5 min, and twice in PBS. Tissue was reduced by immersion in borate buffer (50 mM $NaBH_4$ in PBS, pH 7.6) for 30 min. Slides were washed twice in PBS and blocked with TNB buffer for 1 h followed by the $SRRP_{53608}$-BR binding protocol described above.

To assess binding of $SRRP_{53608}$-BR to HT29-MTX and HT29 cells, confluent monolayers of HT29-MTX cells (passage 51), were washed twice with PBS, fixed with 1:1 acetone: methanol for 10 min at -20°C, then three times with PBS 0.05% BSA and blocked with TNB buffer supplemented with 5% goat serum for 30 min at room temperature. The cells were then washed in PBS 0.05% BSA and incubated with 150 $\mu g.mL^{-1}$ $SRRP_{53608}$-BR in PBS for 2 h at 37 °C. The negative control samples were incubated with PBS only under these conditions. After three washes, the cells were incubated with the custom-made rabbit anti-$SRRP_{53608}$-BR antiserum (BioGenes GmbH, Berlin, Germany; diluted 1:100 in TNB buffer) on its own or simultaneously with anti-mouse MUC5AC antibody (diluted 1:100) for 1 h at room temperature. The cells were then washed three times and secondary antibody goat anti-rabbit IgG-Alexa Fluor 488 (1:400 diluted in PBS) incubated at room temperature in the dark for 1 h. In co-staining experiments, the binding of MUC5AC antibodies was detected with Texas Red-conjugated goat anti-mouse IgG antibody diluted 1:400. The cells were then washed three times in the dark and counterstained with 1 $\mu g.mL^{-1}$ DAPI for 10 min at room temperature in the dark, washed three times and mounted in Vectashield. The same protocol for $SRRP_{53608}$-BR binding was carried out using the non-mucus secreting HT29 cell line (passage 12). The slides were imaged using a Zeiss Axio Imager 2 microscope.

Glycan arrays

Glycan microarrays on nitrocellulose membranes, version Agata 1.0, contained duplicate dilution series (8-1000 $\mu g.mL^{-1}$) of plant-derived polysaccharides and oligosaccharides. The arrays also contained rows of autojet buffer blank spots as negative controls and patterns of dye spots for array orientation (23, 24). All incubations and wash steps were performed in a 6-well Cellstar® cell culture plate (Greiner Bio-One). Arrays were incubated in 5 mL Protein-

Free (PBS) Blocking Buffer (Thermo Scientific Pierce) overnight at 4 °C then washed in either 5 mL PBS (pH 7.4) or 5 mL 10 mM acetate buffer, 150 mM NaCl (acetate buffered saline; ABS, pH 4.0) with 70 rpm rocking for three intervals of 5 min at 25 °C. $SRRP_{53608}$-BR (1 ml of 100 $\mu g.mL^{-1}$) in either PBS (pH 7.4) or ABS (pH 4.0) was incubated with each array on an orbital shaker at 80 rpm for 2 h at 25 °C then three 5 min washes were carried out in the respective 0.05% (v/v) Tween 20-containing buffer solution [PBST (pH 7.4) or ABST (pH 4.0)] with orbital shaking at 90 rpm at 25 °C. All subsequent antibody incubation and washing steps were performed in PBS or PBST (pH 7.4). Arrays were probed with anti-$SRRP_{53608}$-BR primary antibody (raised in rabbit; 1:1000), followed by three 5 min PBST washes, and probed with FITC-anti-rabbit IgG secondary antibody (1:5000). Arrays were washed three times for 5 min each with PBST, dried and imaged in a Pharos FX Plus Molecular Imager (Bio-Rad) using Quantity One® 1-D analysis software (Bio-Rad) and filter settings for FITC.

PCR analysis of *L. reuteri srrp* genes

L. reuteri strains lp167-67, 20-2, 3c6, ATCC 53608 and pg-3b were grown overnight in de Man-Rogosa-Sharpe (MRS) medium at 37 °C. DNA was extracted from 1.5 mL cultures using the gene JET Genomic DNA purification kit (ThermoFischer Scientific, UK), following the manufacturer's instructions for DNA extraction from Gram-positive bacteria, with the following modifications: 10 $U.mL^{-1}$ mutanolysin (Sigma Aldrich, UK) was added to the lysis buffer, and the lysis step was extended to 1 h at 37 °C. For the PCR reactions, primer Lr_1105r2 was used in combination with Lr_1105f2, 3c6-SRRP-rev and 20-2-SRRP-rev, to amplify the *srrp* gene from *L. reuteri* ATCC 53608, lp16-67, 3c6, 20-2 and pg-3b, respectively. Primers 3c6_ps-SRRP-for and 3c6_ps-SRRP-rev were used to amplify the *srrp* pseudogene from *L. reuteri* 3c6 and primers pg3b_ps-SRRP-for and pg3b_ps-SRRP-rev were used to amplify the *srrp* pseudogene from *L. reuteri* pg-3b (see **Table S5**). PCR reactions were performed with Q5® High-Fidelity DNA polymerase (New England Biolabs, UK). DNA was denatured at 98 °C for 10 s, followed by annealing of the primers at 55 °C for the *L. reuteri* 20-2, 3c6 and pg-3b *srrp* genes and 58 °C for the *L. reuteri* lp167-67 *srrp* gene and *L. reuteri* 3c6 and pg-3b *srrp*

pseudogenes. The extension step was performed at 72 ºC for 4 min for 35 cycles. The PCR products were analyses on 1% agarose gels and stained with 2.5 mM ethidium bromide.

Western blot analysis of *L. reuteri* SRRPs

L. reuteri strains lp167-67, 20-2, 3c6, ATCC 53608 and pg-3b were grown overnight in MRS at 37 ºC. Following centrifugation at 4000 *x g* and 4 ºC for 10 min, the supernatant was recovered and concentrated 10-fold by spin filtration using Vivaspin 10 kDa MWCO spin filters (Sartorius, UK). Spent media proteins (15 µg) were separated by SDS-PAGE on a Bis-Tris 4-12% precast NuPAGE gel (ThermoFischer Scientific, UK) for 50 min at 200 V and the gel blotted onto a PVDF membrane at 30 V for 2 h. The blot was blocked with Protein-free (PBS) blocking solution (ThermoFischer Scientific, UK) and incubated with custom-made rabbit anti-$SRRP_{53608}$-BR antiserum (BioGenes GmbH) at 1/2000 dilution in PBS for 1 h at 25 ºC. The blot was then washed three times with PBST and incubated with goat anti-rabbit-IgG-alkaline phosphatase conjugated antibody (Sigma Aldrich, UK) at 1/20000 dilution in PBS for 1 h at 25ºC. After washing three times with PBST, the blot was incubated in a visualisation solution (40 µM $MgCl_2$, 0.1 mM nitroblue tetrazolium, 0.1 mM 5-bromo-4-chloro-3-indolyl phosphate toluidine) for 3 min at 25 ºC. The blot was washed in H_2O and scanned on a GS-800 calibrated densitometer (Bio-Rad, UK).

Native polyacrylamide gel electrophoresis

Samples in Native PAGE Sample Buffer (Novex; Invitrogen Life Technologies) were electrophoresed in a Native PAGE Novex 4-16% gradient Bis-Tris gel with Native PAGE Anode Buffer and Dark Blue Cathode buffer for 105 min at 150 V, fixed according to the manufacturer's protocol then stained with Colloidal Blue (Novex; Invitrogen Life Technologies). Native Mark Unstained Protein Standard MW markers (5 µl) (Novex; Invitrogen Life Technologies) were electrophoresed to estimate protein MW.

Circular Dichroism

$SRRP_{53608}$-BR protein was buffer exchanged into either 10 mM sodium acetate buffer (pH 4.0) or 10 mM sodium phosphate buffer (pH 7.4) and adjusted to a concentration of 1 $mg.mL^{-1}$. Samples were loaded into a 0.1 mm split glass cuvette and run on a JASCO J-710 spectropolarimeter (Great Dunmow, Cambs, UK). UV CD spectra were acquired over a scan range of 180-260 nm at a scan speed of 20 $nm.min^{-1}$ with a band width of 1.0 nm and a response time of 4 s. Data were manipulated including subtraction of blank spectra using the JASCO Spectra Manager 32 v1.40.00a software (Easton, MD, USA) and the DichroWeb online tool (http://dichroweb.cryst.bbk.ac.uk) and CONTIN analysis program. Blank buffers were used as controls.

SI Figures

Figure S1. Neighbour-Joining phylogram generated from a Clustal Omega alignment of 94 BR domains from SRRPs and pseudo-SRRPs displayed using EvolView (http://www.evolgenius.info/evolview/) (25). SRRP-BRs are displayed as follows: circle, from a commensal or non-pathogenic strain; star, from a pathogenic or clinical isolate. Host or source origins of strains are indicated as follows; pink, porcine; red, rodent; purple, human; aqua, avian; lime, bovine; blue, insect; yellow, sourdough; orange, other fermented food or drink. The majority of BRs are from *L. reuteri* strains, in which case only the name of the strain is given.

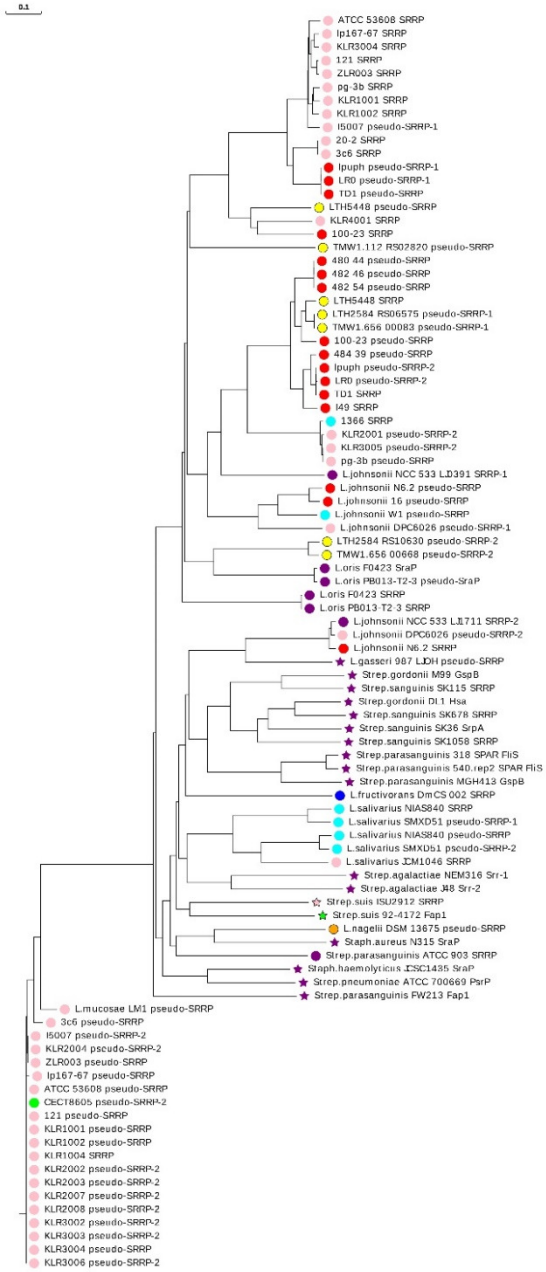


Figure S2. Analysis of SRRP expression in *L. reuteri* strains. (**a**) Analysis of PCR products by agarose gel stained with ethidium bromide, (**b**) Western blot analysis of spent media proteins from *L. reuteri* cultures, probed with anti-$SRRP_{53608}$-BR antibody.

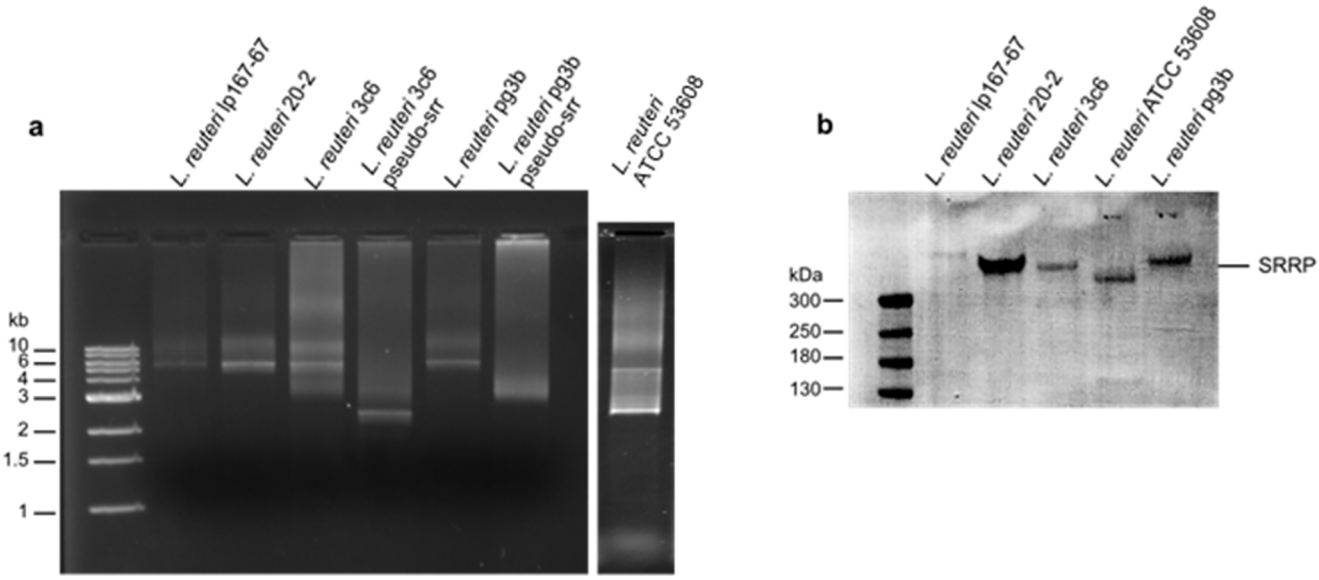


Figure S3. Solvent accessible surface electrostatic potential (EP) maps of TM-Pel and *Lr*SRRP-BRs.

(**a**) TM-Pel in complex with TGA (3ZSC) showing a basic binding pocket. (**b**) $SRRP_{53608}$-$BR_{262-571}$, where PuBS is indicated by a dashed oval, exhibiting an increased basic charge distribution in the region of the aromatic and positive amino acid triads (shown in **Figure 3d**) (**c**) $SRRP_{53608}$-$BR_{262-571}$, where the basicity of the binding pocket is significantly reduced after removal of the lower loop from F411 to T422. (**d**) Structure of $SRRP_{100-23}BR_{257-623}$ which was built without its lower loop, showing an increasingly acidic binding pocket compared to $SRRP_{53608}$-$BR_{262-571}$ ΔF411-T422 in (**c**).

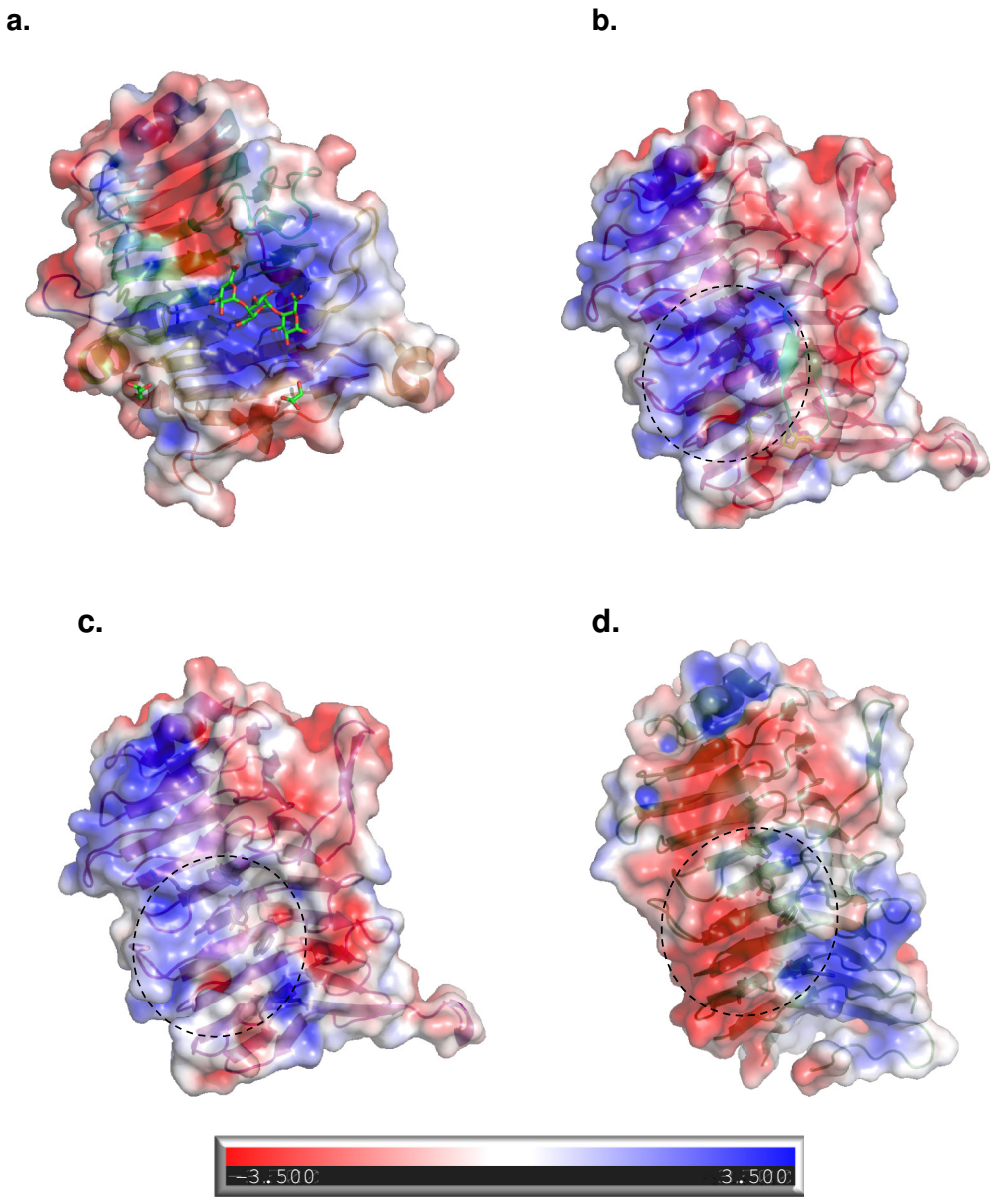


Figure S4. $SRRP_{53608}$-$BR_{262-571}$ distinct predicted protonation states at pH 4.0 and pH 7.4

(**a**) Cartoon representation of $SRRP_{53608}$-$BR_{262-571}$ generated from crystal coordinates. Residues predicted to exhibit different protonation states at pH 4 compared to pH 7.4 are shown as sticks. (**b**) As above, showing the interface between symmetry-related molecules. Differences in protonation state of residues at the interface may prevent crystal formation at pH 4.

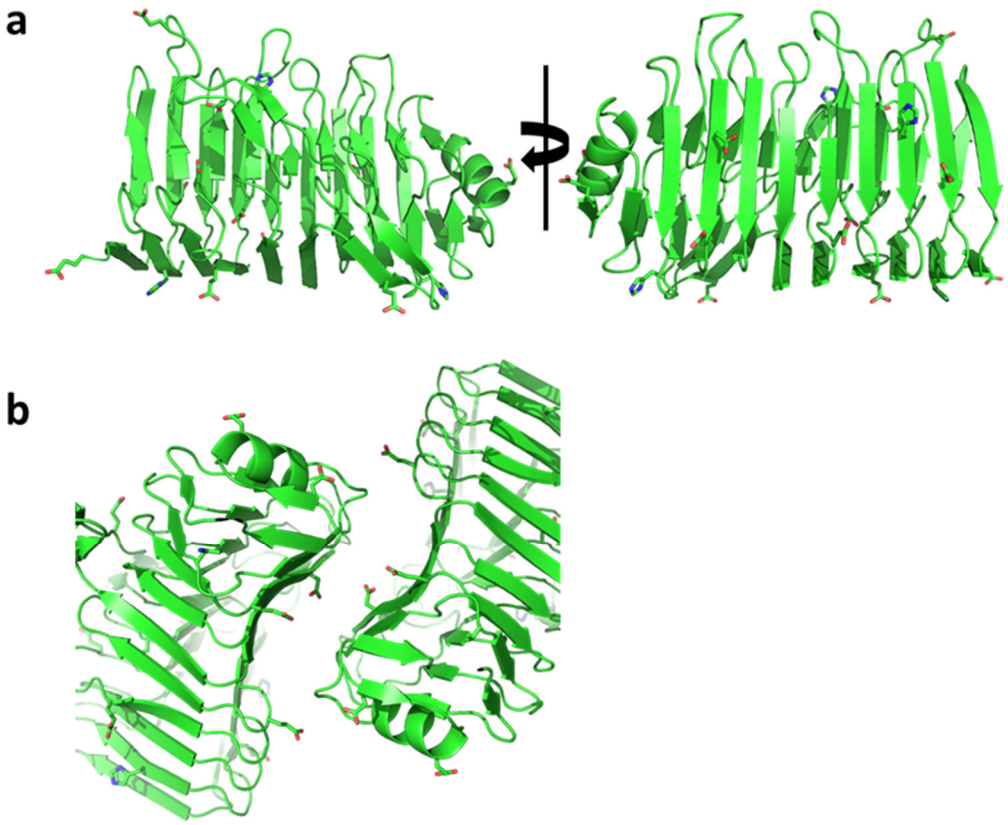


Figure S5. PuBS exhibits a more positive surface electrostatic potential at pH 4.0 compared with pH 7.4. Surface representation of $SRRP_{53608}$-$BR_{262-571}$ at pH 7.4 (**a**) and pH 4 (**b**), showing surface electrostatics of PuBS (circled). At pH 7.4, PuBS exhibits a more negative electrostatic potential. Coordinates were obtained from representative frames of each respective molecular dynamics trajectory (see methods). Surface electrostatics were calculated in PyMOL and coloured as blue (positive), white (neutral), and red (negative).

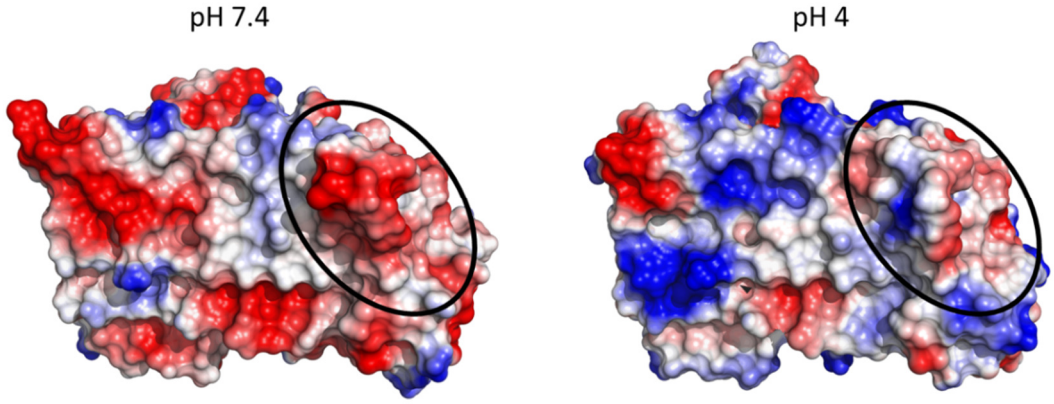


Figure S6. PuBS exhibits a pH-dependent conformational transition. (**a**) Stick representation of $SRRP_{53608}$-$BR_{262-571}$ backbone atoms at pH 7.4 (green carbons) and pH 4 (cyan carbons). Rotation about the I513 C_α-C bond (black arrow) is facilitated by hydrogen bond formation between the protonated carboxylate of D487 and I514 carbonyl (orange dash), leading to a conformational change in the loop between β30 and β31 (β30→β31). (**b**) RMSD of β30→β31 loop backbone as a function of time at pH 7.4 (left) and pH 4 (right). Coordinates from each frame of the MD trajectory were first overlaid by minimising the backbone RMSD of the entire protein before performing a no-fit RMSD calculation, referenced against the crystal coordinates, of the β30→β31 backbone atoms (I513-T522). A notable increase is seen at pH 4 after ca. 0.6 μs (red dash), indicating a conformational change in this region. This is accompanied by a change in the ψ-angle (N_i-C_α-C-N_{i+1}) of I513 (middle), and a decrease in

the distance between the D487 hydroxyl oxygen (OD2) and I514 carbonyl oxygen (O, bottom), indicating hydrogen bond formation.

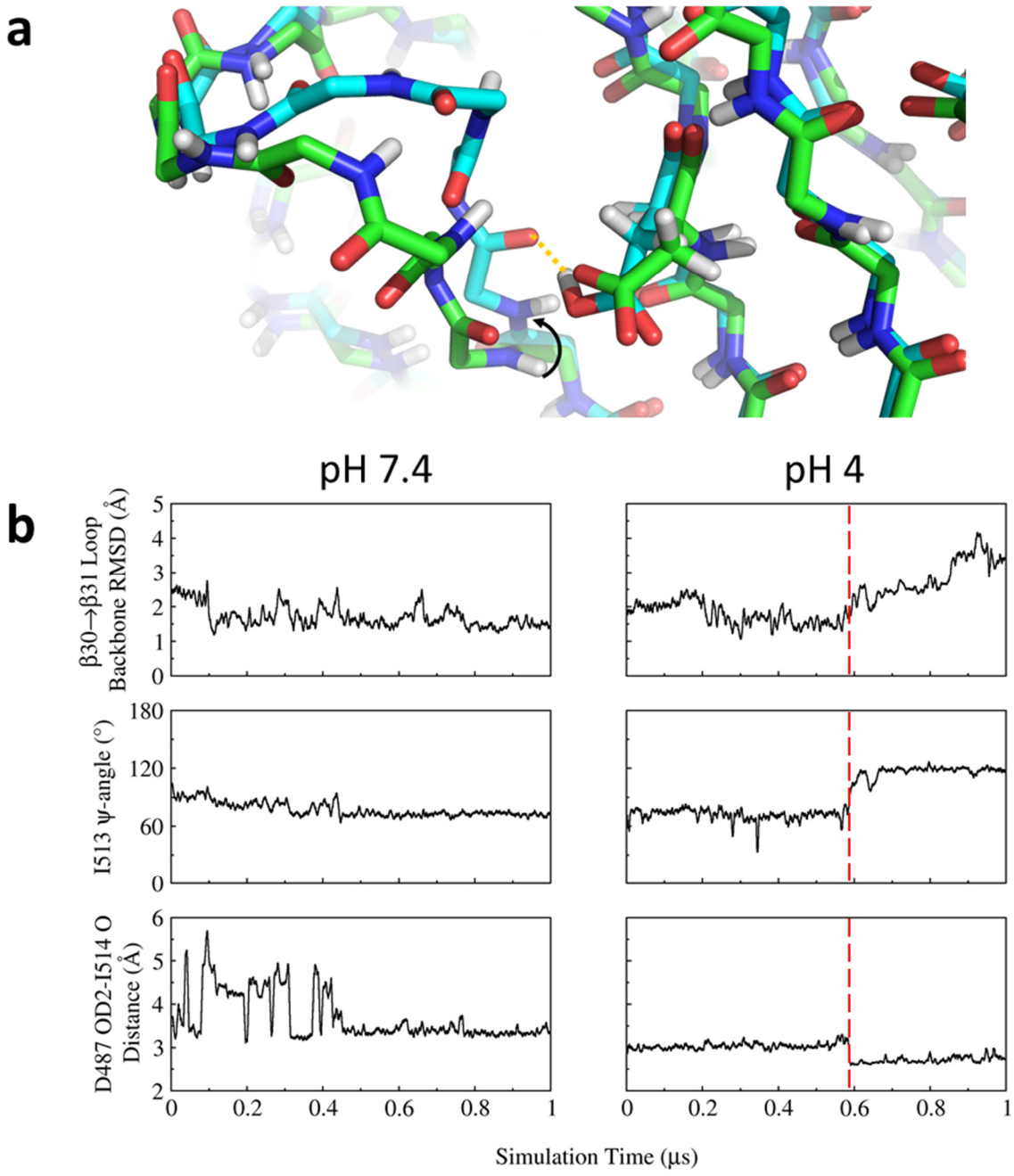


Figure S7. PuBS pH-dependent conformational transition involves rearrangement of key sidechains for anionic ligands binding at low pH. (**a**) Cartoon representation of PuBS in $SRRP_{53608}$-$BR_{262-571}$ at pH 7.4 (green) and pH 4 (cyan) with key residue sidechains shown as sticks. Non-polar hydrogen atoms are omitted for clarity. At pH 4, displacement of the I514 sidechain away from PuBS causes a conformational change in putative binding residue sidechains that may prearrange PuBS for substrate binding. (**b**) RMSD of key PuBS residue sidechains as a function of time (Y482: top, R512: middle, I514: bottom). Coordinates from each frame of the trajectory were first overlaid by minimising the backbone RMSD of the entire protein before performing a no-fit RMSD calculation, referenced against the crystal coordinates, for the heavy atoms of each of the stated sidechains. The conformational change at pH 4 is highlighted with a dashed red line.

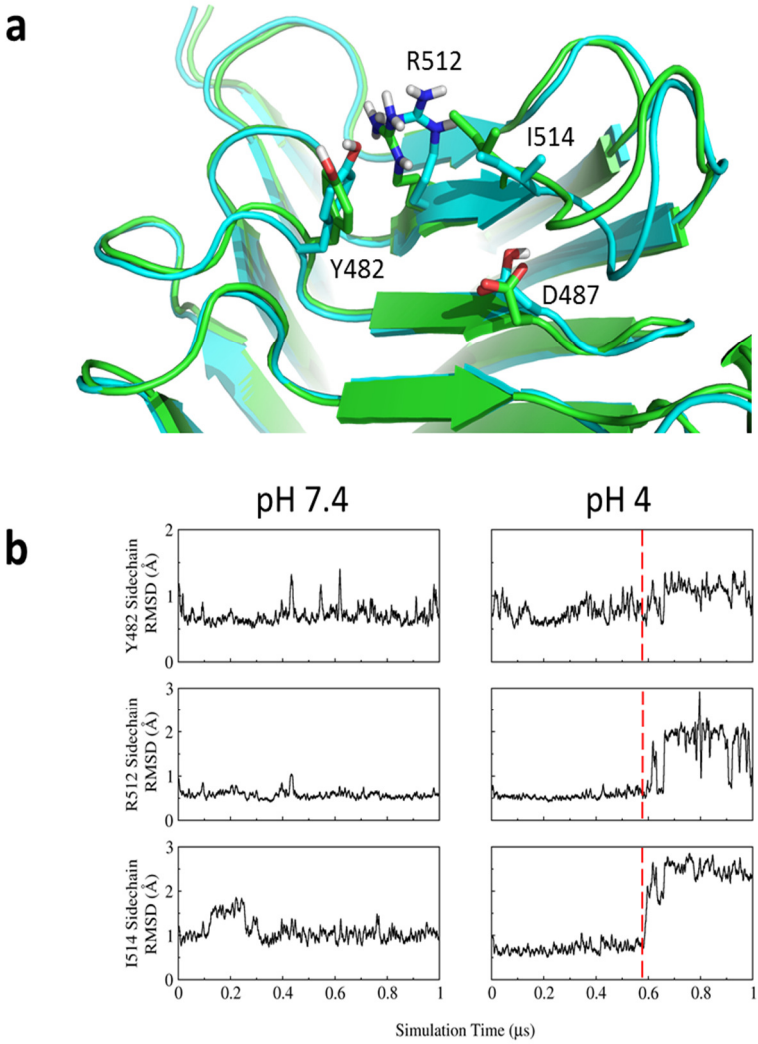


Figure S8. Binding of $SRRP_{53608}$-BR to plant-based glycan arrays version Agata 1.0. Arrays contain duplicate sets of spotted glycan dilution series (8-1000 µg.ml^{-1} glycans) with each identical half separated by a vertical line. The pattern of white spots at the top and bottom of each array are dye spots for array orientation. (**a**) Array probed with $SRRP_{53608}$-BR in ABS (pH 4.0); (**b**) Array probed with $SRRP_{53608}$-BR in PBS (pH 7.4). Arrays were scanned in a Pharos FX Plus Molecular Imager (Bio-Rad). Positive signals are described in **Table S5**. For the list and layout of the glycans that were spotted, see (23, 24).

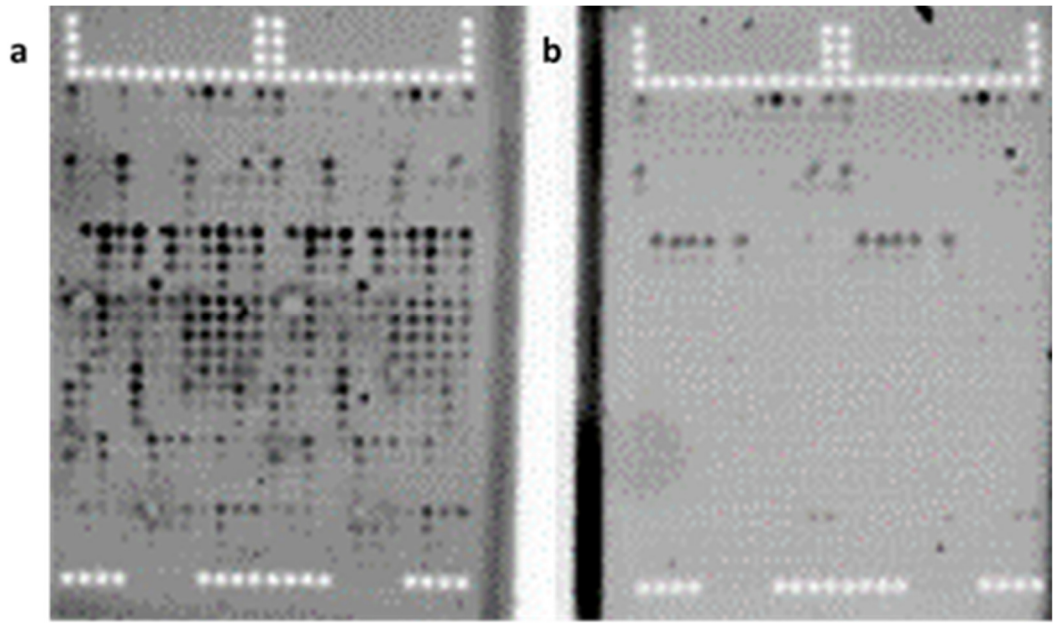


Figure S9. Interaction of $SRRP_{53608}$-BR with pectins by biolayer interferometry. Representative sensorgrams demonstrate the interaction of immobilised $SRRP_{53608}$-BR with (**a**) RGI; (**b**) PGA; (**c**) PECF; and (**d**) chondroitin sulfate A at pH 4.0.

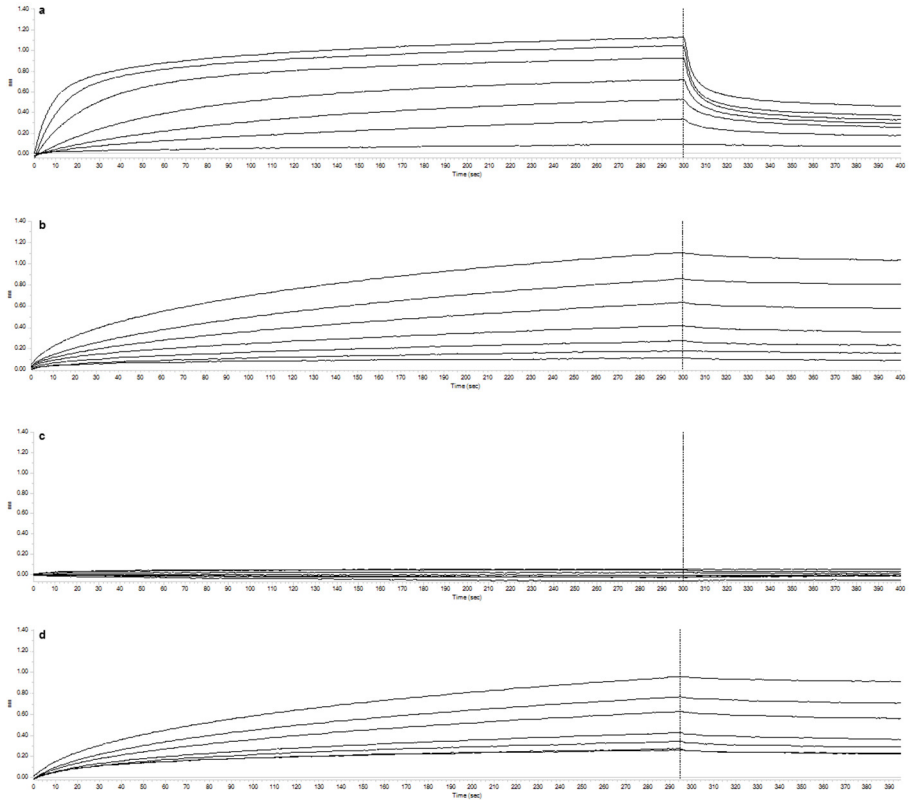
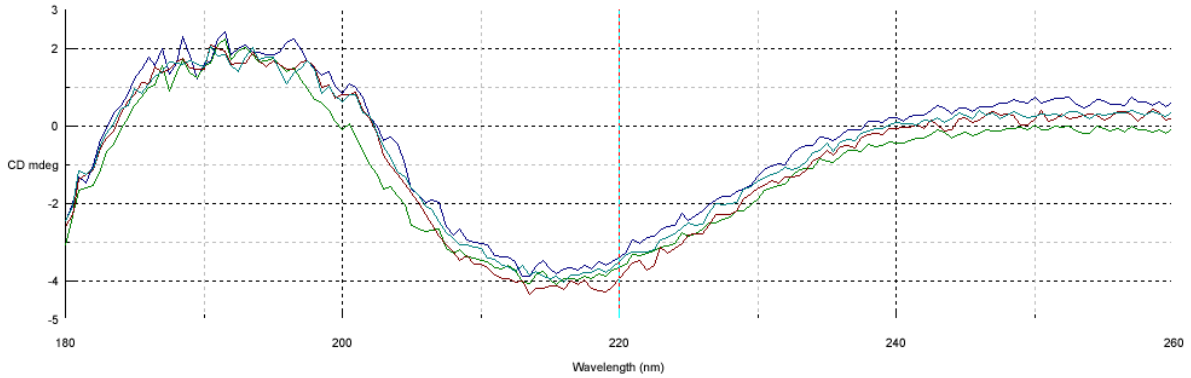


Figure S10. CD spectra of $SRRP_{53608}$-BR wild-type and mutants. (**a**) Monomeric wt $SRRP_{53608}$-BR, pH 7.4 (purple); monomeric wt $SRRP_{53608}$-BR, pH 4.0 (green); oligomeric wt $SRRP_{53608}$-BR, pH 7.4 (teal); oligomeric wt $SRRP_{53608}$-BR (dark red). (**b**) wt $SRRP_{53608}$-BR, pH 7.4 (blue), K377A mutant $SRRP_{53608}$-BR, pH 7.4 (red), R512A mutant $SRRP_{53608}$-BR, pH 7.4 (teal), ΔF411-T422 mutant $SRRP_{53608}$-BR, pH 7.4 (green)

a



b

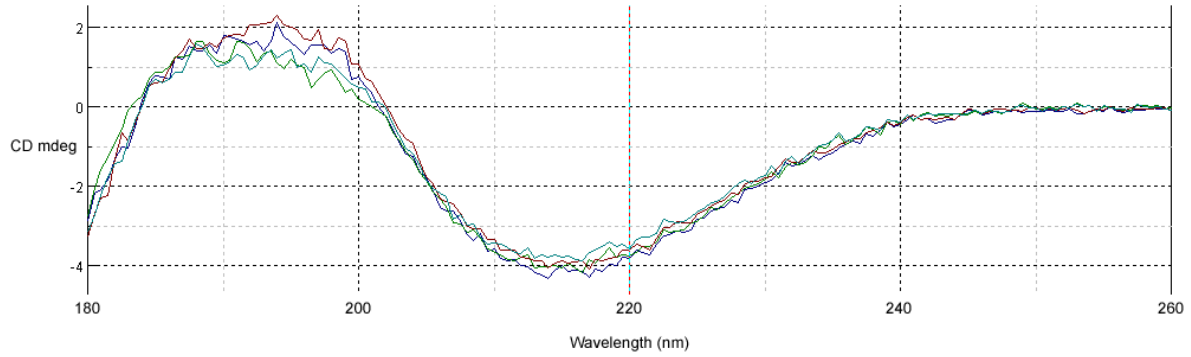


Figure S11. Structures of SRRP$_{53608}$-BR ligands (**a**) RGI; (**b**) PGA; and (**c**) chondroitin sulfate A

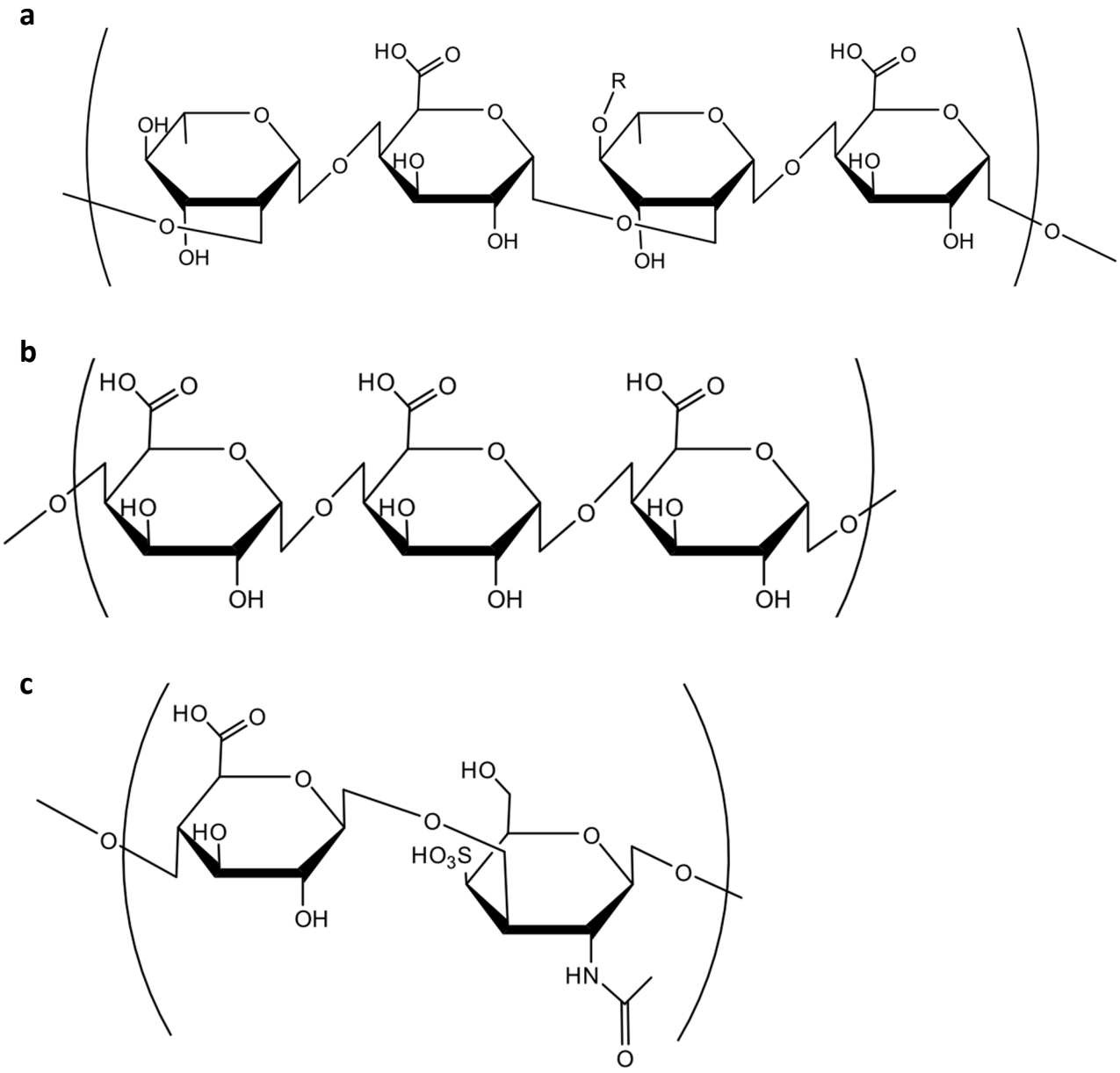
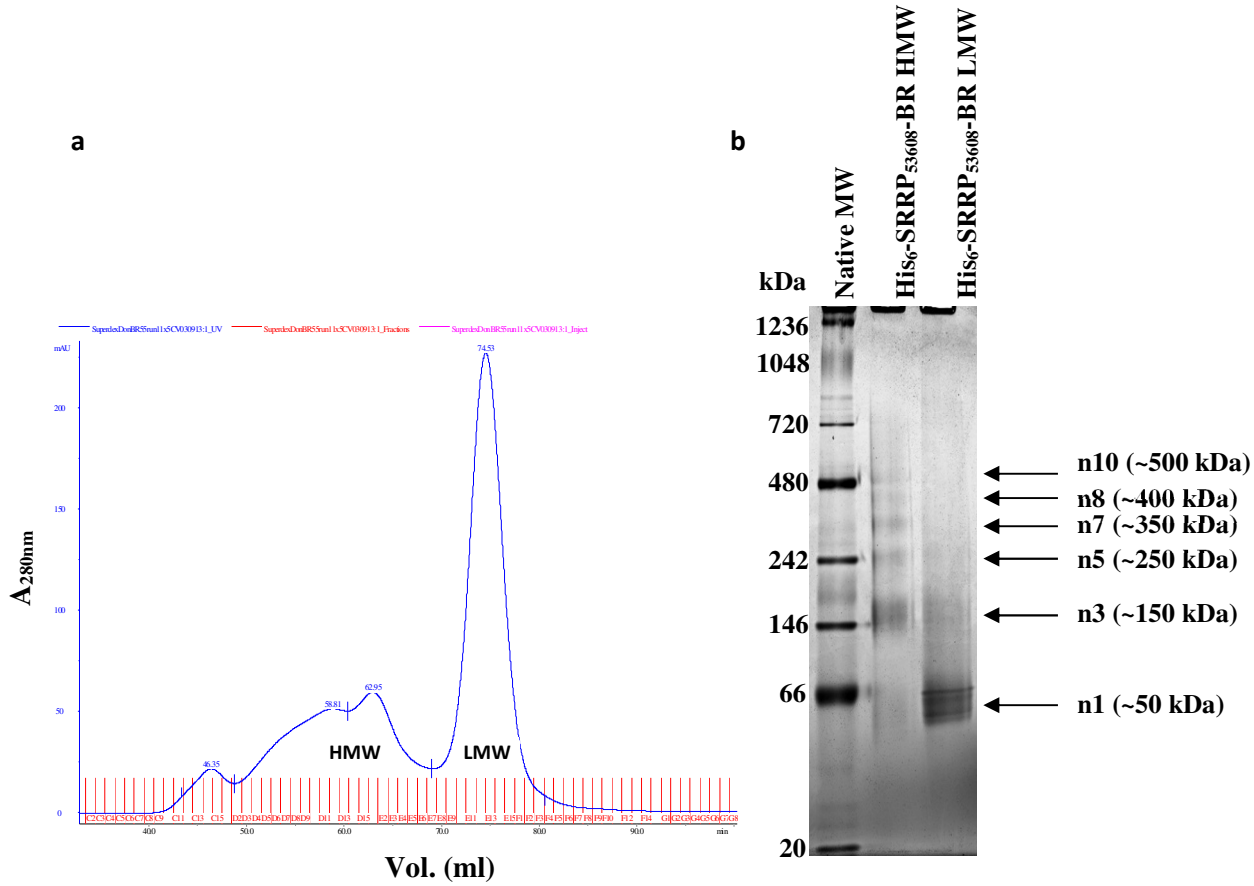


Figure S12. Monomeric and oligomeric forms of His-tagged SRRP$_{53608}$-BR. (**a**) Size exclusion gel filtration (SEGF) chromatography of IMAC-purified and dialysed His$_6$-SRRP$_{53608}$-BR (4-5 mg) through Superdex 200 (16/600) in 20 mM Tris-HCl (pH 8.0), 100 mM NaCl running buffer at 1 mL.min^{-1}, 1.5 column volume, (**b**) Native PAGE Novex 4-16% gradient Bis-Tris gel electrophoresed with native PAGE Anode Buffer and Dark Blue Cathode Buffer for 105 min at 150 V, then fixed and Colloidal Blue stained (Novex; Invitrogen Life Technologies). Samples were 2 µg each of high MW (HMW) and low MW (LMW) SEGF elution fractions. Native Mark Unstained Protein Standard MW markers (5 µL) (Novex; Invitrogen Life Technologies) were electrophoresed to estimate protein MW.



SI Tables

Table S1. Strains of *Lactobacillus* spp. with a SecA2-SecY2 accessory secretion system and SRRP adhesin(s)[a]

Species	Strain	Genome accession no.	SRRP gene ID(s) or locus tag(s)	pseudo-SRRP gene ID(s) or locus tag(s)
L. reuteri	ATCC 53608	LN906634-906636	LRATCC53608_0906	LRATCC53608_0916-_0917
	100-23	IMG: 2500069000	2500070902 ("Lr_70902")	2500070903-2500070904 ("Lr_70903-_70904")
	TD1	CP006603	N134_05915	N134_05970-_05965
	I49	NZ_CP015408	A4V07_RS03935-_RS03925[b]	—
	1366	NBBG01000001-01000093	B6J74_RS03710	—
	LTH5448	NZ_JOOG01000001-01000036	HN00_RS06750[c]	HN00_RS06765-_RS09675[c]
	121	MKQH01000001-01000014	BJI45_RS02085[b]	BJI45_RS02135-_RS02140
	ZLR003	NZ_CP014786	ADV92_RS10755[b]	ADV92_RS10805-_RS10810
	KLR1001	MIME01000001-01000145	BHL74_RS10285[b]	BHL74_RS10235-_RS10230
	KLR1002	MIMF01000001-01000392	BHL85_RS12575[b]	BHL85_RS12625-_RS12630
	KLR1004	MIMH01000001-01000154	BHL89_RS06905-_RS06910[b]	BHL89_RS06855[d]
	KLR3004	MIMT01000001-01000149	BHL81_RS01145[b]	BHL81_RS01195-_RS01200
	KLR4001	MIMW01000001-01000136	BHL84_RS04875	BHL84_RS04815
	3c6	LN887305-887505	LR3C6_01537A-_01537	LR3C6_00253A-_00253
	20-2	LN887506-887693	LR202_00269A-_00269	LR202_00347A-_00347
	lp167-67	LN887694-887827	LRLP167_00243A-_00243	LRLP167_00253-_00254
	pg-3b	LN887201-887304	LRPG3B_00922A-_00922	LRPG3B_00234A-_00234
	lpuph	NZ_AEAX01000001-01000127	—[d]	ECQ_RS11170[d] & ECQ_RS0107090-_RS0107085
	mlc3	NZ_AEAW01000001-01000126	—[d]	ECM_RS10935[d,e]
	480_44	MBLQ01000001-01000154	—[d]	BBP10_RS03300[d]
	482_46	MBLR01000001-01000188	—[d]	BBP11_RS07570[d]
	482_54	MBLS01000001-01000186	—[d]	BBP12_RS00010[d]
	484_32	MBLT01000001-01000523	—[d]	BBP13_RS14455[d]
	484_39	MBLU01000001-01000191	—[d]	BBP14_RS11505[d]
	KLR2001	MIMI01000001-01000149	—[d]	BHL90_RS06020[d] & BHL90_RS06080-_RS06085
	KLR2002	MIMJ01000001-01000169	—[d]	BHL91_RS06735[d] & BHL91_RS06785-_RS06790
	KLR2003	MIMK01000001-01000149	—[d]	BHL92_RS06520[d] & BHL92_RS06570-_RS06575
	KLR2004	MIML01000001-01000140	—[d]	BHL93_RS00715-_RS00720[d] & BHL93_RS00770-_RS00775
	KLR2007	MIMO01000001-01000136	—[d]	BHL76_RS05680-_RS05685[d] & BHL76_RS05735-_RS05740
	KLR2008	MIMP01000001-01000143	—[d]	BHL77_RS05585[d] & BHL77_RS05535-_RS05530
	KLR3002	MIMR01000001-01000228	—[d]	BHL79_RS06455-_RS06450[d] & BHL79_RS06400-_RS06395
	KLR3003[f]	MIMS01000001-01000142	—	BHL80_RS07775-_RS07780
	KLR3005	MIMU01000001-01000172	—[d]	BHL82_RS05895[d] & _BHL82_RS05955[d]
	KLR3006	MIMV01000001-01000257	—[d]	BHL83_RS05150-_RS05145[d] & BHL83_RS05095-_RS05090
	CECT8605	MWVS01000001-01000207	—[d]	B5D07_RS10730-_RS10735[d] & B5D07_RS10785-_RS10790
	I5007	NC_021494-_021504	—	LRI_RS04175-_RS04185 & LRI_RS04235-_RS04240
	LR0	MWIJ01000001-01000075	—	B2G46_RS05900-_RS05885 & B2G46_RS05955-_RS05950
	TMW1.112	NZ_JOKX02000001-02000012	—	HF82_RS02820-_RS02815
	TMW1.656	IMG: 2534682350	—	LR4_00083[g] & LR4_00669-_00668
	LTH2584	NZ_JOSX01000001-01000025	—	LR3_RS06575-_RS10625 & LR3_RS10630-_RS06580
L. oris	F0423	AFTL01000001-01000020	HMPREF9102_0778 & HMPREF9102_0779	—
	PB013-T2-3	AEKL01000001-01000089	HMPREF9265_0662	HMPREF9265_0661
L. salivarius	NIAS840	NZ_AFMN01000001-01000004	NIAS840_RS03930	NIAS840_RS00005[h] & NIAS840_RS10015
	JCM1046	NZ_CP007646-007650	LSJ_RS11575[i]	LSJ_RS00310-_RS11750 & LSJ_RS11780[i]
	L12	IMG: 2540341228	—	L120_00174 & L120_00797
	cp400	NZ_CBVR010000001-010000089	—	LSCP400_RS00005-_RS00010 & LSCP400_RS04945
	SMXD51	NZ_AICL01000001-01000010	—	SMXD51_RS09695-_RS09710 & SMXD51_RS00010[i]
L. johnsonii	NCC 533	AE017198	LJ_0391 & LJ_1711[k]	—
	N6.2	CP006811	T285_07275[j]	T285_01855-_01860
	DPC6026	NC_017477	—	LJP_RS01910-_RS09490 & LJP_RS09630-_RS09450[m]
	L6	IMG: 2529292694	—	L60_01856-_01857
	16	LIGY01000001-01000156	—	LJ16_RS02840
	W1	LSNG01000001-01000049	—	AYJ53_RS01730
L. fructivorans	DmCS_002	JOJZ01000001-01000025	LfDm3_0405	—
L. gasseri	987_LJOH[n]	JUKW01000001-01000081	—	ADF22_RS09755-_RS09750
	JV-V03	GL379580-379587	—	HMPREF0514_10057-_10058
	K7	KL402718-402725	—	LK7_01678
	L3	IMG: 2518645515	—	LGS03_00064-_00066
L. mucosae	LM1[o]	NZ_CP011013	—	LBLM1_RS11745, LBLM1_RS03980[p] & LBLM1_RS11555[p]
	DPC 6426	JSWI01000001-01000072	—	OC62_RS10615
L. murinus	ASF361	KB822402-822412	—	ASF361_01512
L. rhamnosus	L33	IMG: 2518645521	—	LRH33_01715-_01714
L. nagelii	DSM 13675	AZEV01000001-01000044	—	FD45_GL000390[q,q]

[a] Strains that possess a SecA2-SecY2 cluster but lack a SRRP (linked or unlinked to the cluster) have been excluded

[b] Gene incorrectly annotated in genome; full-length SRRP translated in one ORF

[c] Both genes are unlinked to the LTH5448 SecA2-SecY2 cluster (HN00_RS05550-_RS05600)

[d] Possible intact SRRP with partial gene sequence at the end of a draft genome contig, in some cases, annotated incorrectly as two pseudogene fragments

[e] Unlinked to the mlc3 SecA2-SecY2 cluster (ECM_RS0104210-_RS0104260)

[f] Strain KLR3003 has a truncated SecA2-SecY2 cluster with only genes encoding SecA2, GtfA, GtfB, a small hypothetical protein and a pseudo-SRRP

[g] Unlinked to the TMW1.656 SecA2-SecY2 cluster (LR4_00668-_00683)

[h] Unlinked to the SecA2-SecY2 cluster which is split into two regions of the draft genome (NIAS840_RS03905-_RS03930 and NIAS840_RS01240-_RS10015)

[i] SRRP (LSJ_RS11575) and pseudo-SRRP (LSJ_RS11780) are unlinked to the JCM1046 SecA2-SecY2 cluster (LSJ_RS00285-_RS00380)

[j] The SecA2-SecY2 cluster is found on two adjacent contigs in the SMXD51 draft genome (SMXD51_RS00595-_RS09705 and SMXD51_RS09710-_RS00700) but pseudogene SMXD51_RS00010 is unlinked to these
[k] Unlinked to the NCC 533 SecA2-SecY2 cluster (LJ_0384-_0393)
[l] Unlinked to the N6.2 SecA2-SecY2 cluster (T285_01815-_01870) but is linked to four other Gtf genes (T285_07255-_07275)
[m] Unlinked to the DPC6026 SecA2-SecY2 cluster (LJP_RS01870-_RS01945) but is linked to two other Gtf genes (LJP_RS07505-_RS07510)
[n] A clinical isolate from the wound of an intensive care unit patient
[o] Two SecA2-SecY2 clusters present in *L. mucosae* LM1 each with their own pseudo-SRRP(s) but with one cluster lacking Gtf genes (LBLM1_RS03980-_RS04070 & LBLM1_RS11555-_RS04660, respectively)
[p] Pseudogene fragments translated in forward frames (not the reverse complement as indicated in the published genome)
[q] Translated in the opposite reading frame to that annotated in the genome from nt 112572–115183

Table S2. Occurrence of SecA2-SecY2 accessory secretion system and SRRP(s) in genome-sequenced strains of *L. reuteri*

Strain	Host/Origin	Genome accession no.	SecA2-SecY2 gene cluster	SRRP gene ID or locus tag	pseudo-SRRP gene ID(s) or locus tag(s)
ATCC 53608	Pig	LN906634-906636	LRATCC53608_0906-_0917	LRATCC53608_0906	LRATCC53608_0916-_0917
lp167-67	Pig	LN887694-887827	LRLP167_00254-_00243	LRLP167_00243A-_00234	LRLP167_00253-_00254
pg-3b	Pig	LN887201-887304	LRPG3B_00930-_00922	LRPG3B_00922A-_00922	LRPG3B_00234A-_00234
20-2	Pig	LN887506-887693	LR202_00259-_00269	LR202_00269A-_00269	LR202_00347A-_00347
3c6	Pig	LN887305-887505	LR3C6_01062-_01069	LR3C6_01537A_01537	LR3C6_00253A-_00253
121	Pig	MKQH01000001-01000014	BJI45_RS02085-_RS02140	BJI45_RS02085[a]	BJI45_RS02135-_RS02140
ZLR003	Pig	NZ_CP014786	ADV92_RS10755-_RS10810	ADV92_RS10755[a]	ADV92_RS10805-_RS10810
KLR1001	Pig	MIME01000001-01000145	BHL74_RS10230-_RS10285	BHL74_RS10285[a]	BHL74_RS10235-_RS10230
KLR1002	Pig	MIMF01000001-01000392	BHL85_RS12575-_RS12630	BHL85_RS12575[a]	BHL85_RS12625-_RS12630
KLR3004	Pig	MIMT01000001-01000149	BHL81_RS01145-_RS01200	BHL81_RS01145[a]	BHL81_RS01195-_RS01200
KLR4001	Pig	MIMW01000001-01000136	BHL84_RS04815-_RS04875	BHL84_RS04875	BHL84_RS04815
KLR1004	Pig	MIMH01000001-01000154	BHL89_RS06855-_RS06910	BHL89_RS06905-_RS06910[a]	BHL89_RS06855[b]
KLR2001	Pig	MIMI01000001-01000149	BHL90_RS06020-_RS06085	—[b]	BHL90_RS06020[b] & BHL90_RS06080-_RS06085
KLR2002	Pig	MIMJ01000001-01000169	BHL91_RS06735-_RS06790	—[b]	BHL91_RS06735[b] & BHL91_RS06785-_RS06790
KLR2003	Pig	MIMK01000001-01000149	BHL92_RS06520-_RS06575	—[b]	BHL92_RS06520[b] & BHL92_RS06570-_RS06575
KLR2004	Pig	MIML01000001-01000140	BHL93_RS00715-_RS00775	—[b]	BHL93_RS00715-_RS00720[b] & BHL93_RS00770-_RS00775
KLR2007	Pig	MIMO01000001-01000136	BHL76_RS05680-_RS05740	—[b]	BHL76_RS05680-_RS05685[b] & BHL76_RS05735-_RS05740
KLR2008	Pig	MIMP01000001-01000143	BHL77_RS05530-_RS05585	—[b]	BHL77_RS05585[b] & BHL77_RS05535-_RS05530
KLR3002	Pig	MIMR01000001-01000228	BHL79_RS06395-_RS06455	—[b]	BHL79_RS06455-_RS06450[b] & BHL79_RS06400-_RS06395
KLR3003[c]	Pig	MIMS01000001-01000142	BHL80_RS07755-_RS07780[c]	—	BHL80_RS07775-_RS07780
KLR3005	Pig	MIMU01000001-01000172	BHL82_RS05895-_RS05955	—[b]	BHL82_RS05895[b] & BHL82_RS05955[b]
KLR3006	Pig	MIMV01000001-01000257	BHL83_RS05090-_RS05150	—[b]	BHL83_RS05150-_RS05145[b] & BHL83_RS05095-_RS05090
I5007	Pig	NC_021494-_021504	LRI_RS04175-_RS04240	—	LRI_RS04175-_RS04185 & LRI_RS04235-_RS04240
KLR2006	Pig	MIMN01000001-01000096	—	—	—
100-23	Rat	IMG: 2500069000	2500070887-2500070904 ("Lr_70887-_70904")	2500070902 ("Lr_70902")	2500070903-2500070904 ("Lr_70903-_70904")
TD1	Rat	CP006603	N134_05915-_05970	N134_05915	N134_05970-_05965
I49	Mouse	NZ_CP015408	A4V07_RS03925-_RS03975	A4V07_RS03935-_RS03925[a]	—
480_44	Mouse	MBLQ01000001-01000154	BBP10_RS03240-_RS03300	—[b]	BBP10_RS03300[b]
482_46	Mouse	MBLR01000001-01000188	BBP11_RS07510-_RS07570	—[b]	BBP11_RS07570[b]
482_54	Mouse	MBLS01000001-01000186	BBP12_RS00010-_RS00070	—[b]	BBP12_RS00010[b]
484_32	Mouse	MBLT01000001-01000523	BBP13_RS14455-_RS06080	—[b]	BBP13_RS14455[b]
484_39	Mouse	MBLU01000001-01000191	BBP14_RS02325-_RS11505	—[b]	BBP14_RS11505[b]
lpuph	Mouse	NZ_AEAX01000001-01000127	ECQ_RS11170-_RS0107090	—[b]	ECQ_RS11170[b] & ECQ_RS0107090-_RS0107085
mlc3	Mouse	NZ_AEAW01000001-01000126	ECM_RS0104210-_RS0104260	—[b]	ECM_RS10935[b,d]
LR0	Mouse	MWIJ01000001-01000075	B2G46_RS05885-_RS05955	—	B2G46_RS05900-_RS05885 & B2G46_RS05955-_RS05950
LTH5448	Sourdough	NZ_JOOG01000001-01000036	HN00_RS05550-_RS05600	HN00_RS06750[d]	HN00_RS06765-_RS09675[d]
LTH2584	Sourdough	NZ_JOSX01000001-01000025	LR3_RS10625-_RS06705	—	LR3_RS06575-_RS10625 & LR3_RS10630-_RS06580
TMW1.112	Sourdough	NZ_JOKX02000001-02000012	HF82_RS02815-_RS02880	—	HF82_RS02820-_RS02815
TMW1.656	Sourdough	IMG: 2534682350	LR4_00668-_00683	—	LR4_00083[d] & LR4_00669-_00668
CRL 1098	Sourdough	LYWI01000001-01000045	—	—	—
CECT8605	Cow	MWVS01000001-01000207	B5D07_RS10730-_RS10790	—[b]	B5D07_RS10730-_RS10735[b] & B5D07_RS10785-_RS10790
1366	Chicken	NBBG01000001-01000093	B6J74_RS03655-_RS03710	B6J74_RS03710	—
P43	Chicken	MCNS01000001-01000074	BFD03_RS09275-_RS09310	—	—
JCM 1081	Chicken	NBBD01000001-01000061	—	—	—
An71	Chicken	NFHN01000001-01000119	—	—	—
An166	Chicken	NFKV01000001-01000105	—	—	—
CSF8	Chicken	NBBE01000001-01000107	—	—	—
JCM 1112[T]	Human	AP007281	—	—	—
DSM 20016[T]	Human	CP000705	—	—	—
ATCC 55730 (SD2112)	Human	CP002844-002848	—	—	—
ATCC PTA-6475 (MM4-1a)	Human	ACGX02000001-02000007	—	—	—
ATCC PTA-4659 (MM2-3)	Human	GG693756-693850	—	—	—
CF48-3a	Human	GG693664-693755	—	—	—
IRT	Human	CP011024	—	—	—
MD IIE-43	Human	MUZP01000001-01000087	—	—	—
M27U15	Human	NBBH01000001-01000181	—	—	—
MM34-4a	Human	NBBF01000001-01000179	—	—	—
HI24	Curd	MAGB01000001-01000200	—	—	—

[a] Gene incorrectly annotated in genome; full-length SRRP translated in one ORF

[b] Possible intact SRRP with partial gene sequence at the end of a draft genome contig, in some cases, annotated incorrectly as two pseudogene fragments

[c] Strain KLR3003 has a truncated SecA2-SecY2 cluster with only genes encoding SecA2, GtfA, GtfB, a small hypothetical protein and a pseudo-SRRP

[d] Cases where the SRRP or pseudo-SRRP genes are unlinked to the SecA2-SecY2 gene cluster

Table S3: The top ten structural homology results for $SRRP_{53608}$-$BR_{262-571}$ from the Dali Server

The results are sorted in order of decreasing Z-score. Additional measures of agreement between the input and template model is given via the RMSD. 'Lalign' indicates the number of residues that are aligned between the query and template, for which the modelling statistics are provided.

Protein	Z-score	RMSD	Lalign	Sequence ID (%)	PDB ID	Reference
Bordetella bronchiseptica pertactin extracellular domain (Prn-E)	18.5	2.8	233	13	2IOU	(26)
Pectobacterium carotovorum Poly-galacturonase	18.0	3.1	242	13	1BHE	(27)
Shigella phage Sf6 tailspike protein	17.9	3.1	243	8	4URR	(28)
Bordetella pertussis P.69 pertactin	17.8	2.6	214	11	1DAB	(29)
E. coli phage HK620 tail spike protein	17.3 -17.8	3.1	245	11	2X6X, 4XKW, 4XMY, 4X19, 4XON, 4XNF, 4XGF, 4XN3, 2VJJ, 4XOP, 4XLH, 4XLC, 2VJI, 4XNO, 4YEJ, 4XKV, 4XOR, 4XM3, 4XLF, 4XOT, 4XGH, 2X6Y, 4XLE, 4XLA, 2X6W, 4YEL, 4XR6, 4XQI	(30)
T. maritima exopoly-galacturonase	17.8	3.1	246	9	3JUR	(31)

Yersinia enterocolitica Family 28 glycoside hydrolase	17.8	3.2	251	8	2UVE, 2UVF	(32)
Bacillus sp. inulin fructotransferase	17.3	2.9	236	11	2INU, 2INV	(33)
Ruminiclostridium thermocellum poly-saccharide lyase	17.2	2.6	206	12	4PHB	(34)
Azotobacter vinelandii mannuronan C5 epimerase	17.1	3.0	238	12	2PYG, 2PYH	(35)

Table S4: $SRRP_{53608}$-BR probe of Agata array v1.0

Specific Hits	Non-specific Hits (antibody control)
1) Carboxymethyl Cellulose	1) Mannan (ivory nut)
2) Carboxymethyl Cellulose 4M	2) Xylan (beechwood)
3) Lime pectin DE: 15% (B15)	3) MLG Lichenan, β-glucan (1→3),(1→4)-β-D-glucan
4) Lime pectin DE: 15% (B34)	4) β-glucan (Yeast), (1→6),(1→3)-β-D-glucan
5) Lime pectin DE: 43% (B43)	5) β-glucan (oat), (1→3),(1→4)-β-D-glucan
6) Lime pectin DE: 64% (B64)	6) β-1,3-glucan (Euglena gracilis) (β-1→3-Glucan)
7) Lime pectin DE: 71% (B71)	7) Pachyman
8) Lime pectin DE: 31% (F43)	8) Locust bean gum from Ceratonia siliqua seeds
9) Pectin with DE 1% & DA 0%, basic hydrolysis of SBP6230	9) Gum guaic (tree resin)
10) Pectin with DE 9% & DA 15%, basic hydrolysis of SBP6230	10) Karaya gum
11) Pectin with DE 25% & DA 16%, basic hydrolysis of SBP6230	11) Tragacanth gum
12) Pectin with DE 31% & DA 24%, basic hydrolysis of SBP6230	12) Ghatti gum
13) Pectin with DE 46% & DA 26%, basic hydrolysis of SBP6230	13) Gum arabic, Glycoproteins
14) Pectin with DE 53% & DA 26%, basic hydrolysis of SBP6230	14) Amylose (potato)
15) Sugar beet Pectin	15) Amylopectin (potato)
16) RGI (soy bean)	
17) Polygalacturonic acid from citrus pectin (Danisco)	
18) Lemon pectin	
19) Apple pectin	
20) Pectin (CP Kelco)	
21) RGI potato de-galactanated saponified	
22) RGI potato de-galactanated+arabinanased saponified	
23) RGI potato de-arabinanase saponified	
24) Modified hairy regions from apple RGI	
25) Modified hairy regions from sugar-beet RGI	
26) Feruloylated pectin	

Table S5 List of primers used for the amplification of *L. reuteri srrp* genes

OLIGO NAME	SEQUENCE (5' to 3')
3c6-SRRP-rev	GCCAAATTAATTACTTGTTTTGG
20-2-SRRP-rev	GCTTTTTATATAAATTAATTACTTGTTTTGG
pg3b-SRRP-rev	GCTTTTTATATAAACTAATTACTTGTTTTGG
3c6_ps-SRRP-for	GGGTTATCTATGGCAAAAAACAATAAAG
3c6_ps-SRRP-rev	GTATGCATCTAATTATTTATCTTCGTTGTG
pg3b_ps-SRRP-for	GAGTATATGTCTAAACGTAAAAATGAAC
pg3b_ps-SRRP-rev	CTTATGCACCTAATTATTTATCTTCGTTGTG
Lr_1105r2	GGACAATCTATGGTCAAACATAAGCAAG
Lr_1105f2	GTTGAAACTACCCCTTCTTCTTCTTTG

SI references

1. Crost EH*, et al.* (2013) Utilisation of mucin glycans by the human gut symbiont *Ruminococcus gnavus* is strain-dependent. *Plos One* 8(10):e76341.
2. Petersen TN, Brunak S, von Heijne G, & Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8(10):785-786.
3. Sievers F*, et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7.
4. Berrow NS*, et al.* (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* 35(6):e45.
5. Liu H & Naismith JH (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* 8:91.
6. Winter G (2010) xia2: an expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* 43:186-190.
7. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* 62(Pt 9):1002-1011.
8. Battye TG, Kontogiannis L, Johnson O, Powell HR, & Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):271-281.
9. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D* 62:72-82.
10. Winn MD*, et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* 67:235-242.
11. Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D* 67:282-292.
12. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.
13. Langer G, Cohen SX, Lamzin VS, & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7):1171-1179.
14. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126-2132.
15. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240-255.
16. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12-21.

17. Olsson MH, Sondergaard CR, Rostkowski M, & Jensen JH (2011) PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *J Chem Theory Comput* 7(2):525-537.
18. Maier JA*, et al.* (2015) ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* 11(8):3696-3713.
19. Horn HW*, et al.* (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J Chem Phys* 120(20):9665-9678.
20. Hinterdorfer P*, et al.* (2002) Surface attachment of ligands and receptors for molecular recognition force microscopy. *Colloid Surface B* 23(2-3):115-123.
21. Hutter JL & Bechhoefer J (1993) Calibration of atomic-force microscope tips. *Rev Sci Instrum* 64(7):1868-1873.
22. Gunning AP*, et al.* (2013) Mining the "glycocode"-exploring the spatial distribution of glycans in gastrointestinal mucin using force spectroscopy. *FASEB J.* 27(6):2342-235
23 Pedersen HL, *et al.* (2012) Versatile high resolution oligosaccharide microarrays for plant glycobiology and cell wall research. *J Biol Chem* 287(47):39429-39438.
24. Sorensen I, Pedersen HL, & Willats WG (2009) An array of possibilities for pectin. *Carbohydr Res* 344(14):1872-1878.
25. He ZL, *et al.* (2016) Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res* 44(W1):W236-W241.
26. Miller JL*, et al.* (2008) Selective ligand recognition by a diversity-generating retroelement variable protein. *PLoS Biol.* 6(6):1195-1207.
27. Pickersgill R, Smith D, Worboys K, & Jenkins J (1998) Crystal structure of polygalacturonase from *Erwinia carotovora ssp. carotovora*. *J. Biol. Chem.* 273(38):24660-24664.
28. Kang Y*, et al.* (2016) Bacteriophage tailspikes and bacterial O-antigens as a model system to study weak-affinity protein–polysaccharide interactions. *J. Am. Chem. Soc.* 138(29):9109-9118.
29. Emsley P, Charles IG, Fairweather NF, & Isaacs NW (1996) Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* 381(6577):90-92.
30. Broeker NK*, et al.* (2013) Single amino acid exchange in bacteriophage HK620 tailspike protein results in thousand-fold increase of its oligosaccharide affinity. *Glycobiology* 23(1):59-68.
31. Pijning T, van Pouderoyen G, Kluskens L, van der Oost J, & Dijkstra BW (2009) The crystal structure of a hyperthermoactive exopolygalacturonase from *Thermotoga maritima* reveals a unique tetramer. *FEBS Letters* 583(22):3665-3670.
32. Abbott DW & Boraston AB (2007) The structural basis for exopolygalacturonase activity in a family 28 glycoside hydrolase. *J. Mol. Biol.* 368(5):1215-1222.
33. Jung WS*, et al.* (2007) Structural and functional insights into intramolecular fructosyl transfer by inulin fructotransferase. *J. Biol. Chem.* 282(11):8414-8423.
34. Close DW, D'Angelo S, & Bradbury ARM (2014) A new family of [beta]-helix proteins with similarities to the polysaccharide lyases. *Acta Crystallographica Section D* 70(10):2583-2592.
35. Rozeboom HJ*, et al.* (2008) Structural and mutational characterization of the catalytic A-module of the mannuronan C-5-epimerase AlgE4 from *Azotobacter vinelandii*. *J. Biol. Chem.* 283(35):23819-23828.