Structural basis for the role of Serine-Rich Repeat Proteins from *Lactobacillus reuteri* in gut microbe-host interactions

Saannya Sequeira[1], Devon Kavanaugh[2], Donald MacKenzie[2], Tanja Šuligoj[2], Samuel Walpole[1], Charlotte Leclaire[2], A. Gunning[2], Dimitrios Latousakis[2], William Willats[3], Jesus Angulo[1], Changjiang Dong[1], Nathalie Juge[4]

[1]University of East Anglia, [2]Quadram Institute Bioscience, [3]Newcastle University, [4]Quadram Institute

Lactobacillus reuteri, a Gram-positive bacterial species inhabiting the gastrointestinal tract of vertebrates displays remarkable host adaptation. Previous mutational analyses of rodent strain *L. reuteri* 100-23C identified a gene encoding a predicted surface-exposed serine-rich repeat protein (SRRP$_{100\text{-}23}$) that was vital for *L. reuteri* biofilm formation in mice. SRRPs have emerged as an important group of surface proteins on many pathogens but no structural information is available in commensal bacteria. Here we report the 2.00 Å and 1.92 Å crystal structures of the binding regions (BRs) of SRRP$_{100\text{-}23}$ and SRRP$_{53608}$ from *L. reuteri* ATCC 53608, revealing a unique "β-solenoid" fold in this important adhesin family. BR-SRRP$_{53608}$ bound to host epithelial cells and DNA at neutral pH and recognised polygalacturonic acid (PGA), rhamnogalacturonan I or chondroitin sulfate A at acidic pH. Mutagenesis confirmed the role of the BR putative binding site in the interaction of BR-SRRP$_{53608}$ with PGA. Long molecular dynamics simulations showed that SRRP$_{53608}$ undergoes a pH-dependent conformational change. Together, these findings shed new mechanistic insights into the role of SRRPs in host-microbe interactions and open new avenues of research into the use of biofilm-forming probiotics against clinically important pathogens.

SRRP | Lactobacillus reuteri | biofilm | mucin | adhesin

Introduction

The gastrointestinal (GI) tract of vertebrates is colonised by a complex microbial community dominated by bacteria known as the gut microbiota. These bacteria, by having a profound influence on vertebrate physiology, metabolism, and immune functions, play an important role in the health of the host (1). Manipulating the microbiota for the benefit of the host requires an understanding of the molecular mechanisms that govern host-microbe interactions.

Lactobacillus reuteri, a Gram-positive bacterial species that colonises the gut of a range of vertebrate species, has been used as a model to study host adaptation of gut symbionts (2). The ecological strategies of *L. reuteri* are fundamentally different in humans and animals (3). In rodents, pigs, chickens and horses, lactobacilli form large populations in proximal regions of the GI tract, and they adhere directly to the stratified squamous epithelium present at these sites (3). In contrast, stratified squamous epithelia are absent in the human gut where the lactobacilli population is less important and restricted to the mucus layers and intestinal crypts (3). Genome comparisons of *L. reuteri* strains originating from different hosts identified lineage-specific genomic content that reflects the niche differences in the GI tract of vertebrates (4). Experiments in *Lactobacillus*-free mice to measure the ecological fitness of strains originating from different hosts supported host adaptation, as only rodent strains colonised mice efficiently (5). Furthermore the ability of *L. reuteri* to form epithelial biofilms in the mouse forestomach of mono-associated mice was strictly dependent on the strain's host origin (6). Recent studies revealed that rodent strains were particularly successful in colonising mice, confirming previous findings of host adaptation (7).

The ecological significance of a subset of genes from the rodent-specific *L. reuteri* 100-23C strain was demonstrated in the context of the murine gut (6). This mutational analysis revealed that genes encoding proteins involved in epithelial adherence, specialised protein transport, cell aggregation, environmental sensing and cell lysis contributed to biofilm formation and colonisation (6). In particular, the inactivation of a gene encoding a predicted serine-rich repeat protein (Lr_70902; SRRP$_{100\text{-}23}$) surface adhesin completely abrogated colonisation of the mouse forestomach (6). SRRP$_{100\text{-}23}$ is the primary cell wall-associated protein of strain 100-23C that is secreted through an accessory SecA2-SecY2 pathway during *in vivo* growth (6). Recent analysis of the completed genome of pig isolate *L. reuteri* ATCC 53608 revealed the presence of an accessory SecA2-SecY2 secretion system with an associated SRRP that shared the same domain organisation as SRRP$_{100\text{-}23}$ (8).

SRRPs belong to a growing family of adhesins in Gram-positive bacteria that mediate attachment to a variety of host and bacterial surfaces, and many of them are virulence factors that contribute to bacterial pathogenesis and biofilm formation (9). SRRPs are characterised by: 1) two heavily glycosylated serine-rich regions (SRRs), 2) one or two species-unique non-repeat regions (NR domains, which includes the BR domain, referred to as "binding region") towards the N-terminus that facilitates specific interactions with a diverse array of host receptors, and share

---

Significance

Gut bacteria play a key role in health and disease but the molecular mechanisms underpinning their interaction with the host remain elusive. The serine-rich repeat proteins (SRRPs) are a family of adhesins identified in many Gram-positive pathogenic bacteria. We previously showed that beneficial bacterial species found in the gut also express SRRPs and that SRRP was required for the ability of *Lactobacillus reuteri* strain to colonise mice. Here, our structural and biochemical data reveal that *L. reuteri* SRRP adopts a "β-solenoid" fold not observed in other structurally characterised SRRPs, and functions as an adhesin via a pH-dependent mechanism, providing first structural insights into the role of these adhesins in biofilm formation of gut symbionts.

Reserved for Publication Footnotes

137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
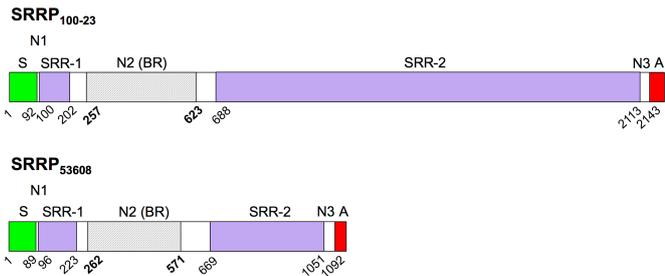195
196
197
198
199
200
201
202
203
204

Fig. 1. **Schematic showing domain organisation of precursor SRRPs from *L. reuteri*.** The two proteins are drawn to scale. Domains are labelled as follows: S, secretion signal sequence; N1, non-repeat region 1; SRR-1, serine-rich region 1; N2 (BR), non-repeat region 2 (putative binding region); SRR-2, serine-rich region 2; N3, non-repeat region 3; A, cell wall anchor including LPXTG motif. The start aa position is indicated below each domain. Regions of the BR that were resolved by crystallography are shaded gray, and span from from aa 257 to 623 for $SRRP_{100-23}$ and from aa 262 to 571 for $SRRP_{53608}$.
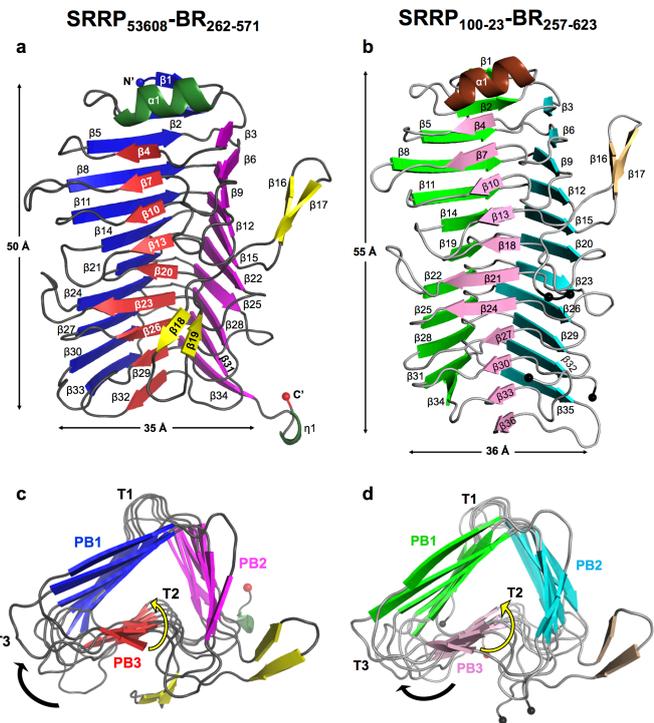

Fig. 2. **Crystal structures of $SRRP_{53608}$-$BR_{262-571}$ and $SRRP_{100-23}$-$BR_{257-623}$.** (a)Longitudinal view of $SRRP_{53608}$-$BR_{262-571}$. β-strands in PB1, PB2 and PB3 are shown in deep blue, magenta and red, respectively; α-helices are dark green and β-strands of the loop are yellow. (b) Longitudinal view of $SRRP_{100-23}$-$BR_{257-623}$. PB1, PB2 and PB3 β-strands are in light green, cyan and pink, respectively. The α-helix is in brown and loop β-strands are in beige. The black spheres indicate the gaps in the model between aa 413-421 and 568-583. Cross-section of (c) $SRRP_{53608}$-$BR_{262-571}$ and (d)$SRRP_{100-23}$-$BR_{257-623}$ β-solenoid superhelices along the helical axis, from N- to C-terminal. β1 and α2 in both are omitted for clarity. The black arrow shows the direction in which the polypeptide chains fold around the helical axis, yielding a right-handed superhelix. The helical twist down each β-sheet is indicated by yellow arrows: along the helical axis, the β-strands in each parallel β-sheet increasingly twist towards the left with respect to each other.

little sequence homology to each other, and 3) a C-terminal cell wall anchor domain (9, 10). Export of SRRPs onto the bacterial surface occurs through a dedicated non-canonical Sec translocase, Sec-Y2A2, following recognition of an extended atypical signal sequence peptide of around 90 aa at the N-terminus (10, 11). The domain organisation of SRRPs are highly conserved

205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272

in pathogenic streptococci and staphylococci, and includes Srr-1 and Srr-2 of *Streptococcus agalactiae*, PsrP of *Strep. pneumoniae*, Fap1 of *Strep. parasanguinis*, GspB and Hsa of *Strep. gordonii*, SraP of *Staphylococcus aureus* and SrpA homologues from *Strep. sanguinis* and *Strep. cristatus* (9, 12, 13). However, the structure and function of SRRPs in commensal bacteria have not yet been determined. Here we used a number of complementary approaches to provide a structural basis for the role of *L. reuteri* SRRPs (*Lr*SRRPs) in bacterial adaptation to the host. We showed that the *Lr*SRRP-BR adopts a right-handed parallel β-helical or "β-solenoid" fold, not observed in other structurally characterised SRRPs, and functions as an adhesin *via* a pH-dependent mechanism. These findings provide new insights into the role of *Lr*SRRPs in biofilm formation, and structural insights into intra- and inter-species adhesins across Gram-positive pathogenic and commensal bacteria.

Results

Bioinformatics analysis of SRRPs from lactobacilli

SRRPs and corresponding specialized secretion systems are being defined in a growing number of pathogens but their occurrence and characterization in commensal bacteria has only been reported infrequently (14, 15). Our bioinformatics analysis of lactobacilli genomes revealed genes encoding fully functional SRRPs and SecA2-SecY2 secretion systems in a number of *Lactobacillus* species, including strains of *Lactobacillus reuteri*, *Lactobacillus oris*, *Lactobacillus salivarius*, *Lactobacillus johnsonii*, and *Lactobacillus fructivorans*, with none found so far in other major lactobacilli species such as *Lactobacillus plantarum* (**SI Appendix, Table S1**). In other cases, strains possessed what appeared to be only an incomplete SecA2-SecY2 gene cluster, a SRRP that lacked a C-terminal cell wall anchor (possibly the result of a pseudogene, capable of exporting a SRRP extracellularly which would not be covalently linked to the cell surface) or an obvious pseudo-SRRP whose domains were encoded by at least two adjacent open reading frames (ORFs). These include strains of *Lactobacillus gasseri*, *Lactobacillus rhamnosus*, *Lactobacillus murinus*, *Lactobacillus nagelii* and *Lactobacillus mucosae*, a species closely related to *L. reuteri* (16) (**SI Appendix, Table S1**). When some strains of lactobacilli harbour two SRRPs, at least one of them is encoded by pseudogene fragment(s), highlighting their loss of function due to lack of selective pressure.

Analysis of 58 available genome-sequenced *L. reuteri* strains showed that homologues of functional SRRPs (and the corresponding linked SecA2-SecY2 gene cluster) were exclusively found in some rodent and pig isolates with the exception of one chicken strain and one sourdough strain (previously reported to be of intestinal origin with a genome content similar to that of the model rodent isolate 100-23 (17)) (**SI Appendix, Table S2**). The putative *Lr*SRRPs from *L. reuteri* rodent strain 100-23C ($SRRP_{100-23}$) and of pig strain ATCC 53608 ($SRRP_{53608}$) possess a LPXTG cell wall anchor and display characteristics of a protein secreted through the SecA2-SecY2 system, i.e. the presence of an unusually long N-terminal signal peptide and two extremely serine-rich regions (SRR-1 and SRR-2), the second of which contains many repeat motifs (**Fig. 1**).

Further comparative sequence analysis of SRRP-BR domains was carried out from a total of 76 lactobacilli SRRPs/pseudo-SRRPs and 18 pathogen/clinical-associated SRRPs to generate a Neighbour-Joining phylogram (**SI Appendix, Fig. S1**). In the case of most *L. reuteri* strains, SRRP-BRs formed groups relating to the host or source from which the strain was isolated, such as the two main groups of pig-derived SRRP-BRs/pseudo-SRRP-BRs and the one main group of rodent/sourdough-derived BRs. Similar relationships were observed for the BRs from *L. oris*, *L. salivarius* and the pathogenic streptococcal BRs. There were a few exceptions where SRRP-BRs/pseudo-SRRP-BRs of *L. reuteri*

273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340

341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408

Table 1. Data collection and refinement statistics for SeMetSRRP$_{53608}$-BR, SRRP$_{53608}$-BR and SRRP$_{100-23}$-BR

Protein	SeMetSRRP$_{53608}$-BR	SRRP$_{53608}$-BR	SRRP$_{100-23}$-BR
Data Collection Statistics			
PDB ID		5NXK	5NY0
Beamline	i04	i03	i04
Wavelength (Å)	0.9792	0.9762	0.9002
Space group	P 3$_2$ 2 1	P 3$_2$ 2 1	H 3 2
Cell parameters: a, b, c (Å)	148.40, 148.40, 110.73	146.70, 146.70, 110.42	162.36, 162.36, 146.78
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	74.20 – 2.73 (2.80 – 2.73)	73.35 –1.92 (1.95 –1.92)	63.40 –2.00 (2.05 –2.00)
I/σI	29.7 (6.2)	9.9 (2.2)	16.8 (2.2)
Unique reflections	37723 (2781)	104897 (5168)	50080 (3650)
Completeness (%)	99.9 (99.9)	100.0 (99.6)	100.0 (99.9)
Multiplicity	39.0 (40.9)	18.0 (15.5)	9.2 (9.5)
R$_{merge}$	0.134 (1.134)	0.278 (3.087)	0.061 (0.945)
R$_{meas}$	0.136 (1.148)	0.286 (3.191)	0.069 (1.060)
R$_{pim}$	0.022 (0.178)	0.067 (0.797)	0.031 (0.472)
CC$_{1/2}$	1.000 (0.984)	0.997 (0.668)	0.998 (0.714)
Refinement Statistics			
Molecules per AU	3	3	1
Total atoms		7896	2786
Water molecules		999	200
R$_{factor}$		0.1570(0.2458)	0.2241(0.2116)
R$_{free}$		0.1825(0.2680)	0.2445(0.2517)
Ramachandran analysis			
Most favoured		95.90	94.10
Allowed		2.99	4.70
Outliers		0.11	1.20
R.M.S.D.			
Bonds (Å)		0.008	0.013
Angles (°)		1.240	1.400
Mean atomic B-factor (Å^2)		30.60	48.40
Molprobity score		1.36 (98th percentile)	1.60 (94rd percentile)
Clashscore		2.98 (99th percentile)	3.98 (99th percentile)

strains crossed this host-specific divide, such as the relatedness of (i) the pseudo-SRRP-BRs from rodent strains lpuph, LR0 and TD1 and a number of pig strain SRRP-BRs; (ii) the SRRP-BR from chicken isolate 1366 and the pseudo-SRRP-BRs from three pig strains KLR2001, KLR3005 and pg-3b; and (iii), most importantly, the BR domain from SRRP$_{100-23}$ of rodent origin and one group of porcine SRRP-BRs that included SRRP$_{53608}$-BR. The SRRP$_{53608}$-BR and SRRP$_{100-23}$-BR shared ~49% aa identity (ID). This compares with the very low aa ID of <15% between SRRP$_{53608}$-BR or SRRP$_{100-23}$-BR and pathogenic bacterial BRs such as GspB-BR, Fap1-NRα, Hsa-BR, PsrP-BR, Srr-1-BR and Srr-2-BR. Typically, SRRP-BRs from different *Lactobacillus* spp. showed low homology between each other but one exception was the pseudo-SRRP-BR from *L. mucosae* pig strain LM1 and the predominantly pseudo-SRRP-BRs from 18 *L. reuteri* pig strains and one *L. reuteri* bovine strain, providing some insight into evolutionary relationships between these two related species (16, 19). We confirmed the presence of the full-length *srrp* gene in five pig strains of *L. reuteri* by PCR and that the encoded SR-RPs were also secreted extracellularly, as previously shown for SRRP$_{100-23}$ during growth *in vitro* (6) (**SI Appendix, Fig. S2**).

Crystal structures of *L. reuteri* SRRP-BR reveals a unique fold within the SRRP family

SRRP$_{53608}$-BR and SRRP$_{100-23}$-BR crystals were obtained via *in situ* limited proteolysis of the full-length *Lr*SRRP-BR proteins with α-chymotrypsin and thermolysin, respectively. The structure of SRRP$_{53608}$-BR was determined at 1.92 Å resolution between residues 262-571, excluding 33 and 101 residues at the N- and C-termini, respectively (**Fig. 2a**). The structure of SRRP$_{100-23}$-BR was determined at a resolution of 2.00 Å. The model contains residues 256-623, excluding 58 and 134 residues from the N- and C-termini, respectively. There are two gaps in the structure indicated by black spheres in **Fig. 2b**, which could not be modelled. The first gap of 9 residues is from aa 413-421 and the second gap of 16 residues is from aa 568-583. The data collection and refinement statistics for SRRP$_{53608}$-BR and SRRP$_{100-23}$-BR are provided in **Table 1**.

Overall fold of LrSRRP-BRs. The structures of SRRP$_{53608}$-BR$_{262-571}$ and SRRP$_{100-23}$-BR$_{257-623}$ share 43% sequence identity with an RMSD of 0.912 over 293 aligned Cα atoms. Both structures adopt a solenoid-type fold comprising β-strands coiled in a repetitive pattern to form a right-handed helix with three parallel β-sheets (**Fig. 2**). A Dali search (20) revealed structural ho-
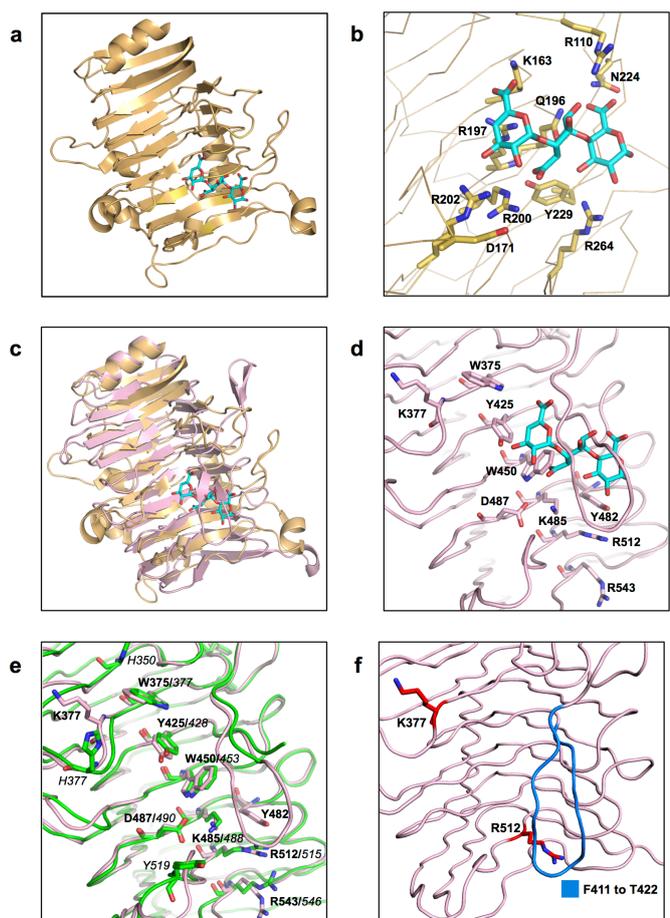
Fig. 3. **Comparison of the TM-Pel binding pocket with the putative binding site (PuBS) of SRRP$_{53608}$-BR and SRRP$_{100-23}$-BR.** (a) Crystal structure of pectate lyase from *Thermotoga maritima* (TM-Pel, 3ZSC), (orange), in complex with TGA (cyan). (b) A close-up view of the TM-Pel binding pocket, with residues involved in TGA binding represented as sticks. (c) Superposition of SRRP$_{53608}$-BR$_{262-571}$ (light pink) and TM-Pel (orange) structures, with an RMSD of 2.63 over 210 residues, show that the PuBS of the former overlaps with that of TM-Pel. (d) Surface exposed aromatic and charged residues on PB3 in SRRP$_{53608}$-BR$_{262-571}$ PuBS (light pink). This includes the aromatic residue triad W375, Y425, W450 and the basic residue triad K485, R512 and R543. (e) Solvent-exposed residues of PuBS in the overlaid structures of SRRP$_{53608}$-BR (light pink) and SRRP$_{100-23}$-BR (green), with an RMSD of 0.912 over 293 residues, showing that several solvent exposed residues in both binding sites are conserved. Residues in bold are from SRRP$_{53608}$-BR and those in italics font are from SRRP$_{100-23}$-BR. (f) Residues mutated for functional analysis of SRRP$_{53608}$-BR. Single mutants were created by substituting the residues in red (K377 or R512) with alanine. Residues in blue from F411 to T422 in the lower loop were deleted to generate the ΔF411-T422 deletion mutant.

Fig. 4. . **pH-dependent conformational change affects putative binding site (PuBS) as predicted by MD simulations.** (a) Combined surface and cartoon models of SRRP$_{53608}$BR$_{262-571}$ (pink) and SRRP$_{100-23}$BR$_{257-623}$ (green). (b) Surface representation of SRRP$_{53608}$-BR$_{262-571}$ at pH 7.4 and pH 4.0 in the same orientation as in a), showing surface electrostatics (top) and surface exposed putative binding residues (bottom, green). At pH 4.0, PuBS exhibits a more positive electrostatic potential as well as an open conformation that exposes a greater number of putative binding residues to the solvent. Coordinates were obtained from representative frames of each respective molecular dynamics trajectory (see methods). Surface electrostatics calculated in PyMOL and coloured as blue (positive), white (neutral), and red (negative).

mology of *Lr*SRRP-BRs with proteins predominantly belonging to CATH superfamilies 2.160.20.10 and 2.160.20.20 (21). These include pectate lyase C (PelC)-like proteins with right-handed β-superhelical topology (22, 23) such as PelC from the plant pathogen *Dickeya didantii*, the first parallel β-helix protein to be reported (24). Other examples include pectate lyases from *Bacillus* spp. (BsPel, Pel15 and Bsp165PelA (25-28)), those from *D. didantii* (previously known as *Erwinia chrysanthemi*) (Pel9A and PelI) (29-31) and from *Thermatoga maritima* (TM-Pel) (32); pectin lyases from *Aspergillus niger* (PelA and PelB) (33, 34) and a rhamnogalacturonase from *Aspergillus aculeatus* (RGaseA) (35). These enzymes all act on pectin, a structural polysaccharide of plant cell walls (36). The structures of *Lr*SRRP-BRs are strickingly similar to those of PelC-like proteins with RMSDs ranging
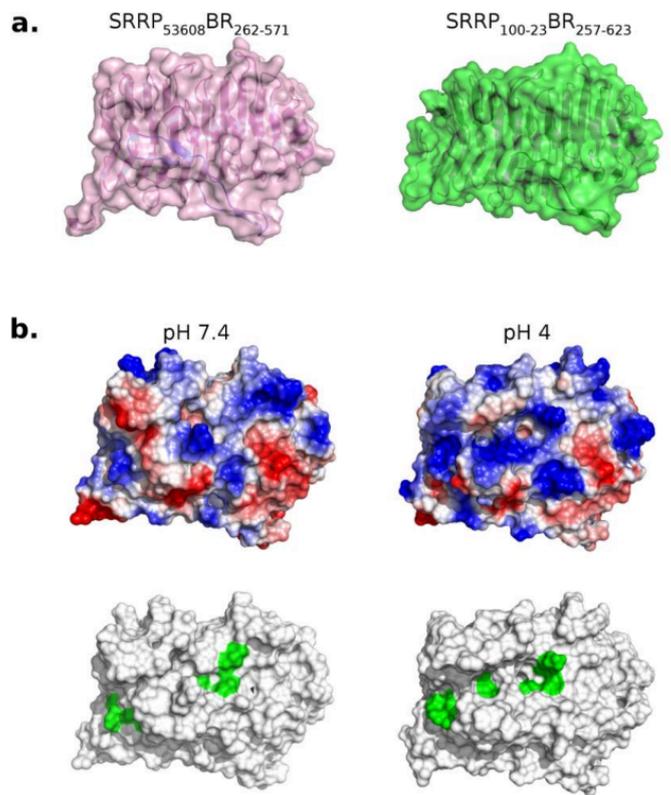
from 2.38 to 3.37 Å over 164 to 233 aligned Cα atoms, despite sharing only 8-23% sequence identity. As shown in Fig. 2, the parallel β-sheets of *Lr*SRRP-BRs are labelled as PB1, PB2 and PB3, and the disordered turns connecting consecutive β-strands are referred to as T1, T2 and T3, following the convention of Yoder et al. (37). Other shared structural features include: i) amino acid stacks; ii) an antiparallel β-sandwich arrangement between PB1 and PB3 to which PB2 is perpendicular; and iii) the protrusion of flexible domains from the core β-solenoid body which is often involved in ligand binding. In SRRP$_{53608}$-BR$_{262-571}$ and SRRP$_{100-23}$-BR$_{257-623}$, this flexible domain takes the form of a "two-winged propeller-like" loop which originates from β15 in PB2 (the "upper loop") and folds back into β20 and β18, respectively, in PB3 (the "lower loop").

The Dali search also revealed structural similarity to certain extracellular adhesive proteins with PelC-like folds, such as pertactins and bacteriophage tail-spike proteins (**SI Appendix, Table S3**). The structure of SRRP$_{53608}$-BR$_{262-571}$ displays an RMSD of 2.81 Å over 220 aligned Cα atoms and 3.48 Å over 190 aligned Cα atoms to P.69 pertactin and phage P22 TSP, respectively. P.69 pertactin from the pathogen *Bordetella pertussis* facilitates adhesion to mammalian epithelial proteins via a conserved 'Arg-Gly-Asp' motif and two proline-rich regions (41). Phage P22 TSP is an endorhamnosidase, acting on the *O*-antigen of *Salmonella typhimurium*; sugar-binding features observed from the structure of the TSP-*O*-antigen complex are hydrophobic stacking of

Table 2. Binding kinetics of SRRP$_{53608}$-BR variants against PGA as analysed by Bio-Layer Interferometry

	K$_D$ (M)	k$_{on}$(1/Ms)	k$_{on}$ Error	k$_{dis}$(1/s)	k$_{dis}$ Error	Full X^2	Full R^2
Wild-type	2.37E-07	3.03E+02	1.64E+00	7.17E-05	4.92E-06	0.231987	0.997077
ΔF411-T422	2.84E-06	3.14E+02	2.71E+00	8.91E-04	9.29E-06	0.314129	0.993248
R512A	4.66E-06	2.46E+02	2.59E+00	1.15E-03	1.03E-05	0.094123	0.992496
K377A	5.24E-06	2.58E+02	3.92E+00	1.35E-03	1.59E-05	0.114659	0.985439
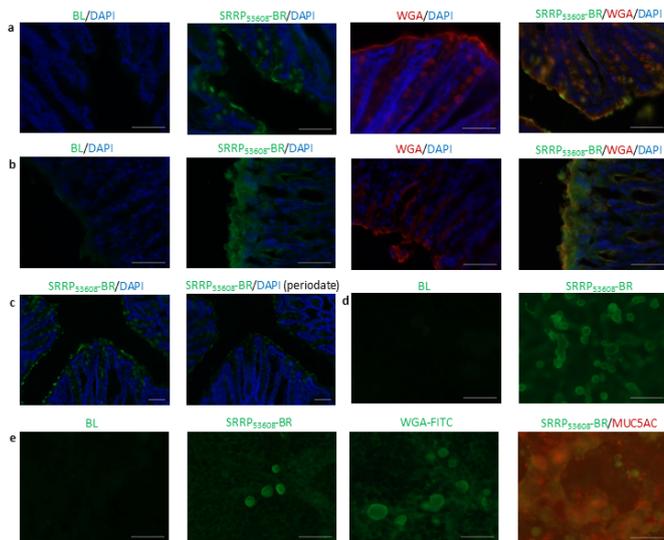
Fig. 5. Adhesion of SRRP$_{53608}$-BR to GI tissue. Immunostaining pattern for SRRP$_{53608}$-BR on cryosections of mouse colon (**a**) and stomach (**b**) correlates with WGA lectin staining. SRRP$_{53608}$-BR binding to mouse colonic sections following sodium periodate treatment at pH 4.5 is significantly reduced (**c**).SRRP$_{53608}$-BR binds to epithelial cells HT29 (**d**) and mucus producing HT29-MTX cells (**e**). Immunostaining pattern for SRRP$_{53608}$-BR on HT29-MTX cells correlates with WGA lectin staining and partly with anti-MUC5AC staining. Cell nuclei of tissue sections were counterstained with DAPI. Scale bars 50 μm.

aromatic sidechains with sugar pyranose, and H-bonds to polar and ionic sidechains (48). Another β-superhelical extracellular adhesive protein is the N-terminal TPS fragment (30 kDa) of a mature filamentous hemagglutinin (FHA) protein (49) from *Bordetella sp.*, which also has an 'Arg-Gly-Asp' motif that recognises macrophage CR3, a heparin-binding domain, as well as a carbohydrate recognition domain, for adhesion to lung epithelial cilia (RMSD of 3.18 Å to SRRP$_{53608}$-BR$_{262-571}$ over 197 aligned Cα atoms) (50). However, no such 'Arg-Gly-Asp' motif was identified in any of the LrSRRP-BR proteins.

Putative binding sites of LrSRRP-BRs. Superposition of the SRRP$_{53608}$-BR$_{262-571}$ structure with that of pectate lyase TM-Pel in complex with trigalacturonic acid (TGA) (**Fig. 3c**) revealed a potential binding site in *Lr*SRRP-BRs. In TM-Pel, predominantly basic residues maintain polar contacts and salt bridges with the acidic TGA molecule as shown in **Fig. 3b**. In SRRP$_{53608}$-BR, the area under the "lower loop" corresponds to TM-Pel's binding site, and has four basic solvent exposed residues (K377 on T3 adjacent to β13 and a triad, K485, R512, R543 on β26, β29 and β32, respectively), one acidic residue (D487 on β26) and four solvent exposed aromatic residues (a triad of W375, Y425, W450 on β13, β20 and β23, respectively, and Y482 on T2, next to β26 – **Fig. 3d**). Given the role of these charged and aromatic residues for sugar binding in PelC like proteins, the aforementioned surface-exposed residues are postulated to form a putative binding site (PuBS) in SRRP$_{53608}$-BR. The positions of Y482 and R512 in the

PuBS are conserved with TM-Pel's binding site. Superposition of SRRP$_{100-23}$-BR$_{257-623}$ upon SRRP$_{53608}$-BR$_{262-571}$ revealed a high degree of conservation of aa residues between their PuBS (depicted in **Fig. 3e**). However, the most notable difference is that the "lower loop" in SRRP$_{100-23}$-BR$_{257-623}$ is five aa longer and includes two aromatic residues, H414 and Y415, although it could not be modelled in the crystal structure. In addition, solvent-accessible surface electrostatic potential maps of SRRP$_{53608}$-BR$_{262-571}$ and SRRP$_{100-23}$-BR$_{257-623}$ revealed that like the TM-Pel binding site, the PuBS in SRRP$_{53608}$-BR$_{262-571}$ is enveloped by positive electrostatic potential (**SI Appendix, Fig. S3**), whereas the corresponding region in SRRP$_{100-23}$-BR$_{257-623}$ is acidic, although this may be due to the absence of the (unmodelled) loop in SRRP$_{100-23}$-BR$_{257-623}$. Indeed, removing the lower loop in the SRRP$_{53608}$-BR$_{262-571}$ model from F411-T422 led to a reduced positive surface charge around the PuBS (**SI Appendix, Fig. S3**), implying that the residues in this loop may play a role in maintaining the basicity of the PuBS.

Molecular Dynamics (MD) simulations suggest a pH-dependent conformational change in *LrSRRP-BR* putative binding site

Starting from the crystal coordinates of SRRP$_{53608}$-BR$_{262-571}$, hydrogen coordinates were added to the structure according to known protein chemistry (see Methods). This included prediction of the protonation state of acidic and basic residues at both pH 4.0 and 7.4. The resulting models differed in that 14 acidic residues (E263, E269, D334, E338, E400, E409, E434, D448, E475, E481, D487, E518, E527, E566) were protonated at the carboxylate sidechain at pH 4.0 (**SI Appendix, Fig. S4**). Furthermore, four histidine residues (H311, H413, H493, H535) were singly protonated at either the δ- or ε-nitrogen positions at pH 7.4, but protonated at both positions at pH 4.0. Of all these residues, five of them (H413, D448, E481, D487, E518) are in close proximity to the PuBS. As a result, the PuBS exhibits a more positive surface electrostatic potential at pH 4.0 compared to pH 7.4 (**Fig. 4, SI Appendix, Fig. S5**), which is expected to facilitate binding to anionic polysaccharides. Furthermore, the majority of differentially protonated residues were found on the surface of SRRP$_{53608}$-BR$_{262-571}$, which predicts an interruption of key interactions between symmetry-related molecules, potentially explaining why no crystallisation could be achieved at pH 4.0 (**SI Appendix, Fig. S4**).

MD simulations in the microsecond timescale revealed a pH-dependent conformational change in the loop connecting β30 and β31, close to the PuBS, resulting in greater solvent exposure of putative binding residues (**Fig. 4**). This change is caused by rotation about the Cα-C bond (ψ) of I514 such as to facilitate hydrogen bond formation between the backbone carbonyl of I514 and the hydroxyl of the D487 carboxylic sidechain (**SI Appendix, Fig. S6**). This is only possible at pH 4.0 due to protonation of the D487 carboxylate group. Furthermore, a resulting reduction in steric interactions between I514 and R512 led to subtle sidechain rearrangements of R512 and Y482 (**SI Appendix, Fig. S7**), which may prearrange these residues for ligand recognition.

SRRP$_{53608}$-BR displays specific binding to polyanionic ligands *via* a low pH mode of adhesion
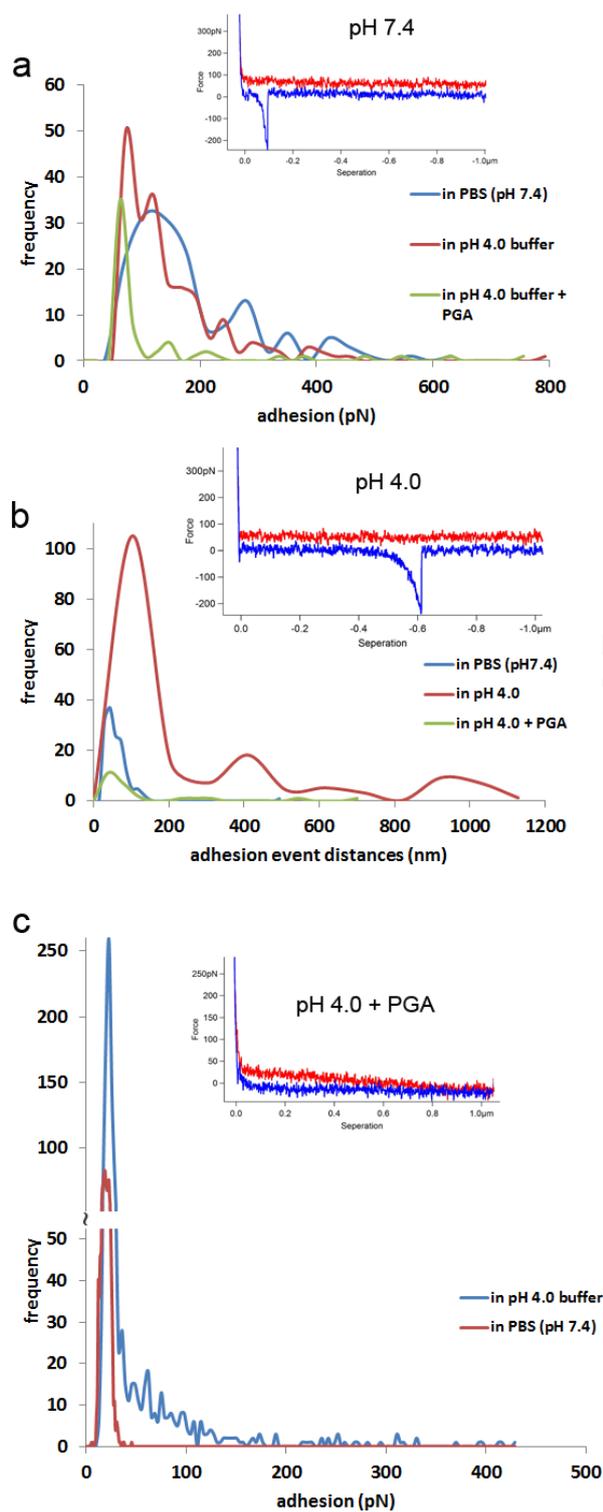
Fig. 6. **AFM force spectroscopy histograms of SRRP$_{53608}$-BR interacting with mucin or PGA** showing insets of example force curve for each assay. **(a-b)** AFM force spectroscopy histogram of SRRP$_{53608}$-BR interacting with mucin in buffers of neutral and acidic pH values or following PGA addition. **(a)** Quantification of adhesion values in buffers of neutral and acidic pH values or following PGA addition. **(b)** Quantification variations of the mucin adhesion event distances. **(c)** Quantification of adhesion values of SRRP$_{53608}$-BR interacting with PGA in buffers of neutral and acidic pH values.

Glycan arrays were first used in an attempt to identify the potential ligands of *Lr*SRRP-BR. Due to the reported binding

specificity of several SRRP-BRs from Gram-positive pathogenic bacteria to sialylated structures, we first tested the binding of SRRP$_{53608}$-BR against a sialoglycan microarray presenting over 70 synthetically recreated, naturally-occurring oligosaccharide structures with diverse sialic acid forms, glycosidic linkages, and underlying glycans (51, 52) using sodium acetate buffer (pH 4.0) or sodium phosphate buffer (pH 7.4). However, no significant binding was observed under the tested conditions. No significant binding was detected at neutral pH using version 5.1 mammalian glycan arrays from the Consortium for Functional Glycomics (CFG) that contain 610 distinct, potential glycan receptors.

Following the structural homology of SRRP$_{53608}$-BR and SRRP$_{100-23}$-BR with PelC-like proteins, binding of SRRP$_{53608}$-BR was performed with polygalacturonic acid (PGA)-containing pectin fragments using a carbohydrate array of well-characterised plant polysaccharides and oligosaccharides produced by partial hydrolysis from polysaccharides (53, 54). Interestingly the binding was found to be pH-dependent with SRRP$_{53608}$-BR showing specific and reproducible binding to pectin structures at pH 4.0 but not at pH 7.4 (**SI Appendix, Table S4** and **Fig. S8**). In most cases, the glycans showing binding were lime pectin fractions with a low degree of methoxylation or esterification (DE ≤ 31%) of the galacturonic acid moieties present or PGA isolated from citrus pectin, whereas pectin fractions with higher DE values gave little or no binding. Such pH-dependent interaction was confirmed by bio-layer interferometry using the Octet system where biotinylated-SRRP$_{53608}$-BR was immobilised on streptavidin-coated optical biosensors and probed with rhamnogalacturonan I (RGI), pectin esterified from citrus fruit (PECF), or PGA as ligand. A sensorgram showing the association and dissociation of SRRP$_{53608}$-BR binding to PGA or RGI at pH 4.0 is shown in **SI Appendix, Fig. S9**. The kinetic parameters for the interaction of SRRP$_{53608}$-BR and PGA were determined through global fitting of raw data using a 1:1 (Langmuir) binding model where $k_a = 3.03 \times 10^2$ M^{-1}s^{-1} ± 0.54% and $k_d = 7.17 \times 10^{-5}$ s^{-1} ± 0.69%, yielding a $K_D = 0.237 \times 10^{-6}$ M with an R^2 of 0.997077 and full $X^2 = 0.23$ (**Table 2**), the latter two values confirming the fitting of the model (**Fig. S9a**). While binding was observed with RGI at pH 4.0, it was not possible to fit an acceptable model to the data, though it is apparent that the interaction dissociates rapidly upon removal of the ligand (**Fig. S9b**). Binding was not observed for two commercially available pectins, PECF and pectin P7536, both obtained from Sigma (**Fig. S9c**). Additionally, no binding could be detected when the experiments were performed at pH 7.4. Preliminary screening assays using pH 7.4, 7.0, 6.5, 6.0, 5.5, 5.0, 4.5, and 4.0 demonstrated the pH-dependent increase in binding to PGA at lower pH (pH 5.0 and below). Initial assays omitting EDTA or Tween 20 in the run buffers resulted in non-specific binding. As shown in Fig. 3, the superposition of the SRRP$_{53608}$-BR$_{262-571}$ and TM-Pel-TGA structures allowed the identification of possible binding residues in SRRP$_{53608}$-BR$_{262-571}$. In order to further investigate the specificity of the interaction, a series of SRRP$_{53608}$-BR mutants were generated by site-directed mutagenesis (see Methods, SI Appendix and **Fig. 3f**) and tested against PGA. These included two alanine-substituted single mutants, K377A and R512A , and one where the lower loop from F411 to T422 was deleted, named ΔF411-T422. K377 and R512 were selected due to their localization at the extremities of the proposed PuBS. Furthermore, MD simulations also indicated a possible role of R512 in ligand binding. The ΔF411-T422 mutant was created to evaluate the importance of the flexible loop in ligand binding. All mutants showed similar circular dichroism spectra at pH 7.4, suggesting correct folding of the recombinant proteins (**SI Appendix, Fig. S10**). Additionally, no differences were observed between wild-type SRRP$_{53608}$-BR at pH 7.4 and pH 4.0 (**Fig. S10**). The mutations led to reduced K$_D$ values, showing similar rates of association but increased rates of dissociation, in comparison
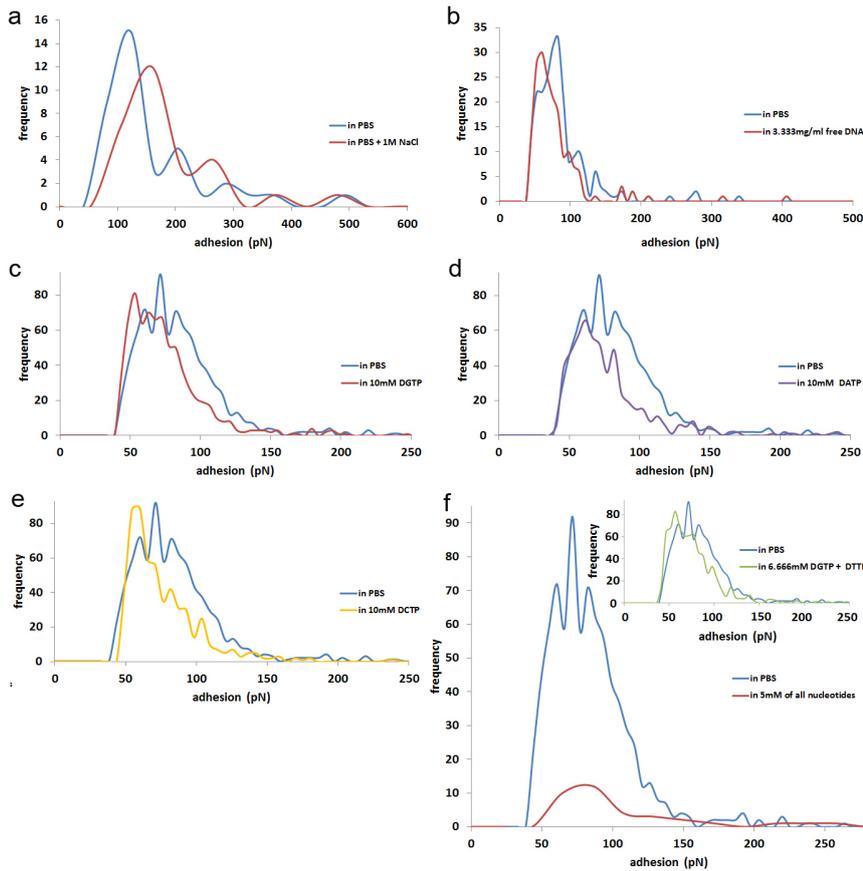
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884

885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952

Fig. 7. **AFM force spectroscopy histograms of SRRP$_{53608}$-BR interacting with DNA.** Quantification of adhesion values of SRRP$_{53608}$-BR interacting with DNA (**a**) in 137 mM NaCl or 1 M NaCl PBS or after addition of (**b**) 'free' DNA, free nucleotides (**c**) dGTP, (**d**) dATP, (**e**) dCTP or (**f**) all four nucleotides (dGTP, dATP, dTTP and dCTP) or two nucleotides (dGTP and dTTP) as shown in the inset.

to the wild-type protein (**Table 2**). Chondroitin sulfate A (from bovine trachea) was also tested against immobilised SRRP$_{53608}$-BR under the above conditions. Similarly, no binding was observed at pH 7.4. However, at pH 4.0, concentration-dependent binding could be observed with chondroitin sulfate A (**Fig. S9d**). Binding analysis of chondroitin sulfate A (as above) determined a $k_a = 9.2 \times 10^1$ M^{-1}s^{-1} ± 1.26% and $k_d = 8.72 \times 10^{-5}$ s^{-1} ± 3.04%, yielding a $K_D = 9.47 \times 10^{-7}$ M with an R^2 = 0.9935 and X^2 = 0.21.

LrSRRP-BR promotes *L. reuteri* adhesion to the intestinal epithelium

To determine the contribution of *Lr*SRRP-BR adherence to the host tissue following the reported ability of *L. reuteri* 100-23 (and not the *L. reuteri* 100-23 *srrp* mutant) to form biofilm *in vivo* (6), we performed adhesion assays to tissue sections of the mouse epithelium using soluble recombinant SRRP$_{53608}$-BR at pH 7.4. Interestingly we could detect binding of SRRP$_{53608}$-BR to both stomach and colonic epithelium. In both types of tissue, the staining patterns correlated with WGA (wheat germ agglutinin) binding. No staining was observed in negative controls (SRRP$_{53608}$-BR free) (**Fig. 5a and b**). Binding of SRRP$_{53608}$-BR to colonic tissue sections was significantly reduced following treatment with periodate at pH 4.0 (**Fig. 5c**), suggesting binding to glycoproteins. In the healthy stomach of mice and humans, the MUC5AC and MUC6 mucins dominate and are produced by the surface epithelium and glands, in line with the mucins produced by human HT-29-MTX cells. Direct binding of *Lr*SRRP-BR to the HT-29-MTX cells was carried out at pH 7.4. The *Lr*SRRP-BR staining pattern correlated with WGA lectin and to some degree also with MUC5AC staining. No staining was observed in the negative control (SRRP$_{53608}$-BR free) (**Fig. 5d**). Furthermore, binding was also observed to HT-29 cells lines (**Fig. 5e**), suggesting

that at pH 7.4, SRRP$_{53608}$-BR can recognise mucins and/or other epithelial receptors.

In order to identify the nature of the ligands, we investigated the binding of SRRP$_{53608}$-BR to individual components of the epithelium and mucus layer including purified mucins and DNA by atomic force microscopy (AFM). As compared to other techniques used to measure the force magnitude of ligand-receptor interactions, AFM provides specific information on the distance of interactions between molecules i. e. the distance to the functionalised tip as it moves along the immobilised ligand and retracts from the surface, as shown in the inset example force curves in Fig. 6. **Fig. 6a** shows quantification of the magnitude of adhesion captured in the retraction curves of the force spectroscopy measurements between SRRP$_{53608}$-BR and mucin. In neutral buffer (pH 7.4) the modal value of adhesion events was 105 pN, whereas bimodal values of adhesion events, at 72 pN and 120 pN, were obtained in acidic buffer (pH 4.0). Addition of PGA appeared to abolish SRRP$_{53608}$-BR mucin interactions at pH 4.0 as PGA only specifically binds to SRRP$_{53608}$ in acidic conditions. Furthermore, despite the adhesion magnitudes being similar in both buffers, the SRRP$_{53608}$-BR and mucin interactions in the acidic buffer produced a significantly larger range of the length of adhesion events compared to the data obtained in neutral conditions (**Fig. 6b**). This suggests that at acidic pH, SRRP$_{53608}$-BR has a better ability to interact along the entire length of the mucin chains compared to neutral pH. **Fig. 6a** and **6b** show a set of the typical force-distance curves from each of the experiments to reveal the distance variations between pH 7.4 and 4.0. In **Fig. 6c**, direct binding of SRRP$_{53608}$-BR to PGA was further confirmed by AFM at pH 4.0. There was a significantly large range of adhesion magnitudes from 33 – 428 pN. The modal value was 36 pN at pH 4.0 and 18 pN in the neutral buffer which corresponds to

the noise level in the force spectra. Adding free PGA to the AFM liquid cell led to a significant reduction in adhesion events due to competitive interactions between free PGA and the glass-attached PGA molecules, as shown in the example force curve, therefore confirming the specificity of the interaction.

Force spectroscopy was also used to investigate binding of $SRRP_{53608}$-BR to DNA (**Fig. 7**). Increasing NaCl molarity from 137 mM to 1 M in the PBS buffer caused a minor reduction of the adhesion frequencies (15 to 12) of $SRRP_{53608}$-BR to DNA but increased the modal values of the adhesion events (123 to 160 pN) (**Fig. 7a**), suggesting that binding was unlikely to be solely due to electrostatic interactions. In order to further assess the specificity of the interaction between $SRRP_{53608}$-BR and DNA, free DNA sample was added into the AFM liquid cell, resulting in a minor reduction of adhesion event frequencies (from 33 to 30) and of modal values (from 83 to 60 pN) (**Fig. 7b**). Adding single nucleotides separately (**Fig. 7c,d,e**) or in tandem (**Fig. 7f inset**) did not cause any major inhibition (only minor reductions of frequencies and modal values of the adhesions were observed). In contrast, the simultaneous addition of all four nucleotides led to a significant inhibition in terms of the adhesion event frequencies (92 to 12) with similar modal values of adhesion (71 and 86 pN respectively) (**Fig. 7f**), which is expected from single ligand-receptor interaction events as measured in the timescale of AFM experiments. These results suggest that all four nucleotides are recognised by $SRRP_{53608}$-BR.

Discussion

Structural differences between SRRPs from lactobacilli and pathogenic streptococci and staphylococci reflect differences in ligand specificity

Bacterial attachment to host surfaces is a pivotal event in the biological and infectious processes of both commensal and pathogenic bacteria, respectively. SRRPs and their associated secretion systems are being defined in a growing number of Gram-positive bacteria, indicating their crucial roles in mediating interaction with the host.

Atomic resolution structures of seven SRRP binding regions have been reported for Gram-positive pathogens to date, highlighting a relationship between their structural folds and binding ligands. This includes *Strep. parasanguinis* Fap1 (2KUB and 2X12) (55), *Strep. gordonii* GspB (3QC5/6) (13)*, Strep. Sanguinis* SrpA (5EQ2) (56), Srr-1 and Srr-2 paralogues of *Strep. agalactiae* (4MBO/R) (57, 58), *Staph. aureus* SraP (4M0(0-3)) (59) and *Strep. pneumoniae* PsrP (3ZGH/I) (60). Srr-1, Srr-2 and PsrP each adopt variations of the Dev-IgG fold (57, 60, 61) and bind to long β-strands in their target proteins, thereby forming a complementary strand along one of the Ig-like domains of the Dev-Ig protein: with Srr-1 binding to cytokeratin-4 (62-64); Srr-1 and Srr-2 binding to fibrinogen Aα (57, 65-68), and PsrP adhering to cytokeratin-10 (69). PsrP also binds to sialylated structures (13, 60, 70, 71) and to DNA (60, 72). Other SRRP-BR regions are composed of two or more sub-domains and include – from N to C-terminal – the helical and CnaA folds for Fap1 (55), the CnaA, siglec and 'unique' subdomains for GspB (13), the siglec and 'unique' subdomains for SrpA (56), and a legume lectin like fold, a β-grasp fold and two eukaryotic cadherin-like modules for SraP (59). The GspB, Hsa, SrpA and SraP SRRP-BRs have all been shown to bind to different types of sialylated ligands (13, 73-80), whereas the binding ligand of Fap1 remains to be identified (55). *Lr*SRRP-BR did not recognise sialylated glycans but was found to bind to host epithelium or pectin-like components in a pH-dependent manner. This difference in ligand specificity can be explained by the *Lr*SRRP-BR right-handed β-solenoid topology, which is typically adopted by extracellular, enzymatic PelC-like proteins. It is also the first time such a fold has been observed in SRRP-BRs and in *L. reuteri* species. Additonally, the *Lr*SRRP-BR

structural data indicate a high representation of Trp, Tyr and basic residues in the PuBS, suggesting an involvement in carbohydrate binding, which may correlate with *Lr*SRRP-BRs' recognition of mucin glycoproteins and plant-derived anionic polysaccharides. In addition, mutagenesis confirmed the importance of the binding loop and residues R512 and K377 within the PuBS in the interaction of $SRRP_{53608}$-BR with PGA.

*Lr*SRRP-BR binds to polyanionic ligands via a pH-dependent binding mode

Three major carbohydrate structures are found in pectin, which include homogalacturonan (HGA), rhamnogalacturonan I (RGI) and rhamnogalacturonan II (RGII). Strong pH dependency was observed when $SRRP_{53608}$-BR binding was tested against chondroitin sulfate A and selected pectin ligands. $SRRP_{53608}$-BR bound PGA and RGI at pH 4.0 but no binding was observed at pH 7.4 or against the two other commercial pectins tested at either pH. A similar preference for acidic pH was observed for binding to chondroitin sulfate A. Long MD simulations (at microsecond timescale) at pH 4.0 and pH 7.4 showed that $SRRP_{53608}$-$BR_{262-571}$ undergoes a pH-dependent conformational change close to the PuBS, such that at pH 4.0 a greater region of the PuBS is solvent-exposed. Coupled with a rearrangement of postulated key sidechains (D487, R512, I514 and Y482) at low pH, this suggests a mechanism for the observed differential binding to anionic PGA, RGI and chondroitin sulfate A polysaccharides (**SI Appendix, Fig. S11**) at pH 4.0 and pH 7.4. Furthermore, the models suggest a notable difference in surface charge distribution in the PuBS at the two pH values, exhibiting a more positive potential at pH 4.0, which would certainly have a marked impact on the ability of $SRRP_{53608}$-$BR_{262-571}$ to bind anionic substrates. These results also suggest that other linear polyanionic glycosaminoglycan (GAG) polysaccharides may be relevant biological ligand for SRRP in the gut. Such binding pH-dependent conformational change has previously been reported for the pectin lyase, PelA (33) and also for the Strep. parasanguinis SRRP adhesin, Fap1, where the low pH-driven conformational change modulates adhesion and likely plays a role in survival in acidic environments (82).

L. reuteri and many other *Lactobacillus* spp. are primary colonizers of the proximal GI tract and therefore exposed to acidic stress in the stomach. The pH values in the rodent forestomach region range between 3.8 to 5.1 depending on feeding (83). Similarly, in the porcine stomach, the pH is relatively low at the oesophageal terminus of the stomach and higher towards the pylorus (84). In addition to a longitudinal pH gradient along the GI tract, there is a pH gradient across mucus, as demonstrated in rodents *in vivo*, supporting a role for this barrier in gastric mucosal protection (85, 86). Our findings that *Lr*SRRP-BR binds to dietary components at low pH and to the mucosal epithelium at higher pH are in line with the observed pH gradient from the lumen to the epithelium surface. It is also worth noting that in the stomach, the *Helicobacter pylori* sialic acid-binding adhesin (SabA) shows a charge/low pH-dependent mechanism likely to play different roles during colonisation of the oral to gastric niches and during long-term infection (87).

Lessons on niche specificity and biofilm formation

SRRPs comprise a large family of adhesins in Gram-positive bacteria (55) which are exported by an accessory Sec system (the SecA2-SecY2 system) (11) and important for biofilm formation (74,87,88). Here we showed that in *L. reuteri* strains, genes encoding homologues of SRRPs are generally co-localised within the SecA2 gene cluster with only a few exceptions of unlinked SRRP genes or pseudogenes. While the overall domain organisation of SRRPs is conserved between Gram-positive bacteria species, the individual SRRP-BR domains are highly diverse with limited to no sequence homology (90). SRRPs have been characterized from numerous streptococci and staphylococci inhabiting differ-

ent niches and contributing to pathogenesis due to their role in host cell adhesion and biofilm formation. Fap1 from Strep. parasanguinis (82, 91, 92) as well as the sialoglycan binding SRR adhesins Hsa and GspB from *Strep. gordonii* strains CH1 and M99, respectively are involved in dental plaque formation and periodontal disease via attachement to salivary components. They are also a virulence factor for infective endocarditis, initiated by their binding to sialoglycans on human platelets. *Strep. sanguinis and Staph. aureus* are other pathogens causing bacterial endocarditis in which their SRR adhesins, SrpA and SraP, respectively, also mediate binding to sialylated receptors on human platelets (73, 93). Srr1 and Srr2 adhesins from the Strep. agalactiae pathogen, causing neonatal meningitis, bind to human fibrinogen and keratin 4. Keratin 4 binding mediates colonisation of the female genital tract, leading to neonatal infection and fibrinogen binding mediates adhesion to human brain microvascular endothelial (57, 65). PsrP from Strep. pneumoniae, causing streptococcal pneumonia, facilitates biofilms formation on lung epithelial cells via self-oligomerisation, DNA binding and adhesion to keratin 10, both facilitated by its BR (60, 69, 88). It is worth noting that SRRPs also occur in commensal streptococci based on sequence analysis and that SRR glycoproteins from *Strep. salivarius* were recently shown to play a major role in host colonization although no structural information is available (14).

Here, we showed that *Lr*SRRP-BR from *L. reuteri* binds to the epithelium and dietary components in a pH-dependent mechanism, which may favour persistence in the GI tract. In addition to its role in adhesion to polysaccharides and glycoproteins, *Lr*SRRP-BR could bind DNA in a specific manner as shown by AFM. DNA from autolysed bacterial cells is a component of many biofilms, helping to form an extracellular network to which live cells can attach. This may therefore contribute to the ability of *L. reuteri* rodent strains (and perhaps other *Lactobacillus* spps. harbouring SRRPs identified by bioinformatics analysis) to form biofilm on the murine forestomach *in vivo* (6). While pathogenic biofilms contribute to states of chronic inflammation, biofilm formation by probiotic bacteria, such as *Lactobacillus* spp. cause a negligible immune response and is considered a beneficial property by promoting colonisation and longer permanence in the host mucosa, and limiting colonization by pathogenic bacteria. Understanding at the molecular level the contribution of lacto-bacillus SRRPs in biofilm formation is needed to fully exploit the functions of this intra- and inter-species family of adhesins across Gram-positive bacteria. These molecular findings may help the rational selection of probiotic strains of lactobacilli that can compete with the SRRP-mediated adhesion of pathogenic streptococci and staphylococci.

Crystallization of SRRP$_{53608}$-BR and SRRP$_{100-23}$-BR. Cloning, purification, crystallization and structure determination were as described in *SI Appendix*.

Molecular Dynamics (MD) simulations. Protein protonation states at pH 4.0 and 7.4 were predicted by Schrodinger's Maestro software suite (PROPKA module), and long molecular dynamics simulations were run using AMBER 14. Details of the simulations are described in *SI Appendix*.

Binding assays. Binding of SRRP$_{53608}$-BR wild-type and mutants to PGA, RGI, pectins, and chondroitin sulfate A was performed by Bio-Layer Interferometry; binding of SRRP$_{53608}$-BR to mucins, DNA and PGA was asessed by force spectroscopy using AFM; immunofluorsecence was used to monitor SRRP$_{53608}$-BR binding to mouse intestinal tissue sections as described in *SI Appendix*.

Full details of all experimental procedures used are described in *SI Appendix*.

1. Lynch SV & Pedersen O (2016) The human intestinal microbiome in health and disease. *New Engl. J. Med.* 375(24):2369-2379.
2. Walter J, Britton RA, & Roos S (2011) Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. *Proc. Natl. Acad. Sci. U. S. A.* 108:4645-4652.
3. Walter J (2008) Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Appl. Environ. Microbiol.* 74(16):4985-4996.
4. Oh PL, *et al.* (2010) Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution. *Isme J.* 4(3):377-387.
5. Frese SA, *et al.* (2011) The evolution of host specialization in the vertebrate gut symbiont *Lactobacillus reuteri*. *Plos Genetics* 7(2).
6. Frese SA, *et al.* (2013) Molecular characterization of host-specific biofilm formation in a vertebrate gut symbiont. *Plos Genetics* 9(12).
7. Duar RM, *et al.* (2017) Experimental Evaluation of Host Adaptation of Lactobacillus reuteri to Different Vertebrate Species. *Appl. Environ. Microbiol.* 83(12).
8. Wegmann U, *et al.* (2015) The pan-genome of *Lactobacillus reuteri* strains originating from the pig gastrointestinal tract. *Bmc Genomics* 16.
9. Lizcano A, Sanchez CJ, & Orihuela CJ (2012) A role for glycosylated serine-rich repeat proteins in gram-positive bacterial pathogenesis. *Mol Oral Microbiol* 27(4):257-269.
10. Feltcher ME & Braunstein M (2012) Emerging themes in SecA2-mediated protein export. *Nat. Rev. Microbiol.* 10(11):779-789.
11. Prabudiansyah I & Driessen AJM (2017) The Canonical and Accessory Sec System of Gram-positive Bacteria. *Curr. Top. Microbiol. Immunol.* 404:45-67.
12. Handley PS, Correia FF, Russell K, Rosan B, & DiRienzo JM (2005) Association of a novel high molecular weight, serine-rich protein (SrpA) with fibril-mediated adhesion of the oral biofilm bacterium *Streptococcus cristatus*. *Oral Microbiol. Immunol.* 20(3):131-140.
13. Pyburn TM, *et al.* (2011) A structural model for binding of the serine-rich repeat adhesin GspB to host carbohydrate receptors. *PLoS Pathog* 7(7):e1002112.
14. Couvigny B, *et al.* (2017) Three glycosylated serine-rich repeat proteins play a pivotal role in adhesion and colonization of the pioneer commensal bacterium, *Streptococcus salivarius*. *Environmental microbiology*.
15. Tytgat HL & de Vos WM (2016) Sugar coating the envelope: Glycoconjugates for microbe-host crosstalk. *Trends Microbiol.* 24(11):853-861.
16. Roos S, Karner F, Axelsson L, & Jonsson H (2000) *Lactobacillus mucosae* sp *nov*., a new species with in vitro mucus-binding activity isolated from pig intestine. *Int. J. Syst. Evol. Microbiol.* 50:251-258.
17. Su MS, Oh PL, Walter J, & Ganzle MG (2012) Intestinal origin of sourdough *Lactobacillus reuteri* isolates as revealed by phylogenetic, genetic, and physiological analysis. *Appl. Environ. Microbiol.* 78(18):6777-6780.
18. 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.
19. Lee JH, *et al.* (2012) Genome sequence of Lactobacillus mucosae LM1, isolated from piglet feces. *J. Bacteriol.* 194(17):4766.
20. Holm L & Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38:W545-W549.
21. Sillitoe I, *et al.* (2015) CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res.* 43(D1):D376-D381.
22. Jenkins J & Pickersgill R (2001) The architecture of parallel beta-helices and related folds. *Prog Biophys Mol Biol* 77(2):111-175.
23. Kobe B & Kajava AV (2000) When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends Biochem. Sci.* 25(10):509-515.
24. Yoder MD, Keen NT, & Jurnak F (1993) New domain motif - the structure of pectate lyase-c, a secreted plant virulence factor. *Science* 260(5113):1503-1507.
25. Akita M, Suzuki A, Kobayashi T, Ito S, & Yamane T (2001) The first structure of pectate lyase belonging to polysaccharide lyase family 3. *Acta Crystallogr D* 57:1786-1792.
26. Pickersgill R, Jenkins J, Harris G, Nasser W, & Robert-Baudouy J (1994) The structure of *Bacillus subtilis* pectate lyase in complex with calcium. *Nat. Struct. Biol.* 1(10):717-723.
27. Seyedarabi A, *et al.* (2010) Structural insights into substrate specificity and the anti beta-elimination mechanism of pectate lyase. *Biochemistry* 49(3):539-546.
28. Zheng YY, *et al.* (2012) Crystal structure and substrate-binding mode of a novel pectate lyase from alkaliphilic *Bacillus sp* N16-5. *Biochem. Biophys. Res. Commun.* 420(2):269-274.
29. Creze C, *et al.* (2008) The crystal structure of pectate lyase peli from soft rot pathogen Erwinia

30. Jenkins J, Shevchik VE, Hugouvieux-Cotte-Pattat N, & Pickersgill RW (2004) The crystal structure of pectate lyase Pel9A from *Erwinia chrysanthemi*. *J. Biol. Chem.* 279(10):9139-9145.

31. Scavetta RD, *et al.* (1999) Structure of a plant cell wall fragment complexed to pectate lyase C. *Plant Cell* 11(6):1081-1092.

32. McDonough MA, *et al.* (2002) Crystallization and preliminary X-ray characterization of a thermostable pectate lyase from *Thermotoga maritima*. *Acta Crystallogr D* 58:709-711.

33. Mayans O, *et al.* (1997) Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases. *Structure* 5(5):677-689.

34. Vitali J, Schick B, Kester HC, Visser J, & Jurnak F (1998) The tree-dimensional structure of aspergillus niger pectin lyase B at 1.7-A resolution. *Plant Physiol.* 116(1):69-80.

35. Petersen TN, Kauppinen S, & Larsen S (1997) The crystal structure of rhamnogalacturonase A from *Aspergillus aculeatus*: a right-handed parallel beta helix. *Structure* 5(4):533-544.

36. Mohnen D (2008) Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* 11(3):266-277.

37. Yoder MD, Lietzke SE, & Jurnak F (1993) Unusual structural features in the parallel beta-helix in pectate lyases. *Structure* 1(4):241-251.

38. Abbott DW & Boraston AB (2007) The structural basis for exopolygalacturonase activity in a family 28 glycoside hydrolase. *J. Mol. Biol.* 368(5):1215-1222.

39. 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.

40. 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.

41. Emsley P, Charles IG, Fairweather NF, & Isaacs NW (1996) Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* 381(6577):90-92.

42. Jung WS, *et al.* (2007) Structural and functional insights into intramolecular fructosyl transfer by inulin fructotransferase. *J. Biol. Chem.* 282(11):8414-8423.

43. 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.

44. Miller JL, *et al.* (2008) Selective ligand recognition by a diversity-generating retroelement variable protein. *PLoS Biol.* 6(6):1195-1207.

45. 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.

46. 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.

47. 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.

48. Steinbacher S, *et al.* (1996) Crystal structure of phage P22 tailspike protein complexed with *Salmonella sp* O-antigen receptors. *Proc. Natl. Acad. Sci. U. S. A.* 93(20):10584-10588.

49. Clantin B, *et al.* (2004) The crystal structure of filamentous hemagglutinin secretion domain and its implications for the two-partner secretion pathway. *Proc Natl Acad Sci U S A* 101(16):6194-6199.

50. Locht C, Bertin P, Menozzi FD, & Renauld G (1993) The filamentous haemagglutinin, a multifaceted adhesion produced by virulent *Bordetella spp*. *Mol. Microbiol.* 9(4):653-660.

51. Deng LQ, Chen X, & Varki A (2013) Exploration of sialic acid diversity and biology using sialoglycan microarrays. *Biopolymers* 99(10):650-665.

52. Padler-Karavani V, *et al.* (2012) Cross-comparison of protein recognition of sialic acid diversity on two novel sialoglycan microarrays. *J. Biol. Chem.* 287(27):22593-22608.

53. Pedersen HL, *et al.* (2012) Versatile high resolution oligosaccharide microarrays for plant glycobiology and cell wall research. *J. Biol. Chem.* 287(47):39429-39438.

54. Sorensen I, Pedersen HL, & Willats WG (2009) An array of possibilities for pectin. *Carbohydr Res* 344(14):1872-1878.

55. Ramboarina S, *et al.* (2010) Structural insights into serine-rich fimbriae from Gram-positive bacteria. *J. Biol. Chem.* 285(42):32446-32457.

56. Bensing BA, *et al.* (2016) Structural basis for sialoglycan binding by the *Streptococcus sanguinis* SrpA adhesin. *J. Biol. Chem.* 291(14):7230-7240.

57. Seo HS, *et al.* (2013) Characterization of fibrinogen binding by glycoproteins Srr1 and Srr2 of Streptococcus agalactiae. *J. Biol. Chem.* 288(50):35982-35996.

58. Sundaresan R, Samen U, & Ponnuraj K (2011) Expression, purification, crystallization and preliminary X-ray diffraction studies of the human keratin 4-binding domain of serine-rich repeat protein 1 from *Streptococcus agalactiae*. *Acta Crystallogr F* 67:1582-1585.

59. Yang YH, *et al.* (2014) Structural insights into SraP-mediated *Staphylococcus aureus* adhesion to host cells. *PLoS Path.* 10(6).

60. Schulte T, *et al.* (2014) The basic keratin 10-binding domain of the virulence-associated pneumococcal serine-rich protein PsrP adopts a novel MSCRAMM fold. *Open Biol* 4(10).

61. Sundaresan R, Samen U, & Ponnuraj K (2015) Structure of KRT4 binding domain of Srr-1 from *Streptococcus agalactiae* reveals a novel beta-sheet complementation. *Int J Biol Macromol* 75:97-105.

62. Samen U, Reinscheid DJ, Reinscheid DJ, & Borges F (2007) The surface protein srr-1 of *Streptococcus agalactiae* binds human keratin 4 and promotes adherence to epithelial HEp-2 cells. *Infect. Immun.* 75(11):5405-5414.

63. Sheen TR, *et al.* (2011) Serine-rich repeat proteins and pili promote *Streptococcus agalactiae* colonization of the vaginal tract. *J. Bacteriol.* 193(24):6834-6842.

64. Wang NY, *et al.* (2014) Group B streptococcal serine-rich repeat proteins promote interaction with fibrinogen and vaginal colonization. *J. Infect. Dis.* 210(6):982-991.

65. Seo HS, Mu R, Kim BJ, Doran KS, & Sullam PM (2012) Binding of glycoprotein srr1 of *Streptococcus agalactiae* to fibrinogen promotes attachment to brain endothelium and the development of meningitis. *PLoS Path.* 8(10).

66. Seo HS, Xiong YQ, & Sullam PM (2013) Role of the serine-rich surface glycoprotein Srr1 of *Streptococcus agalactiae* in the pathogenesis of infective endocarditis. *Plos One* 8(5).

67. Six A, *et al.* (2015) Srr2, a multifaceted adhesin expressed by ST-17 hypervirulent Group B Streptococcus involved in binding to both fibrinogen and plasminogen. *Mol. Microbiol.* 97(6):1209-1222.

68. van Sorge NM, *et al.* (2009) The group B streptococcal serine-rich repeat 1 glycoprotein mediates penetration of the blood-brain barrier. *J. Infect. Dis.* 199(10):1479-1487.

69. Shivshankar P, Sanchez C, Rose LF, & Orihuela CJ (2009) The *Streptococcus pneumoniae* adhesin PsrP binds to Keratin 10 on lung cells. *Mol. Microbiol.* 73(4):663-679.

70. Bensing BA, *et al.* (2016) Novel aspects of sialoglycan recognition by the Siglec-like domains of streptococcal SRR glycoproteins. *Glycobiology* 26(11):1222-1234.

71. Muller YA (2013) Unexpected features in the Protein Data Bank entries 3qd1 and 4i8e: the structural description of the binding of the serine-rich repeat adhesin GspB to host cell carbohydrate receptor is not a solved issue. *Acta Crystallogr F* 69:1071-1076.

72. Sanchez CJ, *et al.* (2010) The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation in vivo and in biofilms. *PLoS Path.* 6(8).

73. Deng L, *et al.* (2014) Oral streptococci utilize a Siglec-like domain of serine-rich repeat adhesins to preferentially target platelet sialoglycans in human blood. *PLoS Pathog* 10(12):e1004540.

74. Kim AR, *et al.* (2016) Serine-rich repeat adhesin gordonii surface protein B is important for *Streptococcus gordonii* biofilm formation. *J Endodont* 42(12):1767-1772.

75. Oguchi R, *et al.* (2016) Contribution of *Streptococcus gordonii* Hsa adhesin to biofilm formation. *Jpn J Infect Dis.*

76. Takahashi Y, *et al.* (2006) Contribution of sialic acid-binding adhesin to pathogenesis of experimental endocarditis caused by *Streptococcus gordonii* DL1. *Infect. Immun.* 74(1):740-743.

77. Takamatsu D, *et al.* (2005) Binding of the Streptococcus gordonii surface glycoproteins GspB and Hsa to specific carbohydrate structures on platelet membrane glycoprotein Ibalpha. *Mol. Microbiol.* 58(2):380-392.

78. Takamatsu D, Bensing BA, Prakobphol A, Fisher SJ, & Sullam PM (2006) Binding of the streptococcal surface glycoproteins GspB and Hsa to human salivary proteins. *Infect. Immun.* 74(3):1933-1940.

79. Urano-Tashiro Y, Takahashi Y, Oguchi R, & Konishi K (2016) Two arginine residues of *Streptococcus gordonii* sialic acid-binding adhesin Hsa are essential for interaction to host cell receptors. *Plos One* 11(4).

80. Yajima A, *et al.* (2008) Hsa, an adhesin of *Streptococcus gordonii* DL1, binds to alpha 2-3-linked sialic acid on glycophorin A of the erythrocyte membrane. *Microbiol. Immunol.* 52(2):69-77.

81. Loukachevitch LV, *et al.* (2016) Structures of the *Streptococcus sanguinis* SrpA binding region with human sialoglycans suggest features of the physiological ligand. *Biochemistry* 55(42):5927-5937.

82. Garnett JA, *et al.* (2012) Structural insight into the role of *Streptococcus parasanguinis* Fap1 within oral biofilm formation. *Biochem. Biophys. Res. Commun.* 417(1):421-426.

83. Ward FW & Coates ME (1987) Gastrointestinal pH measurement in rats - influence of the microbial-flora, diet and fasting. *Lab Anim* 21(3):216-222.

84. Jubb K, Kennedy P, & Palmer N (2013) *Pathology of Domestic Animals 4th Edition* (Academic Press).

85. McNeil NI, Ling KLE, & Wager J (1987) Mucosal surface pH of the large-intestine of the rat and of normal and inflamed large-intestine in man. *Gut* 28(6):707-713.

86. Schreiber S, *et al.* (2004) The spatial orientation of *Helicobacter pylori* in the gastric mucus. *Proc. Natl. Acad. Sci. U. S. A.* 101(14):5024-5029.

87. Linden SK, Wickstrom C, Lindell G, Gilshenan K, & Carlstedt I (2008) Four modes of adhesion are used during *Helicobacter pylori* binding to human mucins in the oral and gastric niches. *Helicobacter* 13(2):81-93.

88. Schulte T, *et al.* (2016) The BR domain of PsrP interacts with extracellular DNA to promote bacterial aggregation; structural insights into pneumococcal biofilm formation. *Sci Rep-Uk* 6.

89. Bandara M, *et al.* (2017) The accessory Sec system (SecY2A2) in *Streptococcus pneumoniae* is involved in export of pneumolysin toxin, adhesion and biofilm formation. *Microbes Infect* 19(7-8):402-412.

90. Zhou MX & Wu H (2009) Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. *Microbiology-Sgm* 155:317-327.

91. Chen Q, Sun B, Wu H, Peng Z, & Fives-Taylor PM (2007) Differential roles of individual domains in selection of secretion route of a *Streptococcus parasanguinis* serine-rich adhesin, Fap1. *J. Bacteriol.* 189(21):7610-7617.

92. Wu H, Zeng M, & Fives-Taylor P (2007) The glycan moieties and the N-terminal polypeptide backbone of a fimbria-associated adhesin, Fap1, play distinct roles in the biofilm development of *Streptococcus parasanguinis*. *Infect. Immun.* 75(5):2181-2188.

93. Siboo IR, Chambers HF, & Sullam PM (2005) Role of SraP, a serine-rich surface protein of *Staphylococcus aureus*, in binding to human platelets. *Infect. Immun.* 73(4):2273-2280.

chrysanthemi in complex with its substrate. *J. Biol. Chem.* 283(26):18260-18268.