

ARTICLE

DOI: 10.1038/s41467-018-05471-3

OPEN

# Structural basis of cell wall anchoring by SLH domains in *Paenibacillus alvei*

Ryan J. Blackler<sup>1,4</sup>, Arturo López-Guzmán<sup>2</sup>, Fiona F. Hager<sup>2</sup>, Bettina Janesch<sup>2</sup>, Gudrun Martinz<sup>3</sup>, Susannah M.L. Gagnon<sup>1</sup>, Omid Haji-Ghassemi<sup>1,5</sup>, Paul Kosma<sup>3</sup>, Paul Messner<sup>2</sup>, Christina Schäffer<sup>2</sup> & Stephen V. Evans<sup>1</sup>

Self-assembling protein surface (S-) layers are common cell envelope structures of prokaryotes and have critical roles from structural maintenance to virulence. S-layers of Gram-positive bacteria are often attached through the interaction of S-layer homology (SLH) domain trimers with peptidoglycan-linked secondary cell wall polymers (SCWPs). Here we present an in-depth characterization of this interaction, with co-crystal structures of the three consecutive SLH domains from the *Paenibacillus alvei* S-layer protein SpaA with defined SCWP ligands. The most highly conserved SLH domain residue SLH-Gly29 is shown to enable a peptide backbone flip essential for SCWP binding in both biophysical and cellular experiments. Furthermore, we find that a significant domain movement mediates binding by two different sites in the SLH domain trimer, which may allow anchoring readjustment to relieve S-layer strain caused by cell growth and division.

<sup>1</sup>Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC V8W 3P6, Canada. <sup>2</sup>Department of NanoBiotechnology, NanoGlycobiology Unit, Universität für Bodenkultur Wien, 1190 Vienna, Austria. <sup>3</sup>Department of Chemistry, Universität für Bodenkultur Wien, 1190 Vienna, Austria. <sup>4</sup>Present address: Zymeworks Inc., Vancouver, BC V6H 3V9, Canada. <sup>5</sup>Present address: Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC V6T 1Z3, Canada. These authors contributed equally: Ryan J. Blackler, Arturo López-Guzmán. Correspondence and requests for materials should be addressed to C.Säf. (email: [christina.schaeffer@boku.ac.at](mailto:christina.schaeffer@boku.ac.at)) or to S.V.E. (email: [svevans@uvic.ca](mailto:svevans@uvic.ca))

Prokaryotic cell envelopes often include proteinaceous surface (S-) layers composed of self-assembling subunits in porous two-dimensional crystalline lattices<sup>1–7</sup>. S-layers can have a variety of critical functions including structural maintenance, environmental protection, adhesion, filtering, and virulence<sup>8–14</sup>. These characteristics make S-layers attractive targets for new antibiotics and provide biotechnological potential for the display of functional epitopes with precise geometries in vaccine design, drug delivery, molecular electronics, and filtration<sup>15–23</sup>. Understanding their three-dimensional structure is key to exploiting the therapeutic or biotechnological potential of S-layers.

The maintenance of S-layers in biological contexts depends on stable and flexible attachment to constantly changing cell surfaces. S-layers have evolved covalent and non-covalent anchor mechanisms to a variety of cell wall components, including the archaeal plasma membrane<sup>22,24,25</sup>, the peptidoglycan (PG) layer of Gram-positive bacteria<sup>9,22,26,27</sup>, and the lipopolysaccharide of Gram-negative bacteria<sup>28–30</sup>.

S-layer anchoring is best characterized for Gram-positive bacteria, where S-layer proteins non-covalently attach to the PG layer via species- and strain-specific non-classical secondary cell wall polymers (SCWPs)<sup>26</sup>. Non-classical SCWPs are neutral or anionic polysaccharides not classifiable as teichoic or teichuronic acids (classical SCWPs), and are covalently linked to muramic acid residues of PG, at least in some cases through phosphodiester bonds. SCWP structures have been reviewed in detail elsewhere<sup>8</sup>.

The most common, but not exclusive, means of SCWP binding is through S-layer homology (SLH) domains, which are often present in triplicate at the termini of S-layer and other extracellular proteins<sup>4,22,30–33</sup>. Within the conserved protein domain family SLH (pfam00395)<sup>34</sup>, there are 14,079 sequences identified as SLH domains across 651 bacterial species, widely in the phyla *Firmicutes* (378 species), *Cyanobacteria* (100 species), *Proteobacteria* (41 species), and *Actinobacteria* (33 species), among others, suggesting the early evolution of the SLH–SCWP interaction for anchoring proteins to the surfaces of bacteria<sup>35</sup>.

The interaction of SLH domains with several SCWPs depends on a common ketal–pyruvate modification of SCWPs that imparts a negative charge<sup>35–37</sup>. SLH domains contain a highly conserved TRAE motif (SLH domain residues 42–45, Fig. 1) where the arginine residue has been shown to be critical to PG binding<sup>27,31,38</sup>, suggesting that it interacts with the negatively charged pyruvate moiety of these SCWPs.

The crystal structure of the SLH domain trimer from the *B. anthracis* S-layer protein Sap revealed the three domains arranged in approximate threefold symmetry, where each SLH domain contributes one helix to a parallel three-helix bundle core and a

second helix rotated ~90° from the core to form three lobes<sup>39</sup>. The conserved TRAE motifs are located at the N-termini of the core helices adjacent to three grooves, suggesting that these grooves are the sites for SCWP binding. The structure of *B. anthracis* SCWP was elucidated subsequently<sup>40</sup>, and a crystal structure of the SLH domains of Sap in complex with a synthetic trisaccharide approximating the terminal repeat of *B. anthracis* SCWP was also recently reported<sup>41</sup>. This structure revealed ligand bound in only one of the three grooves and confirmed that SLH-Arg43 of the conserved TRAE motif interacts with the pyruvate moiety of SCWP. However, questions remain as to the functionality of the remaining two grooves and the roles of additional conserved residues.

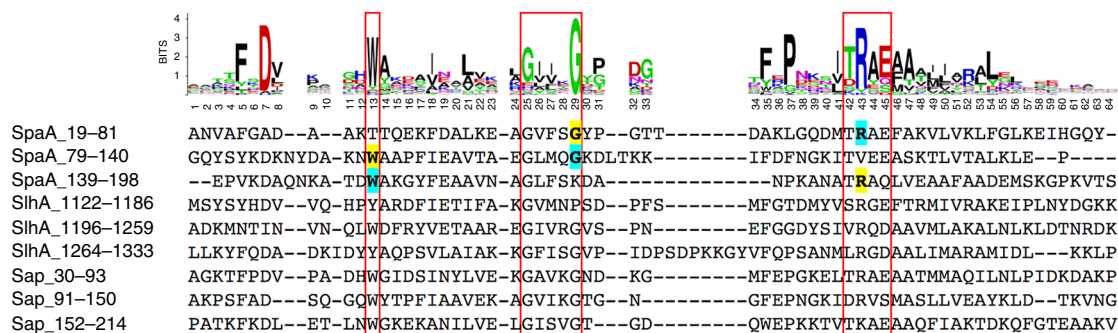
To investigate the molecular details of SCWP binding by SLH domains in depth, we selected the *Paenibacillus alvei* CCM 2015<sup>T</sup> S-layer protein SpaA for structural studies. SpaA possesses three consecutive SLH domains near its N-terminus, followed by a large C-terminal region presumed to harbor the self-assembling domain(s)<sup>5,27,42</sup>. The SCWP from *P. alvei* CCM 2051<sup>T</sup> has been fully defined (Fig. 2)<sup>8,37,43,44</sup>, which enables an investigation of the precise interactions between SLH domains and SCWP.

SpaA possesses the conserved TRAE motif in its first SLH domain (SLH1) and variants of this motif, TVEE in its second (SLH2), and TRAQ in its third (SLH3) (Fig. 1). These motifs were shown to contribute unequally to SCWP binding, where their mutation to TAAA resulted in 37%, 88%, and 50% of wild-type (*wt*) cell wall binding in pelleting assays, respectively<sup>27,45</sup>. These variations of the conserved TRAE motif make SpaA an excellent model to investigate its role in SCWP binding. Similar inequality within SLH domain repeats has been observed in S-layer proteins from other organisms<sup>46,47</sup>, but there is no obvious relationship between primary sequence and SCWP-binding stoichiometry or affinity and a structural foundation is necessary to shed light on evolved differences in SCWP-binding mechanisms.

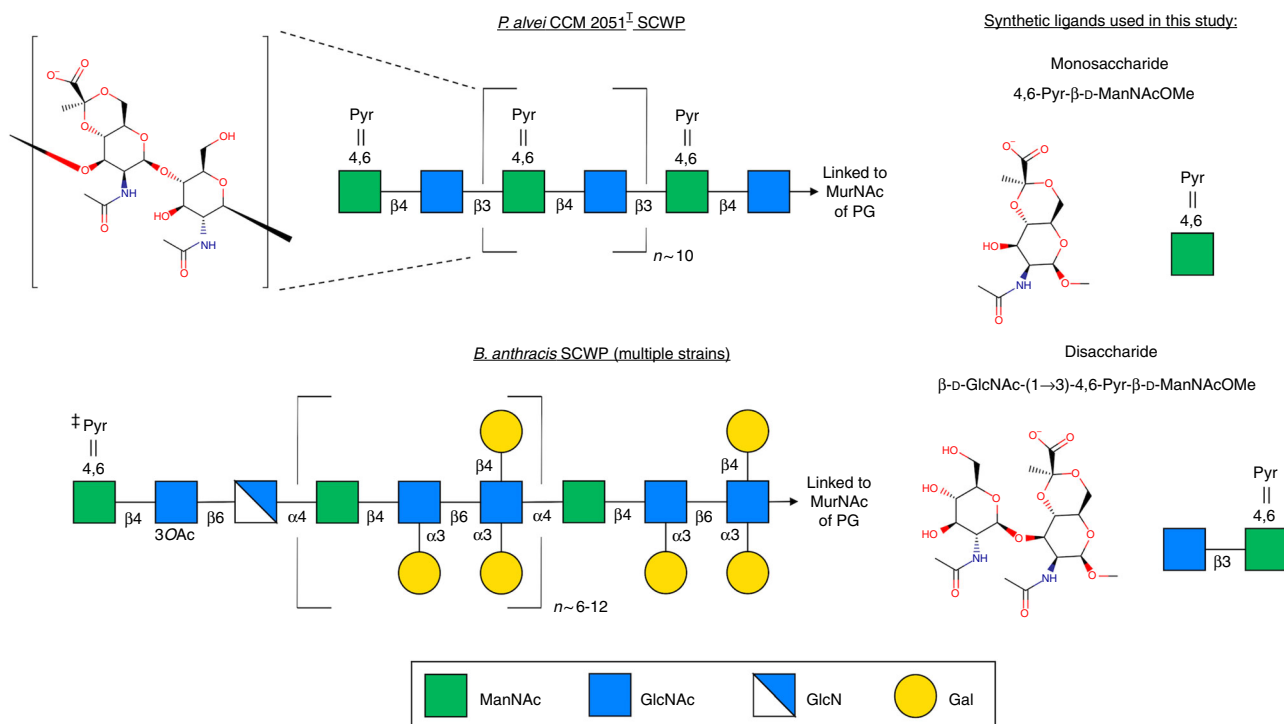
Here we report crystal structures of the SLH domains of SpaA, unliganded and in complex with synthesized monosaccharide and disaccharide building blocks of *P. alvei* CCM 2051<sup>T</sup> SCWP. These structures reveal the contributions of many conserved residues to SCWP binding, notably including the most highly conserved residue SLH-Gly29, which is confirmed by additional co-crystal structures, thermodynamic binding analysis, and microbiological functional assays for SLH-Gly29Ala mutants. Together, these data provide novel molecular insights into this biologically important interaction.

## Results

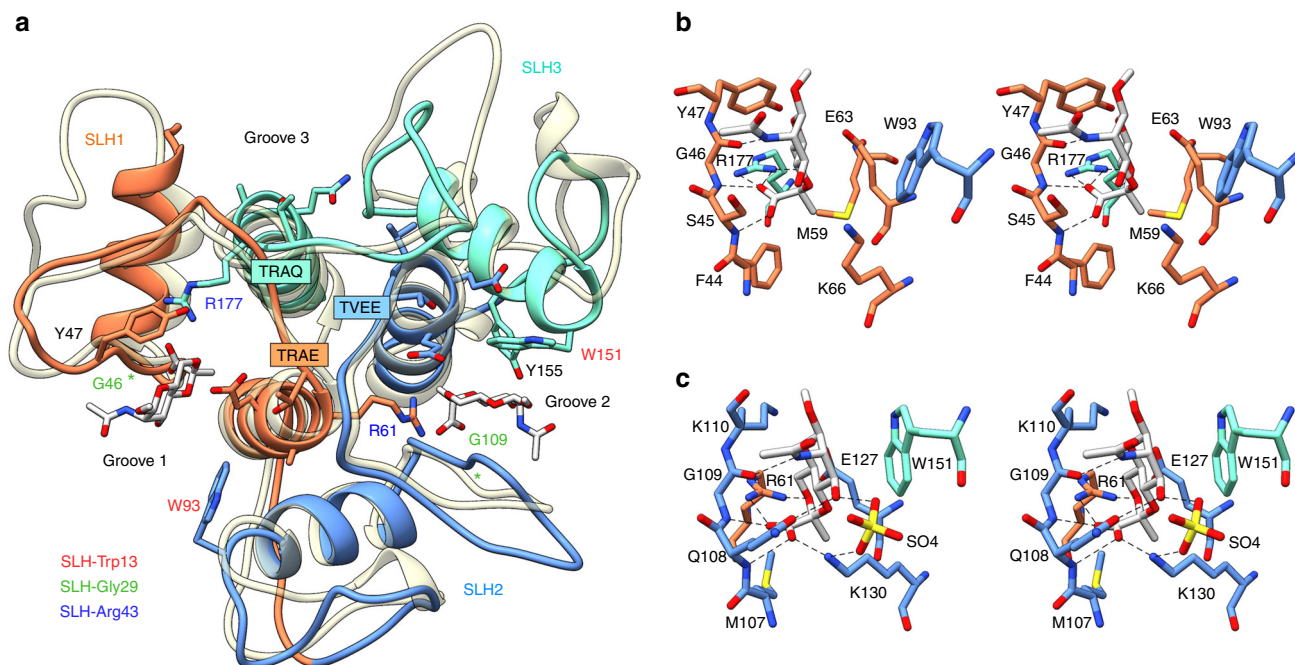
**SpaA<sub>SLH</sub> structure and SCWP binding.** The crystal structure of SpaA<sub>SLH</sub> reveals an overall fold similar to the three-pronged spindle



**Fig. 1** SLH domain sequence alignments. Sequence logo of the SLH domain profile from PROSITE #P551272, with the sequences of the SLH domains from *P. alvei* SpaA, SlhA, and *B. anthracis* Sap aligned underneath. Throughout the manuscript, references to conserved residues of the SLH domain profile are preceded by “SLH-” to differentiate from SpaA<sub>SLH</sub> amino acid labels. The conserved SLH-Trp13, GIIxG motif, and TRAE motif are boxed in red. The conserved SLH-Trp13, SLH-Gly29, and SLH-Arg43 that contribute to ligand binding in SpaA grooves 1 and 2 are highlighted in yellow and cyan, respectively



**Fig. 2** SCWP structures and synthetic ligands used in this study. Schematic diagrams of the SCWP from *P. alvei* CCM 2051<sup>T</sup> and *B. anthracis* (multiple strains) are shown on the left, with a detailed chemical diagram of the repeating unit shown for the former<sup>8, 37, 40, 43, 44, 61</sup>. The synthetic ligands used in this study, representing building blocks of *P. alvei* CCM 2051<sup>T</sup> SCWP, are shown on the right. ‡Pyruvylation, O-acetylation, and N-deacetylation of the terminal trisaccharide repeat were determined for *B. anthracis* Ba684<sup>40</sup>. There remains some uncertainty in the extent of pyruvylation of some *B. anthracis* SCWPs, because the use of HF to cleave SCWPs from PG also liberates the acid-sensitive ketal groups<sup>8</sup>



**Fig. 3** Spa<sub>SLH</sub> structure and SCWP ligand binding. **a** Ribbon diagram of Spa<sub>SLH</sub> with bound monosaccharide (PDB 6CWI; Supplementary Table 1) and *B. anthracis* Sap (PDB 3PYW) overlaid in tan. Spa<sub>SLH</sub> is colored with SLH1 orange, SLH2 blue, and SLH3 aquamarine. Residue labels are colored by conserved SLH domain residue (legend bottom left). Stereo views of monosaccharide binding in G1 and G2 are shown in **b** and **c**

described for the SLH domains of *B. anthracis* Sap<sup>39</sup> (Fig. 3a, Supplementary Fig. 1, and Supplementary Table 1). The TRAE, TVEE, and TRAQ motifs of SpaA are located at the N-termini of the core helices adjacent to three grooves (labeled here G1, G2, and

G3). The TRAE and TRAQ motifs each contribute to two of the three grooves, as their arginine residues thread under the connecting loops of the adjacent SLH domains and into the neighboring grooves. In the case of SLH2, Val125 of the TVEE motif

**Table 1 ITC analyses of 4,6-Pyr- $\beta$ -D-ManNAcOMe binding**

	$-T\Delta S$ (kJ/mol)	$\Delta H$ (kJ/mol)	$\Delta G$ (kJ/mol)	Stoichiometry	$K_A$ ( $M^{-1}$ )	$K_D$ (nM)
SpaA <sub>SLH</sub>	45.74 $\pm$ 16.01	-87.85 $\pm$ 15.62	-42.10 $\pm$ 0.65	0.91 $\pm$ 0.04	3.48 $\times 10^7$ $\pm$ 0.36	29
SpaA <sub>SLH</sub> /G109A	13.33 $\pm$ 7.0	-50.64 $\pm$ 7.29	-37.31 $\pm$ 0.33	0.92 $\pm$ 0.04	4.48 $\times 10^6$ $\pm$ 0.62	226
SpaA <sub>SLH</sub> /G46A/ G109A	No binding					

corresponds in position to the conserved SLH-Arg43 of the TRAE and TRAQ motifs but does not protrude into the neighboring G3. The final residue of each motif, Glu63, Glu127, or Gln179, lines the groove beside its parent helix.

To investigate the molecular details of SCWP binding by SLH domains, we synthesized the monosaccharide 4,6-Pyr- $\beta$ -D-ManNAcOMe (see Supplementary methods and Supplementary Figs. 2–4), which represents the pyruvylated moiety of *P. alvei* SCWP (Fig. 2), and co-crystallized it with SpaA<sub>SLH</sub>. The first co-crystal structure determined reveals ligand bound in G2 only, in a narrow pocket formed by conserved residues from all three SLH domains: Arg61 (corresponding to the conserved SLH-Arg43 of the TRAE motif, Fig. 1) from SLH1; Met107, Gln108, Gly109 (corresponding to the conserved SLH-Gly29), Lys110, Glu127, and Lys130 from SLH2; and Trp151 (corresponding to the conserved SLH-Trp13) from SLH3 (Figs. 1 and 3). The pyruvate moiety of the ligand is bound deep in the pocket through salt bridge interactions to Arg61 and Lys130 and hydrogen bonds to Gln108 and Gly109 backbone amides, while the hydrophobic face of the ManNAc ring stacks against Trp151.

In the initial liganded and unliganded structures, multiple conformations were observed for residues 44–54 that form the connecting loop between the two helices of SLH1 beside G1 (Supplementary Fig. 5). The loop makes unique crystal contacts in each case and has mean B-factors elevated between 4 and 49% of the molecule mean. We investigated the effect of this flexibility on potential ligand binding in G1 by solving additional co-crystal structures in different space groups (P1 and C2, Supplementary Table 1), where further unique conformations of residues 44–54 were observed. Ligand binding in G2 is the same in all cases except for the C2 structure (PDB 6CWI), which shows a sulfate ion coordinated to Lys130 and ligand in G2. Only the C2 structure displays electron density corresponding to a partially occupied site in the G1 pocket of molecule A (out of two molecules A and B in the asymmetric unit [AU]), with G1 of molecule B occluded by molecule A (Supplementary Fig. 6). Nevertheless, this weakly bound 4,6-Pyr- $\beta$ -D-ManNAcOMe in G1 displays similar interactions as the ligand bound in G2 (Fig. 3). The pyruvate moiety is coordinated by a salt bridge from Arg177 and forms hydrogen bonds to Ser45 N and Gly46 N, which are equivalent to residues Arg61, Gln108, and Gly109 in G2. In this case, no stacking interactions are made from Trp93 (equivalent to Trp151 in G2).

To further investigate the activity of these two potential binding sites, we performed isothermal titration calorimetry (ITC) analysis of SpaA<sub>SLH</sub> binding to 4,6-Pyr- $\beta$ -D-ManNAcOMe in solution, which revealed 1:1 binding with an apparent dissociation constant ( $K_D$ ) of 29 nM (Table 1 and Supplementary Fig. 7). This 1:1 binding in solution likely occurs in G2, because this site is occupied in all five crystallographically unique molecules and has better complementarity with ligand than does G1, which was only occupied in one of five molecules. Therefore, the monosaccharide binding observed in G1 is likely to be an artifact of crystallization that does not occur in solution.

**SpaA binds a terminal non-reducing end epitope of SCWP.** To further refine the identity of the SCWP epitope bound by

SpaA<sub>SLH</sub>, we synthesized the disaccharide  $\beta$ -D-GlcNAc-(1  $\rightarrow$  3)-4,6-Pyr- $\beta$ -D-ManNAcOMe for co-crystallization and binding analysis (Fig. 2). This disaccharide represents an internal repeat of *P. alvei* SCWP as opposed to the terminal non-reducing end disaccharide Pyr- $\beta$ -D-ManNAcOMe-(1  $\rightarrow$  4)- $\beta$ -D-GlcNAc, with the former being more tractable to synthesis. The internal disaccharide showed no binding to SpaA<sub>SLH</sub> by ITC (Supplementary Fig. 7), which suggests that SpaA<sub>SLH</sub> is specific for an epitope including the terminal non-reducing end 4,6-Pyr- $\beta$ -D-ManNAc residue of SCWP rather than an internal epitope.

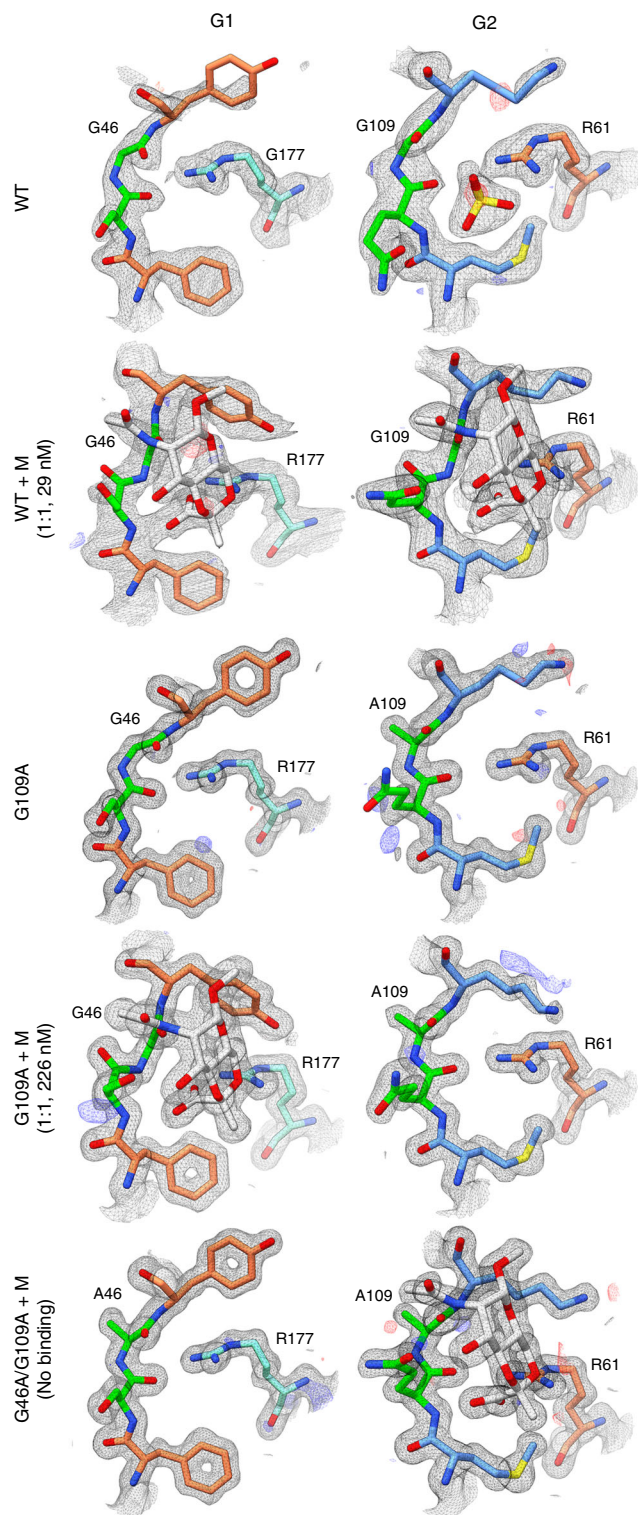
Despite the lack of binding measured by ITC, a co-crystal structure with the internal disaccharide was obtained, again indicating that binding is enhanced significantly in the crystalline environment compared to in solution. In both molecules in the AU of this structure, the 4,6-Pyr- $\beta$ -D-ManNAcOMe moiety of the disaccharide is bound in the G2 pocket in a manner similar to the monosaccharide ligand (Supplementary Fig. 8). The two molecules display fragmented electron density for the GlcNAc moiety in two different conformations, neither of which form hydrogen bonds to the protein (Supplementary Fig. 6). Furthermore, SpaA<sub>SLH</sub> residues 139–152 are disordered in one of two molecules in the AU, whereas they are ordered in all other structures of SpaA<sub>SLH</sub>. This region includes Trp151 (the conserved SLH-Trp13), which normally stacks against the 4,6-Pyr- $\beta$ -D-ManNAcOMe moiety. In the molecule where this loop is disordered, the GlcNAc moiety of the disaccharide occupies the expected location of Trp151, thus suggesting that this region becomes disordered to avoid a clash with the disaccharide ligand.

**SLH-Gly29 enables a backbone flip required for SCWP binding.** In addition to the TRAE motif, SLH domains contain a highly conserved GILxG motif (residues 25–29, Fig. 1), where the second glycine (SLH-Gly29) is the most conserved residue of the SLH domain profile<sup>31,48</sup>. SpaA possesses variants of this motif, with GVFSG, GLMQG, and GLFSK present in SLH1–SLH3. These motifs are located in loops between the two helices of each SLH domain that line one side of each groove (Figs. 1 and 3). The first G of each motif is within a turn exiting  $\alpha$ 1 and transitioning into this loop, and the hydrophobic residues VF, LM, and LF are oriented toward the hydrophobic core of each SLH domain lobe. In both G1 and G2, the non-conserved SLH-28 residue is observed to make a hydrogen bond from its backbone carbonyl to the side chain of the conserved SLH-Arg43 of the TRAE motif.

A remarkable series of structural transformations involving SLH-28 and SLH-Gly29 is observed with SCWP binding. Upon binding in G2 of *wt* SpaA<sub>SLH</sub>, the phi angle of Gly109 (SLH-Gly29) and the psi angle of Gln108 (SLH-28) each flip by  $\sim 180^\circ$  (Fig. 4 and Supplementary Table 2). The backbone carbonyl of Gln108, which hydrogen bonds to Arg61 (the conserved SLH-Arg43) when unliganded, is resultantly replaced in space by the Gly109 backbone amide nitrogen, which forms a hydrogen bond to a pyruvyl carboxyl of ligand. Concurrently, the removal of the hydrogen bond between Gln108 and Arg61 allows Arg61 to make a bidentate interaction with a pyruvyl carboxyl and O6 of ligand.

The hydrogen bond between the SLH-28 backbone carbonyl and SLH-Arg43 in the unliganded state may support the proper folding of the SLH domain trimer, as these residues precisely interlock





**Fig. 4** SLH-Gly29 backbone flip. 2Fo-Fc electron density maps contoured to  $1\sigma$  (black) and Fo-Fc maps contoured to  $3\sigma$  (blue) and  $-3\sigma$  (red) are shown for residues SLH-27 to SLH-30 and SLH-Arg43 in groove 1 (G1) and groove 2 (G2) of SpaA<sub>SLH</sub> structures. Residues are colored according to SLH domain as in Fig. 3, with SLH-28 and SLH-29 in green to highlight the backbone flip. WT wild-type SpaA<sub>SLH</sub>, M monosaccharide ligand, G109A SpaA<sub>SLH</sub>/G109A single mutant, G46A/G109A SpaA<sub>SLH</sub>/G46A/G109A double mutant. Monosaccharide ligand and corresponding electron density are shown for structures where density was observed, and binding stoichiometry and  $K_D$  as determined by ITC are given

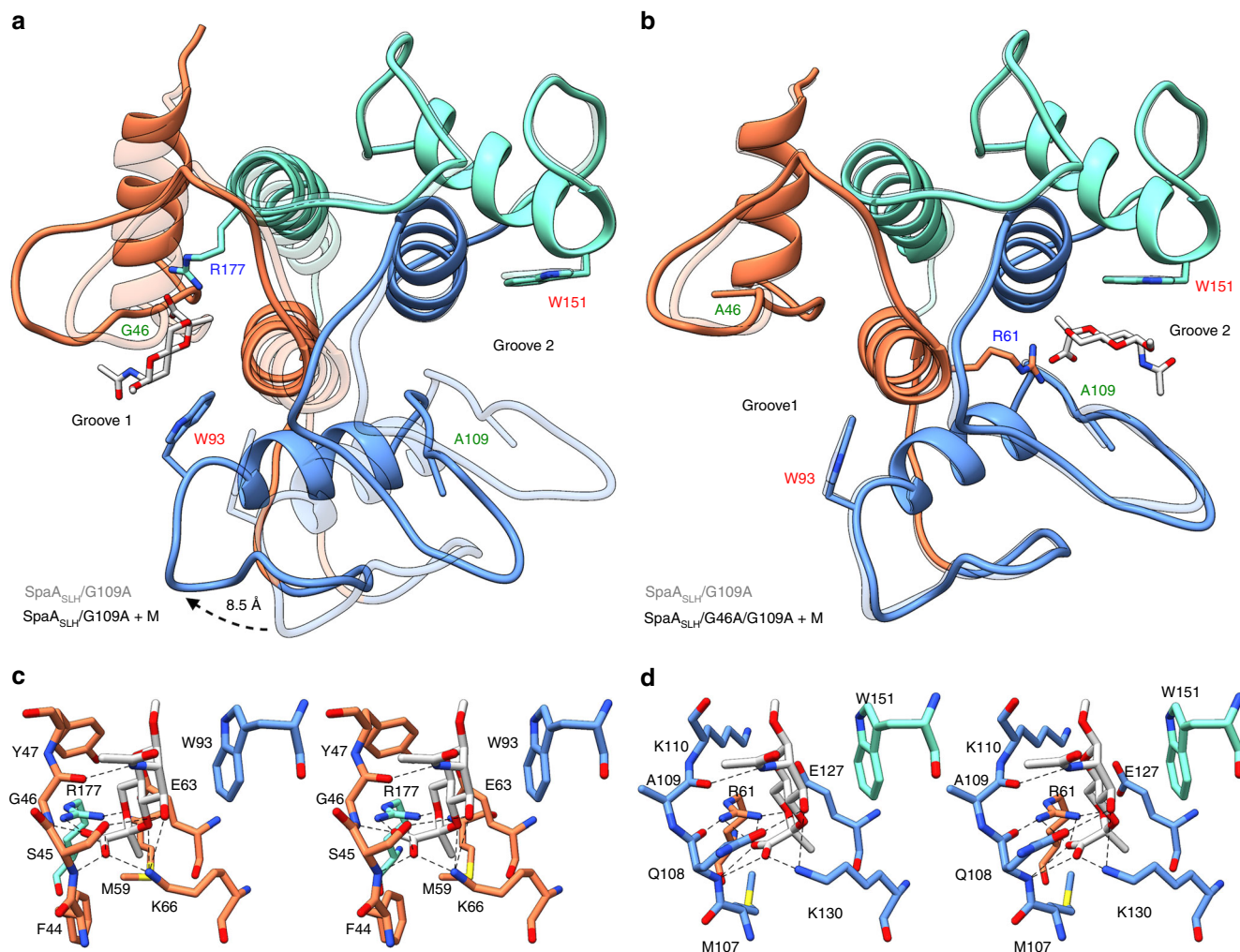
between neighboring SLH domains (Fig. 3). This is consistent with our finding that the TRAE motif is important for the overall folding of the SLH domain trimer, as attempts to express SpaA<sub>SLH</sub>/TAAA/TAAA double mutants resulted in insoluble inclusion bodies (see Supplementary Note 1). However, the presence of this stabilizing hydrogen bond between the SLH-28 backbone carbonyl and SLH-Arg43 would prevent SLH-Arg43 from making an ideal bidentate interaction with ligand, and ligand would be further deprived of the hydrogen bond from the backbone amide of SLH-Gly29. The backbone flip requires the flexibility of SLH-Gly29 in this location as the resulting phi-psi angles are disallowed for all other amino acids. Therefore, the near universal conservation of SLH-Gly29 suggests that this backbone flip is part of a conserved mechanism for SLH-SCWP binding.

To test our hypothesis that any other amino acid would prevent the backbone flip and impede binding, a SpaA<sub>SLH</sub>/G109A mutant was constructed for analyses. ITC revealed 1:1 binding to monosaccharide 4,6-Pyr- $\beta$ -D-ManNAcOMe with 226 nM affinity, nearly an order of magnitude lower than the 26 nM binding observed for *wt* SpaA<sub>SLH</sub> (Table 1 and Supplementary Fig. 7).

To confirm the structural basis for this reduced affinity, we determined the crystal structure of SpaA<sub>SLH</sub>/G109A alone and in complex with 4,6-Pyr- $\beta$ -D-ManNAcOMe. Unliganded SpaA<sub>SLH</sub>/G109A is similar in structure to *wt* SpaA<sub>SLH</sub> (Ca rmsd of 0.34 Å, excluding residues 44–54), and the Gly109Ala mutation is accommodated with no change in the position of Gln108 or Ala109 backbone atoms compared to Gln108 and Gly109 in the *wt* structure, but with a minor shift of residues 111–114. When SpaA<sub>SLH</sub>/G109A crystals were soaked with 4,6-Pyr- $\beta$ -D-ManNAcOMe, no electron density for ligand was observed in the resulting structures. We then co-crystallized SpaA<sub>SLH</sub>/G109A with ligand but attempts to solve the structure by molecular replacement using *wt* SpaA<sub>SLH</sub> structures were unsuccessful. Instead, we solved the structure using SAD phasing from a KI-soaked co-crystal.

The resulting structure of SpaA<sub>SLH</sub>/G109A in complex with 4,6-Pyr- $\beta$ -D-ManNAcOMe reveals a stunning structural change. While no electron density for ligand is observed in the G2 pocket, indicating that the Gly109Ala mutation inactivates binding as hypothesized, the protein is observed to alter significantly its conformation to utilize the G1 pocket instead, where excellent electron density for ligand is observed (Figs. 4 and 5, and Supplementary Fig. 6). The conformational change involves the movement of SLH2 away from SLH3 toward SLH1, resulting in a widening of G2 from 8.1 to 11.9 Å (measured between Trp151 C $\eta$ 2 and Ala109 O) and a narrowing of G1 from 8.9 to 7.6 Å (between Trp93 C $\eta$ 2 and Gly46 O). The overall Ca rmsd between liganded and unliganded SpaA<sub>SLH</sub>/G109A is 1.73 Å (1.52 Å when excluding residues 44–54), and when the SLH3 domain helices are aligned there is an 8.5 Å maximum displacement of SLH2 Ca atoms (measured for Lys91; Fig. 5). This significant structural change explains the failure to observe ligand soaked into crystals of unliganded SpaA<sub>SLH</sub>/G109A, because the structural change required for binding would not be possible in the crystal context, and also explains the initial failure to solve the complex structure by molecular replacement due to the significant difference between the search model and target structure. Ligand binding in G1 of SpaA<sub>SLH</sub>/G109A occurs with the SLH-28 and SLH-Gly29 backbone flip observed through residues Ser45 and Gly46 (Figs. 4 and 5, and Supplementary Table 2), again suggesting that the conservation of SLH-Gly29 is for its role in this binding mechanism.

To demonstrate conclusively the significance of the conserved SLH-Gly29 and the backbone flip for SCWP binding, we created a double mutant SpaA<sub>SLH</sub>/G46A/G109A for binding and structural analysis. As expected, this mutant displays no detectable binding to 4,6-Pyr- $\beta$ -D-ManNAcOMe by ITC (Table 1 and Supplementary Fig. 7). The crystal structure of SpaA<sub>SLH</sub>/G46A/G109A



**Fig. 5** SpaA<sub>SLH</sub>/G109A ligand binding and conformational change. Ribbon diagrams of **a** SpaA<sub>SLH</sub>/G109A with bound monosaccharide overlaid with unliganded SpaA<sub>SLH</sub>/G109A (transparent), and **b** SpaA<sub>SLH</sub>/G46A/G109A with bound monosaccharide overlaid with unliganded SpaA<sub>SLH</sub>/G109A (transparent). Alignments were performed using the helices of SLH3. Stereo views of monosaccharide binding in G1 of SpaA<sub>SLH</sub>/G109A and G2 of SpaA<sub>SLH</sub>/G46A/G109A are shown in **c** and **d**

displays the same groove conformations as *wt* SpaA<sub>SLH</sub> and unliganded SpaA<sub>SLH</sub>/G109A, with G2 narrower than G1 (Fig. 5b). Residues 46–55 in G1 are in a similar conformation as in the unliganded SpaA<sub>SLH</sub>/G109A structure, but with a slight widening of the loop to accommodate the Gly46Ala mutation. Despite the lack of binding detected by ITC, ligand is again observed bound in the G2 pocket, albeit with Gln108 and Ala109 in the unliganded backbone conformation (i.e., not flipped) (Fig. 5d). The monosaccharide is bound with the same interactions as observed for *wt* SpaA<sub>SLH</sub>, except that there is no hydrogen bond between the pyruvyl carboxyl oxygen and Gly109 nitrogen, but instead there is a minor movement of ligand away from the Gln108 carbonyl that results in a less ideal bidentate interaction with Arg61. The observation of bound ligand in the crystal structure despite no binding being detected by ITC, similar to the observation of 2:1 binding in the *wt* SpaA<sub>SLH</sub> crystal structure despite 1:1 binding measured by ITC, indicates again that affinity is markedly increased in the crystal context.

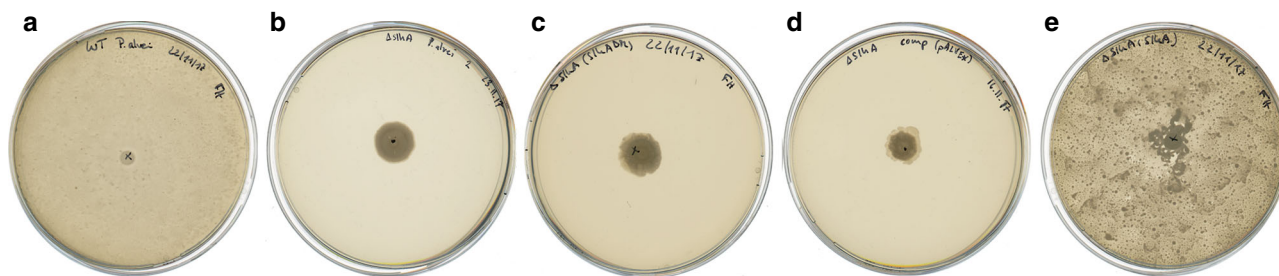
**SLH-Gly29 is required for cellular S-layer anchoring.** To corroborate the structural and biophysical evidence for the importance of the SLH-Gly29 backbone flip to SCWP binding, we investigated the phenotypic effects of SLH-Gly29 mutation in live

cells. Unfortunately, cellular experiments with SpaA are impossible because genetic manipulation of the *spaA* gene in *P. alvei* results in a lethal phenotype. Instead, we probed a different cell surface protein, SlhA, which also possesses three SLH domains with variations of the conserved TRAE and GILxG motifs (Fig. 1). It was shown previously that deletion of the *P. alvei slhA* gene produces changes of the colony morphology, impedes biofilm formation, and results in loss of swarming motility of *P. alvei* cells<sup>49</sup>. Swarming is a known mechanism for the migration of cells or cell clusters on semi-solid surfaces produced by the movement of flagella<sup>50</sup>.

SlhA possesses the conserved SLH-Gly29 of the SLH domain profile (Fig. 1) in domains SLH2 and SLH3, but not SLH1, and possesses the conserved SLH-Arg43 of the TRAE motif and SLH-Trp13 (or Tyr) in all three domains. Presuming the same arrangement of these conserved residues in three binding grooves as observed for SpaA and Sap (Fig. 3), it is likely that SlhA possesses two functional SCWP binding sites in G2 and G3. Therefore, we created a double mutant of SlhA with both Gly1224 and Gly1293 mutated to alanine (corresponding to SLH-Gly29 in SLH2 and SLH3) for cellular testing.

To assess the phenotypic effect of SLH-Gly29 mutation in SlhA, we used a *P. alvei ΔslhA* knockout strain, which shows loss





**Fig. 6** Swarming motility of *P. alvei* CCM 2051<sup>T</sup> variants. LB agar plates showing the growth of **a** *P. alvei* wt, **b** *P. alvei*  $\Delta$ slhA, **c** *P. alvei*  $\Delta$ slhA complemented with slhA/G1224A/G1293A, **d** *P. alvei*  $\Delta$ slhA complemented with empty plasmid pEXALV, and **e** *P. alvei*  $\Delta$ slhA complemented with slhA

of swarming motility on semi-solid agar plates compared to wt *P. alvei* (Fig. 6), and transformed it with plasmids encoding wt SlhA or the SlhA/G1224A/G1293A double mutant. Transformation with plasmid encoding SlhA/G1224A/G1293A resulted in a similar phenotype as transformation with an empty plasmid or as the knockout strain, indicating the loss of swarming motility. Transformation with plasmid encoding wt SlhA restored swarming motility in *P. alvei*  $\Delta$ slhA, as demonstrated previously by Janesch et al.<sup>49</sup> These results suggest that *P. alvei* loses the ability to swarm because SlhA/G1224A/G1293A is unable to attach to the cell surface as a consequence of the mutation of the conserved SLH-Gly29 residues to alanines, and thus confirm the biological importance of the SLH-Gly29 backbone flip to SCWP binding.

#### Structural basis for inequality within SpaA SLH domains.

Previous mutagenesis studies on full-length *P. alvei* CCM 2051<sup>T</sup> SpaA have shown that mutations of the TRAE, TVEE, and TRAQ motifs to TAAA have unequal impacts, where they result in 37%, 88%, and 50% of wt binding to cell wall sacculi, respectively<sup>27</sup>. This inequality of tandem SLH domains in cell wall binding can now be correlated to the complex structures reported here.

The TRAE motif of SLH1 contains Arg61, which forms direct contacts with the pyruvate moiety of 4,6-Pyr- $\beta$ -D-ManNAcOME in the primary binding pocket in G2 (Fig. 3). Glu63 contributes to the secondary binding pocket in G1 and interacts with Arg177 but not with bound ligand (Figs. 3 and 5). The TRAE/TAAA mutation therefore affects both the G1 and G2 sites, consistent with the lowest binding observed when these residues were mutated (37% of wt)<sup>27</sup>.

The TVEE motif of SLH2 lacks the conserved SLH-Arg43, and instead presents Val125 to the non-functional site in G3. Glu127 interacts with Arg61 in the primary binding pocket of G2, but, as observed for Glu63 of the TRAE motif, does not interact with bound ligand (Fig. 3). Given the minor decrease in SCWP binding upon mutation of TVEE to TAAA (88% of wt)<sup>27</sup>, a structural role of Glu127 in the G2 pocket is not critical to binding.

The TRAQ motif of SLH3 contains Arg177, which forms direct contacts with the pyruvate moiety of 4,6-Pyr- $\beta$ -D-ManNAcOME in the secondary binding site in G1 (Figs. 2 and 3). Gln179 of the TRAQ motif lies in the non-functional site in G3. The contribution of the TRAQ motif to only the secondary binding site in G1 explains the lesser impact of the TRAQ/TAAA mutation (50% of wt) than the TRAE/TAAA mutation (37% of wt)<sup>27</sup>.

**Comparison of SpaA<sub>SLH</sub> to other SCWP-binding domains.** The only other published structures of SLH domains are those from the *B. anthracis* S-layer protein Sap, which were published in the unliganded form by Kern et al. (PDB 3PYW)<sup>39</sup>, and in complex with a synthetic trisaccharide 4,6-Pyr- $\beta$ -D-ManNAc-(1  $\rightarrow$  4)- $\beta$ -D-

GlcNAc-(1  $\rightarrow$  6)- $\alpha$ -GlcN (approximating the terminal unit of *B. anthracis* SCWP) by Sychantha et al. during the revision of our manuscript (PDB 6BT4)<sup>41</sup>. Sychantha et al. report ligand binding only in G2 and hypothesize that binding in G1 and G3 are obstructed by crystal packing. The terminal 4,6-Pyr- $\beta$ -D-ManNAc of ligand is bound in G2 with O6 and a pyruvyl carboxyl coordinated by Arg72 (corresponding to SLH-Arg43) and with stacking interactions to Trp164 (corresponding to SLH-Trp13), as observed for 4,6-Pyr- $\beta$ -D-ManNAcOMe bound by SpaA<sub>SLH</sub>, while the base of the binding pocket about the pyruvyl methyl is formed by different hydrophobic residues (Supplementary Fig. 9). Ligand is rotated by  $\sim$ 20° away from SLH-Arg43 in Sap<sub>SLH</sub> compared to SpaA<sub>SLH</sub> so that Trp164 stacks against the 1  $\rightarrow$  4 glycosidic linkage.

Sychantha et al. report that the pyruvyl carboxyl groups of ligand form hydrogen bonds with both the backbone amide and carbonyl oxygen of Lys117 (corresponding to SLH-28). This backbone orientation suggests that the Gly118 (corresponding to SLH-Gly29) backbone amide does not hydrogen bond to ligand and that the backbone flip does not occur with binding in this groove. However, inspection of the associated electron density reveals that the Lys117 backbone carbonyl is modeled outside of 2Fo-Fc density, and there is Fo-Fc density that supports instead a flipped orientation of the peptide bond (Supplementary Fig. 9). When flipped, Gly118 adopts backbone torsions disallowed for all other amino acids, and has its amide positioned to hydrogen bond with the pyruvyl moiety of ligand. This supports our hypothesis that the role of the conserved SLH-Gly29 is to allow this backbone orientation for ligand binding. Unfortunately, the structure of unliganded Sap<sub>SLH</sub><sup>39</sup> displays ambiguous electron density for this region, so it cannot be concluded that SLH-Gly29 also allows a stabilizing interaction between the SLH-28 backbone carbonyl and SLH-Arg43 in the unliganded state in this case. However, G1 of the unliganded Sap<sub>SLH</sub> structure does display well-defined electron density for this loop, and the backbone carbonyl of SLH-28 hydrogen bonds with SLH-Arg43, and SLH-Gly29 is positioned to perform a backbone flip upon ligand binding.

Although binding was observed only in G2 of the Sychantha et al. structure, both G1 and G3 contain conserved SLH-Trp13, SLH-Gly29, and SLH-Arg43 (Lys in G1) residues that were found to indicate active binding sites in SpaA<sub>SLH</sub>. Significantly, both G1 and G3 of Sap<sub>SLH</sub> are in “open” conformations, as observed for G1 of SpaA<sub>SLH</sub>, compared to the “closed” conformations observed for the ligand-bound G2 of SpaA<sub>SLH</sub> or G1 of SpaA<sub>SLH</sub>/G109A (Supplementary Fig. 9). This indicates that if G1 or G3 of Sap<sub>SLH</sub> is capable of binding SCWP, it would require significant structural changes to narrow these grooves, as observed for SpaA<sub>SLH</sub>/G109A. Indeed, Sychantha et al. report that crystals would crack and dissolve when soaked with increasing ligand concentrations, which suggests that conformational changes are occurring in the protein to accommodate binding in alternate

(perhaps even preferred) sites. Given the complexity of SCWP binding by *P. alvei* SpaA<sub>SLH</sub> elucidated by our study, where binding in a secondary site in G1 was accomplished by a significant structural change after deactivating the primary binding site in G2 by SLH-Gly29Ala mutation, it is clear that further investigation is required into the potential activity and cooperativity of the multiple grooves in *B. anthracis* Sap<sub>SLH</sub>.

Besides the SLH domain, the only other known S-layer anchoring module in Gram-positive bacteria is the cell wall binding 2 (CWB2) domain, prevalent in *Clostridia*<sup>30</sup>. The recent crystal structures of the *Clostridium difficile* 630 cell wall proteins Cwp6 and Cwp8 (PDBs 5J72 and 5J6Q) revealed trimeric arrangements of three tandem CWB2 domains that, like SLH domains, possess a parallel three-helix bundle core, but otherwise bear little resemblance<sup>33</sup>. On the sides of the inner three-helix bundle of the CWB2 trimer, there are additional sheets and helices that occupy the spaces where the SLH domain trimer has three grooves and SCWP-binding sites. The binding site for the CWB2 ligand PS-II, an anionic teichoic acid-like polymer consisting of hexa-glycosyl phosphate repeats, was hypothesized to be a shallow V-shaped groove on the surface of the trimer perpendicular to the core helix bundle that outlines the second CWB2 domain. Docking studies placed one hexa-glycosyl phosphate repeat lying in either direction in this groove, with the terminal mannose-1-P phosphate moiety placed near conserved arginine residues<sup>33</sup>. However, it remains possible that the CWB2 domain instead recognizes primarily the terminal mannose-1-P in a manner analogous to the recognition of the terminal pyruvylated residue of SCWP by SLH domains as seen for SpaA<sub>SLH</sub> and Sap<sub>SLH</sub>. Together, the trimeric arrangement of CWB2 and SLH domains and the use of conserved arginine residues to recognize anionic moieties of diverse cell wall polysaccharides may support a common evolutionary origin of these cell wall anchoring modules in Gram-positive bacteria.

## Discussion

The crystal structures and binding analyses presented here of *P. alvei* SpaA<sub>SLH</sub>, SpaA<sub>SLH</sub>/G109A, and SpaA<sub>SLH</sub>/G46A/G109A with synthetic SCWP fragments provide novel insights into the functional contributions of many conserved SLH domain residues, including SLH-Trp13, SLH-Gly29, and SLH-Arg43, which precisely interlock from three different SLH domains to generate two active binding sites. These SLH domains appear to be specific for the non-reducing-end Pyr- $\beta$ -D-ManNAc moiety of *P. alvei* SCWP, as opposed to internal occurrences of Pyr- $\beta$ -D-ManNAc with  $\beta$ -D-GlcNAc-(1  $\rightarrow$  3) linkages. This specificity is consistent with that of *B. anthracis* Sap<sub>SLH</sub> for its SCWP that is pyruvylated only on the terminal ManNAc<sup>40,41</sup>, which may indicate a general trend in the recognition of terminal anionic moieties of SCWPs by SLH domains in other organisms. However, this raises further questions as to why *P. alvei* SCWP is pyruvylated on every repeating unit, or why some SCWP of other organisms are otherwise negatively charged on every repeating unit, rather than on only the terminal moieties utilized for anchoring.

The most highly conserved SLH domain residue SLH-Gly29 is shown to allow a backbone flip of residues SLH-28 and SLH-29 that is critical for SCWP binding in both biophysical and cellular experiments. This backbone flip appears to have evolved as an elegant dual-purpose mechanism that promotes proper folding and stability of the SLH domain trimer in the unliganded state and readjusts to achieve optimal ligand coordination in the bound state. Given the high conservation of SLH-Gly29, we envision that this mechanism applies to the recognition of pyruvylated SCWP by SLH domains in other organisms, and perhaps even other non-pyruvylated anionic SCWPs. Indeed, this backbone flip is

supported by the electron density of the recently published structure of *B. anthracis* Sap<sub>SLH</sub> in complex with 4,6-Pyr- $\beta$ -D-ManNAc-(1  $\rightarrow$  4)- $\beta$ -D-GlcNAc-(1  $\rightarrow$  6)- $\alpha$ -GlcN (PDB 6BT4)<sup>41</sup>.

Our characterization of two functional SCWP-binding sites in SpaA<sub>SLH</sub> reveals a sequence pattern that may predict the functional-binding sites of other SLH domain trimers; a binding site in the groove beside domain SLH<sub>*n*</sub> is predicted by the presence of SLH-Gly29 in SLH<sub>*n*</sub> followed by SLH-Trp13 in SLH<sub>*n+1*</sub> and SLH-Arg43 in SLH<sub>*n+2*</sub> (Figs. 1 and 3), or circular permutations thereof. Indeed, it appears that many proteins containing SLH domain trimers identified in the Pfam family SLH (PF00395) contain either two or three copies of this sequence pattern. However, it should be noted that large structural changes mediating binding between sites may conceal the activity of some sites from analysis, as demonstrated here for *P. alvei* SpaA<sub>SLH</sub>.

When ligand binding in SpaA<sub>SLH</sub> G2 was disrupted by the SLH-Gly29Ala mutation, we observed a significant domain movement that restructured G1 to allow binding. The 1:1 binding to 4,6-Pyr- $\beta$ -D-ManNAcOMe with 29 nM  $K_D$  measured for *wt* SpaA<sub>SLH</sub> corresponds to a primary site in G2, while the 1:1 binding with 226 nM  $K_D$  measured for SpaA<sub>SLH</sub>/G109A corresponds to a secondary site in G1. SpaA<sub>SLH</sub> therefore appears to display a novel form of intramolecular negative cooperativity, where two mutually exclusive binding sites for the same ligand are modulated by a significant structural change. Although binding in G1 could only be detected when G2 was deactivated by the SLH-Gly29Ala mutation, the relationship of the two sites in a biological context may be more fluid. In the paracrystalline protein S-layer, there would be significant crowding between SLH domains of neighboring S-layer proteins and the interspersed SCWP strands, which may stabilize binding in a manner similar to the crystallization-induced binding of monosaccharide in G1 of SpaA<sub>SLH</sub> that was not detected in solution by ITC. Indeed, this crowding in the S-layer context may induce domain movement and the exchange of active binding sites, especially at locations of S-layer strain. Switchable binding between two sites may thus have biological significance in allowing relocation of binding among different SCWP strands to alleviate strain caused by S-layer restructuring during cell growth or division or could facilitate movement of SLH-anchored proteins on the cell surface. It remains possible that the switchable binding observed for SpaA<sub>SLH</sub> is an exception rather than the rule, and that the SLH domain trimer evolved instead for avid binding and has lost functionality in some cases according to the avidity required for certain SCWPs, or for the function of the particular protein being anchored. However, the crystal structure of *B. anthracis* Sap<sub>SLH</sub> reveals one “closed” and two “open” grooves, each possessing the conserved residues that predict active SCWP-binding sites<sup>39,41</sup>, which suggests that it may utilize such a switchable binding mechanism as well.

Given the broad biological importance of S-layers in the survival and virulence of diverse microorganisms, the mechanisms by which these proteins attach to cell surfaces are potential targets for the development of novel antibiotics. This is foreshadowed by the lethality of *spaA* inactivation in *P. alvei* and by our demonstration that swarming motility is impaired by disrupting SLH domain anchoring of the S-layer protein SlhA. Our in-depth characterization of the SpaA<sub>SLH</sub>-SCWP interaction in the model organism *P. alvei* should serve as a benchmark for future studies wishing to elucidate or perhaps inhibit such interactions in other organisms.

## Methods

**Expression and purification of recombinant SpaA<sub>SLH</sub>.** The SpaA<sub>SLH</sub> protein comprising amino acids 21–193, including a Ser-Gly-Ser linker preceding a C-terminal His<sub>6</sub>-tag, was generated by standard PCR amplification from *P. alvei* CCM



2051<sup>T</sup> genomic DNA. Primers for PCR amplification are listed in Supplementary Table 3. The PCR product and pET22b vector were digested with *Nde*I and *Sac*I and ligated. The resulting plasmid, pET22b-*spaAslh*, was transformed into *E. coli* BL21 (DE3-Star) cells (Invitrogen) and plated on LB agar plates containing 100 µg mL<sup>-1</sup> of ampicillin. Single colonies were transferred to 5 mL of LB medium supplied with the antibiotic and grown overnight at 37 °C. This culture was inoculated into four flasks containing 0.5 L of ampicillin-containing LB medium and incubated at 37 °C under shaking (200 rpm). Protein overexpression was induced at an OD<sub>600</sub> ~0.6 by the addition of IPTG at a final concentration of 0.6 mM. Cells were further incubated for 3 h, harvested by centrifugation at 6000 × *g* for 20 min and pellets were stored at -20 °C.

Cell pellets were resuspended in lysis buffer (25 mM Tris-HCl, pH 8.0, 200 mM NaCl, 5 mM imidazole) and disrupted by sonication. The lysate was clarified by centrifugation and applied to a nickel NTA-affinity chromatography column (Qiagen) equilibrated with lysis buffer. The column was washed with ten column volumes of buffer A (25 mM Tris-HCl, pH 8.0, 20 mM imidazole), followed by ten volumes of buffer B (25 mM Tris-HCl, pH 8.0, 50 mM imidazole). His<sub>6</sub>-tagged SpaA<sub>SLH</sub> was eluted with 25 mM Tris-HCl, pH 8.0, 250 mM imidazole (elution buffer). Further purification of SpaA<sub>SLH</sub> was performed by size exclusion chromatography on a Superdex 75 16/60 column equilibrated with 20 mM HEPES, pH 7.5, containing 100 mM NaCl. Fractions were analyzed by SDS-PAGE (10% gel) and proteins were detected by Coomassie brilliant blue (CBB).

**Site-directed mutagenesis.** Site-directed mutagenesis of residues glycine 109 and glycine 46 to alanine were performed by overlapping extension PCR to generate SpaA<sub>SLH</sub>/G109A (single mutant) and SpaA<sub>SLH</sub>/G46A/G109A (double mutant). The upstream and downstream part of the mutation site in SpaA<sub>SLH</sub> were amplified separately using as templates the purified recombinant plasmids pET22b-*spaAslh* for the single mutant and pET22b-*spaAslh*/G109A for the double mutant. The forward primer of the downstream part and the reverse primer of the upstream part were overlapping and included the point mutations that were introduced in both elongations. These two amplicons were mixed and amplified in a second round of PCR. Template DNA was degraded by *Dpn*I. Nicked plasmids were transformed into *E. coli* DH5α. The presence of G46A and G109A mutations was confirmed by direct DNA sequencing. Expression and purification of the single and double mutant proteins were performed as described for the SpaA<sub>SLH</sub> variant (see above). Primers for site-directed mutagenesis are listed in Supplementary Table 3.

**Generation of double TAAA mutants.** SpaA<sub>SLH</sub>/TAAA<sub>12</sub> and SpaA<sub>SLH</sub>/TAAA<sub>13</sub> mutants were generated by standard PCR amplification with the primers Fwd-SpaA<sub>SLH</sub>-*Nde*I and Rev-SpaA<sub>SLH</sub>-*Sac*I utilized to generate SpaA<sub>SLH</sub> (residues 21–93), using as template plasmid pETSMut4H (pET28a carrying His<sub>6</sub>-tagged *spaA* mutated in TRAE and TVEE motifs) for SpaA<sub>SLH</sub>/TAAA<sub>12</sub> and plasmid pET-Mut5H (pET28a carrying His<sub>6</sub>-tagged *spaA* mutated in TRAE and TRAQ motifs) for SpaA<sub>SLH</sub>/TAAA<sub>13</sub>.<sup>27</sup> The PCR products were digested and ligated into linearized pET22b vector using the *Nde*I and *Sac*I restriction enzymes. The presence of the TAAA mutations in SpaA<sub>SLH</sub>/TAAA variants was confirmed by direct DNA sequencing. Expression and purification of SpaA<sub>SLH</sub>/TAAA<sub>12</sub> and SpaA<sub>SLH</sub>/TAAA<sub>13</sub> mutants were carried out as described for the SpaA<sub>SLH</sub> variant.

**Generation of mutant *P. alvei* CCM 2051<sup>T</sup> strains.** In a first round, the SlhA<sub>SLH</sub> protein comprising amino acids 1124–1335 (the three consecutive SLH domains of SlhA) was generated by standard PCR amplification from *P. alvei* CCM 2051<sup>T</sup> genomic DNA. The PCR product was ligated into the pET22b vector using the *Nde*I and *Xho*I restriction sites. Subsequently, site-directed mutagenesis of residues glycine 1224 and glycine 1293 of protein SlhA to alanine was performed by overlapping extension PCR. The upstream and downstream part of the mutation sites were amplified separately using as templates the purified, recombinant plasmids pET22b-*slhA*<sub>SLH</sub> for the single mutant and pET22b-*slhA*<sub>SLH</sub>/G1293A for the double mutant (for more details, see section Site-directed mutagenesis). The presence of G1224 and G1293 mutations was confirmed by direct DNA sequencing. Next, a DNA fragment comprising around 400 bps, where the two glycine point mutations were included, was amplified by standard PCR using as template the purified, recombinant plasmid pET22b-*slhA*<sub>SLH</sub>/G1224A/G1293A. The amplification product was ligated into linearized vector pEXALV\_P(SlhA)\_SlhA<sup>49</sup>, which bears *slhA* native promoter, utilizing the *Pvu*II and *Kpn*I restriction sites. Primers for PCR amplification and site-directed mutagenesis are listed in Supplementary Table 3. The resulting plasmid, pEXALV\_P(SlhA)\_SlhA/G1224A/G1293A, was transformed into *P. alvei* CCM 2051<sup>T</sup> Δ*slhA* cells as described by Zarschler et al.<sup>51</sup>. Briefly, 500 ng of plasmid DNA was added to an aliquot of electro-competent cells, then the cell suspension was transferred into a pre-cooled 1-mm electroporation cuvette. After application of the pulse, the cell suspension was diluted with 4 mL of pre-warmed casein-peptone soymal-peptone broth, containing 250 mM sucrose, 5 mM MgCl<sub>2</sub>, and 5 mM MgSO<sub>4</sub>. Cells were incubated for 2 h at 37 °C, and then plated on LB agar plates supplemented with 10 µg/mL chloramphenicol and incubated overnight at 37 °C. The electroporation conditions were 200 Ω/25 µF/20 kV cm<sup>-1</sup>.

**Synthesis of SCWP ligands.** Described in Supplementary Methods.

**Circular dichroism spectroscopy.** CD spectroscopy for folding and secondary structure analyses of SpaA<sub>SLH</sub> was performed on a Chirascan CD apparatus (Applied Photophysics, Leatherhead, UK) equipped with a thermostatic cell holder and a Peltier element for temperature control. The instrument was flushed with a nitrogen flow at a rate of 5 L min<sup>-1</sup> and measurements were performed in the far UV region (180–260 nm) (Supplementary Fig. 7). The instrument parameters were as follows: path length, 1.0 mm; spectral bandwidth, 3.0 nm; step size, 1.0 nm; scan period, 10 s. Spectra were baseline-corrected to remove birefringence of the quartz cell. SpaA<sub>SLH</sub> was analyzed at a concentration of 10 µM in 20 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.0, at 20 °C.

To analyze the thermal stability of SpaA<sub>SLH</sub>, single wavelength scans were performed at 208 nm using the same protein solution as above, with a scan time per point of 10 s and stepwise temperature increase rate of 1 °C min<sup>-1</sup> over an interval from 20 to 90 °C (Supplementary Fig. 7).

**Isothermal titration calorimetry.** ITC was performed using a MCS titration calorimeter (Microcal, Inc., Northampton, MA). Concentrations of the mono-saccharide and disaccharide ligands were determined based on dry weight and confirmed by high-performance anion exchange chromatography with pulsed electrochemical detection on a PA-1 column (Thermo Fisher)<sup>52</sup>. Protein concentration was determined by UV/Vis light spectroscopy using a molar extinction coefficient of 18450 M<sup>-1</sup> cm<sup>-1</sup> at 280 nm. To avoid sample-related artifacts, protein solutions were freshly prepared prior to each set of titration experiments by dialysis of SpaA<sub>SLH</sub>, SpaA<sub>SLH</sub>/G109A, and SpaA<sub>SLH</sub>/G46A/G109A, respectively, against 20 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.0. The buffer dialysate was used for concentration adjustments and blank titrations. ITC measurements were done at 20 °C in 20 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.0. Protein, ligand, and buffer solutions were degassed prior to running the titrations. The sample cell was filled with 0.020 mM SpaA<sub>SLH</sub> solution (0.25 mL), the reference cell was filled with buffer and the injection syringe contained 20-times concentrated ligand solution. The injection sequence consisted of an initial injection of 0.5 µL of ligand solution to prevent low binding enthalpies arising from the filling of the syringe, followed by injection of 1 µL of ligand solution at 150-s intervals, each, until complete saturation of SpaA<sub>SLH</sub>-binding sites had been reached. SpaA<sub>SLH</sub>/G109A and SpaA<sub>SLH</sub>/G46A/G109A 0.020 mM solutions were titrated with a 20-times concentrated monosaccharide ligand solution, following a similar titration sequence as with *wt* SpaA<sub>SLH</sub>. A blank sample was run in the absence of protein to determine the heat of dilution of the ligand. Data analysis was done with Origin software (Microcal Inc.) by fitting a single-site binding isotherm. The obtained thermodynamic parameters were enthalpy of binding Δ*H*<sup>0</sup>, entropy of binding Δ*S*<sup>0</sup>, and the association constant *K*<sub>a</sub>. Measurements were performed at least in triplicate. Binding isotherms are shown in Supplementary Fig. 7.

**Crystallization and ligand soaking.** Recombinant, purified SpaA<sub>SLH</sub> was concentrated to 30 mg mL<sup>-1</sup> using Amicon ultra centrifugal filter units. Crystal screens were prepared using an Art Robbins Instruments crystal gryphon robot and Hampton 96-well Intelli plates. Crystals were obtained with the Hampton Index screen in condition number 54 (50 mM calcium chloride dihydrate, 0.1 M bis-Tris, pH 6.5, and 30% v/v PEG MME 550), and optimized by hanging drop vapor diffusion in 35 × 10 mm tissue culture dishes in the same condition at 16 °C with higher protein-to-reservoir ratios. Crystals grew within 1–3 days. 4,6-Pyr-β-D-ManNAcOMe was soaked into existing crystals of SpaA<sub>SLH</sub> at a concentration of 10 mM. 4,6-Pyr-β-D-ManNAcOMe was also co-crystallized with SpaA<sub>SLH</sub> in the same crystallization condition at 5 mM and 27 mg mL<sup>-1</sup>, respectively. Additional co-crystals with 4,6-Pyr-β-D-ManNAcOMe were obtained from the Hampton Index screen, with the crystals that provided the C2 structure grown in condition 66 (0.2 M ammonium sulfate, 0.1 M bis-Tris, pH 5.5, and 25% w/v PEG 3350), and the P1 structure in condition 87 (0.2 M sodium malonate, pH 7.0, 20% w/v PEG 3350). Co-crystals of SpaA<sub>SLH</sub> and β-D-GlcNAc-(1 → 3)-4,6-Pyr-β-D-ManNAcOMe were obtained in Hampton Index screen condition number 57 (50 mM ammonium sulfate, 50 mM bis-Tris, pH 6.5, and 30% v/v pentaerythritolthexylate [15/4 EO/OH]). Crystals of SpaA<sub>SLH</sub>/G109A were obtained with the Hampton Index screen in condition number 81 (0.2 M ammonium acetate, 0.1 M Tris, pH 8.5, and 25% w/v PEG 3350). Co-crystals of SpaA<sub>SLH</sub>/G109A with 4,6-Pyr-β-D-ManNAcOMe were obtained with the Hampton Index screen in condition number 95 (0.1 M potassium thiocyanate, pH 6.8, and 30% PEG MME 2000). Co-crystals of SpaA<sub>SLH</sub>/G46A/G109A with 4,6-Pyr-β-D-ManNAcOMe were obtained with the Hampton Index screen in condition number 47 (0.1 M bis-Tris, pH 6.5, and 28% PEG MME 2000).

**Data collection, structure determination, and refinement.** X-ray diffraction data were collected either on a Rigaku R-Axis IV++ area detector with X-rays produced by a Rigaku MM-003 generator, on a Marmosaic CCD300 detector at beamline CMCF-ID at the Canadian Light Source (CLS) synchrotron (Saskatoon, SK), or on a Dectris Pilatus 200 K detector with X-rays produced by a Rigaku Micromax-007 HF generator. Data were scaled, averaged, and integrated using HKL2000<sup>53</sup>. Crystals of unliganded SpaA<sub>SLH</sub> and co-crystals of SpaA<sub>SLH</sub> with 4,6-Pyr-β-D-ManNAcOMe were soaked overnight at 16 °C in mother liquor with the addition of 0.2 M KI prior to data collection. Phenix AutoSol<sup>54</sup> was used to solve

SAD phases from iodide ions and to generate initial models. Additional data sets were collected from native crystals and solved by molecular replacement with Phaser<sup>55</sup> using the lower-resolution KI derivative structures as search models. Data sets from Spa<sub>SLH</sub> crystals soaked and co-crystallized with synthetic 4,6-Pyr-β-D-ManNAcOMe, from co-crystals of Spa<sub>SLH</sub> with β-D-GlcNAc-(1 → 3)-4,6-Pyr-β-D-ManNAcOMe, from unliganded Spa<sub>SLH</sub>/G109A, and from co-crystals of Spa<sub>SLH</sub>/G46A/G109A with 4,6-Pyr-β-D-ManNAcOMe were solved by molecular replacement using the structure of unliganded Spa<sub>SLH</sub> as a search model. All model building and refinement were carried out using Coot<sup>56</sup> and Refmac5 through the CCP4 interface<sup>57,58</sup>.

**Visualization and graphics.** Figure 1 was produced using PROSITE sequence logo for domain profile P851272<sup>59</sup>. The chemical structures in Fig. 2 were prepared using MarvinSketch (version 15.11.30.0; <http://chemaxon.com>). All protein structure and electron density figures were produced with UCSF Chimera. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311)<sup>60</sup>.

**Swarming motility assay.** To investigate the role of the SLH-Gly29 residue in SLH domain-carrying proteins of Gram-positive bacteria within a biological context, an assay based on *P. alvei* CCM 2051<sup>1</sup> swarming ability on agar plates was used<sup>49</sup>. To test the swarming motility of *P. alvei* Δ*slhA* and *P. alvei* Δ*slhA*<sub>comp</sub> strains provided with plasmid-encoded *slhA* and *slhA*/G1224A/G1293A in comparison to *P. alvei* wt, cells were grown overnight (OD<sub>600</sub> ~2.0) and 5 μL of each culture were applied on 1% (semi-solid) LB agar plates. *P. alvei* wt and *P. alvei* Δ*slhA* cells were incubated at 37 °C for 24 h, whereas complemented, pEXALV vector-based *P. alvei* strains were incubated at 37 °C for 48 h. Images were taken using an EPSON PERFECTION V750 PRO scanner. The experiments were repeated at least in triplicates.

**Data availability.** The atomic coordinates and structure factors (Supplementary Table 1) have been deposited in the Protein Data Bank under the accession codes 6CWC, 6CWF, 6CWH, 6CWI, 6CWL, 6CWM, 6CWN, and 6CWR. The data that support the findings of this study are available from the corresponding author on request.

Received: 14 March 2017 Accepted: 19 June 2018

Published online: 07 August 2018

## References

- Sleytr, U. B. et al. S-layers as a tool kit for nanobiotechnological applications. *FEMS Microbiol. Lett.* **267**, 131–144 (2007).
- Messner, P., Schäffer, C. & Kosma, P. Bacterial cell-envelope glycoconjugates. *Adv. Carbohydr. Chem. Biochem.* **69**, 209–272 (2013).
- Sára, M. & Sleytr, U. B. S-layer proteins. *J. Bacteriol.* **182**, 859–868 (2000).
- Messner, P., Schäffer, C., Egelseer, E. & Sleytr, U. B. in *Prokaryotic Cell Wall Compounds* (eds König, H., Claus, H. & Varma, A.) 53–109 (Springer, Berlin Heidelberg, 2010).
- Baranova, E. et al. SbsB structure and lattice reconstruction unveil Ca<sup>2+</sup> triggered S-layer assembly. *Nature* **487**, 119–122 (2012).
- Arbing, M. A. et al. Structure of the surface layer of the methanogenic archaean *Methanosarcina acetivorans*. *Proc. Natl Acad. Sci. USA* **109**, 11812–11817 (2012).
- Chung, S., Shin, S.-H., Bertozzi, C. R. & De Yoreo, J. J. Self-catalyzed growth of S layers via an amorphous-to-crystalline transition limited by folding kinetics. *Proc. Natl Acad. Sci. USA* **107**, 16536–16541 (2010).
- Schäffer, C. & Messner, P. Emerging facets of prokaryotic glycosylation. *FEMS Microbiol. Rev.* **41**, 49–91 (2017).
- Messner, P., Steiner, K., Zarschler, K. & Schäffer, C. S-layer nanoglycobiology of bacteria. *Carbohydr. Res.* **343**, 1934–1951 (2008).
- Zarschler, K. et al. Protein tyrosine O-glycosylation - a rather unexplored prokaryotic glycosylation system. *Glycobiology* **20**, 787–798 (2010).
- Pleschberger, M. et al. An S-layer heavy chain camel antibody fusion protein for generation of a nanopatterned sensing layer to detect the prostate-specific antigen by surface plasmon resonance technology. *Bioconjug. Chem.* **15**, 664–671 (2004).
- Sekot, G. et al. Potential of the *Tannerella forsythia* S-layer to delay the immune response. *J. Dent. Res.* **90**, 109–114 (2011).
- Settem, R. P., Honma, K. & Sharma, A. Neutrophil mobilization by surface-glycan altered Th17-skewing bacteria mitigates periodontal pathogen persistence and associated alveolar bone loss. *PLoS ONE* **9**, e108030 (2014).
- Konstantinov, S. R. et al. S layer protein A of *Lactobacillus acidophilus* NCFM regulates immature dendritic cell and T cell functions. *Proc. Natl Acad. Sci. USA* **105**, 19474–19479 (2008).
- Ferner-Ortner, J., Mader, C., Ilk, N., Sleytr, U. B. & Egelseer, E. M. High-affinity interaction between the S-layer protein SbsC and the secondary cell wall polymer of *Geobacillus stearothermophilus* ATCC 12980 determined by surface plasmon resonance technology. *J. Bacteriol.* **189**, 7154–7158 (2007).
- Huber, C. et al. Heterotetramers formed by an S-layer-streptavidin fusion protein and core-streptavidin as a nanoarrayed template for biochip development. *Small* **2**, 142–150 (2006).
- Schäffer, C. & Messner, P. Surface-layer glycoproteins: an example for the diversity of bacterial glycosylation with promising impacts on nanobiotechnology. *Glycobiology* **14**, 31–42 (2004).
- Schäffer, C. et al. Novel biocatalysts based on S-layer self-assembly of *Geobacillus stearothermophilus* NRS 2004/3a: a nanobiotechnological approach. *Small* **3**, 1549–1559 (2007).
- Steiner, K. et al. Recombinant glycans on an S-layer self-assembly protein: a new dimension for nanopatterned biomaterials. *Small* **4**, 1728–1740 (2008).
- Egelseer, E. M. et al. in *Encyclopedia of Industrial Biotechnology: Bioprocess, Bioseparation, and Cell Technology* (ed Flickinger, M. C.) 7, 4424–4448 (John Wiley & Sons, Inc., Hoboken, 2010).
- Völlenkne, C. et al. Construction of a functional S-layer fusion protein comprising an immunoglobulin G-binding domain for development of specific adsorbents for extracorporeal blood purification. *Appl. Environ. Microbiol.* **70**, 1514–1521 (2004).
- Sleytr, U. B., Schuster, B., Egelseer, E.-M. & Pum, D. S-layers: principles and applications. *FEMS Microbiol. Rev.* **38**, 823–864 (2014).
- Moll, D. et al. S-layer-streptavidin fusion proteins as template for nanopatterned molecular arrays. *Proc. Natl Acad. Sci. USA* **99**, 14646–14651 (2002).
- Albers, S.-V. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426 (2011).
- Kandiba, L. & Eichler, J. Archaeal S-layer glycoproteins: post-translational modification in the face of extremes. *Front. Microbiol.* **5**, 1–5 (2014).
- Schäffer, C. & Messner, P. The structure of secondary cell wall polymers: how Gram-positive bacteria stick their cell walls together. *Microbiology* **151**, 643–651 (2005).
- Janesch, B., Messner, P. & Schäffer, C. Are the surface layer homology domains essential for cell surface display and glycosylation of the S-layer protein from *Paenibacillus alvei* CCM 2051<sup>1</sup>? *J. Bacteriol.* **195**, 565–575 (2013).
- Griffiths, S. G. & Lynch, W. H. Characterization of *Aeromonas salmonicida* variants with altered cell surfaces and their use in studying surface protein assembly. *Arch. Microbiol.* **154**, 308–312 (1990).
- Thomas, S. R. & Trust, T. J. Tyrosine phosphorylation of the tetragonal paracrystalline array of *Aeromonas hydrophila*: molecular cloning and high-level expression of the S-layer protein gene. *J. Mol. Biol.* **245**, 568–581 (1995).
- Fagan, R. P. & Fairweather, N. F. Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* **12**, 211–222 (2014).
- Engelhardt, H. & Peters, J. Structural research on surface layers: a focus on stability, surface layer homology domains, and surface layer-cell wall interactions. *J. Struct. Biol.* **124**, 276–302 (1998).
- Xu, Q. et al. Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. *Sci. Adv.* **2**, e1501254 (2016).
- Usenik, A. et al. The CWB2 cell wall-anchoring module is revealed by the crystal structures of the *Clostridium difficile* cell wall proteins Cwp8 and Cwp6. *Structure* **25**, 1–8 (2017).
- Finn, R. D. et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
- Cava, F., De Pedro, M. A., Schwarz, H., Henne, A. & Berenguer, J. Binding to pyruvylated compounds as an ancestral mechanism to anchor the outer envelope in primitive bacteria. *Mol. Microbiol.* **52**, 677–690 (2004).
- Mesnager, S. et al. Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO J.* **19**, 4473–4484 (2000).
- Schäffer, C. et al. A pyrophosphate bridge links the pyruvate-containing secondary cell wall polymer of *Paenibacillus alvei* CCM 2051 to muramic acid. *Glycoconj. J.* **17**, 681–690 (2000).
- May, A., Pusztahelyi, T., Hoffmann, N., Fischer, R.-J. & Bahl, H. Mutagenesis of conserved charged amino acids in SLH domains of *Thermoanaerobacterium thermosulfurigenes* EM1 affects attachment to cell wall sacculi. *Arch. Microbiol.* **185**, 263–269 (2006).
- Kern, J. et al. Structure of surface layer homology (SLH) domains from *Bacillus anthracis* surface array protein. *J. Biol. Chem.* **286**, 26042–26049 (2011).
- Forsberg, L. S. et al. Localization and structural analysis of a conserved pyruvylated epitope in *Bacillus anthracis* secondary cell wall polysaccharides and characterization of the galactose-deficient wall polysaccharide from avirulent *B. anthracis* CDC 684. *Glycobiology* **22**, 1103–1117 (2012).
- Sychantha, D. et al. Molecular basis for the attachment of S-layer proteins to the cell wall of *Bacillus anthracis*. *Biochemistry* **57**, 1949–1953 (2018).

42. Rünzler, D., Huber, C., Moll, D., Köhler, G. & Sára, M. Biophysical characterization of the entire bacterial surface layer protein SbsB and its two distinct functional domains. *J. Biol. Chem.* **279**, 5207–5215 (2004).
43. Altman, E., Brisson, J.-R., Messner, P. & Sleytr, U. B. Structure of the glycan chain from the surface layer glycoprotein of *Bacillus alvei* CCM 2051. *Biochem. Cell Biol.* **69**, 72–78 (1991).
44. Messner, P., Christian, R., Neuninger, C. & Schulz, G. Similarity of ‘core’ structures in two different glycans of tyrosine-linked eubacterial S-layer glycoproteins. *J. Bacteriol.* **177**, 2188–2193 (1995).
45. Zarschler, K. et al. Cell surface display of chimeric glycoproteins via the S-layer of *Paenibacillus alvei*. *Carbohydr. Res.* **345**, 1422–1431 (2010).
46. Li, J., Hu, X., Yan, J. & Yuan, Z. Species-specific cell wall binding affinity of the S-layer proteins of mosquitocidal bacterium *Bacillus sphaericus* C3-41. *Appl. Environ. Microbiol.* **75**, 3891–3895 (2009).
47. Mader, C., Huber, C., Moll, D., Sleytr, U. B. & Sára, M. Interaction of the crystalline bacterial cell surface layer protein SbsB and the secondary cell wall polymer of *Geobacillus stearothermophilus* PV72 assessed by real-time surface plasmon resonance biosensor technology. *J. Bacteriol.* **186**, 1758–1768 (2004).
48. Zona, R. & Janeček, Š. Relationships between SLH motifs from different glycoside hydrolase families. *Biol. Bratisl.* **60**, 115–121 (2005).
49. Janesch, B., Koerdts, A., Messner, P. & Schäffer, C. The S-layer homology domain-containing protein SlhA from *Paenibacillus alvei* CCM 2051(T) is important for swarming and biofilm formation. *PLoS ONE* **8**, e76566 (2013).
50. Cohen, I., Ron, I. G. & Ben-Jacob, E. From branching to nebula patterning during colonial development of the *Paenibacillus alvei* bacteria. *Phys. A* **286**, 321–336 (2000).
51. Zarschler, K., Janesch, B., Zayni, S., Schäffer, C. & Messner, P. Construction of a gene knockout system for application in *Paenibacillus alvei* CCM 2051T, exemplified by the S-layer glycan biosynthesis initiation enzyme WsP. *Appl. Environ. Microbiol.* **75**, 3077–3085 (2009).
52. Steiner, K. et al. Molecular basis of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus*. *J. Biol. Chem.* **283**, 21120–21133 (2008).
53. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
54. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
55. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
56. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
57. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
58. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
59. Sigrist, C. J. A. et al. New and continuing developments at PROSITE. *Nucleic Acids Res.* **41**, D344–D347 (2013).
60. Pettersen, E. F. et al. UCSF chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
61. Choudhury, B. et al. The structure of the major cell wall polysaccharide of *Bacillus anthracis* is species-specific. *J. Biol. Chem.* **281**, 27932–27941 (2006).

## Acknowledgements

The authors thank Dr. Paul Furtmüller and Dr. Roland Ludwig for assistance with CD and ITC measurements, Dr. Andreas Hofinger-Horvath for recording the NMR spectra, and Dr. Jean-Baptiste Farcet for providing MS data. This work was supported by the Austrian Science Fund FWF, projects P24305-B20 (to P.M.) and P27374-B22 (to C.S.), and by the Natural Sciences and Engineering Research Council of Canada (to R.J.B., S.M.L.G. and S.V.E.).

## Author contributions

R.J.B., A.L.-G., F.F.H., G.M., P.K., S.M.L.G., O.H.-G. and B.J. designed and performed the experiments. P.K., P.M., C.S. and S.V.E. supervised the project. R.J.B., A.L.-G., F.F.H., B. J., G.M., P.K., P.M., C.S. and S.V.E. contributed to writing the paper.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-05471-3>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018