

## Mechanistic basis of staphylococcal interspecies competition for skin colonization

Joseph J. Maciag<sup>1,a</sup>, Constance Chantraine<sup>2†</sup>, Krista B. Mills<sup>3†</sup>, Rahul Yadav<sup>4,b†</sup>, Alexander E. Yarawsky<sup>1,c†</sup>, Catherine T. Chaton<sup>1,d</sup>, Divya Vinod<sup>1,5,e</sup>, Nicholas C. Fitzkee<sup>4</sup>, Marion Mathelié-Guinlet<sup>6</sup>, Yves F. Dufrêne<sup>2</sup>, Paul D. Fey<sup>7</sup>, Alexander R. Horswill<sup>3</sup>, and Andrew B. Herr<sup>1,8,9\*</sup>

<sup>1</sup> Division of Immunobiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

<sup>2</sup> Louvain Institute of Biomolecular Science and Technology, UCLouvain, Louvain-la-Neuve, Belgium

<sup>3</sup> Department of Immunology and Microbiology, University of Colorado School of Medicine, Aurora, CO

<sup>4</sup> Department of Chemistry, Mississippi State University, Mississippi State, MS

<sup>5</sup> Medical Sciences Undergraduate Program, University of Cincinnati, Cincinnati, OH

<sup>6</sup> Institut de Chimie et Biologie des Membranes et des Nano-Objets, CNRS UMR 5248, University of Bordeaux, Pessac, France

<sup>7</sup> Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE

<sup>8</sup> Division of Infectious Diseases, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

<sup>9</sup> Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH

† These authors contributed equally and are listed alphabetically by last name.

<sup>a</sup> Current Affiliation: Department of Molecular Genetics, Biochemistry, & Microbiology, University of Cincinnati College of Medicine, Cincinnati, OH

<sup>b</sup> Current Affiliation: Physical Science Department, University of Arkansas at Fort Smith, Fort Smith, AR

<sup>c</sup> Current Affiliation: BioAnalysis, LLC, Philadelphia, PA

<sup>d</sup> Current Affiliation: Department of Molecular & Cellular Biochemistry and the Center for Structural Biology, University of Kentucky, Lexington, KY

<sup>e</sup> Current Affiliation: Ohio State University College of Medicine, Columbus, OH

\* Correspondence and requests for materials should be addressed to A.B.H. (email: [andrew.herr@cchmc.org](mailto:andrew.herr@cchmc.org))

## Abstract

Staphylococci, whether beneficial commensals or pathogens, often colonize human skin, potentially leading to competition for the same niche. In this multidisciplinary study we investigate the structure, binding specificity, and mechanism of adhesion of the Aap lectin domain required for *Staphylococcus epidermidis* skin colonization and compare its characteristics to the lectin domain from the orthologous *Staphylococcus aureus* adhesin SasG. The Aap structure reveals a legume lectin-like fold with atypical architecture, showing specificity for N-acetyllactosamine and sialyllactosamine. Bacterial adhesion assays using human corneocytes confirmed the biological relevance of these Aap-glycan interactions. Single-cell force spectroscopy experiments measured individual binding events between Aap and corneocytes, revealing an extraordinarily tight adhesion force of nearly 900 nN and a high density of receptors at the corneocyte surface. The SasG lectin domain shares similar structural features, glycan specificity, and corneocyte adhesion behavior. We observe cross-inhibition of Aap- and SasG-mediated staphylococcal adhesion to corneocytes. Together, these data provide insights into staphylococcal interspecies competition for skin colonization and suggest potential avenues for inhibition of *S. aureus* colonization.

The skin microbiome plays an important role in the maintenance of homeostasis and contributes to localized immune responses. Commensal skin bacteria such as *Staphylococcus epidermidis* generally play a beneficial role; for example, *S. epidermidis* and other coagulase-negative staphylococci can suppress invasion by *S. aureus* or other pathogens directly, via secreted factors<sup>1-5</sup>, or indirectly, by cooperating with local immune cells in the skin to induce antimicrobial protein production by keratinocytes<sup>6</sup>. *S. epidermidis* is a near-ubiquitous colonizer of healthy skin, whereas *Staphylococcus aureus* typically colonizes the skin in ~5% of healthy individuals<sup>7-10</sup>. However, up to 90% of individuals with atopic dermatitis (AD) are colonized with *S. aureus*, and AD flares are characterized by dysbiosis of the skin microbiome and overgrowth of *S. aureus* to a striking degree<sup>11</sup>. *S. aureus* colonization of the skin is associated with decreased epidermal barrier function and increased severity of inflammatory outcomes in AD<sup>12-17</sup> and AD patients often progress to other, more threatening, atopic diseases including food allergy, allergic rhinitis, and asthma<sup>13, 14, 18</sup>.

Cell wall surface proteins of Gram-positive bacteria mediate a wide variety of functions related to colonization and virulence. In both *S. epidermidis* and *S. aureus*, a large number of cell wall-anchored (CWA) proteins containing a C-terminal LPXTG sequence motif are covalently attached to peptidoglycan in the cell wall; these CWA proteins mediate processes including adhesion to host tissue, immune evasion, biofilm formation, and a variety of other enzymatic and non-enzymatic functions<sup>19-23</sup>. *S. epidermidis* expresses a CWA protein known as the Accumulation-associated protein (Aap), a large, multidomain protein first identified by its role in the accumulation phase of biofilm formation<sup>24</sup>. In addition to its role in mediating intercellular adhesion in biofilms<sup>25-31</sup>, Aap has been shown to be essential for skin colonization by *S. epidermidis* via adhesion to healthy human corneocytes<sup>32-34</sup>.

The domain architecture of Aap includes the N-terminal A domain, comprising the disordered A-repeat region with 11 imperfect repeats of a 16-residue motif and the following lectin domain. The A domain is immediately followed by 5 to 17 large 128-residue B-repeats that adopt an unusual, highly extended fold made up of unsupported 3-stranded  $\beta$ -sheets with minimal hydrophobic side chain packing<sup>26, 35</sup>. Each B-repeat is comprised of a G5 subdomain followed by a smaller E (or spacer) subdomain with a similar fold<sup>26, 35, 36</sup>. The B-repeat superdomain terminates with a variant G5 domain. The C-terminal end of Aap contains a repeating proline/glycine-rich (PGR) sequence that functions as a highly extended, noncompressible stalk region<sup>37</sup>. The LPXTG anchor is found at the end of the PGR region and is covalently attached to the cell wall by Sortase A<sup>38-40</sup>. The B-repeat superdomain is the region responsible for the intercellular adhesion leading to biofilm formation, but it first requires proteolytic processing by the trypsin-like staphylococcal protease SepA to remove the A domain (which includes the lectin domain)<sup>41</sup>. Once processed, the B-repeat superdomain can engage in self-assembly events

mediated by free  $Zn^{2+}$  to initially form antiparallel dimers that create twisted rope-like strands between staphylococcal cells in the nascent biofilm<sup>25, 26, 35, 42-44</sup>. These dimeric structures can further assemble to form tetrameric species that then nucleate functional amyloid fibers within the biofilm<sup>27, 43, 45</sup>.

While the B-repeat superdomain mediates intercellular adhesion in the biofilm, the A domain plays an important role in adhesion to both biotic and abiotic surfaces<sup>32-34</sup>. Early work demonstrated that the A domain was necessary for adhesion to human corneocytes<sup>33</sup>, and it has also been implicated in biofilm attachment to a catheter surface in a rat infection model<sup>28</sup>. Recently, it was demonstrated by Roy et al. that the Aap lectin subdomain within the A domain is the critical component required for skin colonization, as it is necessary and sufficient for *S. epidermidis* adhesion to healthy human corneocytes<sup>34</sup>. The authors also demonstrated that *S. epidermidis* adhesion could be blocked by enzymatic cleavage of specific glycan moieties, confirming that lectin binding activity was necessary for adhesion<sup>34</sup>.

*S. aureus* expresses an Aap ortholog named SasG that adopts a near-identical domain architecture including A-repeats, lectin domain, B-repeat superdomain, Pro/Gly-rich stalk, and LPXTG sortase motif (Fig. S1)<sup>46</sup>. Similar to Aap, the SasG A domain mediates adhesion to desquamated nasal epithelial cells<sup>47, 48</sup>. Likewise, the B-repeat superdomain of SasG adopts the same fold as Aap B-repeats and can self-assemble in a  $Zn^{2+}$ -dependent manner to mediate both homophilic and heterophilic interactions (with Aap B-repeats)<sup>36, 44, 49</sup>. Many unresolved questions remain about the structure of the Aap and SasG lectin domains, the relative conformation of the lectin with respect to the subsequent B-repeat superdomain, the native ligand specificity of both lectins, and the potential for competition between the two staphylococcal species for colonization of the human skin niche. Here we present multidisciplinary data describing high-resolution structures of both Aap and SasG lectin domains, the dynamics of the lectin fold, specificity for glycan ligands, adhesion to healthy human skin corneocytes, force spectroscopy on the adhesive interaction, and evidence for direct competition between *S. aureus* and *S. epidermidis* for adhesion to corneocytes and colonization of skin.

## Results

**Aap lectin adopts a legume lectin-like fold with an atypical binding site architecture.** The crystal structure of Aap lectin was solved to 1.06 Å resolution; phase determination involved single-wavelength native anomalous diffraction from the  $\delta$  sulfur atoms of two methionine residues (M436

and M593) and a structural calcium ion. Aap lectin adopts a fold similar to a large number of legume lectins, with two 6- or 7-stranded antiparallel  $\beta$ -sheets bridged by an orthogonal 5-stranded sheet along with several long loops, some of which include short  $\alpha$ -helices and one  $3_{10}$  helix (Fig. 1a-b). The 6-stranded  $\beta$ -sheet is relatively flat and solvent-exposed, transitioning on one end into the orthogonal 5-stranded sheet, whereas the 7-stranded sheet forms a concave surface over which four loops curl inward, coming together to create the boundaries of a large pocket on the protein surface with a tyrosine (Y580) at its base (Fig. 1d). Two structural metal ions, a  $\text{Ca}^{2+}$  and a  $\text{Na}^{+}$ , are bound on either side of the Y580 pocket. The Aap fold is most closely related to the lectin domain of SraP, another staphylococcal surface protein<sup>50</sup>; like SraP, Aap is a monomeric lectin (Fig. S2) in contrast to most of the legume lectins that form dimers or higher-order multimers. The location of the putative Aap binding pocket anchored by Y580 is located centrally with regard to the concave  $\beta$ -sheet, which differs from the more distal glycan binding site of SraP and other legume lectins (Fig. 1d-e).

Legume lectins are characterized by a structural  $\text{Ca}^{2+}$  ion adjacent to a conserved central aspartate residue with a *cis* backbone configuration whose side chain is positioned to interact with the glycan ligand<sup>51, 52</sup>. In the case of Aap, the conserved  $\text{Ca}^{2+}$  ion is present; coordination of this  $\text{Ca}^{2+}$  ion constrains one of the loops ('Ca loop') that folds over the concave  $\beta$ -sheet and helps create the Y580 pocket (Fig. 1d, Fig. 2a). The conserved aspartate (D431 in Aap) is also present and forms a water-mediated interaction with the neighboring  $\text{Ca}^{2+}$  ion; however, D431 is in the *trans* configuration in Aap (Fig. 2a-b). In other legume lectins, this conserved aspartate engages the glycan ligand, but D431 in Aap instead forms an H-bond with the backbone nitrogen of Y580, stabilizing another loop ('580 loop') that curls in over the concave  $\beta$ -sheet to help define the binding pocket (Fig. 1d, Fig. 2a). Interestingly, on the opposite side from the  $\text{Ca}^{2+}$  ion, Aap also features a structural  $\text{Na}^{+}$  ion with a coordination number of 5, ligated by 3 backbone carbonyl oxygens, a water, and the side chain oxygen of Q579. Thus, the coordination shells of the  $\text{Ca}^{2+}$  and  $\text{Na}^{+}$  ions provide constraints from both sides that stabilize the unusual configuration of residues Q579-Y580, positioned where the glycan typically binds in other legume lectins (e.g., see transparent sialic acid, Fig. 2b). As a result of these constraints provided by the metal ions, the surface in this region of Aap lectin forms a large pocket with a solvent-accessible volume of  $125 \text{ \AA}^3$  that contains the aromatic ring of Y580 centered in the pocket floor. The walls of the pocket are created by a convergence of loops including the  $\text{Ca}^{2+}$ -bound loop, the '580 loop' downstream of the bend at Y580, and part of the  $\text{Na}^{+}$ -bound loop; the bottom portion of the pocket is defined by the 'basal loop', a long loop and  $\alpha$ -helix 'plug' extending from the base of the  $\beta$ -sheet that folds up across the lower face of the lectin. In contrast, although SraP has a similar loop/helix plug defining the lower boundary of the binding site, in SraP it extends further up to create a binding pocket closer to the top of the lectin domain compared to Aap (Fig. 1d-e). Furthermore, the SraP loops corresponding to Aap's 580 loop and the  $\text{Na}^{+}$ -bound loop are significantly shorter and are

not long enough to curl over the concave  $\beta$ -sheet, providing a shallower surface that accommodates the bound glycan ligand (Fig. 1e).

### **Solution structure of the lectin-B-repeat junction and lectin domain dynamics.**

The N- and C-termini of the Aap lectin domain are adjacent to one another (Fig. 1b), which will direct the intrinsically disordered A-repeat region back toward the B-repeat superdomain. As a result, the lectin domain would be found at the extreme terminus of the highly elongated B-repeat superdomain, consistent with its function in adhesion to host tissue. In order to understand the structural arrangement between the lectin and the B-repeat region, we expressed a lectin-Brpt construct containing the native Aap sequence (351-816) of the lectin through the first 1.5 B-repeats (i.e., the 1<sup>st</sup> two G5 subdomains and the intervening E or spacer subdomain). This B-repeat length corresponds to a stable construct from the C-terminal portion of the B-repeat superdomain that has been characterized extensively<sup>25-27, 35, 42, 43, 45</sup>. The lectin domain alone, the lectin-Brpt construct, and the corresponding N-terminal Brpt1.5 region alone were characterized in solution by small-angle X-ray scattering (SAXS) (Fig. 3a). A dimensionless Kratky plot was consistent with a globular fold for the lectin domain and elongated conformations for both lectin-Brpt and Brpt1.5 alone (Fig. 3b). The lectin domain alone behaved as a globular protein, as expected, with a  $D_{\max}$  of 56 Å from the  $p(r)$  curve (Fig. 3c). The Brpt-lectin construct showed a typical pair distance distribution for an elongated protein with a globular domain at the end, with a  $D_{\max}$  value of 249 Å, compared to 170 Å for the Brpt1.5 construct alone (Fig. 3c). In order to model the solution structure of the lectin-Brpt construct, 50,000 models were generated using SASSIE based on the lectin domain structure and previous structures of other Brpt1.5 constructs<sup>26, 35</sup> as described<sup>43</sup>, allowing flexibility at the hinge between the lectin domain and the start of the B-repeat domains as well as flexion within the B-repeat region. After removal of models with clashes, the 12,000 remaining models were tested for consistency with the observed scattering curves using EOM (Fig. 3d-e). The best-fit models indicate that lectin domain adopts a moderately flexible position at the start of the B-repeat region (furthest from the staphylococcal cell wall), with the Y580 putative binding pocket ideally positioned to interact with a corneocyte receptor (Fig. 3f-g, Fig. S3). Interestingly, this arrangement is reminiscent of chaperone-usher pili from gram-negative bacteria, with a highly extended rod tipped by a lectin subunit for host adhesion<sup>53-55</sup>. However, such pili are formed by the concerted assembly of 5 or more major and minor pilin subunits and translocated by a dedicated membrane usher, whereas Aap acts as a single-molecule, multifunctional pilus-like structure.

The distal end of the lectin domain as it sits on the highly elongated B-repeat superdomain contains a series of loops; three of these are unusually long compared to other legume lectins or SraP, the closest structural homolog to Aap lectin (Fig. 1c, 1f). Two of the long loops (the 380 and 520 loops)

are on the back side of the lectin domain, whereas the 580 loop originates from the Y580 binding pocket (Fig. 1c, 2a). These three loops help to form a broad platform near the distal end of the lectin domain that is likely involved in ligand recognition by Aap (Fig. 1f). To better understand the structural dynamics of these long distal loops as well as the Ca<sup>2+</sup>-binding and Na<sup>+</sup>-binding loops, we conducted multidimensional NMR studies on Aap lectin. Several key residues in and near the Y580 loop could not be directly visualized, presumably because of intermediate conformational exchange; R1, R2, and hetNOE relaxation measurements confirmed that the same regions surrounding Y580 showed a high degree of conformational exchange (Fig. 4, Fig. S4). Conformational exchange in these loops is consistent with a pocket that is able to bind to several different glycans.

**Glycan specificity of the Aap lectin domain.** In order to define the glycan specificity, Aap lectin domain was submitted for a glycan array screen (Fig. 5a). The ten highest-affinity hits are illustrated in Fig. 5b. The majority of the hits were complex biantennary or triantennary N-linked glycans containing a repeating pattern of alternating Gal-GlcNAc residues (i.e, poly-N-acetylactosamine) at the terminal ends of the branching structures, with occasional decoration of the terminal Gal with sialic acid. Other hits included linear or branched glycans that again showed the N-acetylactosamine repeats, occasionally with fucosylation of the GlcNAc residues. The pattern of these top hits from the glycan array are consistent with data from Roy et al. who showed that Aap lectin-mediated adhesion of *S. epidermidis* to human corneocytes was inhibited by enzymatic degradation by galactosidase, neuraminidase, and fucosidase, which cleave galactose, sialic acid, and fucose residues, respectively<sup>34</sup>. To confirm these data, we conducted isothermal titration calorimetry binding experiments using N-acetylactosamine and 3'-sialyl-N-acetylactosamine (sialylactosamine). ITC data showed binding of N-acetylactosamine with ~130  $\mu$ M affinity, whereas sialylactosamine bound with 10-fold higher affinity ( $K_D = 13 \mu$ M) (Fig. 5c). As described above, the putative binding site on Aap is a pocket with the aromatic ring of Y580 centered at the base of the cavity. A common recognition mechanism in lectins involves the van der Waals packing of the nonpolar face of a glycan ring against an aromatic lectin residue, accompanied by H-bonds formed between lectin residues and the axial and equatorial hydroxyl groups from the sugar ring<sup>56, 57</sup>. Thus, Y580 is a likely candidate for a stacking interaction with a bound glycan. We produced a Y580A mutant of Aap lectin and determined that the mutant protein failed to bind either N-acetylactosamine or sialylactosamine (Fig. 5d). Despite repeated attempts, neither lactosamine-based glycan could be co-crystallized with Aap due to crystal packing interactions that sterically blocked the binding site. In addition, NMR titrations consistently showed no chemical shift perturbations in the presence of 20 mM lactosamine (Fig. S5). The relatively low affinity for N-acetylactosamine may be partly due to the flexibility in the region of the binding site detected by NMR (Fig. 4); however, the affinities observed for both glycan ligands are within the range reported for N-acetylactosamine or sialylactosamine-containing N-glycans binding to legume lectins<sup>58, 59</sup>.

### ***S. epidermidis* adhesion to human corneocytes involves glycan binding via the Y580 pocket.**

To validate our structure-guided mutagenesis and binding data, we measured adhesion of a stably GFP-expressing *S. epidermidis* 1457  $\Delta$ ica strain to healthy human corneocytes sampled from healthy volunteers by tape stripping. The  $\Delta$ ica strain of *S. epidermidis* has had the *icaADBC* operon deleted, which is responsible for the biosynthesis of the PNAG (PIA) biofilm polysaccharide; this strain was used to preclude PNAG-mediated adhesion. The *S. epidermidis* 1457  $\Delta$ ica strain adhered well to corneocytes in contrast to the  $\Delta$ ica $\Delta$ aap negative control (Fig. 6a). Consistent with the glycan-binding function of the lectin domain, addition of soluble N-acetyllactosamine inhibited *S. epidermidis* adhesion at all concentrations tested (ranging from 62.5–1000  $\mu$ M) (Fig. 6a); likewise, soluble sialyllactosamine inhibited adhesion at ten-fold lower concentrations (12.5–100  $\mu$ M; 6.25  $\mu$ M showed a trend toward inhibition but did not reach significance) (Fig. 6b). These inhibition data are consistent with the ITC results showing 10-fold higher binding affinity for sialyllactosamine (Fig. 5c). *S. epidermidis* adhesion was also inhibited by the addition of soluble recombinant Aap lectin domain at 5  $\mu$ M, whereas the non-binding Y580A mutant of Aap lectin completely failed to inhibit adhesion of *S. epidermidis* 1457  $\Delta$ ica (Fig. 6c), confirming the role of the Y580 binding pocket of Aap lectin in skin colonization by *S. epidermidis*.

### **Single-cell force spectroscopy reveals tight adhesion of *S. epidermidis* mediated by Aap.**

We quantified the strength of the interaction between the Aap A domain and human corneocytes using AFM-based single-cell force spectroscopy<sup>60, 61</sup>. Living *S. epidermidis* cells expressing variants of cell-surface Aap were immobilized on AFM cantilevers and used to image and force probe corneocytes in the quantitative imaging (QI) mode<sup>62</sup>. The results (Fig. 7a) demonstrated that *S. epidermidis* 1457  $\Delta$ sepA $\Delta$ ica (expressing full-length Aap onto its surface) engages in very strong interactions with corneocytes, with adhesion forces up to 3.5 nN. Adhesive events were densely and homogeneously distributed over the corneocyte surface, with an approximate binding probability of 90% (estimated from the QI map presented in Fig. 7a). In contrast, both *S. epidermidis* 1457  $\Delta$ aap $\Delta$ ica (lacking Aap) and *S. epidermidis* 1457  $\Delta$ sarA $\Delta$ ica (lacking the A domain) showed minimal adhesion, providing direct evidence that the interaction is mediated by the Aap A domain.

Further dissection of the adhesive interactions between *S. epidermidis* 1457  $\Delta$ sepA $\Delta$ ica and corneocytes by means of AFM force volume (FV) analysis (Fig. 7b-c and Fig. S6; total of 1000 adhesive curves from four different cell pairs) showed a large distribution of adhesion forces, ranging from 500 pN to 3,500 pN. These adhesive events were associated with a rupture length centered around 330 nm, matching the length of fully stretched adhesins, as previously demonstrated with Aap homophilic interactions involving the extension of only one Aap adhesin<sup>29</sup> (Fig. 7c). Gaussian fits of the force



distributions revealed three maxima (at ~900, 1,700 and 2,600 pN) that correspond to single, double, and triple bonds (Fig. 7b). Hence, bacteria engage in multiple strong bonds with their corneocyte ligand(s), and we expect that multivalency will strengthen bacterial adhesion. Retraction force profiles (Fig. 7d) featured sawtooth patterns with multiple equally spaced peaks originating from the unfolding of individual E and G5 repeats of the B region, followed by a single strong peak resulting from the rupture of the A domain-corneocyte interaction. Conversely, the *S. epidermidis* 1457  $\Delta aap\Delta ica$  and 1457  $\Delta sarA\Delta ica$  strains showed only few adhesion events coupled with small forces and the curves associated presented slight to no unfolding pattern (Fig. 7e and Fig. S7a,b). These results support even more the key role of the Aap A domain in this interaction. Excitingly, the 900 pN bond strength (obtained using the *S. epidermidis* 1457  $\Delta sepA\Delta ica$  strain) is an order of magnitude higher than that of single lectin-carbohydrate interactions<sup>63</sup>, implying that the A domain does not simply bind through a single “classical” lectin interaction. It is also larger than the strength recently reported for single Aap-Aap bonds (~600 pN)<sup>29</sup>. The G5-E repeats unfold at mechanical forces larger than the unfolding forces of  $\beta$ -fold multidomain proteins (<200 pN), indicating they are highly stable. Under tensile loading, the repeats will act as force buffers capable of relieving mechanical stress. Folded sacrificial domains in *S. epidermidis* adhesive polyproteins have been shown to maximize mechanical work dissipation and were suggested to serve as an adhesion strategy by bacteria<sup>64</sup>. The B-repeat superdomain may play a role in strengthening of the lectin-domain mediated interaction. The unusual mechanostability of the A domain-ligand complexes and of the B repeats is of biological relevance as it allows the pathogen to tightly attach to the skin under mechanical stresses occurring in vivo (e.g., shear, epithelial turnover, cell surface contacts).

### **SasG lectin adopts a similar fold and glycan specificity.**

Similar to *S. epidermidis*, *S. aureus* expresses a large CWA protein named SasG with the same overall domain architecture as Aap that shares many of the same functions (Fig. S1)<sup>46</sup>. Specifically, its A domain is implicated in adhesion to host tissue such as desquamated nasal epithelial cells<sup>46-48</sup> and the B-repeat superdomain engages in homophilic (or heterophilic) assembly interactions with other SasG (or Aap) molecules in a  $Zn^{2+}$ -dependent manner<sup>44, 49</sup>. We solved the crystal structure of the SasG lectin to 1.93 Å resolution, revealing the same basic fold as Aap (Fig. 8a). The structural  $Ca^{2+}$  ion is bound in the corresponding location to the site in Aap, and the  $Ca^{2+}$ -binding loop adopts nearly an identical configuration in the two lectins; similarly, the basal loop-helix ‘plug’ forming the bottom boundary of the binding site is conserved (Fig. 8b). However, the SasG structure does not show a bound structural  $Na^+$  ion, resulting in alternate conformations of the loops corresponding to the Aap  $Na^+$ -binding loop and 580 loop (referred to as the 390 loop in SasG). In place of Y580, SasG features a different aromatic residue, W392, in approximately the same position, which could still

permit stacking of the hydrophobic face of a sugar ring with the planar aromatic rings of the tryptophan side chain. As in Aap, the conserved central aspartate (D240) is in the atypical *trans* backbone configuration, and its carbonyl oxygen forms a water-mediated interaction with the Ca<sup>2+</sup> ion, while its side chain forms an H-bond with the amide nitrogen of W392. Thus, even though the Na<sup>+</sup> ion is not present, the structural Ca<sup>2+</sup> ion and the residues in its coordination shell add constraints that stabilize the atypical configuration of the loop containing W392 (390 loop), forming the base and upper boundary of the presumed binding pocket. We confirmed that SasG lectin is able to bind N-acetyllactosamine by ITC (Fig. 8c) but binding is abrogated in the W392A mutant (Fig. 8d), confirming that this binding pocket supports binding of lactosamine, as seen for Aap. Interestingly, SasG forms a much smaller binding pocket than Aap, raising the possibility that either N-acetyllactosamine binding involves an induced fit conformational change or that Na<sup>+</sup> binding, not observed in the SasG structure, may be needed to provide additional constraints that hold the binding pocket in the open state.

### **SasG mediates staphylococcal adhesion to human corneocytes via glycan binding to the W392 pocket.**

To test the role of SasG in staphylococcal adhesion to human corneocytes, we conducted corneocyte adhesion assays using the non-adherent coagulase-negative species *S. carnosus* engineered to express SasG on its surface, *S. carnosus*/pALC2073-*sasG*<sub>COL</sub> (*S. carnosus*-*sasG*), versus an empty-vector negative control. This approach avoided complications of wild-type *S. aureus* gene regulation and other potentially competing surface adhesins. The *S. carnosus*-*sasG* strain adhered well to healthy human corneocytes whereas the empty-vector control failed to adhere (Fig. 9a). As observed for *S. epidermidis*, N-acetyllactosamine at concentrations ranging from 62.5 to 1000 μM inhibited adhesion of *S. carnosus*-*sasG*, as did sialyllactosamine at ten-fold lower concentrations (12.5–100 μM) (Fig. 9a-b). Furthermore, soluble SasG lectin at 5 μM inhibited adhesion of *S. carnosus*-*sasG*, but the W392A mutant of SasG lectin inhibited adhesion to a significantly lesser degree than wild-type lectin (Fig. 9c). These data confirm the role of the SasG lectin domain in skin colonization and the specific role of the W392 glycan binding site.

### **Lectin-mediated cross-inhibition of *S. aureus* and *S. epidermidis* adhesion to human corneocytes.**

Given the similarities between Aap and SasG lectins in their overall structure, glycan specificity, and binding site architecture containing a prominent aromatic residue, the question arises as to whether *S. aureus* and *S. epidermidis* can compete with one another for adhesion to human corneocytes. We tested each soluble lectin for cross-inhibition of corneocyte adhesion (i.e., soluble SasG inhibiting *S. epidermidis* or soluble Aap inhibiting *S. carnosus*-*sasG*). We observed that 5 μM soluble SasG lectin inhibits adhesion of *S. epidermidis* to a similar degree as soluble Aap, but that the W392A mutant

failed to inhibit adhesion (Fig. 6c). Likewise, adhesion of *S. carnosus*-sasG was inhibited by the addition of 5  $\mu$ M soluble Aap lectin, whereas the Y580 mutant showed a clear trend toward decreased inhibition that did not reach significance; this loss of inhibitory effect was comparable to that seen with the W392A mutant of SasG lectin (Fig. 9c).

## Discussion

The data presented here demonstrate that Aap lectin domain adopts a fold generally similar to legume lectins as well as the staphylococcal CWA protein SraP<sup>50</sup>, but with an atypical binding site architecture due to the central *trans* aspartate residue (D431) that interacts with the amide nitrogen of Y580, forming the floor of the glycan binding pocket. Furthermore, Aap lectin features a structural Na<sup>+</sup> ion that provides additional constraints on the configuration of the glycan binding site. NMR relaxation data revealed a surprising degree of conformational flexibility in regions near the Aap binding site, particularly in the 580 loop that forms part of the glycan binding site (i.e., residue Y580) and extends from Y580 upward into one of three unusually long loops in Aap that diverge from other family members, including SraP. In addition to the long 580 loop, two other long loops extend away from the body of the lectin domain on the upper rear surface of the domain, forming a platform with significant surface area that is unique to Aap. We propose that this surface is likely to represent a binding site for the protein moiety of a corneocyte glycoprotein receptor (Fig. 10a). High-resolution AFM-based imaging showed that the corneocyte receptor recognized by Aap is expressed at high density and is distributed homogeneously across the corneocyte surface (Fig 8a). Single-cell force spectroscopy data revealed surprisingly strong adhesion forces (900 pN) for individual interactions between Aap and the corneocyte receptor; such an adhesion force is an order of magnitude stronger than that expected for typical glycan-lectin interactions. These data are consistent with a bipartite interaction between Aap lectin domain and both protein and N-glycan moieties of a corneocyte receptor rather than a simple lectin-glycan interaction alone. Previous force spectroscopy studies have revealed the remarkably high mechanostability of the B-repeat regions of both Aap and SasG<sup>29, 44</sup>. Given the direct connection between the lectin domain and the start of the B-repeat superdomain (Fig. 3), we speculate that the unique mechanical properties of the unusual B-repeat protein fold<sup>26, 36</sup> may also play a role in strengthening the Aap-corneocyte interactions.

The structure of SasG lectin is highly similar to Aap lectin, with the same atypical architecture resulting from the *trans* aspartate residue (D240) that interacts with the amide nitrogen of W392, placing an aromatic residue similar to Y580 at the base of the binding pocket. The binding pocket is notably smaller in the SasG structure compared to Aap, but this may be due to the lack of a bound structural Na<sup>+</sup> ion that was observed to aid in forming the binding pocket architecture in Aap. Furthermore, the three long loops observed in Aap that diverge from SraP are also present in SasG,

suggestive of similar bipartite recognition of a glycoprotein receptor and consistent with the overlapping specificity for corneocyte adhesion between SasG and Aap.

Taken together, our data strongly suggest that Aap and SasG lectin domains share specificity for a corneocyte receptor, presumably a glycoprotein, that interacts with a large surface patch on the top of the lectin domain formed primarily by the three long divergent loops (Fig. 1c, 1f, 10a). Interestingly, one of these long loops (the 580 loop in Aap), which originates in the Y580 glycan-binding pocket, shows a high degree of conformational exchange by NMR relaxation measurements (Fig. 4); the other two loops (380 loop and 520 loop) on the back side of the lectin also show increased dynamics, albeit to a lesser degree than the 580 loop. If, as expected, the protein moiety of the corneocyte receptor interacts with the distal (top) surface of the lectin, that interface on Aap would likely involve the 380, 520, and/or 580 loops (Fig. 10a). Our glycan array data suggest that the corneocyte receptor features an N-glycan decorated with 3' sialyllactosamine at the terminal end that binds in the Y580 binding pocket. Conformational exchange in the 580 loop may be necessary to facilitate binding of terminal sialyllactosamine groups (Fig. 10a), given the variability observed for N-glycoforms on individual glycoproteins<sup>65-68</sup>. The proposed contribution of both protein and glycan components of the corneocyte receptor to the binding free energy is consistent with the unusually strong adhesion force measured by single-cell force spectroscopy (Fig. 7).

The corneocyte adhesion assays (Figs 6 and 9) provided valuable insights into inter-species competition for colonizing the skin niche. Soluble Aap lectin domain inhibited corneocyte adhesion not only by *S. epidermidis*, but also by *S. carnosus-sasG*; likewise, soluble SasG lectin domain inhibited *S. epidermidis* adhesion. These data indicate that both *S. epidermidis* and *S. aureus* directly compete for the same corneocyte receptor. Both lectins were inhibited by N-acetyllactosamine and sialyllactosamine with similar dose responses. Given the near-ubiquitous colonization of skin by *S. epidermidis*, *S. aureus* would need to out-compete *S. epidermidis* for interaction with the same (or a highly similar) corneocyte receptor (Fig. 10b). However, *S. aureus* is known to express other CWA adhesins with specificity for corneocyte or keratinocyte receptors; co-expression of SasG with one or more of these additional adhesins may provide a critical advantage when competing with established *S. epidermidis* for skin colonization (Fig. 10b). One such *S. aureus* adhesin is the CWA protein SdrD, a member of the serine/aspartate-rich family of proteins. SdrD specifically interacts with desmoglein-1<sup>69</sup>, a corneodesmosomal cadherin that is expressed at high levels in outer epidermal layers. Furthermore, fibronectin-binding protein B (FnBP-B) and ClfB have been shown to interact with loricrin, another characteristic skin-resident protein<sup>70,71</sup>, and both FnBP-B and clumping factor B (ClfB) bind to corneodesmosin<sup>72</sup>. Context-dependent expression of various adhesins with specificity for skin receptors is likely to play an important role in colonization by *S. aureus* under specific

conditions, such as the overgrowth of *S. aureus* during AD flares. Upregulation of particular staphylococcal adhesins or combinations thereof could recognize corneocyte receptors that are differentially expressed (or exposed) under disease conditions. For example, AD skin is characterized by decreased expression levels of the critical structural protein filaggrin, resulting in changes to the architecture of the skin layers and potential exposure of keratinocyte receptors that would be inaccessible in healthy skin. Indeed, FnBP-B and ClfB have been shown to interact with corneodesmosin that is exposed in AD skin as a result of a lack of natural moisturizing factor<sup>72,73</sup>.

The corneocyte adhesion assays also validated the critical role of the Y580 or W392 binding pockets in adhesion via Aap or SasG, demonstrating that small soluble compounds (e.g., N-acetyllactosamine or sialyllactosamine) are efficacious at inhibiting staphylococcal colonization. This opens the door to the development of small-molecule therapeutic agents that target these binding pockets on Aap or particularly on SasG and specifically inhibit adhesion to skin. *S. aureus* produces a wide range of virulence factors including proteases that further damage the skin of AD patients, decrease its barrier function, and increase inflammation, likely predisposing AD patients to develop further atopic conditions<sup>13, 14, 18</sup>. Compounds targeting SasG-mediated adhesion, whether alone or in combination with other inhibitors, could find application in the case of AD or other skin conditions linked to *S. aureus*, including impetigo, cellulitis, or staphylococcal scalded skin syndrome.

## Methods

**Protein Purification.** pDest-HisMBP and pDest17-His plasmids containing the Aap lectin domain (residues 351-605) from *S. epidermidis* RP62A were transformed into BLR (DE3) cells (Novagen). Single colony isolates were transferred to 100 mL of liquid media (LB) supplemented with 50 µg/ml ampicillin sodium salt. Cultures were allowed to grow for 16 hours at 37° C. The overnight (O/N) culture was diluted into 1 L flasks of LB containing 50 µg/mL ampicillin sodium salt to an initial OD600 of 0.03. Cells grew for 3-5 hours, at 37° C, until an OD600 of 0.8-1.0 was reached. Flasks were transferred to an ice bath until the culture attained a temperature of 10° C (to induce a cold-shock response), at which point 22 mLs of ethanol (to maintain cold chaperonin expression) and 300 µM Isopropyl β-D-1-thiogalactopyranoside (IPTG) were added to each flask. Cells were then allowed to grow at 18° C for an additional 16 hours. Cells were harvested via centrifugation at 5,000 RPM for 1 hr. Supernatant was discarded and the cells were resuspended in 20 mM Tris pH 7.5, 150 mM NaCl. The suspension was then frozen overnight. Cells were thawed until a homogeneous suspension was established. The cells were then sonicated in 30 mL aliquots and centrifuged at 14,000 RPM for 1 hr. Insoluble pellet fractions were discarded and the supernatant was separated from the sample. The supernatant was applied to a gravity-flow column containing NiNTA (ThermoFisher) and eluted from the resin with a gradient of 30-300 mM imidazole. Isolated Aap lectin was then dialyzed and incubated with TEV protease to remove His tag or His-tagged fusion protein (MBP). The solution was dialyzed and applied to the NiNTA to remove TEV protease and His-tagged cleaved material. Flow through was then applied to a Superdex 75 column and purified lectin fractions were pooled. Likewise, SasG lectin (residues 163-419 from *S. aureus* 502a) was expressed using the pDest-HisMBP plasmid and purified as described for Aap lectin.

**Protein Crystallization and Data Collection.** All crystals were grown via the hanging drop vapor diffusion method. Protein stocks at a concentration of 20-25 mg/ml in 20 mM BisTris (pH 6.5) and 150 mM NaCl were combined with equal parts of mother liquor (1 µl +1 µL or 2µl +2 µL). Aap lectin crystals formed in various conditions: 1) 20-28% BCS PEG SMEAR Medium (Molecular Dimensions, CalibreScientific); 2) 100 mM Na Acetate (pH 4.5), 22.5% BCS PEG SMEAR High (Molecular Dimensions, CalibreScientific); and 3) 200 mM KCl, 200 mM CaCl<sub>2</sub>, 45% v/v 2-methyl 2-4 pentanediol and 100 mM BisTris (pH 6.5). SasG lectin crystals were grown in 100 mM Tris (pH 7.9), 40 mM Na Formate, 40 mM CaCl<sub>2</sub>, and 14-20% BCS PEG SMEAR Low (Molecular Dimensions, CalibreScientific). Crystals were added to a cryoprotectant consisting of 80% mother liquor and 20% MPD prior to being flash frozen in liquid nitrogen. Data collection occurred at the Advanced Photon Source at Argonne National Lab through the Northeastern Collaborative Access Team (NE-CAT), utilizing both 24-ID-C and 24-ID-E beamlines. Diffraction data were collected on Aap lectin crystals for

S-SAD using an oscillation range of 0.5°, collecting 1440 frames over a resolution range of 200-1.65 Å. Native data sets used an oscillation range of 0.2°, collecting 1800 frames, over a resolution of 200-1.03 Å. SasG data sets used an oscillation range of 0.5°, collecting 360 frames, at a resolution of 200–1.93 Å.

**Phasing, Refinement, and Validation.** Structural determination for Aap lectin was carried out by S-SAD phasing through the detection of three anomalous scattering atoms: sulfur atoms from two methionine residues and a bound calcium. Data was processed by the RAPD server available through NE-CAT (<https://rapd.nec.aps.anl.gov/rapd>); data indexing, space group assignment, scaling, and integration were carried out in RAPD via XDS<sup>74</sup>. Shelx<sup>75</sup> is utilized in the RAPD suite followed by PHENIX AutoSol (HySS, PHASER, RESOLVE, Xtriage, PHENIX-Refinement)<sup>76</sup> to determine anomalous scatterers, scatterer positions, calculation of experimental phases, density modifications, model-building, and refinement. Further refinement was carried out in PHENIX, and model building was done using Coot<sup>77</sup>, on local computers. The 1.06 Å native Aap lectin structure was solved by molecular replacement (MR) using Phaser-MR with the S-SAD-phased Aap as a search model. The SasG structure was solved by MR using Phaser-MR with native Aap lectin as a search model. Anisotropic refinement was carried out using PHENIX Refine. The final refined structures were validated using MolProbity<sup>78</sup>. CASTp (Computed Atlas of Surface Topography of proteins) was used to calculate the volume of surface pockets, including the putative glycan binding pocket<sup>79</sup>.

**Glycan array screening.** Aap lectin was dialyzed into PBS pH 8.0, concentrated to 2 mg/ml, and labeled with AlexaFluor 488 (10 µl of 10 mg/ml stock solution). The reaction was allowed to incubate at 4 °C for 48 h. Labeled Aap lectin was dialyzed into BisTris pH 6.5, 150 mM NaCl to quench the labeling reaction and remove free unbound AlexaFluor. Labeled Aap at 0.5 mg/ml was sent for glycan array screening by the Center for Functional Glycomics, using array version 5.4 consisting of 585 distinct glycans. The binding buffer contained 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2 mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 0.05% Tween 20, and 1% BSA. Data were reported as average RFU from replicate measurements (4 values averaged from 6 original replicates, after discarding the highest and lowest values per condition) along with standard deviation and %CV (%CV = 100x standard deviation / Mean).

**Small-angle X-ray scattering.** Lectin and Lectin-Brpt1.5 SEC-MALS-SAXS experiments were performed at the Advanced Proton Source (APS), Argonne, Illinois, on beam line 18ID, operated by BioCAT. Lectin samples used for SAXS were initially purified as described above. The lectin protein was dialyzed into a 50 mM MOPS (pH 7.2) and 50 mM NaCl buffer and was concentrated to 1 mg/mL prior to being transported to the BioCAT facilities. It was then concentrated on-site to 15-30 mg/mL.

Samples were centrifuged at 12,000 rpm for five minutes prior to injection onto a Superose 200. Pre-filtered sample aliquots of 237  $\mu$ l were applied to a Superose 6 (10/300 GL) column (Cytiva) on an Agilent 1260 Infinity II HPLC at a flow rate of 0.6 ml/min. The samples then passed through the column and were measured using, in order, an Agilent UV detector, a MALLS detector, a Wyatt Technologies Dawn Helios II, and an Optilab RI detector. Intensity measurements of scattering were taken using a Dectris Pilatus3 X 1M detector. Over a time course of 1 second, samples were exposed for 0.5 seconds. Brpt1.5 SAXS data were collected at The Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory using the mail-in HT-SAXS service on SIBYLS beamline 12.3.1<sup>80-83</sup>. Brpt1.5 was dialyzed into 50 mM MOPS pH 7.2, 50 mM NaCl, 2% glycerol and data collected on samples at 5 mg/mL, 6.7 mg/mL, 10 mg/mL, and 15 mg/mL.

SAXS analyses were performed in BioXTAS RAW (version 2.1.0)<sup>84</sup>. The concentration series collected for Brpt1.5 was extrapolated to zero concentration in Primus from ATSAS (version 3.0.3)<sup>85</sup> and then analyzed in BioXTAS RAW alongside the Lectin and Lectin-Brpt1.5 datasets. To model flexibility of Lectin-Brpt1.5, SASSIE's monomer Monte Carlo module<sup>86</sup> was used to generate 50,000 models. After removing clashing models, approximately 12,000 models remained. Flexible regions were allowed to sample up to 30° during each move. Flexible regions included the Lectin-Brpt1.5 hinge region (604-610) and the interfaces between each G5 and spacer domain of Brpt1.5 (686-689 and 735-738). CRY SOL<sup>87</sup> in non-fitting mode was used to generate SAXS scattering curves for each model, while also generating input files for the Ensemble Optimization Method (EOM)<sup>88,89</sup> using a custom python script<sup>43,90</sup>. CRY SOL settings were: max order of harmonics = 99; order of Fibonacci grid = 18; number of points = 101. CRY SOL was also run in fitting mode with the same settings to obtain single-model fits of the experimental dataset.

**Nuclear magnetic resonance spectroscopy.** Uniformly <sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N-labeled Aap lectin was expressed and purified as described above. All NMR relaxation experiments (Heteronuclear <sup>91-15</sup>N NOE, T<sub>1</sub>, and T<sub>1ρ</sub> experiments) were collected at 298K on a Bruker AVANCE III 600 MHz spectrometer with 0.5 mM lectin in NMR buffer (20 mM HEPES, pH 6.8, 50 mM NaCl) containing 10% D<sub>2</sub>O. The NMR data was referenced for <sup>1</sup>H chemical shifts by using 2, 2-dimethyl-2-silapentane-5-sulphonic acid (DSS) at 298 K as a standard. The <sup>13</sup>C and <sup>15</sup>N chemical shifts were referenced indirectly. The steady-state heteronuclear {<sup>1</sup>H}-<sup>15</sup>N NOE was measured as the ratios of the peak intensities with and without proton presaturation period (4 s) applied before the start of the <sup>1</sup>H-<sup>15</sup>N TROSY HSQC experiment. The T<sub>1</sub> and T<sub>1ρ</sub> experiments were recorded using a pulse program described by Lakomek *et al.*<sup>92</sup>. The spin-lattice T<sub>1</sub> experiments were collected as pseudo-three-dimensional (3D) TROSY-HSQC relaxation experiments with eight relaxation delays (0, 80, 160, 240, 320, 400, 640, and 900 ms). The spin-spin relaxation rate, R<sub>2</sub>, was determined from spin-lattice in



rotating frame  $T_{1\rho}$ <sup>93</sup> collected similarly as  $T_1$  with eight relaxation delays (2, 12, 24, 48, 72, 96, 120, and 144 ms). The  $T_{1\rho}$  spin lock power was 2 kHz, and temperature compensation was used for both the  $T_1$  and  $T_{1\rho}$  experiments. Relaxation experiments were acquired in an interleaved manner with  $512 \times 1344$  ( $t_1 \times t_2$ ) with a spectral width of 2128.62 and 9615.38 Hz. Spectra were processed using NMRPipe<sup>94</sup> program, and the relaxation rates,  $R_1$  and  $R_2$  were determined by fitting peak intensities as a function of relaxation decay time using NMRFAM-SPARKY<sup>95</sup>. Titrations in the presence of 20 mM lactosamine were collected at 298 K at 600 MHz under similar conditions as described above, using 0.5 mM <sup>15</sup>N labeled lectin in NMR buffer.

**Isothermal titration calorimetry.** Isothermal titration calorimetry (ITC) experiments were carried out using a MicroCal VP-ITC microcalorimeter. Standard experiments were performed at 25 °C in buffer containing 50 mM phosphate (pH 7.2) and 150 mM NaCl. Briefly, the sample cell contained 20  $\mu$ M of either Aap-lectin or SasG-lectin domains (1.5 mL) while the syringe had 1 mM of glycan (450  $\mu$ L). Twenty injections of glycan were delivered, the first of which was 2  $\mu$ l followed by nineteen injections of 14  $\mu$ l each. Controls experiments were performed with 1 mM glycan in the syringe and used to correct for heat of dilution. All lectin protein concentrations were calculated using either a ThermoScientific NanoDrop or a ThermoScientific BioMate 3S UV Spectrophotometer, where the absorbance at (A280) at 280 nm was measured. Glycan concentrations were calculated based on resuspension of purchased lyophilized glycans. Fitting of all ITC data to binding models was carried out using ORIGIN software.

**Corneocyte collection.** Corneocytes were collected as described in Mills et al. 2022<sup>96</sup>. Corneocytes were collected from either the lower or upper arm near the elbow from human volunteers with healthy skin. The selected skin area was cleaned with an alcohol wipe and allowed to air dry. Corneocytes were removed using clear, adhesive tape stripping discs (d-Squame D100; Clinical & Derm).

**Preparation of bacterial strains.** Bacterial strains used in this study are described in Table S2. Superfolder green fluorescent protein (sGFP)-expressing bacterial cultures for corneocyte adhesion assays were prepared as described in Mills et al. 2022<sup>96</sup>. 5 mL of *S. epidermidis* and *S. carnosus* strains were grown overnight in tryptic soy broth (TSB) at a 1:5 medium-to-culture tube ratio at 37°C with shaking aeration at 220 rpm. Cultures were grown with 10  $\mu$ g/mL chloramphenicol for plasmid maintenance. Cultures were then diluted 1:50 in TSB with chloramphenicol and grown to an OD600 of ~0.75. Cultures were washed once in phosphate-buffered saline (PBS) at a 1:1 ratio and diluted in PBS to a final OD600 of 0.15 (~10<sup>7</sup> CFU/mL). An OD600 of 0.15 allows for sufficient cell enumeration without bacterial clumping<sup>34, 96</sup>.

**Corneocyte adhesion assay with purified glycans and lectins.** Purified N-Acetyl-D-lactosamine (Sigma-Aldrich) was diluted in PBS to 1000  $\mu$ M. Serial dilutions were prepared at a 1:2 ratio in 300  $\mu$ L PBS from 1000  $\mu$ M to 62.5  $\mu$ M. Purified 3'-Sialyl-N-acetyl-lactosamine (Sigma-Aldrich) was diluted in PBS to 100  $\mu$ M. Serial dilutions were prepared at a 1:2 ratio in 300  $\mu$ L PBS from 100  $\mu$ M to 6.25  $\mu$ M. The corneocyte adhesion assays were adapted from Mills et al.<sup>96</sup>. Adhesion of *S. epidermidis*  $\Delta$ ica (which acts as a WT control due to this locus not contributing to intercellular adhesion<sup>34</sup>) and *S. carnosus*-SasG<sub>COL</sub> were tested for adhesion to healthy human corneocytes after co-incubation with purified glycans. 300  $\mu$ L of prepared bacterial cultures and 300  $\mu$ L of individual glycan serial dilutions as described above were mixed and co-incubated at room temperature for 20 minutes. The entire 600  $\mu$ L mixture was then incubated on corneocytes for 45 minutes at 37 °C. Corneocytes were incubated with 300  $\mu$ L of either *S. epidermidis*  $\Delta$ ica  $\Delta$ aap or *S. carnosus*-pALC2073 (EV) as negative controls.

Adhesion of *S. epidermidis*  $\Delta$ ica (and *S. carnosus*-SasG<sub>COL</sub>) were also tested for adhesion to healthy human corneocytes after pre-incubation/blocking with purified lectins. Purified recombinant Aap, Aap  $\Delta$ Y580A, SasG, and SasG  $\Delta$ W392A lectin domains were diluted in 300  $\mu$ L PBS to a concentration of 5  $\mu$ M. 300  $\mu$ L of purified lectins were incubated on corneocytes for 20 minutes at room temperature followed by incubation with 300  $\mu$ L of prepared bacteria (see above). Corneocytes were incubated with 300  $\mu$ L of either *S. epidermidis*  $\Delta$ ica  $\Delta$ aap or *S. carnosus*-pALC2073 (EV) as negative controls.

**Imaging and statistical analysis.** Corneocytes were imaged using fluorescence microscopy at X20 magnification using the bright-field and green channels. Ten images of each disc with corneocytes from three independent experiments (n = 30) were measured for adhesion percent area with Fiji ImageJ and analyzed in GraphPad Prism as described<sup>96</sup>.

**Atomic force microscopy.** Stationary phase cultures of *S. epidermidis* 1457  $\Delta$ sepA $\Delta$ ica,  $\Delta$ aap $\Delta$ ica, and  $\Delta$ sarA $\Delta$ ica bacteria were grown overnight in TSB (+ 10  $\mu$ g/ml trimethoprim for  $\Delta$ aap $\Delta$ ica only) while shaking (180 rpm) at 37 °C. They were then collected by centrifugation (5' at 2000xg), washed two times with PBS and resuspended in PBS, with a 50-fold dilution. A small amount (50  $\mu$ L) of the suspension was then deposited on the left side of a polystyrene Petri dish and the cells were allowed to adhere for 10' then washed with PBS. On the right side of the dish, a piece of corneocytes sample (5 mm x 5 mm) was immobilized using double-sided tape. The Petri dish was filled with 3 mL PBS. For preparing single-cell probes, tipless NPO-10 cantilevers (Bruker) were functionalized on their end with single glued silica microspheres (6.1  $\mu$ m diameter) and incubated for 1h in a polydopamine solution. The nominal spring constant of the colloidal probe cantilever (typically ~0.06 N/m) was determined by the thermal noise method. It was then brought into contact with single *S. epidermidis* bacteria weakly adhered to the surface of the Petri dish and retracted to attach the cell. Next the

probe was transferred to the corneocytes area to perform the force measurement using a JPK NanoWizard® 4 NanoScience AFM and JPK NanoWizard® Control Software v6.1.115 coupled to an optical microscope that allows observation of the sample. Force measurements were subsequently performed first in the quantitative imaging mode, using the following parameters: a contact setpoint of 0.5 nN, a z length of 1  $\mu\text{m}$ , a constant approach and retraction speed of 50  $\mu\text{m}/\text{s}$  and a scan area of 50  $\mu\text{m} \times 50 \mu\text{m}$  (256  $\times$  256 pixels) on top of a corneocyte. A conventional force volume map was also recorded on a smaller area (500 nm  $\times$  500 nm, 16  $\times$  16 pixels) with a contact setpoint of 0.25 nN and a slower constant approach and retraction speed of 1  $\mu\text{m}/\text{s}$ . Data were analyzed using the data analysis software from JPK and Origin software.

### **Acknowledgements**

The authors gratefully acknowledge expert advice and support from Drs. Frank Murphy, Jesse Hopkins, and Srinivas Chakravarthy from Argonne National Laboratory. Y.F.D. thanks Dr. Joan Geoghegan for providing corneocytes for AFM analysis. This research was conducted with NIH funding through NIGMS (R01 GM094363 to ABH) and NIAID (R01 AI162964 to PDF, ARH, and ABH; R01 AI153185 to ARH; and R01 AI139479 to NCF). Work at UCLouvain was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n°693630), and the National Fund for Scientific Research (FNRS). Glycan array screening was carried out by the Protein-Glycan Interaction Resource of the CFG (supported by NIGMS grant R24 GM098791) and the National Center for Functional Glycomics (NCFG) at Beth Israel Deaconess Medical Center, Harvard Medical School (supported by NIGMS grant P41 GM103694). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Diffraction data were collected at the Northeastern Collaborative Access Team (NE-CAT) beamlines, which are funded by the NIH through NIGMS (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by a NIH-ORIP HEI grant (S10OD021527). SEC-MALS-SAXS data were collected at the Biophysics Collaborative Access Team (BioCAT) beamline, supported by NIH grant P41 GM103622 through NIGMS. Use of the Pilatus 3 1M detector was provided by grant 1S10OD018090 from NIGMS. Brpt1.5 SAXS data collection was conducted at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research. Additional support comes from the NIH project ALS-ENABLE (P30 GM124169) and a High-End Instrumentation Grant S10OD018483. The content is solely the responsibility of the authors and does not necessarily reflect

the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

### **Author Contributions**

JJM, CC, KBM, AEY, CTC, NCF, MMG, YFD, PDF, ARH, and ABH designed experiments; JJM, CC, KBM, RY, AEY, CTC, and DV collected data; all authors analyzed data; and JJM, CC, KBM, RY, AEY, NCF, MMG, YFD, PDF, ARH, and ABH wrote the article.

### **Competing interests**

A.B.H. serves as a Scientific Advisory Board member for Hoth Therapeutics, Inc., holds equity in Hoth Therapeutics and Chelexa BioSciences, LLC, and was a co-inventor on seven patents broadly related to the subject matter of this work.

## References

1. Iwase T, *et al.* Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm formation and nasal colonization. *Nature* **465**, 346-349 (2010).
2. Nakatsuji T, *et al.* Antimicrobials from human skin commensal bacteria protect against Staphylococcus aureus and are deficient in atopic dermatitis. *Science translational medicine* **9**, (2017).
3. Cogen AL, *et al.* Selective antimicrobial action is provided by phenol-soluble modulins derived from Staphylococcus epidermidis, a normal resident of the skin. *The Journal of investigative dermatology* **130**, 192-200 (2010).
4. Paharik AE, *et al.* Coagulase-Negative Staphylococcal Strain Prevents Staphylococcus aureus Colonization and Skin Infection by Blocking Quorum Sensing. *Cell host & microbe* **22**, 746-756 e745 (2017).
5. Otto M, Echner H, Voelter W, Gotz F. Pheromone cross-inhibition between Staphylococcus aureus and Staphylococcus epidermidis. *Infect Immun* **69**, 1957-1960 (2001).
6. Naik S, *et al.* Commensal-dendritic-cell interaction specifies a unique protective skin immune signature. *Nature* **520**, 104-108 (2015).
7. Byrd AL, Belkaid Y, Segre JA. The human skin microbiome. *Nat Rev Microbiol* **16**, 143-155 (2018).
8. Ross AA, Muller KM, Weese JS, Neufeld JD. Comprehensive skin microbiome analysis reveals the uniqueness of human skin and evidence for phyllosymbiosis within the class Mammalia. *Proc Natl Acad Sci U S A* **115**, E5786-E5795 (2018).
9. Leung DY. Infection in atopic dermatitis. *Curr Opin Pediatr* **15**, 399-404 (2003).
10. Lin YT, Wang CT, Chiang BL. Role of bacterial pathogens in atopic dermatitis. *Clin Rev Allergy Immunol* **33**, 167-177 (2007).
11. Kong HH, *et al.* Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome Res* **22**, 850-859 (2012).
12. Biagini Myers JM, *et al.* Events in Normal Skin Promote Early-life Atopic Dermatitis - the MPAACH Cohort. *J Allergy Clin Immunol Pract*, (2020).
13. DeVore SB, Gonzalez T, Sherenian MG, Herr AB, Khurana Hershey GK. On the surface: Skin microbial exposure contributes to allergic disease. *Ann Allergy Asthma Immunol* **125**, 628-638 (2020).
14. Gonzalez T, Biagini Myers JM, Herr AB, Khurana Hershey GK. Staphylococcal Biofilms in Atopic Dermatitis. *Current allergy and asthma reports* **17**, 81 (2017).
15. Gonzalez T, *et al.* Biofilm propensity of Staphylococcus aureus skin isolates is associated with increased atopic dermatitis severity and barrier dysfunction in the MPAACH pediatric cohort. *Allergy* **76**, 302-313 (2021).

16. Stevens ML, *et al.* Simultaneous skin biome and keratinocyte genomic capture reveals microbiome differences by depth of sampling. *The Journal of allergy and clinical immunology* **146**, 1442-1445 (2020).
17. Simpson EL, *et al.* Patients with Atopic Dermatitis Colonized with *Staphylococcus aureus* Have a Distinct Phenotype and Endotype. *The Journal of investigative dermatology* **138**, 2224-2233 (2018).
18. Bantz SK, Zhu Z, Zheng T. The Atopic March: Progression from Atopic Dermatitis to Allergic Rhinitis and Asthma. *J Clin Cell Immunol* **5**, (2014).
19. Foster TJ, Geoghegan JA, Ganesh VK, Hook M. Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat Rev Microbiol* **12**, 49-62 (2014).
20. Foster TJ, Hook M. Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol* **6**, 484-488 (1998).
21. Navarre WW, Schneewind O. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* **63**, 174-229 (1999).
22. Paharik AE, Horswill AR. The Staphylococcal Biofilm: Adhesins, Regulation, and Host Response. *Microbiology spectrum* **4**, (2016).
23. Speziale P, Pietrocola G, Foster TJ, Geoghegan JA. Protein-based biofilm matrices in *Staphylococci*. *Frontiers in cellular and infection microbiology* **4**, 171 (2014).
24. Hussain M, Herrmann M, von Eiff C, Perdreau-Remington F, Peters G. A 140-kilodalton extracellular protein is essential for the accumulation of *Staphylococcus epidermidis* strains on surfaces. *Infect Immun* **65**, 519-524 (1997).
25. Conrady DG, Brescia CC, Horii K, Weiss AA, Hassett DJ, Herr AB. A zinc-dependent adhesion module is responsible for intercellular adhesion in staphylococcal biofilms. *Proc Natl Acad Sci U S A* **105**, 19456-19461 (2008).
26. Conrady DG, Wilson JJ, Herr AB. Structural basis for Zn<sup>2+</sup>-dependent intercellular adhesion in staphylococcal biofilms. *Proc Natl Acad Sci U S A* **110**, E202-211 (2013).
27. Yarawsky AE, Johns SL, Schuck P, Herr AB. The biofilm adhesion protein Aap from *Staphylococcus epidermidis* forms zinc-dependent amyloid fibers. *J Biol Chem* **295**, 4411-4427 (2020).
28. Schaeffer CR, *et al.* Accumulation-associated protein enhances *Staphylococcus epidermidis* biofilm formation under dynamic conditions and is required for infection in a rat catheter model. *Infect Immun* **83**, 214-226 (2015).
29. Wang C, *et al.* The staphylococcal biofilm protein Aap mediates cell–cell adhesion through mechanically distinct homophilic and lectin interactions. *PNAS Nexus* **1**, pgac278 (2022).
30. Rohde H, *et al.* Polysaccharide intercellular adhesin or protein factors in biofilm accumulation of *Staphylococcus epidermidis* and *Staphylococcus aureus* isolated from prosthetic hip and knee joint infections. *Biomaterials* **28**, 1711-1720 (2007).

31. Rohde H, *et al.* Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. *Mol Microbiol* **55**, 1883-1895 (2005).
32. Conlon BP, *et al.* Role for the A domain of unprocessed accumulation-associated protein (Aap) in the attachment phase of the *Staphylococcus epidermidis* biofilm phenotype. *J Bacteriol* **196**, 4268-4275 (2014).
33. Macintosh RL, *et al.* The terminal A domain of the fibrillar accumulation-associated protein (Aap) of *Staphylococcus epidermidis* mediates adhesion to human corneocytes. *J Bacteriol* **191**, 7007-7016 (2009).
34. Roy P, Horswill AR, Fey PD. Glycan-Dependent Corneocyte Adherence of *Staphylococcus epidermidis* Mediated by the Lectin Subdomain of Aap. *mBio* **12**, e0290820 (2021).
35. Shelton CL, Conrady DG, Herr AB. Functional consequences of B-repeat sequence variation in the staphylococcal biofilm protein Aap: deciphering the assembly code. *The Biochemical journal* **474**, 427-443 (2017).
36. Gruszka DT, *et al.* Staphylococcal biofilm-forming protein has a contiguous rod-like structure. *Proc Natl Acad Sci U S A* **109**, E1011-1018 (2012).
37. Yarawsky AE, English LR, Whitten ST, Herr AB. The Proline/Glycine-Rich Region of the Biofilm Adhesion Protein Aap Forms an Extended Stalk that Resists Compaction. *J Mol Biol* **429**, 261-279 (2017).
38. Mazmanian SK, Liu G, Ton-That H, Schneewind O. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760-763 (1999).
39. Schneewind O, Missiakas DM. Staphylococcal Protein Secretion and Envelope Assembly. *Microbiology spectrum* **7**, (2019).
40. Ton-That H, Mazmanian SK, Alksne L, Schneewind O. Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Cysteine 184 and histidine 120 of sortase form a thiolate-imidazolium ion pair for catalysis. *J Biol Chem* **277**, 7447-7452 (2002).
41. Paharik AE, *et al.* The metalloprotease SepA governs processing of accumulation-associated protein and shapes intercellular adhesive surface properties in *Staphylococcus epidermidis*. *Mol Microbiol* **103**, 860-874 (2017).
42. Chaton CT, Herr AB. Defining the Metal Specificity of a Multifunctional Biofilm Adhesion Protein. *Protein Sci*, (2017).
43. Yarawsky AE, Hopkins JB, Chatzimagas L, Hub JS, Herr AB. Solution Structural Studies of Pre-amyloid Oligomer States of the Biofilm Protein Aap. *J Mol Biol* **434**, 167708 (2022).
44. Formosa-Dague C, Speziale P, Foster TJ, Geoghegan JA, Dufrene YF. Zinc-dependent mechanical properties of *Staphylococcus aureus* biofilm-forming surface protein SasG. *Proc Natl Acad Sci U S A* **113**, 410-415 (2016).
45. Yarawsky AE, Herr AB. The staphylococcal biofilm protein Aap forms a tetrameric species as a necessary intermediate before amyloidogenesis. *J Biol Chem* **295**, 12840-12850 (2020).

46. Roche FM, *et al.* Characterization of novel LPXTG-containing proteins of *Staphylococcus aureus* identified from genome sequences. *Microbiology* **149**, 643-654 (2003).
47. Corrigan RM, Rigby D, Handley P, Foster TJ. The role of *Staphylococcus aureus* surface protein SasG in adherence and biofilm formation. *Microbiology* **153**, 2435-2446 (2007).
48. Roche FM, Meehan M, Foster TJ. The *Staphylococcus aureus* surface protein SasG and its homologues promote bacterial adherence to human desquamated nasal epithelial cells. *Microbiology* **149**, 2759-2767 (2003).
49. Geoghegan JA, *et al.* Role of surface protein SasG in biofilm formation by *Staphylococcus aureus*. *J Bacteriol* **192**, 5663-5673 (2010).
50. Yang YH, *et al.* Structural insights into SraP-mediated *Staphylococcus aureus* adhesion to host cells. *PLoS Pathog* **10**, e1004169 (2014).
51. Gabius HJ, Andre S, Jimenez-Barbero J, Romero A, Solis D. From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem Sci* **36**, 298-313 (2011).
52. Loris R, Hamelryck T, Bouckaert J, Wyns L. Legume lectin structure. *Biochim Biophys Acta* **1383**, 9-36 (1998).
53. Dodson KW, Pinkner JS, Rose T, Magnusson G, Hultgren SJ, Waksman G. Structural basis of the interaction of the pyelonephritic *E. coli* adhesin to its human kidney receptor. *Cell* **105**, 733-743 (2001).
54. Hospenthal MK, Costa TRD, Waksman G. A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat Rev Microbiol* **15**, 365-379 (2017).
55. Waksman G, Hultgren SJ. Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat Rev Microbiol* **7**, 765-774 (2009).
56. Kolatkar AR, Leung AK, Isecke R, Brossmer R, Drickamer K, Weis WI. Mechanism of N-acetylgalactosamine binding to a C-type animal lectin carbohydrate-recognition domain. *J Biol Chem* **273**, 19502-19508 (1998).
57. Quijcho FA. Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu Rev Biochem* **55**, 287-315 (1986).
58. Haseley SR, Talaga P, Kamerling JP, Vliegenthart JF. Characterization of the carbohydrate binding specificity and kinetic parameters of lectins by using surface plasmon resonance. *Anal Biochem* **274**, 203-210 (1999).
59. Surolia A, Sharon N, Schwarz FP. Thermodynamics of monosaccharide and disaccharide binding to *Erythrina corallodendron* lectin. *J Biol Chem* **271**, 17697-17703 (1996).
60. Dufrene YF. Towards nanomicrobiology using atomic force microscopy. *Nat Rev Microbiol* **6**, 674-680 (2008).
61. Viljoen A, *et al.* Force spectroscopy of single cells using atomic force microscopy. *Nature Reviews Methods Primers* **1**, 63 (2021).

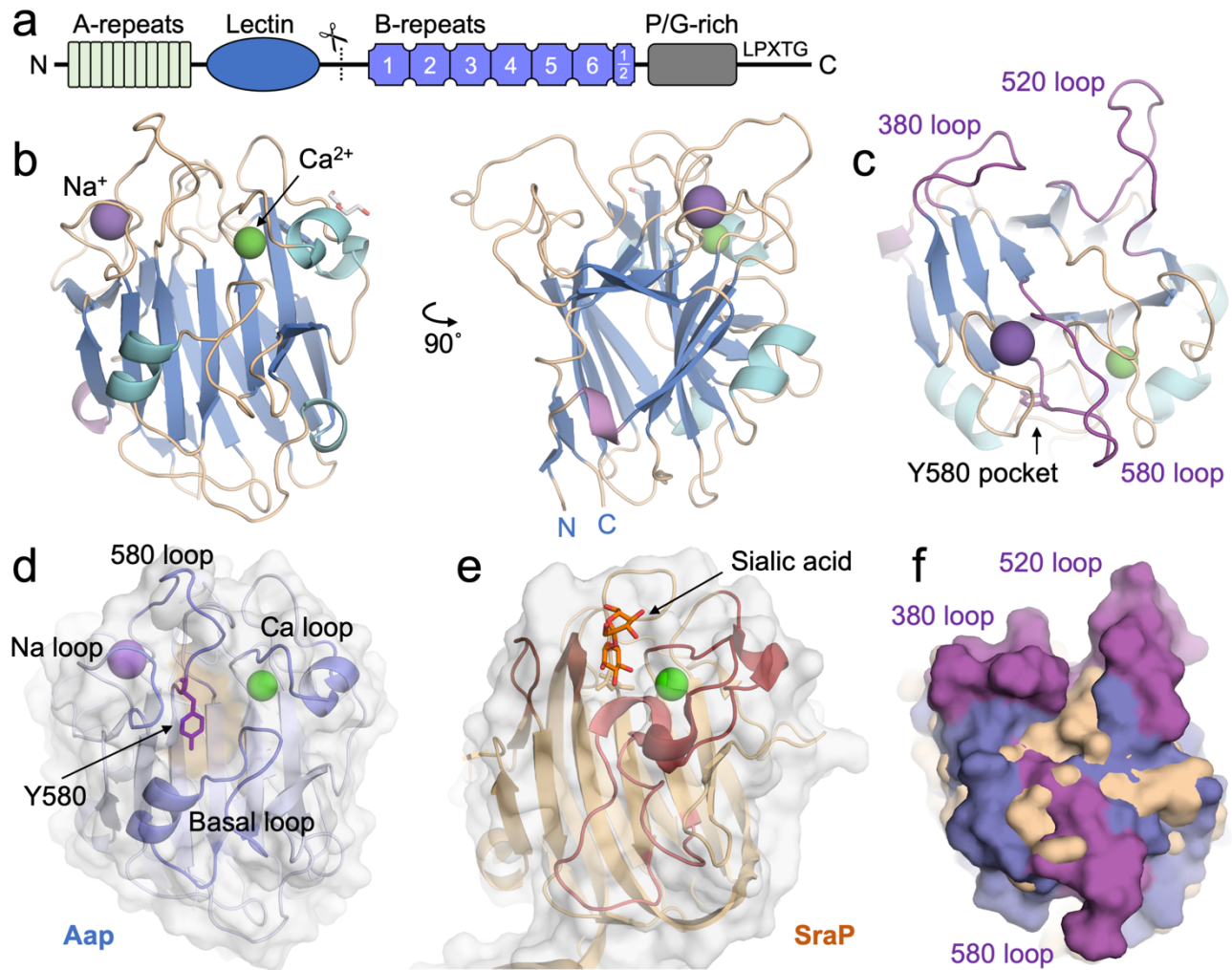


62. Formosa-Dague C, *et al.* Forces between *Staphylococcus aureus* and human skin. *Nanoscale Horiz* **1**, 298-303 (2016).
63. Muller DJ, Helenius J, Alsteens D, Dufrene YF. Force probing surfaces of living cells to molecular resolution. *Nat Chem Biol* **5**, 383-390 (2009).
64. Liu H, Liu Z, Yang B, Lopez Morales J, Nash MA. Optimal Sacrificial Domains in Mechanical Polyproteins: *S. epidermidis* Adhesins Are Tuned for Work Dissipation. *JACS Au* **2**, 1417-1427 (2022).
65. Gomes MM, Wall SB, Takahashi K, Novak J, Renfrow MB, Herr AB. Analysis of IgA1 N-glycosylation and its contribution to Fc $\alpha$ RI binding. *Biochemistry* **47**, 11285-11299 (2008).
66. Cao L, *et al.* Global site-specific analysis of glycoprotein N-glycan processing. *Nature protocols* **13**, 1196-1212 (2018).
67. Kolarich D, Jensen PH, Altmann F, Packer NH. Determination of site-specific glycan heterogeneity on glycoproteins. *Nature protocols* **7**, 1285-1298 (2012).
68. Zaia J. Mass spectrometry and the emerging field of glycomics. *Chemistry & biology* **15**, 881-892 (2008).
69. Askarian F, *et al.* The interaction between *Staphylococcus aureus* SdrD and desmoglein 1 is important for adhesion to host cells. *Scientific reports* **6**, 22134 (2016).
70. da Costa TM, Viljoen A, Towell AM, Dufrene YF, Geoghegan JA. Fibronectin binding protein B binds to loricrin and promotes corneocyte adhesion by *Staphylococcus aureus*. *Nature communications* **13**, 2517 (2022).
71. Vitry P, *et al.* Force-Induced Strengthening of the Interaction between *Staphylococcus aureus* Clumping Factor B and Loricrin. *mBio* **8**, (2017).
72. Towell AM, *et al.* *Staphylococcus aureus* binds to the N-terminal region of corneodesmosin to adhere to the stratum corneum in atopic dermatitis. *Proc Natl Acad Sci U S A* **118**, (2021).
73. Feuillie C, *et al.* Adhesion of *Staphylococcus aureus* to Corneocytes from Atopic Dermatitis Patients Is Controlled by Natural Moisturizing Factor Levels. *mBio* **9**, (2018).
74. Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).
75. Sheldrick GM. A short history of SHELX. *Acta Crystallogr A* **64**, 112-122 (2008).
76. Liebschner D, *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).
77. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
78. Chen VB, *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
79. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res* **46**, W363-W367 (2018).

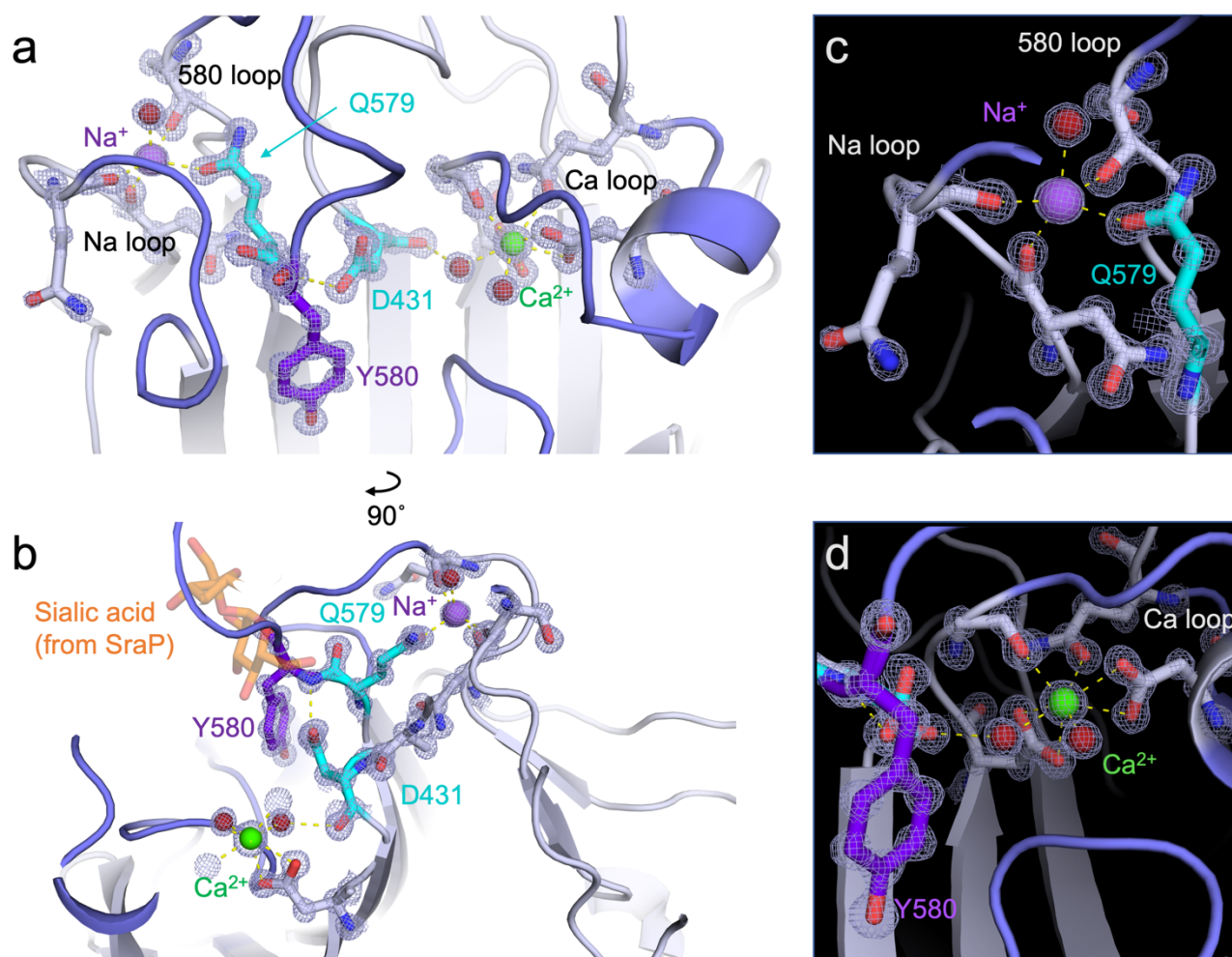
80. Classen S, *et al.* Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source. *J Appl Crystallogr* **46**, 1-13 (2013).
81. Dyer KN, *et al.* High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. *Methods in molecular biology* **1091**, 245-258 (2014).
82. Hura GL, *et al.* Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat Methods* **6**, 606-612 (2009).
83. Putnam CD, Hammel M, Hura GL, Tainer JA. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40**, 191-285 (2007).
84. Hopkins JB, Gillilan RE, Skou S. BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J Appl Crystallogr* **50**, 1545-1553 (2017).
85. Manalastas-Cantos K, *et al.* ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J Appl Crystallogr* **54**, 343-355 (2021).
86. Curtis JE, Raghunandan S, Nanda H, Krueger S. SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints. *Comput Phys Commun* **183**, 382-389 (2012).
87. Svergun DI, Barberato C, Koch MHJ. CRY SOL— a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768-773 (1995).
88. Bernado P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI. Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* **129**, 5656-5664 (2007).
89. Tria G, Mertens HD, Kachala M, Svergun DI. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207-217 (2015).
90. Yarawsky AE, Hopkins JB, Herr AB. Python code to run CRY SOL and generate EOM input files (SAXS analysis.) (2022).
91. Preimesberger MR, *et al.* Direct NMR detection of bifurcated hydrogen bonding in the alpha-helix N-caps of ankyrin repeat proteins. *J Am Chem Soc* **137**, 1008-1011 (2015).
92. Lakomek NA, Ying J, Bax A. Measurement of (1)(5)N relaxation rates in perdeuterated proteins by TROSY-based methods. *J Biomol NMR* **53**, 209-221 (2012).
93. Cavanagh J, Fairbrother WJ, Palmer AG, 3rd, Rance M, Skelton NJ. *Protein NMR Spectroscopy*. Academic Press (2007).
94. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).

95. Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327 (2015).
96. Mills KB, Roy P, Kwiecinski JM, Fey PD, Horswill AR. Staphylococcal Corneocyte Adhesion: Assay Optimization and Roles of Aap and SasG Adhesins in the Establishment of Healthy Skin Colonization. *Microbiology spectrum* **10**, e0246922 (2022).

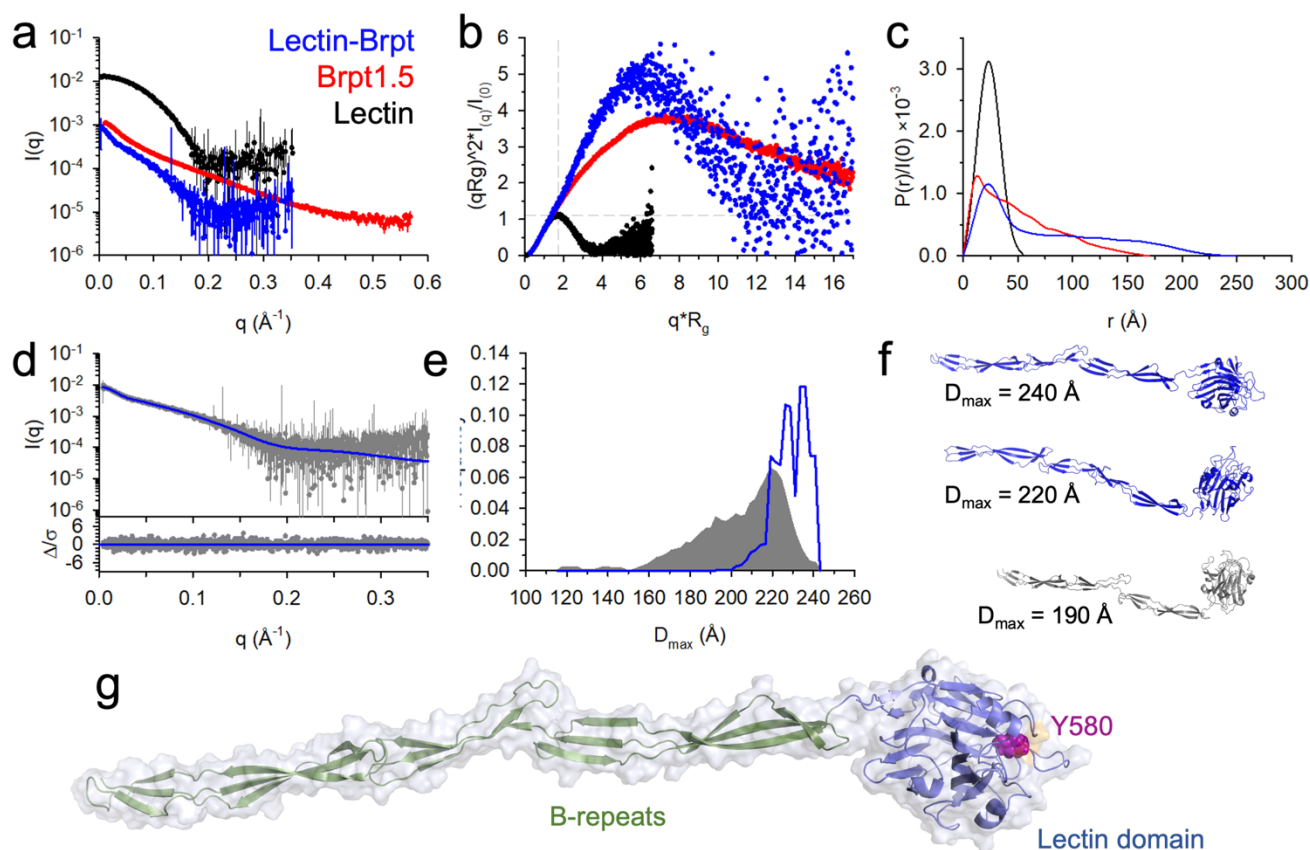
## Figures



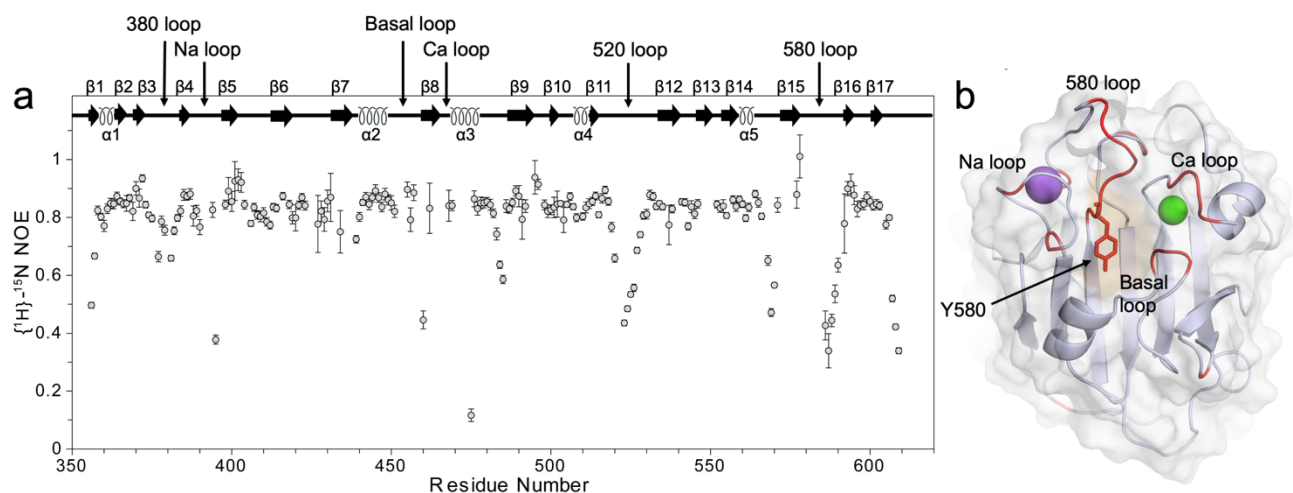
**Fig. 1: Crystal structure of Aap lectin reveals an atypical legume lectin-like fold.** (a) Domain arrangement of Aap from *S. epidermidis* 1457. The scissors icon indicates the cleavage site for SepA that unmasks the B-repeat superdomain and allows protein-dependent biofilm formation. (b) Front and side views of the crystal structure of Aap lectin domain, colored by secondary structural elements.  $\beta$ -strands are colored blue,  $\alpha$ -helices cyan,  $3_{10}$  helices lavender, and loops tan. (c) Top view of Aap lectin, highlighting the divergent long loops (380, 520, and 580 loops) in purple. (d) Front view of Aap lectin with transparent surface, showing the loops that converge on the front face to create a putative glycan binding pocket with Y580 forming its floor. The pocket is colored tan to aid in visualization. (e) Front view of the related lectin domain from the *S. aureus* adhesin SraP, showing the equivalent loops that form its binding pocket, which is located further toward the distal end of the domain compared to Aap. (f) Surface view of the top of the lectin domain of Aap superimposed on SraP. Aap is colored in blue, with the divergent long loops colored purple; SraP is colored tan. The long loops extend beyond the SraP boundary, creating a distinct surface that may be used for target receptor recognition.



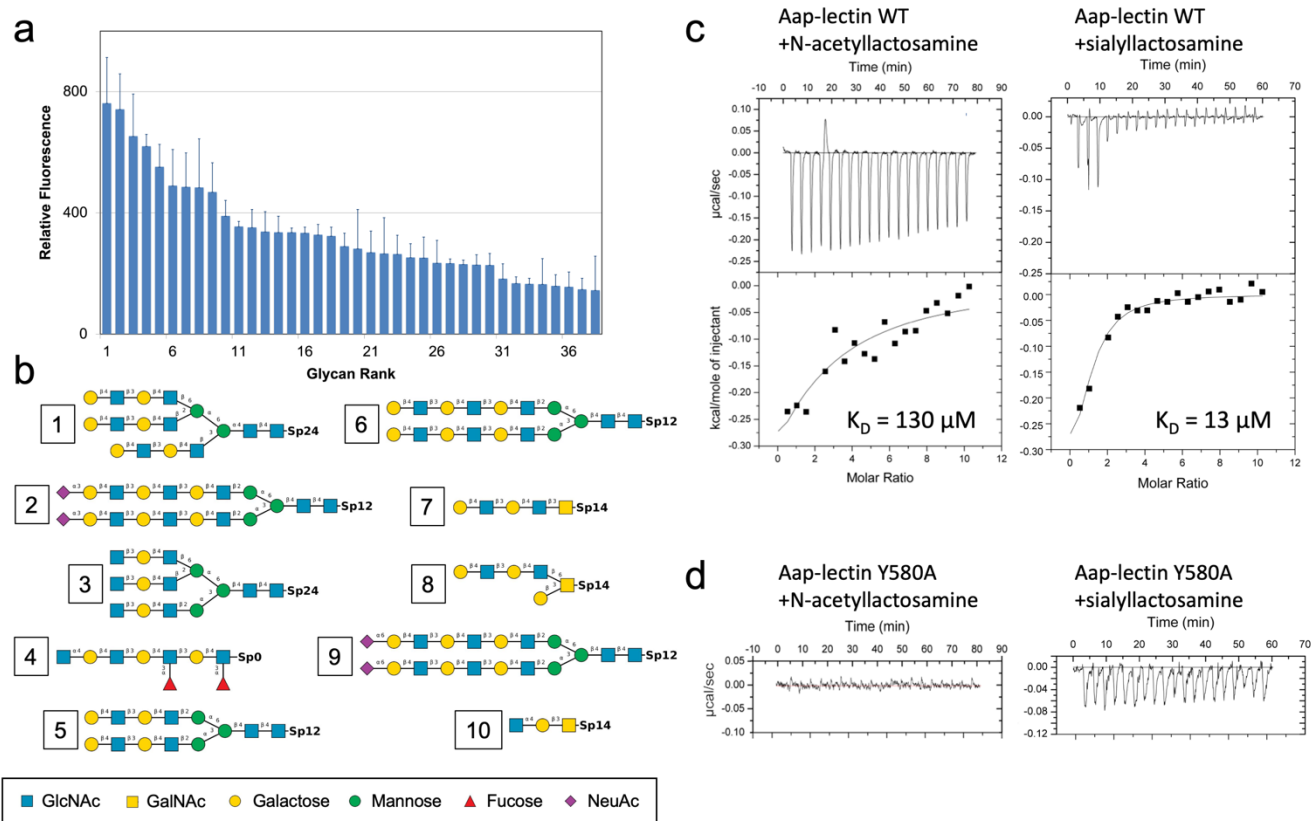
**Fig. 2: Atypical architecture near the Aap glycan binding site.** (a) Model and  $2f_o - f_c$  electron density map illustrating the front view of the Aap lectin in the region of the binding site. The main chain forms a sharp bend at Q579-Y580 that helps create the binding pocket with Y580 as its floor; this bend in the main chain is stabilized on either side via interactions with structural metal ions. D431 is the conserved aspartate residue from legume lectins that typically forms a *cis* peptide bond and whose side chain usually interacts with the glycan ligand. In Aap, D431 forms a *trans* peptide bond and the side chain interacts with the amide nitrogen of Y580, stabilizing the floor of the atypical binding pocket. The carbonyl oxygen of D431 interacts with the structural Ca<sup>2+</sup> ion indirectly via a water-mediated H-bond. Q579 itself forms a direct H-bond with the bound Na<sup>+</sup> ion on the opposite side of the domain from the Ca<sup>2+</sup> binding site. (b) View from behind the Ca<sup>2+</sup> ion looking toward the binding pocket, showing the interaction network involving both Ca<sup>2+</sup> and Na<sup>+</sup> ions, D431, and the Q579-Y580 region of the main chain. For reference, the sialic acid from the SraP binding site is shown as transparent sticks, showing that Y580 occupies the same approximate position as bound glycans in other related lectins. (c) Details of the Na<sup>+</sup> coordination scheme, which shows a typical coordination number of 5. (d) Details of the Ca<sup>2+</sup> coordination scheme, with a typical coordination number of 8.



**Fig. 3: SAXS solution structural characterization of Aap-Brpt.** (a) The SAXS scattering curves for Lectin (black), Brpt1.5 (red), and Lectin-Brpt (blue). (b) The dimensionless Kratky plot. Dashed lines indicate the intersection at which a Gaussian peak should be observed for a globular particle. (c)  $P(r)$  distributions demonstrating overall particle dimensions, where  $D_{\max}$  is determined by the intersection with the x-axis. (d) The EOM fit (blue line) of the experimental data (gray), with the difference plot in the lower panel. (e) The  $D_{\max}$  distribution from EOM showing the distribution of SASSIE-generated models (gray) based on flexibility of the hinge region between the lectin and B-repeat, as well as flexion along the Brpt1.5 region. The distribution of the best-fit EOM models is shown in blue. (f) The top two models (blue) were frequently chosen as part of the EOM ensemble, while the lower model (gray) is from the modeling pool, but not in the range of  $D_{\max}$  of the ensemble. (g) Structure of the most frequently chosen model from the EOM ensemble, highlighting the position of the Y580 binding pocket.

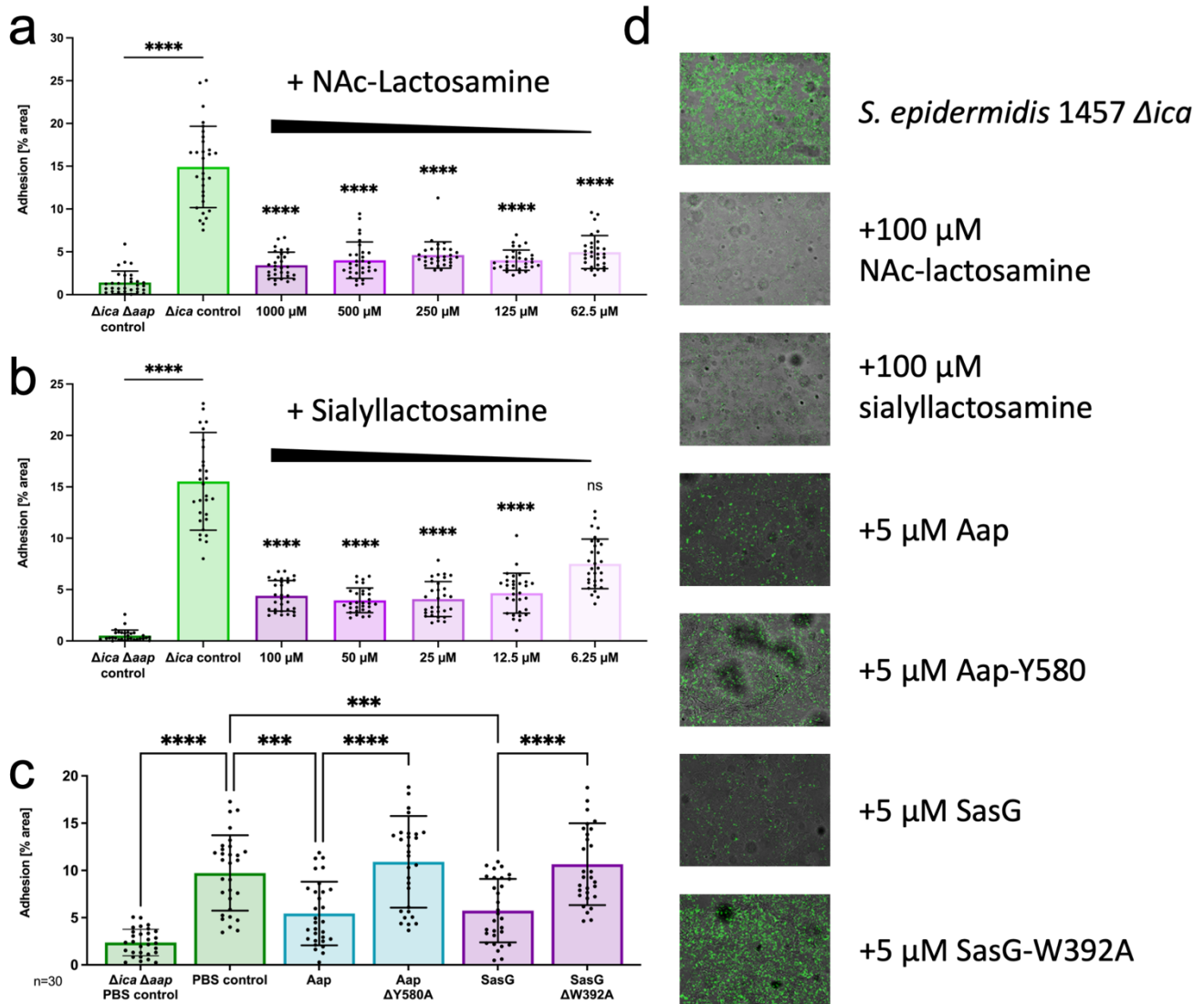


**Fig. 4: NMR reveals conformational exchange of Aap lectin loops.** (a)  $\{^1\text{H}\}-^{15}\text{N}$  heteronuclear NOEs (hetNOE) of Aap-Lectin indicate the flexibility of loops surrounding the glycan binding site. (b) The residues that could not be assigned in the solution NMR study of Aap-lectin are highlighted in red. Residues near the Y580 pocket include Q579 through S585 in the 580 loop; N471 through D473 in the Ca loop; L452-E454 in the basal loop; and residues E391, T396, and T397 in the Na loop.

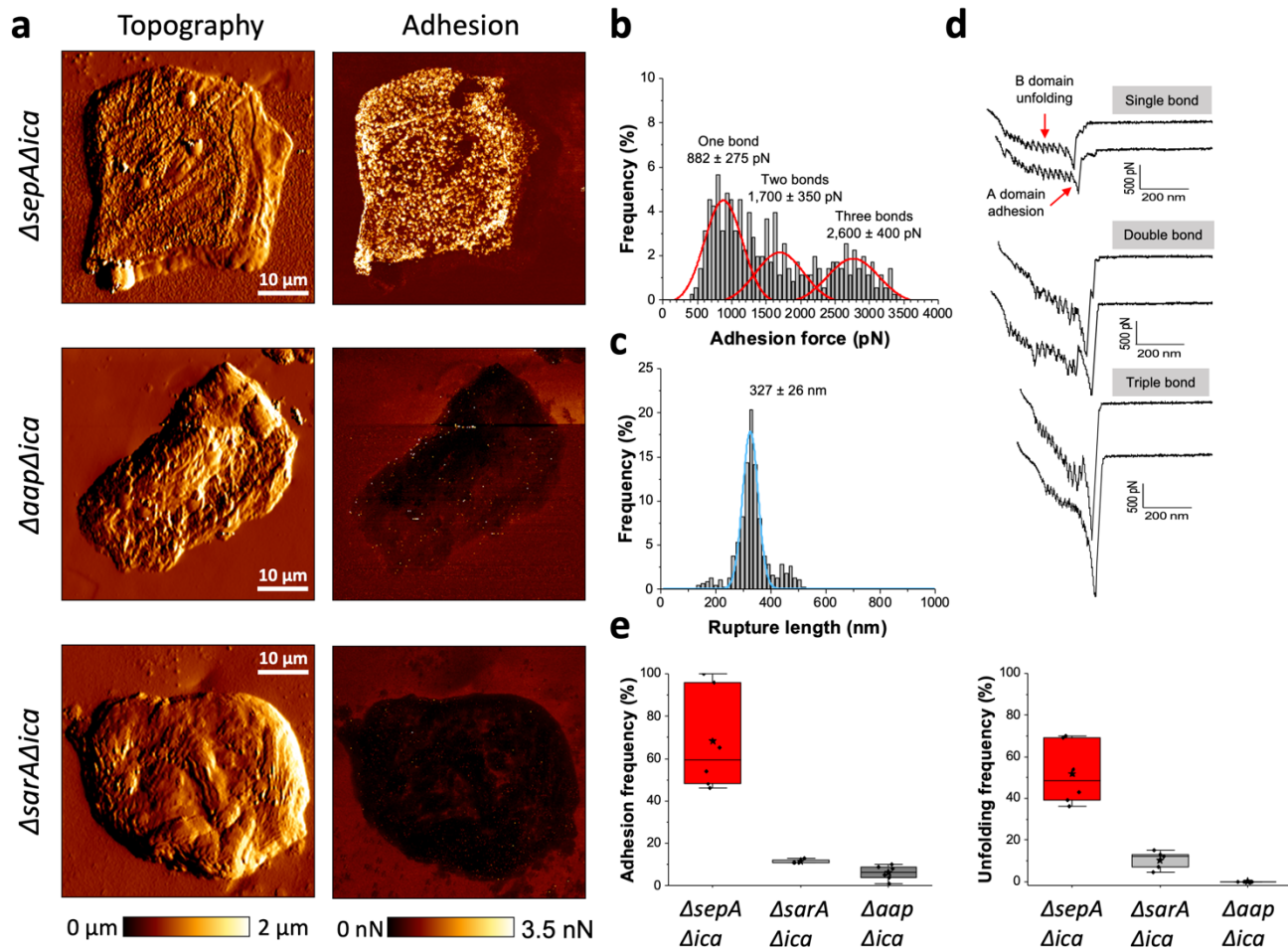


**Fig. 5: Glycan specificity of Aap lectin.** (a) Relative fluorescence measurements for labeled Aap lectin domain binding to a glycan array. (b) Structures of the top 10 glycan hits from the array. Note that all top hits share a repeating Gal-GlcNAc disaccharide (N-acetyllactosamine) that can be modified by sialylation at the terminal end. (c) ITC data for Aap lectin binding to N-acetyllactosamine and sialyllactosamine. (d) ITC data for Aap-Y580A mutant binding to N-acetyllactosamine and sialyllactosamine.

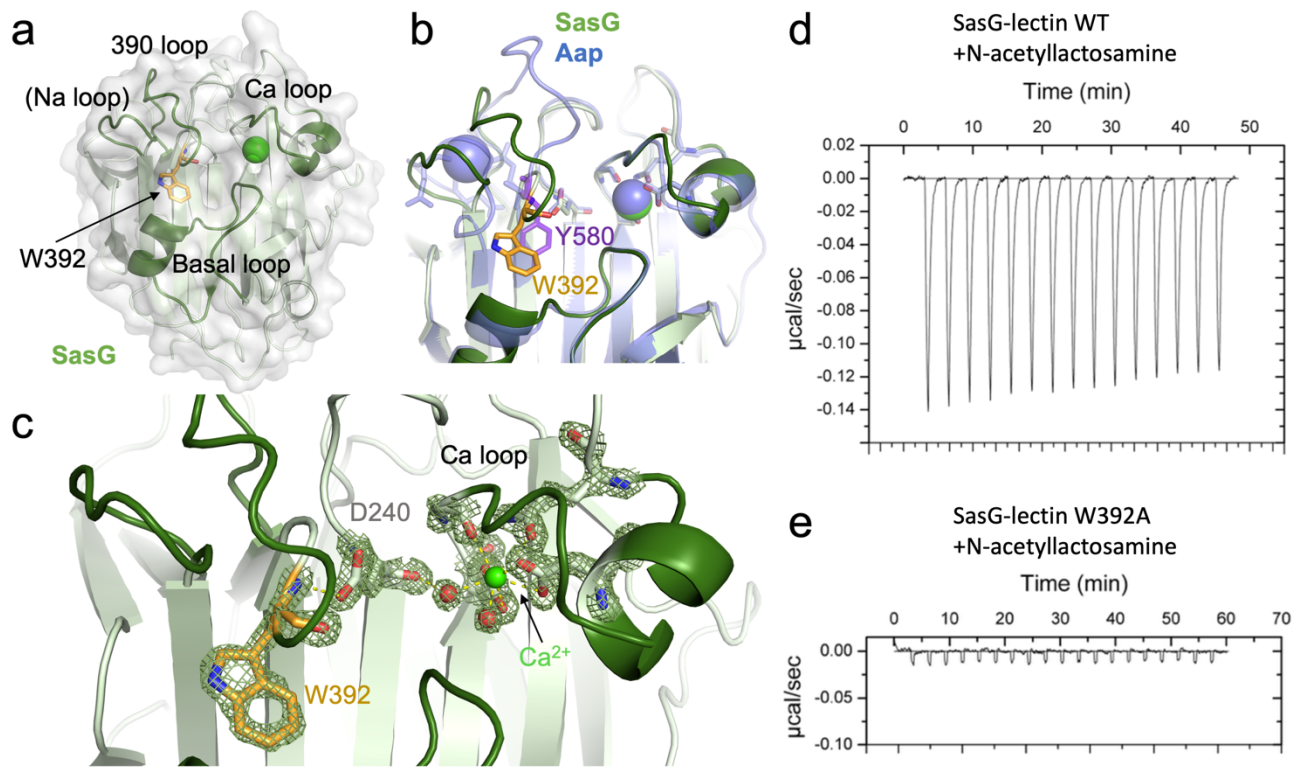




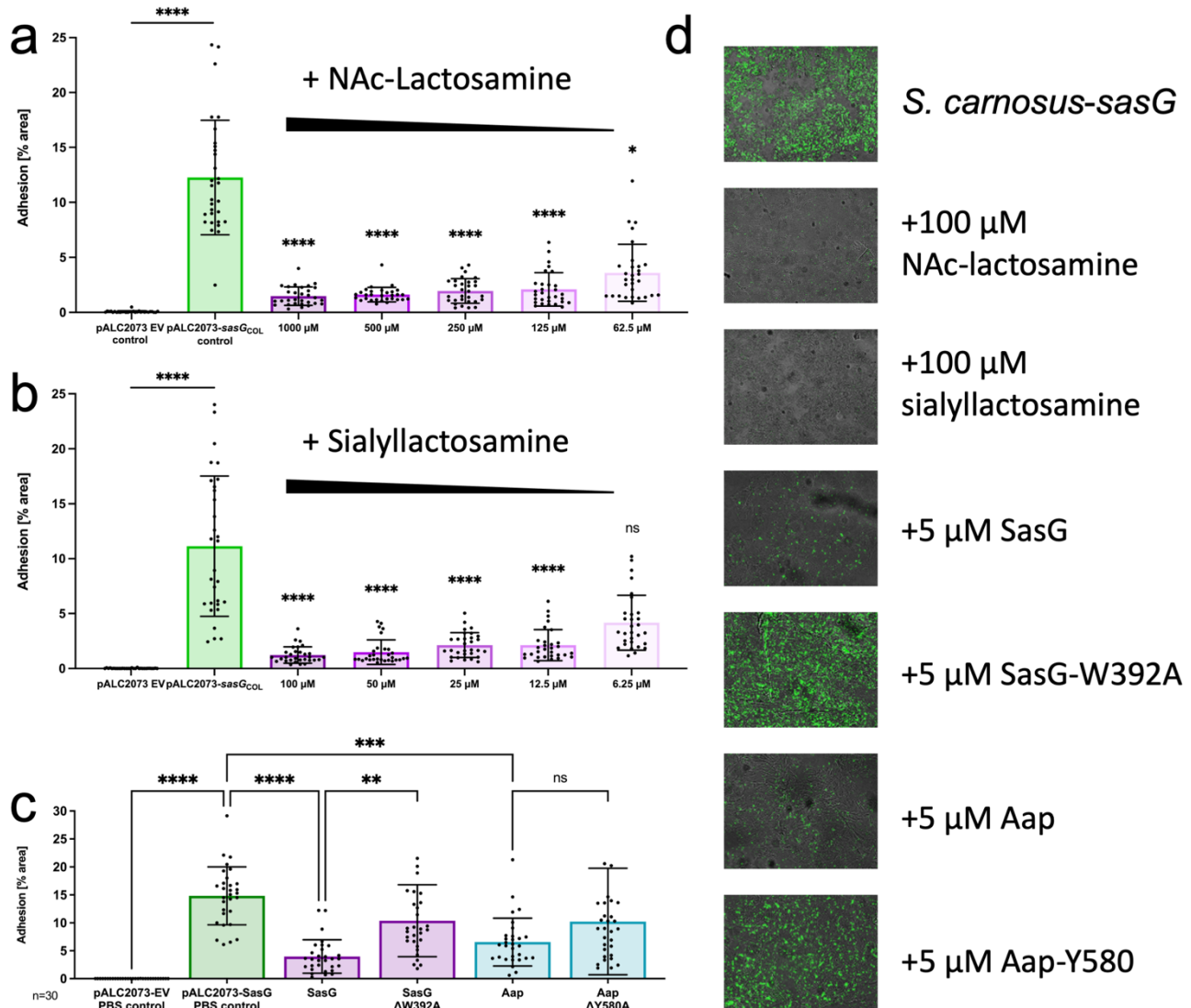
**Fig. 6: Corneocyte adhesion by *S. epidermidis* is mediated by Aap-glycan interactions.** (a) Adhesion of *S. epidermidis* 1457 Δica compared to the negative control *S. epidermidis* 1457 ΔaapΔica, both stably expressing GFP. The Δica background removes production of PNAG (PIA), a polysaccharide involved in intercellular adhesion. N-acetyllactosamine inhibited *S. epidermidis* adhesion at all concentrations tested. (b) Sialyllactosamine inhibited adhesion of *S. epidermidis* at all concentrations except the lowest (6.25 μM). Note that sialyllactosamine was tested at 10-fold lower concentrations than N-acetyllactosamine. (c) *S. epidermidis* adhesion was inhibited by the addition of 5 μM soluble recombinant Aap lectin or recombinant SasG lectin, but inhibition was reduced significantly upon addition of Aap-Y580 or SasG-W392 mutant lectins. (d) Representative microscopy images used for the quantitative results.



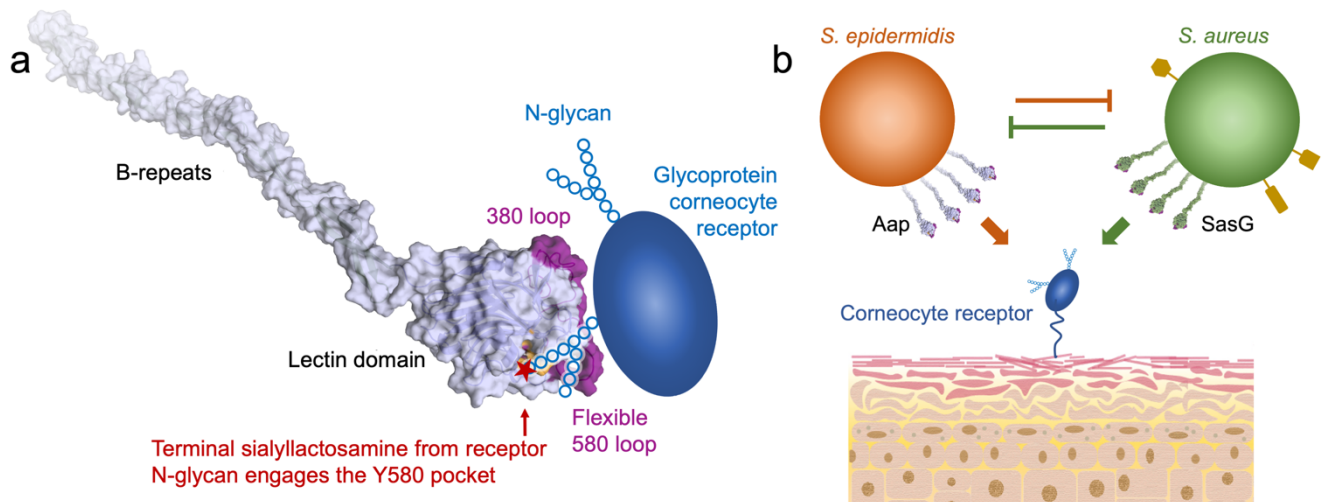
**Fig. 7: Nanoscale analysis of Aap binding to human corneocytes.** (a) High resolution topography (left) and corresponding adhesion (right) images of human corneocytes recorded in PBS, in the quantitative imaging mode, using (from top to bottom) single *S. epidermidis*  $\Delta sepA\Delta ica$ ,  $\Delta aap\Delta ica$  and  $\Delta sarA\Delta ica$  probes. *S. epidermidis* 1457  $\Delta sepA\Delta ica$  lacks the protease SepA that cleaves the A domain from Aap, and the  $\Delta ica$  background removes PNAG production; *S. epidermidis* 1457  $\Delta sarA\Delta ica$  lacks the SarA global transcriptional regulator that suppresses SepA expression, resulting in SepA overexpression and thus complete proteolytic removal of the A domain of Aap. The data shown are representative from a total of 2 different cell-corneocyte pairs for each strain, including independent cultures and preparations. (b, c) Histograms of adhesion forces (b) and rupture lengths (c) of the *S. epidermidis* 1457  $\Delta sepA\Delta ica$  - corneocyte interaction. Histograms present data merged from four different cell-pairs probed in the force volume mode. Maxima were fitted by a Gaussian model to extract mean values ( $\pm$  s.d.). (d) Representative adhesive curves documenting rupture of single, double, and triple bonds. The data in b-d are representative from a total of 6 different cell-corneocyte pairs, including independent cultures and preparations. (e) Box plots showing the binding probability (left) and unfolding frequency (right) of the *S. epidermidis* 1457  $\Delta sepA\Delta ica$  ( $n = 6$ ), *S. epidermidis* 1457  $\Delta sarA\Delta ica$  ( $n = 5$ ) and *S. epidermidis* 1457  $\Delta aap\Delta ica$  ( $n = 5$ ). Stars are the mean values, lines the medians, boxes the 25-75% quartiles and whiskers the s.d.



**Fig. 8: SasG lectin adopts a similar structure and binding specificity to Aap.** (a) Crystal structure of the *S. aureus* SasG lectin domain illustrated with a transparent surface, showing the loops that converge on the front face of the lectin to form the binding pocket. (b) Superposition of SasG and Aap lectin domains. Residues W392 (SasG) and Y580 (Aap) that form the floor of each binding pocket are shown as sticks. Note that SasG features the structural Ca<sup>2+</sup> ion but does not contain the structural Na<sup>+</sup> ion observed in Aap. (c) Details of the interaction network involving Ca<sup>2+</sup>, conserved aspartate D240 (which forms a *trans* peptide bond, similar to D431 in Aap), and W392, all of which stabilize the sharp bend in the main chain near W392 that creates the floor and upper wall of the binding pocket. (d) ITC data for SasG binding to N-acetyllactosamine; the binding was weak and could not be satisfactorily fitted to a binding model. (e) ITC data for SasG-W392A mutant binding to N-acetyllactosamine.



**Fig. 9: Staphylococcal adhesion to corneocytes mediated by SasG.** (a) Adhesion of *S. carnosus* engineered to express SasG, *S. carnosus*/pALC2073-*sasG*<sub>COL</sub> (*S. carnosus-sasG*), compared to the empty-vector negative control. N-acetyllactosamine inhibited *S. carnosus-sasG* adhesion at all concentrations tested. (b) Sialyllactosamine inhibited adhesion of *S. carnosus-sasG* at all concentrations except the lowest (6.25 μM). Note that sialyllactosamine was tested at 10-fold lower concentrations than N-acetyllactosamine. (c) *S. carnosus-sasG* adhesion was inhibited by the addition of 5 μM soluble recombinant SasG lectin but inhibition was reduced significantly upon addition of SasG-W392 mutant lectin. Soluble Aap lectin also inhibited adhesion; addition of Aap-Y580 mutant showed a trend toward reduced inhibition that did not reach significance. (d) Representative microscopy images used for the quantitative results.



**Fig. 10: Model for corneocyte adhesion mediated by Aap and SasG.** (a) Model for the interaction of Aap with a corneocyte glycoprotein receptor. The best-fit structural model of Aap lectin-Brpt from the SAXS EOM analysis is illustrated with a semi-transparent surface. The Y580 binding pocket is colored orange, and the divergent long loops at the top of the lectin domain are colored purple. The cartoon of the glycoprotein receptor illustrates the proposed bipartite binding mechanism in which the divergent long loops (purple) of Aap recognize the protein moiety of the receptor, and a receptor N-glycan containing terminal sialyllactosamine interacts with the Y580 binding pocket. (b) Schematic illustrating competition between *S. epidermidis* and *S. aureus* for corneocyte adhesion and skin colonization. Both Aap and SasG recognize the same (or highly similar) corneocyte receptor(s). Structures illustrated in yellow on *S. aureus* represent additional surface adhesins (e.g., SdrD, FnBP-B, or ClfB) with specificity for corneocyte or keratinocyte receptors that could confer an advantage to *S. aureus* under appropriate conditions.