PNAS

Supporting Information for

Human oral lectin ZG16B acts as a cell wall polysaccharide probe to decode the host-microbe interactions with oral commensals

Soumi Ghosh[1], Christian P. Ahearn[2], Christine R. Isabella[3], Victoria M. Marando[3], Gregory. J. Dodge[1], Helen Bartlett[1], Robert L. McPherson[3], Amanda Dugan[3], Shikha Jain[2], Lubov Neznanova[2], Hervé Tettelin[4], Rachel Putnik[5], Catherine L. Grimes[5], Stefan Ruhl[2], Laura L. Kiessling[3,*] & Barbara Imperiali[1,3,*]

Affiliations:

[1] Dept. of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

[2] Department of Oral Biology, University at Buffalo School of Dental Medicine, Buffalo, NY 14214, USA

[3]Dept. of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

[4]Department of Microbiology and Immunology, Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD 21201, USA

[5]Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, USA

This PDF file includes:

Materials and Methods

Cell line and strains: BL21(DE3) competent cells were obtained from Agilent. *S. vestibularis* strain NCTC 12166 was purchased from ATCC (49124). The *S. mutans* strains UA159, S8BE3 and SJ were kind gifts from Dr. Katharina Ribbeck, Dept. of Biological Engineering, MIT.

Plasmids: The codon-optimized constructs for BL21(DE3) of zymogen granule protein 16 homolog B (ZG16B) (GenBank accession AAH09722) and zymogen granule protein 16 homolog P (ZG16P) (GenBank accession AAC08708), from *H. sapiens* were obtained from IDT as synthetic gBlocks. The ZG16B construct, comprising the carbohydrate recognition domain (CRD) from residues 52-208 with a C-terminal sortase-reaction tag (LPETG) was cloned into sortagging vector pET24a vector within NdeI and XhoI sites with a C-terminal 6x Histidine tag. Residues 21-159 of ZG16P comprising the CRD and C-terminal LPETG linker and 6x Histidine tag were subcloned into pET24a between BglII and XhoI sites.

Expression and batch purification of ZG16B and ZG16P:
The ZG16B and ZG16P constructs were transformed into BL21(DE3) cells and a 5 mL seed culture, supplemented with 50 µg/mL Kanamycin was grown in Luria-Bertani medium at 37°C for 18 hours at 225 rpm for each lectin. The overnight seed cultures were inoculated in 500 mL autoinduction media (0.1% (w/v) tryptone, 0.05% (w/v) yeast extract, 2 mM $MgSO_4$, 0.05% (v/v) glycerol, 0.005% (w/v) glucose, 0.02% (w/v) α-lactose, 2.5 mM Na_2HPO_4, 2.5 mM KH_2PO_4, 5 mM NH_4Cl, 0.5 mM Na_2SO_4) in a baffled flask, supplemented with 30 µg/mL kanamycin. The 500 mL cultures were incubated at 37°C for 4-5 hours at 225 rpm until bacterial growth reached log phase (OD_{600} ~ 0.8-1). The incubation temperature was reduced to 18°C for the autoinduction of the protein expression for 18 hours at 225 rpm. The cells were harvested at 3000 rpm for 25 minutes at 4°C. The pellets were washed with 15 mL phosphate buffer saline and saved at -80°C.

The cell pellet (weight ~8.0 g for ZG16B and ZG16P) was resuspended in ~30 mL of lysis buffer (50 mM HEPES; pH 7.5, 300 mM NaCl, 20 mM imidazole, 10% glycerol, 5 mg lysozyme, 2 mg DNaseI, 2 mM $MgCl_2$) by rotating for 1 h at 4 °C. The cells were homogenized and lysed by sonicating twice for 90 s (1s ON and 2s OFF cycle) at 50% power output. The lysed cells were centrifuged at 30,000 RPM at 4°C for 60 min inTi45 rotor. The supernatant was collected, filtered through a 0.22 µm PES membrane filter and loaded to 3 mL Ni-NTA resin, pre-equilibrated with 10 CV buffer A (50 mM HEPES; pH 7.5, 300 mM NaCl, 5 mM $CaCl_2$, 20 mM Imidazole, 10% glycerol). The supernatant was batch bound for 90 min at 4°C with gentle rotation. The unbound protein was collected by flowing through the protein-resin complex in a gravity column; the supernatant being flowed through thrice to ensure maximum binding of the lectin to the Ni-NTA beads. The 3 mL resin was washed with buffer A, containing 20, 50, 100, 200, and 400 mM of imidazole, 15 mL in each wash. The wash fractions were checked on 4-20% polyacrylamide gel and those containing ZG16B or ZG16P were pooled, and flash-frozen at -80 °C.

Desalting of ZG16B and ZG16P:
Fractions containing Ni-NTA purified ZG16B or ZG16P were thawed on ice. Meanwhile, the 3x 5 mL prepacked Hitrap desalting columns with Sephadex G-25 resin were equilibrated with 3CV of dialysis buffer (25 mM HEPES, pH 7.5, 150 mM NaCl, 10% glycerol), attaching to the AKTA purifier FPLC. The purified lectins are injected into the FPLC using a 5 mL loop and buffer exchanged at the flow rate of 1 mL/min, washing the columns with 4 CV of dialysis buffer. The desalted and buffer exchanged fractions were collected, checked on a 4-20% polyacrylamide gel, and stored at -80°C.

Gravity flow-based sortase-mediated ligation:
The gravity flow method of sortase-mediated ligation (SML) was developed based on the previously published syringe-pump flow method with minor modifications (1). Briefly, 40 µg of 6x histidine-tagged sortase A protein was incubated with 200 µL Ni-NTA resin, pre-equilibrated with sortase reaction buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$), on ice for 10 mins in 600 µL sortase reaction buffer. The sortase-bound resin was poured in a 2 mL regeneration gravity column and excess sortase and buffer were allowed to flow through. A 3 mL reaction mixture of lectin (final concentration 5 µM) and GGGYK-biotin- peptide/GGYK-SulfoCy5-peptide (21st Century Biochemicals) (final concentration 7 µM) was added to the column and the fractions were collected. The column was washed with an excess 3 mL of the peptide (7 µM), followed by a 2 mL sortase buffer wash. Finally, the unreacted lectin was eluted from the column using a buffer containing imidazole (50 mM HEPES; pH 7.5, 300 mM NaCl, 400 mM Imidazole, 10% glycerol). The fractions with biotin-conjugated or fluorophore-conjugated lectin were dialyzed against 25 mM HEPES, pH 7.5, 150 mM NaCl, 10% glycerol, using a 10 kDa MWCO snake-skin dialysis bag. The fractions were saved at -80 °C.
For SML of ZG16P, 160 µg sortase A protein was used to increase the yield of the SulfoCy5-conjugated lectin.

Culture and isolation of bacteria from mixed-species oral biofilms:
Oral biofilm bacteria collected from volunteers were plated onto blood agar, Columbia agar (BD) with 8% v/v defibrinated sheep's red blood cells (Thermo Scientific), then incubated for 48 h at 37 °C under anaerobic conditions using an anaerobic jar and a BD GasPak™ EZ. Select colonies, chosen because they were bound by ZG16B, from the original culture plate were picked, streaked for isolation onto fresh blood agar plates, and incubated as before. Based on colony morphology, colonies were sub-cultured onto fresh blood agar plates until a uniform morphology was observed. To make freezer stocks of oral biofilm isolates, colonies with identical morphology were scraped from the blood agar surface then suspended in sterile 20% glycerol (v/v) (Fisher Scientific) in Brain Heart Infusion (BHI) (BD) broth and maintained at -80°C. For subsequent experiments, frozen stocks of oral biofilm isolates were either cultured on blood agar plates or directly seeded into Brain Heart Infusion (BHI) broth then grown statically for 18 h at 37 °C under anaerobic conditions.

Collection of human oral biofilms:
This study was approved by the University at Buffalo Human Subjects IRB board (study # 030–505616). Informed consent was obtained from all human participants. Oral biofilms were collected as described previously with minor modifications (2). Early oral biofilms, formed for approximately 12 h, were obtained from volunteers who were instructed to clean their teeth the night before collection then refrain from eating, drinking, and brushing their teeth before collection of oral biofilms the next morning. Early dental plaque biofilms were collected by rubbing the surfaces of the teeth in the four quadrants of the dentition with sterile cotton swabs, and by scraping the interproximal region of the molars with a sterilized toothpick while avoiding touching the gingival tissue. Tongue biofilms were collected by rubbing a sterile cotton swab over the dorsum of the tongue. Oral biofilm bacteria from dental plaque or the dorsum of the tongue were each suspended by swirling the swabs and toothpicks in respective 1 mL volumes of sterile phosphate-buffered saline (PBS) (145 mM NaCl, 6.4 mM NaH_2PO_4, 13.5 mM Na_2HPO_4, pH 7.2). The resulting suspensions were vortexed to mix and serial 10-fold dilutions made in PBS were plated onto blood agar and grown as described above.

Screening oral biofilms for bacteria bound by ZG16B-Cy5:
Oral biofilm bacteria bound by ZG16B-Cy5 were identified using a previously described technique (2) that combines the transfer of cultured oral bacteria to a nitrocellulose membrane followed by far-western blotting. Briefly, a round Immobilon®- nitrocellulose membrane 0.45 µM pore size (Millipore-Sigma) was placed onto the agar surface of a mixed oral biofilm culture to transfer bacteria from colonies onto the nitrocellulose membrane. Membranes were then peeled off of the agar surface and washed with Tris-buffered saline containing calcium and magnesium (TBS) (20 mM Tris-HCl, 150 mM NaCl, 1mM $CaCl_2$, 1 mM $MgCl_2$, pH 7.6) to remove visible residues of transferred colonies. Membranes were blocked with blocking buffer, 2% (w/v) polyvinyl alcohol (Millipore-Sigma) in TBS containing 0.05% v/v Tween-20 (TBST) for 1 h statically at room temperature. Blocked membranes were then probed with 200 nM ZG16B-Cy5 in blocking buffer for 1 h statically at room temperature. After membranes were washed thrice with TBST followed by one wash with TBS, they were air-dried, and then scanned using a ChemiDoc™ MP (Bio-Rad) image scanner to detect Cy5 fluorescence. Membranes were then counterstained with Ponceau S stain (0.5% Ponceau S (Sigma) in 1% acetic acid) and then de-stained with ddH_2O to show the location of all bacterial colonies that transferred to the nitrocellulose membrane. Colonies that showed positive ZG16B-Cy5-binding, based on Cy5 fluorescence signal, were picked from the original mixed oral biofilm culture plates, isolated, then used for subsequent lectin binding experiments. In addition to positiveZG16B-Cy5-bound isolates, non-bound isolates that were unlabeled by ZG16B-Cy5 were selected from each volunteer's oral biofilm cultures to serve as negative controls.

ZG16B-Cy5 binding to the oral biofilm isolates was retested. Isolates were grown on blood agar plates, as described before, colonies were then scraped off the plates and the bacteria were washed thrice by centrifugation for 6 min at 4,180 x

rcf, decanting the supernatant, then the pellet was resuspended in PBS. The optical density of the washed isolates was read at 600 nm (OD_{600}), adjusted to OD600 =1 using PBS, then serial 2-fold diluted in PBS before spotting a 1 µL portion of each dilution onto an Immobilon® nitrocellulose membrane followed by far-western blotting with 200 nM of either ZG16B-Cy5 or ZG16P-Cy5.

16S rRNA and whole genome shotgun sequencing:
For 16S rRNA sequencing, dental plaque isolates scraped from glycerol stocks were resuspended and lysed in a buffer with 25 mM NaOH and 2 mM EDTA with repeated cycles of heating at 95 C for 10 min and cooling to 4C for 15 min. The extracted templates were amplified at the V3-V4 region using 16s rRNA sequencing primers (forward primer: AGAGTTTGATCCTGGCTCAG, revers primer: ACGGCTACCTTGTTACGACTT).

For whole genome shotgun sequencing, DNeasy UltraClean microbial kit (Qiagen) was used for DNA extraction. DNA samples were cleaned using a 0.8X SPRI cleanup (Aline Biosciecnes) and purification using a NucleoSpin column (Takara). Illumina libraries were prepared using Nextera XT at 1/5th reaction volume (Illumina) using 14 cycles of PCR. Libraries were sequenced on an Illumina MiSeq using 300nt paired end reads. Sequencing reads were assembled *de novo* by using the latest ABySS 2.0 (3), which was evaluated with the best performance among several similar choices. To achieve the optimal performance particularly on isolates 2, 3, 4, 6, 10, and 11), Pear (4) was used to pre-assemble paired-end reads into merged ones that were submitted to AbySS for further assembly. Assembled contigs greater than 10k in length were submitted to NCBI blast website at https://blast.ncbi.nlm.nih.gov/Blast.cgi (5). A rank-ordered list of target strains was generated based on the sum of the Max-Scores (the highest alignment score calculated from the sum of the rewards for matched nucleotides or amino acids and penalties for mismatches and gaps) from the list of candidate strains returned from the NCBI blast search.
The data from whole genome sequencing results was submitted under the BioProject number PRJNA880893.

Bacterial isolate species and strain identification based on whole genome comparisons:
Contigs ≥500 bases from the assembly of whole genome shotgun sequencing data of newly identified oral biofilm bacterial isolates were used for whole genome sequence-based comparisons to previously sequenced streptococcal or *Gemella* isolates to resolve bacterial species and strain identity. Isolate 13 was excluded from the analysis due to an error in the FASTA file. We downloaded all 34 publicly available (as of November 2021) assembled *Gemella* genomes, >6,000 publicly available *Streptococcu*s species genome assemblies (the many thousands of *S. pneumoniae* assemblies were subsampled using GGrasp (6) to a subset of 258 genomes representing the genomic diversity of the pneumococcal population). We used Mash v2.3 to generate two whole genome pairwise k-mer composition-based distance matrices (7). The matrices were analyzed with MEGA-CC v10.0.5 (8) to generate neighbor-joining phylogenetic trees, one for *Streptococcus* and one for

Gemella. The Newick tree files were uploaded and annotated in iTOL (9). The streptococcal tree was pruned to display clades of streptococcal species that commonly associate with the oral cavity. The tree was further pruned to present species subclades containing the identified ZG16B-bound isolates.

Evaluation of bacterially purified ZG16B binding to the dental plaque isolates using flow cytometry:

Fourteen strains of dental plaque isolate, 12 positive binders, and two negative binders of ZG16B were grown overnight in 5 mL BHI media at 37 ˚C, without shaking. The cells were diluted to OD_{600} 0.2 in PBS. A cell sample (50 µL) was aliquoted into 96-well plates and centrifuged at 4 ˚C at 3280 x g for 10 minutes. Meanwhile, the staining solution was prepared by adding staining dye to the lectins. Streptavidin-Cy5 (AAT Bioquest) was used at 1:250 dilution and SYTO BC was used at 1:1000. All lectins, including bacterially produced Cy5-conjugated ZG16B, were used at 15 µg/mL. The buffer used contained 20 mM HEPES, pH 7.5, 150 mM NaCl, 0.1% Tween-20, 0.1% BSA.

After centrifuging the cells in the 96-well plate, the pellet was resuspended in 50 µL staining solution and incubated for 2 h at 4 ˚C. The cell-lectin complex solution was diluted to 1:3 ratio using appropriate buffers. Data were collected at the flow cytometer using 555 nm and 647 nm laser by drawing 50 µL sample from each well, gating on 10,000 events for bacteria from ssc/fsc and/or 5 minutes of total time per well. The data were analyzed using FlowJo software.

The *S. mutans* strains were grown overnight in 5 mL BHI media at 37 ˚C in 5% carbon dioxide. The cells were stained using the protocol mentioned above and subjected to flow cytometry analysis.

Confocal microscopy of the dental plaque microbes in the presence of ZG16B and MUC7 rich saliva sample:

The dental plaque bacteria were grown and fixed in 4% formaldehyde at 4˚C. The cells at OD_{600} 0.2 were stained with 100 µL staining solution containing SYTO BC (1:1000), ZG16B-Cy5 (15 µg/mL), MUC7 or MUC5B enriched sample (25 µg/mL) and appropriate primary and secondary antibodies for the mucins. MUC7(MyBiosource Inc.) or MUC5B (Sigma-Aldrich) rabbit polyclonal antibody were used at 1:250 dilution, Goat-anti rabbit Alexafluor 555 (Abcam) was used at 1:250 dilution. The bacteria with the staining solution were incubated for 1 h at 4 ºC then plated on the poly-L-lysine coated Mattek glass 24-well plate and incubated overnight at 4 ºC. The cells were imaged at 60x objective, 1.5x magnification with 488, 561, and 645 nm laser using a Confocal ANDOR microscope. The image was analyzed using ImageJ (Fiji) software for the determination of the pixel correlation in colocalization analysis and bacterial cluster size distribution.

Determination of MIC_{50} of ZG16B and brightfield imaging of live cells:

Dental plaque isolates were grown overnight in 5 mL BHI media at 37 ºC, without shaking, and diluted to OD_{600} 0.4 at a final concentration in the same media. Using a 96-well Nunc round-bottom plate, 100 µL of BHI was added to each well. 100 µL

recombinant ZG16B or ZG16P in 25 mM HEPES, pH 7.5, 150 mM NaCl, 10% glycerol, 0.1% BSA at 40 µg/mL was added to the third well, mixed with the medium, and then subjected to a two-fold serial dilution in BHI media to a final concentration of 0.02 µg/mL. The first well corresponded to medium sterility control and the second well corresponded to 100% growth control of the bacteria. Next, all wells (except the first) were inoculated with 100 µL of the bacterial culture. The assay plate was incubated at 37 °C for 12 h and the optical density at 600 nm (OD_{600}) was measured using a microplate reader (Biotek Instruments Inc., VT, USA) every 1 h. The minimal inhibitory concentration (MIC_{50}) was determined as the lowest ZG16B concentration able to promote a reduction of OD_{600} higher or equal to 50%, in comparison with the OD_{600} in no lectin control. For the assay with MUC7-rich saliva sample, the lectin and the mucin were diluted to a final concentration of 1.25 µg/mL and 15 µg/mL in BHI media, respectively, before the addition of the bacteria. Each assay was achieved in three independent experiments.

Lectin Absorption Assay:
The lyophilized peptidoglycan layers extracted from either *S. vestibularis* or *S.oralis* were resuspended in 100 µL resuspension buffer (20 mM Tris-HCl, pH 7.0, 150 mM NaCl, protease inhibitor cocktail mix (EMD Millipore)) at a final concentration of ~1 µg/µL. Then, 2 µg recombinant ZG16B was introduced to the resuspended peptidoglycan layer and the mixture was subjected to gentle rotation for 30 minutes at 4 °C. The peptidoglycan layer was centrifuged at 1200 x g for 5 minutes and washed three times with 200 µL resuspension buffer. The lectins absorbed in the peptidoglycan layer were extracted by boiling with SDS Laemmli dye and quantified from the immunoblot using anti-His_6 mouse monoclonal antibody for ZG16B.

Super-resolution microscopy:
The cells were grown overnight in 5 mL BHI medium with 200 µM 3-[[(7-Hydroxy-2-oxo-2*H*-1-benzopyran-3-yl)carbonyl]amino]-D-alanine hydrochloride (HADA) (R&D Systems) at 37°C, without shaking. The cells were fixed and stained with ZG16B-Cy5 following the same protocol used in the confocal microscopy of the dental plaque microbes in the presence of ZG16B and MUC7 rich saliva sample. For analysis by microscopy, stained cells were pelleted and suspended in PBS supplemented with 1% w/v DABCO (1,4-diazabicyclo[2.2.2]octane) (Sigma-Aldrich) as an antifading reagent. Each sample was spotted onto a glass-bottomed microwell dish (MatTek corporation # P35G-1.5-14-C), allowed to settle overnight, and covered with a pre-cooled 1% (w/v) agarose pad. Images were collected on an Applied Precision DeltaVision-OMXv4 Super-Resolution Microscope (60x/1.42 NA oil immersion lens, sequential imaging on two sCMOS cameras). Structured illumination microscopy (SIM) reconstruction and warp-based image alignments were performed with Applied Precision softWoRx. Multi-color Image alignments were calculated from an Applied Precision grid test slide, checked with TetraSpeck beads (0.1 µm, Molecular Probes), and verified with an Argo-SIM test slide (Axiom Optics) before data collection. Brightness and contrast were identically adjusted

with the open-source Fiji distribution of ImageJ. Images were then converted to an RGB format to preserve normalization and then assembled into panels.

MUC7 enrichment from human saliva:
MUC7 enriched from human saliva was a kind gift from Dr. Molakala S. Reddy, and was purified using the methods detailed in Ramasubbu, N. *et al*. (10). Prior to use, a portion of a -80 °C frozen sample of MUC7 was thawed, then resolved on an SDS-PAGE gel (Novex 8-16% Tris-Glycine) to check MUC7 protein integrity. Samples with a discrete band that migrated to the expected migration position of MUC7, ~150 kDa, in the gel were used.

Lectin blotting on salivary mucins:
15 µg of saliva samples, collected from submandibular/sublingual glands from two donors (volunteer 1 and 2), approved by the University at Buffalo Human Subjects IRB board (study # 030–505616)), were subjected to analysis on a 4-20% polyacrylamide gel (Bio-Rad) at 150 V for 50 minutes. The separated products were transferred onto a nitrocellulose membrane at 100 V for 90 minutes using Towbin buffer. The membrane was blocked with PVA blocking buffer (2% polyvinyl alcohol in TBS, 0.1% Tween-20, 1 mM $CaCl_2$, and 1 mM $MgCl_2$) for 1 h at room temperature and then incubated with 250 nM lectins conjugated with Cy5 or biotin in 1mL blocking buffer overnight at 4 °C in a sealed plastic bag. For the pre-complexed lectins, 250 nM biotinylated lectins were incubated with 62.5 nM streptavidin conjugated to Cy5 (AAT Bioquest) or streptavidin conjugated to alkaline phosphatase (Promega). The samples were kept in 1 mL blocking buffer on ice for 1 h, added to the blot, and then incubated overnight at 4 °C in a sealed plastic bag. Finally, the blots were washed twice with TBS/0.05% Tween-20 at room temperature, each for 10 minutes, and with TBS for 10 minutes. The blots were imaged at ChemiDoc Touch Bio-Rad imager using Cy5 mode or colorimetric mode.

Extraction and fractionation of cell wall capsule from the dental plaque isolates:
The extraction and fractionation of cell wall glycoconjugates from the dental plaque isolates were done following the previously reported protocol with a minor changes (11). Briefly, a 5 mL cell culture of *S. vestibularis* or *S. oralis* was grown overnight at 37 ºC in BHI medium, without shaking. The cells were pelleted and flash-frozen in liquid nitrogen. The cell pellet was resuspended in 0.2 mL 2M NaCl and disrupted by three freeze-thaw cycles. The walls were pelleted by centrifugation (4°C, 14,000 x g, 15 min). The pellet was resuspended in 0.1 mL PBS and collected by centrifugation (4°C, 14,000 x g, 5 min). Noncovalently bound components were removed by resuspending the pellet in 0.2 mL 4% SDS and incubating with gentle stirring at 37 ˚C for 5 hours. The walls were recovered by centrifugation (25 °C, 14,000 x g, 5 min). The pellet was washed three times with 5 mL water. To remove covalently bound protein and nucleic acids, the pellet was incubated at 37 ˚C with gentle stirring with 0.1 mL proteinase K (0.2 mg/mL), RNAse (0.1 mg/mL), DNAse (0.1 mg/mL) (Life Technologies) in 50 mM Tris-HCl pH 7.0, 5 mM $MgCl_2$, 50 µM

$CaCl_2$ for 14 hours. The sample was subjected to centrifugation (4 °C, 14,000 x g, 5 min). The walls were washed with 0.2 mL 1 M Tris-HCl pH 7, 0.2 mL 1 M Tris-HCl, 1 M NaCl, 0.2 mL 1 M Tris-HCl (pH 7), and finally washed three times with 0.2 mL of water. After each wash, the pellet was recovered by centrifugation. To extract the wall teichoic acids, the pellet was resuspended in 0.1 mL ice-cold 10% trichloroacetic acid (TCA) and 4 °C with gentle stirring for 48 hours. The peptidoglycan was removed by centrifugation (21 °C, 5,300 x g, 45 min). The wall teichoic acids were precipitated from the soluble portion with 0.5 mL of 95% ice-cold ethanol at 4 °C for 4 days. The precipitate was pelleted by centrifugation (21 °C, 5,000 x g, 30 min) and then washed with 100 µL of 95% ethanol. The pellet was resuspended in 0.1 mL 95% ethanol and the solvent was evaporated using a speed vac.

The fractions were boiled with 6x SDS Lammeli loading buffer for 10 mins at 100 °C and loaded on a 4-20% polyacrylamide gel. Post gel electrophoresis, the gel was stained with 0.1% w/v Alcian blue stain (Sigma-Aldrich) in 40% ethanol, 60% 20 mM sodium acetate, at pH 4.75.

For the extraction of lipotechoic acid from the dental plaque isolates, cells grown under similar conditions were pelleted and washed once with 2 mL of resuspension buffer (50 mM sodium citrate [pH 4.7]). The cells were disrupted by freeze-thaw cycles, and the cell envelope fragments were pelleted (20,000 x *g*, 4 °C, 1 h). Pellets were suspended in 350 µl of resuspension buffer, and an equal volume of 1-butanol was added. The extraction was carried out on a shaking heat block at 37 °C for 45 min. The crude extract was centrifuged (20,000 x *g*, 4 °C, 1 h) to pellet insoluble materials, and the lower aqueous phase was then extracted into a clean tube and lyophilized overnight.

For analysis using polyacrylamide gel electrophoresis, lipotechoic acid was first enzymatically deacylated with lipase. For the butanol aqueous-phase crude extract, 50 µl of the aqueous layer was combined with 50 µl of 50 mM Tris (free base), and the solution was adjusted to pH 8.5 with 2 µl of 1 M NaOH. The deacylation reaction was carried out by adding 1 µl of Resinase HT (Sterm Chemicals), and the reaction mixture was incubated at 50 °C for 16 h. Lipase-treated samples were then analyzed by PAGE and stained as described above.

Peptidoglycan digestion and turbidity assay:
The lyophilized peptidoglycan layer extracted from *S. vestibularis* or *S. oralis* was resuspended in 200 µL resuspension buffer (20 mM Tris-HCl, pH 7.0, 150 mM NaCl) and incubated with mutanolysin from *Streptomyces globisporous* or chicken egg white lysozyme (Millipore Sigma) at 10 µg/mL for 24 h at 37 °C. The change in turbidity of the peptidoglycan was recorded by measuring OD_{600}.

High-resolution mass spectrometry of soluble peptidoglycan fragments
The soluble peptidoglycan fragments from *S vestibularis* (Isolate 3, 4 8, 12). *G. haemolysans* (Isolate 10), *S. oralis* (Isolate 5), and *S. mutans* (UA159) were processed using a 3 kDa MWCO spin filter to separate the soluble fragments from the non-soluble PG and lyophilized. An LC/MS method previously established was

utilized for analyzing the highly concentrated solution of isolated PG fragments (12). The LC method involved a 0.5 mL min^{-1} linear gradient starting from 0% A (0.1% formic acid in water) to 50% B (0.1% formic acid in acetonitrile) in 4 min. The eluting peaks were subjected to high-resolution mass analysis on the Q-Exactive Orbitrap (Acquity UPLC BEH C18 column 2.1 × 50 mm (Waters) using a Dionex UHPLC coupled to a Q-Exactive Orbitrap (Thermo Fisher Scientific)). Thermo Xcalibur Qual Browser was used to process and analyze the data generated. All species showed the expected mass with correct m/z ratio within ±10 ppm.

Dot-blot and biolayer interferometry:
A 2 µL volume of cell wall capsules extracted from the dental plaque isolates was dotted on a nitrocellulose membrane at no dilution, 1:1 dilution, and 1:2 dilution in 20 mM Tris, pH 7.5, 150 mM NaCl. The membranes were air-dried and baked at 90°C for 1 h. The membranes were blocked with polyvinyl alcohol blocking buffer, used in the lectin blotting assay and incubated with biotinylated lectins at 500 nM in the same buffer overnight at 4 °C. The blots were washed thrice with Tris buffer saline with 0.1% Tween-20 (TBST), incubated with alkaline-phosphatase conjugated Streptavidin at 1:10,000 dilution for 1 h, RT, and washed with TBST and TBS before the incubation with alkaline phosphatase substrate (Thermo Scientific).

For the biolayer interferometry experiment, cell wall capsules at 1:1 dilution in 250 uL of 20 mM Tris, pH 7.5, 150 mM NaCl, 0.1% BSA were placed in a 96-well plate. BLI sensors coated with super streptavidin (SSA sensors) were immersed for 10 minutes in the buffer solution before the loading of the biotinylated lectins. The binding kinetics were analyzed using ForteBio Octet.

Dynamic light scattering (DLS):
The hydrodynamic radii of the recombinant ZG16B or ZG16P resuspended in 25 mM HEPES, pH 7.5 150 mM NaCl, 10% glycerol at 100 µg/mL were determined by the DynaPro Nanostar detector at 25°C. The molecular weight of the lectin forms in the solution was determined from the hydrodynamic radius using DYNAMICS software.

Nano differential scanning fluorimetry (nDSF) of ZG16B and ZG16P:
The change in intrinsic tryptophan and tyrosine fluorescent intensity of recombinant ZG16B and ZG16P as a function of temperature were carried out in a Prometheus NT.48 instrument. Capillaries were filled with 10 µL recombinant lectins at ~5 µM concentration in 25 mM HEPES, pH 7.5 150 mM NaCl. The experiment was conducted using two replicates for each sample. The ratio of the emission intensities (Em_{350nm}/Em_{330nm}), representing the change in intrinsic fluorescence intensity, was recorded as the temperature was increased from 20 °C to 90 °C, taking one measurement per 0.0288°C. Both the raw fluorescence signal and the first derivative (slope) of the fluorescence signal as a function of the temperature were plotted. The melting temperature was calculated from the inflection point of the slope.

Screening of pre-complexed ZG16B on microbial glycan array:
A slide printed with microbial glycans (MGMv2 from Dr. Ryan McBride, Scripps Research Institute) was equilibrated at room temperature after having been warmed from -20 °C to room temperature for 20-30 minutes (13) The slide was placed on a slide holder tube and soaked in wash buffer 20 mM HEPES; pH 7.5, 150 mM NaCl, for 10 minutes. The slide was then placed on a raised slide holder platform in a humidification chamber with wet paper towels, and the printed area (no barcode side) was sealed by placing a rubber guard around it. Biotinylated ZG16B (800 µL) at a concentration of 10 µg/mL (either free or pre-complexed with streptavidin-Cy3) in protein binding buffer (20 mM HEPES; pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 0.1% BSA, 0.1% Tween 20) was added to the slide and incubated for one hour at room temperature. The slide was then washed in protein binding buffer, wash buffer and lastly double-deionized water. The slide was dried in a slide spinner and scanned at the Genepix4000 laser scanner using the 532 nm laser.

Protein Structure and Interaction (PSAIA) analysis of ZG proteins:
The coordinates and symmetry positions for ZG16B (PDB: 3AQG) and ZG16P (PDB: 3APA) were assessed for potential dimerization or higher order oligomerization using the PDBePISA (Proteins, Interfaces, Structures and Assemblies) server (14). No single crystallographic interface had a significantly greater buried surface area, and there were no common interfaces found between the two structures or with structural homologs. The lack of protein-protein interaction sites supports the experimental observation that the ZG proteins exist as monomeric proteins as had been determined by DLS and nDSF.

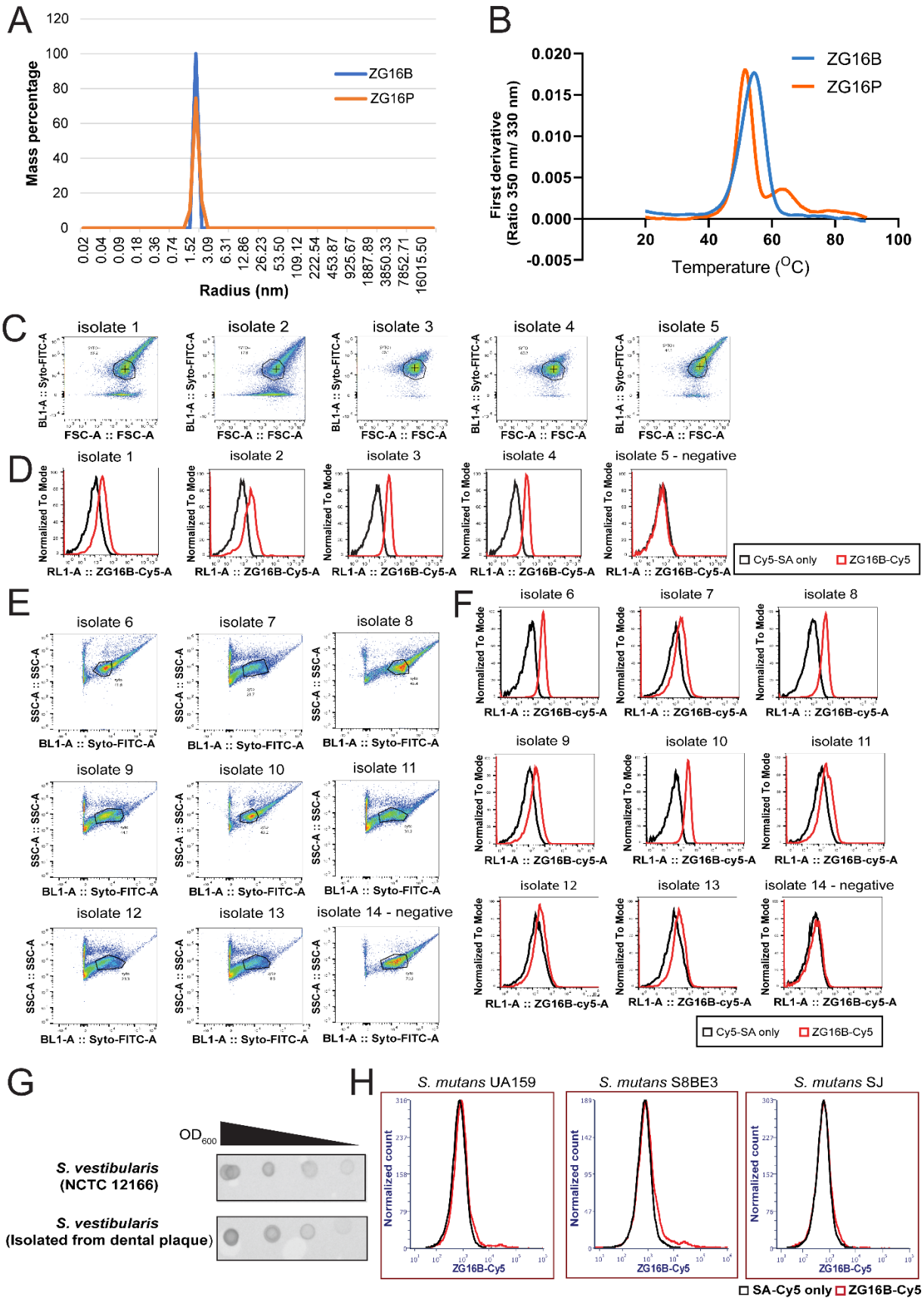

A
Mass percentage
ZG16B
ZG16P
Radius (nm)
B
First derivative (Ratio 350 nm/ 330 nm)
Temperature (°C)
C
isolate 1
isolate 2
isolate 3
isolate 4
isolate 5
BL1-A :: Syto-FITC-A
FSC-A :: FSC-A
D
isolate 5 - negative
Normalized To Mode
RL1-A :: ZG16B-Cy5-A
Cy5-SA only
ZG16B-Cy5
E
isolate 6
isolate 7
isolate 8
isolate 9
isolate 10
isolate 11
isolate 12
isolate 13
isolate 14 - negative
SSC-A :: SSC-A
BL1-A :: Syto-FITC-A
F
RL1-A :: ZG16B-cy5-A
G
OD600
S. vestibularis (NCTC 12166)
S. vestibularis (Isolated from dental plaque)
H
S. mutans UA159
S. mutans S8BE3
S. mutans SJ
Normalized count
ZG16B-Cy5
SA-Cy5 only

Fig. S1. **Recombinant ZG16B, functionalized with fluorophores binds to dental plaque bacteria. (**A) Dynamic light scattering data showing the dispersity of recombinant lectins in solution. The molecular weights of ZG16B and ZG16P were calculated as 15.5 and 18 kDa from the hydrodynamic radii, respectively. (B) The change in the intrinsic fluorescent intensity of recombinant ZG16B and ZG16P as a function of temperature obtained from differential scanning fluorimetry. The melting temperatures of ZG16B and ZG16P were determined to be 54.4 ºC and 51.6 ºC, respectively. Flow cytometry data is represented as (C) scatter plot and (D) histogram showing the binding of Cy5 conjugated ZG16B to the dental plaque microbes collected from donor 4. Flow cytometry data represented as (E) scatter plot and (F) histogram showing the binding of Cy5-conjugated ZG16B to the dental plaque microbes collected from (D) donor 3. (G) Dot-blot analysis of *S. vestibularis* strain NCTC 12166 and isolated from dental plaque spotted at different OD_{600} dilutions and probed with 200 nM ZG16B-biotin. (H) Flow cytometry data is represented as histogram showing the binding of ZG16B-Cy5 to different strains of *S. mutans*. Cy5 conjugated to streptavidin was used instead of ZG16B-Cy5 to show the background signal from the non-specific binding of the microbes to the SYTO BC dye.

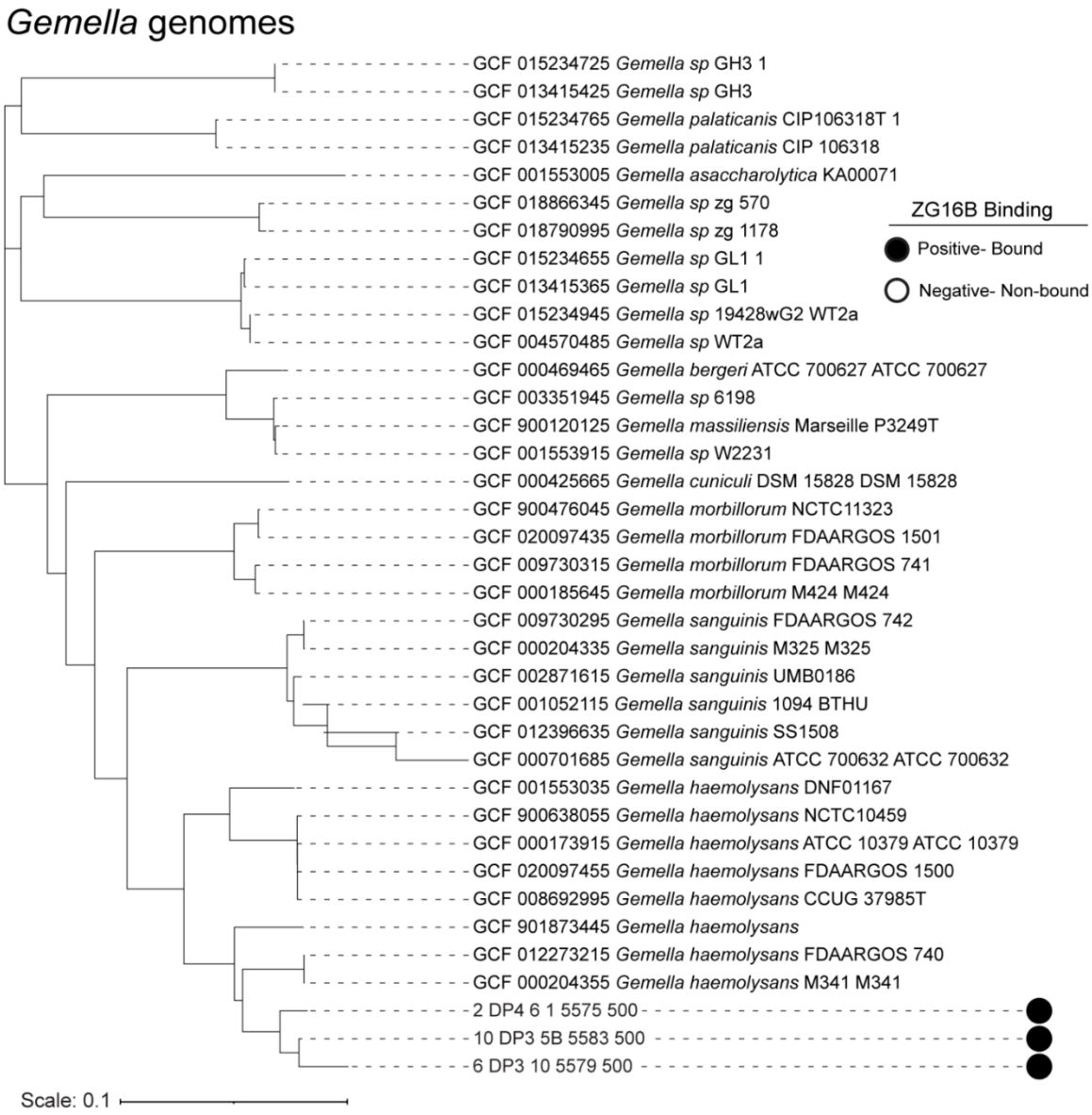


Fig. S2(A). Identification of ZG16B bound strain identity by *Gemella* genome sequence comparisons. The genome sequences of ZG16B bound bacterial isolates were compared to thirty-four publicly available (NCBI) *Gemella* genome sequences using kmer sequence identity matrix method for identification of oral biofilm isolate strain identity. The phylogenetic tree of *Gemella* genomes shows that each of the three of the ZG16B bound oral biofilm isolates were a separate strain located within the clade of the species *Gemella haemolysans*.

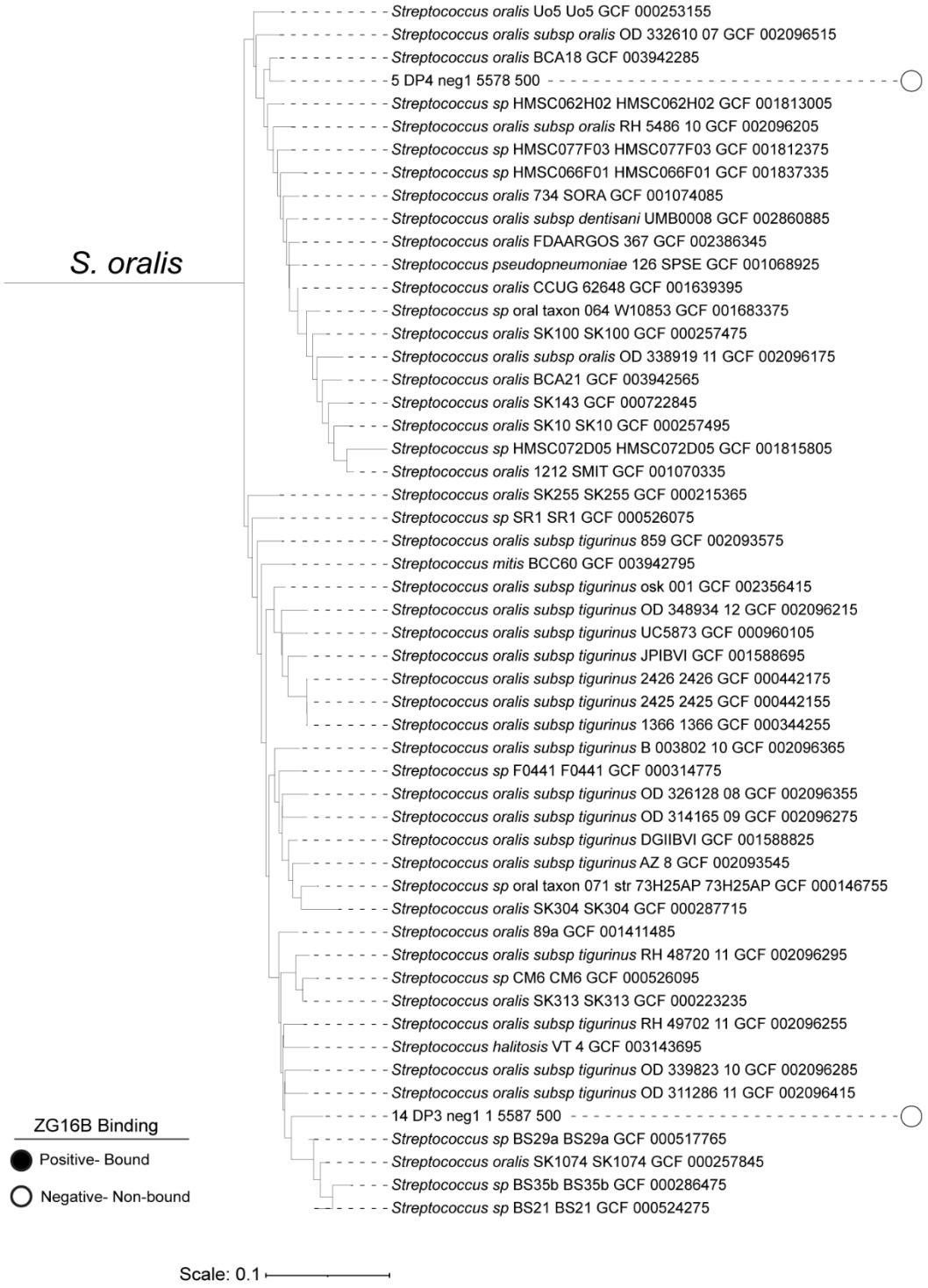


Fig. S2(B). The negative-non-ZG16B-bound oral biofilm isolates are *S. oralis* strains. A phylogenetic subtree of *S. oralis* genomes shows the negative-non-ZG16B-bound isolates (marked black-outlined-open circles) are strains of *S. oralis*.

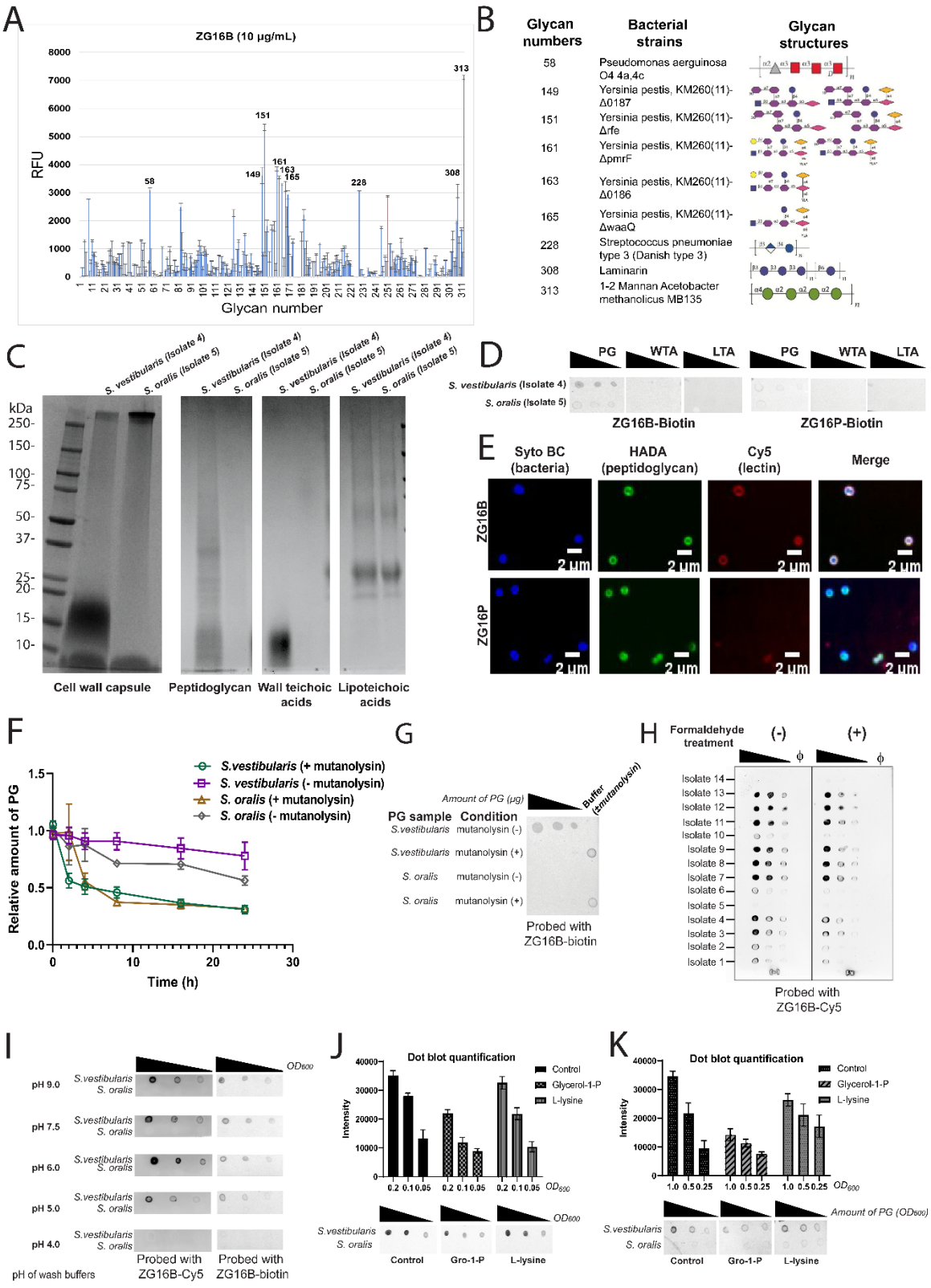

A
ZG16B (10 µg/mL)
RFU
Glycan number
B
Glycan numbers
Bacterial strains
Glycan structures
58 Pseudomonas aeruginosa O4 4a,4c
149 Yersinia pestis, KM260(11)-Δ0187
151 Yersinia pestis, KM260(11)-Δrfe
161 Yersinia pestis, KM260(11)-ΔpmrF
163 Yersinia pestis, KM260(11)-Δ0186
165 Yersinia pestis, KM260(11)-ΔwaaQ
228 Streptococcus pneumoniae type 3 (Danish type 3)
308 Laminarin
313 1-2 Mannan Acetobacter methanolicus MB135
C
S. vestibularis (Isolate 4)
S. oralis (Isolate 5)
kDa
Cell wall capsule
Peptidoglycan
Wall teichoic acids
Lipoteichoic acids
D
PG
WTA
LTA
S. vestibularis (Isolate 4)
S. oralis (Isolate 5)
ZG16B-Biotin
ZG16P-Biotin
E
Syto BC (bacteria)
HADA (peptidoglycan)
Cy5 (lectin)
Merge
ZG16B
ZG16P
2 µm
F
S.vestibularis (+ mutanolysin)
S. vestibularis (- mutanolysin)
S. oralis (+ mutanolysin)
S. oralis (- mutanolysin)
Relative amount of PG
Time (h)
G
Amount of PG (µg)
Buffer (-mutanolysin)
PG sample
Condition
S.vestibularis mutanolysin (-)
S.vestibularis mutanolysin (+)
S. oralis mutanolysin (-)
S. oralis mutanolysin (+)
Probed with ZG16B-biotin
H
Formaldehyde treatment
(-)
(+)
Isolate 14
Isolate 13
Isolate 12
Isolate 11
Isolate 10
Isolate 9
Isolate 8
Isolate 7
Isolate 6
Isolate 5
Isolate 4
Isolate 3
Isolate 2
Isolate 1
Probed with ZG16B-Cy5
I
pH 9.0
pH 7.5
pH 6.0
pH 5.0
pH 4.0
S.vestibularis
S. oralis
OD600
pH of wash buffers
Probed with ZG16B-Cy5
Probed with ZG16B-biotin
J
Dot blot quantification
Intensity
Control
Glycerol-1-P
L-lysine
Gro-1-P
K
Dot blot quantification
Amount of PG (OD600)

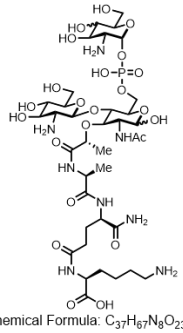
Fig. S3. ZG16B binds to the peptidoglycan layer of *S. vestibularis* (A) Bar-graph representation of microbial glycan array (MGM) screening data of biotinylated ZG16B, represented as the relative fluorescence integrated intensity from n=6 replicates from each glycan structure. (B) Details of selected microbial glycan structures on the microbial glycan array as top hits of ZG16B. Glycan structures taken from MGM array catalog V2. (C) Alcian blue staining of the cell wall capsule of ZG16B binder *S. vestibularis* and non-binder *S. oralis* and the concomitant fractionated cell wall glycoconjugates separated on SDS-PAGE. (D) Dot blot of the fractionated cell wall glycoconjugates probed with biotinylated lectins, showed weak interaction between the peptidoglycans and ZG16B. (E) Confocal microscopic imaging of *S. vestibularis* grown with -[[(7-Hydroxy-2-oxo-2*H*-1-benzopyran-3-yl)carbonyl]amino]-D-alanine hydrochloride HADA and incubated with Cy5 conjugated lectins showed peptidoglycan-mediated binding of ZG16B to *S. vestibularis*, but not for ZG16P. (F) Turbidity measurements at 600 nm during the digestion of peptidoglycan of *S. vestibularis* and *S. oralis* with 10 µg/mL chicken egg lysozyme and 10 µg/mL mutanolysin from *Streptomyces globisporous* at 37^0C for 24 h. (G) Dot blot of the peptidoglycan of *S. vestibularis* or *S. oralis.* Samples were undigested *or* digested with mutanolysin from *Streptomyces globisporous* and probed with biotinylated ZG16B. (H) Lectin blotting assay on 15% formaldehyde-treated dental plaque isolates with ZG16B-Cy5 to assess the effect of lectin binding to the microbes on the denaturation of cell surface proteins, as described previously (2). (I) Lectin blotting assay on *S. vestibularis* and *S. oralis* probed with ZG16B-Cy5 or ZG16B-biotin and washed with buffers at pH 9.0, 7.5, 6.0, 5.0, and 4.0. (J) Lectin blotting assay on *S. vestibularis* and *S. oralis* probed with ZG16B-biotin and washed with 20 mM glycerol-1-phosphate sodium salt or 20 mM L-lysine monohydrochloride salt in wash buffer (20 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl, 0.1% Tween-20). The binding intensity of ZG16B to the microbes at different OD_{600} was quantified from three technical replicates, as represented in the bar graph. (K) Dot blot on the insoluble peptidoglycan of *S. vestibularis* or *S. oralis* probed with ZG16B-biotin and washed with 20 mM glycerol-1-phosphate sodium salt or 20 mM L-lysine monohydrochloride salt in wash buffer (20 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl, 0.1% Tween-20). The binding intensity of ZG16B to the peptidoglycan at different concentrations was quantified from three technical replicates, as represented in the bar graph.

Structures

ESI (+)

ESI (-)

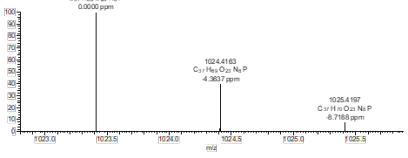
1

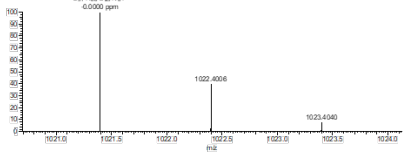


Chemical Formula: $C_{37}H_{67}N_8O_{23}P$
Exact Mass: 1022.4057

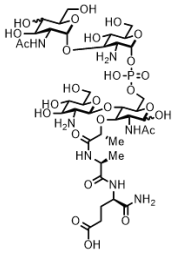
Observed

Simulated





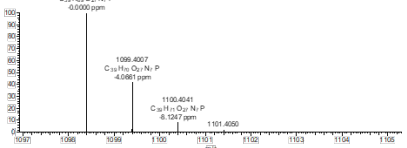
2

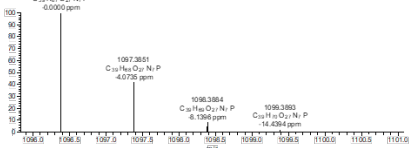


Chemical Formula: $C_{39}H_{68}N_7O_{27}P$
Exact Mass: 1097.3901

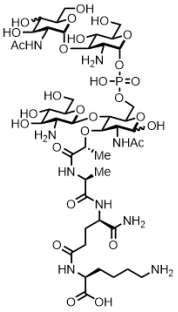
Observed

Simulated





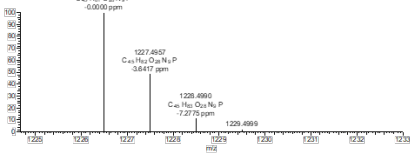
3

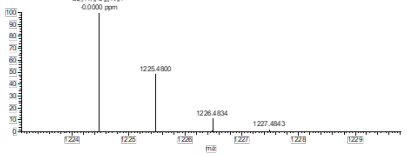


Chemical Formula: $C_{45}H_{80}N_9O_{28}P$
Exact Mass: 1225.4850

Observed

Simulated





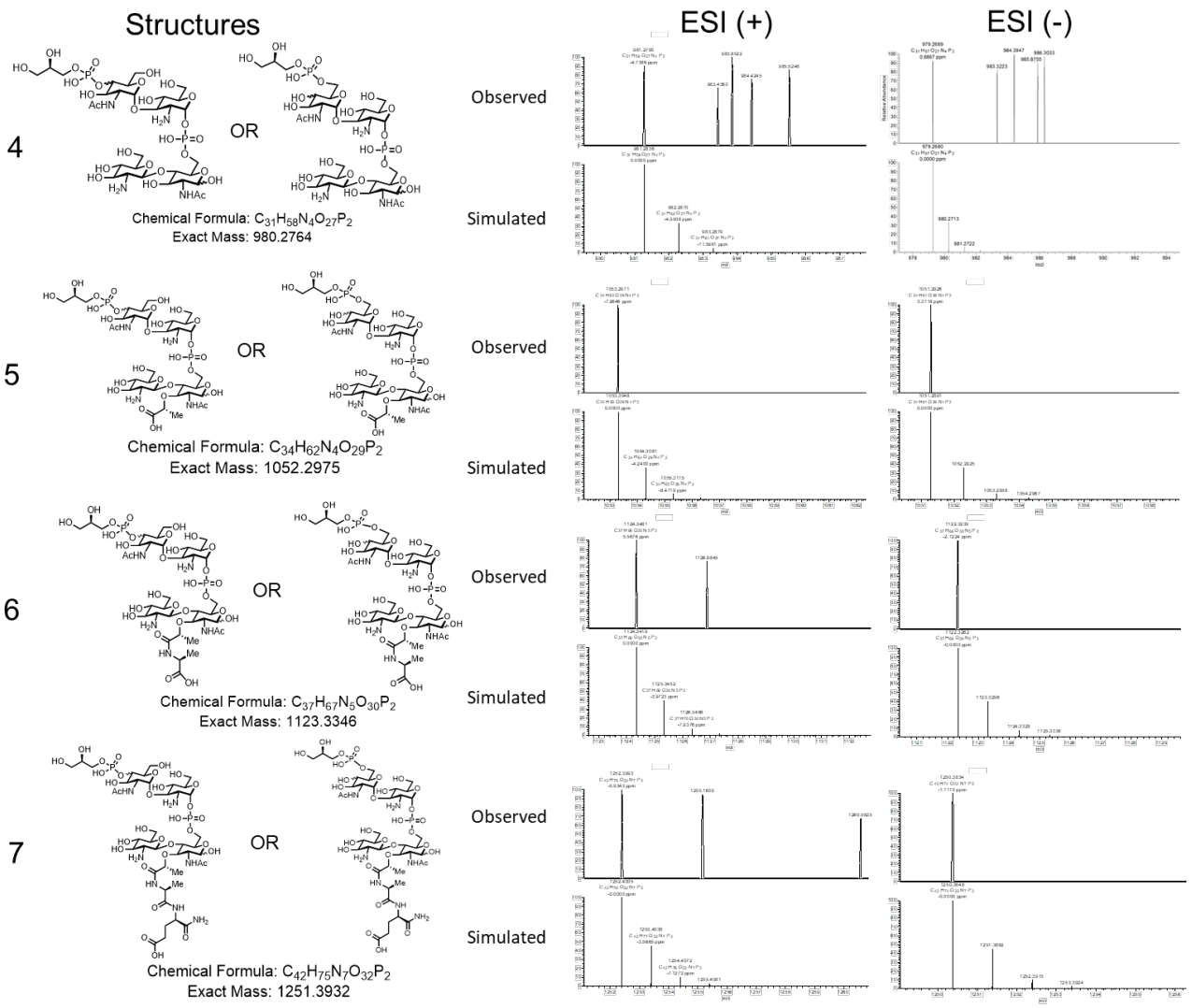
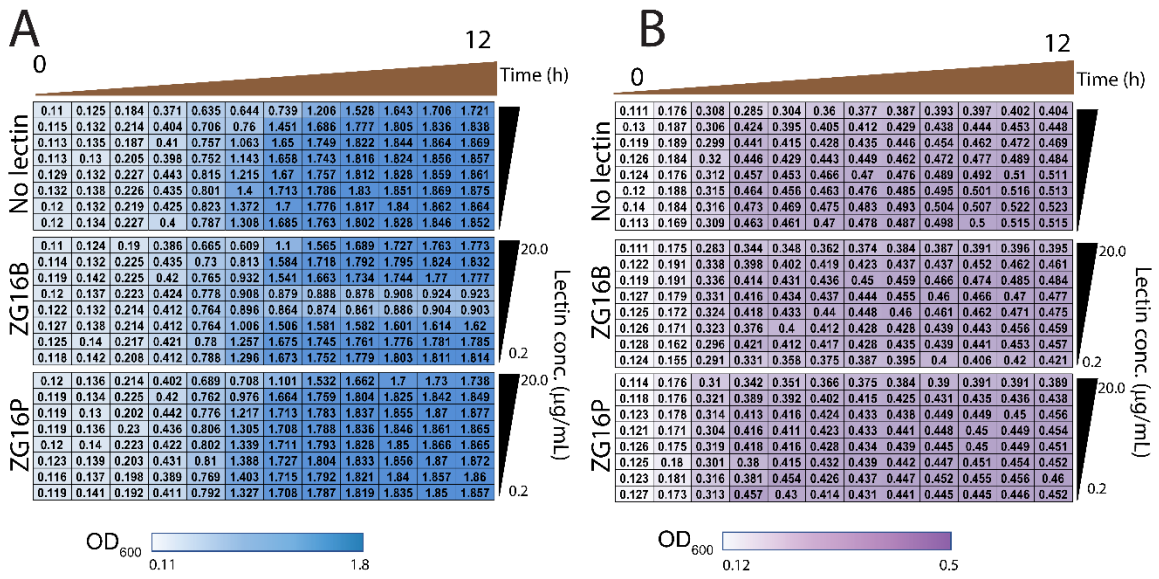
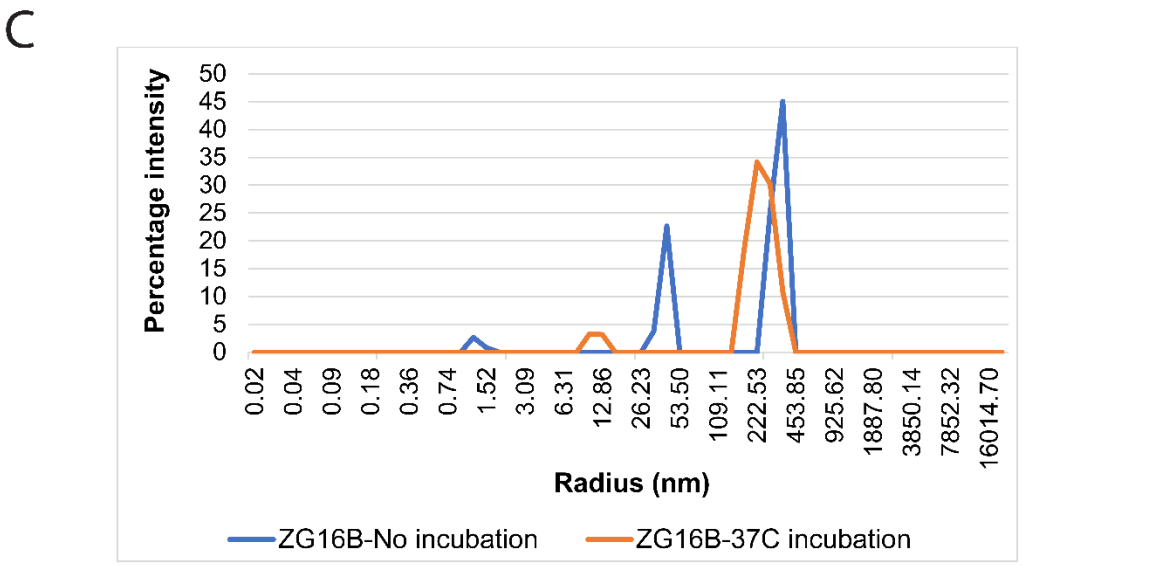


Fig. S4. HRMS spectra of identified peptidoglycan fragments present in ZG16B bound dental plaque bacteria: HRMS spectra obtained in the positive and negative mode ESI for the mutanolysin-digested soluble peptidoglycan fragments. The observed m/z peaks correspond to the simulated spectra of the structures shown here with mass accuracy ±10 ppm. Details of the [M+H] and [M-H] peaks are listed in Table S4.





ZG16B-no incubation				
Peak	Radius (nm)	Mw-R (kDa)	%Intensity	% Mass
Peak 1	1.6	10.2	3.5	100.0
Peak 2	51.9	34760.1	26.4	0.0
Peak 3	419.5	4612637.3	70.1	0.0

ZG16B-37° C incubation				
Peak	Radius (nm)	Mw-R (kDa)	%Intensity	% Mass
Peak 1	14.5	1789.5	6.5	99.8
Peak 2	314.9	2357266.2	93.5	0.2

Figure S5 ZG16B exerts a bacteriostatic effect on species it binds. (A) *S. vestibularis* and (B) *S. oralis* over a 12 h incubation period at 37 °C with or without recombinant lectins (ZG16B or ZG16P) at different concentrations. (C) DLS data show recombinant ZG16B aggregates at a concentration of 20 µg/mL when incubated for 12 h at 37 °C, the growth conditions used for the bacteria tested.

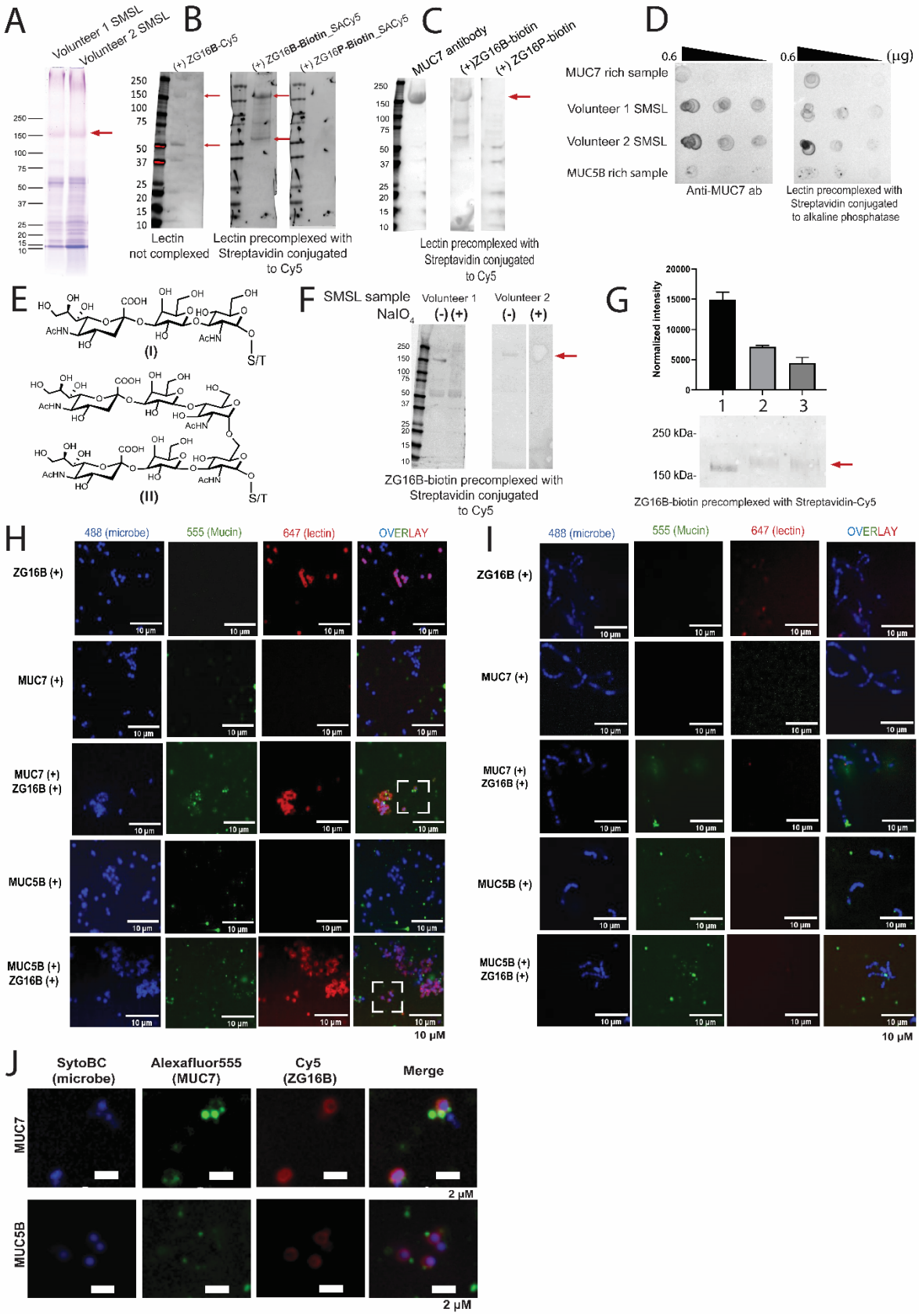

A
Volunteer 1 SMSL
Volunteer 2 SMSL
B
(+) ZG16B-Cy5
(+) ZG16B-Biotin_SACy5
(+) ZG16P-Biotin_SACy5
Lectin not complexed
Lectin precomplexed with Streptavidin conjugated to Cy5
C
MUC7 antibody
(+)ZG16B-biotin
(+) ZG16P-biotin
Lectin precomplexed with Streptavidin conjugated to Cy5
D
0.6
0.6
(µg)
MUC7 rich sample
Volunteer 1 SMSL
Volunteer 2 SMSL
MUC5B rich sample
Anti-MUC7 ab
Lectin precomplexed with Streptavidin conjugated to alkaline phosphatase
E
(I)
(II)
S/T
F
SMSL sample
NaIO4
Volunteer 1
Volunteer 2
(-) (+)
(-) (+)
ZG16B-biotin precomplexed with Streptavidin conjugated to Cy5
G
Normalized intensity
250 kDa-
150 kDa-
ZG16B-biotin precomplexed with Streptavidin-Cy5
H
488 (microbe)
555 (Mucin)
647 (lectin)
OVERLAY
ZG16B (+)
MUC7 (+)
MUC7 (+) ZG16B (+)
MUC5B (+)
MUC5B (+) ZG16B (+)
10 µM
I
488 (microbe)
555 (Mucin)
647 (lectin)
OVERLAY
ZG16B (+)
MUC7 (+)
MUC7 (+) ZG16B (+)
MUC5B (+)
MUC5B (+) ZG16B (+)
10 µM
J
SytoBC (microbe)
Alexafluor555 (MUC7)
Cy5 (ZG16B)
Merge
MUC7
MUC5B
2 µM
2 µM

Fig. S6. ZG16B recruits oral mucin MUC7 on *Streptococcus vestibularis.* (A) SDS-PAGE of submandibular/sublingual (SMSL) saliva samples from the two volunteers used in this study stained with Coomasie-PAS. An arrow indicates the major MUC7 band in the crude samples. (B) Interaction of ZG16B with salivary mucin MUC7 detected by the far-western blot of ST-SMSL salivary sample probed for ZG16B-Cy5 or biotinylated lectins, pre-complexed with Cy5-conjugated Streptavidin. (C) Interaction of ZG16B with salivary mucin MUC7, detected by the far-western blot of enriched MUC7 sample isolated from SMSL saliva, probed with MUC7 antibody or biotinylated lectins, pre-complexed with Cy5-conjugated Streptavidin (D) Dot blot of MUC7-enriched sample, SMSL saliva samples from two donors and MUC5B enriched sample, probed with anti MUC7 antibody and biotinylated ZG16B. (E) Substructures of mono- and (II) di-sialylated O- glycan structures present in MUC7 (I)NeuAcα2→3Galβ1→3GlcNAc-Ser/Thr, (II)NeuAcα2→3Galβ1→3(NeuAcα2→3Galβ1→4GlcNAcβ1→6)GalNAc-Ser/Thr.(F) Abrogation of ZG16B binding to MUC7 in SMSL saliva samples, treated with 1mM sodium periodate for 4 h at 37 ºC, detected by far-Western blotting using biotinylated lectins, pre-complexed with Cy5-conjugated streptavidin (G) The effect on ZG16B binding to MUC7 in SMSL saliva samples, treated with control buffer (1), α2-3,6,8 neuraminidase (2) or with both neuraminidase and O-glycosidase (3) using manufacturer's protocol (NEB E0540S) , detected by far-western blotting using biotinylated lectins, pre-complexed with Cy5-conjugated Streptavidin. The binding intensity of ZG16B to MUC7 at different conditions, normalized to the total MUC7 present in each sample, was quantified from three technical replicates, as represented in the bar graph. Confocal microscopic imaging of Syto BC stained (H) *S. vestibularis* and (I*) S. oralis* in the presence of ZG16B-Cy5 and mucin (MUC7 or MUC5B)-enriched samples probed with Alexa fluor 555, or both. Only MUC7 showed co-localization with ZG16B on the *S.vestibularis* and not MUC5B. (J) 4x zoomed-in ROI (shown in the dashed rectangles) of the confocal images in panels 3 and 5 of Fig S6*H* showing co-localization of MUC7 and ZG16B at the cell wall of the microbes, but not for the control mucin MUC5B.

Table S1. Details of the ZG16B-Cy5 bound bacterial isolates from human donor dental plaque (first round of screening).

Donor	Isolate name*	Sample ID**	Binding to ZG16B	Description (16S rRNA seq)	Whole genome Sequencing	E value	Percentage Identity
4	1	DP4_6_2	+‡	*Streptococcus sp.*	*Streptococcus mitis* group strain ChDC B345	0	94.71
4	2	DP4_6_1	+	*Gemella sp.*	*Gemella haemolysans* strain FDAARGOS_740	0	96.79
4	3	DP4_2A	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	98.05
4	4	DP4_1A	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	98.19
4	5	DP4_neg1	-†	*S. mitis.*	*Streptococcus oralis* strain SF100	0	95.71
3	6	DP3_10	+	*Gemella sp.*	*Gemella haemolysans* strain FDAARGOS_740	0	96.14
3	7	DP3_9B	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	96.8
3	8	DP3_6B	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	96.8
3	9	DP3_6A_2	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	96.8
3	10	DP3_5B	+	*Gemella sp.*	*Gemella haemolysans* strain FDAARGOS_740 chromosome	0	96.14
3	11	DP3_4	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	96.8
3	12	DP3_2B_2	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	96.8

3	13	DP3_1B	+	*S. salivarius*	*Streptococcus vestibularis* strain NCTC12167	0	96.8
3	14	DP3_neg1_1	-	*S. mitis*	*Streptococcus oralis* strain SF100	0	95.71

† = not bound by ZG16B-Cy5

‡ = bound by ZG16B-Cy5

*Isolate name used for reference in the figures and the main text

***Isolate nomenclature:* Dental Plaque (DP)_Donor Number_Isolate number_Letters and numbers indicating a subcultured isolate.

Table S2. Details of the ZG16B-Cy5 bound bacterial isolates from human donor dental plaque (second round of screening)

Donor	Sample ID	Bound by ZG16B-Cy5	Description (16S rRNA seq)	E value	Percentage Identity
8	DP8_Neg_2A	-†	*S. mitis*	0	98.16
8	DP8_1A	+‡	*S. salivarius*	0	99.26
8	DP8_2A	+	*S. salivarius*	0	99.24
8	DP8_3A	+	*S. salivarius*	0	99.62
8	DP8_4A	+	*S. salivarius*	0	98.19
8	DP8_5A	+	*S. salivarius*	0	99.62
8	DP8_6A	+	*S. salivarius*	0	99.62
9	TB9_Neg1_1	-	*Prevotella sp.*	0	92.88
9	TB9_1_1	+	*S. mitis*	0	99.24
9	DP9_10_1_A (Neg)	-	*S. mitis*	0	99.24
9	DP9_1_1	+	*S. salivarius*	0	98.18
9	DP9_2_1	+	*S. salivarius*	0	99.24
9	DP9_3_1	+	*S. salivarius*	0	97.36
9	DP9_4_1	+	*S. salivarius*	0	99.62
9	DP9_5_1	+	*S. salivarius*	0	99.24
9	DP9_6_1	+	*S. salivarius*	0	99.62
9	DP9_7_1	+	*S. salivarius*	0	99.24
9	DP9_8_1_A	+	*S. salivarius*	0	98.53
9	DP9_8_1_B	+	*S. salivarius*	0	98.53
9	DP9_9_1	+	*S. salivarius*	0	98.52
9	DP9_10_1_B	+	*S. salivarius*	0	99.24
9	DP9_11_1	+	*S. salivarius*	0	99.62
9	DP9_12_1_A	+	*S. salivarius*	0	99.24
9	DP9_12_1_B	+	*S. salivarius*	0	99.24
9	DP9_13_1	+	*S. mitis*	0	98.86
9	DP9_14_1	+	*S. salivarius*	0	99.62
9	DP9_15_1	+	*S. salivarius*	0	99.24
9	DP9_16_1	+	*S. salivarius*	0	99.24
9	DP9_17_1_B	+	*S. salivarius*	0	99.24
9	DP9_18_1_B	+	*S. salivarius*	0	98.19
9	DP9_19_1_B	+	*S. salivarius*	0	98.86
9	DP9_20_1_B	+	*S. salivarius*	0	98.89
9	DP9_21_1	+	*S. salivarius*	0	99.24
9	DP9_22_1	+	*S. salivarius*	0	99.24
9	DP9_23_1	+	*S. salivarius*	0	97.79
9	DP9_24_1	+	*S. salivarius*	0	93.29
9	DP9_25_1	+	*S. salivarius*	0	98.54
9	DP9_26_1	+	*S. salivarius*	0	98.53
9	DP9_27_1	+	*S. salivarius*	0	98.19
9	DP9_28_1	+	*S. salivarius*	0	98.88
9	DP9_29_1	+	*S. salivarius*	0	99.24

† = not bound by ZG16B-Cy5, ‡ = bound by ZG16B-Cy5

Table S3. Details of the 16S rRNA sequencing of non-ZG16B-Cy5 bound bacterial isolates from human donor dental plaque

Donor	Sample ID	Bound by ZG16B-Cy5	Description (16S rRNA seq)	E value	Percentage Identity
3	DP3_1A	-†	*Streptococcus pneumoniae*	1.56E-125	99.219
3	DP3_2A	-	*Streptococcus downii*	1.60E-121	99.197
3	DP3_3A	-	*Streptococcus pneumoniae*	1.67E-125	99.219
3	DP3_5A	-	*Salmonella enterica subsp.*	6.97E-127	91.092
3	DP3_8	-	*Streptococcus sp.*	3.54E-123	98.824
3	DP3_9A	-	*Streptococcus downii*	7.85E-125	98.473
3	TB3_neg1	-	*Streptococcus parasanguinis*	3.58E-125	99.216
4	DP4_1B	-	*Actinomyces naeslundii*	1.71E-118	99.583
4	DP4_3	-	*Streptococcus sanguinis*	1.58E-118	99.18
4	DP4_4	-	*Streptococcus sanguinis*	3.88E-124	99.213
4	DP4_5	-	*Streptococcus downii*	1.58E-118	98.031
4	TB4_neg1*	-	*Streptococcus salivarius*	2.20E-62	86.475
4	TB4_1A	-	*Streptococcus parasanguinis*	3.57E-125	99.216
4	TB4_1B	-	*Bacillus subtilis*	4.95E-31	78.696
4	TB4_2A	-	*Streptococcus parasanguinis*	1.18E-126	99.608
4	TB4_2B	-	*Schaalia odontolytica*	1.01E-122	98.824
4	TB4_3A	-	*Streptococcus sinensis*	3.70E-122	98.819
4	TB4_6B	-	*Veillonella nakazawae*	1.01E-115	97.244
7	DP7_neg1	-	*Streptococcus gordonii*	6.10E-123	99.206
7	DP7_1A	-	*Streptococcus downii*	1.27E-124	98.833
7	DP7_1B	-	*Streptococcus sp.*	3.28E-123	98.824
8	DP8_7A	-	*Streptococcus downii*	2.94E-124	99.213
9	DP9_10_1_A	-	*Bacterium NLAE-zl-P627*	5.34E-123	97.07
9	DP9_17_1_A	-	*Bacterium NLAE-zl-P627*	2.29E-121	98.438
9	DP9_18_1_A	-	*Streptococcus downii*	4.35E-124	98.833

† = not bound by ZG16B-Cy5

*Further characterization needed to identify the strain of the microbe belonging to *S. salivarius* group

Table S4: Details of the HRMS spectra of identified peptidoglycan fragments present in ZG16B bound dental plaque bacteria

Structure	Exact Mass	[M+H] expected	[M+H] observed	Mass accuracy (ppm)	[M-H] expected	[M-H] observed	Mass accuracy (ppm)
1	1022.4057	1023.4129	1023.4129	-0.0392	1021.3973	1021.3937	-3.4806
2	1097.3901	1098.3974	1098.3978	0.4322	1096.3817	1096.3816	-0.1008
3	1225.4850	1226.4923	1126.4915	-0.7014	1224.4767	1224.4738	-2.3768
4	980.2764	981.2836	981.2790	-4.7386	979.2680	979.2689	0.8867
5	1052.2975	1053.3048	1053.2971	-7.2646	1051.2891	1051.2926	3.3118
6	1123.3346	1124.3419	1124.3481	5.5676	1122.3262	1122.3239	-2.1224
7	1251.3932	1252.4005	1252.3993	-0.9343	1250.3848	1250.3834	-1.1110

Table S5. Values of Pearson's and Mander's coefficient calculated by the *Colocalization threshold* plugin in Fiji, quantifying the colocalization of ZG16B and MUC7 or non-binding mucin MUC5B on *S.vestibularis* by pixel-to-pixel correlation and co-occurrence.

Parameters	Channels involved	ZG16B, MUC5B (+)	ZG16B, MUC7 (+)
Pearson coefficient (r) (above threshold) *	(Cy5: Alexafluor 555)	0.158	0.437
Mander's coefficient (tM1) (above threshold) *	(Cy5: Alexafluor 555)	0.3571	1.0
Mander's coefficient (tM2) (above threshold) *	(Alexafluor 555: Cy5)	0.134	0.902

*Coste's auto-threshold

Table S6. Determination of the average size of the microbial clusters in *S. vestibularis* in the presence of ZG16B, mucins, or both from image analysis and particle size distribution.

Samples	Avg size of the particles (μm^2)	The area under the curve from the normalized distribution
ZG16B	2.15 ± 0.02	0.83251
MUC7	2.20 ± 0.02	0.90974
ZG16B, MUC7	5.7 ± 0.6	0.707242
MUC5B	2.0 ± 0.02	0.697993
ZG16B, MUC5B	2.97 ± 0.1	0.667081

SI References.

1. R. L. Policarpo et al., Flow-based enzymatic ligation by sortase A. *Angew. Chem. Int. Ed. Engl*. **53**, 9203-9208 (2014).
2. S. Ruhl, A. Eidt, H. Melzl, U. Reischl, J. O. Cisar, Probing of microbial biofilm communities for coadhesion partners. *Appl. Environ. Microbiol.* **80**, 6583-6590 (2014).
3. S. D. Jackman *et al.*, ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. **27**, 768-777 (2017).
4. J. Zhang, K. Kobert, T. Flouri, A. J. B. Stamatakis, PEAR: a fast and accurate Illumina Paired-End reAd mergeR, *Bioinformatics* **30**, 614-620 (2014).
5. N. R. Coordinators, Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **44**, D7-D19 (2015).
6. T. H. Clarke, L. M. Brinkac, G. Sutton, D. E. J. B. Fouts, GGRaSP: a R-package for selecting representative genomes using Gaussian mixture models. *Bioinformatics* **34**, 3032-3034 (2018).
7. B. D. Ondov *et al.*, Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **17**, 1-14 (2016).
8. S. Kumar, G. Stecher, D. Peterson, K. J. B. Tamura, MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. *Bioinformatics* **28**, 2685-2686 (2012).
9. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**, w293-w296 (2021).
10. N. Ramasubbu *et al.*, Large-scale purification and characterization of the major phosphoproteins and mucins of human submandibular-sublingual saliva. *Biochem J* **280** (Pt 2), 341-352 (1991).
11. S. Brown, T. Meredith, J. Swoboda, S. Walker, Staphylococcus aureus and Bacillus subtilis W23 make opolyribitol wall teichoic acids using different enzymatic pathways. *Chem. Biol.* **17**, 1101-1110 (2010).
12. H. Liang *et al.*, Metabolic labelling of the carbohydrate core in bacterial peptidoglycan and its applications. *Nat. Comm.* **8**, 15015 (2017).
13. S. R. Stowell *et al.*, Microbial glycan microarrays define key features of host-microbial interactions. *Nat. Chem. Biol.* **10**, 470-476 (2014).
14. E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774-797 (2007).