

Structural Insights into RbmA, a Biofilm Scaffolding Protein of *V. Cholerae*

Manuel Maestre-Reyna, Wen-Jin Wu, Andrew H.-J. Wang*

Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

Abstract

V. cholerae can form sessile biofilms associated with abiotic surfaces, cyanobacteria, zoo-plankton, mollusks, or crustaceans. Along with the vibrio polysaccharide, secreted proteins of the *rbm* gene cluster are key to the biofilm ultrastructure. Here we provide a thorough structural characterization of RbmA, a protein involved in mediating cell-cell and cell-biofilm contacts. We correlate our structural findings with initial ligand specificity screening results, NMR protein-ligand interaction analysis, and complement our results with a full biocomputational study.

Citation: Maestre-Reyna M, Wu W-J, Wang AH-J (2013) Structural Insights into RbmA, a Biofilm Scaffolding Protein of *V. Cholerae*. PLoS ONE 8(12): e82458. doi:10.1371/journal.pone.0082458

Editor: Holger Rohde, Universitätsklinikum Hamburg-Eppendorf, Germany

Received: May 22, 2013; **Accepted:** October 23, 2013; **Published:** December 5, 2013

Copyright: © 2013 Maestre-Reyna et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The Taiwanese national science council (NSC) indirectly funded this project via the Core Facility for Protein Structural Analysis (grant #NSC 101-2319-B-001-003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: ahjwang@gate.sinica.edu.tw

Introduction

In estuarine and brackish waters[1], *V. cholerae* may swim freely, or grow in sessile biofilms associated with abiotic surfaces, zoo-plankton, mollusks, or crustaceans[2]. *Vibrio cholerae* biofilms are involved in many aspects of the pathogen's life-cycle[3,4], as well as constituting a possible source of antibiotic resistances[5]. Along with the vibrio polysaccharide (VPS)[6], secreted proteins of the *rbm* gene cluster, including RbmA, are key to biofilm ultrastructure[7].

RbmA is a 26.4 KDa protein, with putative carbohydrate binding activity[8], which is found within the biofilm matrix, mediating cell-cell and cell-biofilm contacts[9]. Even though RbmA is not essential for biofilm biogenesis, it confers a high degree of mechanical stability to *V. cholerae* sessile communities by a mechanism which is not well understood. Here we present the RbmA crystal structure, both in its apo form and complexed with an artificial ligand. We have also performed ligand binding screening; and the results were confirmed via saturation-transfer difference (STD) NMR experiments[10]. We then proceeded to define ligand binding mode *in-silico*. We found RbmA to be a two-domain protein, functionally arranged as a dimer, with a high specificity for sialic acid, rhamnose, fucose, galactose and N-acetyl galactosamine. We propose a model for RbmA mediated biofilm-cell cross-linking, based on our structural and functional studies. This model helps advance our understanding of matrix scaffolding in *Vibrio* biofilms, and the study of the mechanisms by which bacteria associate themselves into communities. Our

proposed model may serve as the basis for a wide variety of *in vivo* studies, correlating the molecular with the ultrastructural levels in biofilm architecture. Furthermore, the determination of RbmA specificity is a first step toward the development of scaffolding inhibitors.

Methods

Cloning, production and purification of RbmA

An *E. coli* optimized synthetic RbmA gene based on the sequence from *V. cholerae* O1 (geneID 7855157) was designed, omitting the N-terminal secretion signals, and adding NdeI and XhoI restriction sites for cloning purposes. The secretion signal was predicted using SignalP[11]. The gene was then cloned into the pET28a vector, which was used to transform chemo-competent *E. coli* BL21(DE3) cells. Protein production was carried out in LB medium, via IPTG induction (1 mM final concentration), at 37 °C, and 160 rpm.

Alternatively, Se-Met derived protein destined for single wavelength anomalous diffraction (SAD) phasing was produced using the Overnight Express Autoinduction system 2 (Novagen), as described in the handbook.

E. coli cells carrying RbmA were then harvested, resuspended in loading buffer (20 mM Tris/HCl pH 8, 100 mM NaCl, 5 mM imidazol) complemented with EDTA free protease inhibitor cocktail (Roche), and lysed using a cell disruptor (Constant Systems LTD). Cell debris was removed via centrifugation, and supernatant was filtered and loaded on a 20 mL Ni Sepharose 6 Fast Flow column (GE). After loading, and

washing, the protein was eluted via a linear gradient with elution buffer (loading buffer + 500 mM imidazol). The protein was concentrated using Amicon concentrators, and loaded into a Superose 6 size exclusion chromatography column, pre-equilibrated in crystallization buffer (20 mM Tris/HCl pH, 100 mM NaCl). Main peak fractions were collected and re-concentrated to around 25 mg/mL, and stored at 4 °C.

High-throughput glycan array binding assays

The glycan array used by the Consortium for Functional Glycomics (CFG) consists of different groups of oligosaccharides that are presented by mammalian cells. RbmA was fluorescently labeled using an AlexaFluor 488 SPD kit (Invitrogen) and applied to CFG array V5.1 chips at 200 µg/mL. Alternatively, RbmA was directly applied to the glycan array, with binding activity being detected via fluorescent anti-his-tag antibodies. Chip surfaces were repeatedly washed and remaining fluorescence was measured and quantified.

Each binding event was repeated six times, with the highest and lowest value discarded. The remaining data were averaged, and standard deviations were calculated.

Samples for NMR measurements. The NMR samples contained 20-fold molar excess of sugars added to RbmA (0.1 mM dimer concentration) in pH 7.4 buffers containing 20 mM potassium phosphate, 100 mM NaCl, 8% D₂O (for locking purpose) and 0.01 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, for chemical shift referencing). The pH values for both sugars and RbmA were adjusted to the same value prior to mixing (less than a 0.05 pH unit differences, if any).

Saturation transfer difference (STD) experiments [10]. For the on-resonance irradiation experiment, a train of 50 msec Gaussian shape pulses were applied to the protein signals at -0.37 ppm (up field shifted methyl groups) for two seconds, one second relaxation delay was applied. For the off-resonance irradiation experiment, the same selective pulse was applied to -20 ppm where no signals occurred. A 65 msec spin lock period was employed to filter out protein resonances. Water suppression was achieved using 3919 WATERGATE scheme [12]. 7424-9600 scans were accumulated. The data were collected at 298 K on a Bruker 800 MHz spectrometer equipped with a CryoProbe.

Crystallization of RbmA

RbmA was subjected to an initial sparse matrix screening, followed by optimization of promising hits. Native, and selenomethionine derived RbmA crystals were grown in conditions containing 0.1 M Bis-tris pH 5.5 – 6.5, and 1.5 – 2.5 M NaCl. The best diffracting native RbmA crystal (4be6) was grown in the presence of 10 mM sialic acid. However, no sialic acid could be detected in the electron density map of the RbmA diffraction data. 18-crown-6 co-crystals were obtained in 0.1M HEPES pH 7 – 7.5, 0.2 M CaCl₂, 30% PEG400 (V/V). Native RbmA crystals took between 2 - 3 days to grow fully, while 18-crown-6 co-crystals were visible almost immediately after pipetting, and grew larger within two days.

Table 1. Data collection and Refinement details.

	ApoRbmA	SeMet_RbmA	RbmA Crown
Data collection			
Space group	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2	P3 ₁
Cell dimensions			
a, b, c (Å)	118.22, 118.22, 104.98	118.89, 118.89, 104.78	136.88, 136.88, 116.31
α, β, γ (°)	90, 90, 90	90,90,90	90,90,120
Wavelength (Å)	1.00000	0.96000	0.97622
Resolution (Å)	2.00 - 24 (2.00 - 2.03)	2.46 - 50 (2.46 - 2.5)	2.6 - 22 (2.6 - 2.69)
R _{merge}	6.5 (55.1)	7.2 (37.5)	9.3 (93.1)
I/σ	37.8 (5.2)	45.4 (9.74)	21.6 (2.44)
Completeness (%)	99.7 (100.0)	99.86 (98.97)	100 (100)
Redundancy	13.2 (11.6)	9.8 (9.8)	5.9 (5.8)
Refinement			
Resolution (Å)	2.05	2.46	2.6
No. reflections	50815	26340	75002
R _{work} / R _{free}	13.29/16.09	18.38/23.64	22.09/26.02
No. atoms			
Protein	3570	3562	14166
Water	540	307	290
Crown-ethers			144
B-factors			
Protein	25.90	39.26	55.11
Water	41.10	37.02	45.08
Crown-ethers			44.08
R.m.s deviations			
Bond lengths (Å)	0.012	0.010	0.009
Bond angles (°)	1.46	1.41	1.34

Values for highest resolution shell are in parenthesis.

doi: 10.1371/journal.pone.0082458.t001

Data collection, phase solution, and refinement

Data collection took place at the 13B and C beamlines at the National Synchrotron Radiation Center (Hsinchu, Taiwan), and at the 12B2 beamline Spring-8 (Hyogo, Japan). ApoRbmA crystals and SeMet derivatives belonged to space group P4₁2₁2, and were solved by single anomalous dispersion (SAD) using Shelx[13]. All other crystals, including co-crystals, were solved via molecular replacement using the Phaser[14] software package. Refinement was carried out using Refmac5[15], included in the CCP4 suite[16], and Coot[17]. Secondary structure assignments were performed with STRIDE[18], and figures were rendered using PyMOL[19]. Data collection details and refinement statistics are listed in Table 1.

Docking and molecular dynamics simulations

Suitable binding pockets were first approximated using the InCa-sitefinder server[20]. Intermolecular A/B domain structures were submitted to the server in order to avoid smaller binding pockets being masked by the wide groove (Figure S4 in File S1). Once preliminary binding pockets had been isolated, they were confirmed via docking with sialic acid, using Vina[21]. In order to determine the grid sizes and to

visualize ligand poses, the PyMOL[19] Autodock/Vina plugin was employed[22]. Sialic acid was prepared for docking using a glycam06[23] template. Hydrogen atoms were added to the template using the ADT software package, which is capable of protonating heavy atom according to pH, valence, and charge.

MD simulations were carried out using Amber12[24]. For the 60 ns simulations in the presence and absence of the 18-crown-6 ligand, the protein model based on structure 4bei was surrounded with a 20 Å TIP3PBOX water box. For the modeling of the crown ether, the Antechamber software and the General Amber Force Field (GAFF) were employed. The protein was then subjected to energy minimization for 10000 cycles, followed by equilibration with weak restraints for 100 ps. Finally, the system was equilibrated to 300 K and 1 atm, and run for 60 ns. For the docking experiments, protein-glycan complexes were selected using the docking energy criteria (Table S1 in File S1), but also taking into account the binding mode. Hence, two ligand conformations were selected for the D- and O-loop protein conformations, but only one for the wide groove binding pocket. Further processing of the protein-glycan complexes was analogous to the RbmA•18-crown-6 complex, but the water boxes were kept at 10 Å, and simulations were only 20 ns long. Finally, the RbmA•polylactose complex was generated with Vina by joining the two most favored conformations of a polylactose decamere. The resulting model was then treated following guidelines reported previously in literature for the treatment of protein-glycan complexes[25].

Trajectories were evaluated with VMD[26], which was also used for movie rendering. The RMSD data was plotted with Qtiplot[27].

Multiple alignment of RbmA sequences

The presence of RbmA in the genomes of all sequenced strains presented in reference [28] was first determined via Blastp[29,30]. Once the subset of RbmA+ *V. cholerae* strains was identified, a multiple alignment of all RbmA sequences was performed with COBALT[31] and displayed with ClustalX[32].

Results and Discussion

The RbmA crystal structure (2.05 Å resolution; PDBID 4be6) revealed an all-beta dimer (Figure 1 A), which was similar to recently published results[33], with the main differences at the dimer interfaces. Each monomer is composed of two structurally similar (1.4 Å RMSD), but distinct (24% sequence identity) domains (A- and B-domains). In the dimer, each A-domain interacts loosely with the B-domain of the same chain, forming a wide groove, and with the B-domain of the other polypeptide, forming a tight groove. Within each chain the A and B domains are connected by a flexible hinge loop (Figure 1). The wide groove builds on electrostatic interactions along charged stripes on the inner surface of each of the domains (Figure S1 in File S1), while the tight groove is firmly held together by interlocking β -sheets (Figure 1 A). What our structure shows that the previously published do not, is that, within this region, strand β 5 can show two conformations in the same dimer. In the disordered loop (D-loop) conformation, β 5 is

continuous, and is followed by a long, flexible loop, which was not visible in the electron density map (Figure 1B). On the other hand, in the ordered loop (O-loop) conformation, β 5 is very short, replaced by a well-defined loop, which tightly hugs the B-domain across the tight groove (Figure 1C). The O-loop is followed by the short β 6, which is absent in the D-loop conformation. The O-loop partially shields a positively charged patch on the wide groove from solvent, exposing negative charges on the hitherto fully positive surface of the tight groove, and deepening it into a 417 Å³ pocket (O-loop pocket) (Figure S2 in File S1).

In order to assess the putative glycan binding capacity of RbmA, we undertook high-throughput, semiquantitative binding studies with the Consortium for Functional Glycomics (CFG). The AlexaFluor 488 labeled RbmA binding profile provides a detailed specificity overview against a wide variety of glycans (Figure 2). We further confirmed these binding results, and discarded any false positives or negatives produced by the labeling on primary amines, by repeating the experiment with unlabeled RbmA, which was then detected via fluorescent anti-his antibodies (Figure 2). RbmA seemed to prefer simple mono- to trisaccharides. Sialic acid (Neu5NAc) containing glycans 11 and 223 were the top binders, while fucose-containing glycans also figured prominently (glycans 6 and 80). Finally, galactose, rhamnose, and N-acetyl galactosamine (glycans 1, 8, and 5, respectively) were bound at slightly lower strengths than glycan 80. Glycan 5, in fact, bound solely as α -GalNAc-serine, which is the smallest unit for O-glycosylation[34]. RbmA showed a clear preference for α -linked galactose, while sialic acid was preferably bound in the β form, while the same glycans, with the opposite linkages, remained consistently free (glycans 10, and 11 for α -sialic acid, 12 for β -galactose). Such a preference explains the low number of binding partners for RbmA on the glycan array, since these glycans are underrepresented in it. α -rhamnose is the only rhamnose containing sugar in the array, and only 65 glycans out of the 610 contain α -galactose; the presence of β -sialic acid is even lower, with 6 glycans presenting it. Fucose, on the other hand, is not underrepresented, but every fucoside except glycans 6 and 80 contained a β -galactose moiety, which seems to be detrimental to RbmA binding.

In order to complement our findings on the glycan array, we performed saturation transfer difference (STD) NMR experiments[10]. In these, the saturation (signal attenuation) on a buried methyl group signal propagated throughout the entire protein, and then to the bound-substrate via spin-spin diffusions. Based on the STD effect, it appeared that both sialic acid and sialyllactose (similar to glycan 223) used acetyl methyl groups to bind RbmA, suggesting a very localized binding contact (Figure 3). Further, even though sialyllactose contains α -sialic acid, experiments demonstrate that neither the galactoside, nor the glucoside moieties interact directly with the protein, suggesting that an α -linked negative glycan probably cannot extensively interact with RbmA. Galactose, fucose, and rhamnose STD experiments also demonstrate specific binding to RbmA (Figure S3 in File S1). Galactose and fucose appear to bind mainly via the C3-C4 diol region (Figure S3 A and B in File S1), while the binding of rhamnose is much more

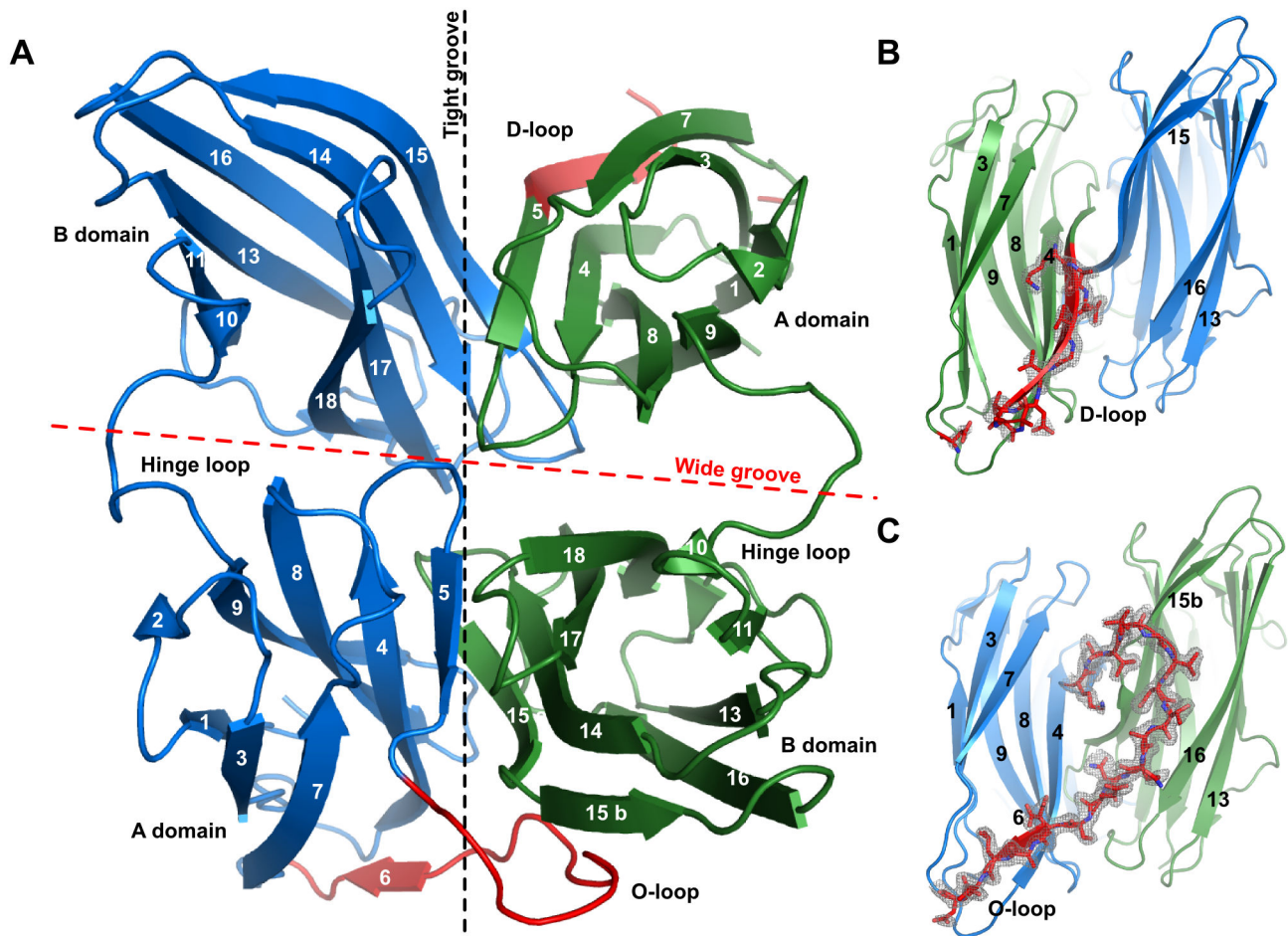


Figure 1. The RbmA structure. A: ApoRbmA general fold. RbmA is a functional dimer (monomers in blue and green). At the tight groove interface (black dotted line), between β -strands five and seven, RbmA can adopt two distinct conformations (red, D- and O-loop). B: Detail of the D-loop conformation, which results in an elongation of $\beta 5$, the loss of $\beta 6$ and the disorganization of the loop beyond $\beta 5$. C: Detail of the O-loop conformation, which presents an elongated loop between $\beta 5$ and $\beta 6$. OMIT electron at 1σ ($0.24 \text{ e}/\text{\AA}^3$), carved to 1.6 \AA .

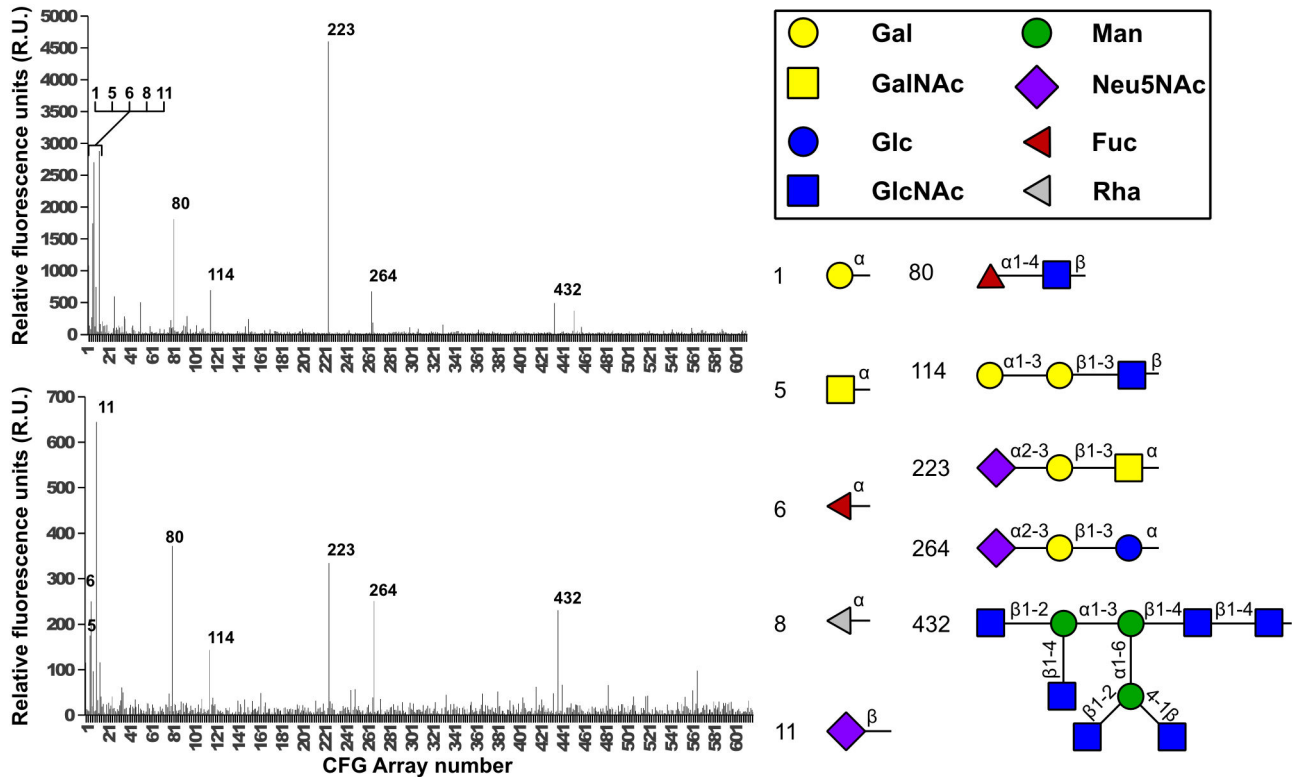
doi: 10.1371/journal.pone.0082458.g001

extensive, with STD signals for the protons on C2, C4, and C5 (Figure S3 C in File S1).

Galactose and GalNAc are building blocks of VPS, whereas sialic acid, fucose, and rhamnose, for which RbmA exhibited preference, are not [35]. On the other hand, all four non-charged sugars can be found in the O-antigen of different *V. cholerae* strains. Sialic acid is represented on the Vibron surface by analogues like 3-Deoxy-D-manno-oct-2-ulosonic acid, or even N-acetylated derivatives of D-perosamine [35–41], hinting at a mechanism for RbmA-cell attachment. *V. cholerae* strains have been mainly characterized as O1, O139 and non-O serotypes. While, within the first two groups, O-antigens are well conserved, the third one is much more diverse. By comparing the first and the third groups [28] and searching for RbmA in their respective genomes, we noticed that many non-O strains did not encode for the *RbmA* gene. Those that did, and some of the O1 serotype strains, presented a series of

mutations within the tight groove area, and at the wide groove interface (Figure S4 in File S1). These mutations directly affect the charge and polarity within the O-loop pocket, and of the complementary stripes within the wide groove, thus altering their specificity. Further, it is not uncommon for microbial adhesins involved in biofilm, and surface adhesion to be very plastic, accepting a relatively diverse set of ligands [42–44], and it is therefore not unreasonable to expect RbmA to be able to accommodate related, yet diverse, ligands.

In RbmA, STD experiments suggested that the region surrounding the positive O-loop pocket (Figure S5 in File S1) is the most likely candidate for strong binding with negatively charged ligands, while the D-loop conformation exposes the positive charges in a flat and open region, allowing for a first interaction with the bacterial surface. With two O-loop pockets per functional unit, RbmA would bidentally recognize two ligands, directly mediating cell-cell contacts. This hypothesis



was further tested via extensive *in-silico* screening of the protein surface. As expected, the O-loop pocket and the wide groove were readily identified as good binding regions by the InCa sitefinder server[20], and via docking with sialic acid (Figure 4A, and C, and S2 and S6 in File S1). For the D-loop conformation, several small pockets were detected (Figure 4B, and S2 in File S1). Further confirmation via molecular dynamics indicated that binding was possible for the first two, but not for the D-loop conformation, which was only marginally stable (Figure 4).

A thorough co-crystallization screening with various carbohydrates was unsuccessful. Furthermore, soaking experiments with a variety of monosaccharides resulted in the crystals immediately dissolving. As an alternative, a co-crystallization of RbmA with the 18-crown-6 ether was attempted. Crown ethers bind proteins, bridging axially bound lysines to polar amino-acids, with equatorial interactions with tryptophanes and other aromatic residues (Lee, C.-C. et al. in press), and have been shown to occlude ligand accessible channels[45]. Since negatively charged glycans can bind to similar pockets[20], a crown ether was used as a probe in co-crystallization experiments (Figure 5A). These dramatically changed the crystallization behavior of RbmA, and resulted in the emergence of a new crystal form with eight molecules per asymmetric unit (PDBID 4bei). Within each chain, one 18-

crown-6 ether molecule was deeply embedded in the wide groove, bound between K51 in the A-domain, and T218 in the B-domain. Furthermore, all dimers containing crown-ethers had their tight grooves protected by the O-loop conformation, perhaps due to the interaction between the 18-crown-6 ethers and Y87 (Figure 5B).

Crown ether dependent protein stability was studied via 60 ns molecular dynamics simulations (Figure 5C and S7 in File S1). In the absence of the crown ether, strong flickering occurs within the wide groove, indicating a change between an open and a closed conformation. In the presence of the crown ether, flickering is reduced, with the protein closing around the ligand, further confirming the wide groove as a good candidate for ligand binding. A similar behavior could be observed when a poly-lactose oligosaccharide was modeled into the wide groove (Figure 5C and S7 in File S1, and movie S3), suggesting that this structural feature is a suitable binding pocket for polysaccharides. Our results support previous data, which had suggested that RbmA adopts a more extended conformation in solution[33]. Our results, however, additionally demonstrate that the wide groove is capable of accommodating large substrates. These stabilize RbmA in the closed, compact conformation, which is preferentially adopted in the crystal.

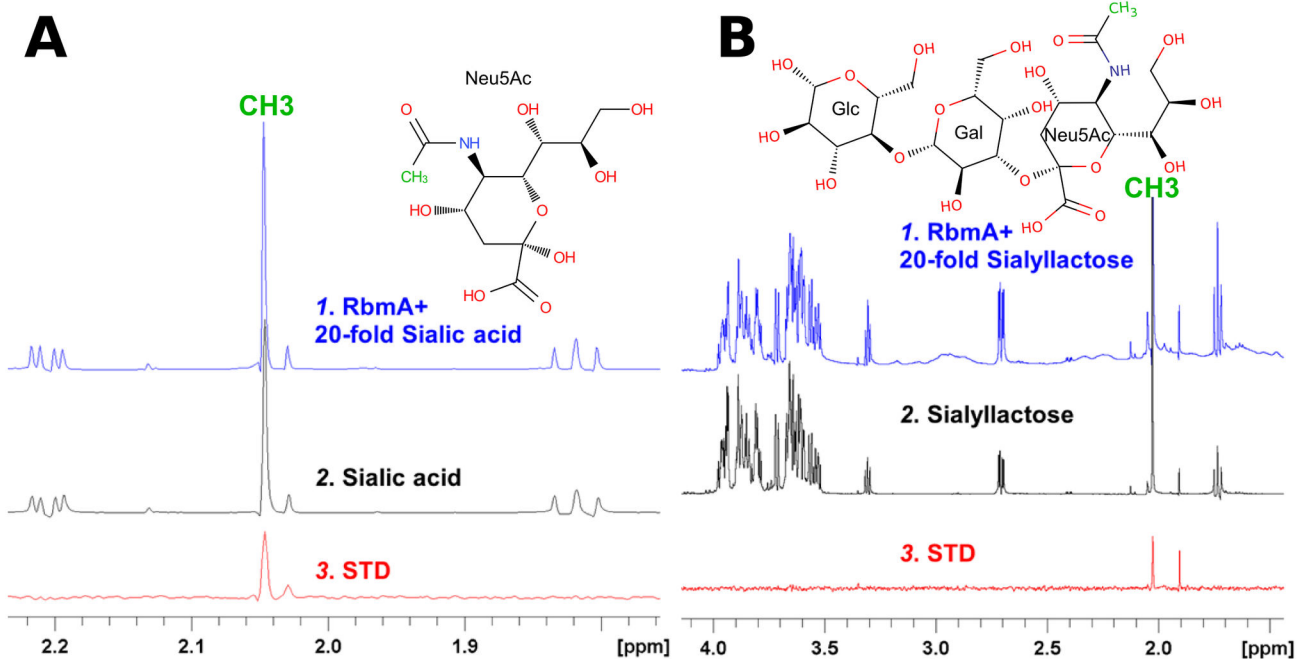


Figure 3. Saturation transfer difference experiments for sialic acid and sialyllactose. A: STD experiments between RbmA and sialic acid. B: STD experiments between RbmA and sialyllactose. The experimental details are described in Methods. The acetyl CH₃ groups exhibiting saturation transfer difference effect are labeled in green in each figure.

doi: 10.1371/journal.pone.0082458.g003

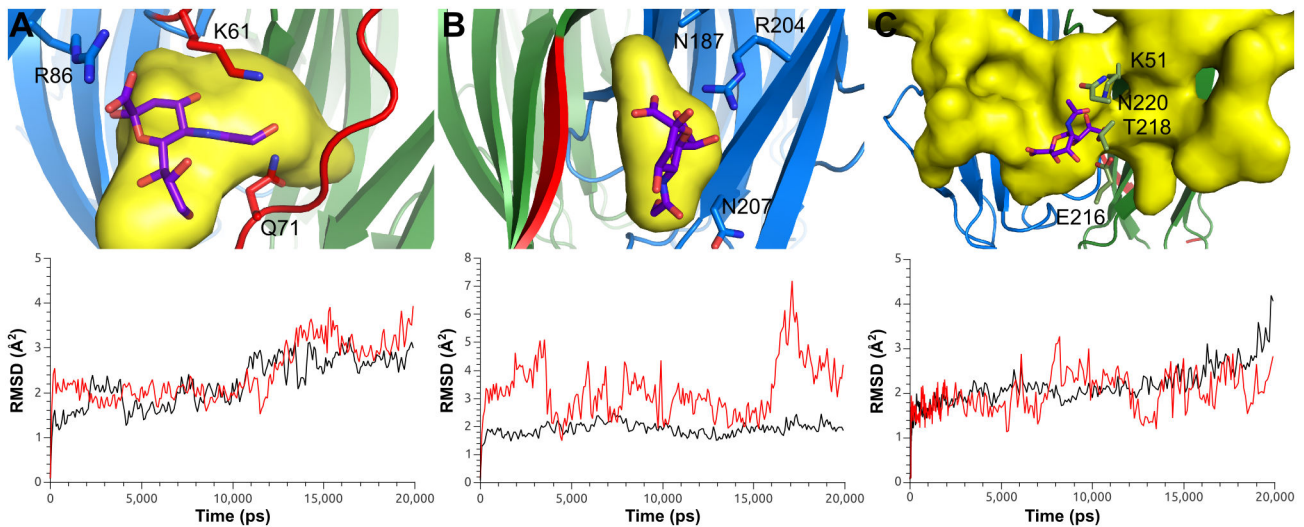


Figure 4. Docking and molecular dynamics simulations of RbmA•sialic acid complexes. Top: RbmA is shown in blue and green, with the O- and D-loop conformations (A and B, respectively) highlighted in red. InCa sitefinder results are shown as yellow surfaces, with the sialic acid docking poses superimposed. Sialic acid is shown in purple, with labeled amino-acids defined as flexible for docking. Bottom: Amber12 md RMSD trajectories for the protein backbone (black) and the glycan (red). A: RbmA•sialic acid within the O-loop pocket. B: RbmA•sialic acid at the D-loop conformation surface. C: RbmA•sialic acid at the wide groove pocket. Other conformations were tested, resulting in less stable binding (Fig. S4 in file S1 and movies S4-S8).

doi: 10.1371/journal.pone.0082458.g004

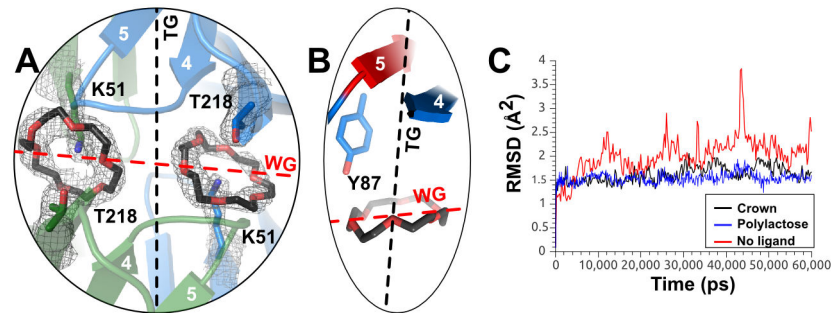


Figure 5. The binding of 18-crown-6 ethers. A: Detail of the RbmA•18-crown-6 binding pocket. One crown ether was bound in the wide groove by each monomer via K51 (A-domain) and T218 (B-domain). B: Detail of the interaction between the 18-crown-6 ether and Y87, which might stabilize $\beta 5$ in the O-loop conformation. C: RMSD traces of free RbmA (red), forming a complex with experimental crown ethers (black), or a modeled poly-lactose chain (blue, movies S2, S1, and S3, respectively). While, in the presence of the 18-crown-6 ligand, or a poly-lactose chain, the average position of the RbmA backbone stays constant for 60 ns, while the apo structure is generally less stable, with strong, occasional flickering every ~ 10 ns.

doi: 10.1371/journal.pone.0082458.g005

Conclusion

Taken together, our results suggest that RbmA presents two distinct binding sites. On the one hand, the wide groove is capable of accommodating large, filamentous substrates, such as VPS. On the other hand, the O-loop pockets are capable of stable binding of negatively charged carbohydrates, which are a hallmark of cellular surfaces. Our glycan array data, confirmed via STD experiments, indicates that RbmA preferentially binds monosaccharides from the VPS and from the LPS. A diverse pool of ligands is a hallmark of other biofilm-related adhesins[42], and may suggest multivalent binding to several components of the matrix and/or cellular surfaces. Furthermore, the importance of the sialic acid amide group in binding suggests that acylated perosamines, a key component in *V. cholerae* O-antigen[40,46], or acidic glycans would be suitable binders.

Our structural, and binding studies, suggest a model for biofilm scaffolding (Figure 6), in which RbmA acts as an anchor

between cells and between cells and the biofilm, by attaching bidentally to the bacterial carbohydrate-rich surface, but also to a filamentous component of the biofilm, in a similar manner as was proposed by biofilm architectural studies[7]. Giglio et al., with their preliminary studies, showed that the positive charge in the tight groove is important for RbmA function, while they were not able to pinpoint the relevance of the loop conformational transitions [33]. However, our data opens new possibilities for further mutagenesis strategies because we were able to assign possible glycan binding pockets, and established a correlation between O-antigen, and RbmA diversity (Figure S4 in File S1). These, in combination with *V. Cholerae in vivo* studies in its sessile habitat may enhance our understanding on RbmA mediated adhesion and RbmA strain-dependent diversity. Additionally, our characterization of glycan specificity will prove very useful in targeting biofilm inhibition assays.

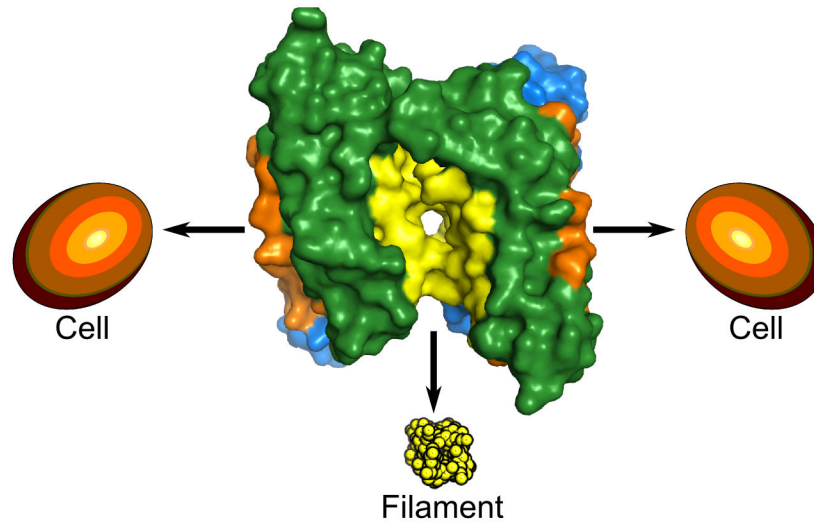


Figure 6. A proposed model for RbmA action. Taking all our results into account, we propose that RbmA (center, green and yellow), is capable of binding long, filamentous substrates in its wide groove (yellow surface area) like, for example polysaccharides (bottom, yellow ball model). On the other hand, the positive charges within O-loop pockets on the tight grooves (orange surface areas), might allow RbmA to bind to cellular surfaces (orange ovoid shapes), which are densely negative.

doi: 10.1371/journal.pone.0082458.g006

Supporting Information

File S1. Supporting Materials. The supporting materials file contains all supporting figures, and supporting tables. (PDF)

Movie S1. Molecular dynamics simulation of the RbmA crown-ether complex. (MP4)

Movie S2. Molecular dynamics simulation of free RbmA. (MP4)

Movie S3. Molecular dynamics simulation of the modeled RbmA -polylactose complex. (AVI)

Movie S4. Molecular dynamics simulation of stable sialic acid-RbmA complex in the O-loop conformation. (MP4)

Movie S5. Molecular dynamics simulation of mildly unstable sialic acid-RbmA complex in the D-loop conformation. (MP4)

Movie S6. Molecular dynamics simulation of stable sialic acid-RbmA complex in the wide groove. (MP4)

Movie S7. Molecular dynamics simulation of unstable sialic acid-RbmA complex in the O-loop conformation.

References

- Colwell RR, Huq A (1994) Environmental reservoir of *Vibrio cholerae*. The causative agent of cholera. *Ann N Y Acad Sci* 740: 44–54. doi: 10.1111/j.1749-6632.1994.tb19852.x. PubMed: 7840478.
- Tamplin ML, Gauzens AL, Huq A, Sack DA, Colwell RR (1990) Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl Environ Microbiol* 56: 1977–1980. PubMed: 2383016.
- Butler SM, Camilli A (2005) Going against the grain: chemotaxis and infection in *Vibrio cholerae*. *Nat Rev Microbiol* 3: 611–620. doi:10.1038/nrmicro1207. PubMed: 16012515.
- Faruque SM, Biswas K, Udden SMN, Ahmad QS, Sack DA et al. (2006) Transmissibility of cholera: in vivo-formed biofilms and their relationship to infectivity and persistence in the environment. *Proc Natl Acad Sci U S A* 103: 6350–6355. doi:10.1073/pnas.0601277103. PubMed: 16601099.
- Meibom KL, Blokesch M, Dolganov NA, Wu C-Y, Schoolnik GK (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science* 310: 1824–1827. doi:10.1126/science.1120096. PubMed: 16357262.
- Fong JCN, Syed KA, Klose KE, Yildiz FH (2010) Role of *Vibrio* polysaccharide (vps) genes in VPS production, biofilm formation and *Vibrio cholerae* pathogenesis. *Microbiology* 156: 2757–2769. doi: 10.1099/mic.0.040196-0. PubMed: 20466768.
- Berk V, Fong JCN, Dempsey GT, Develioglu ON, Zhuang X et al. (2012) Molecular architecture and assembly principles of *Vibrio cholerae* biofilms. *Science* 337: 236–239. doi:10.1126/science.1222981. PubMed: 22798614.
- Absalon C, Van Dellen K, Watnick PI (2011) A communal bacterial adhesin anchors biofilm and bystander cells to surfaces. *PLOS Pathog* 7: e1002210.
- Fong JCN, Karplus K, Schoolnik GK, Yildiz FH (2006) Identification and characterization of RbmA, a novel protein required for the development of rugose colony morphology and biofilm structure in *Vibrio cholerae*. *J Bacteriol* 188: 1049–1059. doi:10.1128/JB.188.3.1049-1059.2006. PubMed: 16428409.
- Mayer M, Meyer B (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew Chem - Int Ed* 38: 1784–1788. doi:10.1002/(SICI)1521-3773(19990614)38:12. PubMed: 16428409.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340: 783–795.
- Sklenar V, Piatto M, Leppik R, Saudek V (1993) Gradient-tailored water suppression for 1H-15N HSQC experiments optimized to retain full sensitivity. *J Magn Reson A* 102: 241–245. doi:10.1006/jmra.1993.1098.
- Sheldrick GM (2002) Macromolecular phasing with SHELXE. *Z Kristallogr* 217: 644–650. doi:10.1524/zkri.217.12.644.20662.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40: 658–674. doi:10.1107/S0021889807021206. PubMed: 19461840.
- Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67: 355–367. doi: 10.1107/S0907444911001314. PubMed: 21460454.
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242. doi:10.1107/S0907444910045749. PubMed: 21460441.
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486–501. doi:10.1107/S0907444910007493. PubMed: 20383002.

(MP4)

Movie S8. Molecular dynamics simulation of highly unstable sialic acid-RbmA complex in the D-loop conformation.

(MP4)

Acknowledgements

The authors would like to thank Academia Sinica for support, the National Synchrotron Radiation Research Center of Taiwan, and Spring-8 for beam time allocations, the Core Facility for Protein Structural Analysis (grant #NSC 101-2319-B-001-003) for use of its crystallization facility, Prof. Lars-Oliver Essen (Philipps University Marburg) and the National Center for High-Performance Computing for computing time allocation, and Dr. David Smith and the Consortium for Functional Glycomics for performing *in vitro* screening of glycan specificity (CFG request # 2666). The NMR spectra were obtained at the Core Facility for Protein Structural Analysis supported by National Core facility Program for Biotechnology.

Author Contributions

Conceived and designed the experiments: AHJW WJW MMR. Performed the experiments: MMR WJW. Analyzed the data: MMR WJW AHJW. Wrote the manuscript: AHJW WJW MMR.

18. Heinig M, Frishman D (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res* 32: W500–W502. doi:10.1093/nar/gkh429. PubMed: 15215436.
19. The PyMOL Molecular Graphics System, Version 1.6, Schrödinger, LLC. Available: <http://www.pymol.org/>. Accessed 2013 January 25
20. Kulharia M, Bridgett SJ, Goody RS, Jackson RM (2009) InCa-SiteFinder: a method for structure-based prediction of inositol and carbohydrate binding sites on proteins. *J Mol Graph Model* 28: 297–303. doi:10.1016/j.jmgm.2009.08.009. PubMed: 19762259.
21. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31: 455–461. PubMed: 19499576.
22. Seeliger D, de Groot BL (2010) Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des* 24: 417–422. doi:10.1007/s10822-010-9352-6. PubMed: 20401516.
23. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR et al. (2008) GLYCAM06: a generalizable biomolecular force field. *Carbohydrates*. *J Comput Chem* 29: 622–655. doi:10.1002/jcc.20820. PubMed: 17849372.
24. Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, et al. (2012) Amber12. University of California, San Francisco.
25. Sapay N, Nurisso A, Imbert A (2013) Simulation of carbohydrates, from molecular docking to dynamics in water. *Methods Mol Biol* 924: 469–483. doi:10.1007/978-1-62703-017-5_18. PubMed: 23034760.
26. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14: 33–38. doi:10.1016/0263-7855(96)00018-5. PubMed: 8744570.
27. QtiPlot: data analysis and scientific visualization. Available: <http://soft.proindependent.com/qtiplot.html>. Accessed 2013 Jan 25.
28. Hasan NA, Choi SY, Eppinger M, Clark PW, Chen A et al. (2012) Genomic diversity of 2010 Haitian cholera outbreak strains. *Proc Natl Acad Sci U S A* 109: E2010–E2017. doi:10.1073/pnas.1207359109. PubMed: 22711841.
29. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402. doi:10.1093/nar/25.17.3389. PubMed: 9254694.
30. Johnson M, Zaretskaya I, Raytselis Y, Merezukh Y, McGinnis S et al. (2008) NCBI BLAST: a better web interface. *Nucleic Acids Res* 36: W5–W9. doi:10.1093/nar/gkn201. PubMed: 18440982.
31. Papadopoulos JS, Agarwala R (2007) COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 23: 1073–1079. PubMed: 17332019.
32. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948. PubMed: 17846036.
33. Giglio KM, Fong JC, Yildiz FH, Sondermann H (2013) Structural basis for biofilm formation via the *Vibrio cholerae* matrix protein RbmA. *J Bacteriol* 195: 3277–3286. doi:10.1128/JB.00374-13. PubMed: 23687270.
34. Schachter H, Brockhausen I (1992) Glycoconjugates. Composition, structure and function. New York: CRC Press. 696 pp.
35. Yildiz FH, Schoolnik GK (1999) *Vibrio cholerae* O1 El Tor: identification of a gene cluster required for the rugose colony type, exopolysaccharide production, chlorine resistance, and biofilm formation. *Proc Natl Acad Sci U S A* 96: 4028–4033. doi:10.1073/pnas.96.7.4028. PubMed: 10097157.
36. Ansari AA, Kenne L, Lindberg B, Gustafsson B, Holme T (1986) Structural studies of the O-antigen from *Vibrio cholerae* O:21. *Carbohydr Res* 150: 213–219. doi:10.1016/0008-6215(86)80017-9. PubMed: 2428496.
37. Chowdhury TA, Jansson P-E, Lindberg B, Lindberg J, Gustafsson B et al. (1991) Structural studies of the *Vibrio cholerae* O:3 O-antigen polysaccharide. *Carbohydr Res* 215: 303–314. doi:10.1016/0008-6215(91)84029-E.
38. Hermansson K, Jansson P-E, Holme T, Gustafsson B (1993) Structural studies of the *Vibrio cholerae* O:5 O-antigen polysaccharide. *Carbohydr Res* 248: 199–211. doi:10.1016/0008-6215(93)84127-R.
39. Knirel YA, Widmalm G, Senchenkova SN, Jansson PE, Weintraub A (1997) Structural studies on the short-chain lipopolysaccharide of *Vibrio cholerae* O139 Bengal. *Eur J Biochem / FEBS* 247: 402–410.
40. Hisatsune K, Kondo S, Isshiki Y, Iguchi T, Kawamata Y, et al. (1993) O-antigenic lipopolysaccharide of *Vibrio cholerae* O139 Bengal, a new epidemic strain for recent cholera in the Indian subcontinent. *Biochem Biophys Res Commun* 196: 1309–1315.
41. Kabir S, Ali S (1983) Characterization of surface properties of *Vibrio cholerae*. *Infect Immun* 39: 1048–1058. PubMed: 6840833.
42. Borlee BR, Goldman AD, Murakami K, Samudrala R, Wozniak DJ et al. (2010) *Pseudomonas aeruginosa* uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol Microbiol* 75: 827–842. doi:10.1111/j.1365-2958.2009.06991.x. PubMed: 20088866.
43. Maestre-Reyna M, Diderrich R, Veelders MS, Eulenburg G, Kalugin V et al. (2012) Structural basis for promiscuity and specificity during *Candida glabrata* invasion of host epithelia. *Proc Natl Acad Sci U S A* 109: 16864–16869. doi:10.1073/pnas.1207653109. PubMed: 23035251.
44. Veelders M, Brückner S, Ott D, Unverzagt C, Mösch H-U et al. (2010) Structural basis of flocculin-mediated social behavior in yeast. *Proc Natl Acad Sci U S A* 107: 22511–22516. doi:10.1073/pnas.1013210108. PubMed: 21149680.
45. Bezrukov SM, Krasilnikov OV, Yuldasheva LN, Berezhkovskii AM, Rodrigues CG (2004) Field-dependent effect of crown ether (18-crown-6) on ionic conductance of alpha-hemolysin channels. *Biophys J* 87: 3162–3171. doi:10.1529/biophysj.104.044453. PubMed: 15507690.
46. Isshiki Y, Haishima Y, Kondo S, Hisatsune K (1992) Serological cross-reaction between intact and chemically modified lipopolysaccharides of O1 *Vibrio cholerae* Inaba and non-O1 *V. cholerae* bio-serogroup Hakata. *Microbiol Immunol* 36: 1201–1205. doi:10.1111/j.1348-0421.1992.tb02123.x. PubMed: 1283438.