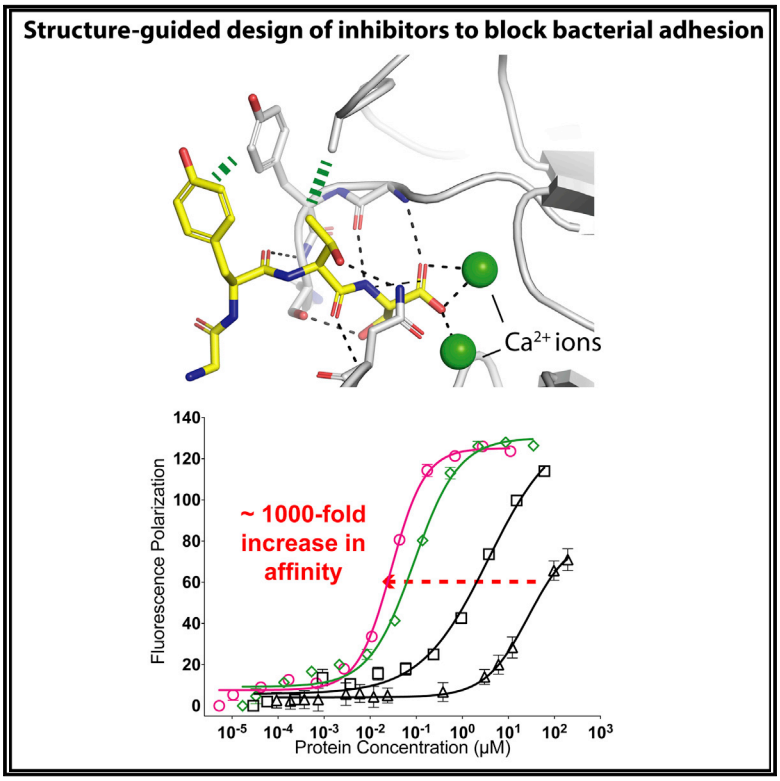
Molecular basis for inhibition of adhesin-mediated bacterial-host interactions through a peptide-binding domain

Cell Reports

Article

Molecular basis for inhibition of adhesin-mediated bacterial-host interactions through a peptide-binding domain

Graphical abstract



Authors

Shuaiqi Guo, Hossein Zahiri, Corey Stevens, ..., Luc Brunsveld, Ilja K. Voets, Peter L. Davies

Correspondence

peter.davies@queensu.ca

In brief

Guo et al. describe peptide inhibitors that can block bacterial adhesion to host cells. They further use a structure-guided approach to enhance the potency of these inhibitors 1,000-fold. Their study gives insight into how bacterial adhesion may be controlled to modulate biofilm formation and treat infections.

Highlights

- Short peptides that can block bacterial adhesion to host cells
- Structural basis for the protein-peptide (inhibitor) interactions
- Development of inhibitors with a 1,000-fold better potency

Guo et al., 2021, Cell Reports *37*, 110002
November 16, 2021
https://doi.org/10.1016/j.celrep.2021.110002

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Molecular basis for inhibition of adhesin-mediated bacterial-host interactions through a peptide-binding domain

Shuaiqi Guo,[1,2,3,5] Hossein Zahiri,[1] Corey Stevens,[1,6] Daniel C. Spaanderman,[3,4,5] Lech-Gustav Milroy,[3,5] Christian Ottmann,[3,5] Luc Brunsveld,[3,5] Ilja K. Voets,[2,4,5] and Peter L. Davies[1,7,*]

[1]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON K7L 3N6, Canada
[2]Laboratory of Self-Organizing Soft Matter, Eindhoven University of Technology PO Box 513, 5600 MB Eindhoven, the Netherlands
[3]Laboratory of Chemical Biology, Eindhoven University of Technology PO Box 513, 5600 MB Eindhoven, the Netherlands
[4]Laboratory of Macromolecular and Organic Chemistry, Eindhoven University of Technology PO Box 513, 5600 MB Eindhoven, the Netherlands
[5]Institute for Complex Molecular Systems, Eindhoven University of Technology PO Box 513, 5600 MB Eindhoven, the Netherlands
[6]Present address: Laboratoire des Polymères, Institut des Matériaux and Institut des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), Batiment MXD, Station 12, 1015 Lausanne, Switzerland
[7]Lead contact
*Correspondence: peter.davies@queensu.ca
https://doi.org/10.1016/j.celrep.2021.110002

SUMMARY

Infections typically begin with pathogens adhering to host cells. For bacteria, this adhesion can occur through specific ligand-binding domains. We identify a 20-kDa peptide-binding domain (PBD) in a 1.5-MDa RTX adhesin of a Gram-negative marine bacterium that colonizes diatoms. The crystal structure of this Ca^{2+}-dependent PBD suggests that it may bind the C termini of host cell-surface proteins. A systematic peptide library analysis reveals an optimal tripeptide sequence with 30-nM affinity for the PBD, and X-ray crystallography details its peptide-protein interactions. Binding of the PBD to the diatom partner of the bacteria can be inhibited or competed away by the peptide, providing a molecular basis for inhibiting bacterium-host interactions. We further show that this PBD is found in other bacteria, including human pathogens such as *Vibrio cholerae* and *Aeromonas veronii*. Here, we produce the PBD ortholog from *A. veronii* and demonstrate, using the same peptide inhibitor, how pathogens may be prevented from adhering to their hosts.

INTRODUCTION

Many microorganisms, including most bacteria found in nature, live in surface-attached multicellular communities called biofilms (Collins et al., 2020; Davey and O'toole, 2000; Teschler et al., 2015; Wozniak and Parsek, 2014; Zheng et al., 2021). During biofilm formation, free-swimming planktonic bacteria transition into a sessile mode by attaching to a surface (Wong et al., 2021). Next, surface-attached bacteria grow into microcolonies that may lead to the formation of mature biofilms in which cells are encased by a protective layer of extracellular polymeric substance (EPS) matrix comprising various proteins, polysaccharides, and extracellular DNAs (Davey and O'toole, 2000; Teschler et al., 2015; Wong et al., 2021). While biofilms of commensal bacteria can be protective for the host, those formed by pathogens are a leading cause of chronic infections in animals and plants (Davey and O'toole, 2000; Han et al., 2010; Hashem et al., 2019; Krachler and Orth, 2013; Lebeaux et al., 2014; Teschler et al., 2015).

Adhesins are bacterial cell-surface appendages required during biofilm formation (Collins et al., 2020; Guo et al., 2019b; Newell et al., 2011; Smith et al., 2018). During early stages, adhesins bind planktonic bacteria stably to various surfaces, tissues, and organisms. Subsequently, adhesins mediate cell-cell interactions between bacteria to form microcolonies and further contribute to the architecture of mature biofilms as they tether bacteria to the EPS matrices (Melia et al., 2021). In both Gram-negative and Gram-positive bacteria, adhesins are typically long, modular proteins with one terminus anchored to the bacterial surface, while the other end is extended out to interact with substrates such as the carbohydrates and proteins on host cell surfaces (Alonso-Caballero et al., 2021; Berne et al., 2015; Chang et al., 2019; Guo et al., 2017, 2018, , 2019b; Kitts et al., 2019; Klemm and Schembri, 2000; Sanchez et al., 2010; Smith et al., 2018). With structures of adhesin-ligand complexes starting to emerge (Moonens et al., 2016, 2018; Vance et al., 2019), researchers are beginning to elucidate the molecular basis of the interactions required for surface colonization that in many cases can lead to biofilm formation. To date, most of these studies have focused on the characterization of lectin-glycan interactions (Moonens et al., 2016; Vance et al., 2019; Wellens et al., 2008) and have resulted in some effective treatments

against bacterial infections (Krachler and Orth, 2013; Ofek et al., 2003). For example, the FimH adhesin of uropathogenic *Escherichia coli* has been successfully targeted by mannose analogs to help treat urinary tract infections (Han et al., 2010; Mydock-McGrane et al., 2017; Sauer et al., 2019; Spaulding et al., 2017; Totsika et al., 2013; Wellens et al., 2008). These studies validate the efficacy of the anti-adhesion strategy, which can be used as an alternative approach to treat bacterial infections without the excessive use of antibiotics.

Marinomonas primoryensis ice-binding protein (*Mp*IBP) is proving to be a useful model system for studying and intervening in adhesin-host interactions. *Mp*IBP is an exceptionally large (~1.5 MDa) repeats-in-toxin (RTX) adhesin found on the surface of its Antarctic bacterium (Guo et al., 2012, 2017). While its N-terminal region I (RI) is responsible for anchoring the adhesin to the outer membrane, its C-terminal ligand-binding RIII and RIV help the bacterium form symbiotic biofilms with diatoms on the underside of lake ice (Guo et al., 2017, 2019b). Specifically, RIV is the ice-binding segment (Garnham et al., 2008, 2011), and antibodies raised to it can prevent the bacterium from binding and colonizing ice (Bar Dolev et al., 2016). *Mp*IBP_RIII is responsible for diatom binding and contains 5 β sheet-rich domains (*Mp*IBP_RIII1–5). The most C-terminal of these, *Mp*IBP_RIII5, was identified as a homolog of a PA14 carbohydrate-binding lectin module. This domain (*Mp*PA14) was recombinantly produced and characterized (Guo et al., 2021). Glycan array analysis showed that it binds to branched polysaccharides such as fucoidans and glucans rich in terminal fucose and related sugar moieties. Consistent with these findings, the adhesion of *Mp*PA14 to diatom cells was selectively blocked by L-fucose. Bioinformatic analyses with BLASTp and the Phyre2 server identified *Mp*IBP_RIII1–4 domains as bacterial immunoglobulin (Big) domains, which are frequently positioned in tandem to function as extenders for ligand-binding domains. However, when this recombinant protein was analyzed by X-ray crystallography, we observed that the three C-terminal amino acid (aa) residues (threonine-proline-aspartate [TPD]) of *Mp*IBP_RIII1-4 were stably anchored in a ligand-binding cavity of a symmetry-related molecule within the crystal unit cell (Guo et al., 2017). This cavity resides at one end of the third domain (*Mp*IBP_RIII3) that folds as an oblong β sandwich. The solvent-exposed ligand-binding cavity into which the peptide protrudes has two strategically placed calcium ions to coordinate the C-terminal carboxyl group of its peptide ligand. *Mp*IBP_RIII1-4 was, therefore, collectively identified as a peptide-binding domain (*Mp*PBD) that may give its bacterium a second option to bind diatoms by interacting with cell-surface proteins, while the adjacent RIII5 binds cell-surface glycans.

Here, we show that the TPD peptide sequence binds *Mp*PBD in solution with an affinity of 26 μM. Using a structure-guided approach informed by co-crystallography with peptide ligands, we optimized the peptidyl sequence by iteratively screening a library of pentapeptides and obtained ligands that bound *Mp*PBD ~1,000-fold more strongly with affinities of ~30 nM. X-ray crystal structures of *Mp*PBD-ligand complexes revealed that the strong protein-peptide interactions originate from a combination of Ca^{2+}-dependent polar interactions and hydrophobic contacts. We further demonstrate that these short peptidyl ligands are not only effective at inhibiting the binding of *Mp*PBD to diatoms but also can block a similar PBD from an RTX adhesin produced by a human pathogen, *Aeromonas veronii*. Thus, results in this work give insight into how microbial adhesion can be disrupted through ligand-based antagonists.

RESULTS AND DISCUSSION

Initial ligand identification

Close inspection of the previously obtained *Mp*IBP_RIII1–4 structure (PDB: 5K8G) revealed the detailed protein-protein interactions involved in the self-association within the crystal. The C-terminal Thr-Pro-Asp (T517-P518-D519) of one molecule bound deep into the pocket of its symmetry-related mate in the unit cell (Figure 1A). The C-terminal aspartate residue was buried inside the ligand-binding pocket of the other molecule, with its terminal α-carboxyl oxygens binding calcium ions 1 and 2 (Ca1 and Ca2) via three ionic bonds (Figure 1B). The protein-protein interaction was further enhanced by hydrogen bonding between the carbonyl and amide groups of the TPD sequence and V293 and N333 from the calcium- and peptide-binding loops 1 and 2 (CPBL1 and -2) of *Mp*IBP_RIII3. The threonine residue at the pre-penultimate position (position 3) was the most solvent-exposed residue of the peptide, and the ambiguous electron density of its side chain was indicative of a high degree of flexibility. The extent and intimacy of this interaction strongly suggested that RIII1–4 is indeed a peptide-binding domain.

This binding interaction was a remarkably serendipitous discovery for several reasons. The C-terminal end of RIII1–4 is the result of splitting up *Mp*IBP into manageable sections for structural characterization, and it does not exist in nature because the whole adhesin is one long polypeptide chain spanning ~130 linked domains. Fortuitously, the C-terminal residue of this section of the adhesin was Asp, which is one of the optimal residues for fitting into the deepest part of the peptide-binding cleft. Also, RIII1–4 happened to crystallize with the symmetry-related molecules in the unit cell in such a way that one could bind its C-terminal end into the binding pocket of the other. Without this fortuitous crystallization, it is unlikely that the function of RIII1–4 would have been recognized from its structure alone. So good was the crystallographic dimer interaction, that this construct failed to yield crystals in the presence of optimal free peptides identified in this study, presumably due to its propensity to self-associate via the C-terminal TPD sequence. To circumvent this oligomerization issue, we designed a new *Mp*IBP_RIII1–4 construct by truncating the carboxyl end of the original protein 12 residues earlier at N507, rather than at D519 (Figures 1A and S1A, green and blue arrows). This truncated construct, referred to as *Mp*PBD, was used for the binding assays described below.

Given that a C-terminal carboxylic acid group appears to be key for the protein-protein interaction, we used fluorescence polarization (FP) to screen a small collection of 15 N-terminal fluorescein isothiocyanate (FITC)-labeled peptides that end in various C-terminal aa residues with a free α-carboxyl group (selected on the basis of availability within the laboratory). In the presence of ~80 μM *Mp*PBD, 2 FITC-labeled and phosphorylated peptides with sequences of IKARAS($_p$S)SPVILVGTHLD (peptide 14) and RHKKLMFK($_p$T)EGPDSD (peptide 15) produced

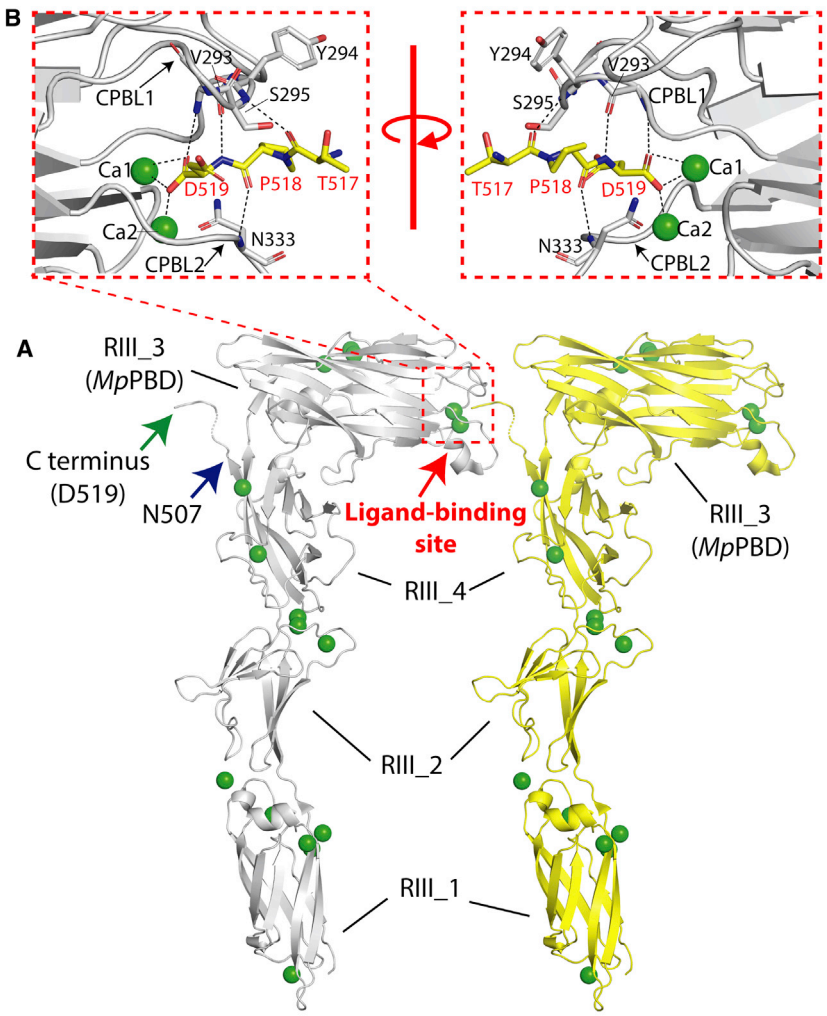


Figure 1. Self-association of *Mp*IBP_RIII1-4 (T1-D519) within the crystal unit cell

(A) *Mp*IBP_RIII1–4 molecule (gray) interacts with a neighboring symmetry mate (yellow) in the crystal via its C-terminal TPD sequence. The C-terminal D519 residue of the original *Mp*IBP_RIII1–4 construct is marked by a green arrow, while the C-terminal N507 residue of the truncated construct used in this study is marked by a blue arrow. The ligand-binding site is indicated by a red box of dashed lines.

(B) Magnified view of the ligand-binding site of *Mp*IBP_RIII1–4 within the crystal showing the atomic details at the protein-protein interaction interface. The right panel shows the interface from a view that is rotated ~180° around a vertical axis from the left. Polar interactions are indicated by black dashed lines. Carbon atoms of the bound TPD sequence are yellow, while those for its symmetry-related protein are gray. Oxygen atoms are red, nitrogen atoms are blue, and Ca^{2+} ions are shown as green spheres. Amino acid residues involved in protein-protein interactions are labeled and shown in stick representation.

See also Figure S1.

significant delta fluorescence polarization values, indicating binding of these peptides to *Mp*PBD (Figure 2A). None of the other 13 peptides appeared to bind *Mp*PBD as they produced negligible delta polarization values. FP with protein titrations further showed that in the presence of 2 mM Ca^{2+}, peptides 14 and 15 bound *Mp*PBD with half-maximal effective concentration (EC_{50}) values of 2.7 and 0.69 μM, respectively. Moreover, the presence of excess EDTA abolished the interaction between peptide and protein (Figure 2B), which validated the Ca^{2+} dependency of the protein stability and the protein-peptide interaction as demonstrated by the structural data. To further verify that the peptide-binding capability of *Mp*PBD was localized to *Mp*IBP_RIII3 as observed in X-ray crystallography, we produced a construct containing only *Mp*IBP_RIII3 (Q206-402G; Figure S1B). The single-domain *Mp*IBP_RIII3 construct bound peptide 15 with an EC_{50} value of 0.80 μM in the presence of 2 mM $CaCl_2$, but no binding was observed in the presence of excess EDTA (Figure S1C). These results are similar to those with the multi-domain *Mp*PBD construct (T1-N507), which verified that *Mp*IBP_RIII3 is the PBD. However, *Mp*IBP_RIII3 has a high propensity to precipitate during the Ni^{2+} affinity purification, which significantly diminished its yield compared to that of the *Mp*PBD construct. Therefore, *Mp*PBD was used for all of the biophysical assays, co-crystallization, and binding experiments described henceforth.

Interestingly, the three binding sequences TPD, peptide 14, and peptide 15 all ended in a C-terminal aspartate residue with a free α-carboxylic acid group. This was consistent with the structural data, which showed that the terminal carboxylic acid group of TPD directly bonded to the two calcium ions in the *Mp*PBD ligand-binding site. However, since the C-terminal aspartate side chain of *Mp*IBP_RIII1–4 had no direct interactions with its neighboring symmetry mate in the crystal, it was not clear how the aspartate residues in these three peptidyl sequences contributed to their binding to *Mp*PBD. To further validate the hypothesis that *Mp*PBD binds specifically to certain C-terminal aa sequences, peptides encompassing only the last 5 aas of the original *Mp*IBP_RIII1-4 construct (FITC-A_βDSTPD) and peptide 15 (FITC-A_βGPDSD) were synthesized and studied. The affinity of *Mp*PBD for these two short peptide ligands, together with two analogs (FITC-A_βDSTD and FITC-A_βGPDD), where the penultimate residues were deleted, was measured by FP (Figure 2C).

Unexpectedly, FITC-A_βDSTPD, which contains the C-terminal TPD binding sequence originally identified in the crystal structure of *Mp*IBP_RIII1–4, showed the weakest interaction to *Mp*PBD out of this series, with an EC_{50} of ~27 μM. In contrast, the peptide FITC-A_βDSTD bound at least 200-fold stronger with an EC_{50} of 0.11 μM (Figure 2C). Similarly,

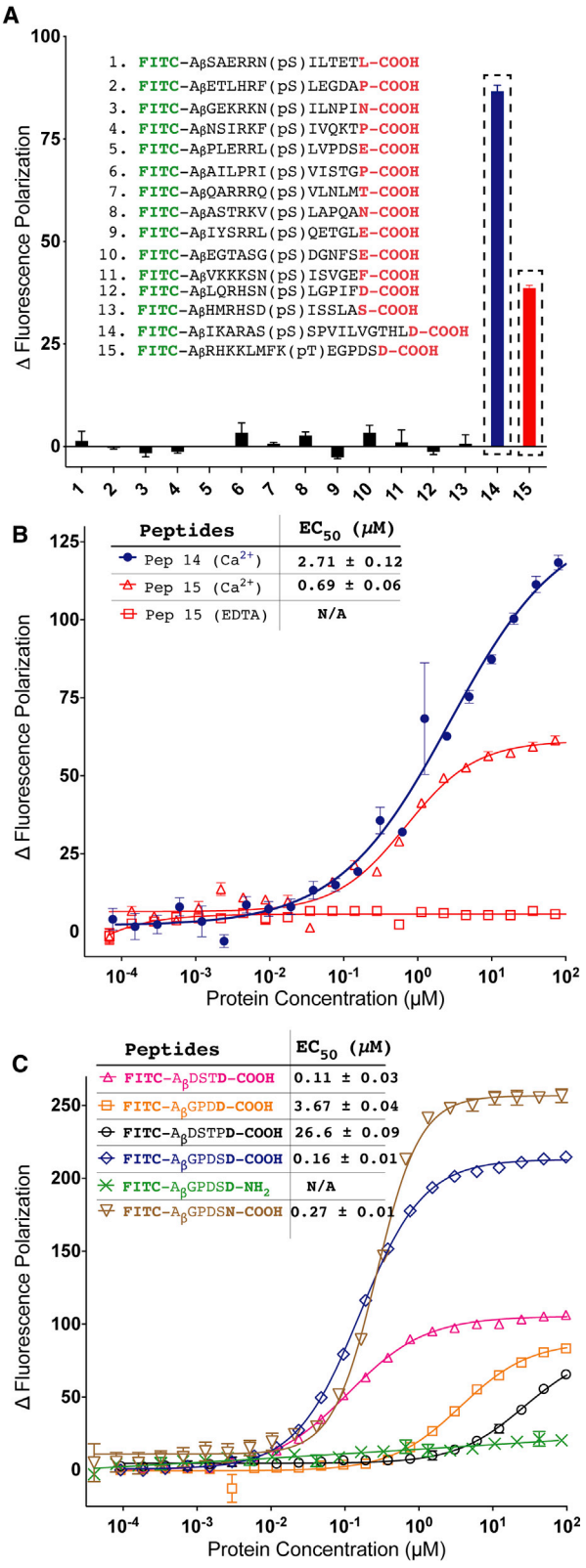


Figure 2. Identification of initial peptidyl ligands of *Mp*PBD

(A) Screening of a collection of 15 FITC-labeled peptides by fluorescence polarization (FP) with *Mp*PBD at a concentration of ~80 μM. Background FP was subtracted from all of the values to result in ΔFP values on the y axis. Peptide sequences are displayed. Data bars for peptides 14 (blue) and 15 (red) are indicated within boxes of dashed lines.

(B) FP assay of peptides 14 (blue closed circles) and 15 (red open triangles) from (A) with titration of *Mp*PBD in the presence of 2 mM $CaCl_2$ and in the presence of excess 2 mM EDTA (red open squares).

(C) FP assay to assess the binding of 6 different FITC-labeled peptides to *Mp*PBD. The mean of 3 experiments was plotted.

Some of the SD error bars are smaller than the data point symbols in (B) and (C). The calculated EC_{50} values and the standard errors (log EC_{50}) for each experiment are shown in (B) and (C).

FITC-A_βGPDSD bound with an EC_{50} of 0.16 μM, which is roughly 4-fold stronger than its longer, phosphorylated counterpart (FITC-RHKKLMFK$_p$TEGPDSD, EC_{50} = 0.69 μM), validating the importance of the C-terminal residues for binding *Mp*PBD. Furthermore, FITC-A_βGPDSD bound the protein 20-fold stronger than its analog FITC-A_βGPDD.

To further ascertain the chemical constituents of the peptide C terminus responsible for binding *Mp*PBD, two additional analogs of the peptide FITC-A_βGPDSD were synthesized. When the C-terminal aspartate was replaced by an asparagine, a residue with a similar length side chain but different chemistry, the affinity of the peptide FITC-A_βGPDSN for *Mp*PBD fell by 40% to an EC_{50} of 0.27 μM (Figure 2C). However, the substitution of the terminal α-carboxylic acid group of FITC-A_βGPDSD by a C-terminal carboxamide group (FITC-A_βGPDSD-NH_2) abolished the protein-peptide interaction altogether (Figure 2C). These results demonstrated the crucial importance of the ionic interaction between the terminal carboxylic acid group and the two Ca^{2+} ions in the ligand-binding site of *Mp*PBD. The electronegative side chain carboxylic acid group of the peptide C-terminal aspartic acid residue bound *Mp*PBD more tightly compared to that of the side chain amide of the peptide bearing an asparagine at the same position. Moreover, the FP results indicate that the peptide aas at the penultimate position (position 2) have a significant impact on the peptide-protein interactions. The strong binders FITC-A_βGPDSD and FITC-A_βDSTD contain the small polar residues of either threonine or serine at this position. The intermediate binder FITC-A_βGPDD has an aspartate, while the weakest binder FITC-A_βDSTPD has a proline at the same position. Although the proline is involved in main chain hydrogen bonding with V293 and N333 of the CPBL1 and CPBL2, respectively, its rigid 5-membered imine ring is unable to interact with the protein (Figure 1B). As for FITC-A_βGPDD, the negative charge of the aspartate side chain at position 2 may be disruptive for the peptide-protein interaction. We therefore sought to acquire detailed structural information to explain the high affinities of FITC-A_βGPDSD and FITC-A_βDSTD for *Mp*PBD.

Structural basis for peptide-MpPBD interactions

To elucidate the atomic details of the peptide-protein interactions, we solved the X-ray crystal structures of *Mp*PBD in complex with the peptides FITC-A_βDSTD and FITC-A_βGPDSD to 1.9- and 2.2-Å resolution, respectively (Table S1). Clear electron densities were observed for the peptide amino acids in the last three positions (Figure 3), while those at the position immediately prior at the N-terminal end appeared to be ambiguous.

At position 2 of FITC-A_βGPDSD, serine has its side chain hydroxyl group hydrogen bonded with the side chain amide of

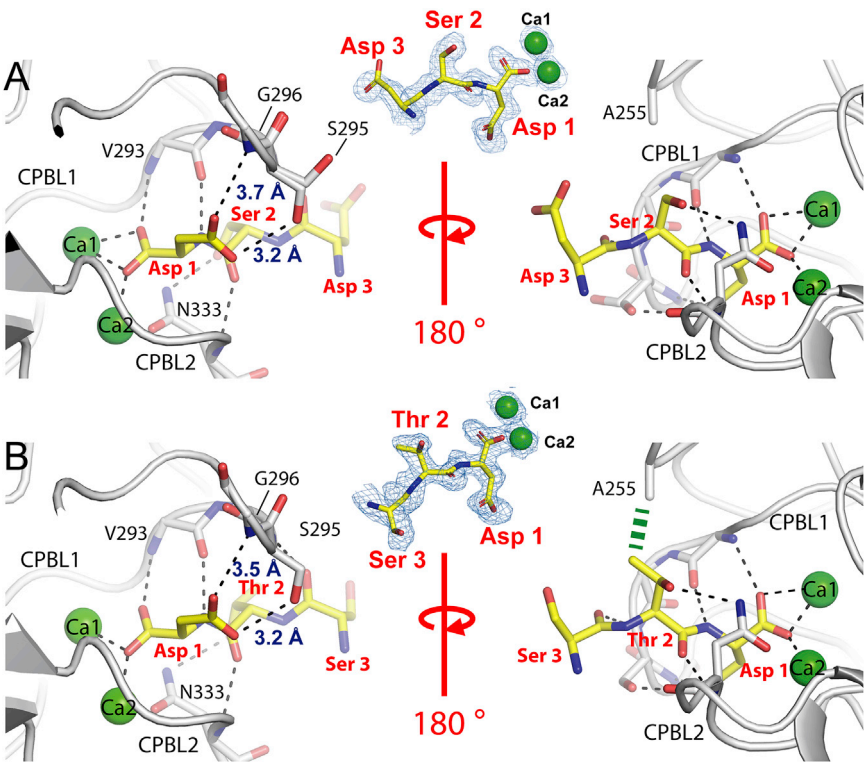


Figure 3. Protein-peptide binding interfaces revealed by X-ray crystallography

Magnified views of the *Mp*PBD ligand-binding site interacting with peptides that end with sequences of DSD (A) and DSTD (B), respectively. The $2F_o–F_c$ electron density maps around the peptide residues and Ca1 and Ca2 are shown as blue mesh (contoured at 1 σ). The right panels in (A) and (B) are views that are rotated ~180° from the left panels. Polar interactions are indicated by black dashed lines. Carbon atoms of the peptides are yellow. Oxygen atoms are red, nitrogen atoms are blue, and Ca^{2+} ions are shown as green spheres.

N333 on CPBL2, while the main chains of these two residues hydrogen bond via their carbonyl and amide groups (Figure 3A). The observed hydrogen bonding interactions help to explain why FITC-A_βGPDSD has a 20-fold higher affinity for *Mp*PBD than that of FITC-A_βGPDD, with the only difference between the two peptides being the presence or absence of serine at position 2. X-ray crystallography indicates that the larger aspartate residue at position 2 would clash with N333 and destabilize the protein-peptide interactions. In comparison, FITC-A_βDSTD has a similar binding mode to FITC-A_βGPDSD, with the threonine side chain hydroxyl hydrogen bonded with the amide of N333 side chain at the edge of the peptide-binding cavity. In addition, the threonine side chain methyl group is involved in hydrophobic contact with residue A255 of the protein (Figure 3B). This helps restrain the free rotation of the threonine side chain hydroxyl, locking it into a favorable conformation for polar interactions. The observed additional hydrophobic interactions may help explain why FITC-A_βDSTD bound *Mp*PBD with slightly higher affinity than did FITC-A_βGPDSD.

Interestingly, the gain of hydrogen bonding interactions at the peptide 2nd position had an impact on how the C-terminal aspartate of the peptides bound *Mp*PBD. The aspartate of the TPD sequence observed in the ligand-binding cavity showed no interaction between its side chain γ-carboxylic acid group and the protein. In contrast, when FITC-A_βDSTD and FITC-A_βGPDSD were complexed with *Mp*PBD, the side chain hydroxyl of S295 within CPBL1 pointed downward to hydrogen bond with the γ-carbonyl of the peptide aspartate, which helped G296 to interact with the γ-hydroxyl of the same peptide aspartate via its main chain amide (Figures 3A and 3B, left panels). Taken together with the added interactions from the peptide position 2, the structural data gave a molecular explanation of why FITC-A_βDSTD and FITC-A_βGPDSD bound *Mp*PBD with substantially higher affinity than their analogs, FITC-A_βDSTPD and FITC-A_βGPDD, respectively.

Structure-guided screening identified ligands with low nanomolar affinity

Guided by the structural insight that the C-terminal residues are crucial for the peptide-protein interactions, we next systematically screened libraries of pentapeptides by FP to identify stronger ligands for *Mp*PBD.

The first round of screening involved 20 pentapeptides with a consensus sequence of FITC-AGAGX in which the first four amino acids of the peptides were alternating alanine and glycine residues. The ultimate (C-terminal) residue "X" represented 1 of the 20 naturally occurring amino acids (Figures 4A and S2). Consistent with the results shown above, the peptide with an aspartate at the C-terminal position (FITC-AGAGD) bound *Mp*PBD the tightest, with a moderate EC_{50} of 3.2 μM. This was followed by those peptides with large hydrophobic side chains at position 1, such as FITC-AGAGI, FITC-AGAGY, and FITC-AGAGF, which produced EC_{50} values of 4.4, 4.6, and 7.2 μM, respectively, when bound to *Mp*PBD (Figure S2; Table S2). Most of the other peptides bound *Mp*PBD with an affinity in the 10-μM range, while those with basic side chains (FITC-AGAGK and FITC-AGAGR) had negligible interaction with *Mp*PBD.

Having identified FITC-AGAGD as the strongest ligand for *Mp*PBD in the first round of screening, we proceeded with a second set of 20 pentapeptides that had a consensus sequence of FITC-AGAXD. All of the peptides except for AGAPD (EC_{50} = 4.6 μM) bound more strongly than FITC-AGAGD (EC_{50} = 1.5 μM). Consistent with the structural studies reported earlier (Figure 1B), neither proline nor glycine at position 2 can have side chain interactions with *Mp*PBD, explaining their lower affinity for the protein. Furthermore, the peptide with threonine in the second position (FITC-AGATD) has the highest affinity for

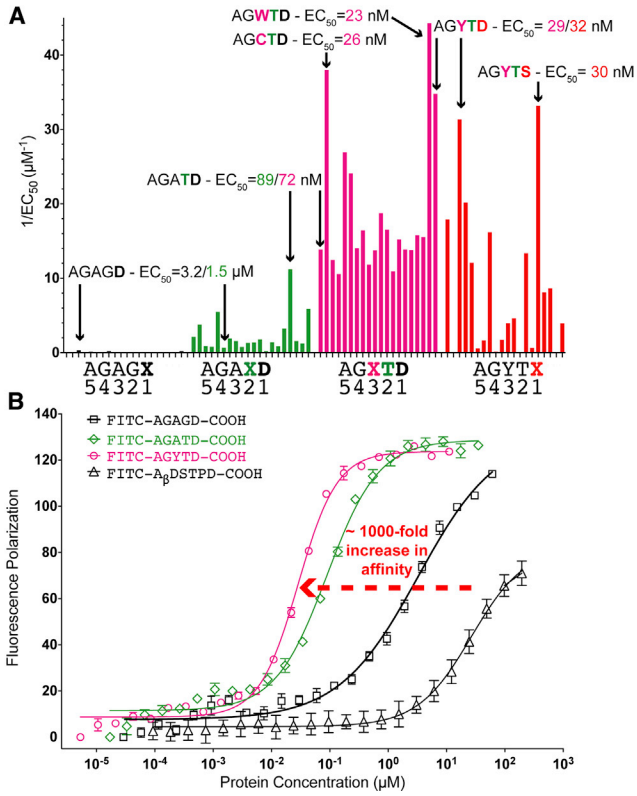


Figure 4. Structure-guided optimization of peptidyl ligands of *Mp*PBD

(A) Overview of the binding of the 80 FITC-labeled pentapeptides to *Mp*PBD in 4 rounds of screening. Data are plotted as average $1/EC_{50}$ values calculated from FP assays of each individual FITC-labeled peptide (in 3 replicates). All EC_{50} values are listed in Tables S2 and S3, and the corresponding FP titration plots are shown in Figure S2. Data bars for peptides in the first round of screening are black, while those for 2nd, 3rd and 4th rounds are green, magenta, and red, respectively. The strongest binders from each round of screening are marked by arrows and their calculated EC_{50} values are indicated.

(B) FP titration plots of representative peptides showing the progressive enhancement of their affinity for *Mp*PBD.

*Mp*PBD, with an EC_{50} of 89 nM (Figures 4A, green and S2; Table S2), which is up to 35-fold tighter than those of FITC-AGAGD. Peptides with aromatic residues at the same position, including FITC-AGAFD and FITC-AGAYD, bound slightly weaker than did AGATD, with EC_{50} values of 180 and 170 nM, respectively. Consistent with the structural characterization of FITC-A_βGPDSD and FITC-A_βDSTD (Figure 3), FITC-AGATD bound *Mp*PBD 3-fold more strongly than did FITC-AGASD (EC_{50} = 310 nM), validating the importance of the threonine methyl group in restraining the hydroxyl group to a favorable conformation for interacting with N333.

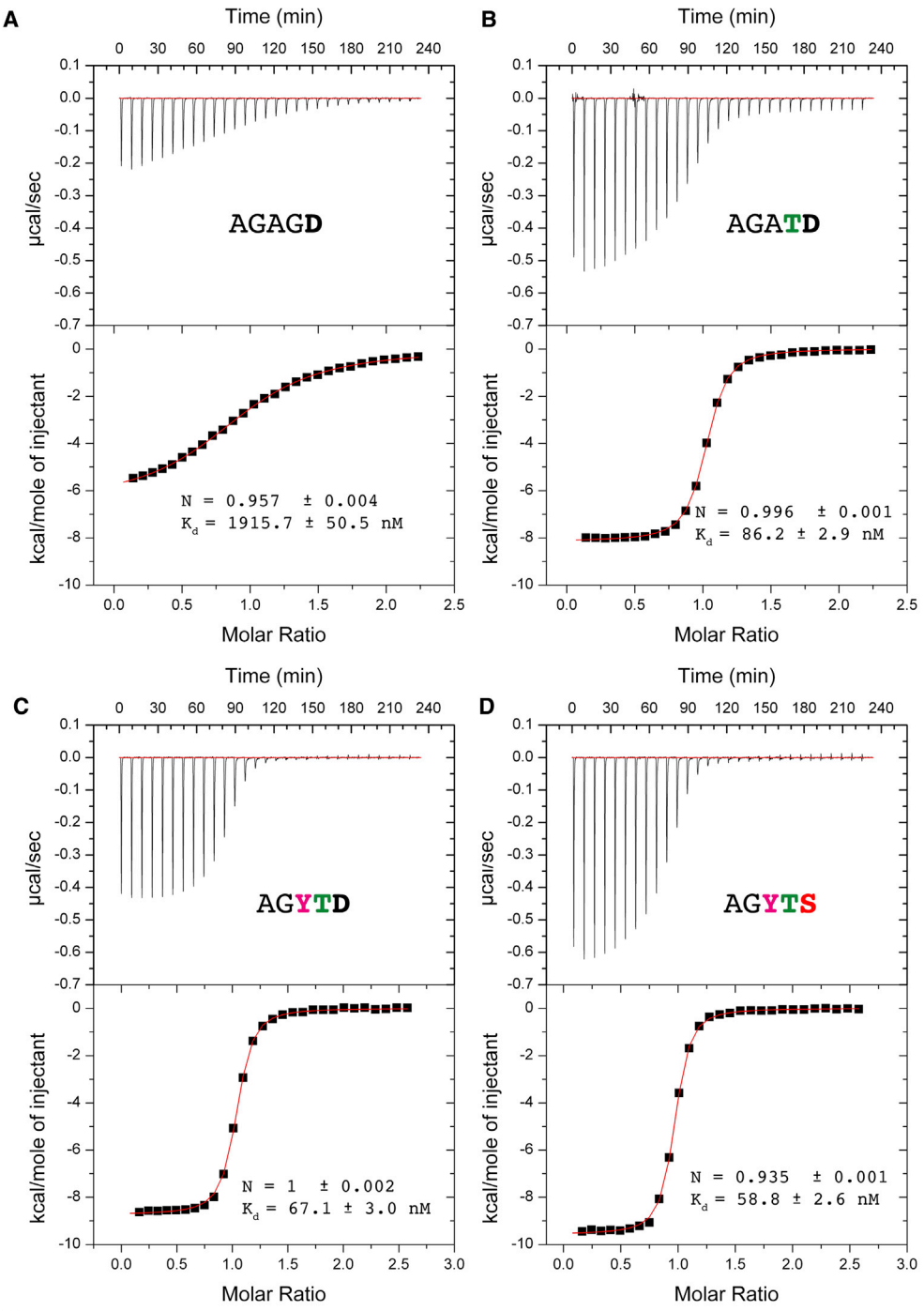
With the last two C-terminal residues defined as Thr and Asp, we proceeded to screen for the optimal aa in the 3rd position from the C terminus (FITC-AGXTD). While the majority of the 20 peptides of FITC-AGXTD bound *Mp*PBD more strongly than did FITC-AGATD (72 nM in this round of screening), with affinity in the nanomolar range, the three with aromatic side chains and cysteine stood out. In particular, FITC-AGYTD, FITC-AGWTD, and FITC-AGCTD produced EC_{50} values of 29, 23, and 26 nM (Figure 4A, magenta; Table S3; Figure S2), respectively. We reason that these aas with hydrophobic side chains at the 3rd position likely help the solvent-exposed residue interact with the protein via hydrophobic interactions, contributing to their higher affinity for *Mp*PBD.

Having demonstrated that residues at the 2nd and 3rd positions of the peptides have an impact on how the C-terminal aa interacted with *Mp*PBD, we performed a final round of screening, with a consensus sequence of FITC-AGYTX (Figure 4A, red). The reasons for selecting the tyrosine-containing peptide sequence over the tryptophan- or the cysteine-containing peptide include that FITC-AGYTD is more soluble than FITC-AGWTD and that FITC-AGYTD is not subject to dimerization the way that FITC-AGCTD can be linked by cysteine-dependent disulfide formation, which may confound the results of the binding studies. The results of the screening revealed that all but one of the FITC-AGYTX peptides bound *Mp*PBD more weakly than did FITC-AGYTD, with their EC_{50} values ranging from high nanomolar to micromolar concentrations. This validated the importance of the C-terminal aspartate residue for the protein-peptide interaction. The only peptide with a comparable affinity to FITC-AGYTD is FITC-AGYTS, which produced a calculated EC_{50} of 30 nM (Figures 4A, red, and S2; Table S3). Given their similar affinities for *Mp*PBD, AGYTD and AGYTS were considered two of the optimal peptidyl ligands, which bound *Mp*PBD roughly 1,000-fold tighter than the initially identified TPD binding sequence (Figure 4B).

Isothermal titration calorimetry validated the nanomolar affinity of MpPBD ligands

To further validate the peptidyl ligands identified by FP, we used isothermal titration calorimetry (ITC) to directly measure the interactions between unlabeled peptides (i.e., lacking FITC labels) and *Mp*PBD. ITC measurements of *Mp*PBD (~20 µM) titrated with the 4 different ligands (~200 µM) of AGAGD, AGATD, AGYTD, and AGYTS yielded sigmoidal-shaped curves with calculated stoichiometry values (N) close to 1 (Figures 5A–5D). This was consistent with the X-ray crystallography data that demonstrated *Mp*PBD has only one ligand-binding site. The interaction of the moderate binder, AGAGD, with *Mp*PBD showed a gradually transitioning sigmoidal curve, indicating a relatively slow saturation rate of the *Mp*PBD ligand-binding sites. This resulted in a calculated K_d of 1.9 µM, which is comparable to the EC_{50} values of FITC-AGAGD obtained from the FP experiments (1.5 or 3.2 µM). The result suggested that the FITC label did not have a significant impact on the peptide-protein interactions. The three strong binders AGATD, ATYTD, and ATYTS resulted in sigmoidal-shaped ITC curves with much steeper transitions. The 3 peptides had calculated K_d values of 86, 67, and 59 nM, respectively. Thus, the ITC results of the representative peptides showed the same trend in binding *Mp*PBD as did FP, with comparable K_d and EC_{50} values. Taken together, binding studies by FP and ITC have identified short peptidyl ligands with nanomolar affinities to *Mp*PBD (Figure 5E).

ITC indicated that significantly larger negative enthalpic (ΔH) contributions were involved in the binding of the potent ligands AGATD, AGYTD, and AGYTS to *Mp*PBD than for the moderate



E

Peptides (FITC-labelled or unlabelled)	Enthalpy ΔH (cal/mol)	Entropy ΔS (cal/mol/K)	Free energy ΔG (ΔH - TΔS)	K_d by ITC (nM)	EC_{50} by FP (nM)	
(FITC-)AGYTS	-9565 ± 26.2	1.53	-10028.8	58.8 ± 2.6	30	~ 1,000x
(FITC-)AGYTD	-8727 ± 24.9	4.03	-9948.7	67.1 ± 3.0	29/32	
(FITC-)AGATD	-8134 ± 18.0	5.49	-9798.3	86.2 ± 2.9	89/72	
(FITC-)AGAGD	-6448 ± 42.7	4.89	-7930.4	1915 ± 50.5	3200/1500	
(FITC-)AβDSTPD	N/A	N/A	N/A	N/A	~26600	

(legend on next page)

binder AGAGD (Figure 5E). In contrast, the positive entropic contributions (ΔS) were smaller for AGYTD and AGYTS when compared to AGAGD. The calculated thermodynamic profiles suggested that the enhancement of the peptide-protein interaction is primarily a result of a gain in polar interactions compared to that of a hydrophobic effect. This was supported by the structural comparison between the protein-peptide complex structures, which showed threonine or serine at peptide position 2 gained side chain hydrogen bonding interactions with the protein compared to that of proline or glycine at the same position (Figure 3 versus Figure 1). However, additional high-resolution structural information was required to elucidate the basis for the stronger interactions between *Mp*PBD and its peptidyl ligands AGYTD and AGYTS.

Molecular basis for potent binding by MpPBD ligands

To further reveal the molecular basis for the potent *Mp*PBD ligands, we solved X-ray crystal structures of *Mp*PBD in complex with the peptides AGYTS and AGYTD to resolutions of 1.6 and 1.8 Å, respectively (Figures 6 and S3; Table S1). Consistent with the other *Mp*PBD-peptide complexes, AGYTD and AGYTS had their terminal α-carboxyl group in contact with Ca1 and Ca2 via 3 ionic bonds with average lengths of ∼2.5 Å (Figure 6; Table S4). The threonine residues in position 2 had their hydroxyl group bond with the side chain amide of N333, as seen with FITC-A_{β}DSTD and FITC-A_{β}GPDSD (Figure 6). However, the YTD- and YTS-containing peptides bound *Mp*PBD with more hydrogen bonds of shorter lengths on average compared to the weaker binders (Table S4), which is consistent with the large negative enthalpic contributions for the binding indicated by ITC. For example, the side chain of the terminal aspartate of AGYTD interacted with *Mp*PBD S295 and G296, with bond lengths of 2.8 and 3 Å, respectively (Figure 6A; Table S4), while those at the same position for FITC-A_{β}GPDSD and FITC-A_{β}DSTD were at 3.2 and 3.7/3.5 Å, respectively (Figure 3). In contrast, the initially identified binding sequence that ends with TPD lacked these two bonds when bound by *Mp*PBD (Figure 1B; Table S4).

The tyrosine residues at position 3 of the YTD- or YTS-containing peptides played a key role in their tight interactions with *Mp*PBD. The aryl side chain of the solvent-exposed peptide tyrosine appears to pack against Y294 of *Mp*PBD, which helps CPBLs to clench the peptide more tightly in the ligand-binding cavity (Figures 6A and 6B). The hydrogen bond lengths between the peptide tyrosine main chain oxygen (Tyr3-O) and protein S295 main chain nitrogen (S295-N) is 3 Å for AGYTD and 2.8 Å for AGYTS, which are shorter than those at the same position for FITC-A_{β}DSTD and FITC-A_{β}GPDSD (3.1 and 3.3 Å, respectively) (Table S4). These subtle structural differences underlined the molecular basis for the potent binding to *Mp*PBD by the peptides that end with YTD or YTS compared to the weaker ligands (Figures 6C and 6D).

The presence of a tyrosine residue at position 3 of AGYTS and AGYTD helped strengthen the peptide-protein interactions. This correlated with the results that other FITC-AGXTD peptides with aromatic residues at the 3rd position were also strong ligands for *Mp*PBD (Figure 4A; Table S3). These atomic details explain why our systematic screening approach in conjunction with structure-guided ligand analysis was extremely effective in obtaining potent ligands with nanomolar affinities for *Mp*PBD.

AGYTD blocks MpPBD binding to diatoms

Having identified optimal peptidyl ligands for *Mp*PBD, we next tested their potential as antagonists to disrupt the protein-mediated bacteria-diatom interactions that led to the discovery and characterization of this ligand-binding domain (Guo et al., 2017). Here, we tested whether AGYTD can block fluorescently labeled *Mp*PBD from binding to the Antarctic marine diatom *Chaetoceros neogracile* (Gwak et al., 2010). The porous silica cell wall (frustule) of *C. neogracile* is rectangular in shape, with a length of roughly 10 μm and a width of 3–4 μm, with 1–4 projections protruding from the corners (Figures 7 and S4). The binding of tetramethylrhodamine isothiocyanate (TRITC)-labeled *Mp*PBD to *C. neogracile* resulted in red fluorescence evenly distributed around the cell centrally located inside of the diatom frustule (Figure 7B). At concentrations of ≥37.5 μM, AGYTD was extremely effective at blocking the accumulation of the protein on the diatom, as the peptide outcompeted the cell surface proteins for binding TRITC-labeled *Mp*PBD and displaced ∼95% of the fluorescent signal (Figures 7C and 7D, black bars). The inhibitory effect of AGYTD fell off by ∼30% at 3.75 μM but was still more potent compared to the weaker *Mp*PBD ligand, AGAGD, even at a 100-fold higher concentration of 375 μM (Figure 7D, black and blue bars). While 37.5 μM AGAGD had a minimal effect on the binding of *Mp*PBD to diatoms, the potency of AGYTD diminished in the nanomolar concentrations. These results indicated that the effective inhibitory concentrations of the *Mp*PBD peptide ligands are significantly higher than the K_d values calculated for the protein-peptide interactions. This is likely due to the multivalent effect originating from the presence of many nearby protein binding sites on the diatom cell membrane. We have observed a similar phenomenon when using fucose to compete off fluorescently tagged RIII5 lectin from binding to surface glycans on the same diatom (Guo et al., 2021). However, the lectin-glycan interactions are significantly weaker and require millimolar concentrations of ligand antagonists, whereas the peptide inhibitors function in the micromolar range.

Remarkably, AGYTD proved to be a potent antagonist that can displace *Mp*PBD pre-bound to the diatoms. The addition of 37.5 μM AGYTD resulted in the dissociation of roughly 70% of the TRITC-labeled *Mp*PBD from the diatoms (Figure 7D, gray bars). Although the peptide was more effective at the same concentration when used as a prophylactic (Figure 7D, black bar),

Figure 5. Isothermal titration calorimetry results of the binding of 4 unlabeled peptidyl ligands to *Mp*PBD

(A–D) Four unlabeled peptidyl ligands: AGAGD (A), AGATD (B), AGYTD (C), and AGYTS (D) to *Mp*PBD.

(E) Table showing the thermodynamic parameters of the binding of the 4 unlabeled peptidyl ligands to *Mp*PBD calculated from ITC. K_d by ITC and EC_{50} by FP for various peptides are shown. Red arrow compares the binding of the initially identified ligand that ends with TPD with the optimal sequences that end with YTS and YTD obtained from the structure-guided ligand optimization.

See also Figure S6.

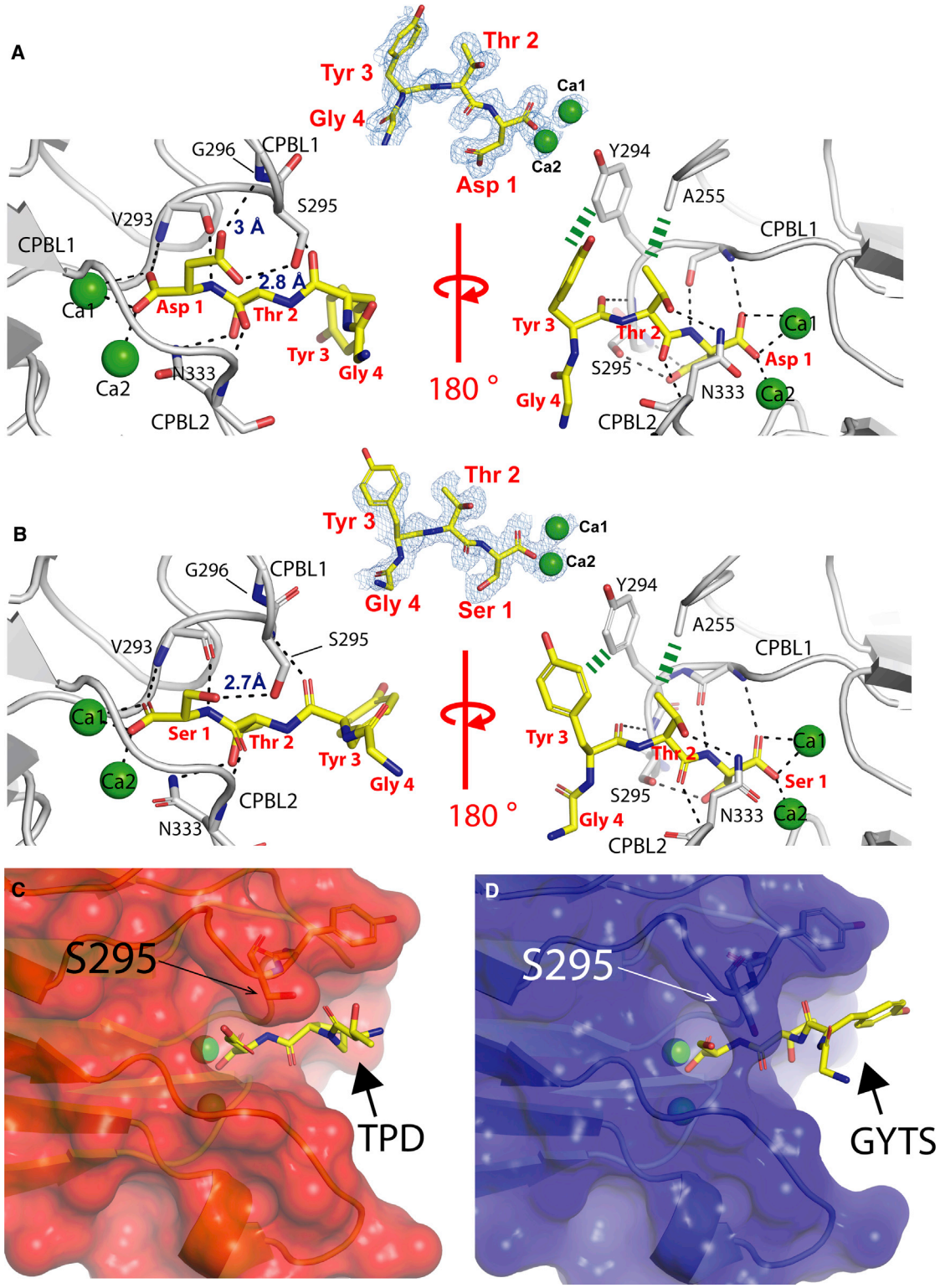


Figure 6. Atomic details of the interactions between *Mp*PBD and its nanomolar-affinity peptidyl ligands revealed by X-ray crystallography

(A and B) Nanomolar-affinity peptidyl ligands AGYTD (A) and AGYTS (B). The color scheme is the same as in Figure 3. Thick, green-dashed lines indicate hydrophobic contact between the protein and peptide tyrosine residues.

(legend continued on next page)

this important result demonstrated the potential of the peptide antagonists in disrupting preexisting bacteria-host interactions, like those involved in forming early-stage biofilms.

Ligand antagonist cocktail blocks adhesin-host interactions

We have previously shown that *Marinomonas primoryenesis* binds to diatoms via both the PBD and a sugar-binding PA14 domain (*Mp*IBP_RIII5) (Guo et al., 2017). We demonstrated that fucose is a potent antagonist that can block *Mp*PA14 binding to diatoms at low millimolar concentrations (Guo et al., 2021). Having more than one adhesion domain to the same host cell can enhance the chance of bacteria colonizing a specific niche. Given that we have identified specific ligands for both the *Mp*PBD and *Mp*PA14, we tested whether the combination of peptide and sugar inhibitors was more potent in blocking adhesin-host interactions than either ligand alone.

The diatoms bound RIII1–5 tagged with an N-terminal GFP (GFP-RIII1–5) in the absence of competing ligands and showed fluorescence roughly 20 times brighter than their intrinsic diatom autofluorescence (Figures 7E, 7F, and 7J). The addition of 5 mM fucose alone reduced the fluorescence signal produced by GFP-RIII1–5 binding by >30%, while the addition of 37.5 μM peptide AGYTD alone competed off >40% of GFP-RIII1–5 (Figures 7G, 7H, and 7J). However, the combination of both ligands reduced GFP-RIII1–5 binding by ~90%, which is clearly more effective than either ligand by themselves (Figures 7I and 7J). These experiments demonstrated that the effectiveness of a targeted anti-adhesion strategy requires detailed molecular knowledge of all of the ligand-binding domains in the adhesin.

MpPBD homologs are found in pathogens

Inhibition of pathogenic bacterial adhesion to human cells has yielded promising results in combating infections (Cusumano et al., 2011; Huebinger et al., 2020; Krachler and Orth, 2013; Mydock-McGrane et al., 2017; Solanki et al., 2018; Totsika et al., 2013; Wellens et al., 2008). The anti-adhesion strategy works by preventing and clearing the accumulation of bacteria at the sites of infections. In contrast to the conventional small-molecule antibiotics, the adhesin antagonists do not kill the bacteria; thus, they are less likely to raise resistance and are not required to penetrate the cells. These key attributes justified the pursuit of adhesin modulators to disrupt protein-mediated bacterial adhesion to hosts to treat bacterial infections. Homologs of *Mp*PBD with a conserved ligand-binding site are found in human pathogens such as *Vibrio cholerae* (Guo et al., 2017; Guo et al., 2019b; Syed et al., 2009) and other pathogens such as leech-symbiont *Aeromonas veronii*, which can cause serious infections ranging from diarrhea to wound infections and sepsis (Hickman-Brenner et al., 1987; Whitaker et al., 2009), and the flesh-eating bacterium *Vibrio vulnificus* (Hollis et al., 1976) (Figure S5).

As a proof of concept for the anti-adhesion strategy that targets PBD-mediated microbial-host interactions, we produced and purified the PBD from *A. veronii* (*Av*PBD), which is part of an extremely large adhesin (6,090 aas; GenBank: WP_181926999). Although the Antarctic diatom *C. neogracile* is unlikely to be a natural host for *A. veronii*, we reasoned that *Av*PBD should bind *C. neogracile* since it has an identical fold to *Mp*PBD (63% sequence identity), with key residues conserved in their ligand-binding sites (Figure S5). Consistent with our structural and binding analyses, FITC-labeled *Av*PBD bound to the cell membrane of *C. neogracile* inside the frustule in the same manner as *Mp*PBD, and this binding was more effectively blocked by 37.5 μM AGYTD than by AGAGD (Figures 7K–7N). This key result demonstrates that our structure-aided approach to identifying and optimizing peptide antagonists can be broadly applied to target harmful adhesion by pathogens.

One major challenge for the development of anti-adhesin strategies is a lack of starting points for ligand identification. Here, a serendipitously discovered peptidyl ligand TPD has served as a lead for the development of a class of short peptides that bound *Mp*PBD with 1,000-fold higher affinity. This was largely achieved by systematically screening small libraries totaling 80 peptides that informed on the optimal residue in each of three terminal positions and was confirmed by X-ray co-crystallography. Remarkably, these high-affinity ligands can serve as antagonists to disrupt *Mp*PBD from binding to the diatom cells that form symbiotic associations with the Antarctic bacterium. We further showed that an *Mp*PBD homolog, *Av*PBD, from the human pathogen *A. veronii* can be blocked from binding to diatoms cells in the same manner. This proof of principle has established the value of antagonists in disrupting adhesion involved in bacterial infections. With the emergence of antibiotic-resistant pathogens, the present study gives insight into how microbes may be controlled through the modulation of adhesion to their hosts.

Limitations of the study

Although we demonstrate that peptide antagonists can block the PBD from binding to the diatom, *C. neogracile*, we have not yet identified the protein target(s) on the host species. As shown in Figure 4, the PBD can bind to a range of different C-terminal sequences with different affinities, which may make identification of the ligand on diatom cells difficult to discern. Since PBD homologs are found in adhesins from a variety of bacteria, it will be worthwhile to identify their protein ligands on various hosts to see whether they are related proteins that can be blocked by the same antagonists. This information will be relevant to the prevention of bacterial colonization and treatment of infections.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY

(C and D) The plasticity of the *Mp*PBD ligand-binding site is illustrated by the differences in complexing peptides that end with TPD (C) and YTS (D). Residue S295, which pointed to the side in (C) and downward in (D), is indicated by black arrows. The protein is shown in surface representation, while the peptide is shown in stick representation.
See also Figure S3.

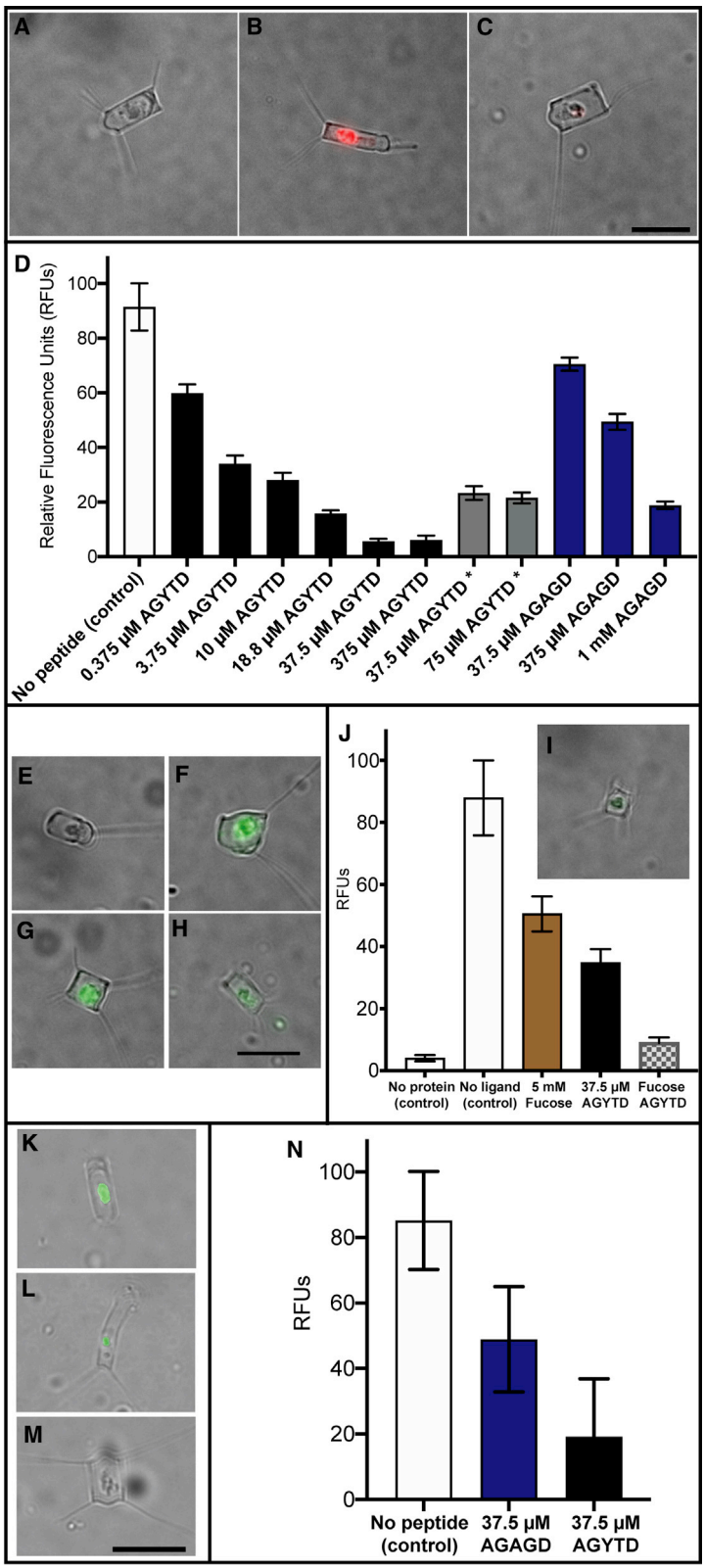


Figure 7. Inhibition of binding of *Mp*PBD to diatoms by competing peptides

(A–C) Representative images showing (A) an untreated diatom, *C. neogracile*; (B) a diatom treated with TRITC-labeled *Mp*PBD; and (C) a diatom treated with TRITC-labeled *Mp*PBD in the presence of 37.5 μM AGYTD. The scale bar in (C) indicates 10 μm (same magnification as [A] and [B]).

(D) Fluorescence levels shown by the diatom cells alone and those treated with TRITC-labeled *Mp*PBD and various concentrations of AGYTD and AGAGD, respectively. The white bar represents the control experiment in which diatoms were treated with TRITC-labeled *Mp*PBD in the absence of peptide. Black bars represent experiments in which diatoms were treated with TRITC-labeled *Mp*PBD and indicated concentrations of AGYTD. Gray bars with an asterisk underneath represent experiments in which diatoms were incubated with TRITC-labeled *Mp*PBD before AGYTD was added. Blue bars represent experiments in which diatoms were treated with TRITC-labeled *Mp*PBD and indicated concentrations of AGAGD.

(E–I) Representative images showing (E) an untreated diatom; (F) a diatom treated with GFP-RIII1–5; (G) a diatom treated with GFP-RIII1–5 in the presence of 5 mM L-fucose; (H) a diatom treated with GFP-RIII1–5 in the presence of 37.5 μM AGYTD; and (I) a diatom treated with GFP-RIII1–5 in the presence of both 37.5 μM AGYTD and 5 mM L-fucose.

(J) Fluorescence levels shown by untreated diatoms and those treated with GFP-RIII1–5 and the addition of L-fucose, AGYTD, and both L-fucose and AGYTD together.

(K–M) Representative images showing (K) a diatom treated with FITC-labeled *Av*PBD; (L) a diatom treated with FITC-labeled *Av*PBD in the presence of 37.5 μM AGAGD; and (M) a diatom treated with FITC-labeled *Av*PBD in the presence of 37.5 μM AGYTD. (K)–(M) are at the same magnification, with the black scale bar in (M) indicating 10 μm.

(N) Fluorescence levels shown by those diatoms treated with FITC-labeled *Av*PBD and those with the additional 37.5 μM of AGAGD and AGYTD, respectively. Additional representative images are shown in Figure S4. Each bar represents the quantification of average fluorescence from 30 individual diatoms. The error bars represent SEMs.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.110002.

ACKNOWLEDGMENTS

We thank Dr. John Allingham for the use of the home source X-ray diffractometer at Queen's University, as well as staff members at the Canadian Light Source in Saskatoon, Canada, and the PETRA III facility at DESY in Hamburg, Germany, for access to data collection at these synchrotrons. We are grateful to Dr. EonSeon Jin, Hanyang University, Seoul, for the gift of the diatom, *C. neogracile*, and to Dr. Saeed Rismani Yazdi for assistance with the diatom cultures. We are indebted to Mr. Kim Munro at the Protein Function Discovery Unit at Queen's University for assistance with acquiring and interpreting ITC data, and to Mr. Rob Eves for completing the binding studies with *Av*PBD. We thank Ms. Irene van Oekel for preliminary tests on the binding of peptidyl ligands to *Mp*PBD and Mr. Joost van Dongen for analytical support. This project was funded by a Natural Science and Engineering Research Council (NSERC) Discovery Grant (RGPIN-2016-04810, to P.L.D., who holds the Canadian Research Chair in Protein Engineering). I.K.V. acknowledges financial support from the European Union (ERC-2014-StG contract no. 635928) and the Dutch Science Foundation (NWO ECHO Grant No. 712.016.002).

AUTHOR CONTRIBUTIONS

S.G. and P.L.D. conceived the study, designed the experiments, and wrote the manuscript. S.G. performed the co-crystallization, data collection, and structure determination of the X-ray crystal structures. S.G. performed the FP binding experiments and analyzed the data with the assistance of D.C.S. L.-G.M. and D.C.S. performed the peptide synthesis and purification. H.Z. and C.S. performed the diatom binding experiments and analyzed the data. C.O., L.B., and I.K.V. provided critical feedback to S.G. throughout the project and contributed to the critical editing of drafts of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 13, 2021
Revised: August 26, 2021
Accepted: October 25, 2021
Published: November 16, 2021

REFERENCES

Alonso-Caballero, A., Echelman, D.J., Tapia-Rojo, R., Haldar, S., Eckels, E.C., and Fernandez, J.M. (2021). Protein folding modulates the chemical reactivity of a Gram-positive adhesin. Nat. Chem. *13*, 172–181.

Bar Dolev, M., Bernheim, R., Guo, S., Davies, P.L., and Braslavsky, I. (2016). Putting life on ice: bacteria that bind to frozen water. J. R. Soc. Interface *13*. https://doi.org/10.1098/rsif.2016.0210.

Beilsten-Edmands, J., Winter, G., Gildea, R., Parkhurst, J., Waterman, D., and Evans, G. (2020). Scaling diffraction data in the DIALS software package: algorithms and new approaches for multi-crystal scaling. Acta Crystallogr. D Struct. Biol. *76*, 385–399.

Berne, C., Ducret, A., Hardy, G.G., and Brun, Y.V. (2015). Adhesins Involved in Attachment to Abiotic Surfaces by Gram-Negative Bacteria. Microbiol. Spectr. *3*. https://doi.org/10.1128/microbiolspec.MB-0018-2015.

Burgess, A., Vigneron, S., Brioudes, E., Labbé, J.C., Lorca, T., and Castro, A. (2010). Loss of human Greatwall results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B-Cdc2/PP2A balance. Proc. Natl. Acad. Sci. USA *107*, 12564–12569.

Chang, C., Wu, C., Osipiuk, J., Siegel, S.D., Zhu, S., Liu, X., Joachimiak, A., Clubb, R.T., Das, A., and Ton-That, H. (2019). Cell-to-cell interaction requires optimal positioning of a pilus tip adhesin modulated by gram-positive transpeptidase enzymes. Proc. Natl. Acad. Sci. USA *116*, 18041–18049.

Collins, A.J., Smith, T.J., Sondermann, H., and O'Toole, G.A. (2020). From Input to Output: The Lap/c-di-GMP Biofilm Regulatory Circuit. Annu. Rev. Microbiol. *74*, 607–631.

Cusumano, C.K., Pinkner, J.S., Han, Z., Greene, S.E., Ford, B.A., Crowley, J.R., Henderson, J.P., Janetka, J.W., and Hultgren, S.J. (2011). Treatment and prevention of urinary tract infection with orally active FimH inhibitors. Sci. Transl. Med. *3*, 109ra115.

Davey, M.E., and O'toole, G.A. (2000). Microbial biofilms: from ecology to molecular genetics. Microbiol. Mol. Biol. Rev. *64*, 847–867.

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. *62*, 72–82.

Garnham, C.P., Gilbert, J.A., Hartman, C.P., Campbell, R.L., Laybourn-Parry, J., and Davies, P.L. (2008). A Ca2+-dependent bacterial antifreeze protein domain has a novel beta-helical ice-binding fold. Biochem. J. *411*, 171–180.

Garnham, C.P., Campbell, R.L., and Davies, P.L. (2011). Anchored clathrate waters bind antifreeze proteins to ice. Proc. Natl. Acad. Sci. USA *108*, 7363–7367.

Guo, S., Garnham, C.P., Whitney, J.C., Graham, L.A., and Davies, P.L. (2012). Re-evaluation of a bacterial antifreeze protein as an adhesin with ice-binding activity. PLoS ONE *7*, e48805.

Guo, S., Stevens, C.A., Vance, T.D.R., Olijve, L.L.C., Graham, L.A., Campbell, R.L., Yazdi, S.R., Escobedo, C., Bar-Dolev, M., Yashunsky, V., et al. (2017). Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. Sci. Adv. *3*, e1701440.

Guo, S., Langelaan, D.N., Phippen, S.W., Smith, S.P., Voets, I.K., and Davies, P.L. (2018). Conserved structural features anchor biofilm-associated RTX-adhesins to the outer membrane of bacteria. FEBS J. *285*, 1812–1826.

Guo, S., Campbell, R., Davies, P.L., and Allingham, J.S. (2019a). Phasing with calcium at home. Acta Crystallogr. F Struct. Biol. Commun. *75*, 377–384.

Guo, S., Vance, T.D.R., Stevens, C.A., Voets, I.K., and Davies, P.L. (2019b). RTX Adhesins are Key Bacterial Surface Megaproteins in the Formation of Biofilms. Trends Microbiol. *27*, 470.

Guo, S., Vance, T.D.R., Zahiri, H., Eves, R., Stevens, C., Hehemann, J.H., Vidal-Melgosa, S., and Davies, P.L. (2021). Structural Basis of Ligand Selectivity by a Bacterial Adhesin Lectin Involved in Multispecies Biofilm Formation. MBio *12*, e00130-21.

Gwak, I.G., Jung, W.S., Kim, H.J., Kang, S.H., and Jin, E. (2010). Antifreeze protein in Antarctic marine diatom, Chaetoceros neogracile. Mar. Biotechnol. (NY) *12*, 630–639.

Han, Z., Pinkner, J.S., Ford, B., Obermann, R., Nolan, W., Wildman, S.A., Hobbs, D., Ellenberger, T., Cusumano, C.K., Hultgren, S.J., and Janetka, J.W. (2010). Structure-based drug design and optimization of mannoside bacterial FimH antagonists. J. Med. Chem. *53*, 4779–4792.

Hashem, A., Tabassum, B., and Fathi Abd Allah, E. (2019). *Bacillus subtilis*: a plant-growth promoting rhizobacterium that also impacts biotic stress. Saudi J. Biol. Sci. *26*, 1291–1297.

Hickman-Brenner, F.W., MacDonald, K.L., Steigerwalt, A.G., Fanning, G.R., Brenner, D.J., and Farmer, J.J., 3rd. (1987). Aeromonas veronii, a new ornithine decarboxylase-positive species that may cause diarrhea. J. Clin. Microbiol. *25*, 900–906.

Hollis, D.G., Weaver, R.E., Baker, C.N., and Thornsberry, C. (1976). Halophilic Vibrio species isolated from blood cultures. J. Clin. Microbiol. *3*, 425–431.

Huebinger, R.M., Do, D.H., Carlson, D.L., Yao, X., Stones, D.H., De Souza Santos, M., Vaz, D.P., Keen, E., Wolf, S.E., Minei, J.P., et al. (2020). Bacterial adhesion inhibitor prevents infection in a rodent surgical incision model. Virulence *11*, 695–706.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. *66*, 133–144.

Kitts, G., Giglio, K.M., Zamorano-Sánchez, D., Park, J.H., Townsley, L., Cooley, R.B., Wucher, B.R., Klose, K.E., Nadell, C.D., Yildiz, F.H., and Sondermann, H. (2019). A Conserved Regulatory Circuit Controls Large Adhesins in Vibrio cholerae. MBio *10*, e02822, 19.

Klemm, P., and Schembri, M.A. (2000). Bacterial adhesins: function and structure. Int. J. Med. Microbiol. *290*, 27–35.

Krachler, A.M., and Orth, K. (2013). Targeting the bacteria-host interface: strategies in anti-adhesion therapy. Virulence *4*, 284–294.

Lebeaux, D., Ghigo, J.M., and Beloin, C. (2014). Biofilm-related infections: bridging the gap between clinical management and fundamental aspects of recalcitrance toward antibiotics. Microbiol. Mol. Biol. Rev. *78*, 510–543.

Melia, C.E., Bolla, J.R., Katharios-Lanwermeyer, S., Mihaylov, D.B., Hoffmann, P.C., Huo, J., Wozny, M.R., Elfari, L.M., Böhning, J., Morgan, A.N., et al. (2021). Architecture of cell-cell junctions in situ reveals a mechanism for bacterial biofilm inhibition. Proc. Natl. Acad. Sci. USA *118*, e2109940118.

Moonens, K., Gideonsson, P., Subedi, S., Bugaytsova, J., Romaõ, E., Mendez, M., Nordén, J., Fallah, M., Rakhimova, L., Shevtsova, A., et al. (2016). Structural Insights into Polymorphic ABO Glycan Binding by Helicobacter pylori. Cell Host Microbe *19*, 55–66.

Moonens, K., Hamway, Y., Neddermann, M., Reschke, M., Tegtmeyer, N., Kruse, T., Kammerer, R., Mejías-Luque, R., Singer, B.B., Backert, S., et al. (2018). *Helicobacter pylori* adhesin HopQ disrupts *trans* dimerization in human CEACAMs. EMBO J. *37*, e98665.

Mydock-McGrane, L.K., Hannan, T.J., and Janetka, J.W. (2017). Rational design strategies for FimH antagonists: new drugs on the horizon for urinary tract infection and Crohn's disease. Expert Opin. Drug Discov. *12*, 711–731.

Newell, P.D., Boyd, C.D., Sondermann, H., and O'Toole, G.A. (2011). A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. PLoS Biol. *9*, e1000587.

Ofek, I., Hasty, D.L., and Sharon, N. (2003). Anti-adhesion therapy of bacterial diseases: prospects and problems. FEMS Immunol. Med. Microbiol. *38*, 181–191.

Sanchez, C.J., Shivshankar, P., Stol, K., Trakhtenbroit, S., Sullam, P.M., Sauer, K., Hermans, P.W., and Orihuela, C.J. (2010). The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation in vivo and in biofilms. PLoS Pathog. *6*, e1001044.

Sauer, M.M., Jakob, R.P., Luber, T., Canonica, F., Navarra, G., Ernst, B., Unverzagt, C., Maier, T., and Glockshuber, R. (2019). Binding of the Bacterial Adhesin FimH to Its Natural, Multivalent High-Mannose Type Glycan Targets. J. Am. Chem. Soc. *141*, 936–944.

Smith, T.J., Sondermann, H., and O'Toole, G.A. (2018). Type 1 Does the Two-Step: Type 1 Secretion Substrates with a Functional Periplasmic Intermediate. J. Bacteriol. *200*, e00168-18.

Solanki, V., Tiwari, M., and Tiwari, V. (2018). Host-bacteria interaction and adhesin study for development of therapeutics. Int. J. Biol. Macromol. *112*, 54–64.

Spaulding, C.N., Klein, R.D., Ruer, S., Kau, A.L., Schreiber, H.L., Cusumano, Z.T., Dodson, K.W., Pinkner, J.S., Fremont, D.H., Janetka, J.W., et al. (2017). Selective depletion of uropathogenic E. coli from the gut by a FimH antagonist. Nature *546*, 528–532.

Syed, K.A., Beyhan, S., Correa, N., Queen, J., Liu, J., Peng, F., Satchell, K.J., Yildiz, F., and Klose, K.E. (2009). The Vibrio cholerae flagellar regulatory hierarchy controls expression of virulence factors. J. Bacteriol. *191*, 6555–6570.

Teschler, J.K., Zamorano-Sánchez, D., Utada, A.S., Warner, C.J., Wong, G.C., Linington, R.G., and Yildiz, F.H. (2015). Living in the matrix: assembly and control of Vibrio cholerae biofilms. Nat. Rev. Microbiol. *13*, 255–268.

Totsika, M., Kostakioti, M., Hannan, T.J., Upton, M., Beatson, S.A., Janetka, J.W., Hultgren, S.J., and Schembri, M.A. (2013). A FimH inhibitor prevents acute bladder infection and treats chronic cystitis caused by multidrug-resistant uropathogenic Escherichia coli ST131. J. Infect. Dis. *208*, 921–928.

Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G.N. (2004). REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr. D Biol. Crystallogr. *60*, 2184–2195.

Vance, T.D.R., Guo, S.Q., Assaie-Ardakany, S., Conroy, B., and Davies, P.L. (2019). Structure and functional analysis of a bacterial adhesin sugar-binding domain. PLoS ONE *14*, e0220045.

Wellens, A., Garofalo, C., Nguyen, H., Van Gerven, N., Slättegård, R., Hernalsteens, J.P., Wyns, L., Oscarson, S., De Greve, H., Hultgren, S., and Bouckaert, J. (2008). Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the FimH-oligomannose-3 complex. PLoS ONE *3*, e2040.

Whitaker, I.S., Kamya, C., Azzopardi, E.A., Graf, J., Kon, M., and Lineaweaver, W.C. (2009). Preventing infective complications following leech therapy: is practice keeping pace with current research? Microsurgery *29*, 619–625.

Winter, G., Lobley, C.M., and Prince, S.M. (2013). Decision making in xia2. Acta Crystallogr. D Biol. Crystallogr. *69*, 1260–1273.

Wong, G.C.L., Antani, J.D., Lele, P.P., Chen, J., Nan, B., Kühn, M.J., Persat, A., Bru, J.L., Høyland-Kroghsbo, N.M., Siryaporn, A., et al. (2021). Roadmap on emerging concepts in the physical biology of bacterial biofilms: from surface sensing to community formation. Phys. Biol. *18*. https://doi.org/10.1088/1478-3975/abdc0e.

Wozniak, D.J., and Parsek, M.R. (2014). Surface-associated microbes continue to surprise us in their sophisticated strategies for assembling biofilm communities. F1000Prime Rep. *6*, 26.

Zheng, S., Bawazir, M., Dhall, A., Kim, H.E., He, L., Heo, J., and Hwang, G. (2021). Implication of Surface Properties, Bacterial Motility, and Hydrodynamic Conditions on Bacterial Surface Sensing and Their Initial Adhesion. Front. Bioeng. Biotechnol. *9*, 643722.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli BL21 (DE3)	ThermoFisher	Catalog# EC0114
Chemicals, peptides, and recombinant proteins		
FITC	Sigma-Aldrich	CAS# 3326-32-7
Unlabelled peptides	GenicBio	http://www.genicbio.com/
Fmoc amino acids	Novabiochem	https://p3bio.com/amino-acids/fmoc-amino-acids/?gclid=Cj0KCQjw5JSLBhCxARIsAHgO2SdK-skpdiZsXIdjMczeURwv-kVIfRVpBYsFEhmwjwRW8BFbqB7p7T0aAuPFEALw_wcB
Deposited data		
MpIBP_RIII1-4 in complex with FITC-A_{β}GPDSD	Protein Data Bank	6X6Q
MpIBP_RIII1-4 in complex with FITC-A_{β}DSTD	Protein Data Bank	6X6M
MpIBP_RIII1-4 in complex with AGYTD	Protein Data Bank	6X5W
MpIBP_RIII1-4 in complex with AGYTS	Protein Data Bank	6X5V
Experimental models: organisms/strains		
Chaetoceros neogracile	(Guo et al., 2017)	NCBI GenBank: EL621146.1
Escherichia coli BL21 (DE3)	ThermoFisher	Catalog# EC0114
Recombinant DNA		
*Mp*IBP_RIII1-4 construct	(Guo et al., 2017)	GenBank: ABL74378.1
Software and algorithms		
GraphPad Prism 8	GraphPad	https://www.graphpad.com/
CCP4i2 Software suite	CCP4	http://legacy.ccp4.ac.uk/ccp4i_main.php
ImageJ	ImageJ	https://imagej.nih.gov/ij/
Other		
Intavis MultiPep RSi peptide synthesizer	Intavis Inc.	https://www.selectscience.net/products/multipep-rsi/?prodID=171693
Tecan Infinite F500 plate reader	Eindhoven University of Technology	https://lifesciences.tecan.com/?utm_term=tecan&utm_campaign=SO-Brand&utm_source=adwords&utm_medium=ppc&hsa_net=adwords&hsa_tgt=kwd-386882420855&hsa_ad=514074167501&hsa_acc=9279258943&hsa_grp=121576178140&hsa_mt=p&hsa_cam=12736641030&hsa_kw=tecan&hsa_ver=3&hsa_src=g&gclid=Cj0KCQjw5JSLBhCxARIsAHgO2SfpkZx47QU-sLWEUTdhcgo1jeMZHqrv6fy9yyU6qYSUnaprCeqPAz0aAnZTEALw_wcB
P11 beamline	PETRA III facility in Hamburg, Germany	https://photon-science.desy.de/facilities/petra_iii/beamlines/p11_bio_imaging_and_diffraction/index_eng.html
08ID-1 beamline	Canadian Light Source synchrotron	https://cmcf.lightsource.ca/news/beamline-08id1-upgraded/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MicroCal VP-ITC	Malvern	https://www.malvernpanalytical.com/en/products/product-range/microcal-range?campaignid=1071234466&adgroupid=54122082362&creative=349133914257&keyword=malvern%20microcal&matchtype=e&network=g&device=c&gclid=Cj0KCQjw5JSLBhCxARIsAHgO2Scv-ec0CcuTgxIM8ZClfHFZ2kIwGRRIVTkok0s9BPuJlDFReaaUl28aAmD-EALw_wcB
Olympus IX83 inverted fluorescence microscope w/ Andor Zyla 4.2+ Camera	Olympus	https://www.olympus-lifescience.com/en/microscopes/inverted/ixplore-pro/?gclid=Cj0KCQjw5JSLBhCxARIsAHgO2SfWr_2y46Q7DOdGQC8CeVJkSAmxgnGurZfkxm43tEtwgSpiF69VSeoaAmelEALw_wcB

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Peter Davies (peter.davies@queensu.ca).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- X-ray crystal structure coordinates solved in this study have been deposited in the Protein Data Bank with accession codes of PDB: 6X6Q (*Mp*PBD - FITC-A_{β}GPDSD), PDB: 6X6M (*Mp*PBD - FITC-A_{β}DSTD), PDB: 6X5W (*Mp*PBD - AGYTD), and PDB: 6X5V (*Mp*PBD - AGYTS). Data S1 that includes Figures S2E–S2P and S6B–S6M has been deposited in Mendeley (https://data.mendeley.com/datasets/y6yg4cmp8s/1)
- This paper does not report original code.
- The data that support the findings of this study are available from the corresponding author P.L.D upon reasonable request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Diatom cell culture

Chaetoceros neogracile is a psychrophilic marine diatom isolated from Antarctica (GenBank Accession: EL621146.1). *C. neogracile* used in this study was a gift by Dr. E. Jin from the Research Institute for Natural Sciences at Hanyang University(Guo et al., 2017). *C. neogracile* was grown in F/2 medium at 4°C with light and shaking.

Protein overexpression by *E. coli*

E. coli BL21 (DE3) (ThermoFisher) cells were grown in LB medium at 37°C or 25°C as previously described (Guo et al., 2017, 2019a).

METHOD DETAILS

Peptide synthesis

FITC-labeled peptides FITC-A_{β}GPDSD, FITC-A_{β}DSTPD, FITC-A_{β}DSTD and FITC-A_{β}GPDD were synthesized by Fmoc solid-phase peptide synthesis, either manually or using an automated Intavis MultiPep RSi peptide synthesizer. The protected amino acids (linked to Wang resins) and FITC were purchased from Novabiochem and Sigma-Aldrich. Crude peptides were then analyzed and purified by high-pressure liquid chromatography (HPLC) using a preparative reverse-phase column with MS detection. The peptides were freeze-dried and stored at – 30°C. All other FITC-labeled peptides used in the structure-guided ligand optimization procedures by fluorescence polarization (FP) and the four unlabeled peptides (AGAGD, AGYTD, ATYTS and AGATD) used in the ITC measurements and X-ray crystallography were purchased from GenicBio (Shanghai, China). HPLC and MS spectra for representative peptides are shown in Figure S6.

Fluorescence polarization

To ensure the solubility of FITC-labeled short peptides used in FP assays, they were first dissolved in a small amount of DMSO before being diluted by the FP buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 0.1% Tween 20, 1 mg/mL BSA) to a final concentration of 10 nM. Dilution series of *Mp*PBD were made on round-bottom 384-well plates (Corning, Black). The protein-peptide mixture in the plate was incubated at room temperature for at least 30 min before the fluorescence polarization was measured using a Tecan Infinite F500 plate reader (excitation 485 nm, emission 535nm).

Co-crystallization, X-ray diffraction, and structure solutions of MpPA14 with peptide ligands

The original *Mp*IBP_RIII1-4 construct was overexpressed in *E.coli* BL21(DE3) and purified as previously described (Guo et al., 2017, 2019a). The protein self-associated in the crystal by inserting its C-terminal "TPD" sequence into the ligand-binding site of a symmetry-related molecule. This crystal contact competed with free peptides for binding *Mp*PBD and interfered with the crystallization of peptide-protein complexes. Thus, *Mp*IBP_RIII1-4 protein (referred to as *Mp*PBD throughout the manuscript) used for co-crystallography was truncated by 12 amino acids from the original construct, which ended at the residue N507 instead of D519. In all other respects, *Mp*PBD was produced, purified, and crystallized as previously described (Guo et al., 2017, 2019a). Co-crystallization of *Mp*PBD with various peptides was performed using the "microbatch-under-oil" method by mixing equal volumes of ~5 mg/mL protein with a precipitant solution composed of approximately 0.1 M calcium chloride, 0.1 M sodium acetate (pH 4.6), 30% (w/v) PEG400 and 1-2 mM of different peptides, including FITC-A_{β}GPDSD, FITC-A_{β}DSTD, and the unlabeled peptides AGYTD and AGYTS.

X-ray crystallographic data were collected either at the P11 beamline of the PETRA III facility at DESY (Hamburg, Germany) or at the 08ID-1 beamline of the Canadian Light Source synchrotron facility via remote access. Data were indexed and integrated with X-ray Detector Software (XDS) (Kabsch, 2010) and CCP4-Aimless (Evans, 2006) or the DIALS/xia2 (Beilsten-Edmands et al., 2020; Winter et al., 2013) in the CCP4i2 software suite. The structure solutions for all complexes were obtained by using molecular replacement using the *Mp*IBP_RIII1-4 structure (PBD: 5K8G) as the search model (Guo et al., 2017). The structures were refined using CCP4-Refmac5 (Vagin et al., 2004).

Isothermal titration calorimetry

Isothermal calorimetric titration (ITC) measurements were performed at 30°C using a MicroCal VP-ITC instrument (Malvern). *Mp*PBD was dialyzed overnight in a buffer of 50 mM Tris-HCl, pH 9, 150 mM NaCl, 5 mM $CaCl_2$. Next, the protein was diluted to approximately 20 μM and was mixed with serial 5 μL aliquots of 200 μM of each of the four peptide solutions (AGAGD, AGATD, AGYTD and AGYTS; all readily soluble in the buffer). Peptide solutions were automatically added by a rotating syringe (400 RPM) at 5-min intervals into the *Mp*PBD solution for a total of 50 injections.

Diatom binding experiments

TRITC-labeled *Mp*PBD (TRITC-*Mp*PBD, 0.2 mg/mL) in the presence or absence of peptides was incubated with diatoms in buffer (50 mM Tris-HCl pH 9, 300 mM NaCl, 5 mM $CaCl_2$) with gentle mixing for 2 h. Next, diatoms were pelleted by centrifugation for 3 min at 4,500 x g, and the resulting supernatant was discarded. This procedure was repeated three times to wash out unbound TRITC-*Mp*PBD before the diatom pellet was finally resuspended in 20 μL of buffer, which was then examined on slides by fluorescence microscopy. In parallel experiments to test if the strong ligand AGYTD could compete off the TRITC-*Mp*PBD that was already bound to diatoms, TRITC-*Mp*PBD was incubated with diatom for 1.5 h before the peptide AGYTD was added. The remainder of the experiment followed the same procedure as described above.Images were obtained using an Olympus IX83 inverted fluorescence microscope equipped with an Andor Zyla 4.2 Plus camera. For experiments shown in Figures 7E–7J, a GFP-RIII (3.5 μM) construct was used as opposed to TRITC-*Mp*PBD. All other experimental procedures were the same as described above.

QUANTIFICATION AND STATISTICAL ANALYSIS

Fluorescence polarization

All measurements were performed in triplicate, and the data were plotted with the GraphPad Prism 8 software using a non-linear regression analysis method (single-site binding model).

Isothermal titration calorimetry

The data were analyzed by Origin software Version 5.0 (MicroCal).

Diatom binding experiments

Quantification of the fluorescence intensity was done using Fiji ImageJ. The corrected total cell fluorescence (CTCF) was calculated using the formula: CTCF = Integrated Density – (Area of selected cell x Mean Fluorescence of the background) (Burgess et al., 2010). Quantification of 30 individual diatom cells was done for each treatment. As diatom cell aggregates produce overexposed fluorescence while those cells lacking a silica frustule have damaged plasma membranes necessary for protein binding, they were excluded from the measurements. Graphs were made using GraphPad Prism 8.