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

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

[*]Corresponding author: Peter L. Davies, peter.davies@queensu.ca

[1]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON Canada K7L 3N6

[2]Laboratory of Self-Organizing Soft Matter,

[3]Laboratory of Chemical Biology,

[4]Laboratory of Macromolecular and Organic Chemistry,

[5]Institute for Complex Molecular Systems,

[2-5]Address: Eindhoven University of Technology P.O. Box 513, 5600 MB Eindhoven, Netherlands Tel: +31 (0)40 247 5303

[#]Author 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

Abstract:

Modulation of protein-protein interactions (PPIs) with small-molecules is a promising conceptual approach in drug discovery. In the area of bacterial colonization, PPIs contribute to adhesin-mediated biofilm formation that cause most infections. However, the molecular basis underlying these adhesin-ligand interactions is largely unknown. The 1.5-MDa adhesion protein, *Mp*IBP, uses a peptide-binding domain (*Mp*PBD) to help its Antarctic bacterium form symbiotic biofilms on sea ice with microalgae such as diatoms. X-ray crystallography revealed *Mp*PBD uses Ca^{2+}-dependent interactions to self-associate with a crystal symmetry mate via the C-terminal threonine-proline-aspartate sequence. Structure-guided optimization derived penta-peptide ligands that bound *Mp*PBD 1,000-fold more tightly, with affinities in the nano-molar range. These ligands act as potent antagonists to block *Mp*PBD from binding to the diatom cells. Since adhesins of some human pathogens contain peptide-binding module homologs of *Mp*PBD, this same conceptual approach could help develop ligand-based PPI modulators to disrupt harmful bacteria-host interactions.

Introduction:

Protein-protein interactions (PPIs) are central to most biochemical processes such as actin and tubulin polymerization, ATP production through the electron transport chain, and signal transduction via G protein-coupled receptors(1-3). Given PPIs' vital role in the well-being of cells, their modulation by drug-like molecules is an attractive approach that holds great promise for the development of new strategies to treat various diseases. Over the past two decades, significant advances in structural biology and medicinal chemistry have enabled some fruitful developments in drug discovery to target PPIs that govern mammalian cellular processes(4-6). For example, the structure of B-cell lymphoma extra-large (Bcl-xL) protein in complex with a peptide derived from the Bcl-2 Antagonist/Killer (Bak) protein inspired the fragment-based identification of the first Bcl-2 family inhibitors(7). This ultimately led to the FDA approval of Venetoclax in 2016 to treat lymphocytic leukemia and lymphoma(8). In addition to such small-molecule inhibitors, peptide antagonists have been developed to target their partner proteins' ligand-binding sites to disrupt PPIs of Bcl-2 proteins(9) and those involved in other diseases(10-13). Although the use of small-molecule and peptide therapeutics to target PPIs is an active and highly promising field(12-14), most of the studies published to date have been focused on cancer-related, human PPIs. Investigations of PPIs that mediate bacterial infections have taken second place to antibiotic development. It is only now when antibiotics are losing some of their potency that other strategies for inhibiting bacteria are being intensively studied(15-18).

Bacterial adhesins are a key class of virulence factors that bind bacteria to host cells(19-22), and subsequently help develop multi-cellular communities called

biofilms responsible for over 80% of chronic infections in humans(23). There is currently a shortage of effective treatments against biofilm-related infections(23-25). With the increasing spread of antibiotic-resistant pathogenic bacteria, there is an urgent need for the development of PPI disruptors to block the adhesins of bacteria, which might prevent the colonization and persistent infections caused by biofilms. 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 like the carbohydrates and proteins on host cell surfaces(19, 26-31). With structures of adhesin-ligand complexes starting to emerge(32-34), researchers are beginning to elucidate the molecular basis of the interactions required to build a biofilm. To date, most of these studies have focused on the characterization of lectin-glycan interactions(32, 33, 35), and have resulted in some effective treatments against bacterial infections(17, 36). For example, the FimH adhesin of uropathogenic *Escherichia coli* has been successfully targeted by mannose analogs to help treat urinary tract infections(35, 37-41). 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. While PPIs are heavily involved in bacteria-host interactions, examples that detail the molecular basis of these interactions are scarce. Given that PPIs are typically many fold stronger than those of lectin-glycan interactions(42), the development of PPI inhibitors to disrupt bacteria-host interactions will be of great interest and potential utility.

Marinomonas primoryensis ice-binding protein (*Mp*IBP) is an exceptionally large (~ 1.5 MDa) Repeats-In-Toxin (RTX) adhesin found on the surface of its Antarctic bacterium(27, 43). While its N-terminal Region I (RI) is responsible for anchoring the adhesin to the outer-membrane, its C-terminal ligand-binding Regions III and IV (RIII

and RIV) help the bacterium form symbiotic biofilms with diatoms on the underside of lake ice(27, 28). *Mp*IBP_RIII contains five β-sheet-rich domains, including a carbohydrate-binding lectin module (RIII_5, also referred as *Mp*PA14) and a Peptide-Binding Domain (RIII_3, *Mp*PBD), both responsible for binding diatoms by interacting with the cell-surface glycans and proteins, respectively. We reported the X-ray crystal structure of *Mp*IBP_RIII1-4, which revealed that *Mp*PBD folds as an oblong β-sandwich with a shallow, solvent-exposed ligand-binding cavity on the periphery of the structure(27). Importantly, it was observed in the crystal unit cell that the three C-terminal amino-acid residues (threonine-proline-aspartate, TPD) of *Mp*IBP_RIII1-4 were stably anchored in the ligand-binding cavity of a symmetry-related molecule within the crystal unit cell. The observed crystal contacts led us to postulate that *Mp*PBD binds its bacterium to diatoms by interacting with surface proteins via short tripeptide sequences at the C-termini.

Here, we show that the TPD peptide sequence binds *Mp*PBD in solution with an affinity of 26 μM. Using a structure-guided approach, we optimized the peptidyl sequence by iteratively screening a library of pentapeptides and obtained ligands that bound *Mp*PBD ~ 1000-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 effective at inhibiting the binding of *Mp*PBD to diatoms, which gives insight into how microbial adhesion can be disrupted through ligand-based antagonists.

Results and discussion

Initial ligand identification

Close inspection of the *Mp*IBP_RIII1-4 structure (PDB code: 5K8G) revealed the detailed PPIs involved in its self-association within the crystal by binding the C-terminal Thr-Pro-Asp (TPD) of a symmetry-related molecule (Figure 1A). The C-terminal aspartate residue is buried inside the ligand-binding pocket of the other molecule, with its terminal α-carboxyl oxygens binding Ca1 and Ca2 via three ionic bonds (Figure 1B). The protein-protein interaction is further enhanced by hydrogen bonding between 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*PBD. The threonine residue at the pre-penultimate position (position 3) is 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 fortuitous crystallisation of the self-interacting *Mp*PBD_RIII1-4 indicated that peptidyl sequences are the likely ligands of *Mp*PBD. To prevent the self-association through the extended C-terminal "TPD" sequence, and facilitate the binding of free peptide to *Mp*PBD, we designed a new *Mp*IBP_RIII1-4 construct by truncating the carboxyl end of the original protein from D519 to N507 (Figure 1A, Supplementary file 1-Figure S1; green and blue arrows). This new *Mp*IBP_RIII1-4 construct was used for the subsequent binding and structural studies described below.

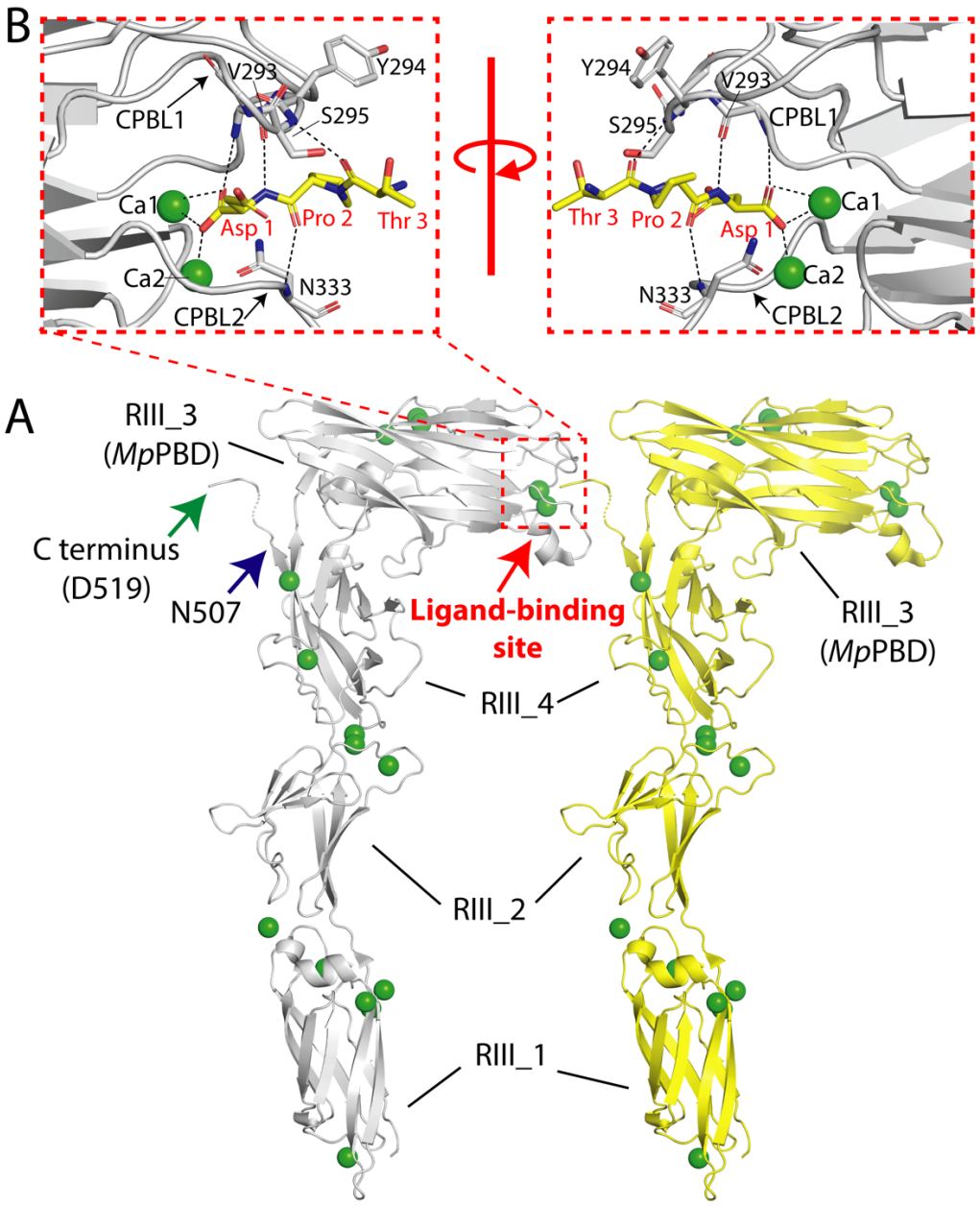


Figure 1: Self-association of *Mp*IBP_RIII1-4 within the crystal unit cell. A) *Mp*IBP_RIII1-4 molecule (grey) interacts with a neighboring symmetry mate (yellow) in the crystal via its C-terminal "TPD" sequence. The C-terminal D519 residue of the original *Mp*IBP_RIII1-4 construct is marked by a green 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) Zoomed-in view of the ligand-binding site of *Mp*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 approximately 180° around a vertical axis from the left. Polar interactions are indicated by black dashed lines. Carbon atoms of the bound "TPD" sequence are colored in yellow while those for its symmetry-related protein are colored in grey. 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 labelled and shown in stick representation.

Given that a C-terminal carboxylic acid group appears to be key for the protein-protein interaction, we used fluorescence polarisation (FP) to screen a small collection of 15 N-terminally FITC-labelled peptides that end in various C-terminal amino acid residues with a free α-carboxyl group (selected on the basis of availability within the laboratory). In the presence of ~ 80 μM of *Mp*IBP_RIII1-4, two FITC-labelled and phosphorylated peptides with sequences of IKARAS($_p$S)SPVILVGTHLD (pep 14) and RHKKLMFK($_p$T)EGPDSD (pep 15) produced significant delta fluorescence polarisation 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 polarisation values. FP with protein titrations further showed that in the presence of 2 mM Ca^{2+}, peptides 14 and 15 bound *Mp*PBD with EC_{50} values of 2.7 μM 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-peptide interaction as demonstrated by the structural data.

Interestingly, the three binding sequences “TPD”, pep 14 and pep 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 neighbouring 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 amino-acid sequences, peptides encompassing only the last five

196 amino acids of the original *Mp*IBP_RIII1-4 construct (FITC-A_βDSTPD) and pep 15
197 (FITC-A_βGPDSD) were synthesized and studied. The affinity of *Mp*PBD for these two
198 short peptide ligands, together with two analogs (FITC-A_βDSTD and FITC-A_βGPDD)
199 was measured by FP (Figure 2C).
200
201 Unexpectedly, FITC-A_βDSTPD, which contains the C-terminal "TPD" binding
202 sequence originally identified in the crystal structure of *Mp*IBP_RIII1-4, showed the
203 weakest interaction to *Mp*PBD out of this series, with an EC_{50} of ~ 27 μM. In
204 contrast, the peptide, FITC-A_βDSTD, bound at least 200-fold stronger with an EC_{50} of
205 0.11 μM (Figure 2C). Similarly, FITC-A_βGPDSD bound with an EC_{50} of 0.16 μM,
206 which is roughly 4-fold stronger than its longer, phosphorylated counterpart (FITC-
207 RHKKLMFK_pTEGPDSD, EC_{50} = 0.69 μM), validating the importance of the C-
208 terminal residues for binding *Mp*PBD. Furthermore, FITC-A_βGPDSD bound the
209 protein 20-fold stronger than its analog FITC-A_βGPDD.
210
211 To further ascertain the chemical constituents of the peptide C-terminus responsible
212 for binding *Mp*PBD, two additional analogs of the peptide FITC-A_βGPDSD were
213 synthesized. When the C-terminal aspartate was replaced by an asparagine, a
214 residue with similar length side chain but different chemistry, the affinity of the
215 peptide FITC-A_βGPDSN for *Mp*PBD fell by 40% to an EC_{50} of 0.27 μM (Figure 2C).
216 However, the substitution of the terminal α-carboxylic acid group of FITC-A_βGPDSD
217 by a C-terminal carboxamide group (FITC-A_βGPDSD-NH_2) abolished the protein-
218 peptide interaction all together (Figure 2C). These results demonstrated the crucial
219 importance of the ionic interaction between the terminal carboxylic acid group and
220 the two Ca^{2+} ions in the ligand-binding site of *Mp*PBD. The side-chain carboxylic acid

group of the peptide C-terminal aspartic acid residue gave more favourable binding to *Mp*PBD 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 amino acids at the penultimate position (position 2) have a significant impact on the peptide-protein interactions. The strong binders, FITC-A$_\beta$GPDSD and FITC-A$_\beta$DSTD, contain the small polar residues of either threonine or serine at this position. The intermediate binder FITC-A$_\beta$GPDSD has an aspartate while the weakest binder FITC-A$_\beta$DSTPD 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 five-membered imine ring is unable to interact with the protein (Figure 1B). As for FITC-A$_\beta$GPDD, 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$_\beta$GPDSD and FITC-A$_\beta$DSTD for *Mp*PBD.

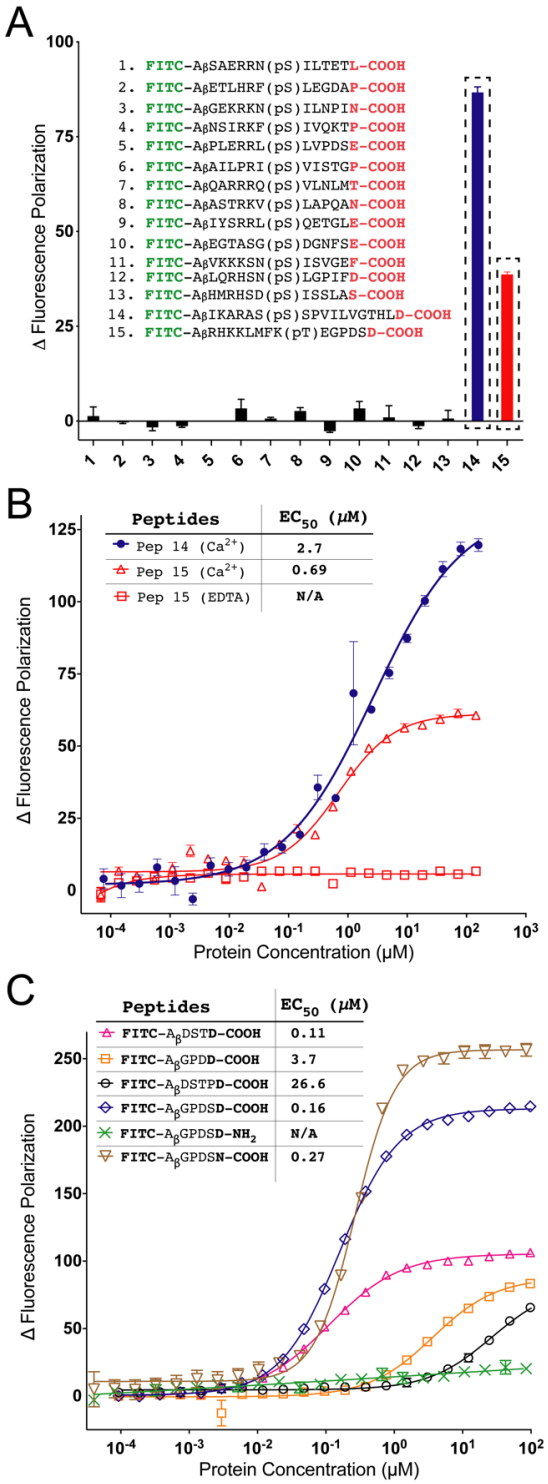


Figure 2: Identification of initial peptidyl ligands of *Mp*PBD. (A) Screening of a collection of 15 FITC-labelled peptides by fluorescence polarisation (FP) with *Mp*IBP_RIII1-4 at a concentration of ~ 80 μM. Background polarization was subtracted from all values to result in Δ polarisation values on 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*IBP_RIII1-4 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 six different FITC-labelled peptides to *Mp*PBD. The mean of three experiments were plotted. Some of the SD error bars are smaller than the data point symbols in B) and C). The calculated EC_{50} values for each experiment are shown in B) and C).

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*IBP_RIII1-4 in complex with the peptides FITC-A_βDSTD and FITC-A_βGPDSD to 2-Å resolution (Supplementary file 1-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 beforehand at the N-terminal end appeared to be ambiguous. No electron density could be observed beyond position 3 suggesting these residues are highly flexible because they are fully exposed to the solvent and are thus unlikely to interact with the protein.

At position 2 of FITC-A_βGPDSD, serine has its side-chain hydroxyl group hydrogen bonded with the side-chain amide of 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. Indeed, 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 favourable conformation for polar interactions. The observed additional

hydrophobic interactions might 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 (Figure 3A, B, left panels). Taken together with the gained interactions from the peptide position 2, the structural data gave the 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 FITCA_βGPDD, respectively.

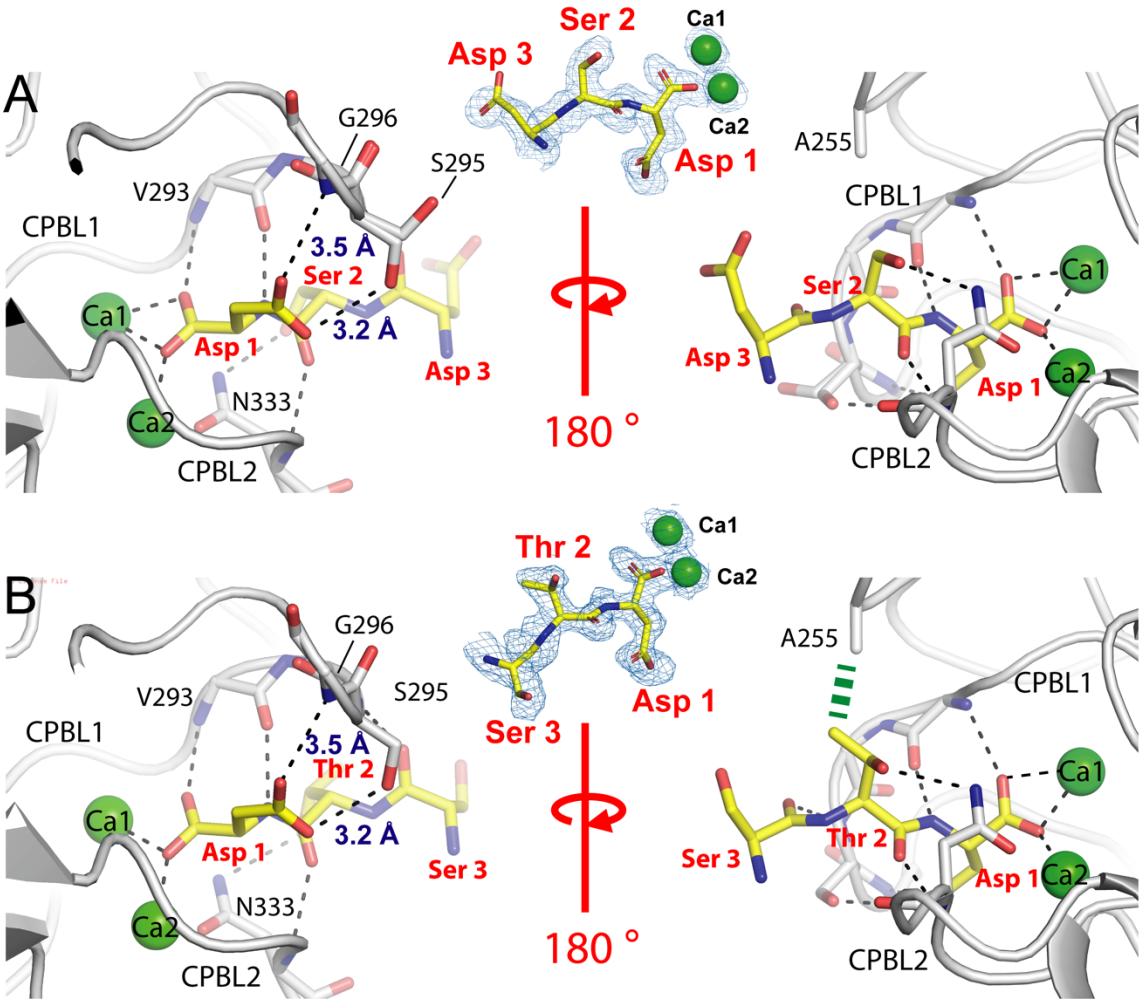


Figure 3: Protein-peptide binding interfaces revealed by X-ray crystallography. Zoomed-in views of the *Mp*PBD ligand-binding site interacting with peptides that end with sequences of DSD (A) and DSTD (B), respectively. The $2F_o - F_c$ electron density maps around the peptide residues and Ca1 and 2 are shown as blue mesh (contoured at 1 σ). The right panels in A) and B) are views that are rotated around approximately 180° from the left panels. Polar interactions are indicated by black dashed lines. Carbon atoms of the peptides are colored in yellow while those for the protein are colored in grey. Oxygen atoms are red, nitrogen atoms are blue, while Ca^{2+} ions are shown as green spheres. Amino acids involved in protein-peptide interactions are shown in stick representation. Peptide amino-acid residues are indicated in red 3-letter code, while those for the protein are labelled in black one-letter code. CIF files can be found in Figure 3 – source data 1 and has been deposited in the Protein Data Bank (PDB: "Code")

312
313 ***Structure-guided screening identified ligands with low-nano-molar affinity***

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

317 The first round of screening involved 20 pentapeptides with a consensus sequence
318 of FITC-AGAGX where the first four amino acids of the peptides were alternating
319 alanine and glycine residues. The ultimate (C-terminal) residue "X" represented one
320 of the 20 naturally occurring amino acids (Figure 4A). Consistent with the results
321 shown above, the peptide with an aspartate at the C-terminal position (FITC-
322 AGAGD) bound *Mp*PBD the tightest, with a moderate EC_{50} of 3.2 μM. This was
323 followed by those peptides with large hydrophobic side-chains at position 1, such as
324 FITC-AGAGI, FITC-AGAGY and FITC-AGAGF, which produced EC_{50} values of 4.4
325 μM, 4.6 μM, and 7.2 μM, respectively, when bound to *Mp*PBD (Supplementary file 1-
326 Table S2, Figure S2). Most of the other peptides bound *Mp*PBD with an affinity in the
327 10 μM range, while those with basic side chains (FITC-AGAGK and FITC-AGAGR)
328 had negligible interaction with *Mp*PBD.

329 Having identified FITC-AGAGD as the strongest ligand for *Mp*PBD in the first round
330 of screening, we proceeded with a second set of 20 pentapeptides that had a
331 consensus sequence of FITC-AGAXD. All peptides except for AGAPD (EC_{50} = 4.6
332 μM) bound more strongly than FITC-AGAGD (EC_{50} =1.5 μM). Consistent with the
333 structural studies reported earlier (Figure 1B), neither proline nor glycine at position 2
334 can have side-chain interactions with *Mp*PBD, explaining their lower affinity for the
335 protein. Furthermore, the peptide with threonine in the second position (FITC-
336 AGATD) has the highest affinity for *Mp*PBD with an EC_{50} of 89 nM (Figure 4A, green;

Supplementary file 1-Table S2; Figure S3), 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 nM and 170 nM, respectively. Consistent with the characterisation of FITC-A_{β}GPDSD and FITC-A_{β}DSTD, FITC-AGATD bound *Mp*PBD more than 3-fold stronger than did FITC-AGASD (EC_{50} = 310 nM), validating the importance of the threonine methyl group in restraining the hydroxyl group to a favourable conformation for interacting with N333.

With the last two C-terminal residues defined as Thr and Asp, we proceeded to screen for the optimal amino acid 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 nM, 23 nM, and 23 nM (Figure 4A, magenta; Supplementary file 1-Table S3; Figure S4), respectively. We reason that these amino acids 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 amino-acid 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: FITC-AGYTD is more soluble than FITC-AGWTD; FITC-AGYTD is not subject to dimerization the way FITC-

AGCTD can be linked by cysteine-dependent disulphide formation, which might 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 weaker than did FITC-AGYTD, with their EC_{50} values ranging from high nano- to micro-molar 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 (Figure 4A, red; Supplementary file 1-Table S3, Figure S5). 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 “TPD” binding sequence initially identified (Figure 4B).

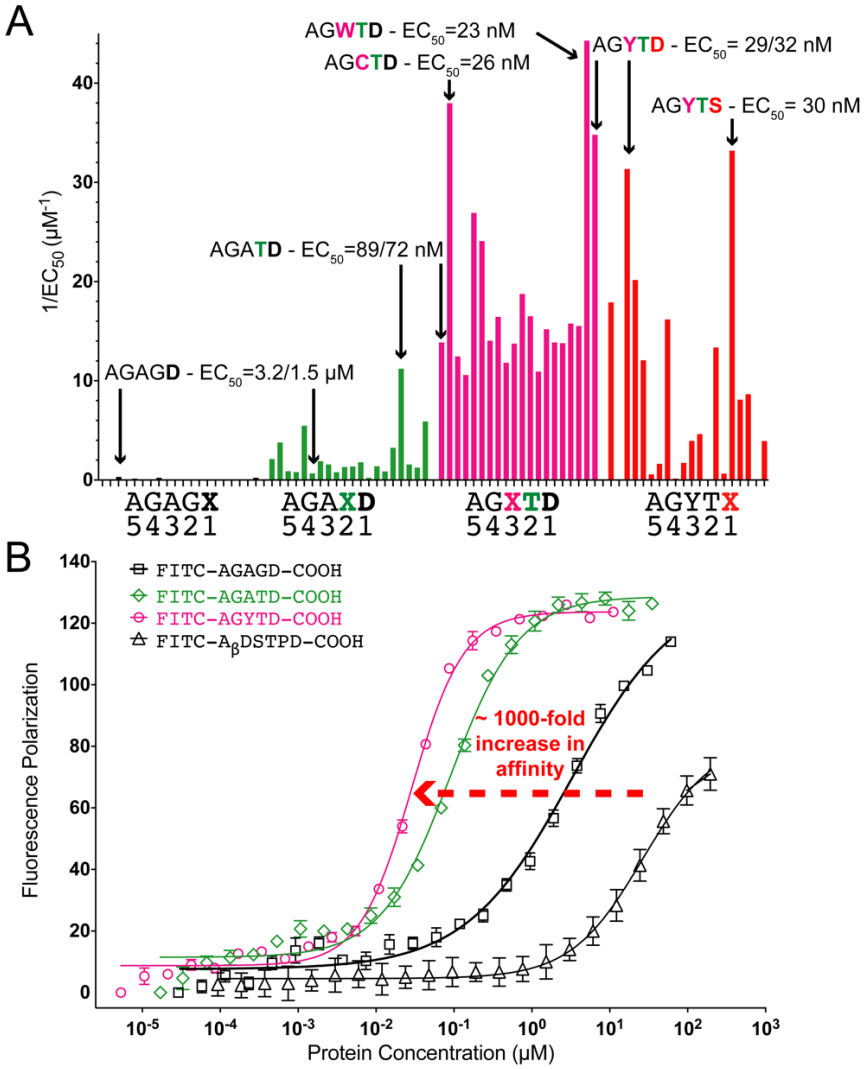


Figure 4: Structure-guided optimization of peptidyl ligands of *Mp*PBD. (A) Overview of the binding of the 80 FITC-labelled pentapeptides to *Mp*PBD in four rounds of screening. Data are plotted as average 1/EC_{50} values calculated from FP assays of each individual FITC-labelled peptide (in three replicates). All EC_{50} values are listed in Tables S2, S3, and the corresponding FP titration plots are shown in Figures S2-S5. Data bars for peptides in the first round of screening (FITC-AGAGX) are colored black, while those for second (FITC-AGAXD), third (FITC-AGXTD) and fourth (FITC-AGYTX) rounds are colored green, magenta and red, respectively (only some amino-acid sequences are labelled in the graph for short). The strongest *Mp*PBD 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 ligands' affinity for *Mp*PBD.

Isothermal titration calorimetry validated the nano-molar affinity of MpPBD ligands

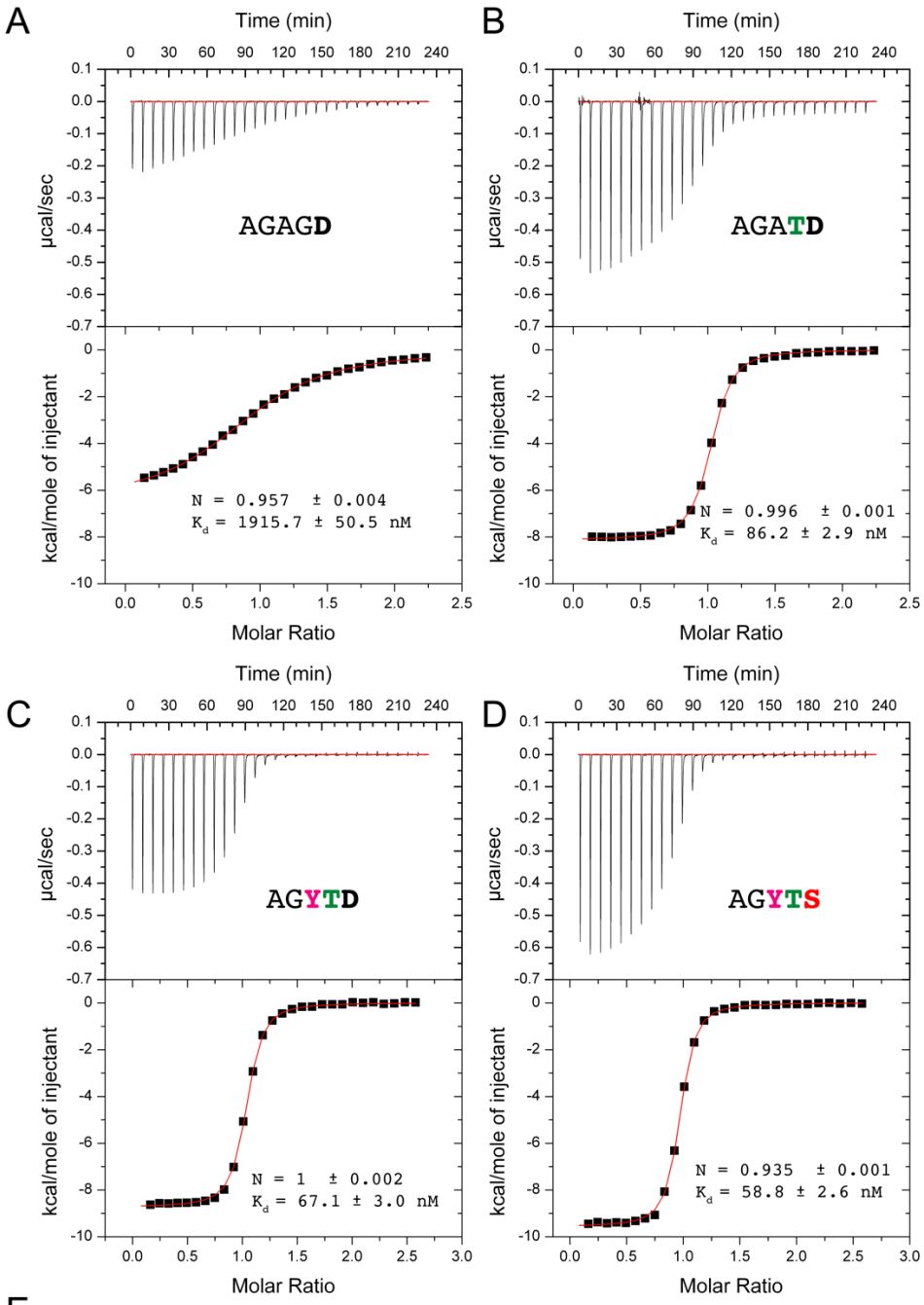
To further validate the peptidyl ligands identified by FP, we used isothermal titration calorimetry (ITC) to directly measure the interactions between unlabelled peptides (i.e. lacking FITC labels) and *Mp*PBD. ITC measurements of *Mp*IBP_RIII1-4 (~ 20 μM) titrated with the four different ligands (~ 200 μM) of AGAGD, AGATD, AGYTD,

and AGYTS all yielded sigmoidal-shaped curves with calculated stoichiometry values (N) close to 1 (Figure 5A-D). 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 μM 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, all resulted in sigmoidal-shaped ITC curves with much steeper transitions. The three peptides had calculated K_d values of 86 nM, 67 nM 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 nano-molar 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 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 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

415 protein compared to that of proline or glycine at the same position (Figure 3 vs
416 Figure 1). However, additional high-resolution structural information was required to
417 elucidate the basis for the stronger interactions between *Mp*PBD and its peptidyl
418 ligands AGYTD and AGYTS.

419



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
(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-)A_βDSTPD	N/A	N/A	N/A	N/A	~26600

~ 1,000x

420

Figure 5: Isothermal titration calorimetry results of the binding of four unlabelled peptidyl ligands AGAGD (A), AGATD (B), AGYTD (C), and AGYTS (D) to *Mp*IBP_RIII1-4. (E) Table showing the thermodynamic parameters of the binding of the four unlabeled peptidyl ligands to *Mp*PBD calculated from ITC. K_d by ITC and EC_{50} by FP for the five representative peptidyl ligands of *Mp*PBD 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.

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*IBP_RIII1-4 in complex with the peptides AGYTS and AGYTD to a resolution of 1.8 Å (Supplementary file 1-Table S1, Figure 6, Supplementary file 1-Figure S6). Consistent with the other *Mp*PBD-peptide complexes, AGYTD and AGYTS had their terminal α-carboxyl group in contact with Ca1 and Ca2 via three ionic bonds with average lengths of approximately 2.5 Å (Figure 6, Figure 7A; Supplementary file 1-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_βDSTD and FITC-A_βGPDSD (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 (Figure 7B; Supplementary file 1-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*IBP_RIII1-4 S295 and G296 with bond lengths of 2.5 Å and 2.8 Å, respectively (Figure 6A; Supplementary file 1-Table S4), while those at the same position for FITC-A_βGPDSD and FITC-A_βDSTD were at 3.2 Å and 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).

451 The tyrosine residues at position 3 of the YTD or YTS-containing peptides played a
452 key role in their tight interactions with *Mp*PBD. Given the hydrophobic nature of its
453 aryl side chain, the solvent-exposed peptide tyrosine appears to pack against Y294
454 of *Mp*PBD, which helps CPBLs to clench the peptide more tightly in the ligand-
455 binding cavity (Figure 6A, B). Indeed, the hydrogen bond lengths between the
456 peptide tyrosine main-chain oxygen (Tyr3-O) and protein S295 main-chain nitrogen
457 (S295-N) is 2.9 Å for AGYTD and 3 Å for AGYTS, which are shorter than those at
458 the same position for FITC-A$_\beta$DSTD (3.2 Å) and FITC-A$_\beta$GPDSD (3.3 Å)
459 (Supplementary file 1-Table S4). These subtle structural differences underlined the
460 molecular basis for the potent binding to *Mp*PBD by the peptides that end with YTD
461 or YTS compared to the weaker ligands (Figure 6C, D).
462
463 Taken together, the presence of a tyrosine residue at position 3 of AGYTS and
464 AGYTD stabilized the peptide-protein complexes via hydrophobic interactions. This
465 corroborated with the results that other FITC-AGXTD peptides with aromatic
466 residues at the 3rd position were also strong ligands for *Mp*PBD (Supplementary file
467 1-Table S3; Figure 4A). These atomic details explain why our structure-guided
468 screening approach was extremely effective in obtaining potent ligands with nano-
469 molar affinities for *Mp*PBD.
470

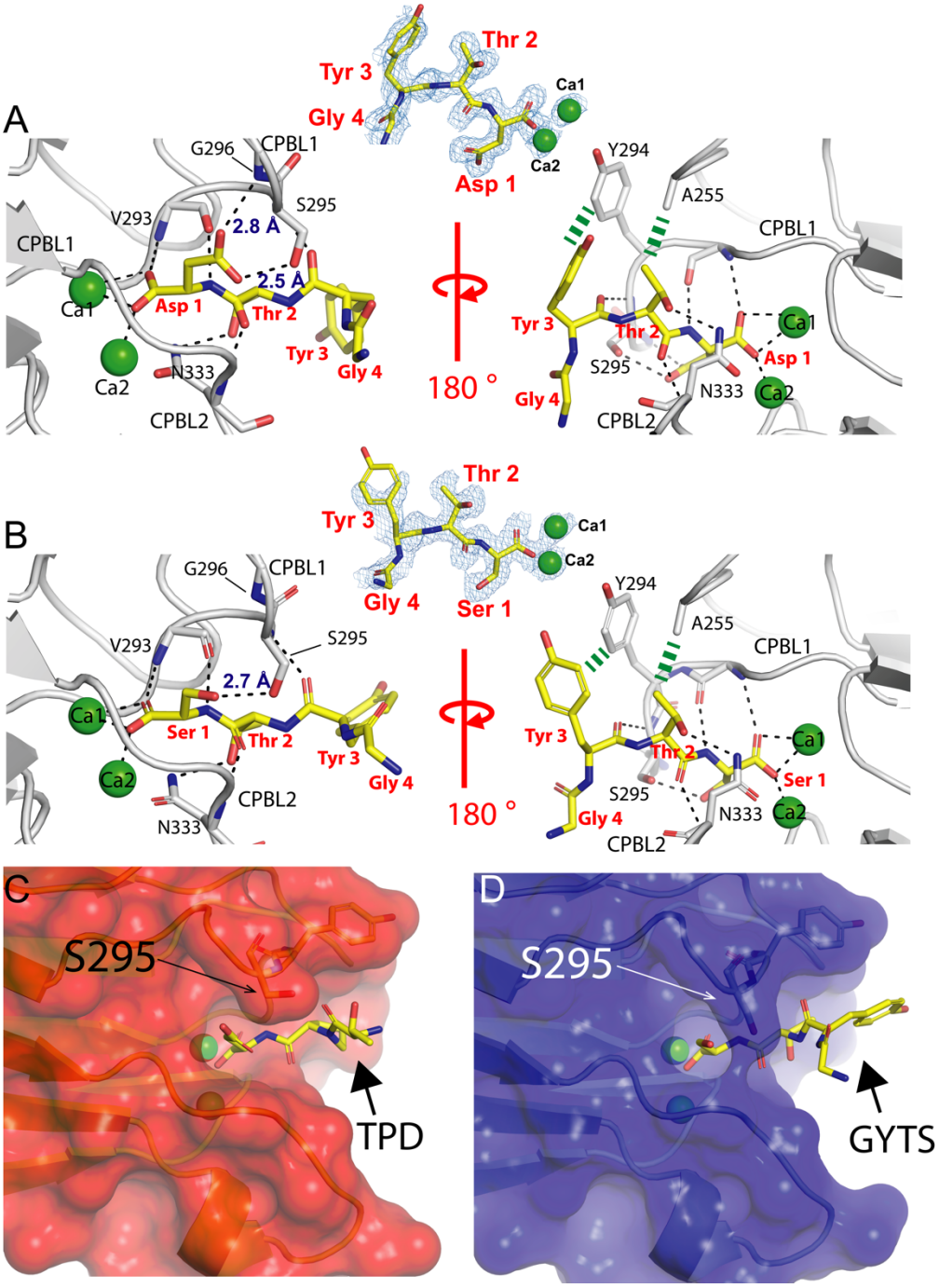


Figure 6: Atomic details of the interactions between *Mp*PBD and its nano-molar-affinity peptidyl ligands AGYTD (A) and AGYTS (B) revealed by X-ray crystallography. Color scheme is the same as in Figure 3. Thick green-dashed lines indicate hydrophobic contact between the protein and peptide tyrosine residues. The plasticity of the *Mp*PBD ligand-binding site is illustrated by their differences in complexing peptides that end with "TPD" (C) and "YTS" (D). Residue S295 that 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. CIF files can be found in Figure 6 – source data 2

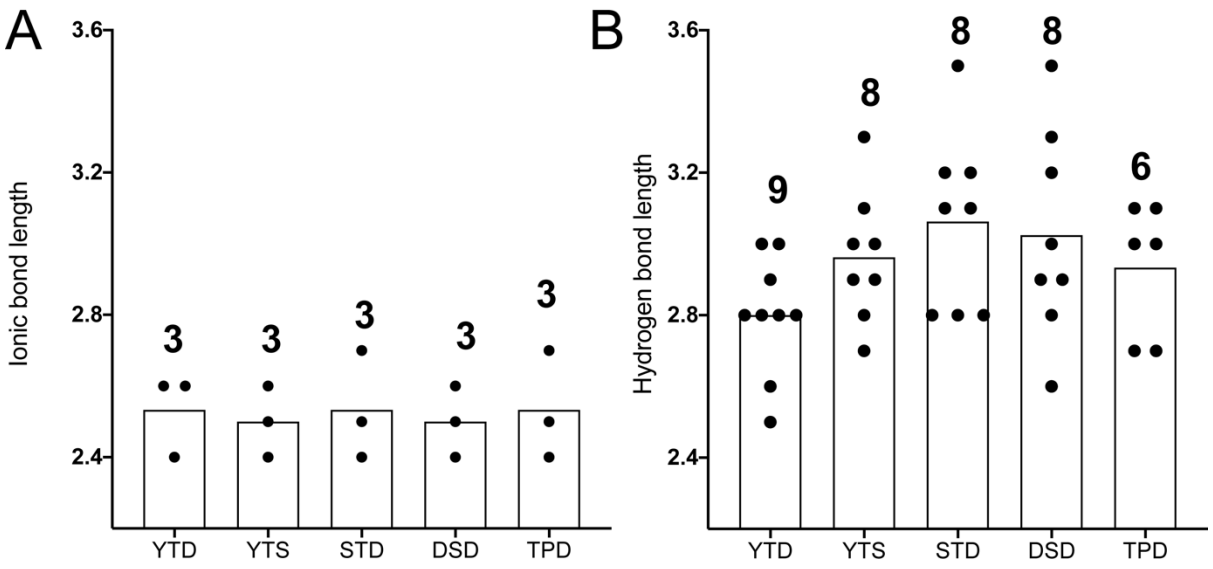


Figure 7: Length of the polar bonds between *Mp*PBD and its peptidyl ligands revealed by X-ray crystal structures of protein-peptide complexes. Data bars are the average lengths for ionic (A) and hydrogen (B) bonds. Black data points indicate the length of individual bonds between the protein and peptide. Numbers above the data bars indicate the number of bonds at the protein-peptide interfaces. Details of the bonds are listed in Supplementary file 1-Table S4.

AGYTD blocks MpPBD binding to diatoms

Having identified optimal peptidyl ligands for *Mp*PBD, we next tested their potential as antagonists to disrupt the PPI-mediated bacteria-diatom interactions that led to the discovery and characterization of this protein(27). Here we tested if AGYTD can block fluorescently labelled *Mp*IBP_RIII1-4 from binding to the Antarctic marine diatom *C. neogracile*(44). 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 (Figure 8A, Supplementary file 1-Figure S7). The binding of TRITC-labelled *Mp*IBP_RIII1-4 to *C. neogracile* resulted in red fluorescence evenly distributed around the cell centrally located inside of the diatom frustule (Figure 8B). At concentrations of 37.5 µM or above, AGYTD was extremely effective at blocking accumulation of the protein on the diatom, as the peptide outcompeted the cell surface proteins for binding *Mp*PBD and displaced ~ 95% of the fluorescent signal (Figure 8C, D, 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 8D, 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 nano-molar 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 multi-valent effect originating from the presence of many nearby protein binding sites on the diatom cell membrane.

Remarkably, AGYTD proved to be a potent antagonist that can displace *Mp*PBD pre-bound to the diatoms. The addition of 37.5 μM of AGYTD resulted in the dissociation of roughly 70% of the TRITC-labelled *Mp*IBP_RIII1-4 from the diatoms (Figure 8D, grey bars). Although the peptide was more effective at the same concentration when used as a prophylactic (Figure 8D, black bar), this important result demonstrated the potential of the peptide antagonists in disrupting pre-existing bacteria-host interactions, like those involved in biofilms.

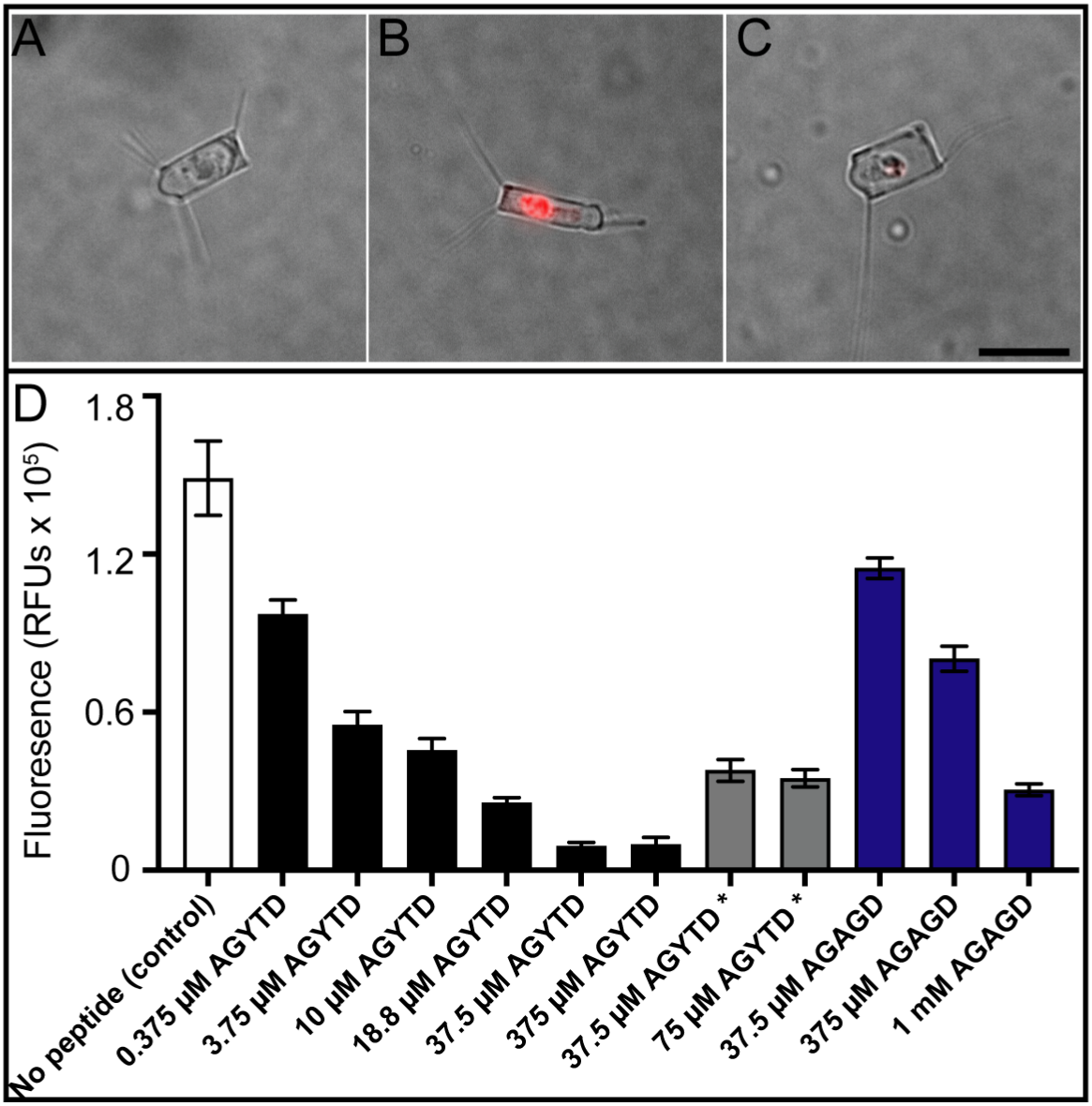


Figure 8: Inhibition of binding of *Mp*PBD to diatoms by competing peptides. Representative images showing: (A) an untreated *C. neogracile*; (B) a *C. neogracile* cell treated with 0.2 mg/mL TRITC-labelled *Mp*IBP_RIII1-4: (C) a *C. neogracile* cell treated with 0.2 mg/mL TRITC-labelled *Mp*IBP_RIII1-4 in the presence of 375 µM AGYTD. All three images (A) – (C) are at the same magnification with the black scale bar in (C) indicating 10 µm. (D) Fluorescence levels shown by the *C. neogracile* cells alone and those treated with TRITC-labelled *Mp*IBP_RIII1-4 and various concentrations of a strong or weak ligand, AGYTD and AGAGD, respectively. Each bar represents the quantification of average fluorescence from 30 individual diatoms. The white bar represents the control experiment where diatoms were treated with TRITC-labelled *Mp*IBP_RIII1-4 in the absence of peptide. Black bars represent experiments where diatoms were treated with 0.2 mg/mL TRITC-labelled *Mp*IBP_RIII1-4 and indicated concentrations of AGYTD. Grey bars with an asterisk underneath represent experiments where diatoms were incubated with 0.2 mg/mL TRITC-labelled *Mp*IBP_RIII1-4 before AGYTD was added. Blue bars represent experiments where diatoms were treated with 0.2 mg/mL TRITC-labelled *Mp*IBP_RIII1-4 and indicated concentrations of AGAGD. Additional representative images are shown in Fig.S7.

Conclusion and outlook:

One major challenge in the field of PPI modulator discovery is a lack of starting points for the initial 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. 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.

Inhibition of pathogenic bacterial adhesion to human cells has yielded promising results in combating infections(17, 35, 38, 40, 45-47). 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. Furthermore, given that adhesins are localized to the cell surfaces, their antagonists are not required to penetrate or be taken into the cells, simplifying their applications. These key attributes justified the pursuit of novel adhesin modulators to disrupt PPI-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*(27, 28, 48), the causative agent of cholera and others such as *Aeromonas veronii* that cause infections ranging from diarrhoea to wound infections and sepsis(49, 50), and the flesh-eating bacterium *Vibrio vulnificus*(51) (Figure S8). Therefore, there is an opportunity to use this conceptual approach to devise antagonists to disrupt the PPIs involved in bacterial infections. With the emergence of antibiotic-resistant pathogens, this work gives insight into how microbes might be

567 controlled through the modulation of PPI-related adhesion to their hosts. In future
568 research, it will be of interest to design adhesin modulators consisting of D-peptides,
569 which are non-metabolizable and are thus resistant to proteolytic degradation, as
570 well as cyclic peptides that might achieve higher affinity via favourable entropic
571 contributions.

572 **Materials and Methods:**

573

574 **Peptide synthesis**

575 FITC-labelled peptides FITC-A$_\beta$GPDSD, FITC-A$_\beta$DSTPD, FITC-A$_\beta$DSTD and FITC-
576 A$_\beta$GPDD were synthesized by Fmoc solid-phase peptide synthesis, either manually
577 or using an automated Intavis MultiPep RSi peptide synthesizer. The protected
578 amino acids (linked to Wang resins) and FITC were purchased from Novabiochem
579 and Sigma-Aldrich. Crude peptides were then analyzed and purified by high-
580 pressure liquid chromatography (HPLC) using a preparative reversed-phase column
581 with MS detection. The peptides were freeze-dried and stored at – 30 ℃. All other
582 FITC-labelled peptides used in the structure-guided ligand optimization procedures
583 by fluorescence polarization (FP) and the four unlabeled peptides (AGAGD, AGYTD,
584 ATYTS and AGATD) used in the ITC measurements and X-ray crystallography were
585 purchased from GenicBio (Shanghai, China). HPLC and MS spectra for
586 representative peptides are shown in Supplementary file 1-Figure S9-S21.

587

588 **Fluorescence polarization**

589 The FITC-labelled peptides were dissolved in FP buffer (10 mM HEPES, pH 7.4, 150
590 mM NaCl, 0.1% Tween 20, 1 mg/mL BSA) to a final concentration of 10 nM. Dilution
591 series of *Mp*IBP_RIII1-4 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). 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).

Co-crystallization, X-ray diffraction and structure solutions of *Mp*PA14 with peptide ligands

The original *Mp*IBP_RIII1-4 construct 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 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*IBP_RIII1-4 was produced, purified, and crystallized as previously described(27, 52). Co-crystallization of *Mp*IBP_RIII1-4 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)(53) and CCP4-Aimless(54) or the

617 DIALS/xia2(55, 56) in the CCP4i2 software suite. The structure solutions for all
618 complexes were obtained by using molecular replacement using the *Mp*IBP_RIII1-4
619 structure (PBD: 5K8G) as the search model(27). The structures were refined using
620 CCP4-Refmac5(57).

621 **Isothermal Titration Calorimetry**

622 Isothermal calorimetric titration (ITC) measurements were performed at 30 ℃ using a
623 MicroCal VP-ITC instrument (Malvern). *Mp*IBP_RIII1-4 was dialyzed overnight in a
624 buffer of 50 mM Tris-HCl, pH 9, 150 mM NaCl, 5 mM $CaCl_2$. Next, the protein was
625 diluted to approximately 20 µM and was mixed with serial 5-µl aliquots of 200 µM of
626 each of the four peptide solutions (AGAGD, AGATD, AGYTD and AGYTS). Peptide
627 solutions were automatically added by a rotating syringe (400 RPM) at 5-min
628 intervals into the *Mp*IBP_RIII1-4 solution for a total of 50 injections. The data were
629 analyzed by Origin software Version 5.0 (MicroCal).
630

631 **Diatom binding experiments**

632 The Antarctic diatom, *Chaetoceros neogracile,* was cultured as previously
633 described(27, 44). TRITC-labelled *Mp*IBP_RIII1-4 (TRITC-*Mp*IBP_RIII1-4, 0.2
634 mg/mL) in the presence or absence of peptides was incubated with diatoms in buffer
635 (50 mM Tris-HCl pH 9, 300 mM NaCl, 5 mM $CaCl_2$) with gentle mixing for 2 h. Next,
636 diatoms were pelleted by centrifugation for 3 min at 4,500 x g, and the resulting
637 supernatant was discarded. This procedure was repeated three times to wash out
638 unbound TRITC-*Mp*IBP_RIII1-4 before the diatom pellet was finally resuspended in
639 20 µL of buffer, which was then examined on slides by fluorescence microscopy. In
640 parallel experiments to test if the strong ligand AGYTD could compete off the TRITC-
641 *Mp*IBP_RIII1-4 that was already bound to diatoms, TRITC-*Mp*IBP_RIII1-4 was

642 incubated with diatom for 1.5 h before the peptide AGYTD was added. The
643 remainder of the experiment followed the same procedure as described above.
644
645 Images were obtained using an Olympus IX83 inverted fluorescence microscope
646 equipped with an Andor Zyla 4.2 Plus camera. Quantification of the fluorescence
647 intensity was done using Fiji ImageJ. The corrected total cell fluorescence (CTCF)
648 was calculated using the formula: CTCF = Integrated Density – (Area of selected cell
649 x Mean Fluorescence of the background)(58). Quantification of 30 individual diatom
650 cells was done for each treatment. As diatom cell aggregates produce overexposed
651 fluorescence while those cells lacking a silica frustule have damaged plasma
652 membranes necessary for protein binding, they were excluded from the
653 measurements. Graphs were made using Graphpad Prism.
654
655 **Data availability**
656 X-ray crystal structure coordinates solved in this study have been deposited in the
657 Protein Data Bank with accession codes of 6X6Q (*Mp*IBP_RIII1-4 - FITC-
658 A_{β}GPDSD), 6X6M (*Mp*IBP_RIII1-4 - FITC-A_{β}DSTD), 6X5W (*Mp*IBP_RIII1-4 -
659 AGYTD), and 6X5V (*Mp*IBP_RIII1-4 - AGYTS). The data that support the findings of
660 this study are available from the corresponding author P.L.D upon reasonable
661 request.
662
663 **Acknowledgements**
664 We thank Dr. John Allingham for access to his home source X-ray diffractometer at
665 Queen's University, and to staff members at the Canadian Light Source in
666 Saskatoon, Canada and the PETRA III facility at DESY in Hamburg, Germany for

667 access to data collection at these synchrotrons. We are grateful to Dr. EonSeon Jin,
668 Hanyang University, Seoul, for the gift of the diatom, *Chaetoceros neogracile* and Dr.
669 Saeed Rismani Yazdi for assistance with the diatom cultures. We are indebted to Mr.
670 Kim Munro at the Protein Function Discovery unit at Queen's University for his
671 assistance with acquiring and interpreting ITC data. We thank Ms. Irene van Oekel
672 for preliminary tests on the binding of peptidyl ligands to *Mp*PBD and Mr. Joost van
673 Dongen for analytical support. This project was funded by a Natural Science and
674 Engineering Research Council (NSERC, http://www.nserc-crsng.gc.ca/index_eng.
675 asp) Discovery Grant (RGPIN-2016-04810) to PLD who holds the Canadian
676 Research Chair in Protein Engineering. IKV acknowledges financial support by the
677 European Union (ERC-2014-StG Contract No.635928) and the Dutch Science
678 Foundation (NWO ECHO Grant No. 712.016.002).

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

687 **Competing interests:** The authors declare that they have no competing interests.

688 **References:**
689
690 1.Winder SJ, Ayscough KR. Actin-binding proteins. J Cell Sci. 2005;118(Pt 4):651-4.
691 2.Cunha CA, Romao MJ, Sadeghi SJ, Valetti F, Gilardi G, Soares CM. Effects of protein-protein
692 interactions on electron transfer: docking and electron transfer calculations for complexes
693 between flavodoxin and c-type cytochromes. J Biol Inorg Chem. 1999;4(3):360-74.

3.Pawson T, Nash P. Protein-protein interactions define specificity in signal transduction. Genes Dev. 2000;14(9):1027-47.
4.Arkin MR, Tang Y, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. Chem Biol. 2014;21(9):1102-14.
5.Andrei SA, Sijbesma E, Hann M, Davis J, O'Mahony G, Perry MWD, et al. Stabilization of protein-protein interactions in drug discovery. Expert Opin Drug Discov. 2017;12(9):925-40.
6.Milroy LG, Grossmann TN, Hennig S, Brunsveld L, Ottmann C. Modulators of protein-protein interactions. Chem Rev. 2014;114(9):4695-748.
7.Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, et al. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature. 2005;435(7042):677-81.
8.Roberts AW, Davids MS, Pagel JM, Kahl BS, Puvvada SD, Gerecitano JF, et al. Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. N Engl J Med. 2016;374(4):311-22.
9.Dutta S, Ryan J, Chen TS, Kougentakis C, Letai A, Keating AE. Potent and specific peptide inhibitors of human pro-survival protein Bcl-xL. J Mol Biol. 2015;427(6 Pt B):1241-53.
10.Chang YS, Graves B, Guerlavais V, Tovar C, Packman K, To KH, et al. Stapled alpha-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. Proc Natl Acad Sci U S A. 2013;110(36):E3445-54.
11.Brancaccio D, Di Maro S, Cerofolini L, Giuntini S, Fragai M, Luchinat C, et al. HOPPI-NMR: Hot-Peptide-Based Screening Assay for Inhibitors of Protein-Protein Interactions by NMR. ACS Med Chem Lett. 2020;11(5):1047-53.
12.Drucker DJ. Advances in oral peptide therapeutics. Nat Rev Drug Discov. 2020;19(4):277-89.
13.Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. Drug Discov Today. 2015;20(1):122-8.
14.Wojcik P, Berlicki L. Peptide-based inhibitors of protein-protein interactions. Bioorg Med Chem Lett. 2016;26(3):707-13.
15.Davies J, Davies D. Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev. 2010;74(3):417-33.
16.Frieri M, Kumar K, Boutin A. Antibiotic resistance. J Infect Public Health. 2017;10(4):369-78.
17.Krachler AM, Orth K. Targeting the bacteria-host interface: Strategies in anti-adhesion therapy. Virulence. 2013;4(4):284-94.
18.Ventola CL. The antibiotic resistance crisis: part 1: causes and threats. P T. 2015;40(4):277-83.
19.Klemm P, Schembri MA. Bacterial adhesins: function and structure. Int J Med Microbiol. 2000;290(1):27-35.
20.Kline KA, Falker S, Dahlberg S, Normark S, Henriques-Normark B. Bacterial adhesins in host-microbe interactions. Cell Host Microbe. 2009;5(6):580-92.
21.Stones DH, Krachler AM. Fatal attraction: how bacterial adhesins affect host signaling and what we can learn from them. Int J Mol Sci. 2015;16(2):2626-40.
22.Newell PD, Boyd CD, Sondermann H, O'Toole GA. A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. PLoS Biol. 2011;9(2):e1000587.
23.Jamal M, Ahmad W, Andleeb S, Jalil F, Imran M, Nawaz MA, et al. Bacterial biofilm and associated infections. J Chin Med Assoc. 2018;81(1):7-11.

741 24.Lebeaux D, Ghigo JM, Beloin C. Biofilm-related infections: bridging the gap between
742 clinical management and fundamental aspects of recalcitrance toward antibiotics. Microbiol
743 Mol Biol Rev. 2014;78(3):510-43.
744 25.Romling U, Balsalobre C. Biofilm infections, their resilience to therapy and innovative
745 treatment strategies. J Intern Med. 2012;272(6):541-61.
746 26.Guo S, Langelaan DN, Phippen SW, Smith SP, Voets IK, Davies PL. Conserved structural
747 features anchor biofilm-associated RTX-adhesins to the outer membrane of bacteria. FEBS J.
748 2018;285(10):1812-26.
749 27.Guo S, Stevens CA, Vance TDR, Olijve LLC, Graham LA, Campbell RL, et al. Structure of a
750 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. Sci Adv.
751 2017;3(8):e1701440.
752 28.Guo S, Vance TDR, Stevens CA, Voets IK, Davies PL. RTX Adhesins are Key Bacterial
753 Surface Megaproteins in the Formation of Biofilms. Trends Microbiol. 2019;27(5):470.
754 29.Smith TJ, Sondermann H, O'Toole GA. Type 1 Does the Two-Step: Type 1 Secretion
755 Substrates with a Functional Periplasmic Intermediate. J Bacteriol. 2018;200(18).
756 30.Berne C, Ducret A, Hardy GG, Brun YV. Adhesins Involved in Attachment to Abiotic
757 Surfaces by Gram-Negative Bacteria. Microbiol Spectr. 2015;3(4).
758 31.Kitts G, Giglio KM, Zamorano-Sanchez D, Park JH, Townsley L, Cooley RB, et al. A
759 Conserved Regulatory Circuit Controls Large Adhesins in Vibrio cholerae. mBio. 2019;10(6).
760 32.Vance TDR, Guo SQ, Assaie-Ardakany S, Conroy B, Davies PL. Structure and functional
761 analysis of a bacterial adhesin sugar-binding domain (vol 14, e0220045, 2019). Plos One.
762 2019;14(8).
763 33.Moonens K, Gideonsson P, Subedi S, Bugaytsova J, Romao E, Mendez M, et al. Structural
764 Insights into Polymorphic ABO Glycan Binding by Helicobacter pylori. Cell Host Microbe.
765 2016;19(1):55-66.
766 34.Moonens K, Hamway Y, Neddermann M, Reschke M, Tegtmeyer N, Kruse T, et al.
767 Helicobacter pylori adhesin HopQ disrupts trans dimerization in human CEACAMs. EMBO J.
768 2018;37(13).
769 35.Wellens A, Garofalo C, Nguyen H, Van Gerven N, Slattegard R, Hernalsteens JP, et al.
770 Intervening with urinary tract infections using anti-adhesives based on the crystal structure
771 of the FimH-oligomannose-3 complex. Plos One. 2008;3(4):e2040.
772 36.Ofek I, Hasty DL, Sharon N. Anti-adhesion therapy of bacterial diseases: prospects and
773 problems. FEMS Immunol Med Microbiol. 2003;38(3):181-91.
774 37.Han Z, Pinkner JS, Ford B, Obermann R, Nolan W, Wildman SA, et al. Structure-based
775 drug design and optimization of mannoside bacterial FimH antagonists. J Med Chem.
776 2010;53(12):4779-92.
777 38.Mydock-McGrane LK, Hannan TJ, Janetka JW. Rational design strategies for FimH
778 antagonists: new drugs on the horizon for urinary tract infection and Crohn's disease. Expert
779 Opin Drug Discov. 2017;12(7):711-31.
780 39.Sauer MM, Jakob RP, Luber T, Canonica F, Navarra G, Ernst B, et al. Binding of the
781 Bacterial Adhesin FimH to Its Natural, Multivalent High-Mannose Type Glycan Targets. J Am
782 Chem Soc. 2019;141(2):936-44.
783 40.Totsika M, Kostakioti M, Hannan TJ, Upton M, Beatson SA, Janetka JW, et al. A FimH
784 inhibitor prevents acute bladder infection and treats chronic cystitis caused by multidrug-
785 resistant uropathogenic Escherichia coli ST131. J Infect Dis. 2013;208(6):921-8.

786 41.Spaulding CN, Klein RD, Ruer S, Kau AL, Schreiber HL, Cusumano ZT, et al. Selective
787 depletion of uropathogenic E. coli from the gut by a FimH antagonist. Nature.
788 2017;546(7659):528-32.
789 42.Kiessling LL. Chemistry-driven glycoscience. Bioorg Med Chem. 2018;26(19):5229-38.
790 43.Guo S, Garnham CP, Whitney JC, Graham LA, Davies PL. Re-evaluation of a bacterial
791 antifreeze protein as an adhesin with ice-binding activity. Plos One. 2012;7(11):e48805.
792 44.Gwak IG, Jung WS, Kim HJ, Kang SH, Jin E. Antifreeze Protein in Antarctic Marine Diatom,
793 Chaetoceros neogracile. Mar Biotechnol. 2010;12(6):630-9.
794 45.Huebinger RM, Do DH, Carlson DL, Yao X, Stones DH, De Souza Santos M, et al. Bacterial
795 adhesion inhibitor prevents infection in a rodent surgical incision model. Virulence.
796 2020;11(01):695-706.
797 46.Cusumano CK, Pinkner JS, Han Z, Greene SE, Ford BA, Crowley JR, et al. Treatment and
798 prevention of urinary tract infection with orally active FimH inhibitors. Sci Transl Med.
799 2011;3(109):109ra15.
800 47.Solanki V, Tiwari M, Tiwari V. Host-bacteria interaction and adhesin study for
801 development of therapeutics. Int J Biol Macromol. 2018;112:54-64.
802 48.Syed KA, Beyhan S, Correa N, Queen J, Liu J, Peng F, et al. The Vibrio cholerae flagellar
803 regulatory hierarchy controls expression of virulence factors. J Bacteriol.
804 2009;191(21):6555-70.
805 49.Hickman-Brenner FW, MacDonald KL, Steigerwalt AG, Fanning GR, Brenner DJ, Farmer JJ,
806 3rd. Aeromonas veronii, a new ornithine decarboxylase-positive species that may cause
807 diarrhea. J Clin Microbiol. 1987;25(5):900-6.
808 50.Whitaker IS, Kamya C, Azzopardi EA, Graf J, Kon M, Lineaweaver WC. Preventing infective
809 complications following leech therapy: is practice keeping pace with current research?
810 Microsurgery. 2009;29(8):619-25.
811 51.Hollis DG, Weaver RE, Baker CN, Thornsberry C. Halophilic Vibrio species isolated from
812 blood cultures. J Clin Microbiol. 1976;3(4):425-31.
813 52.Guo S, Campbell R, Davies PL, Allingham JS. Phasing with calcium at home. Acta
814 Crystallogr F Struct Biol Commun. 2019;75(Pt 5):377-84.
815 53.Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta
816 Crystallogr D Biol Crystallogr. 2010;66(Pt 2):133-44.
817 54.Evans P. Scaling and assessment of data quality. Acta Crystallogr D. 2006;62:72-82.
818 55.Beilsten-Edmands J, Winter G, Gildea R, Parkhurst J, Waterman D, Evans G. Scaling
819 diffraction data in the DIALS software package: algorithms and new approaches for multi-
820 crystal scaling. Acta Crystallogr D Struct Biol. 2020;76(Pt 4):385-99.
821 56.Winter G, Lobley CM, Prince SM. Decision making in xia2. Acta Crystallogr D Biol
822 Crystallogr. 2013;69(Pt 7):1260-73.
823 57.Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, et al. REFMAC5
824 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta
825 Crystallogr D. 2004;60:2184-95.
826 58.Burgess A, Vigneron S, Brioudes E, Labbe JC, Lorca T, Castro A. Loss of human Greatwall
827 results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B-
828 Cdc2/PP2A balance. P Natl Acad Sci USA. 2010;107(28):12564-9.
829