Re-Evaluation of a Bacterial Antifreeze Protein as an Adhesin with Ice-Binding Activity

Shuaiqi Guo, Christopher P. Garnhama, John C. Whitneyb, Laurie A. Graham, Peter L. Davies*

Protein Function Discovery Group and the Department of Biomedical and Molecular Sciences, Queen’s University, Kingston, Ontario, Canada

Abstract

A novel role for antifreeze proteins (AFPs) may reside in an exceptionally large 1.5-MDa adhesin isolated from an Antarctic Gram-negative bacterium, Marinomonas primoryensis. MpAFP was purified from bacterial lysates by ice adsorption and gel electrophoresis. We have previously reported that two highly repetitive sequences, region II (RII) and region IV (RIV), divide MpAFP into five distinct regions, all of which require mM Ca2+ levels for correct folding. Also, the antifreeze activity is confined to the 322-residue RIV, which forms a Ca2+-bound beta-helix containing thirteen Repeats-in-Toxin (RTX)-like repeats. RII accounts for approximately 90% of the mass of MpAFP and is made up of ~120 tandem 104-residue repeats. Because these repeats are identical in DNA sequence, their number was estimated here by pulsed-field gel electrophoresis. Structural homology analysis by the Protein Homology/analogy Recognition Engine (Phyre2) server indicates that the 104-residue RII repeat adopts an immunoglobulin beta-sandwich fold that is typical of many secreted adhesion proteins. Additional RTX-like repeats in RIV may serve as a non-cleavable signal sequence for the type I secretion pathway. Immunodetection shows both repeated regions are uniformly distributed over the cell surface. We suggest that the development of an AFP-like domain within this adhesin attached to the bacterial outer surface serves to transiently bind the host bacteria to ice. This association would keep the bacteria within the upper reaches of the water column where oxygen and nutrients are potentially more abundant. This novel envirotactic role would give AFPS a third function, after freeze avoidance and freeze tolerance: that of transiently binding an organism to ice.

Citation: Guo S, Garnham CP, Whitney JC, Graham LA, Davies PL (2012) Re-Evaluation of a Bacterial Antifreeze Protein as an Adhesin with Ice-Binding Activity. PLoS ONE 7(11): e48805. doi:10.1371/journal.pone.0048805

Editor: Michael Hensel, University of Osnabrueck, Germany

Received August 2, 2012; Accepted October 5, 2012; Published November 7, 2012

Copyright: © 2012 Guo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by a grant from the Canadian Institutes of Health Research (http://www.cihr-irsc.gc.ca/e/193.html) to PLD. CPG was the recipient of an Natural Sciences and Engineering Research Council-PGSD3 scholarship (http://www.nserc-crsng.gc.ca/index_eng.asp) and an R. Samuel McLaughlin fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: peter.davies@queensu.ca

† Current address: National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, United States of America

‡ Current address: Molecular Structure and Function, The Hospital for Sick Children, Toronto, Ontario, Canada, and Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada

Introduction

Antifreeze proteins (AFPs) were initially characterized in marine fishes [1], [2] where they protect their hosts from freezing by binding to, and preventing the growth of, seed ice crystals [3]. AFPs lower the freezing temperature of a solution containing ice below the melting point of the ice. This difference between the freezing and melting temperatures is called thermal hysteresis (TH) and is used as a measure of antifreeze activity. AFPs were subsequently found in freeze-tolerant organisms [4], [5] where, rather than preventing freezing, they stop ice crystals in frozen tissues from growing larger through the process of ice recrystallization (IRI) [6].

The bacterium Marinomonas primoryensis, isolated from a brackish, ice-covered lake in Antarctica, produces an exceptionally large protein (ca. 1.5 MDa) with Ca2+-dependent antifreeze activity [7], [8]. The protein contains two highly repetitive segments, Regions II and IV (RII and RIV), that divide it into five distinct regions (RI–V) [7]. RIV, which contains thirteen 19-aa repeats in tandem, comprising ~2% of the entire protein, is the only region with antifreeze activity. We recently solved the X-ray crystal structure of MpAFP Region IV to 1.7 Å [9]. This segment of the protein folds as an extended, Ca2+-bound right-handed beta-helix whose ice-binding site (IBS) consists of a flat, repetitive array of outward-projecting Thr and Asx residues. The IBS organizes water molecules into a regular ice-like lattice that matches, at a minimum, the primary prism and basal planes of ice. Experimental observation of these ordered surface waters provided strong physical evidence for a mechanism of ice binding that was originally predicted by molecular modeling [10], [11], [12], [13]. This may well be a general mechanism of action for all ice-binding proteins where the IBS orders surface waters into an ice-like “anchored clathrate” pattern that then helps “freeze” the AFPs to the ice surface [9].

Both the TH and IRI activities of AFPs are concentration dependent. Fish typically produce 10-20 mg/ml of AFPs to depress their freezing by the ~1°C needed to survive in ice-laden seawater [14]. Although fusion proteins can be effective AFPs if the added domain does not occlude the IBS [15], from the perspective of biological efficiency, AFPs are typically small, single domain proteins produced from large multi-gene families. Only in this way can the host produce the millimolar AFP concentrations needed to prevent freezing. It is telling, therefore, that the AFP-active RIV makes up only 2% of the residues in the whole protein.
This strongly suggests the primary function of the protein is not that of TH or IRI. To better understand the function of *MpAFP*, we have examined other regions of the protein.

Here we report the purification and characterization of wild-type *MpAFP*. Two lambda clones, one from each end of the gene, were sequenced, to reveal both the flanking genes and the domain structure of the AFP. We used Southern blotting of pulsed-field gel separated DNA to ascertain the full extent of RII, and show it makes up more than 90% of the mass of this incredibly large protein. We have derived partial or complete homology models for each of the five regions of *MpAFP* consistent with its role as an adhesin secreted via its RTX repeats. Using immunofluorescence we have confirmed the location of the protein on the outer surface of the host bacterium, and suggest that mutation and amplification of tandem RTX repeats within Region IV of the adhesin has formed the AFP-like domain that serves to dock the bacterium to ice. This suggests a novel function for an AFP: that of simply binding to ice, rather than preventing its growth or recrystallization.

**Materials and Methods**

*M. primoryensis* was cultured and the crude lysate prepared as previously described [8], with the exception that the lysis buffer contained 25 mM Tris-HCl (pH 8.0) and 20 mM CaCl₂. The soluble portion of this lysate was adjusted to 70% ammonium sulfate then centrifuged at 13,000 × g for 30 min at 4°C. The precipitate was resuspended in lysis buffer (50 ml) and dialysed against the same buffer. This material was subjected to ice affinity purification (IAP) [16]. Ice was slowly grown on a cold finger held at −0.5°C for 30 min, then the temperature was gradually decreased to −2.5°C over 48 h until ~50% of the original volume was frozen. The ice fraction was melted and adjusted to 25 mM Tris-HCl (pH 8.0) and 20 mM CaCl₂, before being subjected to a second round of IAP, as above. The second ice fraction was then concentrated to 2 ml by dry dialysis in 3,500 molecular weight cut-off dialysis tubing exposed to PEG 8000. This concentrate was then analyzed by standard PAGE under both native and denaturing conditions and the AFP detected using either the cationic carbocyanine dye “Stains-All” (Sigma) or Coomassie blue. Stains-All has been shown to stain Ca²⁺-binding proteins dark blue or purple while staining other proteins red or pink [17].

**Tandem mass spectrometry analysis**

Pure *MpAFP* was resolved by standard SDS-PAGE (10% (w/v) stacking and 4% (w/v) resolving gels) and visualized by Coomassie blue staining. The AFP band was excised and the gel plugs were trypsin digested on a Waters MassPREP station using method 5.7S with reduction (dithiothreitol) and alkylation (iodoacetamide) of trypsin digested on a Waters MassPREP station using method 5.7S. The digestion was followed by drying the gel plugs and resuspension in 0.1% (v/v) formic acid. An initial TOF-MS survey scan was acquired over the range m/z 400–1600 (from which tryptic fragment masses were determined), and the Q-TOF was programmed to ignore singly charged ions while collecting MS-MS data on up to three co-eluting species. The spectra were smoothed, converted to a list of m/z centroids and submitted to the MASCOT (www.matrixscience.com) and Proteinlynx Global Server 2.0 (www.waters.com) search engines for database comparison. Those peaks that were not identified as trypsin or keratin were manually sequenced.

**Genomic DNA extraction**

The CTAB method of chromosomal DNA extraction [18] was used on a 50-ml culture of *Marinomonas primoryensis* grown for 5 days at 4°C in 50% (w/v) SWB (19 g/l sea salt (Sigma); 1 g/l Tryptone; 1 g/l yeast extract) as above. This DNA was used in subsequent PCR reactions and in the construction of a genomic library.

**Amplification of a fragment of *MpAFP* sequence**

Two fully-degenerate primers were designed based on amino acid sequences determined above, The sense primer 5’-GAYGC-NACNNTYGARGCNCGNAA-3’ corresponds to DATFEEAN. The antisense primer 5’-TCRTCRTTNCNCGTNCCNGCRTG-3’ corresponds to DAGTGNDE. PCR conditions using 3 μM of each primer were as follows: 30 cycles of 95°C for 30 s, 50°C for 1 min and 72°C for 90 sec with a final extension at 72°C for 8 min. The resulting product was purified by gel extraction (Qiagen gel extraction kit), cloned using the TOPO TA kit (Invitrogen), and sequenced at the Cortec DNA Service Laboratories, Kingston, Ontario. Additional sequence was obtained by inverse PCR but ultimately, a more complete sequence was obtained as described below.

**Genomic Lambda library construction and analysis**

A genomic Lambda Dash II library was constructed from *M. primoryensis* DNA partially digested with Sau3A I (BioS&T, Montréal, Canada). It afforded ~16-fold coverage of the genome as it had a titre of 4 × 10⁶ pfu/ml with an average insert size of 20 kb. A PCR product, corresponding to bases 3634–4053 of GenBank Accession ABL74377, was labeled with [α-³²P]-dCTP using a random priming DNA labeling system (Invitrogen, Carlsbad, California) and was used to screen the library by standard methodologies [19]. Phage DNA was isolated by the CsCl gradient technique [10]. DNA insert size was determined by digestion with SalI. After this clone was sequenced (below) it was necessary to isolate a second clone to obtain the 5’ end of the *MpAFP* gene. A PCR product corresponding to two repeats from the highly-repetitive section RII found at the end of the first clone was amplified (bases 1946–2569 of Genbank Accession ABL74378). A clone which hybridized to this probe, but not to the probe used earlier, was isolated.

**Sequencing of lambda clones**

Lambda phage DNA was mechanically sheared and shotgun cloned into pUC19 vector (Genome Quebec, McGill University, Montréal, Canada). A total of 288 randomly selected clones were initially sequenced using the M13 forward primer. Gaps in the sequence were closed by sequencing relevant clones with either the M13 reverse primer or by sequence walking. In total, 390 sequence reads were performed to ensure double coverage of all regions except the highly repetitive region.

**Pulsed-field gel electrophoresis**

The CHEF Bacterial Genomic DNA Plug Kit (catalog 170–3592) (Bio-Rad laboratories, Hercules, California) was used to prepare agarose plugs containing *M. primoryensis* for in-gel restriction endonuclease digestion. The kit was used according to manufacturer's instructions except that the cells were resuspended in a higher salt buffer (10 mM Tris-HCl (pH 7.2), 330 mM NaCl, and 150 mM EDTA (pH 8.0)) prior to agarose addition. Digests
were also performed according to kit instructions using the restriction enzymes PstI, AelI, MboI, AluI and MpiI (New England Biolabs). After washing the plugs in gel buffer (0.5X TBE), they were embedded in a 20-cm-long 1% (w/v) agarose gel. The gel was run in a CHEF-DR® II Pulsed Field Electrophoresis System (Bio-Rad laboratories, Hercules, California) at 120 V for 22 h with a linearly ramped 50-90 s switch time during the length of the run. The temperature was maintained at ca. 4°C. The gel was stained with a 1 μg/ml solution of ethidium bromide to allow visualization of DNA samples and standards.

Southern blotting

The PFGE gel from above was blotted onto a Zeta-Probe® membrane (Bio-Rad laboratories, Hercules, California) by the alkaline capillary method [18] and probed with the repeat-containing fragment as above. After washing to remove excess probe, Kodak BioMax XAR film was exposed to the membrane for 16 h.

Homology modeling of MpAFP domains

Sections of the MpAFP sequence were submitted to the Phyre2 server [20]. Phyre2 uses the hidden Markov method to generate alignments of a submitted protein sequence against proteins with published structures [21]. The resulting alignments are then used to produce homology-based models of the query sequence to predict its three-dimensional structure. In addition, Phyre2 uses an ab initio folding simulation called Poing to model regions of a query with no detectable similarities to known structures [22]. Poing combines multiple templates of known structures to produce the final model of the query sequence. The model is judged to be accurate when over 90% of the submitted residues are modeled at greater than 90% confidence [20].

Production of polyclonal antibodies to MpAFP RII and RIV

Two recombinant proteins, corresponding to RII (beginning at residues TTGS and ending at GNTVD) and RIV (beginning at residues NVSQ and ending at MTVV) from MpAFP (Genbank ABL74378.1) were produced in E. coli with N-terminal His-tags. Once the His-tags were removed via thrombin cleavage, aliquots (750 μg) were emulsified using TiterMax® (Cedarlane, Burlington, Canada) and used as separate antigens for the production of polyclonal antibodies. Single doses were injected into rabbits, and sera were collected approximately 6 weeks later.

Immunodetection and fluorescence microscopy imaging of MpAFP

An aliquot (0.5 mL) of an M. primoryensis culture in its stationary growth phase (OD600 = 1.3) was centrifuged at 2,000×g for 10 min. The cell pellet was resuspended in 1 mL of 0.85% (w/v) NaCl and an aliquot (10 μl) was pipetted onto a round coverslip. The cells were air dried for 30 min then fixed in 1% (v/v) paraformaldehyde for 20 min. After three 10-min washes in 0.85% (w/v) NaCl, the coverslips were incubated with a 1:200 dilution of anti-sera against either MpAFP_RII or RIV in the same solution at room temperature for 1 h. After three washes as above, the coverslips were incubated in the dark with a 1:200 dilution of goat anti-rabbit Alexa Fluor 350 secondary antibodies (Invitrogen) for 1 h at room temperature. To test the specificity of the secondary antibody, a control experiment was also carried out in which the fixed cells were incubated with fluorescent secondary antibodies in absence of the primary antibodies. After three more washes, coverslips were incubated in the dark with 0.05 mM SYTO 9 (Invitrogen) for 30 min to stain DNA. After three final
Nevertheless, the two genomic lambda clones revealed the entire identical (except the final one which has a single base difference), it these repeats, but as the DNA sequences of all repeats were 100% the insert. A large number of shotgun clones (45 in total) contained repeats (Fig. 3), as well as a total of twenty flanking genes (11 that this region of RTX repeats folds as a Ca\textsuperscript{2+}-u represents a large hydrophobic residue). We have determined xGTGNDuxGxuxGxux (where x can be any amino acid and conserved 19–21 residue repeat with the consensus sequence aa repeats mentioned above. Region IV is 322 amino acids (aa) which (II and IV) are highly repetitive. Region II contains the 104-\textsuperscript{aa} repeats in tandem. This makes \textit{MpAFP} a massive protein (ca. 1.5 MDa), with RII accounting for roughly 90% of its size (\textasciitilde 12,480 aa) compared to a total of 1435 aa for the other four regions combined.

Bioinformatics analysis of \textit{MpAFP} reveals homologous proteins with similar domain architectures

BLASTp searches performed using \textit{MpAFP} identified matches to many outer membrane adhesion proteins in Gram-negative bacteria (see Text S1 for details). Moreover, in conserved domain analyses, RII was identified as a poorly characterized repeat found in bacteria (expect value \textasciitilde 10\textsuperscript{-23}, pfam13753), which is similar to the bacterial Immunoglobulin (Ig)-like fold found in a variety of bacterial surface proteins (expect value \textasciitilde 10\textsuperscript{-3}, Pfam PF13754). The second match was between RV and Pfam PF08348 (expect value \textasciitilde 10\textsuperscript{-3}), a serralysin C-terminal domain thought to be important for secretion through the bacterial cell wall via the type I secretion pathway (TISS) [23].

The homologous model of a single 104-aa repeat of \textit{MpAFP} folds as an immunoglobulin-like beta-sandwich, confirming its identification as an adhesion protein

Structures of proteins with high sequence similarity to \textit{MpAFP} have not been described, but the conserved domain search suggested that the Protein Homology/analogy Recognition Engine (Phyre 2) might be able to generate homology models. A single 104-aa repeat segment of \textit{MpAFP}_RII was submitted to the Phyre2 server. It was modeled by the suite of programs as an S-type immunoglobulin (Ig)-like beta-sandwich (Table 1) with seven alphabetically listed beta-strands arranged in a Greek-key topology [24]. The N-terminal A and B strands hydrogen bond with strand E to form sheet I that packs against sheet II consisting of the other four strands (G, F, C and D) (Fig. 6). Although the six template structures for the final model show only ca. 20% sequence identity to the 104-aa RII, 92% of the protein’s residues were modeled at greater than 90% confidence.

The template structures used to model this domain consist of the two divergent repeats from the fibronectin (II)-like module from \textit{Clostridium thermocellum} (PDB accession 3PE9 and 3PDG) and a sulfite oxidase from chicken (PDB: 2A9D). In addition, several other structures also demonstrated high quality structural matches (confidence greater than 80%) to the single repeat of 104-aa RII, including regions of a collagen adhesin from \textit{Staphylococcus aureus} [25]. All of these matches were to regions of these proteins that adopt the Ig-like beta-sandwich fold. Identical results were obtained using the I-TASSER server in place of Phyre2.
The Ig-like fold also extends into RI and RIII. The Ig-like fold also extends into RI and RIII. The Ig-like fold also extends into RI and RIII. The Ig-like fold also extends into RI and RIII. The Ig-like fold also extends into RI and RIII. The Ig-like fold also extends into RI and RIII. The Ig-like fold also extends into RI and RIII. The Ig-like fold also extends into RI and RIII. The fold of RV is predicted to include a Ca2+-dependent beta-roll. The fold of RV is predicted to include a Ca2+-dependent beta-roll. The fold of RV is predicted to include a Ca2+-dependent beta-roll. The fold of RV is predicted to include a Ca2+-dependent beta-roll. The fold of RV is predicted to include a Ca2+-dependent beta-roll. The fold of RV is predicted to include a Ca2+-dependent beta-roll. The fold of RV is predicted to include a Ca2+-dependent beta-roll. The fold of RV is predicted to include a Ca2+-dependent beta-roll. Phyre2 predicted that the 249-aa MpAFP_RV adopts a Ca2+-bound beta-roll-containing structure in which 90% of the residues were modeled at over 90% confidence, despite only 13% sequence identity. The final model was constructed using extracellular lipases from *Pseudomonas sp. mis38* (PDB: 2ZJ6) and *Serratia marcescens* (PDB: 2QUQ). Both of the lipases belong to the I.3 family and consist of an N-terminal catalytic domain that is rich in alpha-helices and a C-terminal beta-roll containing RTX repeats with Ca2+-ions coordinated in the turns. *MpAFP_RV* was modeled based on the C-terminal domain. Several other modules also demonstrate excellent structural alignment (>95% confidence) to *MpAFP_RV*. These include the C-terminal domain of *Serratia*-like metalloproteases (PDB: 1G9K and 1K7G), a region from the secreted protease C (PDB: 1K7Q) and even the crystal structure of *MpAFP_RIV*. Again, these are all RTX-like, Ca2+-dependent beta-rolls, although models produced by Phyre2 do not include metal ions.

**MpAFP is localized to the cell surface of *M. primoryensis***

Evidence that *MpAFP* is localized to the exterior of the cell surface includes: 1) the presence of non-cleavable secretion signals near the C terminus (RTX repeats) indicates that *MpAFP* is secreted via TISS. This secretion system will allow *MpAFP* to be transported directly through the bacterial membranes without forming periplasmic intermediates. 2) No antifreeze activity is released into the cell culture medium of *M. primoryensis* [8]. This suggests that although *MpAFP* is secreted via TISS, it remains bound to the cells. 3) Bioinformatics analyses outlined above show that *MpAFP* has the hallmarks of an outer membrane adhesion protein. 4) Circular dichroism analyses demonstrate that region IV of *MpAFP* will only take on its beta-rich structure in the presence of the millimolar Ca2+ levels found in the bacterium’s natural environment [8]. The other regions of *MpAFP* also show dependence on millimolar Ca2+ for folding, and are predominantly random coil in the presence of excess EDTA (data not shown). The low cytosolic Ca2+ concentration (high nanomolar range) is insufficient to fold *MpAFP*, thus eliminating the possibility of it being functional inside the cell.

Here we have also used immunodetection to confirm the localization of *MpAFP* on the cell surface. We used antisera specific for RII or RIV, followed by a secondary antibody conjugated to a blue fluorophore, to screen for the presence of *MpAFP* on intact whole cells affixed to coverslips. The cells were counterstained with the cell-permeable nucleic acid dye SYTO 9 to differentiate the bacteria from debris. The green fluorescence of the nucleic acid stain clearly showed the bacteria as rod-shaped cells 2–3 μm in length (Fig. 7B and 7E). Blue fluorescence, indicating the presence of *MpAFP*, was co-localized to these cells when antiserum to RII (Fig. 7A) or RIV (Fig. 7D) was used. This was confirmed by the cyan coloration of all of the cells in the image overlays (Fig. 7C and 7F). Controls, from which the antiserum to RII or RIV was omitted, failed to show any blue fluorescence with just the secondary antibody present (not shown).
In another control reaction, *E. coli* fixed to glass slides stained green with SYTO 9 but showed negligible blue fluorescence from the anti-RII and anti-RIV antibodies used in conjunction with the labeled anti-rabbit second antibody (Fig. S3). Taken together, these results demonstrate that *Mp* AFP is uniformly distributed over the *Marinomonas primoryensis* cell surface and that both the large repetitive RII and the antifreeze domain are exposed to the extracellular environment, consistent with what would be expected for an adhesion protein.

**Discussion**

Here we report the purification of wild-type *Mp* AFP, its full amino acid sequence, and show that its tremendous size (ca. 1.5 MDa) is the result of ca. 120 tandem copies of an identical 104-aa repeat that is predicted to form an Ig-like beta-sandwich domain. The region is flanked by one or two non-identical repeats (65% identity) on each side. All of the other repeats are 100% identical at the DNA level, as shown by sequencing of numerous genomic clones. This suggests that their expansion in *M. primoryensis*, likely by duplication followed by multiple rounds of unequal recombination [27], is a relatively recent event. Each of the 120 Ig-like domains is expected to fold as an independent unit, forming a chain as observed for a combinatorial model of mouse cadherin based on electron tomography [28] and X-ray crystallography of five repeats [29]. As the monomers of cadherin and those predicted for *Mp* AFP are spaced approximately 50 Å apart, *Mp* AFP could be over 0.6 mm long, or 20% of the length of the cell. This implies that extreme length is a necessary property of this protein.
The only portions of MpAFP that showed high sequence similarity to proteins in the PDB database were the RTX repeats of RIV and RV. Thus, it was a helpful development to find such convincing structural homologs through the Phyre2 analyses. The different immunoglobulin (Ig)-like domains found in RI through RIII are all similar to the arrays found in well-characterized adhesin proteins like cadherin, fibronectin, and bacterial pili [26], [30], [31], strongly supporting MpAFP’s role in adhesion. MpAFP is particularly similar to some large adhesion proteins (LAPs) from the RTX family such as *Pseudomonas putida* LapA and LapF [32], [33], [34]. Not only do the individual regions of MpAFP show

![Figure 4. Estimation of 312-bp repeat copy number by Southern blotting.](A) Schematic diagram of the MpAFP coding region illustrating the relative positions of restriction enzyme cut sites. Cross-hatched lines indicate the break between the two segments of MpAFP coding region described in Fig. 3. Four restriction enzymes, including *Mse*, *Ase*, *Pst* and *Alu* cut only outside the 312-bp repeats. Since the cut site of *Mse* (*T/TAA*) is present within that of *Ase* (*AT/TAAT*), *Mse* is indicated in parentheses beside *Ase*. *Msp* is the only restriction enzyme in this set that cuts within the 312-bp repeats in RII, and it also cuts once in RI. (B) Southern blot of the digests using a 2.749-bp repeat from RII as the probe. Undigested DNA (lane 6) remained near the well, whereas the four restriction enzymes *Alu*, *Pst*, *Ase* and *Mse* (lane 1, 2, 4 and 5) that cut outside the repeats produced a fragment of approximately 37,500 bp in length. The *Msp* (lane 3) partial digest produced a ladder of bands at 312-bp intervals and those containing between 2 (624 bp) and 13 (4056 bp) repeats are marked on this blot. DNA length markers (kb) are indicated on the left of the blot.

doi:10.1371/journal.pone.0048805.g004

![Figure 5. Schematic representations of MpAFP and other large adhesion proteins.](A) Proteins are drawn to scale. Each Ig-like unit of the large repetitive region is shown as a green rectangle. The RTX-repeats in the C-terminal region are illustrated as brown blocks. The schematic diagrams are based on the following sequences from NCBI: *Pseudomonas putida* KT2440 LapA (NP_742337) and LapF (NP_742967), *M. sp. MWYL1* hemolysin-type calcium-binding protein (YP.001340713), *M. posidonica* outer membrane adhesin-like protein (YP.004881992) and the *M. mediterranea* homologues (YP.004313542+YP.004313541) that are separated by one nucleotide preceding an irregular GTG start codon. (B) Weblogo representations of 9-aa RTX repeats. Consensus plots are shown for: i) 12 RTX repeats in MpAFP_RIV; ii) 29 RTX repeats from homologs of MpAFP from three other *Marinomonas* species shown in Fig. 5A; iii) 10 RTX repeats in adhesion proteins LapA and LapF. As shown in the weblogo plots, the RTX repeats from the three homologs of MpAFP and the LAPs follow the consensus of the conventional nine-residue RTX repeats of GGxGxDxUx (Fig. 5Bii and 5Biii). The RTX-like repeats in MpAFP_RIV deviate from the conventional RTX repeats by introducing conserved ice-binding residues at positions 3 and 5 (Thr and Asx, Fig. 5B). Residues are colored black except for Gly (orange), Thr (red), and Ala (green).

doi:10.1371/journal.pone.0048805.g005
structural similarity to the domains of these adhesion molecules, but *Mp*AFP bears an overall domain architecture that matches the arrangement seen in these LAPs (Fig. 5A). This arrangement is characterized by an extremely large repetitive region of Ig-like folds near the N terminus (RII) followed by a non-repetitive region (RIII) and several RTX repeats near to the C terminus (RIV and RV). Intriguingly, no ice-binding characteristics (parallel arrays of Thr and Asx) were present in the RTX repeats of these other adhesion proteins, suggesting *Mp*AFP has divergently evolved specifically for ice binding (Fig. 5B). A comparison of the Weblogo plots for the RTX repeats suggests that the transition to an ice-binding role is relatively minor and involves conversion of the third residue to Thr. The fifth residue already has a high incidence of Asx. The only other difference of note is the lower incidence of Gly in position 1 in the AFP.

A second commonality among these LAPs is that they are all secreted via the TISS. The RTX-like repeats of RIV participate in ice-binding, and it is not clear if they can also serve in a secretion role. However the RTX repeats of RV are likely to act as accessory secretion signals, similar to the C-terminal RTX-repeats in adenylate cyclin toxin [35]. Many proteins secreted by the TISS are encoded within operons containing three transport proteins including an ATP-binding-cassette transporter, a membrane fusion protein and an outer membrane protein [23], but these

### Table 1. Statistics of homology modeling studies on *Mp*AFP domains.

| Regions | Residues modelled at 90% confidence (%) | Templates | Sequence identity to the templates (%) | Organisms |
|---------|-----------------------------------------|-----------|----------------------------------------|-----------|
| RI (394 aa) | 36 | Fibronectin (III)-like | 15 | *Clostridium thermocellum* |
| RII (104 aa) | 92 | Fibronectin (III)-like | 20 | *Clostridium thermocellum* |
| RIII (788 aa) | 16 | Cadherin-1 | 9 | Mouse |
| RV (249 aa) | 98 | I.3 lipase | 12 | *Pseudomonas sp. Mis38* |

![Figure 6. Homology models produced for domains of *Mp*AFP using Phyre2.](image)

The linear map of the regions of *Mp*AFP is colored as in Fig. 3. Hatched lines indicate the break between the two continguously sequenced segments of *Mp*AFP. Uncolored regions could only be modeled *ab initio* by the Phyre2 server and did not produce reliable results. Colored regions were modeled at greater than 90% confidence and their structures in ribbon representation are shown above and below the map of *Mp*AFP. The X-ray crystal structure of RIV is shown [9]. Residues for each model/structure are numbered according to the sequence in Fig. 3. The protein folds are shown in the colors of the rainbow from the N terminus (blue) to the C terminus (red).

doi:10.1371/journal.pone.0048805.g006

![Figure 7. Localization of *Mp*AFP by immunofluorescence.](image)

Immobilized *M. primoryensis* cells were fixed in 1% paraformaldehyde before being incubated with anti-RII (A, B and C) or anti-RIV (D, E and F) polyclonal antibodies, followed by Alexa Fluor 350 - conjugated goat anti-rabbit secondary antibody (blue in A and D) and SYTO 9 (green in B and E). Images (C) and (F) are composite images of (A) and (B), and (D) and (E), respectively. A 10-micron scale marker is shown by the white horizontal line in panels C and F.

doi:10.1371/journal.pone.0048805.g007
ligand. This resembles the role of surface; whereas the adhesive tip of a pilus is used to bind to its pilus from Gram-negative bacteria, which is a surface adhesin that can remain near the water surface where there is more oxygen and that ice growth will shear off the cells and release them back into step growth. Given the even larger size of the bacteria it is likely appear to be sheared off the surface as ice fronts move laterally by proteins are not included into a slowly growing mass of ice but We have observed that phage displaying AFP on their coat may prevent this and allow the cell to retain a shell of water. Thus, we speculate that the cell-ice interaction could potentially be a way of shielding the bacterium from the harmful effect of ice, and/or a chemotactic (envirotactic) type of response in which the bacteria keep themselves in the upper reaches of lakes for better access to oxygen. This represents a novel function for an antifreeze protein – that of binding an organism to ice. It would give AFPs a third role, distinct from thermal hysteresis (freeze avoidance) and ice recrystallization inhibition (freeze tolerance).

Supporting Information

Figure S1 Representative MS/MS spectrum of the IDAGTGNDIEYIK tryptic peptide. m/z values are shown for the abundant fragments in the mass spectrum above the corresponding peaks. The y series fragments have a charge of +1 and all extend to include the C-terminal lysine displayed on the left-hand side. The sequence of the fragment is displayed at the top of the spectrum. (TIF)

Figure S2 Amino acid alignment of MpAFP_RII with RII-like repeats. The alignment includes two tandem sequences from the C terminus of RII (RI-1: 184*-287* and RI-2: 288*-394*) and one from the N-terminal sequence of RII (RII-1, 209-310). These three sequences are aligned against the 104-aa repeat in RII. The residues shaded black corresponds to residues identical to those in MpAFP_RII, whereas the ones shaded grey mark conservative substitutions. (TIF)

Figure S3 Control experiment testing the reactivity of E. coli to anti-RII and anti-RIV polyclonal antibodies. Immobilized E. coli cells were fixed in 1% paraformaldehyde before being incubated with anti-RII (A, B and C) or anti-RIV (D, E and F) polyclonal antibodies, followed by Alexa Fluor 350-conjugated goat anti-rabbit secondary antibody (blue in A and D) and SYTO 9 (green in B and E). Images (C) and (F) are composite images of (A) and (B), and (D) and (E), respectively. A 10-micron scale marker is shown by the white horizontal line in panels C and F. (TIF)

Table S1 Comparison of monoisotopic [M+H]+ masses and sequences of tryptic peptides from native M. primoryensis AFP to those predicted from the gene sequence. (DOCX)

Table S2 Proteins immediately upstream and downstream of MpAFP. (DOCX)

Text S1 Bioinformatics analyses of MpAFP via BLASTp. When BLAST searches were performed using MpAFP, three sequences from other Marinomonas species were detected that are flanked on either side by the same two genes that lie adjacent to the MpAFP gene. The AFP homologues share a similar domain structure in that they possess internal MpAFP_RII-like repeats of ~100 aas, albeit far fewer than in MpAFP and they also contain C-terminal regions of RTX repeats that are similar to those of MpAFP_RIV and RV (Fig. 5A). However, the RII-like repeats, as well as the bulk of the protein, are variably conserved between species. For example, the only regions where MpAFP and the M. psidonica homolog contain over 50% identity are within the first and last ~150 aa. These homologs also contain variable numbers of RTX repeats near their C-termini that are similar to those from MpAFP_RIV. However, they lack the ice-binding Thr residues in position 3 of the repeat (Fig. 5B i), which suggests these proteins do
not bind to ice. Similar domain structures were also detected in other large RTX proteins, including two adshes from *Pseudomonas putida* (Fig. 5A) that contain many RII-like repeats along with RJV-like RTX repeats (PP00353) that again lack the ice-binding residues (Fig. 3B ii and iii).

**Acknowledgments**

We are grateful to microbiologists Drs. Nancy Martin and Keith Poole for advice, to Dr. Scott Davey for access to the PFGE apparatus, and to Dr.

**References**

1. DeVries AL, Wohlschld Dr (1969) Freezing Resistance in Some Antarctic Fishes. Science 163: 1073-8.
2. Duman JG, DeVries AL (1974) Freezing Resistance in Winter Flounder *Pseudopleuronectes americanus*. Nature 247: 237-238.
3. Raymond JA, DeVries AL (1977) Adsortion Inhibition as a Mechanism of Freezing Resistance in Polar Fishes. PNAS 74: 2389-2393.
4. Sidebottom C, Buckley S, Pulley P, Twigg S, Jarman C, et al. (2000) Phenytoinchemistry - Heat-stable antifreeze protein from grass. Nature 406: 256-256.
5. Worrall D, Elias L, Ashford D, Smallwood M, Sidebottom C, et al. (1998) A carrot leucine-rich-repeat protein that inhibits ice recrystallization. Science 282: 113-117.
6. Knight CA, DeVries AL, Oulson LD (1984) Fish Antifreeze Protein and the Freezing and Recrystallization of Ice. Nature 308: 295-296.
7. Garnham CP, Gilbert JA, Hartman CP, Campbell RL, Laybourn-Parry J, et al. (2000) A Ca2+-dependent bacterial antifreeze protein domain has a novel beta-helical ice-binding fold. Biochem J 341: 171-180.
8. Gilbert JA, Davies PL, Laybourn-Parry J (2005) A hyperactive, Ca2+-dependent antifreeze protein in an Antarctic bacterium. FEBS Microbiol Lett 245: 67-72.
9. Garnham CP, Campbell RL, Davies PL (2011) Anchored chloral waters bind antifreeze proteins to ice. PNAS 108: 7363-7367.
10. Gallagher KR, Sharp KA (2003) Analysis of thermal hysteresis protein hydration using the random network model. Biophys Chem 103: 195-209.
11. Nutt DR, Smith JC (2008) Function of the hydration layer around an antifreeze protein revealed by atomistic molecular dynamics simulations. J Am Chem Soc 130: 13066-13073.
12. Yang C, Sharp KA (2004) The mechanism of the type III antifreeze protein action: a computational study. Biophys Chem 109: 137-148.
13. Yang C, Sharp KA (2005) Phodyntropic tendency of polar group hydration as a major force in type I antifreeze protein recognition. Proteins 59: 266-274.
14. Fletcher GL, Hee CL, Davies PL (2001) Antifreeze proteins of tilasid fishes. Annu Rev Physiol 63: 359-390.
15. DeLuca CI, Comley R, Davies PL (1998) Antifreeze proteins bind independently to ice. Biophys J 74: 1502-1508.
16. Kupper MJ, Lankin C, Gauthier SY, Walker VK, Davies PL (2003) Purification of antifreeze proteins by adsorption to ice. Biophys Res Comm 300: 645-648.
17. Campbell KP, MacIemman DH, Jorgensen AO (1983) Staining of the Ca2+ binding proteins, calcequstin, calmodulin, troponin C, and S-100, with the cationic carbocyanine dye ''Stains-all''. J Biol Chem 258: 11267-11273.
18. Orlin S (1990) Molecular Cloning - A Laboratory Manual, 2nd Edition - Sambrook J, Fritsch EF, Maniatis T. Nature 343: 604-605.
19. Hisson JE (1980) Short Protocols in Molecular Biology - Ausubel FM, Brent R, Kingston RE, Moore DM, Seidman JG, Smith JA, Struhl K, Wang, Zbonerion, F, Bonitz SG, Am J Hum Biol 2: 172-173.
20. Kelley LA, Sternberg MJE (2000) Protein structure prediction on the web: a case study using the Phyre server. Nature Protocols 4: 363-371.
21. Soding J (2003) Protein homology detection by HHM HHM comparison. Bioinformatics 21: 951-960.
22. Jefferys BR, Kelley LA, Sternberg MJE (2010) Protein Folding Requires Crowd Control in a Simulated Cell. J Mol Biol 397: 1329-1338.
23. Delpeule P (2004) Type I secretion in Gram-negative bacteria. Biochim Biophys Acta 1694: 149-161.
24. Bork P, Holm L, Sander C (1994) The Immunoglobulin Fold - Structural Classification, Sequence Patterns and Common Core. J Mol Biol 242: 309-320.
25.Devnanayagam CCS, Rich RL, Carson M, Owens RT, Daltulvari S, et al. (2000) Novel fold and assembly of the repetitive B region of the Staphylococcus aureus collagen-binding surface protein. Structure 8: 67-78.
26. Shapiro L, Weis WI (2009) Structure and Biochemistry of Cadherins and Catenins. Cold Spring Harb Perspect Biol 1.
27. Baymek M, Lovert ST (2001) Instability of repetitive DNA sequences: The role of replication in multiple mechanisms. PNAS 98: 8319-8323.
28. He WZ, Woin P, Stokes DL (2005) Untangling desmosomal knots with electron tomography. Science 302: 1089-1103.
29. Boggon TJ, Murray J, Chappuis-Flament S, Wong E, Gumbiner BM, et al. (2002) C-cadherin extracellular structure and implications for cell adhesion mechanisms. Science 296: 1306-1313.
30. Pankov R, Yamada KM (2002) Fibronectin at a glance. J Cell Sci 115: 3861-3863.
31. Proby T, Baker EN (2009) PI3K in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. Cell Mol Life Sci 66: 613-635.
32. Espinoso-Urgel M, Salido A, Ramos JL (2000) Genetic analysis of functions involved in adhesion of Pseudomonas putida to seeds. J Bacteriol 182: 2363-2369.
33. Martinez-Gil M, Yousef-Coronado F, Espinoso-Urgel M (2010) LapF, the second largest Pseudomonas putida protein, contributes to plant root colonization and determines biofilm architecture. Mol Microbiol 77: 549-561.
34. Satchell KJF (2011) Structure and Function of MARTX Toxins and Other Large Repeat-forming RTX Proteins. Annu Rev Microbiol 65: 59-80.
35. Sebo P, Ladant D (1993) Repeat Sequences in the Bordetella-Pertussis Plul (PF00353) that again lack the ice-binding residues (Fig. 3B ii and iii).

**Author Contributions**

Conceived and designed the experiments: PLD SG CPG LAG. Performed the experiments: SG CPG JCW LAG. Analyzed the data: SG CPG LAG. Contributed reagents/materials/analysis tools: PLD. Wrote the paper: PLD SG CPG LAG.