

A Ca²⁺-dependent bacterial antifreeze protein domain has a novel β -helical ice-binding fold

Christopher P. GARNHAM*, Jack A. GILBERT*¹, Christopher P. HARTMAN*, Robert L. CAMPBELL*, Johanna LAYBOURN-PARRY^{†2} and Peter L. DAVIES*³

*Protein Function Discovery Group and the Department of Biochemistry, Queen's University, Kingston, ON, Canada K7L 3N6, and [†]Institute for the Environment, Physical Sciences and Applied Mathematics, University of Keele, Keele, Staffordshire, ST5 5BG, U.K.

AFPs (antifreeze proteins) are produced by many organisms that inhabit ice-laden environments. They facilitate survival at sub-zero temperatures by binding to, and inhibiting, the growth of ice crystals in solution. The Antarctic bacterium *Marinomonas primoryensis* produces an exceptionally large (>1 MDa) hyperactive Ca²⁺-dependent AFP. We have cloned, expressed and characterized a 322-amino-acid region of the protein where the antifreeze activity is localized that shows similarity to the RTX (repeats-in-toxin) family of proteins. The recombinant protein requires Ca²⁺ for structure and activity, and it is capable of depressing the freezing point of a solution in excess of 2°C at a concentration of 0.5 mg/ml, therefore classifying it as a hyperactive AFP. We have developed a homology-guided model of the antifreeze region based partly on the Ca²⁺-bound β -roll from alkaline protease. The model has identified both a novel

β -helical fold and an ice-binding site. The interior of the β -helix contains a single row of bound Ca²⁺ ions down one side of the structure and a hydrophobic core down the opposite side. The ice-binding surface consists of parallel repetitive arrays of threonine and aspartic acid/asparagine residues located down the Ca²⁺-bound side of the structure. The model was tested and validated by site-directed mutagenesis. It explains the Ca²⁺-dependency of the region, as well its hyperactive antifreeze activity. This is the first bacterial AFP to be structurally characterized and is one of only five hyperactive AFPs identified to date.

Key words: Antarctic bacterium, antifreeze protein, calcium binding, β -helix, *Marinomonas primoryensis*, repeats-in-toxin proteins (RTX proteins).

INTRODUCTION

Since their discovery over 30 years ago in Antarctic fish [1], antifreeze proteins (AFPs) have been found in a wide array of organisms that inhabit freezing environments. AFPs enhance survival at sub-zero temperatures by inhibiting the growth of potentially harmful ice crystals within the organism. If left uncontrolled, ice growth can lead to cellular dehydration, tissue damage and death of the organism [2]. By binding to the surface of ice crystals, AFPs are able to depress the freezing point of a solution below its melting point [3]. The difference between these temperatures is termed 'thermal hysteresis' (TH) and is used as a measure of antifreeze activity.

Fish and insects which avoid freezing typically have AFPs with high TH activity [4]. Plants and micro-organisms that are unable to avoid freezing are thought to use AFPs with weaker TH activity to inhibit recrystallization of ice [5,6]. It is remarkable, therefore, that the Antarctic bacterium *Marinomonas primoryensis* produces a Ca²⁺-dependent AFP with a strong TH activity [7]. This could potentially help it avoid freezing in the brackish ice-covered lakes from which it was isolated [8]. MpAFP (*M. primoryensis* AFP) is an extremely large protein with unusual properties that make it refractory to conventional purification techniques. Nevertheless, we were able to sequence a number of tryptic peptides from gel-purified MpAFP, some of which resemble sequences from the Ca²⁺-binding regions of RTX (repeats-in-toxin) proteins (J.A. Gilbert, C.P. Garnham, L.A. Graham, J. Laybourn-Parry and P.L. Davies, unpublished work).

A genomic lambda library of *M. primoryensis* was constructed and probed with a PCR-amplified DNA sequence developed from tryptic peptides in the RTX-like region. Preliminary analysis shows the gene encodes a huge (>1 MDa) protein that contains two repetitive sequences that divide the protein into five distinct regions (Figure 1A). Region II of the protein contains an unknown number of tandem repeats of a 104-aa (104-amino-acid) sequence. Each 104-aa repeat is perfectly conserved, even at the DNA level. This has made it difficult to determine the exact copy number of 104-aa repeats, and therefore difficult to estimate the size of MpAFP.

In the interim, we have focused our research on Region IV, the second repetitive region of MpAFP. Region IV consists of tandem 19-aa repeats, each beginning with the consensus sequence XGTGND, where X is usually alanine or glycine. This region of MpAFP is clearly homologous with the RTX proteins. RTX proteins are a diverse family of secreted virulence factors found in Gram-negative gammaproteobacteria [9]. They are so called because of the presence of a tandemly repeated nonapeptide motif located towards their C-terminus. The motif has a consensus sequence of GGXGXDXUX, where X represents any residue and U is a large hydrophobic residue. AP (alkaline protease), an RTX-like virulence factor secreted by the Gram-negative bacterium *Pseudomonas aeruginosa*, contains five copies of this nonapeptide motif. The X-ray-crystallographic structure of this protein revealed that these repeats fold into a parallel β -roll stabilized by internal Ca²⁺ ions sandwiched between the turns of the structure [10].

Abbreviations used: aa, amino acid; AFP, antifreeze protein; AP, alkaline protease; His₆, hexahistidine; IBF, ice-binding face; IBP, ice-binding protein; MpAFP, *Marinomonas primoryensis* AFP; Ni-NTA, Ni²⁺-nitrilotriacetate; PDB, Protein Data Bank; Q-TOF, quadrupole time-of-flight; RTX, repeats-in-toxin; TH, thermal hysteresis.

¹ Present address: Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth, PL1 3DH, U.K.

² Present address: Private Bag 51, University of Tasmania, Hobart, Tasmania 701, Australia.

³ To whom correspondence should be addressed (email peter.davies@queensu.ca).

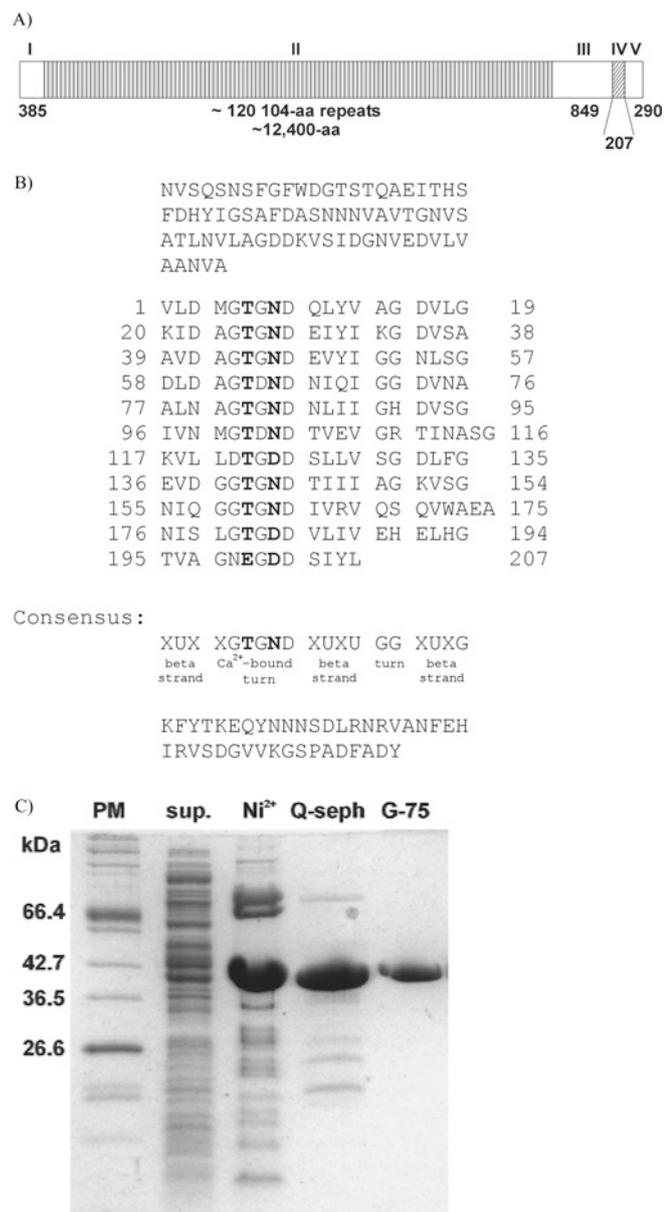


Figure 1 Sequence, expression and purification of *MpAFP* Region IV

(A) *MpAFP* contains two highly repetitive sequences that divide the protein into five distinct regions (I–V) with the number of residues in each region shown below. The approx. ~120–104-aa repeats of *MpAFP* are in light grey. The 207-aa Region IV of *MpAFP* is indicated by the diagonal hatching. **(B)** Alignment of the eleven 19-aa repeats of Region IV of *MpAFP*. Note that two of the repeats have 21 aa and the eleventh repeat is incomplete. The additional 74 N-terminal and 41 C-terminal residues of Region IV that were used in the expression construct are also shown. The sequence is spaced according to secondary structure with the positions of the XGTGND Ca^{2+} -binding turn and three β -strands indicated below the consensus sequence. Indicated in bold are the threonine and aspartic acid/asparagine residues that constitute the IBF of the protein. As in the text, X means any residue and U means a large hydrophobic residue. **(C)** SDS/10%-(w/v)-PAGE analysis of the purification of *MpAFP* Region IV. Lanes from left to right are identified as follows: PM, protein markers; sup, supernatant from the *E. coli* lysate; Ni^{2+} , pooled fractions from the Ni-NTA-agarose affinity-chromatographic step; Q-Sep, pooled fractions from the Q-Sepharose ion-exchange chromatography step; G-75, pooled fractions from size-exclusion chromatography on Superdex G-75.

The wild-type *MpAFP* is dependent on Ca^{2+} for TH activity, and this activity is resistant to proteolysis by trypsin only if Ca^{2+} is present [7]. This suggests a structural role for the cation, as opposed to its cofactor role in the Ca^{2+} -dependent type II fish

AFP from the Atlantic herring (*Clupea harengus*) [11,12]. On the basis of these data and the similarity of Region IV to the Ca^{2+} -binding β -roll from AP, we hypothesized that Region IV was the antifreeze-active region of *MpAFP* and that its structure would be similar to the β -roll of AP. Indeed, cloning and expression of Region IV in *Escherichia coli* revealed it to be a potent AFP, capable of depressing the freezing temperature of a solution in excess of 2°C at a concentration of 0.5 mg/ml. This level of TH is quantitatively the same as the wild-type *MpAFP* and indicated that Region IV could account for the antifreeze activity of *MpAFP*.

Attempts to solve the three-dimensional structure of *MpAFP* Region IV have been unsuccessful. The isolated domain has a propensity to aggregate and precipitate above 1 mg/ml, a concentration well below that needed for solution NMR or X-ray crystallography. However, the repetitive nature of Region IV and its homology with the known β -roll structure in AP make it an ideal candidate for molecular modelling.

In this paper we present a β -helical model of *MpAFP* Region IV based partly on the Ca^{2+} -bound β -roll from the RTX-like AP. The model contains both a novel ice-binding motif and β -helical fold that explains the Ca^{2+} -dependency and hyperactivity of the protein. The modelled structure revealed a putative ice-binding site along one edge of the β -helix that is flat and is comprised of parallel ranks of equally spaced threonine and asparagine residues. The threonine and asparagine residues originate from each XGTGND repeat that forms a Ca^{2+} -bound turn in the β -helix similar to the GGXGXD Ca^{2+} -bound turns in AP. The IBF (ice-binding face) of the model is similar to the IBFs of the hyperactive β -helical insect AFPs. They each contain a repetitive TXT motif located in a single flat β -sheet on one side of their respective structures [13,14]. It is hypothesized that this allows the insect AFPs to bind both primary prism and basal planes of an ice crystal by mimicking the distance of oxygen atoms in the ice lattice. Our model of Region IV predicts that it, too, should bind both basal and prism planes of an ice crystal lattice by the same mechanism. However, the location of the IBF in the Ca^{2+} -bound turn of the model, as well as the use of threonine and asparagine/aspartic acid as ice-binding residues, is novel compared with the insect AFPs and highlights the plasticity of the β -helix as an AFP structural fold. Extensive site-directed mutagenesis was performed in order to validate the putative IBF of the model and its overall fold. Mutation of specific threonine and asparagine residues on the putative IBF of the protein to tyrosine resulted in large decreases in TH activity, whereas TH activity was left unchanged when similar substitutions were made off the IBF.

MATERIALS AND METHODS

Cloning and expression of *MpAFP* Region IV

Two primers were designed to clone and express a 322-aa portion of *MpAFP* as a C-terminally His₆ (hexahistidine)-tagged protein. The 322-aa segment corresponds to residues 997–1318 inclusively of the sequence deposited in the NCBI protein database under accession number ABL74378. This portion contained a short segment of Region III (74 aa), followed by the entire Region IV (207 aa, corresponding to residues 1071–1277 of ABL74378) and a short segment of Region V (41 aa). The forward primer was 5'-ACGTCATATGAATGTGTCGCAATCAAATTCG-3', which contained an NdeI site, and the reverse primer was 5'-TGCACTCGAGATAGTCAGCAAAGTCCGAGG-3', which contained an XhoI site. Following digestion with NdeI and XhoI, the PCR product obtained from the genomic DNA template was cloned into the corresponding sites of the expression vector pET-24a. Positive clones were inoculated into 1 litre of Luria-Bertani

medium with kanamycin (100 μ g/ml) and grown at 37 °C and at 200 rev./min agitation until the attenuation (D_{600}) reached 0.5. The cells were allowed to grow for another 0.5 h at 23 °C (or until the D_{600} reached 1.0), then isopropyl β -D-thiogalactoside was added to a final concentration of 1 mM to induce expression overnight at 23 °C. Following incubation, the cells were recovered by centrifugation (2500 g for 30 min at 4 °C) and resuspended in 25 ml of buffer A (50 mM Tris/HCl, pH 7.5, 150 mM NaCl, and 2 mM CaCl₂). The resuspension was sonicated [Fisher Scientific (Waltham, MA, U.S.A.) Sonic Dismembrator, model 5000; five bursts of 45 s each at 50 % amplitude] to break open the cells and centrifuged at 21 000 g and 4 °C for 1 h in a JA20 rotor to remove cellular debris.

Purification of His₆-tagged Region IV of MpAFP

The crude cellular lysate was mixed with 10 ml of Ni-NTA (Ni²⁺-nitrilotriacetate) resin (Qiagen, Mississauga, ON, Canada) and stirred at 4 °C for 30 min, loaded into a column, washed with Buffer N [0.5 M NaCl, 50 mM Tris/HCl, pH 7.5, 2 % (v/v) glycerol, 2 mM CaCl₂ and 5 mM imidazole] and eluted with buffer N containing 250 mM imidazole. Fractions displaying TH were pooled and diluted 10-fold in buffer A to decrease the NaCl concentration. This material was applied to a HiLoad 16/10 Q-Sepharose High Performance column (Amersham) equilibrated with buffer A and eluted with a 0–1 M NaCl gradient in buffer A over 10 column vol.. Samples displaying TH activity were again pooled and loaded on to a HiLoad 16/60 Superdex 75 prep grade size-exclusion column (Amersham) equilibrated in buffer A. Fractions were analysed for purity by SDS/PAGE.

Mutagenesis

All mutagenesis was performed by using the QuikChange® II XL Site Directed Mutagenesis Kit as described by the manufacturer (Stratagene, La Jolla, CA, U.S.A.) and checked by DNA sequencing (Cortec, Kingston, ON, Canada). Each mutant of Region IV was purified in the same manner as wild-type recombinant Region IV.

CD

CD spectra were measured on an Olis Rapid Scanning Monochromator with a DSM (digital subtractive method) CD module (Olis, Bogart, GA, U.S.A.) using a 121-QS 1-mm-pathlength quartz cuvette (Hellma Optik GmbH, Jena, Germany). The temperature of the cuvette during scans was maintained at 4 °C using a Peltier platform. Initial CD experiments on MpAFP Region IV in the presence of 5 mM Tris/HCl, pH 7.5, and 2 mM CaCl₂ or 2 mM EDTA were carried out at a concentration of about 0.4 mg/ml. Since the exact concentration of the samples was unknown, the results are plotted solely as ellipticity versus wavelength. All subsequent CD experiments on Region IV and its mutants were carried out at a protein concentration of 1 mg/ml in buffer containing 5 mM Tris/HCl, pH 7.5, and 2 mM CaCl₂. Protein concentration was determined by measuring the A_{280} and using the protein's molar absorption coefficient (0.666 mol⁻¹ · litre⁻¹ · cm⁻¹).

Modelling and molecular dynamics

Modelling was performed using the program Modeller8v2 [15]. Input alignments were manually created as outlined in the Results section. PyMol v0.99 [16] was used to assemble and visualize modelled segments. Molecular-dynamics simulations were performed using Gromacs v3.3 [17]. The model was solvated in a cube containing 9105 water molecules, with all sides of the cube 6.5 nm in length. Because the overall charge of the

model was –10, ten water molecules were replaced by ten Na⁺ ions to neutralize the system. The solvated model was subjected to energy minimization by steepest-descents position-restrained molecular dynamics to relax the solvent, followed by full molecular dynamics. The GROMOS96 43a1 force field was chosen. Berendsen temperature coupling was applied. Simulations were performed at 277 K (4 °C) for 2 ns.

TH measurements

TH activity was measured using a Clifton Nanolitre Osmometer (Clifton Technical Physics, Hartford, NY, U.S.A.) as previously described [18]. Digital images of ice crystals were recorded using a Nikon COOLPIX 4500 digital camera mounted on a Leitz Dialux 22 microscope with a Leitz Wetzlar 160/- EF 32/0.40 objective.

Ice-crystal morphology produced by Region IV was examined using a 0.08 mg/ml solution of MpAFP Region IV in 25 mM Tris/HCl (pH 7.5)/2 mM CaCl₂. TH measurements of Region IV and its mutants were conducted in the same buffer at concentrations from 0.9 down to 0.025 mg/ml. The TH activity of Region IV as a function of Ca²⁺ concentration was determined by dialysing a 0.1 mg/ml solution of Region IV against 25 mM Tris/HCl (pH 7.5)/100 μ M EDTA. Increasing amounts of Ca²⁺ were titrated into the solution and measurements were taken at each concentration.

Q-TOF (quadrupole time-of-flight) MS

The mass of MpAFP Region IV in EDTA and Ca²⁺ was determined using a Waters Q-TOF Global MS instrument. A 100 μ M solution of Region IV was dialysed against 10 mM Tris/HCl, pH 7.5, and either 2 mM CaCl₂ or 2 mM EDTA. Both the EDTA and CaCl₂ samples were diluted to 10 pmol/ μ l in 10 mM ammonium formate and were subjected to Q-TOF MS in the static-spray mode to determine their respective masses. The Waters MassLynx 4.0 software package was used for data analysis and acquisition.

RESULTS

Cloning, expression and purification of MpAFP Region IV

Shotgun sequencing of selected clones from a genomic lambda library of *M. primoryensis* identified an extremely large protein that contained five distinct regions (Figure 1A). The 207-aa Region IV was tentatively identified as the antifreeze-active portion of the protein because of its perceived dependence on Ca²⁺ for structural integrity based on homology with the Ca²⁺-dependent RTX family of proteins and, in particular, the Ca²⁺-bound β -roll from the RTX-like AP. The sequence of Region IV is shown in Figure 1(B). Region IV was cloned as a C-terminal His₆-tagged protein with an extra 74 N-terminal residues and 41 C-terminal residues to help ensure that enough sequence was present at either end of Region IV to allow for proper folding of that domain. Region IV was purified to homogeneity in three sequential chromatographic steps: Ni-NTA-agarose, Q-Sepharose and Superdex G-75. An SDS/PAGE analysis of the purification shows that the nickel column chromatography produced a very substantial enrichment of a 40 kDa band from the lysate supernatant (Figure 1C). The next two chromatographic steps removed minor impurities such that the Superdex G-75 fraction contained a single intense protein band at 40 kDa.

The TH activity of MpAFP Region IV is hyperactive and Ca²⁺-dependent

The purified Region IV was tested for TH activity as a function of its concentration in buffer containing 2 mM CaCl₂

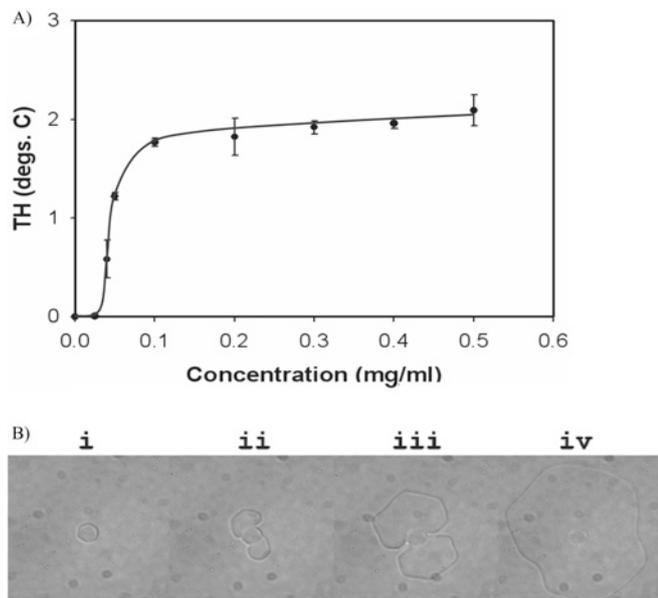


Figure 2 Antifreeze activity of *MpAFP* Region IV

(A) TH activity [$^{\circ}\text{C}$ ('degs.C')] was plotted as a function of protein concentration (mg/ml). Each reading was taken three times and averaged. Vertical bars represent \pm S.D. (B) Ice-crystal morphology and 'burst' (sudden and rapid ice growth) produced by Region IV. i, Picture taken pre-burst; ii–iv, pictures taken of the burst at 0.4 s intervals. The crystals burst at $\sim 1.8^{\circ}\text{C}$ below the colligative freezing point of the solution.

(Figure 2A). TH levels of just under 2°C were obtained at a concentration of 0.1 mg/ml ($2.9\ \mu\text{M}$). At this concentration (low μM levels), a typical fish AFP would only produce about 0.1°C of TH, whereas hyperactive AFPs produce TH levels at least 10 times greater [4]. The overall shape of the activity curve was sigmoidal. At concentrations below 0.025 mg/ml, TH activity was undetectable, but above 0.025 mg/ml, TH activity increased rapidly and approached a plateau at about 2°C at concentrations above 0.1 mg/ml. Region IV produced hexagonally shaped ice crystals that 'burst' out of the sides of the hexagon at temperatures below the non-equilibrium freezing point (Figure 2B). This same phenomenon, namely of hexagonally shaped ice crystals that burst along the *a*-axes, was first documented with the β -helical AFP from the tortricid moth *Choristoneura fumiferana* (spruce budworm) [19] and is characteristic of all hyperactive AFPs [4].

The TH activity of Region IV was also tested as a function of Ca^{2+} concentration (Figure 3). All TH activity was lost in the presence of EDTA. A rapid hyperbolic increase in activity was seen as the Ca^{2+} concentration increased, levelling out at about 10 mM Ca^{2+} . Even 100 μM Ca^{2+} produced appreciable TH activity ($>1^{\circ}\text{C}$).

MpAFP Region IV has Ca^{2+} -dependent β -sheet structure

The secondary structure of Region IV in the presence of either Ca^{2+} or EDTA was monitored via CD (Figure 4). In the presence of Ca^{2+} , the spectrum was similar to that observed with proteins containing primarily β -sheet, indicated by the maximum at 196 nm and the minimum at 218 nm. The spectrum in the presence of EDTA was that of a random coil, with a large minimum at 200 nm. This indicated that Region IV adopted a Ca^{2+} -dependent predominantly β -sheet fold that could be disrupted by EDTA.

On the basis of its hyperactivity, Ca^{2+} -dependence, hexagonal ice-crystal shape and CD spectrum, we hypothesized that Region IV of *MpAFP* folded as a β -helix similar to the hyperactive insect

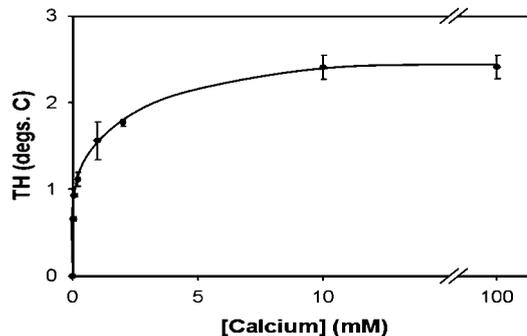


Figure 3 Ca^{2+} -dependency of *MpAFP* Region IV

TH activity [$^{\circ}\text{C}$ ('degs. C')] was plotted as a function of Ca^{2+} concentration. Each reading was taken three times and averaged. Vertical bars represent \pm S.D. A break in the graph between 10 mM and 100 mM CaCl_2 is indicated.

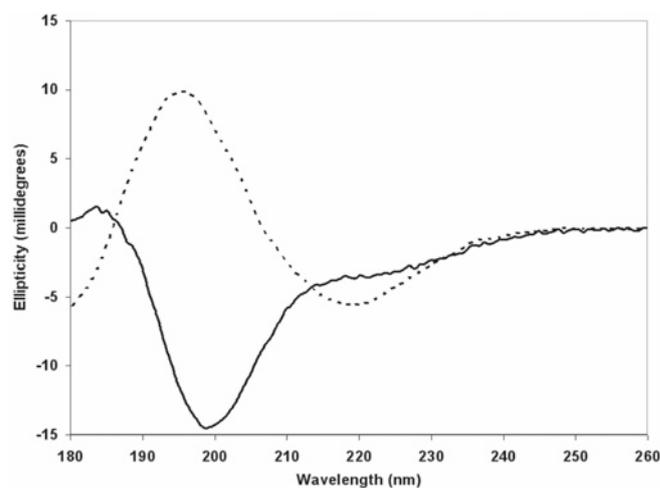


Figure 4 CD spectra of *MpAFP* Region IV in Ca^{2+} and EDTA

The far-UV CD spectra of Region IV in the presence of either 2 mM EDTA (continuous line) or 2 mM Ca^{2+} (broken line) were plotted as ellipticity versus wavelength.

AFP, but one that also requires Ca^{2+} for structural stability. The similarity of Region IV to the Ca^{2+} -bound β -roll from AP (Figure 5A), and in particular each XGTGND repeat of Region IV to the GGXGXD repeats of AP, suggested that the β -roll from AP could be used as a valid starting template for a homology-based model of Region IV.

Description of β -roll from AP and comparison to *MpAFP* Region IV

Each loop in the β -roll of AP consists of two identical GGXGDXUX nonapeptide motifs repeated in tandem (Figure 5B). This creates one loop in the structure 18 aa in length (Figure 5A). The GGXGXD sequence within each nonapeptide motif of AP forms a turn stabilized by internal Ca^{2+} ions sandwiched between loops of the β -roll. This Ca^{2+} -bound turn is followed by a three-residue XUX β -strand, the X residues of which point outwards and the U residues of which point inwards, contributing to the hydrophobic core of the β -roll. Each nonapeptide motif creates one half of an 18-aa loop, with an ensuing nonapeptide motif completing the second half. This creates a β -roll with Ca^{2+} bound down both sides of the structure and with a narrow hydrophobic core in between, consisting of the U residues from each XUX β -strand.

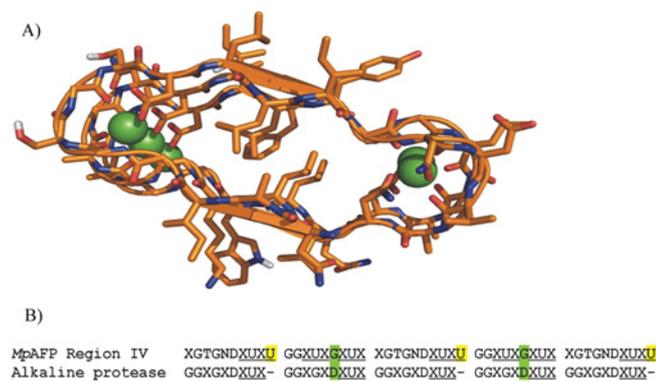


Figure 5 Ca^{2+} -bound β -roll from AP

(A) Cross-section of the Ca^{2+} -bound β -roll from AP (PDB code 1KAP) used as the starting template for the Region IV model. Key to colours: orange, carbon; red, oxygen; blue, nitrogen; green, Ca^{2+} . (B) Consensus aa alignment of Region IV to the β -roll from AP. Highlighted in yellow is the extra hydrophobic residue of Region IV following its XGTGND Ca^{2+} -bound turn. As in the text, X is any residue and U is a large hydrophobic residue. Highlighted in green is the aspartic acid-to-glycine substitution in Region IV. The asterisk (*) denotes the conserved hydrophobic U residue located in the second β -strand of each 19-aa loop of Region IV. Residues underlined (i.e.: XUX) represent the parts of the sequence that fold as a β -strand.

Each 19-aa repeat of MpAFP Region IV consists of a decapeptide motif with a consensus sequence of XGTGNDXUXU, followed by a nonapeptide motif the consensus sequence of which is GGXUXGXUX (Figure 1B). Each one of these motifs can be viewed as one half of a loop in a β -helical structure, similar to the β -roll of AP, yet differing in three key areas (Figure 5B). The initial decapeptide motif of each 19-aa repeat of Region IV contains a Ca^{2+} -binding XGTGND sequence, followed by a four-residue XUXU β -strand. This is very similar to the nonapeptide motif of AP (GGXGDXUX), except for the presence of an extra hydrophobic residue at the end of the β -strand in Region IV (i.e. XUXU). The ensuing GGXUXGXUX nonapeptide motif of each 19-aa repeat differs from the nonapeptide repeat of AP by the presence of a hydrophobic U residue at position 4 of the repeat (GGXUXGXUX) and by an aspartic acid-to-glycine substitution at position 6 of the motif (GGXUXGXUX). This creates a nine-residue sequence that contains two three-residue XUX β -strands (GGXUXGXUX), as opposed to the nonapeptide motif of AP, which contains a GGXGDX Ca^{2+} -bound turn followed by a single three-residue XUX β -strand.

Initial modelling of MpAFP Region IV

The program Modeller was used to model MpAFP Region IV. The β -roll segment from AP (residues 330–378) was excised on the computer and used as the template. It contains five GGXGDX Ca^{2+} -bound turns that create a $2\frac{1}{2}$ -loop β -roll (Figure 5A). This template posed two potential problems for the model, as it was much shorter than the 11-loop MpAFP Region IV and it contained Ca^{2+} ions bound down both sides of the structure. The aspartic acid-to-glycine substitution at position 6 of the nonapeptide motif in each 19-aa repeat (GGXUXGXUX) of Region IV as well as the presence of two XUX β -strands in that motif (GGXUXGXUX) as compared with AP (GGXGDXUX) indicated that Ca^{2+} was likely not to be bound down that side of the model. However, the sequence of each 19-aa repeat of MpAFP Region IV was initially threaded on to the β -roll of AP by aligning each XGTGND sequence with a GGXGDX Ca^{2+} -bound turn in AP (Figure 5B). We also introduced a gap in the alignment to accommodate the extra hydrophobic residue of each

loop of Region IV (XGTGNDXUXU). Tandem $2\frac{1}{2}$ -loop segments of Region IV were modelled in an iterative fashion with Ca^{2+} present down both sides of the structure. Each segment was joined manually via a peptide bond using the program PyMol and, once complete, the model was energy-minimized and subjected to a 2 ns solvated molecular-dynamics simulation run at 4 °C using the program GROMACS. It became immediately apparent from the simulation that the model was unable to support a second row of Ca^{2+} ions, as they were ejected during the course of the simulation (results not shown). However, each XGTGND turn of the model was able properly to co-ordinate and retain a Ca^{2+} ion during the simulation, and the β -strands both immediately before and after the XGTGND turn (XUXXGTGNDXUXUGGXUXG) were stable, with their U residues creating a hydrophobic core for the protein. The main problem was with the first XUX β -strand of each nonapeptide motif (GGXUXGXUX). When aligned with AP, it was constrained as a Ca^{2+} -bound turn instead of an XUX β -strand. This placed the side chain of each U residue in a solvent-exposed orientation, away from the hydrophobic core of the model, a situation that had a destabilizing effect on the structure. To remedy the problem, the original alignment was kept the same; however, the first XUX sequence from each nonapeptide motif of Region IV was restrained in Modeller to a β -strand conformation rather than a Ca^{2+} -bound turn. This pointed the U residue of the XUX motif towards the interior of the structure, allowing it to interact with the other hydrophobic residues from the other β -strands.

Description of the model

The final model of MpAFP Region IV is a right-handed β -helix made up of 11 Ca^{2+} -binding loops (Figure 6A). Each loop contains one XGTGND Ca^{2+} -binding turn followed by three β -strands, with the U residues from each β -strand contributing to the hydrophobic core of the model (Figure 6B). This precludes a second row of bound Ca^{2+} ions as is seen in AP. There are three main factors that stabilize the structure:

- (1) Ca^{2+} ions which co-ordinate the proper folding of each XGTGND turn
- (2) a conserved hydrophobic core located throughout the length of the model
- (3) hydrogen bonding between backbone carbonyl oxygen atoms and the amide nitrogen atoms from the parallel β -strands of each loop

The architecture of the model also allows for main-chain hydrogen bonding in the turns between the β -strands of each loop, therefore increasing the stability of each turn and hence the overall structure (Figure 6B). The most striking feature of the model is the conserved row of threonine and aspartic acid/asparagine residues that align down the Ca^{2+} -bound side of the model (Figure 6C). This immediately drew our attention, because flat conserved arrays of threonine residues have been identified as the IBFs of the β -helical insect AFPs [13,14].

A solvated 2 ns molecular-dynamics simulation of the model was performed at 4 °C to test the feasibility of the structure (Figure 6A). The overall architecture of the model was maintained throughout the course of the simulation, with Ca^{2+} remaining bound in each XGTGND turn and the hydrophobic U residues of each β -strand creating the hydrophobic core. The slight twist present in the model prior to the simulation was relieved during the course of the simulation. The maintenance of the fold demonstrated that the model was stable and feasible. This was encouraging, as it identified a novel ice-binding motif (the threonine and aspartic acid/asparagine residues of each XGTGND

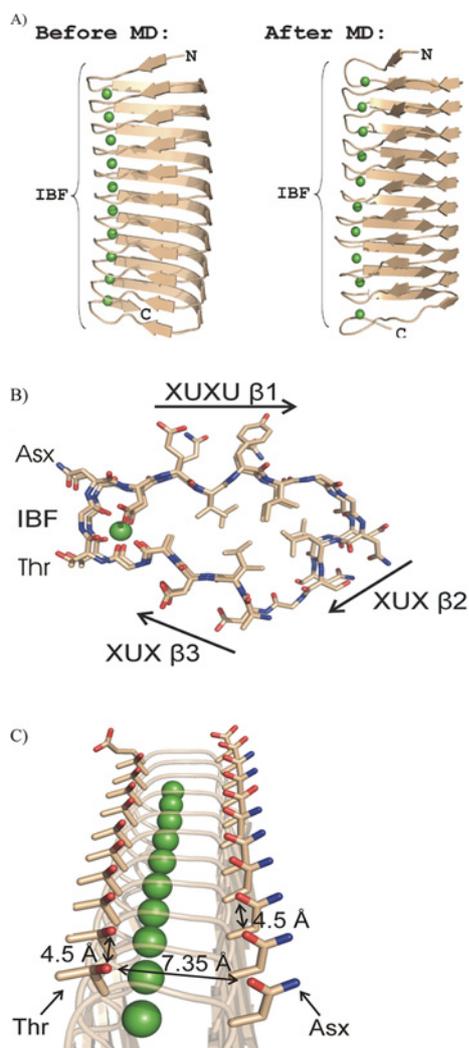


Figure 6 β -Helix model of *MpAFP* Region IV

(A) Cartoon depiction of *MpAFP* Region IV model. The model consists of an 11-loop β -helix with ten Ca^{2+} ions (shown in green) bound down the side of the IBF. Abbreviations: After MD, after 2 ns molecular-dynamics simulation; Before MD, before 2 ns molecular-dynamics simulation; C, C-terminus. N, N-terminus. The model retains its overall fold after the simulation, with only a minor amount of unfolding present at the C-terminus of the structure. (B) Cross-section of *MpAFP* Region IV residues 39–75. Key to colours: beige, carbon; red, oxygen; blue, nitrogen; green, Ca^{2+} . The IBF with the ice-binding threonine and aspartic acid/asparagine residues is indicated. The three β -strands of each 19-aa repeat are indicated with black arrows, along with their consensus sequence. (C) IBF of *MpAFP* Region IV. Each XGTGND turn of the β -helix contributes a threonine and an aspartic acid/asparagine residue to the IBF of the structure. Ca^{2+} ions co-ordinate the proper folding of the XGTGND repeats and align the threonine and aspartic acid/asparagine residues in a flat repetitive array.

Ca^{2+} -bound turn) as well as a novel β -helical fold (Ca^{2+} bound solely down one side of the model; hydrophobic core located throughout the rest). We therefore decided to test the validity of the model using site-directed mutagenesis to confirm the location of the putative IBF and to verify its overall fold.

Site-directed mutagenesis and TH values for mutants

The interaction between an AFP and ice has been described using the analogy of a receptor (AFP) binding to its ligand (ice) [20]. This binding is dependent upon the intimate surface/surface complementarity between the protein and the ice lattice.

Therefore, the introduction of a large bulky residue to the relatively flat IBF of an AFP is very effective at lowering antifreeze activity by hindering the simultaneous engagement of all residues that normally contact the surface of an ice crystal [21–23]. We therefore made a series of sterically hindering tyrosine mutations to residues located both on and off the putative IBF of *MpAFP* Region IV to confirm the location of the IBF and to validate the structure of the theoretical model.

Four IBF mutants were created [two single IBF mutants (T81Y and N65Y) and two double IBF mutants (T81Y/N65Y and T81Y/T141Y)] (Figure 7A). Each mutant was purified in the same manner as wild-type Region IV and assayed for TH activity over a similar range of concentrations (0.9 down to 0.025 mg/ml). As seen in Figure 7(B), both single IBF mutants (T81Y and N65Y) showed TH activity to be decreased by 50% compared with that of wild-type Region IV. This indicated that both the threonine and aspartic acid/asparagine residues of each XGTGND Ca^{2+} -bound turn contribute to the antifreeze activity of the protein. The double IBF mutants (T81Y/T141Y and T81Y/N65Y) each showed a decrease in the residual TH activity of a further 50%. Two single non-IBF mutants were also made (T105Y and T111Y). Thr¹⁰⁵ lies just off the IBF and Thr¹¹¹ is on the opposite surface (Figure 7A). Neither mutant showed essentially any change in TH activity as compared with the wild-type, with values of about 2 °C at a concentration of 0.5 mg/ml. This was as expected, as they are located away from the XGTGND Ca^{2+} -bound turn of the model and are therefore not part of the IBF.

A final mutant (V93R) was created in order to probe the orientation of the side chain of the U residue at position 4 of the nonapeptide motif in the 19-aa loop (GGXUXGXUX) (Figure 7A). We were originally unsure as to the orientation of the side chain of this residue, as it is normally present as glycine in AP (ie: GGXGXDXUX), a situation that allows for the proper folding of a Ca^{2+} -bound turn. However, if the side chain of the U residue points towards the interior of the structure as part of a β -strand instead of Ca^{2+} -bound turn, changing it to a long positively charged arginine residue should disrupt the hydrophobic core of the protein and affect its folding. This would diminish the intimate surface/surface complementarity required for *MpAFP* Region IV to bind an ice crystal and therefore decrease its TH activity. TH values of the V93R mutant were the lowest of any mutant tested, with a decrease to less than 25% of wild-type values at a concentration of 0.9 mg/ml.

To validate the mutagenesis data further, the folding of wild-type Region IV and the T105Y, T81Y/N65Y and V93R mutants was analysed via CD spectroscopy (Figure 8). Wild-type Region IV displayed a distinctive β -sheet spectrum, with a maximum molar ellipticity at 196 nm and minimum at 218 nm. Both the double IBF mutant T81Y/N65Y and the single non-IBF mutant T105Y displayed CD spectra virtually identical with that of wild-type Region IV. This indicated correct folding of the mutated constructs. The spectrum of the V93R mutant showed a dramatic change, with a large decrease in the amplitude of both the maximum and minimum molar ellipticity. This indicated a decrease in the β -sheet content of the protein and an increase in the random-coil content. These data strongly suggested that the side chain of the U residue at position 4 of the nonapeptide motif of each 19-aa loop (GGXUXGXUX) points towards the interior of the structure, therefore confirming the notion that, unlike the situation in AP, Ca^{2+} is not present in that side of the model. As further evidence for the destabilizing effect of the V93R mutation on the structure of Region IV, the mutant was eluted from a Superdex G-75 size-exclusion column a full 10 ml earlier (50 ml as against 60 ml) than any other mutant or the wild-type (results not shown). Clearly *MpAFP* Region IV is destabilized by the

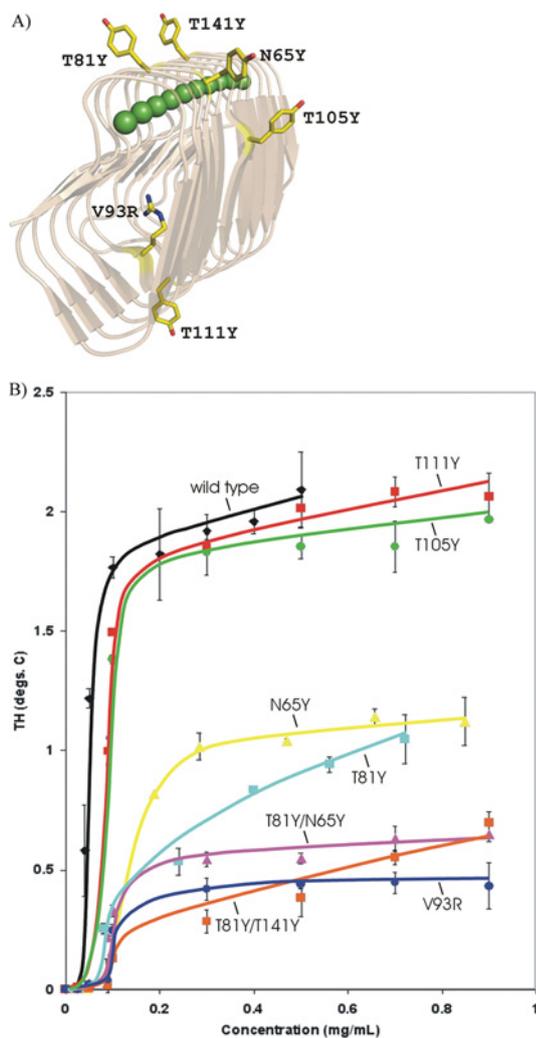


Figure 7 TH activity of *MpAFP* Region IV mutants

(A) The location of each mutation performed on Region IV is shown with yellow-coloured residues. (B) The TH of each Region IV mutant was plotted as a function of concentration. Values are the averages \pm S.D. of three measurements. Key to coloured curves: black, wild-type Region IV; red, T111Y; green, T105Y; yellow, N65Y; aqua, T81Y; orange, T81Y/N65Y; pink, T81Y/T141Y; dark blue, V93R. Abbreviation: degs. C, $^{\circ}$ C.

V93R mutation, with a resulting increase in the Stokes radius of the protein.

Q-TOF MS on *MpAFP* Region IV

Further confirmation for the presence of only a single row of Ca^{2+} in the β -helix of *MpAFP* Region IV was obtained by Q-TOF MS (Table 1). Two distinct masses were observed for Region IV in the presence of EDTA. The first mass of 34 405.6 Da corresponds to the predicted mass of Region IV (34 405 Da). The second mass of 34 423.4 Da corresponds to the predicted mass of Region IV with a single ammonium adduct (+18 Da, due to the presence of NH_4^+ ions in the sample buffer). A distribution of masses was observed for Region IV in the presence of CaCl_2 centering on 34 822.1 Da and increasing and decreasing by 40 Da, the mass of a single Ca^{2+} ion. The theoretical average mass of Region IV with ten Ca^{2+} ions bound (the number of Ca^{2+} ions predicted by the model) is 34 805 Da (an additional 10×40 Da). Each peak observed with the CaCl_2 sample had a single NH_4^+ ion bound, increasing the mass of each species by 18 Da. Therefore the difference between

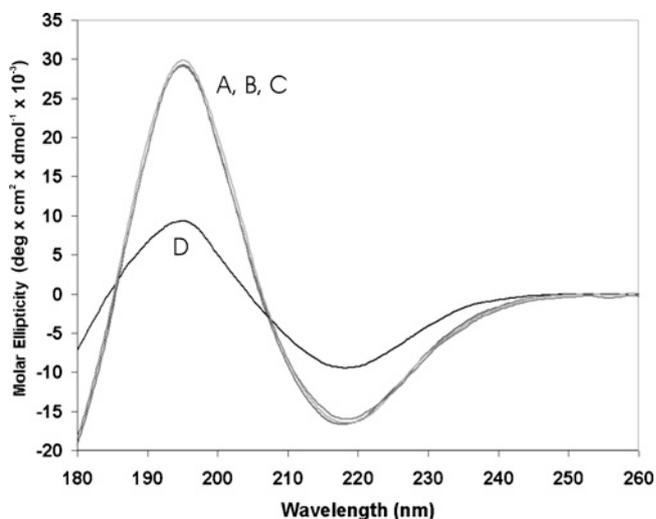


Figure 8 CD spectra of *MpAFP* Region IV mutants

The folding of mutants V93R (D), T105Y (A), T81Y/N65Y (B) and wild-type recombinant Region IV (C) was monitored by CD. The spectra were plotted as $10^3 \times$ molar ellipticity (degrees \cdot $\text{cm}^2 \cdot \text{dmol}^{-1}$) versus wavelength.

the mass of the predominant species in the CaCl_2 sample and the mass of the EDTA sample with an ammonium adduct is 399 Da. This mass divided by the mass of a single Ca^{2+} ion is 9.9, which corresponds to 10 Ca^{2+} ions bound in the model. This is exactly the number predicted for Region IV. The distribution of peaks observed with the CaCl_2 sample is most likely the result of dynamic exchange of Ca^{2+} ions at the ends of the β helix. (Note: there would be half Ca^{2+} binding sites at each end of the β -helix.) Also, the mass value corresponding to 11 Ca^{2+} ions bound would indicate that the sequence before or after Region IV is able to bind Ca^{2+} as well. However, the manner in which it does so cannot be predicted by homology-based modelling. If indeed *MpAFP* Region IV did bind two Ca^{2+} ions per loop of the structure, a distribution of masses centering on 35 205 Da would be expected (additional $20 \times 40 = 800$ Da). This is clearly not the case.

DISCUSSION

General discussion of the model and comparison with hyperactive insect AFPs

Our model of *MpAFP* Region IV, based partly on the Ca^{2+} -bound β -roll from AP, is able to explain both the Ca^{2+} -dependency and hyperactivity of this AFP. Ca^{2+} is necessary to co-ordinate the proper folding of each XGTGND turn in the model. This, in combination with the repetitive nature of Region IV, creates a long flat IBF on the protein, consisting of the outward-projecting threonine and aspartic acid/asparagine residues from each XGTGND Ca^{2+} -bound turn on that side of the β -helix. The threonine and aspartic acid/asparagine residues create a novel ice-binding motif that is similar to the TXT ice-binding motif of the hyperactive β -helical insect AFPs from both *C. fumiferana* [13] and that of the yellow mealworm beetle, *Tenebrio molitor* [14]. The hyperactivity of the insect AFPs is attributed to their ability to bind the basal planes of ice in addition to prism planes [13,14,24], a phenomenon not attributed to the moderately active fish AFPs [4]. The insect AFPs contain a repetitive array of TXT β -strands that align in a flat parallel β -sheet on one side of the protein [13,14,24]. The distance between side-chain

Table 1 Peak list and masses from Q-TOF MS on Region IV

A solution of Region IV was dialysed against 10 mM Tris/HCl and either 2 mM CaCl₂ or 2 mM EDTA. Both samples were diluted in ammonium formate and subjected to Q-TOF MS in the static-spray mode to determine their respective masses. For further details, see the text. N/A, not applicable.

Sample	Peak no.	Mass observed (± 1 Da)	NH ₄ ⁺ adduct (+18 Da)	Predicted mass (Da)	Difference from observed mass of Region IV in EDTA+1NH ₄ ⁺ adduct (Da)	Difference divided by mass of single Ca ²⁺ ion	Number of Ca ²⁺ ions bound
2 mM EDTA	1	34405.6	No	34405	N/A		0
	2	34423.4	Yes	34423	0		0
2 mM CaCl ₂	1	34783.9	Yes	34783	360.5	9.0	9
	2	34822.1	Yes	34823	398.7	9.9	10
	3	34860.6	Yes	34863	437.2	10.9	11

oxygen atoms of threonine in the same β -strand is about 7.35 Å (1 Å = 0.1 nm), whereas the distance between side-chain oxygen atoms of threonine from adjacent β -strands is about 4.5 Å. These distances match the spacing of oxygen atoms in an ice lattice on both the primary prism and basal planes. Oxygen atoms repeat at 4.52 Å along the a -axis in both the primary prism and basal planes. Oxygen atoms also repeat at 7.35 Å along the c -axis of the primary prism plane and at 7.8 Å along the a -axis of the basal plane. In the model of *MpAFP* Region IV, the spacing between threonine and asparagine side-chain oxygen atoms in the same Ca²⁺-bound turn is about 7.35 Å, whereas the spacing between the side-chain oxygen atoms of both threonine and asparagine from adjacent Ca²⁺-bound turns is about 4.5 Å. This spacing, which could only occur in the Ca²⁺-bound form of the protein, would allow Region IV to bind both the primary prism and basal planes of an ice crystal, therefore explaining its Ca²⁺-dependent hyperactive antifreeze activity.

We have chosen this uniform orientation of the threonine residues that places the methyl groups on the exterior of the IBF, because the spacing between oxygen atoms matches that seen with the two insect AFPs. However, although this orientation is consistent with the molecular-dynamics simulations performed, there is no clear indication that it is preferred over one where the methyl groups are on the inside in place of the hydroxy groups. Another reason to choose the first orientation is that the linear array of threonine hydroxy groups is then flanked, and perhaps insulated, by two hydrophobic zones, one formed by the outer rank of threonine methyl groups and the other by the rank of backbone methylene groups from the C α atoms of the glycine residues located between the threonine and aspartic acid/asparagine residues in the Ca²⁺-bound turns of the β -helix. Here the methylene groups may serve a similar function to that of the methyl groups of the second threonine array in the insect AFPs [13,14]. Thus, although at first sight the IBF of *MpAFP* appears to be dominated by hydrogen-bonding groups, the threonine hydroxy groups are flanked by hydrophobic zones, as are the threonine hydroxy groups of the longest threonine array in the insect AFPs. The relative contribution of hydrogen-bonding and hydrophobic interactions to ice binding continues to be a contentious issue and is an active area of investigation in the antifreeze field [25–28].

The difference in location and aa composition of the IBF of *MpAFP* Region IV compared with that in the hyperactive β -helical insect AFPs is noteworthy. *MpAFP* Region IV is the first putative β -helical AFP to locate its IBF in a Ca²⁺-bound turn as opposed to a flat parallel β -sheet, as is seen with the insect AFPs. Region IV is also the first AFP to use a TGN ice-binding motif. This is in contrast with the TXT sequences that define the ice-binding motifs of the insect AFPs. The X residue of each insect

TXT motif points towards the interior of the structure, helping to stabilize the β -helix and to avoid steric clashes on the IBF of the protein. Glycine is necessary in the TGN ice-binding motif of Region IV to allow proper folding of each XGTGND Ca²⁺-bound turn and to maintain a flat ice-binding surface. The fact that such different ice-binding motifs in different locations of a β -helix are able to lower the freezing point of a solution in a hyperactive manner is a testament to the plasticity of the β -helix as an AFP structural fold.

The aa composition of *MpAFP* Region IV is rich in acidic residues (pI 3.67, calculated from the aa composition). Aspartic acid accounts for 13.5% of the residues, whereas lysine and arginine combined account for only 2.9% of the total residues. Glycine is the most prevalent aa, at 18.4%. The region is also high in threonine, serine, (iso)leucine and valine, and it does not contain a single proline or cysteine residue. Proteins high in these specific residues have a propensity to fold as β -helices [29]. The high proportion of glycine is necessary to allow proper turns in the helix, while valine and (iso)leucine have high β -sheet propensities and their side chains create the hydrophobic core for the protein. Proline residues are under-represented, owing to their destabilizing effects on β -sheets. The actual pI of Ca²⁺-bound Region IV is probably higher than pH 3.67, owing to the fact that 11 of its 28 aspartic acid residues point inwards to co-ordinate Ca²⁺ ions in the XGTGND turns. Their charge is offset by the Ca²⁺ ions and they are not exposed to the solvent.

We were originally unsure of the contribution that aspartic acid/asparagine residues of the XGTGND Ca²⁺-bound turns made towards ice binding. However, the mutagenesis data clearly show that the threonine and aspartic acid/asparagine residues contribute roughly equally to the ice-binding ability of this protein. Single mutations of residues located in the middle of the flat IBF, namely T81Y or N65Y, both lowered TH levels by 50%. The double threonine or threonine/asparagine mutants, namely T81Y/N65Y and T81Y/T141Y, both lowered TH levels to 25%. This clearly implicates the aspartic acid/asparagine residues of each XGTGND Ca²⁺-bound turn in ice-binding, therefore making this the first report of a repetitive array of both threonine and aspartic acid/asparagine constituting the IBF of an AFP.

Structural basis for the novel β -helical fold

The presence of a single row of bound Ca²⁺ ions in Region IV, as opposed to two, is understandable when its aa sequence is compared with that of AP. Each 19-aa loop of Region IV contains only one XGTGND Ca²⁺-binding motif. This is in contrast with each 18-aa loop of AP, which contains two GGXGXD Ca²⁺-binding motifs. Region IV only requires Ca²⁺ down the side of the IBF, because it is stabilized by a more extensive hydrophobic

core down the opposite side of the structure. Every 18-aa loop of AP contains two inward-pointing hydrophobic U residues that are found in each XUX β -strand. The 19-aa loops of Region IV contain 4 hydrophobic U residues that point inwards. These two extra hydrophobic residues are located at the end of each decapeptide motif (that is, XGTGNDXUXU) and in the first β -strand of each nonapeptide motif (that is, GGXUXGXUX). The aspartic acid-to-glycine mutation at position 6 of each nonapeptide motif of Region IV (that is, GGXUXGXUX) was perhaps the strongest indicator that Ca^{2+} cannot be bound down the side opposite the IBF, since aspartic acid residues are crucial in the co-ordination of Ca^{2+} in a GGXGX $\underline{\text{D}}$ turn. Molecular-dynamics simulations suggested that, in the absence of the aspartic acid residues, the remaining interactions with the carbonyl oxygen atom of the glycine and 'X' residues are insufficient to keep the Ca^{2+} in place (results not shown). Instead, the conserved glycine residue at that position facilitates the turn between XUX β -strands in each GGXUXGXUX nonapeptide motif of Region IV.

Differences between Region IV and other β -helical proteins

MpAFP is the first reported instance of a β -helix that contains a single Ca^{2+} -binding turn per loop. There are over 100 structures in the PDB (Protein Data Bank) today that contain some variation of a β -helix fold [29]. The only β -helical proteins identified to date that bind Ca^{2+} and utilize it in a structural role are AP and its homologues, as well as C5-epimerase from *Azotobacter vinlandii* [30]. They all contain the archetypal RTX nonapeptide motif of a GGXGX $\underline{\text{D}}$ Ca^{2+} -binding turn followed by a three-residue XUX β -strand. The 19-aa repeats of *MpAFP* Region IV therefore represent a novel β -helical fold that contains a single RTX Ca^{2+} -binding motif per loop of the structure. The prevalence of these novel β -helical 19-aa repeats in other proteins in the database is much lower as compared with the classic tandemly repeated RTX nonapeptide motifs. The top three BLASTp hits to *MpAFP* Region IV are RTX-like proteins produced by alphaproteobacteria from the *Magnetospirillum* genus. All three contain 12 copies of the 19-aa β -helix repeat that are nearly identical with *MpAFP* Region IV. However, they do not contain the conserved threonine and asparagine residues at the XGTGND position of the Ca^{2+} -bound turn that is seen in Region IV. We have cloned and recombinantly expressed the top BLASTp hit to Region IV (gi = 83311100) in *E. coli* as a C-terminal His₆-tagged protein (results not shown). The protein did not display any TH activity, further demonstrating the necessity of the conserved threonine and aspartic acid/asparagine residues for ice binding.

Comparison of *MpAFP* with models of plant AFPs

Asparagine residues have been implicated in ice binding in a carrot (*Daucus carota*) AFP [31]. A theoretical model exists for the protein, which has been predicted to fold as a β -helical leucine-rich repeat with loops of 23–25 aa in length. A well-conserved leucine/isoleucine hydrophobic core stabilizes the structure, and the putative IBF of the model consists of an asparagine residue in each loop flanked by hydrophilic residues on either side. The TH activity of the wild-type protein is very low, however, with TH levels of about 0.4 °C at a concentration of 2.5 mg/ml. It does have appreciable ice re-crystallization inhibition activity, however. The authors of [31] support the contention that the low TH activity of the protein is the result of the plant having evolved a freeze-tolerance strategy as opposed to a freeze-avoidance strategy, where high levels of TH are actually detrimental to the organism.

A theoretical β -roll model also exists for the AFP from *Lolium perenne* (perennial ryegrass) [32]. It contains a conserved

hydrophobic core of valine residues, with an asparagine 'ladder' at each end of the roll that helps to stabilize the structure. These asparagine residues are internal and make hydrogen-bond contacts with the β -roll backbone. The model contains two potential ice-binding surfaces on either side of the structure, which consist of two rows of solvent-accessible residues similar to those of the insect AFPs, but less well conserved. As was the case for the carrot AFP, this AFP has very low TH activity, but is very effective at inhibiting ice recrystallization, a common trend displayed among plant AFPs.

Comparison of *MpAFP* with other bacterial AFPs

Our biophysical characterization, modelling and site-directed mutagenesis of *MpAFP* Region IV has generated the first plausible structure of a bacterial AFP. Prior to this presently reported study, only two other bacterial AFPs had been characterized at the sequence level. The arctic plant-growth-promoting rhizobacterium *Pseudomonas putida* GR12-2 secretes a 164 kDa lipoglycoprotein (afpA) that contains five RTX-like XGXGX $\underline{\text{D}}$ repeats within its sequence [33]. AfpA shows affinity for ice by being able to shape ice crystals, but its antifreeze activity is questionable, since no TH value has been reported. It is unclear whether afpA's XGXGX $\underline{\text{D}}$ repeats play an active role in its ability to shape ice. TH assays of the recombinantly expressed protein were performed in 1mM EDTA, which, in the case of *MpAFP* Region IV, would have destroyed its ice-shaping ability. However, afpA was still able to shape ice crystals. It is therefore possible that afpA's XGXGX $\underline{\text{D}}$ repeats perform some function involved in the secretion and subsequent extracellular targeting of the protein and not in its antifreeze activity. This idea is bolstered by the fact that these repeats do not contain the requisite threonine and aspartic acid/asparagine residues at the corresponding XGXGX $\underline{\text{D}}$ positions of the motif that have been shown to be vital for TH activity in *MpAFP* Region IV. Also, if indeed afpA's XGXGX $\underline{\text{D}}$ repeats fold as a Ca^{2+} -stabilized β -roll, the size of the roll, and hence the putative IBF, of the structure would likely be too small to effectively inhibit the growth of ice.

More recently, an Antarctic strain of the Gram-negative bacterium *Colwellia* sp. SLW05 was found to produce a 26 kDa IBP (ice-binding protein) that it secretes into the medium [5]. The protein is able to shape ice crystals, but it displays extremely low TH values (< 0.1 °C). The aa sequence of the protein is similar to those of IBPs of other sea-ice bacteria (*Polaribacter irgensii* and *Psychromonas ingrahamii*), a sea-ice diatom (*Navicula glaciei*) and a snow-mould fungus (*Typhula ishikariensis*). There is, however, no similarity between it and any region of *MpAFP*. Like the plant AFPs from *D. carota* and *L. perenne*, it is hypothesized to function as an inhibitor of ice recrystallization, where it can effectively limit the growth of large ice crystals at the expense of smaller ones.

In conclusion, we have identified a novel β -helical ice-binding fold found in the region of *MpAFP* responsible for its antifreeze activity. On the basis, partly, of the β -roll from AP, the model for Region IV identified a repetitive array of threonine and aspartic acid/asparagine found in each XGTGND Ca^{2+} -bound turn of the β -helix that constitutes the IBF of the protein. A single row of Ca^{2+} ions is required in the structure to co-ordinate properly the folding of each XGTGND turn. This turn is followed by three β -strands in each loop, the hydrophobic residues of which point inwards, creating a structurally stabilizing hydrophobic core for the protein. The Ca^{2+} -dependent hyperactivity of the protein is explained by the ability of the side-chain oxygen atoms of the threonine and aspartic acid/asparagine residues located in the Ca^{2+} -bound turns of the β -helix to mimic the distance of oxygen atoms in an ice

crystal in multiple planes of the ice lattice, a mechanism similar to that employed by the hyperactive β -helical insect AFPs.

Note added in proof (Received 7 February 2008)

Since this paper was proofed out, the crystal structure of an extracellular lipase from *Serratia marcescens* was solved [34]. The structure reveals a β -roll with Ca^{2+} ions bound internally down one side of the roll, lending further support to the plausibility of the model put forward in the present paper.

We thank Mr Kim Munro and Dr David Hyndman from the Protein Function Discovery facility at Queen's University for help with acquiring and interpreting CD and MS data respectively. We thank Mrs Sherry Gauthier of this Department for her expert technical assistance with the mutagenesis, and Dr Laurie A. Graham, also of this Department, for her thoughtful critique of this manuscript before its submission. This research was funded by a grant to P.L.D. from the Canadian Institutes for Health Research. P.L.D. holds a Canada Research Chair in Protein Engineering. C.P.G. was supported by a Natural Sciences and Engineering Research Council of Canada 3-Year Postgraduate Doctoral Scholarship (NSERC-PGSD3).

REFERENCES

- DeVries, A. L. and Wohlschlag, D. E. (1969) Freezing resistance in some Antarctic fishes. *Science* **163**, 1073–1075
- Webb, M. S., Uemura, M. and Steponkus, P. L. (1994) A comparison of freezing injury in oat and rye: two cereals at the extremes of freezing tolerance. *Plant Physiol.* **104**, 467–478
- Kristiansen, E. and Zachariassen, K. E. (2005) The mechanism by which fish antifreeze proteins cause thermal hysteresis. *Cryobiology* **51**, 262–280
- Scotter, A. J., Marshall, C. B., Graham, L. A., Gilbert, J. A., Garnham, C. P. and Davies, P. L. (2006) The basis for hyperactivity of antifreeze proteins. *Cryobiology* **53**, 229–239
- Raymond, J. A., Fritsen, C. and Shen, K. (2007) An ice-binding protein from an Antarctic sea ice bacterium. *FEMS Microbiol. Ecol.* **61**, 214–221
- Urrutia, M. E., Duman, J. G. and Knight, C. A. (1992) Plant thermal hysteresis proteins. *Biochim. Biophys. Acta* **1121**, 199–206
- Gilbert, J. A., Davies, P. L. and Laybourn-Parry, J. (2005) A hyperactive, Ca^{2+} -dependent antifreeze protein in an Antarctic bacterium. *FEMS Microbiol. Lett.* **245**, 67–72
- Gilbert, J. A., Hill, P. J., Dodd, C. E. and Laybourn-Parry, J. (2004) Demonstration of antifreeze protein activity in Antarctic lake bacteria. *Microbiology* **150**, 171–180
- Coote, J. G. (1992) Structural and functional relationships among the RTX toxin determinants of Gram-negative bacteria. *FEMS Microbiol. Rev.* **8**, 137–161
- Baumann, U., Wu, S., Flaherty, K. M. and McKay, D. B. (1993) Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel β roll motif. *EMBO J.* **12**, 3357–3364
- Li, Z., Lin, Q., Yang, D. S., Ewart, K. V. and Hew, C. L. (2004) The role of Ca^{2+} -coordinating residues of herring antifreeze protein in antifreeze activity. *Biochemistry* **43**, 14547–14554
- Liu, Y., Li, Z., Lin, Q., Kosinski, J., Seetharaman, J., Bujnicki, J. M., Sivaraman, J. and Hew, C. L. (2007) Structure and evolutionary origin of Ca^{2+} -dependent herring type II antifreeze protein. *PLoS ONE* **2**, e548
- Graether, S. P., Kuiper, M. J., Gagné, S. M., Walker, V. K., Jia, Z., Sykes, B. D. and Davies, P. L. (2000) β -Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature* **406**, 325–328
- Liou, Y. C., Tocilj, A., Davies, P. L. and Jia, Z. (2000) Mimicry of ice structure by surface hydroxyls and water of a β -helix antifreeze protein. *Nature* **406**, 322–324
- Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
- DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA (<http://www.pymol.org>)
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718
- Chakrabarty, A. and Hew, C. L. (1991) The effect of enhanced α -helicity on the activity of a winter flounder antifreeze polypeptide. *Eur. J. Biochem.* **202**, 1057–1063
- Tyshenko, M. G., Doucet, D., Davies, P. L. and Walker, V. K. (1997) The antifreeze potential of the spruce budworm thermal hysteresis protein. *Nat. Biotechnol.* **15**, 887–890
- Jia, Z. and Davies, P. L. (2002) Antifreeze proteins: an unusual receptor–ligand interaction. *Trends Biochem. Sci.* **27**, 101–106
- Baardsnes, J., Kondejewski, L. H., Hodges, R. S., Chao, H., Kay, C. M. and Davies, P. L. (1999) New ice-binding face for type I antifreeze protein. *FEBS Lett.* **463**, 87–91
- Baardsnes, J., Jelokhani-Niaraki, M., Kondejewski, L. H., Kuiper, M. J., Kay, C. M., Hodges, R. S. and Davies, P. L. (2001) Antifreeze proteins from shorthorn sculpin: identification of the ice-binding surface. *Protein Sci.* **10**, 2566–2576
- Marshall, C. B., Daley, M. E., Graham, L. A., Sykes, B. D. and Davies, P. L. (2002) Identification of the ice-binding face of antifreeze protein from *Tenebrio molitor*. *FEBS Lett.* **529**, 261–267
- Graether, S. P. and Sykes, B. D. (2004) Cold survival in freeze-intolerant insects: the structure and function of β -helical antifreeze proteins. *Eur. J. Biochem.* **271**, 3285–3296
- Jorov, A., Zhorov, B. S. and Yang, D. S. (2004) Theoretical study of interaction of winter flounder antifreeze protein with ice. *Protein Sci.* **13**, 1524–1537
- Yang, C. and Sharp, K. A. (2005) Hydrophobic tendency of polar group hydration as a major force in type I antifreeze protein recognition. *Proteins* **59**, 266–274
- Doxey, A. C., Yaish, M. W., Griffith, M. and McConkey, B. J. (2006) Ordered surface carbons distinguish antifreeze proteins and the ice-binding regions. *Nat. Biotechnol.* **24**, 852–855
- Wierzbicki, A., Dala, P., Cheatham, 3rd, T. E., Knickelbein, J. E., Haymet, A. D. and Madura, J. D. (2007) Antifreeze proteins at the ice/water interface: three calculated discriminating properties for orientation of type I proteins. *Biophys. J.* **93**, 1442–1451
- Kajava, A. V. and Steven, A. C. (2006) β -Rolls, β -helices, and other β -solenoid proteins. *Adv. Protein Chem.* **73**, 55–96
- Aachmann, F. L., Svanem, B. I., Güntert, P., Petersen, S. B., Valla, S. and Wimmer, R. (2006) NMR structure of the R-module: a parallel β -roll subunit from an *Azotobacter vinelandii* mannuronan C-5 epimerase. *J. Biol. Chem.* **281**, 7350–7356
- Zhang, D. Q., Liu, B., Feng, D. R., He, Y. M., Wang, S. Q., Wang, H. B. and Wang, J. F. (2004) Significance of conservative asparagine residues in the thermal hysteresis activity of carrot antifreeze protein. *Biochem. J.* **377**, 589–595
- Kuiper, M. J., Davies, P. L. and Walker, V. K. (2001) A theoretical model of a plant antifreeze protein from *Lolium perenne*. *Biophys. J.* **81**, 3560–3565
- Muryoi, N., Sato, M., Kaneko, S., Kawahara, H., Obata, H., Yaish, M. W., Griffith, M. and Glick, B. R. (2004) Cloning and expression of *afpA*, a gene encoding an antifreeze protein from the arctic plant growth-promoting rhizobacterium *Pseudomonas putida* GR12–2. *J. Bacteriol.* **186**, 5661–5671
- Meier, R., Drepper, T., Svensson, V., Jaeger, K. E. and Baumann, U. (2007) A calcium-gated lid and a large β -roll sandwich are revealed by the crystal structure of extracellular lipase from *Serratia marcescens*. *J. Biol. Chem.* **282**, 31477–31483

Received 8 October 2007/7 December 2007; accepted 20 December 2007

Published as BJ Immediate Publication 20 December 2007, doi:10.1042/BJ20071372