

# $\beta$ -Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect

Steffen P. Graether\*<sup>†</sup>, Michael J. Kuiper<sup>‡</sup>, Stéphane M. Gagné<sup>§</sup>, Virginia K. Walker<sup>‡</sup>, Zongchao Jia\*, Brian D. Sykes<sup>§</sup> & Peter L. Davies\*<sup>‡</sup>

\* Department of Biochemistry, Queen's University, Kingston, Ontario K7L 3N6, Canada

<sup>‡</sup> Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada

<sup>§</sup> Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

<sup>†</sup> Present address: Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada.

Insect antifreeze proteins (AFP) are considerably more active at inhibiting ice crystal growth than AFP from fish or plants. Several insect AFPs, also known as thermal hysteresis proteins, have been cloned<sup>1–3</sup> and expressed<sup>1,2</sup>. Their maximum activity is 3–4 times that of fish AFPs<sup>1</sup> and they are 10–100 times more effective at micromolar concentrations. Here we report the solution structure of spruce budworm (*Choristoneura fumiferana*) AFP and characterize its ice-binding properties. The 9-kDa AFP is a  $\beta$ -helix with a triangular cross-section and rectangular sides that form stacked parallel  $\beta$ -sheets; a fold which is distinct from the three known fish AFP structures. The ice-binding side contains 9 of the 14 surface-accessible threonines organized in a regular array of TXT motifs that match the ice lattice on both prism and basal planes. In support of this model, ice crystal morphology and ice-etching experiments are consistent with AFP binding to both of these planes and thus may explain the greater activity of the spruce budworm antifreeze.

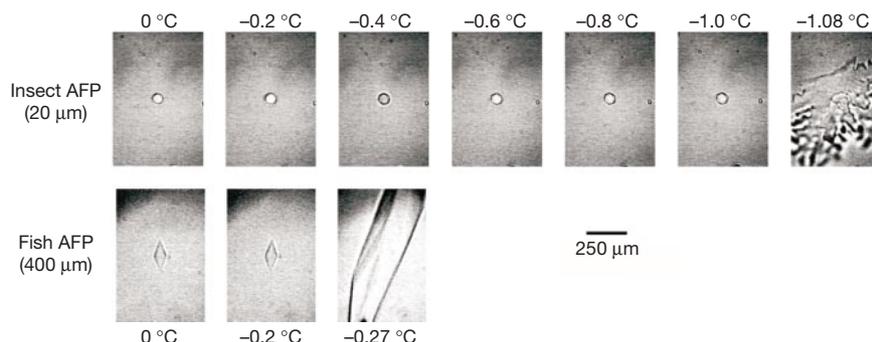
Recombinant spruce budworm antifreeze protein (sbwAFP) is as hyperactive as the native protein (Fig. 1), and was used for structure determination. We obtained 1014 nuclear Overhauser effect (NOE) distance and 88 dihedral angle restraints from several nuclear magnetic resonance (NMR) experiments and used them to generate an ensemble of structures (Fig. 2a). The overall fold is that of a left-handed  $\beta$ -helix (Fig. 2b), with one partly structured loop at the amino terminus (loop 1), four complete left-handed loops (loops

2–5) and one right-handed loop at the carboxy terminus (loop 6). The protein fold is stabilized by a hydrophobic core, and by disulphide bonds and parallel  $\beta$ -sheets between the loops. The left-handed loops are 15 residues long, and are spaced 4.5–5.7 Å apart. The right-handed loop is 17 residues long and has two anti-parallel  $\beta$ -sheets that cap the C-terminal end of the protein. In addition, the N, C $\alpha$  and C' atoms of the C-terminal residues 88–90 have a root mean square deviation (r.m.s.d.) of 0.53 Å, which is the same as the r.m.s.d. of the whole protein (Table 1). The cross-section of the protein perpendicular to the axis of the  $\beta$ -helix has a triangular shape. Of the rectangular sides, at least two consist of

**Table 1 Structural statistics for the ensemble of 20 refined structures of sbwAFP**

	(SA)*	(SA) <sup>†</sup>
r.m.s.d. from NOE distances (Å) (1014) <sup>¶</sup>	0.012 ± 0.002	0.014
r.m.s.d. deviations from dihedral angles (degrees) (88)	0.42 ± 0.07	0.44
r.m.s.d. deviations from carbon chemical shift		
$\delta C^{\alpha}$ (p.p.m.) (59)	1.43 ± 0.06	1.45
$\delta C^{\beta}$ (p.p.m.) (59)	1.62 ± 0.06	1.46
r.m.s.d. from idealized covalent geometry		
Bonds (Å)	0.0018 ± 0.0001	0.0019
Angles (degrees)	0.54 ± 0.01	0.55
Impropers (degrees)	0.46 ± 0.02	0.48
Final energies # (kcal mol <sup>-1</sup> )		
$F_{NOE}$	8 ± 2	9.8
$F_{cdih}$	0.9 ± 0.2	1.1
$F_{repel}$	12 ± 1	12.8
$E_{L-J}^{\ddagger}$	-240 ± 30	-244
Cartesian coordinate r.m.s.d. (Å) § (SA) versus (SA) <sub>avg</sub> <sup>  </sup>	N, C $\alpha$ and C'	All heavy atoms
	0.52 ± 0.09	1.0 ± 0.1
Ramachandran plot <sup>24</sup>	All structures	(SA) <sub>r</sub>
Residues in most favoured region	47.9	48.0
Residues in additionally allowed region	50.2	49.3

\* Represents the average root mean square deviations (r.m.s.d.) for the ensemble.  
<sup>†</sup> (SA)<sub>r</sub> is the single member of the ensemble that is closest to the average structure.  
<sup>‡</sup> The Lennard-Jones potential was not used in the refinement process.  
<sup>§</sup> Residues 4–5, 7–18, 23–69, 71–72, 80–90.  
<sup>||</sup> (SA)<sub>avg</sub> is the geometric average structure of the ensemble.  
<sup>¶</sup> Numbers in brackets represent the number of constraints used in the structure calculation.  
<sup>#</sup> Final energies are calculated based on contributions from nuclear Overhauser effect ( $F_{NOE}$ ), dihedral angles ( $F_{cdih}$ ), van der Waals repulsion ( $F_{repel}$ ) and Lennard-Jones potential ( $E_{L-J}$ ).



**Figure 1** Hyperactivity of insect antifreeze protein compared to a fish antifreeze protein. Ice crystals in the presence of 20  $\mu$ m spruce budworm (insect) antifreeze protein (sbwAFP) or 400  $\mu$ m winter flounder type I (fish) AFP were observed by video-microscopy during thermal hysteresis measurements at -0.2 °C intervals of cooling, which occurred at a rate of ~0.07 °C per min. There was no observable change in the ice crystals as the temperature was lowered until the non-equilibrium freezing point was exceeded at -1.08 °C for sbwAFP, and at -0.27 °C for type I AFP. Thus, at one-twentieth of the fish

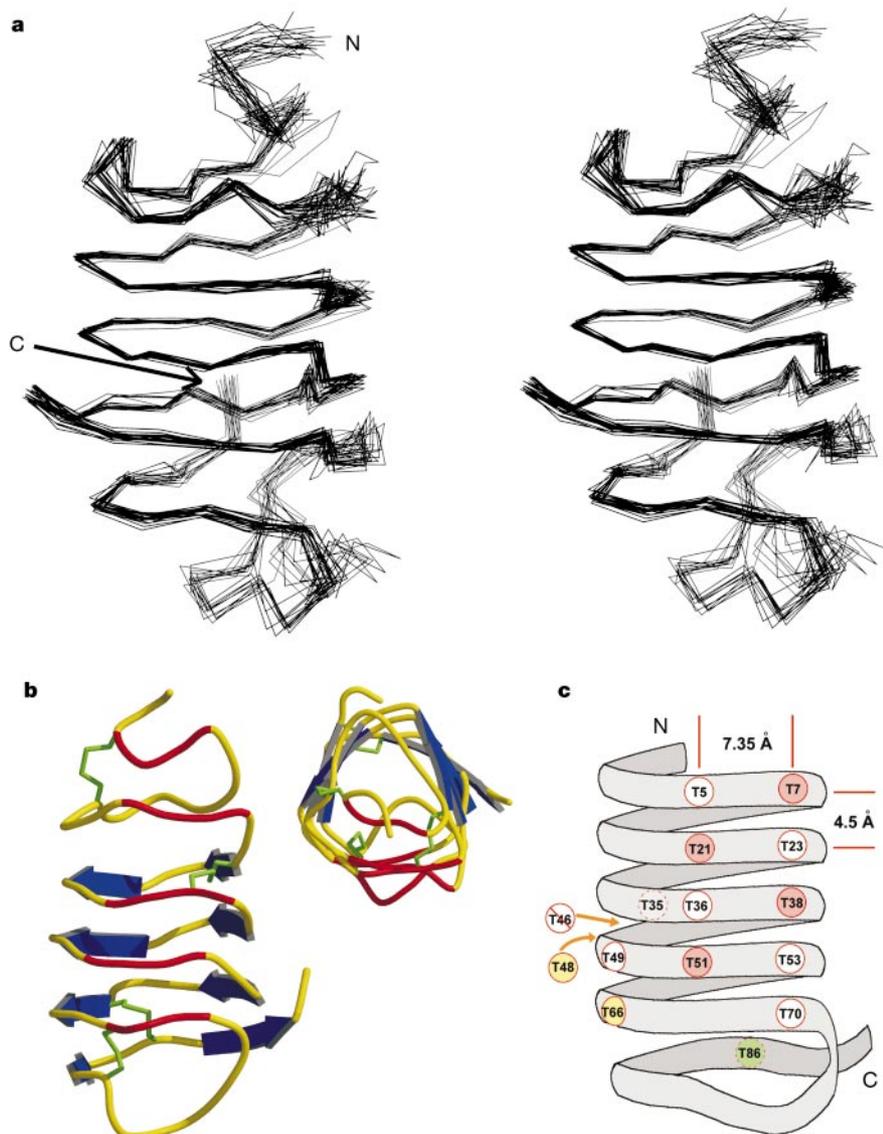
AFP molar concentration, sbwAFP causes about four times the depression of the freezing point (1.08 °C versus 0.27 °C). The hexagonal ice crystal obtained with sbwAFP lies with its *c*-axis perpendicular to the plane of the figure, and the bipyramidal crystal produced by type I AFP has its *c*-axis in the plane of the figure, parallel to its long dimension. Below the non-equilibrium freezing point, uncontrolled growth of the crystal in the presence of sbwAFP occurs laterally along the *a*-axis. In contrast, in the presence of type I fish AFP the crystal grows rapidly along the *c*-axis to form a spicule.

stacked, parallel  $\beta$ -sheets (Fig. 2b). With respect to the third side, 20–30% of the members of the ensemble have  $\beta$ -strand content, suggesting that there is some  $\beta$ -sheeting along this face as well. This third side is threonine-rich due to the clustering and alignment of four TXT motifs (where X is any inward pointing amino acid) at positions 5–7, 21–23, 36–38 and 51–53 (Fig. 2b and c). A fifth motif at 68–70 in the longer loop 5 is imperfect (IXT).

The left-handed  $\beta$ -helices characterized to date consist mostly of 18 residues per loop, but also have a triangular cross-section<sup>4,5</sup>. The 15-residue loop in sbwAFP is three residues shorter, and corresponds to the removal of one residue from each of the three sides of the triangle. In addition, other  $\beta$ -helices contain secondary structural elements and inserts that extend outward from the helix.

Therefore, sbwAFP and beetle AFP<sup>6</sup>, represent the first isolated  $\beta$ -helix structures determined.

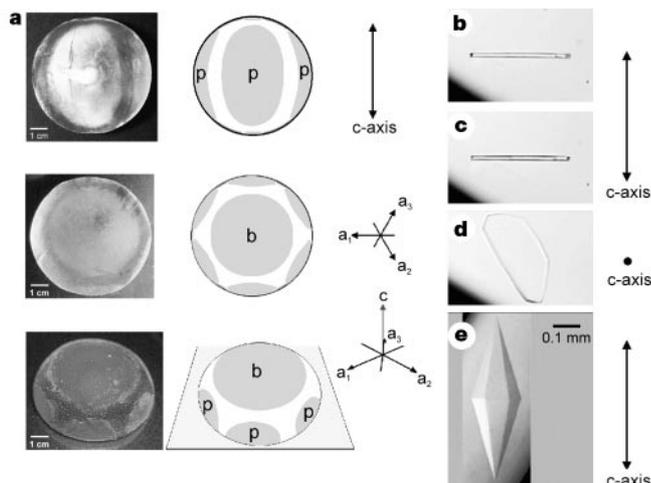
Given the importance of threonines for ice binding in other AFPs<sup>7–9</sup>, several of these residues in the TXT motifs were mutated to leucine to alter both the polarity and the shape of the protein surface. Mutants T7L, T21L, T38L and T51L each showed an 80–90% loss of activity (Gauthier *et al.*, unpublished results; Fig. 2c). Subsequently, threonines not present in the array of TXT motifs were also mutated. Mutants T48L and T66L, adjacent to the TXT motifs, showed only a moderate decrease in activity (30% and 35%, respectively). Mutation of threonine 86 to leucine, located opposite the TXT side, resulted in no loss of activity. Thus, we propose that the TXT face constitutes the ice-binding site of sbwAFP.



**Figure 2** Structure of spruce budworm antifreeze protein. **a**, Stereoview of C $\alpha$  atoms of residues 1–90 from the 20 lowest-energy structures of sbwAFP (of the 40 that converged from the 50 computed). No distance restraint was violated by more than 0.3 Å and no dihedral angle restraint by more than 4°. The structural data statistics are shown in Table 1. Residues in regions 5–8, 19–22, 35–38, 49–53 and 68–70 appear to be exchange-broadened as suggested by preliminary <sup>15</sup>N-relaxation experiments (data not shown), resulting in poorly defined residue positions. **b**, Ribbon representation of the sbwAFP structure closest to the mean. The  $\beta$ -sheets, shown in blue, are located at residues 25–28, 30–33, 40–43, 45–48, 55–58, 62–65, 67–69, 81–83 and 85–89. Disulphide bonds, shown in green, are located between loops 1 and 2, loops 2 and 3, and two

between loops 5 and 6. Coils shown in red indicate the threonine-rich face. The inset shows the view down the  $\beta$ -helix axis. **c**, Schematic representation of sbwAFP showing the relative positions of threonine residues. The putative ice-binding threonines (residues 5, 7, 21, 23, 36, 38, 51, 53 and 70) are aligned in a grid-like arrangement spaced 7.4 Å and 4.5 Å apart. The coloured circles indicate the following percentage antifreeze activities relative to wild-type protein: red, <30%; yellow, <60%; green, 100%; white, not determined. Dotted circles indicate projection of the threonine side chain on the other side of the coil. The line through T46 signifies that the T46L mutation is a folding mutant. The N and C termini are marked.

We performed ice etching<sup>10</sup> and crystal habit studies to determine which ice planes are bound by sbwAFP. Ice hemispheres were grown in the presence of dilute sbwAFP to incorporate the protein into the ice lattice on a binding surface. These binding planes were revealed by partial sublimation of the ice hemisphere at  $-20^{\circ}\text{C}$  as 'etches' of AFP accumulation left behind by the receding ice (Fig. 3a). Views of three different hemispheres show sbwAFP binding as distinct etches centred on the six equivalent primary prism planes (p) as well as the basal planes (b) of ice. The etched regions are shaded in the

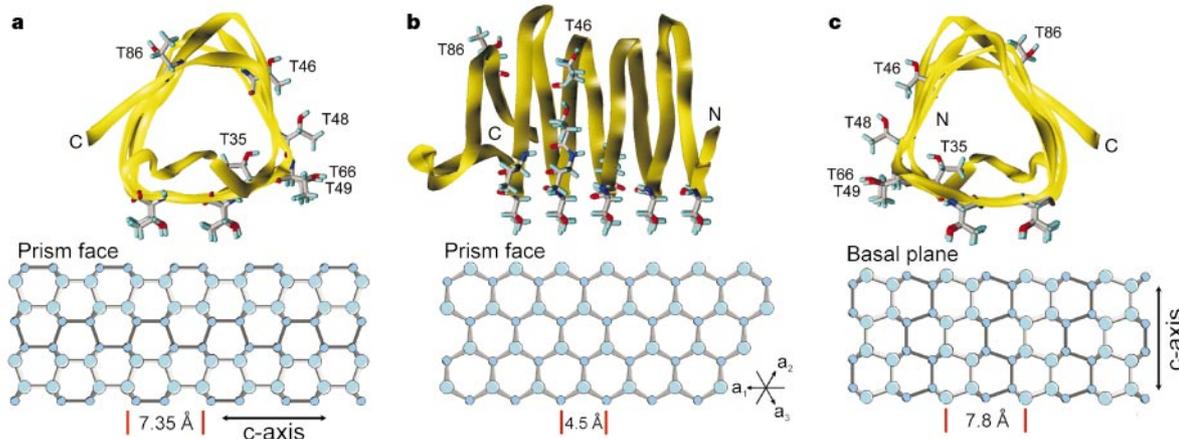


**Figure 3** Ice hemispheres and an ice crystal grown in the presence of slow AFP. **a**, The three ice hemispheres were each grown from a single orientated seed crystal in the presence of dilute sbwAFP ( $\sim 0.02\text{ mg ml}^{-1}$ ). They were subsequently allowed to sublime to reveal bound AFP as a proteinaceous coating (white etch on the ice surface). Alongside are the corresponding interpretive sketches in which the etched surfaces are shaded and identified as a primary prism plane (p) or basal plane (b). The identification as primary prism planes was confirmed by orientation to the facets formed on a drilled cavity after sublimation of the ice hemisphere. The unshaded regions in the sketches represent ice surfaces not bound by AFP, which are most evident in the oblique view in the bottom panel. The relative orientation of the *c*- and *a*-axes of ice is shown on the right. **b**, Single ice crystal grown in the presence of sbwAFP ( $0.2\text{ mg ml}^{-1}$ ) by manipulation of the supercooling temperature and then maintained at  $-0.2^{\circ}\text{C}$ . **c**, The same crystal after 24 h. **d**, A tilted view of the same crystal. **e**, Hexagonal bipyramidal morphology produced by  $2\text{ mg ml}^{-1}$  of fish type I AFP from the winter flounder. In **b–e**, the direction of the *c*-axis is indicated by arrows, which in **d** is normal to the plane of the figure.

interpretive sketches. The top panel shows three of the six equivalent prism plane etches. The middle panel is a view down the *c*-axis at the basal plane etch, with the tips of the six prism plane etches visible at the periphery. The bottom panel presents an oblique view of an etched hemisphere to show the distinct separation between prism and basal plane etches. Basal plane etching by other AFPs has not been observed before, and we cannot rule out the possibility that protein accumulation on this surface is due to prism plane binding at the risers of steps on the basal plane.

Ice morphology studies at higher sbwAFP concentrations ( $0.2\text{ mg ml}^{-1}$ ) and low supercooling ( $-0.20^{\circ}\text{C}$ ) also show simultaneous prism-face and basal-plane inhibition (Fig. 3b–d). As crystal habit between the freezing and melting points (thermal hysteresis gap) is the result of growth-inhibited faces, AFPs binding simultaneously to these surfaces typically produce the hexagonal ice crystals seen during thermal hysteresis measurements (Fig. 1, top panels). By teasing some growth out of these crystals before returning them to the thermal hysteresis gap, it is possible to see the growth-inhibited surfaces more clearly. The single crystal viewed perpendicular to the *c*-axis (Fig. 3b) demonstrates that the prism faces are at right angles to the flat basal planes. This crystal showed no growth or shrinkage within the thermal hysteresis gap for at least 24 h (Fig. 3c). The expression and stability of such extensive basal planes is unique to sbwAFP and has not been observed with any of the fish AFPs. Indeed, fish AFPs typically bind to pyramidal or prism planes to form ice crystals with hexagonal bipyramidal or trapezohedral morphology (Fig. 3e)<sup>10,11</sup>.

Experimental evidence and theories from the fish AFP literature suggest that a specific ice-lattice match is required for AFP binding to ice<sup>12,13</sup>. The nature of the binding force is not known, but is assumed to involve both hydrogen bonds and non-polar forces<sup>8,14–17</sup>. Combining NMR structural information and experimental evidence for prism and basal plane binding, we have developed a model to demonstrate the specificity of sbwAFP for these planes (Fig. 4). On the ice-binding site of sbwAFP, the loops are a series of structural repeats, similar to those found in other left-handed  $\beta$ -helix proteins<sup>4,5</sup>, and bring the TXT motifs into alignment so that the threonines form two parallel arrays. In the model, the spacing between threonines in the neighbouring TXT motifs makes an ideal match to the  $4.5\text{ \AA}$  repeats along the *a*-axis in both the prism and basal planes of ice. At right angles to this repeat, the distance between threonines in the TXT motifs makes an ideal match to the  $7.35\text{ \AA}$  spacing on the prism plane, and a fairly close match to the  $7.8\text{ \AA}$  spacing on the basal plane.



**Figure 4** SbwAFP model showing surface complementarity with the prism and basal planes of ice. **a**, A side view of the protein aligned above an ice prism plane. Circles depict the lattice positions of water oxygen atoms in ice. **b**, A view perpendicular to that of **a**. This view down the *c*-axis illustrates the match of adjacent loops to the ice surface. As oxygen

atoms have the same interatomic distance along the *a*-axis in both basal and prism planes, the match to the protein is the same for both planes. **c**, The sbwAFP aligned above the basal plane, depicting the analogous surface match to that of the prism plane.

This simple model demonstrates the geometric feasibility of an ice lattice match between the protein and the primary prism and basal planes of ice without requiring knowledge of the forces involved in the interaction. It is consistent with the mutagenesis, ice-etching and ice-morphology data and shows how the ice-binding site threonines in the aligned TXT motifs could form nine contacts with ice, considerably more than the four contacts modelled between winter flounder AFP and ice<sup>18</sup>. The greater number of contacts and the ability to bind both prism and basal planes of ice may explain the greater activity of this insect AFP over fish AFPs.

The sbwAFP and beetle AFP<sup>6</sup> structures are the first insect AFP structures determined, and the  $\beta$ -helix fold represents a new AFP structural motif. Several different folds have been found for the AFP proteins, suggesting that nature has evolved different structures to form similar ice-binding surfaces. Given the controversy surrounding the nature of the forces involved between fish AFPs and ice<sup>8,14–17</sup>, the  $\beta$ -helical insect AFPs provide an independent model with which to test binding hypotheses. □

## Methods

### NMR data collection

Recombinant <sup>15</sup>N- and <sup>13</sup>C/<sup>15</sup>N-labelled sbwAFP was produced using BioExpress media and purified as described<sup>19</sup>. The protein was dissolved at a concentration of 18 mg ml<sup>-1</sup> in 90% H<sub>2</sub>O/10% D<sub>2</sub>O or 100% D<sub>2</sub>O as necessary, and the pH was adjusted to 5.5 using NaOD or DCl as required. CBCA(CO)NNH, HNCACB, <sup>13</sup>C/<sup>15</sup>N-edited NOESY, <sup>15</sup>N-edited nuclear Overhauser effect enhancement spectroscopy (NOESY), <sup>15</sup>N-edited total correlation spectroscopy (TOCSY), HCCH-TOCSY, HNHA and 2D-NOESY spectra were collected on a Varian 500 or 600 MHz spectrometer with the temperature control set to 30 °C. Data were processed and analysed using NMRPIPE<sup>20</sup> and PIPP, and the ensemble of structures was generated in X-PLOR<sup>21</sup> using distance and dihedral restraints from NOESY and HNHA experiments and empirical restraints from <sup>13</sup>C-chemical shifts<sup>22</sup>. Analysis of the ensemble was performed using VADAR<sup>23</sup> and PROCHECK\_NMR<sup>24</sup>.

### Ice crystal morphology

Ice hemispheres, initially grown from a single ice crystal, were allowed further growth in the presence of 0.02 mg ml<sup>-1</sup> of sbwAFP and permitted to etch as described<sup>10</sup>. Ice crystal morphology in the presence of 0.2 mg ml<sup>-1</sup> sbwAFP was monitored at -0.2 °C over a period of five days.

### Figure preparation

Figures were generated using MOLSCRIPT<sup>25</sup>, RASTER3D<sup>26</sup> and SYBYL<sup>27</sup>.

Received 22 March; accepted 15 May 2000.

1. Tyshenko, M. G., Doucet, D., Davies, P. L. & Walker, V. K. The antifreeze potential of the spruce budworm thermal hysteresis protein. *Nature Biotechnol.* **15**, 887–890 (1997).
2. Graham, L. A., Liou, Y. C., Walker, V. K. & Davies, P. L. Hyperactive antifreeze protein from beetles. *Nature* **388**, 727–728 (1997).
3. Duman, J. G. *et al.* Molecular characterization and sequencing of antifreeze proteins from larvae of the beetle *Dendroides canadensis*. *J. Comp. Physiol. B* **168**, 225–232 (1998).
4. Beaman, T. W., Sugantino, M. & Roderick, S. L. Structure of the hexapeptide xenobiotic acetyltransferase from *Pseudomonas aeruginosa*. *Biochemistry* **37**, 6689–6696 (1998).
5. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G. & Rees, D. C. A left-hand beta-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon *Methanosarcina thermophila*. *Embo J.* **15**, 2323–2330 (1996).

6. Liou, Y. -C., Tocilj, A., Davies, P. L. & Jia, Z. Mimicry of ice structure by surface hydroxyls and water of a  $\beta$ -helix antifreeze protein. *Nature* **406**, 322–324 (2000).
7. Sicheri, F. & Yang, D. S. Ice-binding structure and mechanism of an antifreeze protein from winter flounder. *Nature* **375**, 427–431 (1995).
8. Sönnichsen, F. D., DeLuca, C. I., Davies, P. L. & Sykes, B. D. Refined solution structure of type III antifreeze protein: hydrophobic groups may be involved in the energetics of the protein-ice interaction. *Structure* **4**, 1325–1337 (1996).
9. Jia, Z., DeLuca, C. I., Chao, H. & Davies, P. L. Structural basis for the binding of a globular antifreeze protein to ice. *Nature* **384**, 285–288 (1996).
10. Knight, C. A., Cheng, C. C. & DeVries, A. L. Adsorption of alpha-helical antifreeze peptides on specific ice crystal surface planes. *Biophys. J.* **59**, 409–418 (1991).
11. Deng, G., Andrews, D. W. & Laursen, R. A. Amino acid sequence of a new type of antifreeze protein, from the longhorn sculpin *Myoxocephalus octodecimpinosus*. *FEBS Lett.* **402**, 17–20 (1997).
12. Wen, D. & Laursen, R. A. A model for binding of an antifreeze polypeptide to ice. *Biophys. J.* **63**, 1659–1662 (1992).
13. Knight, C. A., Driggers, E. & DeVries, A. L. Adsorption to ice of fish antifreeze glycopeptides 7 and 8. *Biophys. J.* **64**, 252–259 (1993).
14. Haymet, A. D., Ward, L. G., Harding, M. M. & Knight, C. A. Valine substituted winter flounder 'antifreeze': preservation of ice growth hysteresis. *FEBS Lett.* **430**, 301–306 (1998).
15. Zhang, W. & Laursen, R. A. Structure-function relationships in a type I antifreeze polypeptide—The role of threonine methyl and hydroxyl groups in antifreeze activity. *J. Biol. Chem.* **273**, 34806–34812 (1998).
16. Chao, H. M. *et al.* A diminished role for hydrogen bonds in antifreeze protein binding to ice. *Biochemistry* **36**, 14652–14660 (1997).
17. Graether, S. P. *et al.* Quantitative and qualitative analysis of type III antifreeze protein structure and function. *J. Biol. Chem.* **274**, 11842–11847 (1999).
18. Chou, K. C. Energy-optimized structure of antifreeze protein and its binding mechanism. *J. Mol. Biol.* **223**, 509–517 (1992).
19. Gauthier, S. Y., Kay, C. M., Sykes, B. D., Walker, V. K. & Davies, P. L. Disulfide bond mapping and structural characterization of spruce budworm antifreeze protein. *Eur. J. Biochem.* **258**, 445–453 (1998).
20. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX Pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
21. Brünger, A. T. *X-PLOR Manual Version 3.1. A system for X-ray crystallography and NMR*. (Yale Univ. Press, New Haven, Connecticut, 1992).
22. Kuszewski, J., Qin, J., Gronenborn, A. M. & Clore, G. M. The impact of direct refinement against <sup>13</sup>C alpha and <sup>13</sup>C beta chemical shifts on protein structure determination by NMR. *J. Magn. Reson. B* **106**, 92–96 (1995).
23. Willard, L., Wishart, D. S. & Sykes, B. D. *VADAR Version 1.2* (Univ. Alberta, Edmonton, Alberta, Canada, 1997).
24. Laskowski, R. A., MacArthur, M. W., Moss, D. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
25. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950 (1991).
26. Merrit, E. A. & Murphy, M. E. P. A program for photorealistic molecular graphics. *Raster3D Version 2.0. Acta Cryst. D* **50**, 869–873 (1994).
27. SYBYL Molecular modelling package. Version 6. 5 (Tripos Software, St. Louis, Missouri, 1998).

## Acknowledgements

We thank S. Gauthier and D. Doucet for help with the mutagenesis, and L. Saltibus and G. McQuaid for their excellent technical assistance. Special thanks to L. Spyropoulos for helping in the analysis of the NMR data and structures. We thank Dan Garrett and the Laboratory of Chemical Physics at the National Institutes of Health for making available the program PIPP that was used in analysing our NMR data. This work was supported by MRC grants to P.L.D., B.D.S. and Z.J. and an NSERC grant to V.K.W.; Z.J. is an MRC Scholar and P.L.D. is a Killam Research Fellow.

Correspondence and requests for materials should be addressed to P. L. D. (e-mail: daviesp@post.queensu.ca). The atomic coordinates have been deposited in the Brookhaven Protein Database under accession number 1EWW.