Identification of the Ice-Binding Surface on a Type III Antifreeze Protein with a "Flatness Function" Algorithm

Daniel S. C. Yang,*# Wai-Ching Hon,* Steve Bubanko,* Yiqi Xue,* J. Seetharaman,* Choy L. Hew,§ and Frank Sicheri*

*Department of Biochemistry, Faculty of Health Science, McMaster University, Hamilton, Ontario L8N 3Z5, Canada; #BioCrystallography Laboratory, VA Medical Centre, Pittsburgh, Pennsylvania 15240 USA; and §Research Institute, Hospital for Sick Children, Toronto, and Departments of Clinical Biochemistry and Biochemistry, University of Toronto, Toronto, Ontario M5G 1L5, Canada

ABSTRACT Antifreeze proteins (AFPs) adsorb to surfaces of growing ice crystals, thereby arresting their growth. The prevailing hypothesis explains the nature of adsorption in terms of a match between the hydrophilic side chains on the AFP's ice-binding surface (IBS) and the water molecules on the ice surface. The number and spatial arrangement of hydrogen bonds thus formed have been proposed to account, respectively, for the binding affinity and specificity. The crystal structure of a type III AFP from ocean pout (isoform HPLC-3) has been determined to 2.0-Å resolution. The structure reveals an internal dyad motif formed by two 19-residue, loop-shaped elements. Based on of the flatness observed on the type I α-helical AFP's IBS, an automated algorithm was developed to analyze the surface planarity of the globular type III AFP and was used to identify the IBS on this protein. The surface with the highest flatness score is formed by one loop of the dyad motif and is identical to the IBS deduced from earlier mutagenesis studies. Interestingly, 67% of this surface contains nonpolar solvent-accessible surface area. The success of our approach to identifying the IBS on an AFP, without considering the presence of polar side chains, indicates that flatness is the first approximation of an IBS. We further propose that the specificity of interactions between an IBS and a particular ice-crystallographic plane arises from surface complementarity.

INTRODUCTION

Antifreeze proteins were first discovered in the blood sera of fish that inhabit icy polar waters (DeVries et al., 1970; Davies and Hew, 1990; DeVries, 1986; Yang, 1993). Proteins with similar properties have also been found in overwintering arthropods (Duman, 1982), bacteria, fungi (Duman and Olsen, 1993), and plants (Griffith et al., 1997; Urrutia et al., 1992), which must survive subzero temperatures. AFPs exhibit three general characteristics in vitro. First, they adsorb to the nonbasal planes of ice such that ice crystals grown in solution containing AFPs express hexagonal habits (Raymond et al., 1989; Yang, 1993). The adsorption results in inhibition of ice growth. Second, they depress the freezing temperature of solution in a noncolligative manner, with the melting temperature staying close to the equilibrium freezing point (Davies and Hew, 1990; DeVries, 1983). The difference between the freezing and melting temperatures is termed "thermal hysteresis," which currently offers the only quantitative measurement of antifreeze activity. Third, AFPs inhibit the recrystallization of ice (Knight et al., 1984). It is remarkable that these properties are displayed to a similar extent by all AFPs despite their widely different protein sizes, amino acid sequences, and tertiary structures (Jia et al., 1996; Sicheri and Yang, 1995; Sonnichsen et al., 1996). This indicates the presence of a common ice-binding mechanism that is not dictated by primary or secondary protein structures.

The ice-binding mechanism was first modeled on the alanine-rich, winter flounder AFP. In the absence of atomic details of side-chain conformation in the first 2.5-Å crystal structure (Yang et al., 1988), the lone amphiphilic helix presented itself as a simple and attractive model for the elucidation of the mechanism of adsorption. The presence of hydrophilic residues (Thr, Asn, and Asp) on one face of the helix initiated the concept of IBS being present in all AFPs. The later 1.5-Å resolution crystal structure revealed the flat and rigid nature of this IBS (Sicheri and Yang, 1995). This observation disputed not only the hypothesis that the hydrophilic side chains possessed torsional freedom to facilitate hydrogen bonding with the ice surface (Yang et al., 1988), but also the ice-binding model in which the polar groups of these side chains extend into the ice lattice, adopting water positions (Knight et al., 1993). The presence of a flat and rigid IBS was therefore proposed to be a general feature of AFPs (Sicheri and Yang, 1995).

The fish AFPs show specificities in their binding to ice-crystallographic planes, as demonstrated by Knight's unique ice-etching studies (Knight et al., 1991). For instance, within the class of alanine-rich, α-helical type I AFPs, those from winter flounder and Alaskan plaice showed preference for the pyramidal planes of ice $\{20\bar{2}1\}$, whereas sculpin AFPs adsorb to the secondary prism planes $\{2\bar{1}\bar{1}0\}$. Because of the 11-residue repeats, each ending in a hydrophilic residue, these AFPs were suggested to align along the $\langle 01\bar{1}2\rangle$ axis on these ice planes. This was based on the close match between the repeat spacing and the water spacing along the axis $\langle 01\bar{1}2\rangle$. This finding provided an

Received for publication 18 August 1997 and in final form 12 November 1997.

Address reprint requests to Dr. Daniel S. C. Yang, Department of Biochemistry, Faculty of Health Science, McMaster University, Hamilton, Ontario L8N 3Z5, Canada. Tel: 905-525-9140, ext. 22455; Fax: 905-522-9033; E-mail: yang@xtliris.csu.mcmaster.ca.

0006-3495/98/05/2142/10 $2.00

ice-binding mechanism for AFPs, namely, a structural match through hydrogen bonding with the ice lattice. The importance of spatial arrangement of polar groups on the IBS for ice binding has largely been upheld. Nevertheless, this model cannot explain satisfactorily the specific binding of individual AFP to a particular ice plane.

To investigate the generality of our previous hypothesis that a flat and rigid IBS is a critical feature of all AFPs (Sicheri and Yang, 1995), we have determined the crystal structure of a type III AFP from ocean pout (isoform HPLC-3). This structure was derived independently of the recently published crystal structure (Jia et al., 1996) and the refined, high-precision NMR structure of the HPLC-12 isoform (Sonnichsen et al., 1996). The two isoforms shared 62% of amino acid sequence identity and have very similar three-dimensional structures. Before publication of these two groups' findings, we had developed an algorithm to compute a "flatness function" to describe the complete surface of a protein, with the aim of identifying the IBS objectively and analytically. The surface that ranked the highest in our flatness function is identical to the consensus IBS identified by these two studies, based on the results from mutagenesis work targeted to conserved polar residues. In this report, we will present a refined ice-binding model of AFPs, combining structural and mutagenesis evidence from the type III AFPs.

MATERIALS AND METHODS

Cloning, protein purification, and crystallisation

The native protein was purified from fish serum and crystallized as previously described (Xue et al., 1994). The gene coding for the isoform HPLC-6 was subcloned from the original λPCIIL-rHPLC6 vector (Li et al., 1991) into the pET15b vector (Novagen, Madison, WI). The resulting construct has the hexahistidine tag at the N-terminus. The His-rHPLC6 was overexpressed in *Escherichia coli*, and purified by the following steps. The lysate was first passed through a DE52 anion-exchange column (Whatman, Clifton, NJ), and the majority of the contaminating proteins in the eluent were removed by heat denaturation (60°C) and centrifugation. The His-tagged protein was then isolated on a Ni-affinity column and proteolyzed on the column with thrombin to remove the His-tag, and rHPLC-6 was eluted. This final product has the extra amino acids GSHMK, and K, at the N- and C-termini of the wild-type HPLC-6, respectively. Both His-rHPLC6 and rHPLC6 proteins expressed full antifreeze activity. Large diffraction-quality crystals of dimensions $1.2 \times 0.57 \times 0.15$ mm could be grown at room temperature over 3 days by the hanging drop vapor diffusion method. The crystallization reagents consisted of 68 mM Tris-HCl (pH 8.0) and 2.72 M ammonium sulfate, with an initial protein concentration of 8 mg ml^{-1}.

Preparation and isolation of monoiodinated rHPLC-6

Initial attempts to prepare heavy atom derivatives by soaking technique failed. We therefore produced a single derivative by iodinating the only tyrosine residue (Tyr^{63}) in rHPLC-6 (Means and Feeney, 1971). The chemical reaction was carried out by incubating 16 mg ml^{-1} of protein in 0.1 M Tris-HCl (pH 7.1), 0.075 mM I_2, and 0.15 mM KI for 2 h at room temperature. The reaction products contained unreacted, singly and doubly iodinated species, which could be readily resolved on a C_{18}, reversed-phase high-performance liquid chromatography (HPLC) column, using a 33–36% gradient of acetonitrile with 0.1% trifluoroacetic acid. The identities of the eluted peaks were confirmed by MALDI mass spectrometric analysis (McMaster Regional Centre for Mass Spectrometry, McMaster University). The peak corresponding to the singly iodinated species was used for crystallization and structure determination.

Data collection and processing

Diffraction experiments were performed on a Raxis IIC area detector, using Cu $K\alpha$ radiation generated by a Rigaku RU200 rotating anode generator operating at 3 kW (60 mA × 50 kV), with a 0.2-mm cathode and focused by Supper double focusing mirrors. The native data set was collected at 4°C and processed with the Raxis processing software. With the aim of solving the structure of the iodinated rHPLC-6 using the Iterative Single-wavelength Anomalous Scattering (ISAS) technique (Wang, 1985), the data collection strategy was optimized to produce accurate anomalous signals. The crystal, belonging to the space group $P2_12_12_1$, was aligned with its *b* axis perpendicular to the x-ray beam such that Bijvoet's reflections could be collected on the same frame. To maximize the signal-to-noise ratio, fine-slicing data collection was practiced. The crystal was exposed for 15 min/0.25° oscillation for a total 90° sweep, and at −165°C in the absence of cryoprotectant. The missing reflections in the blind zone were not used in the determination of either the heavy atom substructure or the protein structure, because of the high value of merging statistics. The diffraction data were indexed and measured using MADNES (Messerschmidt and Pflugrath, 1987), and then scaled with the Bijvoet's reflections separated using SCALEPACK (Otwinowski, 1993).

Protein phasing, model building, and refinement

Positions for two iodine atoms were initially identified from the anomalous difference Patterson function using HASSP (Terwilliger et al., 1987). Four additional sites were found in the subsequent difference Fourier map. Using these six sites, the ISAS procedure as implemented in the PHASES package (Furey and Swaminathan, 1997) was carried out, using anomalous data up to 3.5-Å resolution. The phases obtained were extended to 3.0 Å, and the electron density map thus generated clearly showed the boundary of the four protomers in the asymmetric unit. This map was used to identify the noncrystallographic symmetry elements. Subsequent fourfold noncrystallographic symmetry map averaging produced an excellent quality map, from which the model of the entire protein could be built using O (Jones et al., 1991).

The model was refined using the positional powell minimization and simulated annealing procedures as implemented in X-PLOR (Brunger, 1993). On examining the residue-omitted electron density map, we found multiple conformations for the iodo-Tyr^{60} side chain. This explained the presence of six heavy atom sites, two more than the expected four.

The model built from the iodo-rHPLC-6 isoform was used to determine the 2.0-Å crystal structure of the native HPLC-3 isoform using the technique of molecular replacement. Successive cycles of simulated annealing refinement and manual model rebuilding using residue-omitted maps were carried out. Using a bulk solvent model and all reflection data, the model was refined to an R-factor of 21.7% and R-free of 29.1%. Solvent molecules were then located using two cycles of water selection and refinement. Residual peak heights, hydrogen-bonding distances, steric contacts, individual refined B-factors, and overall R-factors were taken into account during the selection process. The automatically selected water molecules were inspected on a 2Fo-Fc sigma a-weighted map to ensure good electron density peaks. A total of 33 water molecules were incorporated into the final model, with a final R-factor of 20.0% and R-free of 25.9%.

Computation of the flatness function

We define "flatness" as a local surface property of a protein molecule. The flatness at any point on a protein surface is estimated by counting the

number of contacts between the molecule and a planar surface. It is directly related to the solvent-accessible surface area buried when a planar surface is brought into tangential contact with the protein at that particular point, and indirectly related to the root mean square deviations in distance between all of the plane-contacting atoms and the plane. To avoid over-reliance on the contribution from any single residue, an additional parameter considered is the number of atoms and residues that are involved in forming the surface.

To describe the molecular surface uniquely, the protein is first aligned with its three gyration axes coinciding with the Cartesian axes X, Y, and Z. Any point on the protein surface is then denoted by the spherical polar angles, ϕ (rotation angle about the z axis, 0–180°) and ψ (inclination angle from the z axis, 0–360°). Rotation operations on the x-y plane about the Z and X axes with angles ϕ and ψ, respectively, will bring it parallel to the protein surface at the point ϕ, ψ. After the rotation operations, the plane is translated away from the protein centroid toward the surface, until the outermost atom resides on the plane. Fig. 1 presents the concept of such a representation of a general protein molecule. In this study, the flatness function for the AFP, $F(\phi, \psi)$, was calculated for a total of 64,800 (180×360) points on the protein surface, sampling through every degree on ϕ and ψ. A crucial parameter in the flatness function is the cutoff distance used for identifying atoms that are in contact with the rotated plane. To arrive at a noncommittal estimate, two cut-offs were used. A stringent distance of 0.5 Å and a distance approximating the solvent accessibility probe radius of 1.6 Å are used. Therefore, two slabs of atoms within 0.5 and 1.6 Å of the plane were selected. For each slab of atoms, a number of parameters were evaluated. These included $F_{\text{NUM-A}}$, the number of atoms; $F_{\text{NUM-R}}$, the number of residues; $F_{\text{RMS-A}}$, a reciprocal function of root mean square deviations of atoms from the least-squares plane; and $F_{\text{SUF-A}}$, the sum of atomic solvent-accessible surface area. The values for each of these parameters were normalized with respect to the standard deviations among the 64,800 grid points. The flatness function $F(\phi, \varphi)$ was then computed by taking the dot product of the different functions, i.e., $F(\phi, \psi) = F(\phi, \psi)_{\text{NUM-A}} * F(\phi, \psi)_{\text{NUM-R}} * F(\phi, \psi)_{\text{RMS-A}} * F(\phi, \psi)_{\text{SUF-A}}$. A program was used to identify peaks that have flatness scores higher than three standard deviations.

RESULTS AND DISCUSSION

Quality of model

The 2.0-Å protein model of the ocean pout AFP (isoform HPLC-3) was determined by a molecular replacement method, based on our 3.0-Å crystal structure of an iodinated recombinant pout AFP (isoform HPLC-6), which was phased using the technique of ISAS (Iterative Single-wavelength Anomalous Scattering) (Wang, 1985). In sigma a-weighted 2Fo-Fc electron density maps (Read, 1994), densities of the main chain and side chains could be clearly seen for all 63 residues, except for the side chain of Thr^{47}. The main chain has excellent geometry, as 96.5% of the residues were placed in the most favored region in the Ramachandran plot (Laskowski et al., 1993). Tyr^{63} was placed in the additionally allowed region. The model consists of one *cis*-peptide bond, formed between Asn^{28} and Pro^{29}. Other refinement statistics are presented in Materials and Methods.

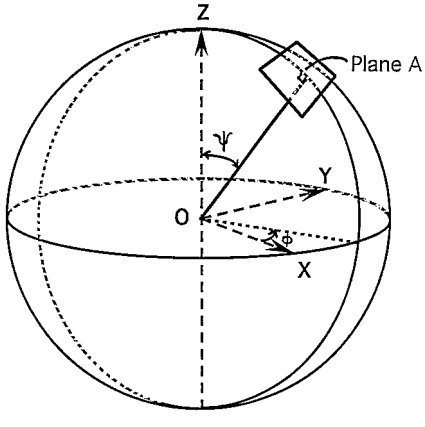

FIGURE 1 Schematic diagram illustrating the coordinate system for the flatness function $F(\psi, \phi)$. The protein molecule is treated as a sphere. Each point on the protein surface is represented by the spherical polar angles ψ and ϕ. ψ and ϕ are the inclination angle to the z axis and the rotational angle about the z axis, respectively. Plane A denotes an imaginary plane brought into tangential contact at point (ψ, ϕ).

Overall description of protein structure

The overall structure of isoform HPLC-3 is very similar to those of isoform HPLC-12 (Jia et al., 1996; Sonnichsen et al., 1996). It has a very compact fold with very few secondary structural elements. It basically contains seven antiparallel β-strands, three 3_{10}-helices, eight β-turns, and numerous interconnecting coils (Fig. 4). (Secondary structure assignment was performed by DSSP as implemented in RASMOL (Sayle and Millner-White, 1995).)

We observed a previously unnoticed internal twofold symmetry motif, formed by two antiparallel, 19-residue loop-shaped elements (Val^5 to Glu^{23} and Lys^{43} to Lys^{61}) (Fig. 2). When superimposed, the two loops show a main-chain RMS deviation of only 0.59 Å. Each loop contains the β-turn-β-turn-β-3_{10}-helix-β structure. The two loops are held together in a back-to-back manner by nine main-chain hydrogen bonds, six of which form three pairs of antiparallel, one-residue, pseudo-β-strands (Fig. 3). This structural symmetry, however, is scarcely reflected by the sequence, as the two loops share only 33% sequence identity (Fig. 4 *a*). Loop 1 also has the highest degree of local sequence identity among the pout AFP isoforms. As will be discussed later, the surface formed by loop 1 is the putative ice-binding surface, whereas the ice accessibility of the surface formed by loop 2 is blocked by the flanking residues 24–32 (loop 3) (Fig. 5). This loop-loop dimerization is likely to be important in providing a rigid support for the putative IBS. To our knowledge, this motif has not been described in the literature. We name it the "pretzel fold," based on its topographic appearance (Fig. 2 *a*).

Identification of the putative ice-binding surface

Unlike most protein-ligand interactions, the definitive identification of the ligand-binding site on an AFP through structural means is difficult, as it is impossible to produce single crystals of protein-ice complex. The unique sequence repeats and the α-helical structure of the winter flounder type I AFP provide obvious clues to the site of the IBS. By visual inspection, the ice-binding residues on the winter flounder AFP form a flat plane, with little protrusion of side

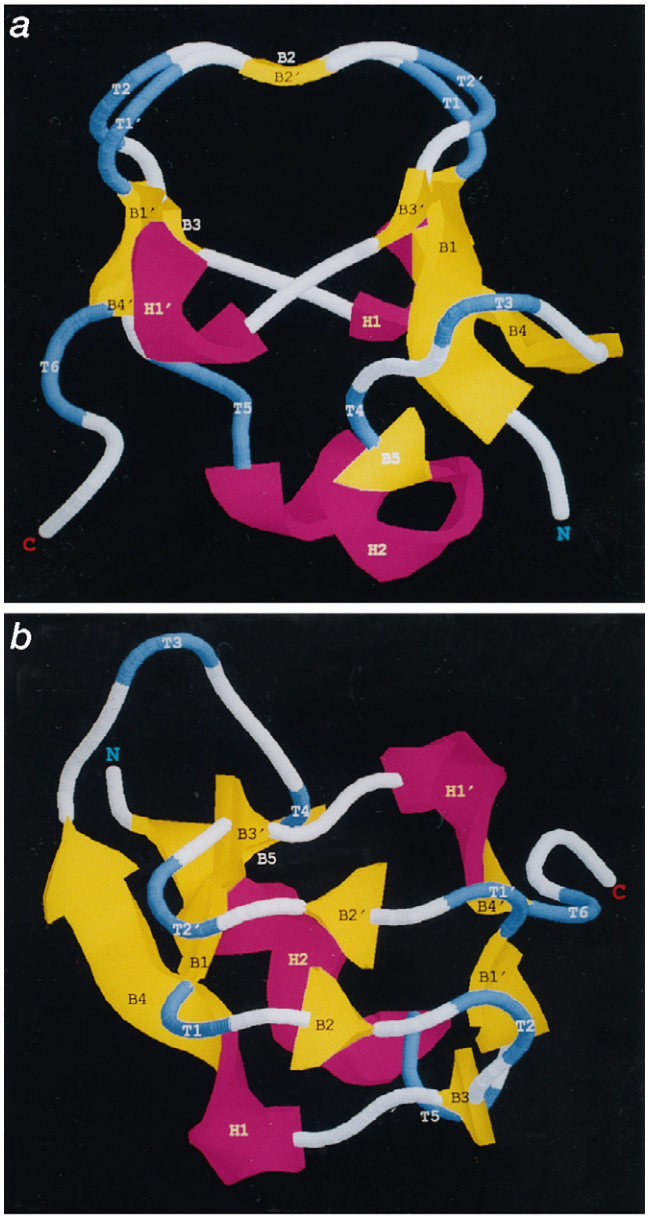

FIGURE 2 Ribbon representation of the pout AFP molecule. Secondary structure was assigned by the program DSSP (Kabsch and Sander, 1983) as implemented in RASMOL (Sayle and Millner-White, 1995). The color scheme is as follows: magenta, helix; yellow, β-strand; blue, turn; white, coil. A number of single-residue β-strands are included to reflect the internal twofold symmetry of the "pretzel" fold. The putative ice-binding surface is formed mainly by loop 1, which consists of T1, B2, T2, B3, and H1. The equivalent secondary structural elements on loop 2 are labeled with primed numbers. Views perpendicular (*a*) and parallel (*b*) to the twofold axis of the dimeric pretzel fold are shown.

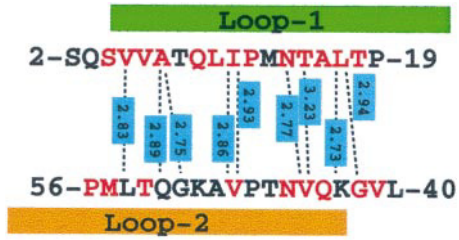

FIGURE 3 Main-chain hydrogen bonding scheme between loop 1 and loop 2 of the pretzel fold. The color boxes correspond to the secondary structure assignment as shown in Fig. 2. The boxes marking the ranges of loop 1 and loop 2 are colored in the same manner as in Fig. 5. Residues in red are invariantly conserved among the different isoforms of type III AFPs (Davies and Hew, 1990). The numbering scheme of the residues conforms to that used for the eelpout AFP, isoform HPLC-12 (Chao et al., 1994). The numbers in the blue boxes represent the hydrogen-bonding distances in Å.

chain atoms from the IBS (Sicheri and Yang, 1995). Because the surface of ice crystal is macroscopically flat, by shape complementarity consideration, it is reasonable to assume that this flat feature of the IBS extends to all other AFPs, regardless of their overall three-dimensional structures. To locate this flat surface on the type III pout AFP in an objective and analytical manner, we developed an algorithm to compute a "flatness function" for all possible points on the entire protein surface. The flattest surface on the protein would have the highest flatness score, and would likely form the ice-binding surface. We first tested our reasoning on the crystal structure of the winter flounder AFP. The results ranked the previously identified IBS the first, with a flatness score of 106, which is 75 standard deviations higher than the surface that ranked the second. Details of the technique are described in Materials and Methods.

Results of the computation of the flatness function on our pout AFP structure are shown in Fig. 4 *b*. The surface that ranked the highest is formed by the largest number of invariantly conserved residues predicted to be in contact with the ice plane (Jia et al., 1996; Sonnichsen et al., 1996). Eight of these residues reside on loop 1 and one on loop 2. Met^{13}, which contributes a main-chain O atom, is the only residue not invariantly conserved. The flatness score of this surface is 155, and is 88 standard deviations from the second surface. In the analysis of our structure, we do not observe another flat surface that is orthogonal to this plane, and which has been conjectured to bind to the basal plane of ice crystal (Jia et al., 1996; Sonnichsen et al., 1996).

This surface is remarkably similar to the IBS recently reported for the isoform HPLC-12 (Jia et al., 1996; Sonnichsen et al., 1996). The derivation of the IBS by these groups was based on the results of extensive mutagenesis work targeted mainly to the conserved polar residues. It should be noted that using the same approach but with fewer mutagenesis data, Chao et al. (1994) predicted earlier that a different surface adjacent to this consensus IBS is responsible for antifreeze activity.

Structural features of the ice-binding surface

The IBS is formed by the largely coil part of loop 1, which encircles a large surface area (Fig. 2). This is distinct from the winter flounder's AFP, whose IBS is formed along the entire length of one side of a perfect α-helix. The composition of the putative ice-contacting residues on the type III AFP is more varied than that of the type I AFP, which is rich

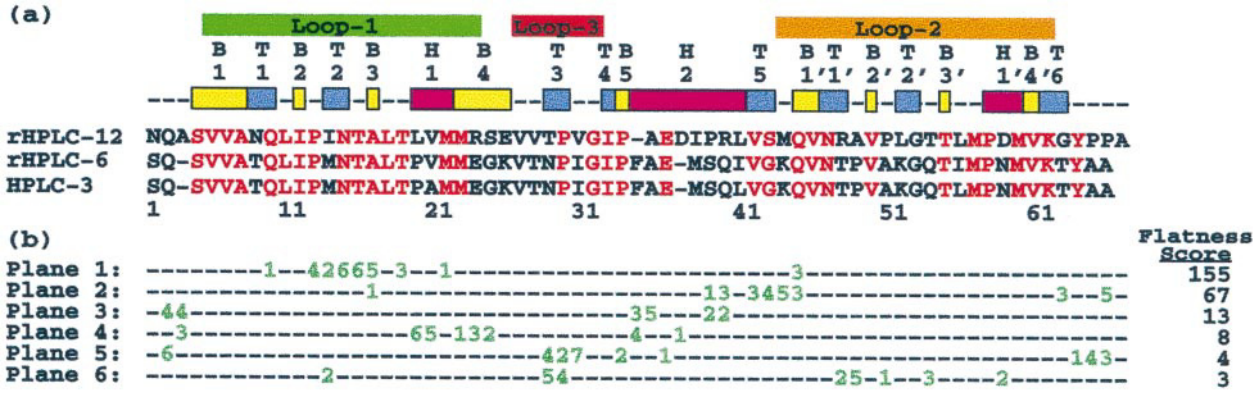

FIGURE 4 Amino acid sequence alignment of three isoforms of type III AFPs (*a*) and results of ice-binding surface search on the structure of pout AFP isoform HPLC-3 (*b*). (*a*) The color schemes are the same as those described in Fig. 3. Loop 1 and loop 2 are the two elements composing the dimeric pretzel fold, and loop 3 is the intervening structure that blocks the ice accessibility of loop 2. (*b*) The six protein surface planes that score the highest on the maximum-likelihood flatness function are presented. The residues composing each plane are marked by green numbers, which represent the number of plane-contacting atoms for the corresponding residues as listed in *a*. The flatness scores are in units of standard deviation of the entire flatness function.

in Leu, Asn, and Ala. Table 1 shows a detailed breakdown of all of the predicted ice-contacting atoms and their properties. There are a total of 27 predicted ice-contacting atoms, of which 82% are from side-chain atoms, and 59% are nonpolar. The average distances from an imaginary plane in tangential contact are 0.91 Å for all atoms, and 0.96 Å and 0.81 Å for polar and nonpolar atoms, respectively. The total solvent-accessible surface area was estimated to be 430 $Å^2$, 67% of which is hydrophobic. Taken together, polar atoms do not make a dominant contribution to the area of the IBS, nor are they in particularly advantageous positions to make contact with the ice plane.

To evaluate the uniqueness of the structural features of the putative IBS, we compare the two symmetry-related regions on loop 1 and loop 2 (residues 9–21 and 47–59). Although the β-turn-β-turn-β-3_{10}-helix-β motif provides a basic framework, it seems that the side-chain conformations are also important for the formation of a flat surface. With the exception of Gln^{44}, the Cα atom of which resides much further away from the least-square plane of the IBS, the longer, more flexible side chains (Gln^9, Asn^{14}, and Met^{21}) of the putative IBS invariably adopt rotamers with the long axes lying close to the least-square plane. This results in maximizing the number of nonpolar atoms that can contact ice, and hence the ice-binding surface area. This mode of side-chain packing is not observed for the symmetry-related region on loop 2 (Fig. 5 *b*).

A similar kind of "flatness-oriented" restraint of side-chain conformations has also been reported on the type I AFP structure (Sicheri and Yang, 1995). This phenomenon was dubbed "rigidity" and was distinct from the conventional connotation of the term, namely, low crystallographic temperature factors. In this structure, the average side-chain B value of the putative ice-binding residues is 19.2 $Å^2$, which is almost the same as that of the symmetry-related residues on loop 2 (19.0 $Å^2$). Side-chain rigidity was implicit in our calculations of the flatness function. The side-chain conformations as seen in the crystal structure were assumed to be the most stable and the most represented conformers in solution, and hence the most likely structure competent for ice adsorption.

Comparisons with reported mutagenesis studies

Our ability to locate the ice-binding sites on the protein surfaces of a type I and a type III AFP, without considering the hydrophilicity or the thermal factors of the side chains, leads us to conclude that "flatness" is the first approximation of an IBS on ice-binding proteins. Although this hypothesis has never been directly tested in published work, it is supported by existing mutagenesis data. These involve conversions of ice-binding residues to bulkier side chains, resulting in disruption of the flatness of the IBS. Depending on the degree of increased bulkiness, these mutants (A16T, A16H, and T18N) retain from 10 to 75% of wild-type antifreeze (Chao et al., 1994; Jia et al., 1996). In parallel, mutations of Ala^{17} to either Gln or Leu were reported to result in complete loss of antifreeze activity in the winter flounder AFP (Wen and Laursen, 1993). Ala^{17} forms part of the contiguous IBS that spans the entire length of the α-helical AFP, although it was not explicitly identified as one of the ice-binding residues (Sicheri and Yang, 1995).

There are two other types of mutants whose reduced activity was attributed to loss of hydrogen-bonding groups with ice (Jia et al., 1996). One type (Q9T, T15A, and T18A) retains 70–80% of wild-type activity, whereas the other type (Q44T, N14S, and N14S/Q44T) results in relatively drastic loss of activity (50%, 25%, and 10% wild-type activity). Among all of the reported single mutations of polar residues, N14S is the most potent in activity reduction. Hence Asn^{14} was described as the most important residue for ice binding (Jia et al., 1996). Our analysis indicates, however, that the latter type of mutation could be compli-

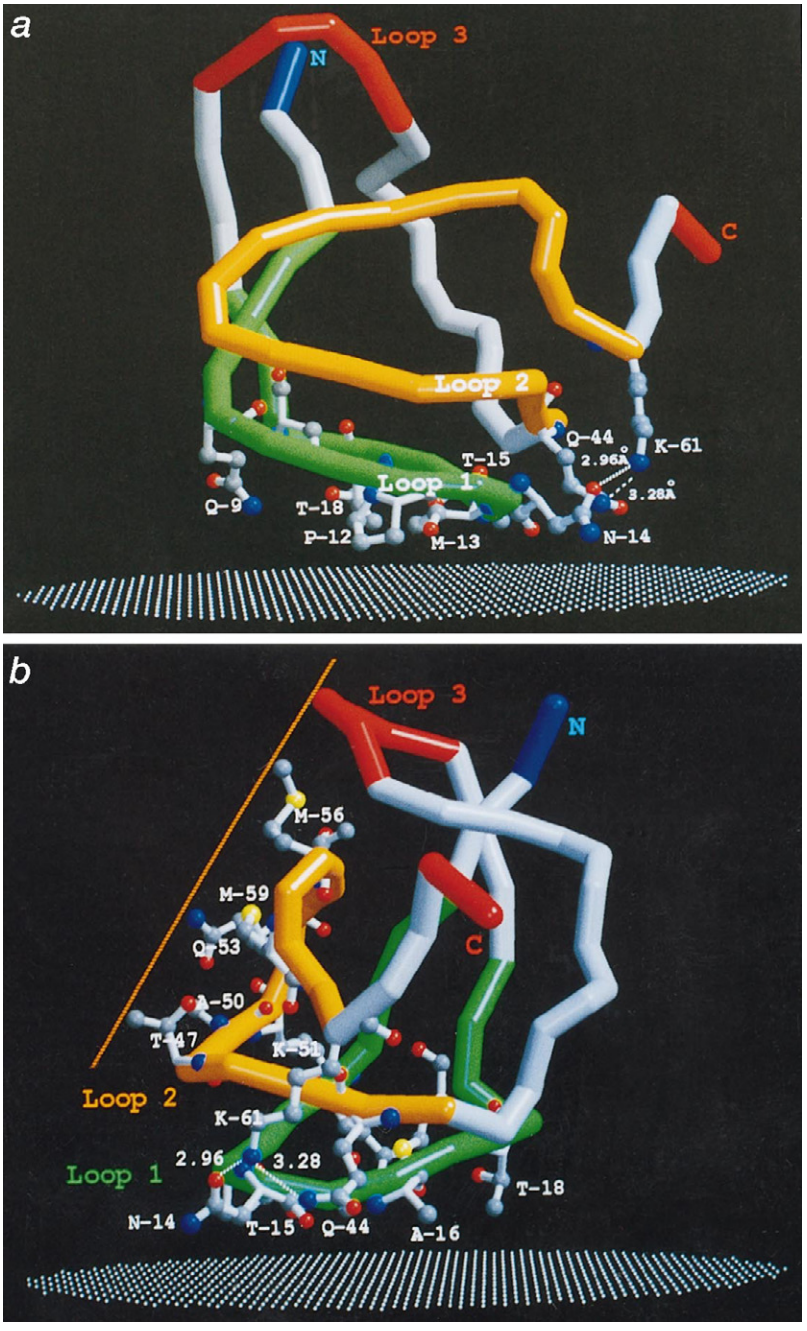

FIGURE 5 Backbone structure of the AFP molecule in a putative ice-binding mode. The ice surface is represented by a plane of white dots. (*a*) Side chains of residues contacting the ice plane are shown in ball-and-stick representation, depicting their nonprotruding nature. (*b*) Side view of the plane parallel to the surface formed by loop 2. Side chains of residues on loop 2 are shown in parallel to demonstrate their different propensities for rotamer conformations. (Figures were generated with MOLSCRIPT (Kraulis, 1991) and rendered with Raster3D (Merrit and Murphy, 1994).)

cated by structural instability. Asn^{14}, Gln^{44}, and Lys^{61} form a tripartite hydrogen-bonding network (Fig. 5). This hydrogen-bonding network seems to stabilize the protein fold by linking loop 1 (Asn^{14}), loop 2 (Gln^{44}), and the C-terminus (Lys^{61}) together. Lys^{61} is not one of the putative ice-contacting residues. We predict that Lys^{61} mutants incapable of maintaining the tripartite hydrogen-bonding network will also result in a significant reduction in antifreeze activity.

In summary, disregarding structural mutants, it is evident that single mutations that cause major activity reduction are due to significant steric disruption of the ice-binding surface. Ice binding is likely a collective activity contributed by all ice-contacting atoms, and it is improbable that any one residue would play a dominant role.

Implications for a general ice-binding mechanism

The flatness of the IBS observed on both the type I and type III AFP protein structures clearly excludes the possibility of extension of polar groups into the ice lattice to take up water positions. A couple of ice-binding models have been presented (Jia et al., 1996; Sonnichsen et al., 1996) that dock the flat IBSs to the ice planes that Knight identified by ice-etching experiments (Cheng and DeVries, 1991). These docking models opt for optimal structural matches through hydrogen bonding, with the polar atoms on the IBSs riding directly on the O atoms of the ice surface. Although the bonding distances are within an acceptable range, these hydrogen bonds were found to have less than perfect bond

TABLE 1 List of protein atoms within 3 Å of the ice plane in putative ice-binding mode

Nonpolar				Polar			
Residue	Atom type	Distance* (Å)	Surface area# (Å^2)	Residue	Atom type	Distance* (Å)	Surface area# (Å^2)
Pro^{12}	CA	1.76	2.46	Gln^{9}	NE2	1.68	26.50
	CB	0.24	33.26	Met^{13}	O	0.67	10.78
	CD	0.96	11.65	Asn^{14}	O	0.07	10.07
	CG	0.03	37.36		ND2	0.00	37.33
Asn^{14}	C	0.86	0.86		OD1	1.88	0.91
	CB	0.30	23.98	Thr^{15}	OG1	0.36	4.42
	CG	0.79	9.11	Ala^{16}	N	0.76	0.05
Thr^{15}	CA	0.78	12.11		O	1.35	1.18
	CB	1.01	0.67	Thr^{18}	OG1	1.90	1.04
	CG2	0.48	23.32	Gln^{44}	NE2	1.09	5.18
Ala^{16}	CB	0.31	45.74		OE1	0.76	42.66
Thr^{18}	CB	1.70	4.89				
	CG2	0.46	52.22				
Met^{21}	CE	1.41	19.16				
Gln^{44}	CD	1.19	12.04				
	CG	1.92	1.95				

*Distance of atoms from an imaginary plane tangent to the AFP molecule centred on atom ND2 of Asn^{14}.
#Solvent-accessible surface areas calculated by routines as implemented in X-PLOR (Brunger, 1993).

angles between the donors and acceptors (Sicheri and Yang, 1995; Sonnichsen et al., 1996).

Consider the process of ice crystal growth in a solution of AFP as a competition between water and AFP molecules for binding sites on the ice lattice. Because the majority of ice water molecules buried by the mainly hydrophobic IBS do not have hydrogen-bonding partners, the number of hydrogen bonds that can be formed is less when an AFP molecule

TABLE 2 Statistics of data collection, phase determination, and model refinement

	Iodo-rHPLC-6	HPLC-3
Data collection		
Space group	$P2_12_12_1$	$P2_1$
Cell dimensions		
(Å)	$a = 35.8, b = 75.0, c = 108.6$	$a = 23.1, b = 40.8, c = 30.0$
(°)	$\alpha = \beta = \gamma = 90$	$\alpha = \gamma = 90\ \beta = 100.6$
Number of molecules per asymmetric unit	4	1
Resolution limit (Å)	2.5	2.0
Number of unique reflections	8551	3124
Completeness (%) [overall/last shell]	79.8/68.3	82.4/54.9
R_{sym}* (%)	4.4	6.6
Phase determination (5–3.5 Å)		
No. of sites	6	—
Phasing power#	2.0	—
Mean figure of merit	0.447	—
R-factor§ (%)	27.0	—
Ranomalous¶		
Model refinement (40–2.0 Å)		
No. of protein atoms	—	648
No. of solvent atoms	—	33
R-factor (%)	—	20.0
R-factor§		
R-free (%)	—	25.9
RMSD in bond length (Å)	—	0.014
RMSD in bond angle (°)	—	1.7
Average B-factor (Å^2) [protein/water]	—	17.4/35.1

*$R_{sym} = \Sigma(\ |\ I_i - \langle I\rangle\ |\ )/\Sigma\langle I\rangle$ where I_i is the intensity of an individual reflection, and $\langle I\rangle$ is the mean intensity of that reflection.
#Phasing power = $\langle F_h\rangle/\langle E\rangle$, where $\langle F_h\rangle$ is the mean calculated heavy-atom structure factor amplitude, and $\langle E\rangle$ is the mean estimated lack of closure.
§R-factor = $\Sigma\|F_o\ |\ -\ |\ F_c\|/\Sigma\ |\ F_o\ |$ where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. Before crystallographic refinement, 10% of F_o was randomly selected and set aside for monitoring R-free.
§Ranomalous = $\Sigma\|\Delta F_o\ |\ -\ |\ \Delta F_c\|/\Sigma\ |\ \Delta F_o\ |$ where ΔF_o and ΔF_c are the observed and calculated structure factor amplitudes differences between Friedel pairs, respectively.

is bound to an ice plane. Furthermore, perfect hydrogen bonds are formed when water molecules are incorporated into the ice lattice as ice crystals grow, but not when AFPs bind to ice as mentioned above. Hence it seems that hydrogen bonding cannot account satisfactorily for AFP's high-affinity binding of ice.

To date, all docking studies have focused on the two-dimensional arrangement of water molecules on an ice plane. The different crystallographic ice planes available for AFP binding have diverse atomic structural features. Fig. 6 shows the three-dimensional views of a few of these ice planes, revealing the surface textures. We propose an alternative mode of structural match between AFP and ice that is driven by van der Waals surface complementarity. Each IBS, albeit macroscopically flat, has a unique atomic contour. The specific binding of an AFP could be due to positional matches between the relatively protruding atoms on the IBS and the grooves and holes on a particular ice surface, so that the optimal protein-ice surface contact could be formed. The dimensions of these grooves and holes are large enough to accommodate the relatively protruding ice-contacting atoms so that there is no need for the displacement of water molecules from the ice planes. The polar atoms on the AFP would still be within acceptable hydrogen-bonding distance to ice, because of the high density of water molecules and their relatively short interatomic distances on the ice crystallographic planes (Fig. 6 *b*). We believe that this model provides a reasonable explanation for the specificity of ice binding. However, it is difficult to delineate the different kinds of contributions to binding affinity. Although the binding energy offered by hydrogen bonding is higher than that offered by van der Waals interaction, the less than optimal hydrogen bonds that can be formed between AFP and ice likely reduce their total contributions. Moreover, we have yet to determine the minimum flat surface area that is required for ice binding.

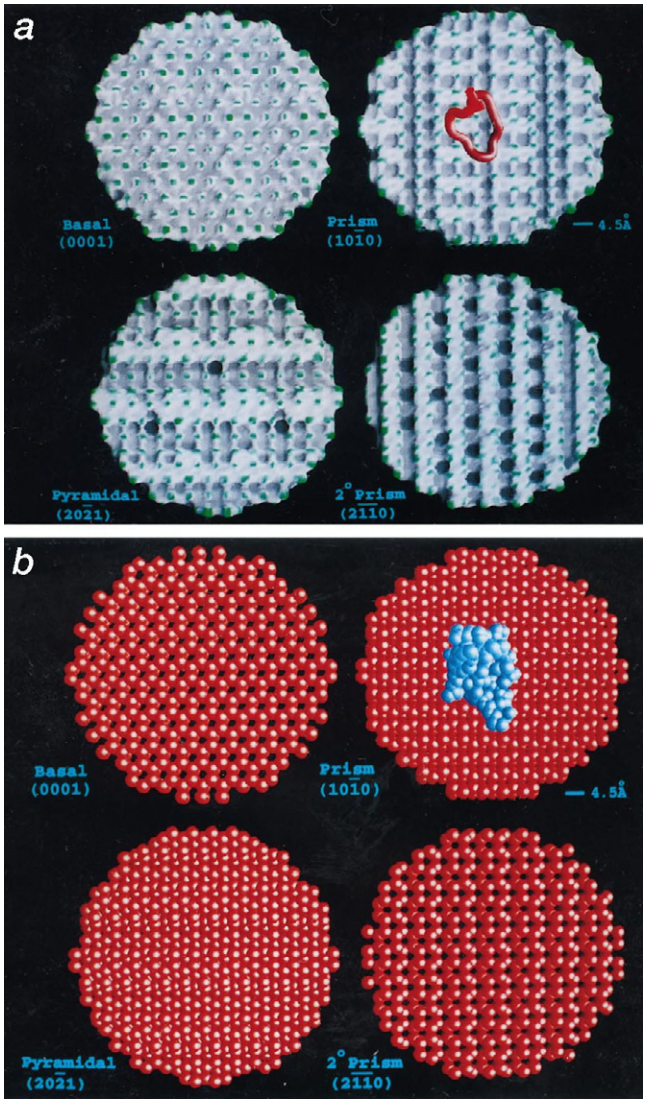

FIGURE 6 Diagrams of various crystallographic planes of ice I_h. (*a*) Molecular surfaces colored in terms of curvature. (*b*) Space filling representation of the oxygen atoms in the ice crystal demonstrating the hydrogen-bonding potential of the surface. The backbone (*a*) and the space-filling model (*b*) of loop 1 are placed on the prism face in an arbitrary orientation, with the putative ice-binding surface of the loop parallel to the ice-plane. (Figures were generated with GRASP (Nicholls et al., 1991).)

CONCLUSION

Antifreeze proteins are a unique class of macromolecules classified by their ability to bind to ice crystals and arrest their growth. They have been found in a vast variety of poikilothermic organisms that must survive subzero temperatures. The amino acid sequences of these proteins are diverse, and in a few cases have been found to belong to classes of proteins with other distinct functions (Chen et al., 1997; Ewart et al., 1992; Hon et al., 1995; Ng and Hew, 1992). These diverse protein structures, however, exhibit similar properties in inhibiting ice growth. This reflects a relatively loose but common structural requirement for ice binding. From our analysis of a type I and a type III AFP crystal structure, we predict that a flat surface is the prerequisite for an AFP. The consensus has so far been that the formation of perfect hydrogen bonds is the dominant energetic that contributes to protein-ice interactions. This seemingly intuitive consideration imposes a rather strict requirement of a perfect positional match between the polar atoms on the AFP's ice-binding surface and the oxygen atoms on the ice surface. This seems to be unnecessary, in view of our success in identifying the ice-binding surface on the type III AFP without considering the presence of polar atoms. Instead of thinking purely in chemical terms, we would like to consider a more physical aspect of AFP-ice interaction, i.e., that of atomic surface complementarity through van der Waals interactions.

Our understanding of protein-ice interactions at the atomic level can be extended to other systems of biologically controlled crystal formation. These include normal and pathological calcification processes in bone, teeth, soft

tissues, and cartilage (Daulsi et al., 1997). For example, the proteins osteocalcin and osteopontin are likely to be involved in controlling the growth of hydroxyapatite in bone (Hunter et al., 1996), several salivary proteins are involved in the inhibition of crystallization of calcium phosphate in the oral cavity (Verdier, 1993), and the protein lithostathine is responsible for the inhibition of pancreatic stone formation (Geider et al., 1996). Although these mineralization processes are of physiological importance, very little is known about the nature of the underlying protein-mineral interactions. The principle governing the interaction between AFP and ice surface as presented in this study may also be applicable to the mentioned mineralization systems.

ACCESSION NUMBERS

Coordinates and structure factors of the ocean pout antifreeze protein (isoform HPLC-3) are being deposited with the Brookhaven Data Bank. PDB ID codes are pending for processing.

We thank Dr. G. Fletcher for providing partially purified protein; Dr. V. S. Ananthanarayanan for stimulating discussion, critically reviewing the manuscript, and suggesting the term "pretzel fold"; Dr. A. Berghuis for assistance with graphics programs; and R. Satmaka for technical assistance.

We acknowledge the Canadian Space Agency (DSCY), NSERC Canada (DSCY), and MRC Canada (CLH) for funding; and the National Cancer Institute for allocation of computing time and staff support at the Biomedical Supercomputing Center of the Frederick Cancer Research and Development Center.

REFERENCES

Brunger, A. T. 1993. X-PLOR, Verson 3.1: A System for X-ray Crystallography and NMR. Yale University Press, New Haven, CT.

Chao, H., F. D. Sonnichsen, C. I. DeLuca, B. D. Sykes, and P. L. Davies. 1994. Structure-function relationship in the globular type III antifreeze protein: identification of a cluster of surface residues required for binding to ice. *Protein Sci.* 3:1760–1769.

Chen, L., A. L. DeVries, and C. H. C. Cheng. 1997. Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proc. Natl. Acad. Sci. USA.* 94:3811–3816.

Cheng, C. H. C., and A. L. DeVries. 1991. The role of antifreeze glycopeptides and peptides in freezing avoidable of cold-water fish. *In* Life under Extreme Conditions. G. d. Prisco, editor. Springer-Verlag, Berlin. 1–14.

Daulsi, G., J. M. Bouler, and R. Z. LeGeros. 1997. Adaptive crystal formation in normal and pathological calcifications in synthetic calcium phosphate and related biomaterials. *Int. Rev. Cytol.* 172:129–191.

Davies, P. L., and C. L. Hew. 1990. Biochemistry of fish antifreeze proteins. *FASEB J.* 4:2460–2468.

DeVries, A. L. 1983. Antifreeze peptides and glycopeptides in cold-water fishes. *Annu. Rev. Physiol.* 45:245–260.

DeVries, A. L. 1986. Antifreeze glycopeptides and peptides: interactions with ice and water. *Methods Enzymol.* 127:293–303.

DeVries, A. L., S. K. Komatsu, and R. E. Feeney. 1970. Chemical and physical properties of freezing-point depressing glycoproteins from Antarctic fishes. *J. Biol. Chem.* 245:2901–2908.

Duman, J. G. 1982. Insect antifreezes and ice-nucleating agents. *Cryobiology.* 19:613–627.

Duman, J. G., and T. M. Olsen. 1993. Thermal hysteresis protein activity in bacteria, fungi and phylogenetically diverse plants. *Cryobiology.* 30:322–328.

Ewart, K. V., B. Rubinsky, and G. L. Fletcher. 1992. Structural and functional similarity between fish antifreeze proteins and calcium-dependent lectins. *Biophys. Biochim. Res. Commun.* 185:335–340.

Furey, W., and S. Swaminathan. 1997. PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules. *Methods Enzymol.* 277:590–620.

Geider, S., A. Baronnet, C. Cerini, S. Nitsche, J. P. Astier, R. Michel, R. Biostelle, Y. Berland, J. C. Dagorn, and J. M. Verdier. 1996. Pancreatic lithostathine as a calcite habit modifier. *J. Biol. Chem.* 271: 26302–26306.

Griffith, M., X. M. Yu, M. Antikainen, W. C. Hon, D. S. C. Yang, K. Pihakaski-Maunsbach, and J. U. Chun. 1997. Antifreeze proteins in winter rye. *Physiol. Plant.* 100:327–332.

Hon, W. C., M. Griffith, A. Mlynarz, Y. C. Kwok, and D. S. C. Yang. 1995. Antifreeze proteins in winter rye are similar to pathogenesis-related proteins. *Plant Physiol.* 109:879–889.

Hunter, G. K., P. V. Hauschka, A. R. Poole, L. C. Rosenberg, and H. A. Goldberg. 1996. Nucleation and inhibition of hydroxyapatite formation by mineralized tissue proteins. *Biochem. J.* 17:59–64.

Jia, Z., C. I. DeLuca, H. Chao, and P. L. Davies. 1996. Structural basis for the binding of a globular antifreeze protein to ice. *Nature.* 384:285–288.

Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta. Crystallogr.* 47:110–119.

Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.

Knight, C. A., C. C. Cheng, and A. L. DeVries. 1991. Adsorption of alpha-helical antifreeze peptides on specific ice crystal sruface planes. *Biophys. J.* 59:409–418.

Knight, C. A., A. L. DeVries, and L. D. Oolman. 1984. Fish antifreeze protein and the freezing and recrystallization of ice. *Nature.* 308: 295–296.

Knight, C. A., E. Driggers, and A. L. DeVries. 1993. Adsorption to ice of fish antifreeze glycopeptides 7 and 8. *Biophys. J.* 64:252–259.

Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.

Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemistry of protein structures. *J. Appl. Crystallogr.* 26:283–291.

Li, X. M., K. Y. Trinh, and C. L. Hew. 1991. Expression and characterization of an active and thermally more stable recombinant antifreeze polypeptide from ocean pout, *Macrozoarces americanus*, in *Escherichia coli*: improved expression by the modification of the secondary structure of the mRNA. *Protein Eng.* 4:995–1002.

Means, G. E., and R. E. Feeney. 1971. Chemical modification of proteins. Holden-Day, San Francisco.

Merrit, E. A., and M. E. P. Murphy. 1994. Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.* D50: 869–873.

Messerschmidt, A., and J. W. Pflugrath. 1987. Crystal orientation and x-ray pattern prediction routines for area-detector diffractometer systems in macromolecular crystallography. *J. Appl. Crystallogr.* 20:306–315.

Ng, N. F., and C. L. Hew. 1992. Structure of an antifreeze polypeptide from the sea raven. Disulfide bonds and similarity to lectin-binding proteins. *J. Biol. Chem.* 267:16069–16075.

Nicholls, A., K. Sharp, and B. Honig. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins.* 11:281–296.

Otwinowski, Z. 1993. Oscillation data reduction program. *In* Proceedings of the CCP4 Study Weekend: Data Collection and Processing. L. Sawyer, N. Issacs, and S. Bailey, editors. SERC Daresbury Laboratory, Warrington, England. 56–62.

Raymond, J. A., P. Wilson, and A. L. DeVries. 1989. Inhibition of growth of nonbasal planes in ice by fish antifreezes. *Proc. Natl. Acad. Sci. USA.* 86:881–885.

Read, R. 1994. Model bias and phase combination. *In* Proceedings of the CCP4 Study Weekend. S. Bailey, R. Hubbard, and D. Waller, editors. Daresbury Laboratory, Warrington, England. 31–40.

Sayle, R. A., and E. J. Millner-White. 1995. Rasmol—biomolecular graphics for all. *Trends Biochem. Sci.* 20:374–376.

Sicheri, F., and D. S. C. Yang. 1995. Ice-binding structure and mechanism of an antifreeze protein from winter flounder. *Nature.* 375:427–431.

Sonnichsen, F. D., C. I. DeLuca, P. L. Davies, and B. D. Sykes. 1996. 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.

Terwilliger, T. C., S. H. Kim, and D. Eisenberg. 1987. Generalized method of determining heavy-atom positions using the difference Patterson function. *Acta Crystallogr.* A43:1–5.

Urrutia, M. E., J. G. Duman, and C. A. Knight. 1992. Plant thermal hysteresis proteins. *Biochim. Biophys. Acta.* 1121:199–206.

Verdier, J. M. 1993. Macromolecular inhibitors of crystallization in saliva and bile. *Nephrologie.* 14:251–255.

Wang, B. C. 1985. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* 114:90–117.

Wen, D., and R. A. Laursen. 1993. Structure-function relationships in an antifreeze polypeptide: the effect of added bulky groups on activity. *J. Biol. Chem.* 268:16401–16405.

Xue, Y. Q., F. Sicheri, P. Ala, C. L. Hew, and D. S. C. Yang. 1994. Single crystals of a type III antifreeze polypeptide from ocean pout. *J. Mol. Biol.* 237:351–352.

Yang, D. S. C. 1993. Protein-ice interactions: antifreeze proteins. *In* The Amphipathic Helix. R. M. Epand, editor. CRC Press, Boca Raton, FL.

Yang, D. S. C., M. Sax, A. Chakrabartty, and C. L. Hew. 1988. Crystal structure of an antifreeze polypeptide and its mechanistic implications. *Nature.* 333:232–237.