BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 2593–2603 2593
DOI: 10.1021/bi901929n

Structural Basis for the Superior Activity of the Large Isoform of Snow Flea Antifreeze Protein[†,‡]

Yee-Foong Mok,[§] Feng-Hsu Lin,[§] Laurie A. Graham,[§] Yeliz Celik,[∥] Ido Braslavsky,[∥] and Peter L. Davies*[,§]

[§]*Department of Biochemistry and Protein Function Discovery Group, Queen's University, Kingston, Ontario K7L 3N6, Canada, and* [∥]*Department of Physics and Astronomy, Ohio University, Athens, Ohio 45701*

Received November 10, 2009; Revised Manuscript Received January 22, 2010

ABSTRACT: The snow flea (*Hypogastrum harveyi*) is protected from freezing at sub-zero temperatures by a glycine-rich antifreeze protein (AFP) that binds to seed ice crystals and prevents them from growing larger. This AFP is hyperactive and comprises two isoforms [Graham, L. A., and Davies, P. L. (2005) *Science 310*, 461]. The larger isoform (15.7 kDa) exhibits several-fold higher activity than the smaller isoform (6.5 kDa), although it is considerably less abundant. To establish the molecular basis for this difference in activity, we determined the sequence of the large isoform. The primary sequences of these two isoforms are surprisingly divergent. However, both contain tripeptide repeats and turn motifs that enabled us to build a three-dimensional model of the large isoform based upon the six-polyproline helix structure of the small isoform. Our model contains 13 polyproline type II helices connected by proline-containing loops stacked into two flat sheets oriented antiparallel to one another. The structure is strictly amphipathic, with a hydrophilic surface on one side and a hydrophobic, putative ice-binding surface on the other. The putative ice-binding site is approximately twice as large in area as that of the small isoform, providing an explanation for the difference in activity that is consistent with other examples noted. By tagging the recombinant AFP with green fluorescent protein, we observed its binding to multiple planes of ice, especially the basal plane. This finding supports the correlation between AFP hyperactivity and basal plane binding first observed with spruce budworm AFP.

Antifreeze proteins (AFPs)[1] are structurally diverse macromolecules found in a variety of organisms that live in freezing habitats (*1*−*5*). AFPs make crucial contributions to the existence of life in these environments, where organisms can be killed by uncontrolled ice crystal growth (*6*, *7*). AFPs function at the ice crystal surface by an absorption−inhibition mechanism (*8*), creating microcurvatures in the ice surface that, due to the Kelvin effect, energetically disfavor the further addition of water molecules (*9*, *10*). This results in the depression of a solution's freezing temperature below its equilibrium melting point, a phenomenon that is termed "thermal hysteresis" (TH) (*11*). In this way, AFPs inhibit the growth of ice crystals within organisms and enhance their ability to survive at sub-zero temperatures.

To date, AFPs have been isolated from many species of fish, insects, plants, and microorganisms. For proteins that serve the same function, AFPs have a remarkable diversity of structures, including a number of folds that are novel (*3*). One common feature, however, appears to be an ice-binding face that is relatively flat and more hydrophobic than other surfaces of the AFPs and that in several cases contains a regular array of threonine residues (*12*−*14*). Nevertheless, the molecular basis of AFP binding to ice is still not well understood, and a major driving factor for the modeling and elucidation of new AFP structures is the potential for discerning clues about the exact mechanism by which AFPs exert their effect.

A highly potent, glycine-rich AFP was discovered in snow fleas (*Hypogastrum harveyi*) collected from Eastern Ontario, Canada, in late winter (*15*). Crude extracts from these primitive arthropods were able to depress the freezing temperature of a solution in the presence of ice ~6 °C below its melting temperature. The substantial TH activity exhibited by the extracts places snow flea AFP in the class of hyperactive AFPs, with activity significantly higher than those found in most fish and plants (*16*). This feature potentially helps snow fleas avoid freezing in the exposed, snow-laden ground from where they forage. The AFP responsible for the observed TH activity was subsequently purified from the extracts by two rounds of ice affinity purification (*17*). Reversed-phase high-performance liquid chromatography (HPLC) resolved the protein into two isoforms, each with a single mass (*15*). The larger (15.7 kDa) of the two isoforms was considerably less abundant than the smaller (6.5 kDa) one. However, the 15.7 kDa isoform exhibited significantly greater antifreeze activity. Both isoforms had an unusual amino acid composition, in that glycine made up ~45% of the residues. The smaller isoform contained no aromatic or long chain aliphatic amino acids, while the larger isoform had a small number of these residues. The small isoform contained four internally disulfide bonded cysteines, while the large isoform had two.

When clones from a snow flea cDNA library were randomly sequenced, four of 57 sequences matched the amino acid composition of the small isoform and had a predicted mass (after

[†]P.L.D. holds a Canada Research Chair in Protein Engineering. F.-H. L. is funded by an Ontario Graduate Scholarship. Y.C. acknowledges support from the Condensed Matter and Surface Science program at Ohio University. This research was funded by a grant to P.L.D. from the Canadian Institutes for Health Research and by a grant to I.B. from the National Science Foundation (Grant CHE-0848081).

[‡]The DNA sequence of the large isoform of snow flea antifreeze protein has been submitted to GenBank (accession number GU169329).

*To whom correspondence should be addressed. E-mail: peter.davies@queensu.ca. Telephone: (613) 533-2983. Fax: (613) 533-2497.

[1]Abbreviations: AFP, antifreeze protein; GFP, green fluorescent protein; Cy5, cyanine-5; PPII, polyproline type II; rmsd, root-mean-square deviation.

Published on Web 02/16/2010

removal of a signal sequence) equal to that experimentally determined for the naturally isolated AFP (*15*). None of the 57 sequences encoded the large isoform. The small isoform had a prominent tripeptide repeat pattern where glycine was present in the first position. The second position was occupied by either glycine or an amino acid with a small side chain. The third position was the most variable, but even here there was a pattern, where runs of charged and hydrophilic residues alternated with hydrophobic residues separated by discontinuities in the tripeptide repeat pattern that often included a proline residue. These observations led to the development of a model for the three-dimensional (3D) structure of the small isoform in which the 81-residue protein was folded into six polyproline type II helices that alternated in direction at the proline-containing turns (*18*). The core of the protein was stabilized by hydrogen bonds between the backbone amides which were made possible by the regular pairing of glycines on opposing faces of the internal coils. The model predicted a novel structure that had several of the attributes of AFPs. These include amphipathic character and an elongated, relatively flat hydrophobic surface with surface regularity that could potentially dock to ice. A high-resolution crystal structure of the small isoform obtained via complete chemical synthesis of the protein (*19*) was recently reported and is in excellent agreement with the model (*20*).

The large isoform of snow flea AFP is of interest because of its superior antifreeze activity compared to that of the small isoform. To determine the molecular basis of this difference in activity, we set out to sequence the large isoform. Given the similarity in amino acid composition between the two AFPs, we hypothesized that the 15.7 kDa protein was an isoform of the smaller AFP and would also fold into polyproline type II antiparallel sheets to give a larger version of the small isoform structure. In this paper, we present the primary structure of the large isoform as well as a 3D model based on the small isoform structure of snow flea AFP. The primary sequences of the two isoforms are surprisingly divergent, but the 3D models are highly similar. The hydrophobic face of the structure is at least twice as large as the equivalent face on the small isoform. Chemical synthesis of the gene for the large isoform then allowed us to express the protein recombinantly in sufficient amounts to confirm its superior activity. By tagging this protein with green fluorescent protein (GFP), we found that it binds to multiple planes of ice, including the basal plane, which is consistent with its hyperactivity.

MATERIALS AND METHODS

N-Terminal and Internal Sequencing of the 15.7 kDa Protein. N-Terminal sequence analysis was performed by Edman degradation on an Applied Biosystems Procise Sequencer at the Peptide Sequencing Facility of the Advanced Protein Technology Centre (Hospital for Sick Children, Toronto, ON) using protein from the 15.7 kDa HPLC peak. A tryptic mass fingerprint, as well as a second, internal protein sequence determined by MS/MS sequencing, were obtained from the Protein Function and Discovery facility (Queen's University).

Cloning of the DNA Sequence Using the Polymerase Chain Reaction. Snow flea mRNA (100 ng per reaction) was reverse transcribed with either oligo$(dT)_{17}$-CSX or random hexamer primers using the Thermoscript RT-PCR System (Invitrogen, Carlsbad, CA) by following the manufacturer's instructions. PCR was performed using 2 μL of cDNA ($^1/_{10}$ of that in the reaction mentioned above) in a 50 μL volume using Taq DNA polymerase with Q-solution (for GC-rich templates) (Qiagen, Germantown, MD), together with partially degenerate primers (final concentration of 0.4 μM) corresponding to the peptide sequences DGRSNGE (5′-GAYGGNAGRAGYAAY-GGCGAA-3′) and RGGDGAN (5′-TTNGCNCCRTCNCC-TCCACG-3′). The cDNA was denatured at 95 °C for 5 min, followed by 30 cycles at 95 °C for 1 min, 60 °C for 1 min, and 72 °C for 2 min. An aliquot (1 μL) was reamplified as described above for 25 cycles. All of the PCR bands described were TA cloned (TOPO TA-cloning kit, Invitrogen) and sequenced (Cortec, Kingston, ON). Note that the degeneracy of the primers described above was reduced on the basis of the sequence of short PCR fragments obtained using degenerate primers based on the N-terminal sequence APNGADG (5′-GCWCCHAAYGGWG-CWGAYGG-3′) and the internal sequence GGDGANG (5′-NC-CRTTNGCNCCRTCNCCNCC-3′).

The 3′ end of the cDNA was cloned from an aliquot of the cDNA library (*15*) by anchor PCR. The gene-specific primer (5′-ACAACCCGGTGGTAATGGTGGAAAC-3′) and the T7 extended primer (5′-ACGACTCACTATAGGGCGAATTGG-3′) were used in the first reaction, and the second reaction was nested at the 5′ end using a second gene-specific primer (5′-CGGGAC-TAGGTGGTGATAGTGTAA-3′).

Cloning of the 5′ Portion of the Gene by Inverse PCR. Aliquots (2 μg) of snow flea genomic DNA were digested in a 250 μL volume with 40 units of either *Hin*dIII, *Ase*I, or *Hae*III for 4 h at 37 °C and then overnight following the addition of an identical aliquot of enzyme. Following denaturation of the restriction enzymes at 68 °C, 200 ng of digested DNA was ligated in a 100 μL volume at 16 °C overnight. PCR was performed as previously described but with an annealing temperature of 65 °C. The reaction was nested at the 3′ end using the same primers used in anchor PCR and at the 5′ end using the following two primers, 5′-CAGACCCTGTTCCAGGGAACCCAT-3′ and 5′-ATCAC-AACCATTTGCCCCAGCAGTA-3′. A second portion of genomic sequence was obtained using genomic DNA in PCRs as described above, using one of the second gene-specific primers described above and a primer overlapping the stop codon (5′-TTACGCTCCACCGCCGCCACC-3′). The sequences of the two isoforms were aligned using DNAMAN version 6 (Lynnon Corp., Pointe-Claire, QC) followed by minor manual adjustments.

Modeling Building and Molecular Dynamics Simulation. Modeling of the large isoform employed the same concepts and strategy described for the small isoform (*18*) and was an intuitive, iterative process. Briefly, it began with the recognition of a tripeptide-repeating pattern throughout the protein suggestive of a 3-fold helical repeat. It continued with the identification of regularly spaced discontinuities in the repeat pattern that might correspond to bends or turns in the chains, thus dividing the sequence into 13 segments. The presence of a disulfide bond and irregular distribution of Gly-Gly pairs were also taken into account. A physical model was constructed using the HGS Biochemistry Molecular Model from Himomoto Plastics (Tokyo, Japan) before the model was rendered with PyMOL (*21*).

To perform molecular dynamics simulations, the model was computationally solvated in a 8.4 nm × 6.3 nm × 5.2 nm box of water containing 8409 water molecules and had a single net positive charge. To neutralize the charge of the system and to provide an effective concentration of 0.1 M NaCl, 33 water molecules were replaced with 16 Na^+ and 17 Cl^- ions. The system was subject to energy minimization by steepest descents,

position-restrained molecular dynamics to relax the solvent, followed by full molecular dynamics. All molecular dynamics calculations were done with GROMACS version 4.0.4 (*22*). The program uses a triclinic unit cell for its periodic boundary conditions. For the calculation of short-range nonbonded interactions, only the nearest image was considered. Long-range electrostatics was treated with the particle mesh Ewald method. The simulations were conducted under isothermal conditions using V-rescale temperature coupling. The GROMOS96 43a1 force field was employed. Simulations were conducted at 4 °C using a time step of 2 fs and for a total duration of 10 ns.

Expression of Recombinant Snow Flea AFP and GFP-Linked Snow Flea AFP. A codon-optimized gene sequence for the large isoform, flanked by *Nde*I and *Xho*I restriction sites, was designed by and purchased from GeneArt (Regensberg, Germany). The sequence was excised from the supplied vector with *Nde*I and *Xho*I and cloned into the pET28a vector for expression in *Escherichia coli* BL21(DE3). The gene sequence includes an N-terminal His tag and a stop codon that removes the C-terminal His tag present in the pET28a vector. To generate GFP-tagged large isoform AFP, the sequence encoding GFP was PCR amplified from pEGFP (catalog no. 6077-1, Clontech, Mountain View, CA) using a 5′ primer (5′-CTATAGCATATGGTGAGCAAGGGCGAGGAG-3′) and a 3′ primer (5′-AGCTCATATGACCTGCACCCTTGTACAGCTCGTCCATGC-3′) each containing an *Nde*I restriction site. Additionally, the 3′ primer was designed to encode the Gly-Ala-Gly linker sequence to better separate the GFP from the AFP. The resulting fragment was gel-purified and digested with *Nde*I. The expression construct described above was digested with *Nde*I and dephosphorylated with alkaline phosphatase. Following ligation, colonies were screened by PCR to determine the orientation of the GFP sequence using the T7 promoter and 3′ GFP primers. Clones that appeared to contain a correctly oriented insert were sequenced to ensure the fidelity of the sequence encoding the fusion protein.

Cultures were grown in 2 L of Luria-Bertani medium at 37 °C with kanamycin (100 μg/mL) until an OD_{600} of ~0.8 was reached. The culture was kept at 16 °C for ~1 h prior to induction of protein expression with the addition of isopropyl β-D-thiogalactoside at a final concentration of 1 mM. Expression was allowed to proceed at 16 °C for 24 h. Cells were then harvested by centrifugation (2500*g* for 40 min at 4 °C) and resuspended in 50 mL of 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl containing an EDTA-free protease inhibitor cocktail (Roche, Indianapolis, IN). The resuspension was sonicated in a Sonic Dismembrator 5000 (Fisher Scientific, Waltham, MA), using 10 s bursts at 30% amplitude. The suspension was allowed to cool between each burst to ensure that the temperature stayed below 12 °C at all times. The suspension was clarified by centrifugation at 21000*g* and 4 °C for 40 min in a JA20 rotor.

Ice Affinity Purification. Ice affinity purification was performed as described previously (*17*). Briefly, the supernate was diluted to 100 mL with the resuspension buffer used above. A brass "finger" filled with circulating coolant was immersed into the diluted supernate, and the temperature of the coolant was decreased from −1.0 to −4.0 °C over 48 h, at which point approximately half of the sample was frozen around the ice finger. This ice fraction was melted, diluted to a volume of 100 mL as described above, and subjected to a second round of ice affinity purification. Aliquots were concentrated in an Amicon Ultra centrifugal filter (Millipore, Bellerica, MA), and the AFP concentration was determined by amino acid analysis (Peptide Sequencing Facility, Hospital for Sick Children). The yield for the purified protein is approximately 10 mg/L of expression medium.

TH Measurements. TH measurements were taken using a Clifton (Hartford, NY) nanoliter osmometer as previously described (*23*), but using a cooling rate of 0.09 °C/min (equivalent to 50 mOsm/min). Ice crystals were examined with a Leitz Dialux microscope with a 160/-EF 32/0.40 objective and photographed with a Nikon Coolpix 4500 digital camera mounted on the microscope.

Ice Hemisphere and Ice Crystal Fluorescence Experiments. A single ice crystal was grown from degassed ddH_2O (MilliQ) and mounted on a brass coldfinger to seed the growth of a single-crystal hemisphere (*24*). The *c*-axis was oriented normal to the coldfinger. The hemisphere was then immersed in a dilute solution of GFP and large isoform AFP in an insulated cup. The temperature of the coldfinger was gradually lowered at a rate of −0.1 °C/h to effect ice growth. The hemisphere was removed from the solution after ~12 h before the ice grew to the edges of the cup and then allowed to melt back slightly at room temperature for the removal of nonspecifically bound protein. The ice hemisphere was illuminated with a UV lightbox and photographed.

Fluorescent microscopic ice crystal experiments were performed as described previously (*25*). Briefly, a sample was held between two glass coverslips. A 22 mm square coverslip was used as the base of the cell onto which a drop of AFP solution was placed. An 18 mm circular coverslip was set down on the liquid and was used as the top of the cell. The periphery of the coverslip was sealed with uncured polydimethylsiloxane. Coverslips were placed on a temperature-controlled metal plate containing an array of light holes with a diameter of 125 μm. Samples consisted of GFP-tagged AFP and a free dye, cyanine-5 (Cy5). Nonconjugated Cy5 dye does not interact with ice and is not incorporated into the ice body. The sample was imaged using a confocal fluorescence microscope (model LSM 510, Zeiss, Thornwood, NY) equipped with 488 and 633 nm illumination lines and filters for the detection of GFP and Cy5 fluorescence. To image the fluorescent signal being emitted by the samples, the image of nonconjugated Cy5 dye was subtracted from the GFP−AFP image as explained in detail previously.

RESULTS

The Large Isoform of Snow Flea AFP Is a 203-Residue Glycine-Rich Secreted Protein with a Low Level of Sequence Identity to the Small Isoform. Initial attempts to clone the long isoform by screening the cDNA library with labeled short isoform cDNA were unsuccessful because of the limited sequence identity between the two isoforms (see below). Therefore, the purified protein was subjected to N-terminal sequencing by Edman degradation from which the following sequence was obtained: CKAPNGADGRSNGEAGGA (underlined in Figure 1). A second sequence was obtained following tandem mass spectrometry of a tryptic fragment which was interpreted as GGNGANGGSGHGNPxR (underlined in Figure 1). Since this second peptide did not overlap with the N-terminal sequence, we reasoned that there was probably sufficient distance between the two peptide sequences to generate a useful PCR product. Internal fragments were obtained using degenerate or partially degenerate primers, and additional 3′ sequence was obtained by anchor PCR and 5′ sequence by inverse PCR. A portion of the gene containing

```
L   1   mlpsriigv--llfcciggiisCKAPNGADG   29
S   1   maqmkfilvaflvvlavswanaCKGADGAHG   31

L   30  R--SNGEAGGAGTAGANGCDGGDGGNGFPGT   58
S   32  VNGCPGTAGAAGSVGGPGCDGGHGGNGGNGN   62

L   59  GSAAGGAGGIGGVGGAGAKGGNGGKGGIGAS   89
S   63  PGCAGGVGGAGGASGGTGVGGRGGKGGSG--   91

L   90  SESAAGGAGGAGGAAGAGSSGPGGQGGEGGK  120
S   92  ---TPKGADGAPGAP----------------  103

L   121 GGKANGVTGTGGAGGIGGVGGAGSGGQPGGN  151
S       ------------------------------

L   152 GGNAGLGGDSVTAGGAGGTGGAGGAGTPGGR  182
S       ------------------------------

L   183 GGNGANGGSGHGNPGGRGLPGAAGLPVGGGG  213
S       ------------------------------

L   214 AGGIGGGGGA                      223
S       ----------
```

FIGURE 1: Sequence alignment of snow flea AFP large and small isoforms. L and S denote the large and small isoform sequences, respectively. Identical residues are highlighted in gray, and gaps in the alignment are indicated by dashes. The signal peptide is in lowercase, and residues identified by N-terminal or tandem mass spectrometry sequencing are shown in bold. An underlined Lys or Arg residue indicates that the mass of the tryptic cleavage fragment was observed by MALDI mass spectrometry. Italicized residues indicate short tryptic fragments for which individual masses were not observed but that were observed as a missed cleavage product with the preceding and/or succeeding fragment. Sections underlined in gray denote the sequences obtained from N-terminal sequencing and MS/MS sequencing.

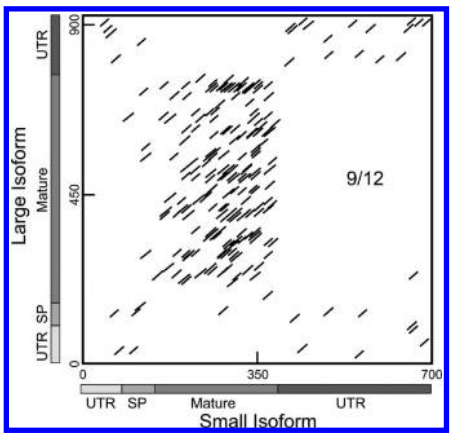

FIGURE 2: Dot matrix plot of the large and small isoform nucleotide sequences. The large isoform nucleotide sequence was plotted against that of the small isoform using a match of 9 of 12 residues. Regions of the snow flea gene, including the signal peptide (SP) and untranslated regions (UTR), are indicated by gray bars parallel to the axes. The nucleotide sequence lengths are indicated by numbers.

two introns was also amplified. In the interest of brevity, the considerable difficulties we encountered in attempting to clone this sequence have not been described in detail. Suffice to say that we were unable to amplify the entire sequence in one piece and frequently obtained fragments containing internal deletions, presumably due to the extreme G richness (46%) and repetitiveness of the coding sequence, which probably led to inefficient amplification with partially extended products acting as primers at ectopic sites. Hindsight revealed these fragments arose either by annealing of the second primer to the position encoding GGEGGKG, likely as a result of our misidentification of Asn for Asp in the MS/MS sequence, or by PCR recombination within repetitive glycine-encoding G-rich regions (data not shown).

The complete cDNA sequence deduced from the assembled sequence was found to encode a mature, 203-residue protein that was very glycine-rich (50%) and contained the same tripeptide repeat pattern, runs of Gly-Gly-x or Gly-x-x, as found in the small (81-residue) isoform (Figure 1). Like the small isoform, it is preceded by a classic signal polypeptide for protein export, and both mature proteins start with the sequence Cys-Lys. Despite these similarities, the two protein sequences are difficult to align with any confidence because of the glycine richness and repeat pattern. In addition, the signal peptides differ in length and are not obviously similar. Also, the larger isoform contains two fewer cysteine residues, suggesting that it contains only one disulfide bond.

The composite DNA sequence (Figure 1 of the Supporting Information) of 1162 nucleotides includes 20 nucleotides of poly(A+) tail the appropriate distance downstream of an AATAAA polyadenylation signal. The 5′ region, cloned by inverse PCR, is predicted to contain a potential TATA box sequence, transcription start site, and 95 bases of the 5′ untranslated region. Two introns, 84 and 99 nucleotides in length, containing canonical GT-AG splice sites were also found when a second portion of genomic DNA was amplified and sequenced. The deduced transcript (GenBank accession number GU169329) is 927 nucleotides in length, including a 160-nucleotide 3′ untranslated region.

Several pieces of evidence provide confidence that the composite sequence is that of the large snow flea AFP isoform. The deduced sequence matches the N-terminal sequence determined by Edman degradation at every position (Figure 1). Its amino acid composition agrees very closely with that determined experimentally and predicted a length of 202 residues (*15*). Also, the experimentally determined average mass of the large isoform (15709 ± 5 Da) corresponds to that calculated for the sequence in Figure 1, with 2 Da removed for internal disulfide bond formation (15712 Da). More importantly, the tryptic fragment masses observed by MALDI-MS (Table 1 of the Supporting Information) match the masses of the predicted tryptic peptides. These fragments span all but the C-terminal fragment, which was likely not observed because it lacks a basic residue.

Given the low degree of similarity between the two AFP isoforms (< 50% sequence identity), a dot matrix analysis comparison of the cDNA sequences, including untranslated regions, was performed (Figure 2). When a match of 9 of 12 nucleotides was searched, there was no obvious diagonal pattern generated. There were many short alignments within the coding region but at multiple sites along the sequence to generate a rectangular pattern, typical of that seen when highly repetitive or low-complexity sequences are compared. This can be attributed to the many Gly-Gly-x repeats containing the G-rich glycine codon, GGN. The 5′ and 3′ untranslated regions and the signal sequence had surprisingly few matches.

Molecular Modeling of the Large Isoform as a Polyproline Type II Helix Structure. Despite the low degree of sequence identity between the two AFP isoforms, a tripeptide repeat with glycine present in the first position and often present in the second position was common to both sequences. Also, it was possible to pick out in the large isoform a similar pattern of

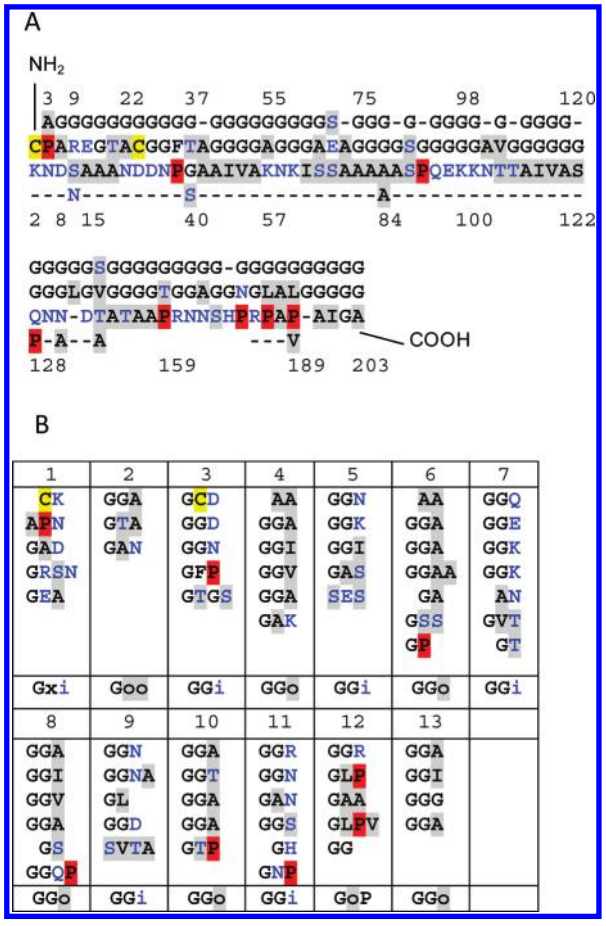

FIGURE 3: Primary structure of the snow flea AFP large isoform. (A) Amino acid sequence of the large isoform of snow flea AFP written from top to bottom and from left to right to emphasize the three-amino acid repeats. A red background denotes Pro residues predicted to disrupt the PPII helix. A yellow background marks cysteine residues. A gray background denotes small residues, including Ala, Val, Ser, and Thr. Blue letters indicate charged or hydrophilic residues, including Arg, Lys, Asp, Asn, Ser, Thr, and His. (B) Predicted segments arranged into columns, with the tripeptide consensus repeat for each column indicated below. The symbol o denotes small residues, the symbol i hydrophilic or large residues, and the symbol x any residue.

interruptions to the tripeptide repeats that putatively divide the sequence into 13 alternating segments (Figure 3A,B). These interruptions were often but not always associated with a proline. In organizing the sequence, we predicted that segment 2 would be shorter than the others because such an arrangement would bring the cysteine residues in segments 1 and 3 into proximity to form a disulfide bond in the 3D structure. Segments 4−13 were organized on the basis of interruptions of the sequence by a proline or a break in the repeating Gly-Gly-x motif that suggest the presence of a turn or loop.

Once the organization of the segments was established, it was possible to model each segment as a polyproline type II (PPII) helix and connect these by a hydrogen bonding network to produce two antiparallel sheets of coils, which was done for the small isoform (*18*). In the model, each Gly-x1-x2 repeat makes a helical turn consisting of three residues with the hydrogen bonding groups of the peptide bonds facing outward from the helical axis (Figure 4A). The 13 segments form individual PPII regions and make backbone−backbone hydrogen bonding interactions with each other. The consecutive segments are connected by loops that change the direction of the polypeptide chain. This

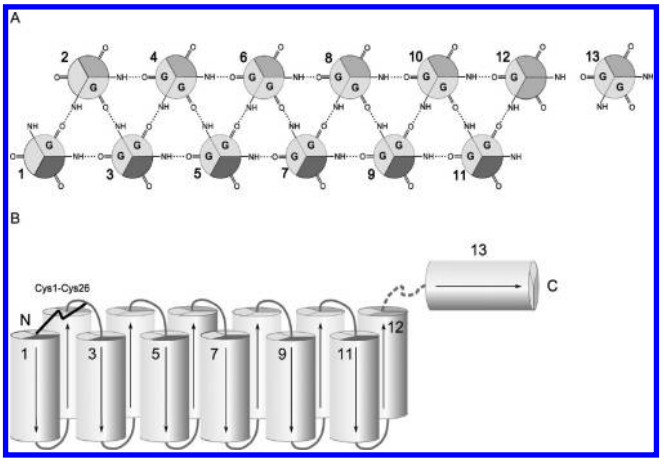

FIGURE 4: Polyproline type II-like fold of the snow flea AFP large isoform. (A) Top-down view, with the hydrogen bonding pattern between coils shown as dotted lines. The segments are numbered 1−13 and divided into thirds to represent the stacking of amino acids. G represents glycine stacks. Hydrophobic residue stacks are shaded light gray, and hydrophilic residue stacks are shaded dark gray. (B) Side view of the polyproline type II coils with the disulfide bond drawn as a thick black line. Arrows illustrate the alternating direction of the segments. Loops connecting segments are colored gray. The loop connecting coil 12 to coil 13 is represented by a dashed line, indicating the uncertainty of the packing of coil 13 relative to the other coils.

arrangement brings Cys 1 and Cys 26 close enough for disulfide bonding, which links the N-terminal end of segment 1 to the loop between segments 2 and 3. Segments 1, 3, 5, 7, 9, and 11 are parallel (forming one face of the protein) and antiparallel to segments 2, 4, 6, 8, 10, 12, and 13 (which form the other face) (Figure 4B). There is a possibility that segment 13 actually loops back such that it is antiparallel to segment 12, but we think it is more likely to orient toward segments 8 and 9 (and parallel to segment 12) where the presence of Ser 124, Ser 141, and Thr 143 provides side chains for hydrogen bonding.

In the small isoform, the central segments (segments 3−5) typically have glycine in the second position of their tripeptide repeats, whereas the more peripheral segments often have alanine (*18*). The central segments are in contact with four other strands, and it is critical for the fold that they contain two glycine-rich faces. Having glycine in the second position ensures that there is no steric barrier to making hydrogen bonds to the neighboring PPII polypeptide segments that abut a large proportion of their surface area. In the large isoform, the tripeptide repeats within the more extensive central core (segments 3−11, six more segments than the smaller isoform) favor glycine in the first and second position. The third position (x2) is generally occupied by hydrophilic residues in odd-numbered segments 1, 3, 5, 7, 9, and 11 and by small hydrophobic residues in even-numbered segments 2, 4, 6, 8, 10, 12, and 13 (Figures 3B and 4). This arrangement of hydrophilic and hydrophobic residues in the third position of the repeat causes the face formed by the even-numbered segments to be quite hydrophobic, whereas the odd-numbered segments form a relatively hydrophilic surface (Figures 4B and 5). As with the small isoform, the central structure of the protein was maintained not by a hydrophobic core but by an extensive network of hydrogen bonds between the neighboring peptide backbones.

The Large Isoform Is a Flat Amphipathic Platelike Structure. The PPII-like fold was built into a physical model, as was done with the small isoform, to determine three-dimensional

coordinates. An energy-minimized model is represented in Figure 5 and in space-filling mode in Figure 6. The modeled protein is a disk with a length of approximately 34 Å, a width of 46 Å, and a thickness of 12 Å. The overall fold of the large isoform is similar to that of the small isoform, but its larger size extends the twin sheets of alternating helices to effectively double the length of the structure, which also substantially emphasizes its flatness. The alternation of the coils introduces pronounced amphipathic character. All the charged groups (seven basic and seven acidic) are on one side or at the edges of the disk. The other side has a very extensive hydrophobic patch with a few peripheral hydrophilic groups (Figure 6). The hydrophobic side chains on this face form a regular pattern of equally spaced ridges that together with the hydrophobicity and flatness make the hydrophobic patch a potential ice-binding site. Interestingly, the predicted hydrophobic patch extends to the edge of the disk on the C-terminal end (Figure 6A, right panel), whereas the equivalent region on the small isoform does not contain such hydrophobic character.

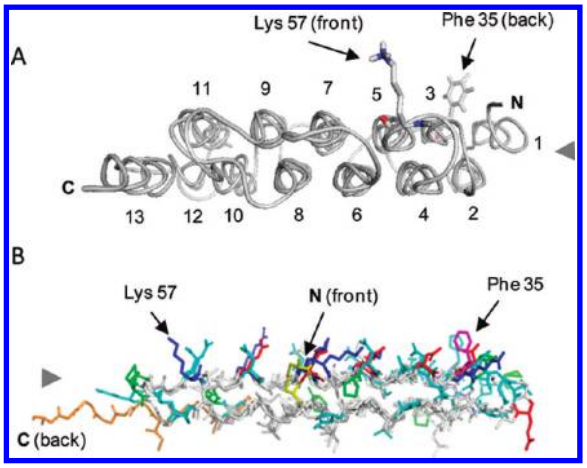

FIGURE 5: Overall structural model of the large isoform of snow flea AFP. (A) Representation of the structure as a tube backbone, with the 13 coils numbered from the N- to C-termini. The side chains of Phe 35 and Lys 57 are shown as stick figures for the sake of orientation. The gray arrowhead indicates the side of the molecule viewed in panel B. (B) Edge-on view of the structure, with side chains shown and Phe 35 and Lys 57 identified. Coils 1 and 2 are on top. The gray arrowhead indicates the direction viewed in panel A. Positively and negatively charged residues are colored blue and red, respectively. Uncharged polar residues are colored cyan, and prolines are colored green. Phenylalanine is colored purple. Cysteines are colored yellow. Hydrophobic and glycine residues are colored light gray, and coil 13 is colored tan. The N- and C-termini are denoted with N and C, respectively.

The stability of the large isoform model was tested by a 10 ns molecular dynamics simulation at 4 °C. The total energy of the system at each time step of the simulation, the rmsd, and the hydrogen bonding between peptide backbones per residue are plotted in Figure 7. The total energy and rmsd remained relatively constant beyond 2 ns, indicating that the structure has reached equilibrium. The proportion of hydrogen bonds formed per residue for the large isoform was similar to that for the small isoform.

Expression and Characterization of the Large Isoform. Determination of the full sequence of the large isoform of snow flea AFP allowed us to obtain a chemically synthesized gene for the protein in which the codons have been optimized for expression in *E. coli*. Accordingly, the protein was recombinantly expressed and the lysate supernate subjected to two rounds of ice affinity purification (*17*). A sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) analysis of the purification procedure shows moderate enrichment of a protein that was preferentially incorporated into the ice fraction after ice affinity purification, with an apparent molecular mass of 30 kDa (Figure 8A). It should be noted that the large isoform contains only four Arg residues and no aromatic residues and thus binds poorly with gel-staining dyes like Coomassie blue. Therefore, the intensity of the observed band very likely underrepresents the amount of large isoform protein present in the ice fractions.

The final ice fraction was examined and tested for TH activity. Ice crystals formed in a concentrate of this fraction loosely resembled grains of rice (Figure 8B) as did the ones observed

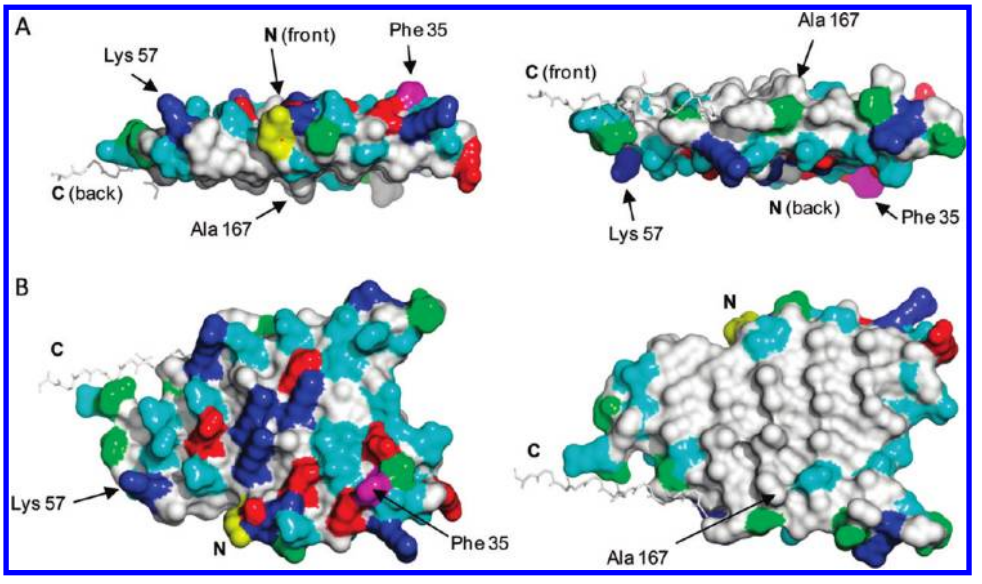

FIGURE 6: Space-filling surface representation of the large isoform of snow flea AFP. (A) Edge view of the protein: (left) viewing coils 1 and 2 edge-on with the hydrophilic side facing upward and (right) viewing coils 12 and 13 edge-on with the hydrophobic side facing upward. (B) Top-down views of the protein showing the hydrophilic side (left) and the hydrophobic side (right). Residues are colored as described in the legend of Figure 5. N- and C-termini are denoted with N and C, respectively. Phe 35, Lys 57, and Ala 167 are indicated for the sake of orientation.

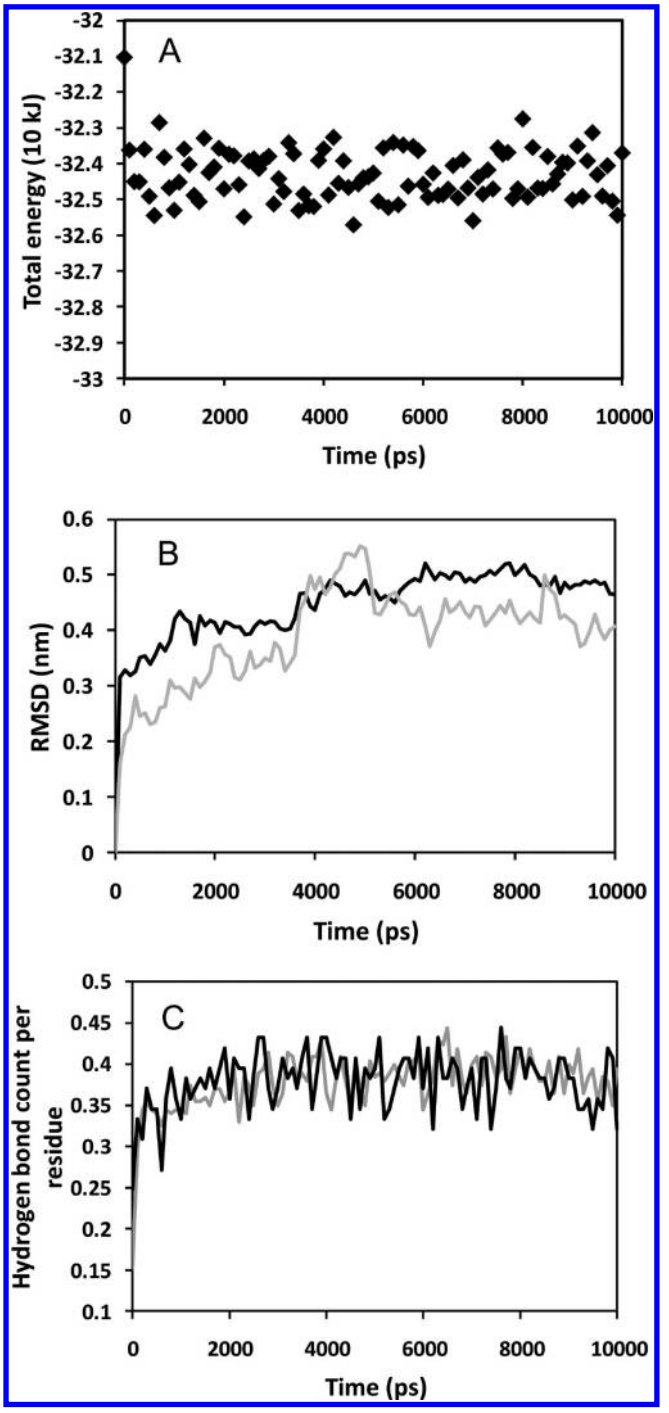

FIGURE 7: Molecular dynamics simulation of the large isoform of snow flea AFP. Total energy (A), root-mean-square deviation (rmsd) (B), and number of hydrogen bonds formed per residue (C), at each time step of the simulation performed on the model for 10 ns at 4 °C. For panels B and C, data for the large isoform are colored black and data for the small isoform gray.

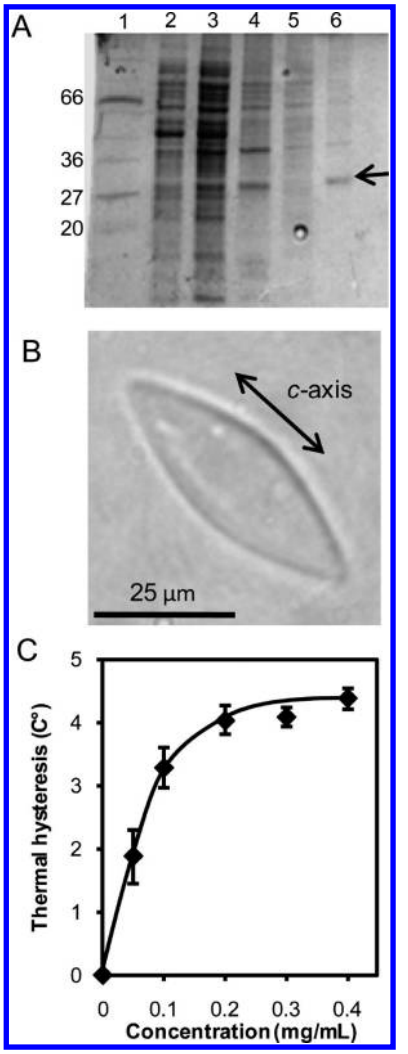

FIGURE 8: Expression and characterization of the recombinant large isoform of snow flea AFP. (A) SDS−PAGE analysis of the recombinant expression and purification process: lane 1, molecular mass markers; lane 2, *E. coli* lysate after induction of expression and sonication; lane 3, concentrate of the lysate for ice affinity purification; lane 4, ice fraction after one round of ice affinity purification; lane 5, liquid fraction after one round of ice affinity purification; lane 6, ice fraction after two rounds of ice affinity purification. The arrow indicates the band showing enrichment of the putative large isoform. (B) Single ice crystal, showing grain-shaped morphology, grown in a solution of the large isoform protein and held at a temperature just below its equilibrium melting point. (C) Thermal hysteresis activity plotted as a function of protein concentration. Three replicates of each concentration were measured and averaged. Vertical bars represent the standard deviation.

in the snow flea homogenate (*15*). At a concentration of 0.4 mg/mL (25 μM), almost 4 °C of TH was observed (Figure 8C). A high degree of TH was maintained as the concentrations decreased from 0.3 to 0.1 mg/mL, until activity was reduced to ~2 °C at 0.05 mg/mL. The remarkable overall activity of the recombinant large isoform AFP is comparable at equivalent concentrations to a long isoform of spruce budworm antifreeze protein (CfAFP-501) (*26*) that has TH activity 3 times greater than that of short AFP isoforms from that species. The activity of the recombinant protein is also similar to the observed activity of the large isoform purified from snow flea homogenate and approximately twice as high as the activity of the small isoform (~4 °C at 25 μM for the large isoform compared to ~2 °C at the same concentration for the small isoform). These findings provide confidence that the ~30 kDa enriched band after ice affinity purification contained recombinant protein of the snow flea large isoform.

To visualize the binding of the large isoform to ice and determine the ice crystal planes with which the protein interacts, we expressed the large isoform attached to green fluorescent protein and conducted experiments adapted from Knight et al. (*24*). A single crystal ice hemisphere was immersed in a solution containing GFP large isoform AFP and allowed to grow to a diameter of approximately 5 cm. The views of the hemisphere under UV light down either the *c*-axis (viewing the basal plane) or

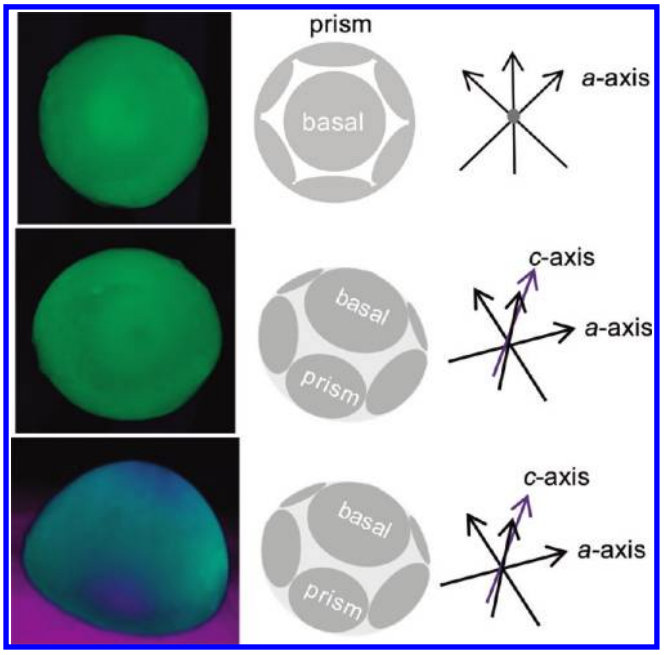

FIGURE 9: Ice hemispheres grown in the presence of the GFP-tagged large isoform of snow flea AFP. The top panel shows the top-down view of an ice hemisphere grown in ~0.04 mg/mL protein with the *c*-axis normal to the plane of the page and looking directly at the basal plane. The middle panel shows an angled view of the same ice hemisphere looking directly at a prism plane. The bottom panel shows an angled view of an ice hemisphere grown in ~0.01 mg/mL protein looking at a prism plane. The color differences in this image relative to the images in the first two panels are due to the wider aperture and slower shutter speed that were employed during photography to compensate for the reduced fluorescence intensity. Cartoon representations in the second column and arrow diagrams in the third column illustrate the orientation of the ice hemispheres.

one of the *a*-axes (viewing the prism or pyramidal planes) show uniform binding of the GFP-tagged large isoform (Figure 9, top panels) on all hemisphere surfaces. The same result was observed when the concentration of the GFP large isoform employed in the solution for ice hemisphere growth was reduced by one-quarter (Figure 9, bottom panel).

To further document the ice plane binding of the large isoform, we also examined microscopic ice crystals formed in a solution of the GFP-tagged large isoform using a confocal fluorescence microscope. The solution was rapidly frozen and then slowly melted back until separate, single crystals ~10−20 μm in size were formed (Figure 10A). The crystals adopted a rice grain morphology similar to the morphology of those formed in a solution of non-GFP-labeled protein. The representative crystal shown in this figure is tilted on its side with the *c*-axis in the plane of the image. The pattern of fluorescence on the crystal indicates clearly that the GFP large isoform binds on all surfaces of the crystal. Importantly, there appears to be slightly more protein bound on and nearby the basal plane than elsewhere. Furthermore, when the temperature is lowered, ice crystals occasionally grow along the *a*-axes to expose long segments of the basal plane. Substantial binding of the protein to the basal plane of these crystals is similarly observed (Figure 10B).

DISCUSSION

In this study, we have determined after considerable difficulty the sequence of the large isoform of snow flea AFP. This enabled us to model the 3D structure of the protein by adopting the strategy used on the small isoform. The recombinant protein was successfully expressed and its hyperactive antifreeze activity confirmed. By tagging the protein with GFP, we found that the large isoform binds to multiple planes of ice, including the basal plane, which supports the correlation between hyperactivity and basal plane binding.

Of the AFPs identified to date, five hyperactive AFPs have been structurally characterized: three from arthropods (*12*, *13*, *20*), one from an Antarctic bacterium (*14*), and one from winter flounder (*27*). The large AFP isoform from the snow flea has several structural attributes in common with these other hyperactive AFPs. The molecule is amphipathic and has at least one relatively flat surface. The more hydrophobic side, which is the putative ice-binding site, possesses some regularity due to ridges formed by aligned alanine side chains and presents a large percentage of the protein's surface area for binding to ice. These shared structural features lend credence to our model of the large isoform.

A further piece of evidence supporting the model is the location of the two cysteines in the large isoform. There are four of these residues in the small isoform that are linked via coils 1 and 3 and coils 2 and 4. The first of these is the N-terminal residue, which is also present in the large isoform. The other cysteine in the large isoform is in a position equivalent to that of the third cysteine, and not the second or fourth. Thus, there was a one in three chance that the cysteine pair would be the same in both isoforms. The 1−3 pairing (between the top of coil 1 and the top of coil 3) offers support for the model because it is the only combination that is allowed by the modeled fold.

The putative hydrophobic face on the large isoform is flat and contains a regular pattern of five equally spaced ridges formed by the hydrophobic side chains of coils 2, 4, 6, 10, and 12. We propose that these features make the hydrophobic face of the protein the putative ice-binding site. A flat surface is common to the ice-binding sites of all AFPs, including the globular type III AFPs (*28*, *29*), the α-helical type I AFPs from winter flounder and shorthorn sculpin (*30*, *31*), and the β-helical insect AFPs (*12*, *13*). In addition, the most hydrophobic regions on AFPs are often found to be the ice-binding site, including those on type III AFPs and type I AFPs (*31*, *32*). Finally, regularly spaced surface features on putative AFP ice-binding sites similar to ones observed on snow flea AFP have been modeled to make highly complementary fits to the lattice structure on the ice planes to which AFPs bind (*13*, *33*, *34*).

While recombinant expression of the small isoform of snow flea AFP has been fraught with difficulty due to its thermolability and very low yields of active protein (*18*), it has been possible to produce moderate amounts of large isoform protein for TH and ice binding analysis. One possible reason for this may be the presence of just one disulfide bond in the large isoform rather than two in the small isoform. The large isoform is also likely to be more stable than the small isoform because of its more extensively hydrogen bonded core. In contrast to the latter, the tripeptide repeats for every coil except two in the large isoform have glycine in the first and second positions. This allows hydrogen bonding contacts all down the coils from two of the three peptide bonds per repeat. This proposal is consistent with the low total energy observed in the molecular dynamics simulation of the large isoform and the relatively stable number of hydrogen bonds formed for the entire length of the simulation.

It is also noteworthy that the large isoform is at least 2-fold more active in thermal hysteresis than the small isoform (*15*). Again, the structure of the large isoform offers an explanation for

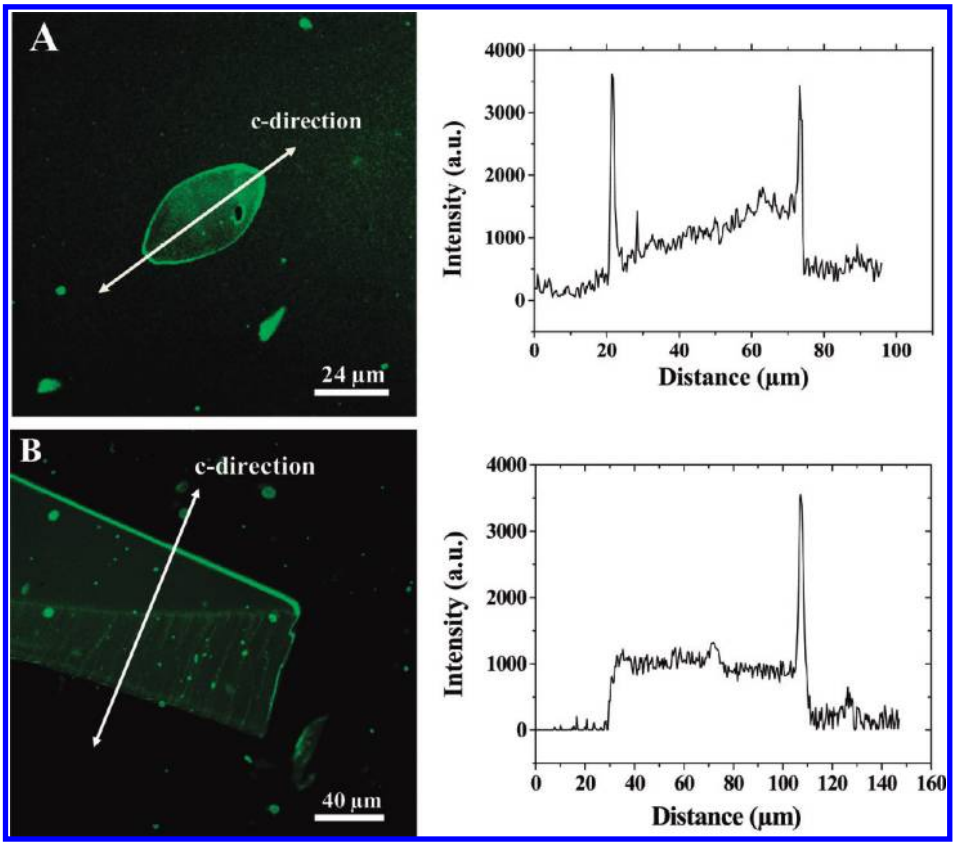

FIGURE 10: Fluorescent ice crystals formed in the presence of the GFP-tagged large isoform of snow flea AFP. (A) The left panel shows the grain-shaped ice crystal with bound protein on its surface. The *c*-axis is parallel to the plane of the page, as indicated. The right panel shows a plot of the fluorescence intensity along the white arrow shown in the left panel. Peak intensities correspond to the fluorescence at the basal planes. (B) The left panel shows the ice crystal after the temperature has been further lowered, showing bound protein along the basal plane. The *c*-axis is parallel to the plane of the page, as indicated. The right panel shows a plot of the fluorescence intensity along the white arrow shown in the left panel. The peak intensity corresponds to the fluorescence on the basal plane.

this observation. The area of the putative ice binding site in the large isoform is approximately double that found in the small isoform (*18*). It has been generally observed that the larger isoforms of natural AFPs are more active than the small ones. This was originally noted for the repetitive glycoprotein antifreezes (*35*, *36*). Rock cod, for example, contains a large number of glycoprotein antifreeze isoforms, and those with higher molecular masses are consistently several times more active than those with lower molecular masses. Even relatively small increases in area can produce severalfold differences in antifreeze activity. For instance, the α-helical type I AFP typically has three 11-amino acid repeats. An isoform with four 11-amino acid repeats has 2−3 times the activity of the three 11-amino acid repeat isoforms (*37*). Similarly, a natural variant of the spruce budworm AFP with two additional 15-amino acid repeats that extend its ice-binding surface has a similarly enhanced activity (*26*). There are other examples in which the area of the ice-binding site has been manipulated to study its effect on activity, as in the addition and deletion of 12-amino acid repeats from the beetle AFP (*38*). Via addition of three or four coils to a minimized beetle β-helical AFP, the ice-binding surface was expanded and activity was increased up to 100-fold. Similar experiments performed with dimers and oligomers of type III AFP (*39*, *40*) are particularly instructive in light of the doubling of the ice-binding area in the large isoform to the small isoform. Tandem duplication of the ice-binding area in type III AFP through dimerization of the recombinant protein effectively doubles its TH activity, while inactivation of one of the ice-binding sites reduces its activity to close to that of the monomer (*26*).

Macroscopic ice hemisphere and microscopic ice crystal experiments both show that the large isoform directly adheres to all surfaces of ice, including prism and basal planes. It has been proposed that a key difference between hyperactive and nonhyperactive AFPs is that the former bind the basal plane of ice whereas the latter do not, and that this behavior underlies the hyperactivity observed in certain AFPs (*16*). Ice etching experiments show no evidence of basal plane binding by nonhyperactive fish AFPs and antifreeze glycoproteins (*24*, *41*). On the other hand, direct binding to the basal plane has been visualized in hyperactive AFPs from spruce budworm (*41*), *Tenebrio* beetles, and the Antarctic bacterium *Marinomonas primoryensis* (C. Garnham, personal communication). We propose that the binding of the large isoform of snow flea AFP to the basal plane is the basis for its hyperactivity. The structure of snow flea AFP, however, is clearly divergent from the β-helical rolls of the beetle, spruce budworm, and Antarctic bacterium AFPs (*12*, *14*, *42*) as well as the extended α-helical bundle of the hyperactive type I AFP (*43*), and it will be interesting to see how the ice-binding sites of these very different structures can all match multiple planes of ice.

Comparison of the large isoform sequence with that of the small isoform shows a very low level of sequence identity at the amino acid and nucleotide levels. This explains the failure to detect the long isoform sequence in the library by hybridization using the small isoform sequence as a probe. It also suggests that

the separation between the large and small isoforms is ancient. Of note is the observation that the small isoform is considerably more abundant in snow flea homogenate than the large isoform (*15*). It is not clear why two isoforms with significantly different levels of activity have evolved in the snow flea. However, there are other systems in which large and small AFP isoforms have coevolved and coexist. For example, there are the large isoforms of the antifreeze glycoproteins in both the Northern cods and Antarctic Nototheniids that are accompanied by larger quantities of much less active smaller isoforms. Similarly, in the winter flounder, a hyperactive 32 kDa large isoform is present along with much higher concentrations of the three- and four-repeat type I AFP isoforms. Previous studies have indicated that multiple AFP isoforms can act cooperatively to enhance the overall TH activity (*44*, *45*), although the mechanism by which such synergy arises is not understood. The work presented here extends previous structural studies of the small isoform of snow flea AFP to the more active large isoform and provides a foundation for further investigation into their mechanism of ice binding and how they may act in concert to produce antifreeze hyperactivity.

ACKNOWLEDGMENT

We are extremely grateful to Sherry Gauthier for technical assistance.

SUPPORTING INFORMATION AVAILABLE

DNA sequence encoding the large isoform of snow flea AFP and a table of the observed and predicted tryptic peptide masses of the protein from MS sequencing. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Duman, J., Bennett, V., Sformo, T., Hochstrasser, R., and Barnes, B. (2004) Antifreeze proteins in Alaskan insects and spiders. *J. Insect Physiol. 50*, 259–266.
2. Fletcher, G., Hew, C., and Davies, P. (2001) Antifreeze proteins of teleost fishes. *Annu. Rev. Physiol. 63*, 359–390.
3. Jia, Z., and Davies, P. (2002) Antifreeze proteins: An unusual receptor-ligand interaction. *Trends Biochem. Sci. 27*, 101–106.
4. Ewart, K. V., Lin, Q., and Hew, C. L. (1999) Structure, function and evolution of antifreeze proteins. *Cell. Mol. Life Sci. 55*, 271–283.
5. Doucet, C. J., Byass, L., Elias, L., Worrall, D., Smallwood, M., and Bowles, D. J. (2000) Distribution and characterization of recrystallization inhibitor activity in plant and lichen species from the UK and maritime Antarctic. *Cryobiology 40*, 218–227.
6. Fletcher, G. L., Kao, M. H., and Fourney, R. M. (1986) Antifreeze peptides confer freezing resistance to fish. *Can. J. Zool. 64*, 1897–1901.
7. Rubinsky, B., Arav, A., and Fletcher, G. (1991) Hypothermic protection: A fundamental property of "antifreeze" proteins. *Biochem. Biophys. Res. Commun. 180*, 566–571.
8. Raymond, J., and DeVries, A. (1977) Adsorption inhibition as a mechanism of freezing resistance in polar fishes. *Proc. Natl. Acad. Sci. U.S.A. 74*, 2589–2593.
9. Hew, C. L., and Yang, D. S. (1992) Protein interaction with ice. *Eur. J. Biochem. 203*, 33–42.
10. Knight, C. A. (2000) Structural biology. Adding to the antifreeze agenda. *Nature 406*, 249–251.
11. Kristiansen, E., and Zachariassen, K. E. (2005) The mechanism by which fish antifreeze proteins cause thermal hysteresis. *Cryobiology 51*, 262–280.
12. Graether, S., Kuiper, M., Gagné, S., Walker, V., Jia, Z., Sykes, B., and Davies, P. (2000) β-Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature 406*, 325–328.
13. Liou, Y., Tocilj, A., Davies, P., and Jia, Z. (2000) Mimicry of ice structure by surface hydroxyls and water of a β-helix antifreeze protein. *Nature 406*, 322–324.
14. Garnham, C., Gilbert, J., Hartman, C., Campbell, R., Laybourn-Parry, J., and Davies, P. (2008) A Ca^{2+}-dependent bacterial antifreeze protein domain has a novel β-helical ice-binding fold. *Biochem. J. 411*, 171–180.
15. Graham, L., and Davies, P. (2005) Glycine-rich antifreeze proteins from snow fleas. *Science 310*, 461.
16. Scotter, A., Marshall, C., Graham, L., Gilbert, J., Garnham, C., and Davies, P. (2006) The basis for hyperactivity of antifreeze proteins. *Cryobiology 53*, 229–239.
17. Kuiper, M., Lankin, C., Gauthier, S., Walker, V., and Davies, P. (2003) Purification of antifreeze proteins by adsorption to ice. *Biochem. Biophys. Res. Commun. 300*, 645–648.
18. Lin, F., Graham, L., Campbell, R., and Davies, P. (2007) Structural modeling of snow flea antifreeze protein. *Biophys. J. 92*, 1717–1723.
19. Pentelute, B. L., Gates, Z. P., Dashnau, J. L., Vanderkooi, J. M., and Kent, S. B. (2008) Mirror image forms of snow flea antifreeze protein prepared by total chemical synthesis have identical antifreeze activities. *J. Am. Chem. Soc. 130*, 9702–9707.
20. Pentelute, B., Gates, Z., Tereshko, V., Dashnau, J., Vanderkooi, J., Kossiakoff, A., and Kent, S. (2008) X-ray structure of snow flea antifreeze protein determined by racemic crystallization of synthetic protein enantiomers. *J. Am. Chem. Soc. 130*, 9695–9701.
21. DeLano, W. L. The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org.
22. 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.
23. Chakrabartty, A., and Hew, C. (1991) The effect of enhanced α-helicity on the activity of a winter flounder antifreeze polypeptide. *Eur. J. Biochem. 202*, 1057–1063.
24. Knight, C., Cheng, C., and DeVries, A. (1991) Adsorption of α-helical antifreeze peptides on specific ice crystal surface planes. *Biophys. J. 59*, 409–418.
25. Pertaya, N., Marshall, C., DiPrinzio, C., Wilen, L., Thomson, E., Wettlaufer, J., Davies, P., and Braslavsky, I. (2007) Fluorescence microscopy evidence for quasi-permanent attachment of antifreeze proteins to ice surfaces. *Biophys. J. 92*, 3663–3673.
26. Leinala, E. K., Davies, P. L., Doucet, D., Tyshenko, M. G., Walker, V. K., and Jia, Z. (2002) A β-helical antifreeze protein isoform with increased activity. Structural and functional insights. *J. Biol. Chem. 277*, 33349–33352.
27. Graham, L. A., Marshall, C. B., Lin, F. H., Campbell, R. L., and Davies, P. L. (2008) Hyperactive antifreeze protein from fish contains multiple ice-binding sites. *Biochemistry 47*, 2051–2063.
28. Jia, Z., DeLuca, C., Chao, H., and Davies, P. (1996) Structural basis for the binding of a globular antifreeze protein to ice. *Nature 384*, 285–288.
29. Yang, D. S., Hon, W. C., Bubanko, S., Xue, Y., Seetharaman, J., Hew, C. L., and Sicheri, F. (1998) Identification of the ice-binding surface on a type III antifreeze protein with a "flatness function" algorithm. *Biophys. J. 74*, 2142–2151.
30. Baardsnes, J., Kondejewski, L., Hodges, R., Chao, H., Kay, C., and Davies, P. (1999) New ice-binding face for type I antifreeze protein. *FEBS Lett. 463*, 87–91.
31. Baardsnes, J., Jelokhani-Niaraki, M., Kondejewski, L., Kuiper, M., Kay, C., Hodges, R., and Davies, P. (2001) Antifreeze protein from shorthorn sculpin: Identification of the ice-binding surface. *Protein Sci. 10*, 2566–2576.
32. Sönnichsen, F., DeLuca, C., Davies, P., and Sykes, B. (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.
33. Davies, P., Baardsnes, J., Kuiper, M., and Walker, V. (2002) Structure and function of antifreeze proteins. *Philos. Trans. R. Soc. London, Ser. B 357*, 927–935.
34. Leinala, E. K., Davies, P. L., and Jia, Z. (2002) Crystal structure of β-helical antifreeze protein points to a general ice binding model. *Structure 10*, 619–627.
35. Schrag, J. D., O'Grady, S. M., and DeVries, A. L. (1982) Relationship of amino acid composition and molecular weight of antifreeze glycopeptides to non-colligative freezing point depression. *Biochim. Biophys. Acta 717*, 322–326.
36. Wu, Y., Banoub, J., Goddard, S. V., Kao, M. H., and Fletcher, G. L. (2001) Antifreeze glycoproteins: Relationship between molecular weight, thermal hysteresis and the inhibition of leakage from liposomes during thermotropic phase transition. *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 128*, 265–273.
37. Chao, H., Hodges, R. S., Kay, C. M., Gauthier, S. Y., and Davies, P. L. (1996) A natural variant of type I antifreeze protein with four ice-binding repeats is a particularly potent antifreeze. *Protein Sci. 5*, 1150–1156.

38. Marshall, C., Daley, M., Sykes, B., and Davies, P. (2004) Enhancing the activity of a β-helical antifreeze protein by the engineered addition of coils. *Biochemistry 43*, 11637–11646.
39. Baardsnes, J., Kuiper, M., and Davies, P. (2003) Antifreeze protein dimer: When two ice-binding faces are better than one. *J. Biol. Chem. 278*, 38942–38947.
40. Miura, K., Ohgiya, S., Hoshino, T., Nemoto, N., Suetake, T., Miura, A., Spyracopoulos, L., Kondo, H., and Tsuda, S. (2001) NMR analysis of type III antifreeze protein intramolecular dimer. Structural basis for enhanced activity. *J. Biol. Chem. 276*, 1304–1310.
41. Pertaya, N., Marshall, C., Celik, Y., Davies, P., and Braslavsky, I. (2008) Direct visualization of spruce budworm antifreeze protein interacting with ice crystals: Basal plane affinity confers hyperactivity. *Biophys. J. 95*, 333–341.
42. Liou, Y., Daley, M., Graham, L., Kay, C., Walker, V., Sykes, B., and Davies, P. (2000) Folding and structural characterization of highly disulfide-bonded beetle antifreeze protein produced in bacteria. *Protein Expression Purif. 19*, 148–157.
43. Marshall, C., Chakrabartty, A., and Davies, P. (2005) Hyperactive antifreeze protein from winter flounder is a very long rod-like dimer of α-helices. *J. Biol. Chem. 280*, 17920–17929.
44. Wang, L., and Duman, J. G. (2005) Antifreeze proteins of the beetle *Dendroides canadensis* enhance one another's activities. *Biochemistry 44*, 10305–10312.
45. Nishimiya, Y., Sato, R., Takamichi, M., Miura, A., and Tsuda, S. (2005) Co-operative effect of the isoforms of type III antifreeze protein expressed in Notched-fin eelpout, *Zoarces elongatus* Kner. *FEBS J. 272*, 482–492.