Structural Modeling of Snow Flea Antifreeze Protein

Feng-Hsu Lin,* Laurie A. Graham,* Robert L. Campbell,* and Peter L. Davies†
*Department of Biochemistry and the †Protein Function Discovery Group, Queen's University, Kingston, Ontario K7L 3N6, Canada

ABSTRACT The glycine-rich antifreeze protein recently discovered in snow fleas exhibits strong freezing point depression activity without significantly changing the melting point of its solution (thermal hysteresis). BLAST searches did not detect any protein with significant similarity in current databases. Based on its circular dichroism spectrum, discontinuities in its tripeptide repeat pattern, and intramolecular disulfide bonding, a detailed theoretical model is proposed for the 6.5-kDa isoform. In the model, the 81-residue protein is organized into a bundle of six short polyproline type II helices connected (with one exception) by proline-containing turns. This structure forms two sheets of three parallel helices, oriented antiparallel to each other. The central helices are particularly rich in glycines that facilitate backbone carbonyl-amide hydrogen bonding to four neighboring helices. The modeled structure has similarities to polyglycine II proposed by Crick and Rich in 1955 and is a close match to the polyproline type II antiparallel sheet structure determined by Traub in 1969 for $(Pro\text{-}Gly\text{-}Gly)_n$. Whereas the latter two structures are formed by intermolecular interactions, the snow flea antifreeze is stabilized by intramolecular interactions between the helices facilitated by the regularly spaced turns and disulfide bonds. Like several other antifreeze proteins, this modeled protein is amphipathic with a putative hydrophobic ice-binding face.

INTRODUCTION

A newly discovered antifreeze protein (AFP) from the snow flea, *Hypogastrura harveyi* Folsom, is thought to help these primitive arthropods to survive freezing temperatures by inhibiting the growth of ice (1). At subzero temperatures, uncontrolled ice growth can lead to the dehydration or rupture of cells, loss of tissue functions, and, ultimately, the organism's death. AFPs, also known as thermal hysteresis proteins (2), bind to a growing ice crystal surface (3) by an adsorption-inhibition mechanism (4) creating microcurvature of the ice front that makes it less favorable for water molecules to add to the crystal due to the Kelvin effect (5). In this way, AFP increases the energy requirement for ice growth and the resulting inhibition of growth prevents freezing damage. Although surface adsorption of AFPs to ice is generally accepted, the details of the antifreeze mechanism are uncertain. See the recent review by Prabhu and Sharp (6) for a discussion of some of the outstanding issues.

Two antifreeze isoforms were isolated from the snow flea, which differ in mass (15.7 kDa and 6.5 kDa) (1). Both of these proteins had potent thermal hysteresis activities of >2 C° at micromolar concentrations. They have very similar amino acid compositions. The short isoform (6.5 kDa) of snow flea AFP (sfAFP) contains 81 residues organized into a prominent repeat of Gly-x_1-x_2 where x_1 is often a glycine, and x_2 varies between a charged/hydrophilic residue and a small, hydrophobic residue (alanine/valine). This protein contains 37 glycines (45.7%) and the second most abundant residue is alanine (13.6%). The protein also contains four cysteines, which were found to form two disulfide bonds based on the mass difference after reduction and alkylation, although the bonding pattern is not known. Interestingly, the larger isoform has only two cysteines, and these are also disulfide bonded.

The amount of sfAFP that can be readily purified from natural sources is insufficient for conventional structure determination. It has proven difficult to collect even gram amounts of this organism (5–10,000 individuals/g) from which microgram quantities of AFP can be purified. Moreover, it is not easy to produce the AFP as a well-folded recombinant protein because of its unusual properties and thermal instability (data not shown). In the interim, we set out to model the protein structure ab initio.

METHODS

Model building

Progress toward the model was an intuitive, iterative process. It began with recognition of a tripeptide-repeating pattern throughout the protein that was highly suggestive of a threefold 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, splitting it into six roughly equal segments. Since the AFP is monomeric, a major constraint to be satisfied was the introduction of two intramolecular disulfide bonds. Lastly, an additional aspect of the tripeptide repeat, the irregular distribution of Gly-Gly, was taken into consideration.

A physical model was constructed using the HGS Biochemistry Molecular Model from Hinomoto Plastics (Tokyo, Japan).

Molecular dynamics

A virtual model was built using PyMOL 0.98 (7). The model was solvated in a 5.1 nm $\times$ 3.6 nm $\times$ 3.4 nm box of water containing 1776 waters and had a net single positive charge. To neutralize the charge of the system and to provide an effective concentration of 0.1M NaCl, nine molecules of water

Submitted July 14, 2006, and accepted for publication November 7, 2006.

Address reprint requests to Peter L. Davies, Dept. of Biochemistry, Queen's University, Kingston, ON K7L 3N6, Canada. Tel.: 613-533-2983; Fax: 613-533-2497; E-mail: daviesp@post.queensu.ca.

0006-3495/07/03/1717/07 $2.00

doi: 10.1529/biophysj.106.093435

were replaced by 4 Na^+ and 5 Cl^- ions. The system was subjected to energy minimization by steepest descents, position restrained molecular dynamics to relax the solvent, followed by full molecular dynamics. All molecular mechanics calculations were done with GROMACS 3.3 (8). GROMACS uses a triclinic unit cell for its periodic boundary conditions. For calculating short-range nonbonded interactions, only the nearest image is considered (8). Long-range electrostatics is treated with the particle mesh Ewald method. The simulations were done under isothermal conditions using Berendsen temperature coupling. The GROMOS96 43a1 force field was chosen. The simulations were done independently at 4°C and 25°C using a timestep of 2 fs for a total duration of 10 ns at each temperature.

RESULTS AND DISCUSSION

Model prediction

The most prominent features of this protein are its glycine-rich composition and its organization into Gly-x_1-x_2 tripeptide repeats as shown in Fig. 1 *A*. It is notable that the sfAFP sequence is periodically interrupted by prolines that accompany a minor change in the tripeptide repeat pattern. For example, the tripeptide repeats around Pro-41 (Gly-Asn-Pro and Gly-Cys-Ala) break a long series of Gly-Gly-x_2 repeats. Based on the distribution of prolines, the AFP sequence can be split into five segments: four of approximately equal length and one segment that is twice the length of the others. The longer segment contains an irregularity starting at Ser-54 that disrupts the tripeptide repeat pattern at the middle of the sequence between Pro-41 and Pro-71, shown in Fig. 1, *A* and *B*. Thus the repetitive sfAFP sequence is interrupted five times and can be broken down into six segments, each comprising three to five Gly-x_1-x_2 repeats (Fig. 1 *B*).

In our model, each G-x_1-x_2 repeat makes a helical turn consisting of exactly three residues with the hydrogen-bonding groups of the peptide bonds facing outward from the helical axis. This is graphically represented in Fig. 2 *A*. It is proposed that the six segments form individual polyproline type II (PPII) regions and make backbone-backbone hydrogen-bonding interactions with each other. Segments 1, 3, and 5 are parallel (forming one face of the protein) and antiparallel to segments 2, 4, and 6 (which form the other face) (Fig. 2 *B*). The consecutive segments are connected by loops that change the direction of the polypeptide chain. This arrangement brings two pairs of Cys close enough for disulfide bonding. These are Cys-1 and Cys-29, which link the N-terminal end of segment 1 to the loop between segments 2 and 3, and Cys-13 and Cys-44, which link the loops between segments 1 and 2, and segments 3 and 4.

Although the core sequences of each segment contain G-x_1-x_2 repeats, there are noticeable preferences for the x residues within each segment that differ from one to another. Segments 3, 4, and 5 have Gly in the second position (x_1), whereas the other segments typically have Ala (Fig. 1 *B*). Central segments 3 and 4 are in contact with four other strands and it is critical for the fold that they contain two glycine-rich faces. Having Gly in the second position (x_1) ensures that there is no steric barrier to making hydrogen

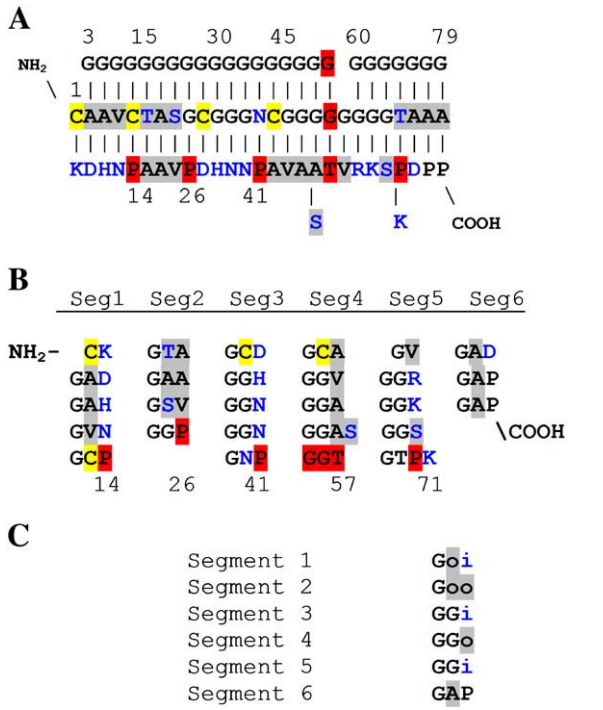

FIGURE 1 (*A*) Amino acid sequence of the 81-residue isoform of sfAFP. Red background indicates residues predicted to disrupt the PPII helix. Yellow background denotes cysteine residues. Gray background marks small residues including alanine, valine, serine, and threonine. Blue letters are charged or hydrophilic residues including arginine, lysine, aspartic acid, asparagine, serine, threonine, and histidine. (*B*) Segments arranged in columns. (*C*) Consensus repeat in each of the six segments. The symbol o represents small residues. The symbol i represents hydrophilic, large residues and some prolines.

bonds to the neighboring PPII polypeptide segments that abut two-thirds of their surface area (Fig. 2 *A*). Glycine-rich surfaces are buried in the core of the bundle and are not exposed to the solvent.

The third position (x_2) is generally occupied by hydrophilic residues in segments 1, 3, and 5 and by small hydrophobic residues in segments 2, 4, and 6 (Fig. 1, *B* and *C*). 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 relatively hydrophobic, whereas the odd-numbered segments form a relatively hydrophilic surface (Fig. 2 *A*).

Circular dichroism spectrum analysis

The only biophysical evidence we have for this model (given the tiny amount of natural protein available) is its circular dichroism (CD) spectrum (Fig. 3). Deconvolution of this CD spectrum suggested a significant amount of random coil (1). However, Sreerama and Woody (9) have pointed out the difficulties of deconvoluting CD spectra of PPII-containing structures and in particular of distinguishing random coil or

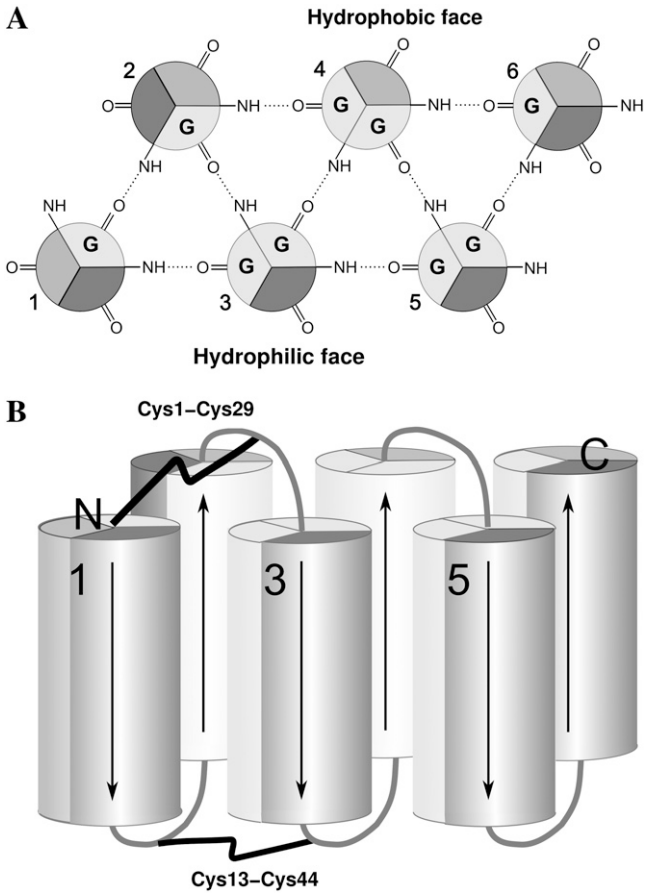

FIGURE 2 PPII-like fold of sfAFP. (*A*) Top view, with hydrogen-bonding pattern shown as dotted lines. The segments are numbered 1–6 and divided into thirds to represent the stacking of amino acids. G represents glycine stacks. Hydrophobic residue stacks are shaded light gray, and hydrophilic residue stacks are shaded dark gray. (*B*) Side view with disulfides drawn as black lines and arrows to illustrate the alternating direction of the segments. Loops connecting segments are colored gray.

unstructured protein from PPII helix structure. Interestingly, there is increasing evidence that what has traditionally been considered as random coil may instead contain a significant amount of PPII structure (10). The sfAFP CD spectrum is similar to those of other PPII-type structures (Fig. 3), especially since different peptides that contain PPII structure display a variety of minima and maxima in their CD spectra. Therefore, although we cannot conclude from the CD spectrum that sfAFP is predominantly formed of PPII helices, we can say the CD spectrum is consistent with such a structure.

Similarity of the sfAFP model to synthetic glycine-rich protein structures

The glycine-rich composition of sfAFP means that this natural protein has considerable sequence identity to both poly-(ProGlyGly) and poly-glycine, two synthetic polymers. The polyglycine II structure, which was first proposed in 1955 by Crick and Rich (11) is made up of parallel

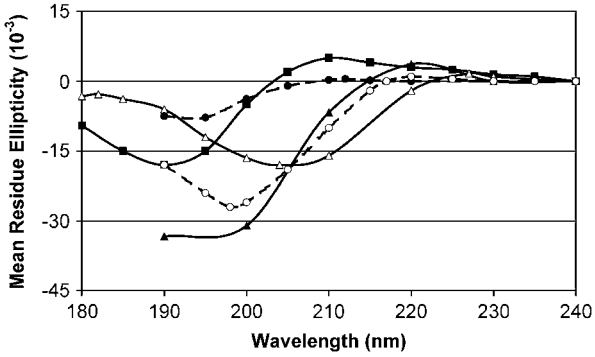

FIGURE 3 CD spectra of proteins proposed to have a PPII fold. Data were compiled for the small isoform of sfAFP at 4°C (1)—solid squares on a solid line; AFGP-8, which is the smallest of the naturally occurring AFGPs with four tandem repeats of the glycotripeptide unit (7)—solid triangles on a solid line; collagen CB2-MG, which is a type I collagen trimer modified to form a homogenous solution when dissolved (24)—open circles on a dashed line; polyproline (25)—open triangles on a solid line; GAGG repeat at 0°C (26) solid circles on a dashed line. The amplitude of the published spectrum for AFGP-8 was reduced by a factor of three to be consistent with mean residue ellipticity.

polyglycine chains that have a threefold screw axis with a vertical repeat (pitch) of 9.3 Å. The lack of side chains allows the main chains to be packed into a hexagonal array with each chain hydrogen-bonded to six neighbors through backbone amide groups (C=O and N-H) that project normal to the helical axis. The lack of chirality around the Cα atom means that an antiparallel chain could fit into the bundle and still make the same hydrogen-bonding connections to its neighbors. In contrast, the poly-(ProGlyGly) structure has more constraints due to the presence of Pro in every third position (12). It has characteristics of both the polyglycine II structure and the PPII structure, which is a left-handed helix with 3.0 residues per turn. Poly-(ProGlyGly) is also a left-handed helix with 3.0 residues per turn. It differs from polyglycine II in being only able to form two backbone-backbone NH. . .O hydrogen bonds per tripeptide to neighboring helical chains. This constrains the structure into double-layered sheets where the polarity of the strands in each sheet is the same but antiparallel to the other sheet. The crucial difference between the poly-(ProGlyGly) fold and that of the sfAFP model is that the former is held together by intermolecular interactions and the latter's interactions are intramolecular, including two pairs of disulfide bonds.

sfAFP does not adopt a collagen-like structure

The skewed amino acid composition of sfAFP means that it also has considerable sequence similarity to collagen (13). Indeed, the CD spectra of sfAFP and collagen are alike, although the AFP spectrum is blue shifted (Fig. 3). The spectral similarities likely reflect similar backbone conformations. As in collagen, our model makes a left-handed helix with 3.0 residues per turn and hydrogen bonds to neighboring

helices. However, in collagen three identical chains are coiled around each other with a common axis to form the collagen triple-helix. The contacting surface of the three collagen monomers is composed of glycine from the G-x-P repeat. This allows a compact structure to form between coils due to low steric hindrance.

We considered the theoretical possibility that sfAFP adopts a collagen-like triple-helix fold. However coiling three sfAFP around each other to form the triple-helix would obviously position the cysteines too far apart for intrachain disulfide bonding. Mass spectrometry clearly shows that sfAFP makes two intrachain disulfide bonds, with no hint of interchain Cys-Cys connections (1).

To investigate the possibility that sfAFP forms an intramolecular triple-helix, the protein was subdivided into three segments, 1 + 2, 3 + 4, and 5 + 6. Three-segment collagen folding would require the middle segment to run antiparallel to the other two, whereas natural collagen monomers polymerize in a parallel fashion. This model would not be compatible with disulfide-bond formation. The spacing between the cysteines does not allow any combination of disulfide bonding to form in a three-segmented sequence. A scenario where the six segments form two independent collagen-like folds seems unlikely for similar reasons. We conclude from this analysis that sfAFP does not adopt a collagen-like fold.

Comparison with antifreeze glycoprotein

Another naturally occurring protein with which sfAFP can be usefully compared is antifreeze glycoprotein (AFGP). AFGPs exist as multiple isoforms of varying lengths of tandem repeats of the consensus sequence AlaAlaThr with the threonine glycosylated by a disaccharide moiety (Galb1-3GalNAca1-). Although the sequence itself is not similar to that of sfAFP, it does contain a tripeptide repeat. In the NMR structure analysis of synthetic AFGPs by Nishimura's lab, this simple tripeptide repeat assumes a PPII-like fold with the glycosylated threonine side chain extending out on one side and the two alanines on the opposite side of the PPII barrel (14). This structure generates an amphipathic molecule where the alanine side chains form the hydrophobic face and the disaccharides make a more hydrophilic surface. AFGPs exhibit CD spectra resembling those of collagen and sfAFP, albeit with a much greater amplitude in the 190–210-nm range (Fig. 3).

Details of the sfAFP model

The conceptual model was built into a physical model to determine three-dimensional coordinates. The final minimized structure is represented in stereo in Fig. 4 and in space-filling mode in Fig. 5. The modeled polypeptide fold closely matches the fold proposed by Traub for poly-(ProGlyGly) but is limited to the six internally hydrogen-bonded PPII helices. In the model, an extensive regular

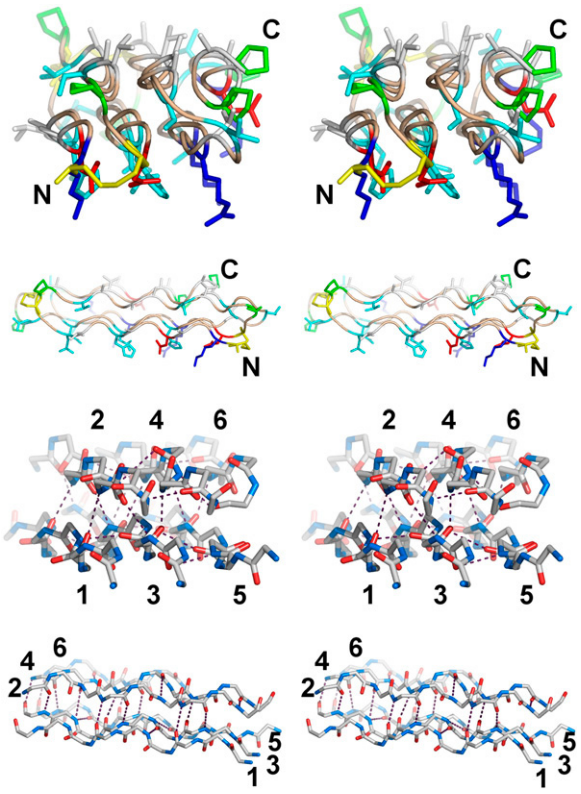

FIGURE 4 Overall structure of sfAFP shown as stereo views. (*Upper two panels*) Representation of the structure as a tube backbone with side chains. Positively and negatively charged residues are colored blue and red, respectively. Uncharged polar residues are cyan, prolines are green, cysteines are yellow, hydrophobic residues are white, and glycines are tan. The N- and C-termini are indicated by N and C, respectively. (*Lower two panels*) Stick representation of the highly repetitive backbone in stereo. Intrasegment hydrogen bonds are shown as dotted lines. Segments are numbered 1–6. Red atoms are oxygens; blue atoms are nitrogens. (Note that only helices are represented; the loops have been left out.)

pattern of hydrogen bonding can be observed between helices as shown in Fig. 4. As indicated earlier, helices 3 and 4 each make hydrogen bonds to four neighboring helices, due to their high glycine content. Throughout molecular dynamics simulations, especially at 4°C, Lys-2 is predominantly in proximity to Asp-5 and Lys-72 is in proximity to Asp-75. This suggests the potential for salt links between these pairs of residues with an i, $i + 3$ relationship, corresponding to one turn down the PPII helices. All potential salt links are located on the hydrophilic side of the AFP. During simulations, the loops connecting the PPII helices are flexible relative to the helices themselves, but their movements do not distort the flat surfaces formed by the core of the structure.

Predicted ice-binding site

When viewed in space-filling mode, the energy-minimized structure displays a hydrophilic surface on one side (segments

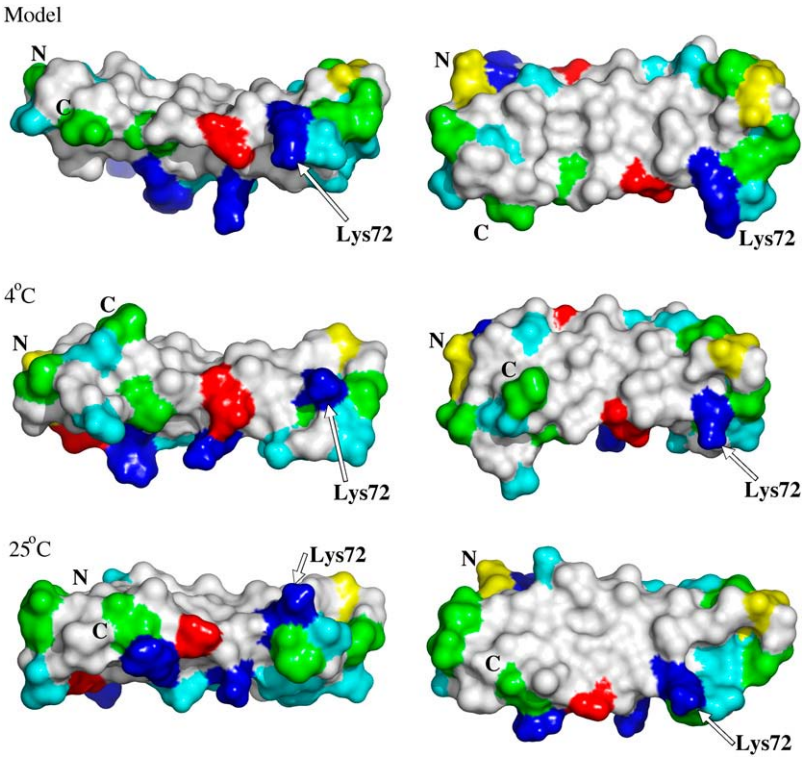

FIGURE 5 Surface representation of sfAFP before and after molecular dynamics at 4°C and 25°C. The left-hand panels present a side view facing segments 5 and 6, and the right-hand panels display the hydrophobic (putative ice-binding) surface (of segments 2, 4, and 6) to the front. The top panels are the modeled structure before molecular dynamics. The middle and bottom panels show the modeled structure after molecular dynamics at 4°C and 25°C, respectively. Residues are colored as described in the legend to Fig. 4. Asp-5 and Lys-72 are marked for orientation. N- and C-termini are indicated by N and C, respectively.

1, 3, and 5), and a relatively flat, hydrophobic surface on the other (segments 2, 4, and 6) (Fig. 4). These features were maintained throughout the molecular dynamics simulation at 4°C. The putative hydrophobic face contains a regular pattern of three equally spaced ridges formed by the hydrophobic side chains of segments 2, 4, and 6. These attributes of the hydrophobic surface of sfAFP make it quite likely to be the ice-binding site. Flatness is common to all AFP ice-binding sites that have been defined in detail beginning with type III AFP (15,16) and extending to the α-helical type I AFPs of winter flounder (17) and shorthorn sculpin (18) and the insect β-helical AFPs (19,20). As first pointed out by Sönnichsen et al. (21), the ice-binding site is the most hydrophobic face of the protein and in some cases is very hydrophobic, like the alanine-rich ice-binding surface of winter flounder and shorthorn sculpin type I AFPs (18). For the structurally repetitive AFPs, regularly spaced surface features on the ice-binding site have been modeled to make snug fits to complementary surface features on the planes of ice to which the AFPs bind (19,22,23).

Molecular dynamics: 25°C simulation

During the 10-ns molecular dynamic simulation of the model at 25°C, the loop between segments 4 and 5 gradually breaks away from the hydrogen-bonded core. In addition, because the hydrogen bonding between internal glycines becomes more irregular than in the 4°C structure, the pattern of ridges of hydrophobic residues becomes blurred (Fig. 5). This result is consistent with the sfAFP's lack of stability at room temperature (data not shown). Despite the increased flexibility of the backbone fold in the 25°C model, salt links suggested by the 4°C simulation appear to be maintained.

Based on the simulations, the total energy of each system was plotted, and hydrogen bonding between peptide backbones was plotted in Fig. 6. Despite the fact that the total energy of the system and hydrogen bonding within the backbone chain remained constant throughout the simulations, significant differences can be observed between the structures at 4°C and 25°C. We suggest that the conformation(s) adopted at 25°C results in loss of ice surface complementarity at the ice-binding face. Moreover, the abundance of hydrogen bonding at 25°C makes it likely that this conformation will remain stable (locked in place) at lower temperatures, and it might account for some irreversible loss in thermal hysteresis after the protein is heated to 25°C.

Ramachandran plot analysis

A Ramachandran plot of the energy-minimized model was compared to that of synthetic collagen (Fig. 7). The majority of ϕ and ψ angles for the model are distributed in the most favored region of the upper left quadrant, which is where those of the optimized collagen monomer cluster. The tight clustering of the collagen data points reflects the extreme regularity of this polymer. Predicted points for sfAFP were more scattered and differ slightly from the synthetic collagen

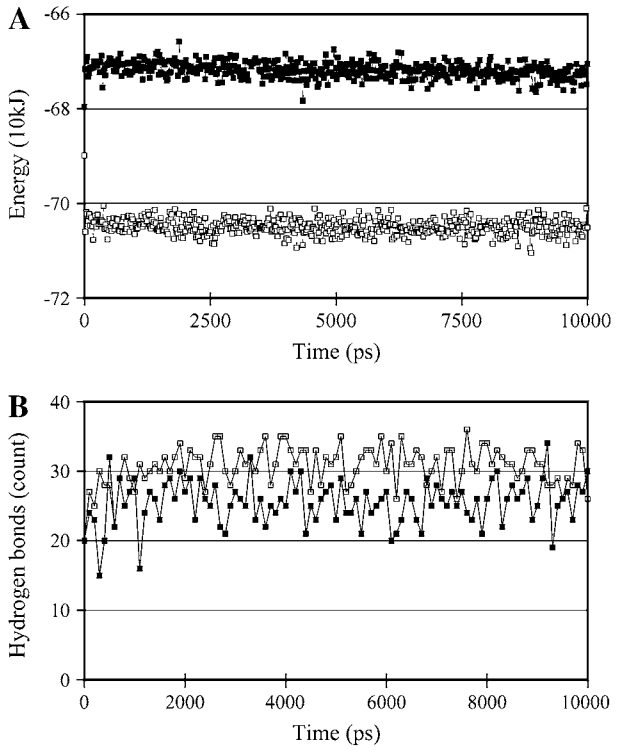

FIGURE 6 Analysis of the molecular dynamics simulations showing (*A*) total energy, and (*B*) number of hydrogen bonds formed at each time step. Solid squares indicate the model during 25°C simulation and open squares during the 4°C simulation.

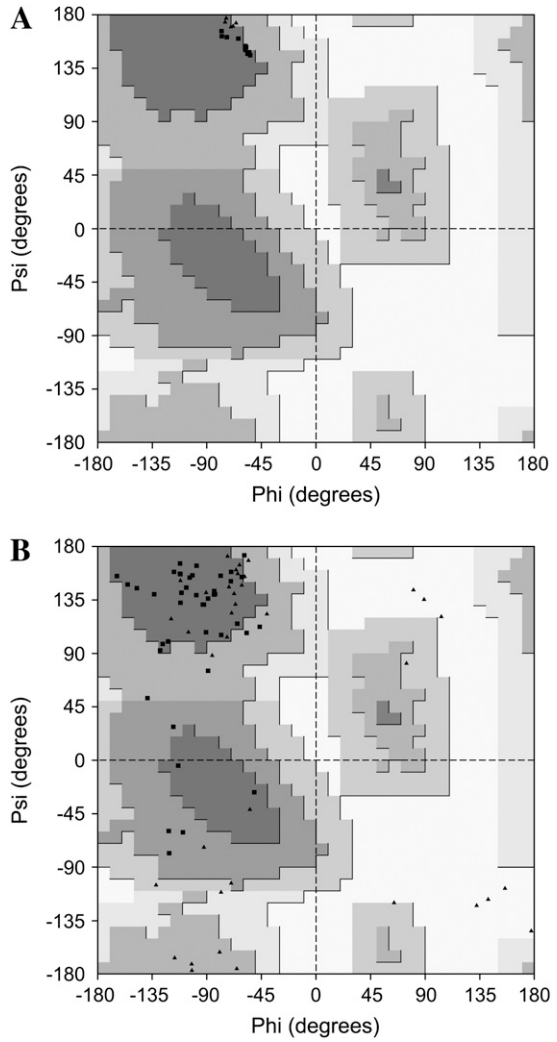

FIGURE 7 Ramachandran plots. (*A*) Plot of synthetic collagen model. Solid triangles represent glycines. (*B*) Plot of energy-minimized sfAFP model (27).

peptides. The bottom left quadrant residues include some glycines that are located close to the end of the helical segments. In less favored regions of the two right quadrants are residues that are present in the loops of the predicted structure. The flexibility of these loops at both 4°C and 25°C is illustrated in Fig. 8 by a plot of root mean-square fluctuation for each residue in the sequence. Fluctuation is greatest for residues in the loops and least for residues in the middle of the coiled segments. As mentioned earlier, the loop between segments 4 and 5 is particularly unstable and shows the greatest root mean-square fluctuation of all the loops.

SUMMARY OF THE SFAFP MODEL

The 81-residue glycine-rich AFP from snow fleas has been modeled as a novel fold in which six short segments of PPII helix alternate direction to form two antiparallel sheets (Fig. 4). The 180° change in direction of the polypeptide chain is facilitated at four out of the five turns by prolines, (Figs. 1, 4, and 5). The putative structure is held together by hydrogen bonding between backbone peptide amides. These bonds are normal to the axis of the helix and are of the appropriate length because the internal-facing residues are glycines that permit the individual helical segments to pack closely. In the model, the structure is reinforced by i), two disulfide bonds that link the N-terminal end of segment 1 to the turn between segments 2 and 3, and the turn between segments 1 and 2 to the turn between segments 3 and 4; and ii), by potential salt bridges on the hydrophilic side of the protein.

The modeled protein has dimensions 47 Å × 17 Å × 13 Å and has distinct amphipathic character. As an AFP, the hydrophilic surface formed by the outer residues of segments 1, 3, and 5 is most likely to present to the solvent, and the hydrophobic surface formed by the outer residues of segments 2, 4, and 6 will bind to ice. On this hydrophobic surface, there are projections with regular spacing that offer the potential to interact with a crystalline surface with complementary spacing.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

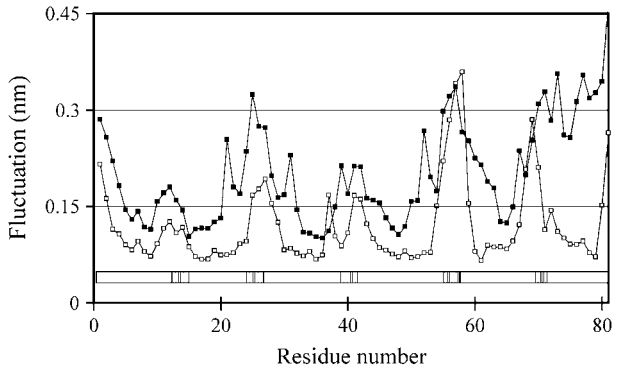

FIGURE 8 Root mean-squared fluctuation of each atom of the sfAFP model during 4°C and 25°C simulations. Solid squares represent the model during 4°C simulation and open squares the model during 25°C simulation. In the bar diagram below the plot, blank regions correspond to sequences predicted to be in the PPII fold. Shaded regions represent the locations of predicted loops.

We are very grateful to Dr. Christopher B. Marshall for constructive criticisms of the model and manuscript.

This research was funded by a grant to P.L.D. from the Canadian Institutes for Health Research. P.L.D. holds a Canada Research Chair in Protein Engineering.

REFERENCES

1. Graham, L. A., and P. L. Davies. 2005. Glycine-rich antifreeze proteins from snow fleas. *Science*. 310:461.
2. Kristiansen, E., and K. E. Zachariassen. 2005. The mechanism by which fish antifreeze proteins cause thermal hysteresis. *Cryobiology*. 51:262–280.
3. Knight, C. A., C. C. Cheng, and A. L. DeVries. 1991. Adsorption of alpha-helical antifreeze peptides on specific ice crystal surface planes. *Biophys. J.* 59:409–418.
4. Raymond, J. A., and A. L. DeVries. 1977. Adsorption inhibition as a mechanism of freezing resistance in polar fishes. *Proc. Natl. Acad. Sci. USA*. 74:2589–2593.
5. Knight, C. A., and A. Wierzbicki. 2001. Adsorption of biomolecules to ice and their effects upon ice growth. 2. A discussion of the basic mechanism of ''antifreeze'' phenomena. *Cryst. Growth Des.* 1:439–446.
6. Prabhu, N., and K. Sharp. 2006. Protein-solvent interactions. *Chem. Rev.* 106:1616–1623.
7. DeLano, W. L. 2002. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA. http://www.pymol.org.
8. Lindahl, E., B. Hess, and D. van der Spoel. 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model. (Online).* 7:306–317.
9. Sreerama, N., and R. W. Woody. 2004. Computation and analysis of protein circular dichroism spectra. *Methods Enzymol.* 383:318–351.
10. Rath, A., A. R. Davidson, and C. M. Deber. 2005. The structure of ''unstructured'' regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. *Biopolymers*. 80:179–185.
11. Crick, F. H. C., and A. Rich. 1955. Polyglycine II. *Nature*. 4486:780–781.
12. Traub, W. 1969. Polymers of tripeptides as collagen models. V. An x-ray study of poly(L-prolyl-glycyl-glycine). *J. Mol. Biol.* 43:479–485.
13. Creamer, T. P., and M. N. Campbell. 2002. Determinants of the polyproline II helix from modeling studies. *Adv. Protein Chem.* 62: 263–282.
14. Tachibana, Y., G. L. Fletcher, N. Fujitani, S. Tsuda, K. Monde, and S. Nishimura. 2004. Antifreeze glycoproteins: elucidation of the structural motifs that are essential for antifreeze activity. *Angew. Chem. Int. Ed. Engl.* 43:856–862.
15. 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.
16. Yang, D. S., W. C. Hon, S. Bubanko, Y. Xue, J. Seetharaman, C. L. Hew, and F. Sicheri. 1998. Identification of the ice-binding surface on a type III antifreeze protein with a ''flatness function'' algorithm. *Biophys. J.* 74:2142–2151.
17. Baardsnes, J., L. H. Kondejewski, R. S. Hodges, H. Chao, C. Kay, and P. L. Davies. 1999. New ice-binding face for type I antifreeze protein. *FEBS Lett.* 463:87–91.
18. Baardsnes, J., M. Jelokhani-Niaraki, L. H. Kondejewski, M. J. Kuiper, C. M. Kay, R. S. Hodges, and P. L. Davies. 2001. Antifreeze protein from shorthorn sculpin: identification of the ice-binding surface. *Protein Sci.* 10:2566–2576.
19. Liou, Y. C., A. Tocilj, P. L. Davies, and Z. Jia. 2000. Mimicry of ice structure by surface hydroxyls and water of a beta-helix antifreeze protein. *Nature*. 406:322–324.
20. Graether, S. P., M. J. Kuiper, S. M. Gagne, V. K. Walker, Z. Jia, B. D. Sykes, and P. L. Davies. 2000. Beta-helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature*. 406:325–328.
21. 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.
22. Davies, P. L., J. Baardsnes, M. J. Kuiper, and V. K. Walker. 2002. Structure and function of antifreeze proteins. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357:927–935.
23. Leinala, E. K., P. L. Davies, and Z. Jia. 2002. Crystal structure of beta-helical antifreeze protein points to a general ice binding model. *Structure*. 10:619–627.
24. Consonni, R., L. Zetta, R. Longhi, L. Toma, G. Zanaboni, and R. Tenni. 2000. Conformational analysis and stability of collagen peptides by CD and H- and C-NMR spectroscopies. *Biopolymers*. 53: 99–111.
25. Jeness, D. D., C. Sprecher, and W. Curtis. 1976. Circular dichroism of collagen, gelatin, and poly(proline) II in the vacuum ultraviolet. *Biopolymers*. 15:513–521.
26. Bochicchio, B., A. Pepe, and A. M. Tamburro. 2005. Circular dichroism studies on repeating polypeptide sequences of abductin. *Chirality*. 17:364–372.
27. Okuyama, K., C. Hongo, R. Fukushima, G. Wu, H. Narita, K. Noguchi, Y. Tanaka, and N. Nishino. 2004. Crystal structures of collagen model peptides with pro-hyp-gly repeating sequence at 1.26 Å resolution: implications for proline ring puckering. *Biopolymers*. 76:367–377.