Mechanistic Analyses of Two Structurally Disparate Antifreeze Proteins: Types I and II

by

Michèle Christine Loewen

A thesis submitted to the Department of Biochemistry
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
August, 1997

To My Family

ABSTRACT

To learn more about how antifreeze proteins (AFP) bind to ice, structural and mutagenic studies were carried out on two different fish AFPs: Types I and II. The Type I AFP is a long, single α-helix protein. Type II AFP is an extensively disulfide-bonded, globular protein.

In order to understand structure/function relationships in Type II AFP, an expression system was developed, using the yeast *Pichia pastoris* as a host, to produce the sea raven AFP (rSRAFP). Gross yields of rSRAFP in flask culture were 3mg/L, which was improved 10-fold by the use of fermentation techniques. This protein was used for NMR and crystallographic structural studies, and as a basis for analysis by site-directed mutagenesis. Mutations to probe the structure-function relationships in SRAFP were based on an earlier computer generated 3D-model of the protein. Single site changes, insertions, and a peptide loop swap were made to assess the ice-binding site (IBS1), proposed in the model. IBS1 corresponds to the region that comprises the Ca^{2+}-binding site on the carbohydrate-binding domain of C-type lectins, which are structural homologs of the AFP. An inability to completely eliminate activity in this region, prompted a search for a second ice binding site on the molecule. Ultimately another putative ice-binding surface (IBS2) on a different face, adjacent to IBS1, was identified and its role was confirmed by mutagenesis. Overall this study provides the first evidence for the presence of more than one ice-binding site on a single AFP.

In the second study, the roles of individual amino acids in the putative ice-binding motif (LTAAN) of Type I AFP were probed. Variations of this motif are repeated 4

times along the length of the helical protein. Variants were made by solid-phase synthesis to have identical changes to the inner two identical ice-binding motifs. Substitution of Thr with either Ser or Val caused 90% and 10% losses of activity respectively, indicating a significantly diminished role for hydrogen-bonds and an augmented role for van der Waals contacts in the ice-AFP interaction. This is contrary to all previous models. Replacements of Leu suggests that it plays a role in maintaining solubility. Replacement of Asn with Thr caused a significant increase in thermal hysteresis activity (120%) and slightly altered ice crystal morphology. Minimization of the peptide from 37 residues to 15, with conversion of the salt bridge into a lactam bridge demonstrated that only two of the four ice binding motifs are required for surface-specific binding and that full-length AFP might bind to ice in stages rather than all at once.

Overall this work has provided a number of significant findings about AFP structure/function relationships, that have altered views on some of the very basic concepts in this field.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Peter Davies, whose patience and enthusiasm for science were a constant inspiration to me throughout a somewhat difficult and trying project. I would also like to thank my co-supervisor Dr. Andrew Daugulis, for his instruction in fermentation technique, as well as his never ending optimism, even in the face of disappointment and the occasional disaster.

I want to acknowledge Dr. Wolfram Gronwald, Dr. Frank Sönnichsen and Dr. Brian Sykes for their NMR work and patience throughout this project, as well as Dr. Zongchao Jia for his expertise in protein crystallography. I am grateful for the opportunity to have worked with Dr. Heman Chao and Dr. Michael Houston and acknowledge all their contributions to the Type I project, in discussions, peptide synthesis, CD and sedimentation analyses. I thank Dr. Choy Hew for the gift of purified native sea raven AFP.

Stimulating discussions and help from all the members, past and present, of the Davies group including, Sherry Gauthier, Dr. Carl DeLuca, Dr. Cathy Watson, Dr. Kim Kenward, Dr. Bernard Duncker, Yih-Cherng Liou, Jason Baardnes, Dr. Laurie Graham, Dr. Simon Arthur, Dr. John Elce, Chris Hosfield and Steffen Graether was appreciated and will always be remembered.

Finally I wish to acknowledge NSERC and MRC for providing me with studentships and research funding to PLD for this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

μ	micro
3D	three-dimensional
Å	angstrom
AFGP	antifreeze glycoprotein
AFP	antifreeze protein
amino acids	3 letter and single letter abbrevations
AOX	alcohol oxidase
ATP	Adenosine triphosphate
BMGY	buffered minimal glycerol-complex medium
BMMY	buffered minimal methanol-complex medium
bp	base pair
BSA	bovine serum albumin
°C	degrees Celcius
CD	circular dichroism
cDNA	complimentary dideoxyribonucleic acid
CIL	Cambridge Isotope Laboratories
CDW	cell dry weight
CRD	carbohydrate recognition domain
C-Type	Ca^{2+}-dependent
Da	dalton
DEAE	diethylaminoethyl
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DO	dissolved oxygen
DTT	dithiothreitol
E.coli	*Escherichia coli*
EDTA	ethylene-diamine tetra-acetic acid
FPLC	fast performance liquid chromatography
g.	genus
g	gram
h	hour
HIS4	gene encoding histidinol dehydrogenase
HIS4$^-$	HIS4 gene deleted
HIS4$^+$	HIS4 gene present
(*his4*)	HIS4 gene deleted
HPLC	high pressure liquid chromatography
IBM	ice binding motif
IBS	ice binding surface

I_h	hexagonal ice
Kbp	kilo base pair
kDa	kilo dalton
L	liter
LB	Luria Bertani
M	molar
m	milli
MALDI	matrix assisted laser desorption/adsorption ionization
MeOH	methanol
MBP	mannose binding protein
MD	minimal dextrose medium
MG	minimal glycerol medium
min	minute
MM	minimal methanol medium
Mut^+	methanol utilization normal
Mut^s	methanol utilization slow
n	nano
NEB	New England Biolabs
NEN	NEN Research Products
NMR	nuclear magnetic resonance
NOE	nucear overhauser effect
nSRAFP	native sea raven AFP
OD	optical density
rSRAFP	recombinant sea raven AFP
PAGE	poly-acrylamide gel electrophoresis
PCR	polymerase chain reaction
pI	isoelectric point
P. pastoris	*Pichia pastoris*
PSP	pancreatic stone protein
PTR	proteoglycan tandem repeat
PVDF	polyvinylidene difluoride
SDS	sodium dodecyl sulfate
SRAFP	sea raven antifreeze protein
ssDNA	single stranded deoxyribonucleic acid
TFA	trifluoro acetic acid
USB	United States Biologicals

CHAPTER 1

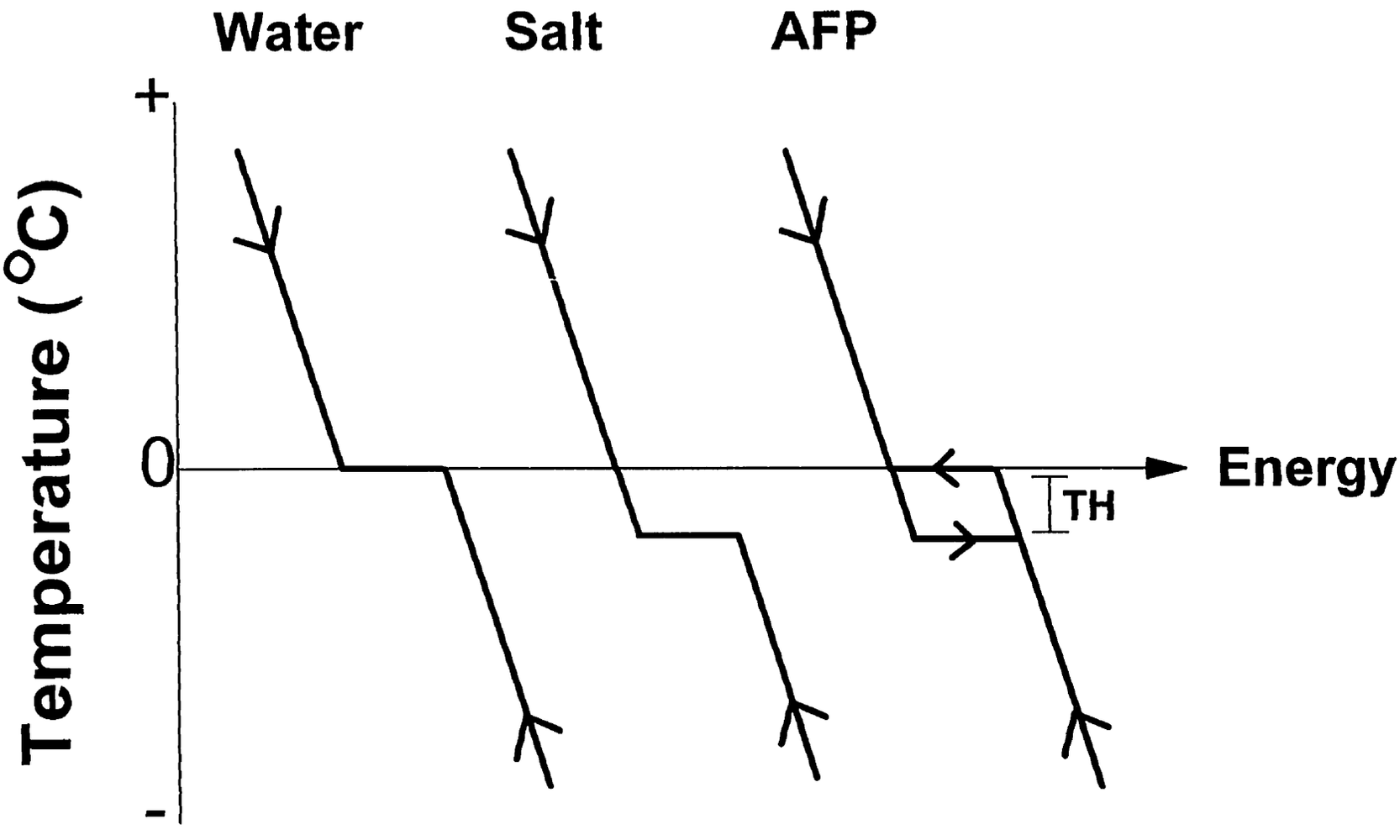
INTRODUCTION

ANTIFREEZE PROTEINS

Antifreeze proteins (AFPs) have been isolated from a number of organisms including fish, insects, plants and bacteria (1)(2)(3)(4). These macromolecular antifreezes serve to depress the freezing point of fluids within the organism in order to protect the organism from freezing to death during periods of sub zero °C temperatures. The viability of freeze tolerant organisms is presumably helped by the ability of AFPs to inhibit ice-recrystallization (5). Dissimilarity in amino acid composition and protein structure have been observed between AFPs from different organisms. Despite this dissimilarity, all AFPs characterized to date, act in a non-colligative manner to depress the freezing point of solutions, some by as much as five to six °C (6)(7) (Figure 1). The melting point is lowered only by the colligative property of the AFP, which is a trivial change because the molar concentration of macromolecules is almost negligible (8)(9)(10). The resulting difference between the melting and freezing points is termed thermal hysteresis, and is a function of AFP concentration.

Another characteristic of these proteins is that they alter ice crystal morphology in specific ways. Normal ice growth at mild under-cooling, occurs 100 times faster along the a-axis (prism face) than the c-axis (unordered basal plane), producing a circular

Fig. 1 **The effects of salt and AFP on freezing and melting points.** The equilibrium melting and freezing point of pure water is 0°C. Salt causes a colligative depression of the equilibrium melting/freezing point. AFP causes a depression of the non-equilibrium freezing point below the equilibrium melting point. The difference is termed thermal hysteresis (TH).


Water
Salt
AFP
+
0
-
Temperature (°C)
Energy
TH

disk-like ice crystal morphology (11)(12)(13)(14). Fish AFPs alter the shape of this growing ice crystal from the circular plate to one of a number of distinct hexagonal bipyramidal forms, depending on the type and concentration of AFP involved (15)(16)(Figure 2). This is indicative of AFP binding to a derivative of the prism plane, inhibiting the energetically favourable a-axis growth, allowing only the energetically unfavourable growth along the c-axis of the crystal (17).

FISH ANTIFREEZE PROTEINS

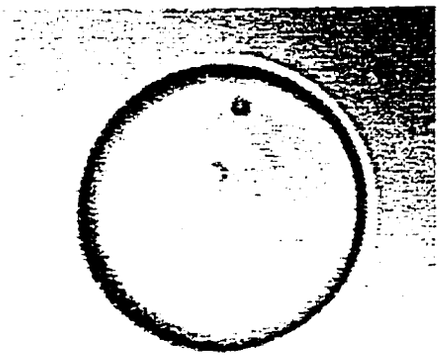
Antifreeze glycoproteins

The antifreeze glycoproteins (AFGPs) were the first identified macromolecular antifreeze. They were isolated from the serum of nototheniid fishes in the Antarctic Ocean (18). These glycoproteins are polymers of tripeptide repeating units (Ala-Ala-Thr)$_n$ to which the disaccharide beta-D-galactosyl(1-3)-alpha-N-acetyl-D galactosamine is linked via the hydroxyl group of the threonyl residue of each repeat (19). Gel electrophoresis has resolved as many as eight distinct components, ranging in mass from 2600 to 34000 Daltons (Da). AFGPs 1-5 represent the five largest AFGPs ranging from 50 repeats (AFGP1) to 17 repeats (AFGP5) (20). AFGPs 6-8 are the three shorter, less active components (4-7 repeats), in which Pro occasionally substitutes for Ala on the C-terminal side of Thr (21).

A gene for AFGP has been cloned and sequenced. It contains 46 tandemly repeated segments each encoding one AFGP peptide plus a 3-amino-acid-long spacer

Fig. 2 **Ice crystals formed in the absence and presence of AFPs.** Example of the disk-like ice crystal morphology produced in pure water. Examples of the hexagonal bipyramidal ice crystal morphology, and variations thereof, produced by AFPs from fish (Type I, II and III).

WATER



Type I

Type II

Type III

(22). This polypeptide is post-translationally processed by proteolysis and glycosylation to produce active AFGPs. Recently, comparison of the gene structure of the notothenioid AFGPs, indicates up to 96% homology in the signal peptide and 3' untranslated region with the notothenioid pancreatic trypsinogen gene (23). This suggests trypsinogen is the likely ancestral gene precursor of this particular AFGP. It is hypothesized that the *de novo* amplification of a nine nucleotide Thr-Ala-Ala element from the trypsinogen produced the backbone of the AFGP. Other AFGPs from cod do not maintain this homology to trypsinogen, even though their AFGP structures are very similar.

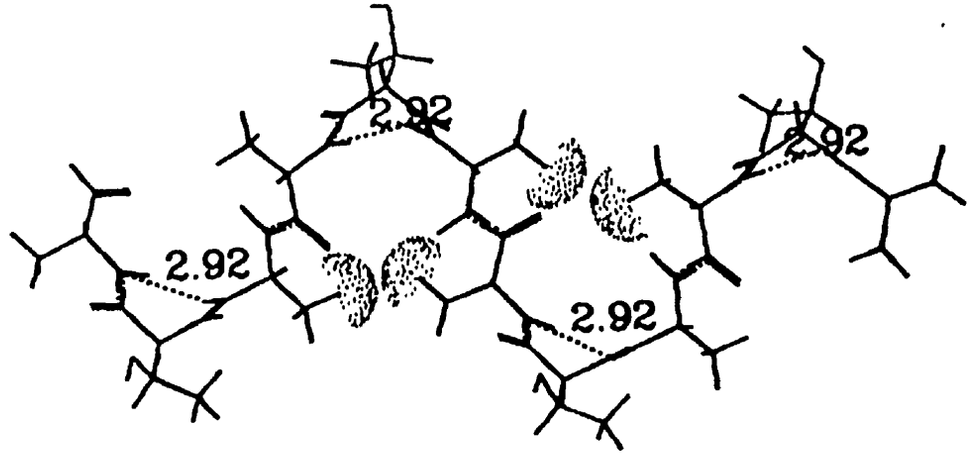
A polyproline type II helix has been the most widely accepted conformation for AFGP, as determined by NMR (24)(25). However, recent studies using Raman and infra-red spectrometry, combined with molecular modelling, now suggest a γ-turn motif as the most plausible conformation for active AFGPs (26) (Figure 3). This places all of the disaccharides on the same side of the molecule spaced 9.0Å apart, twice the distance of the 4.5Å spacing of alternate oxygens on the prism face of the ice crystal to which AFGP has been shown to bind by ice etching studies (27)(28).

Type I Antifreeze Protein

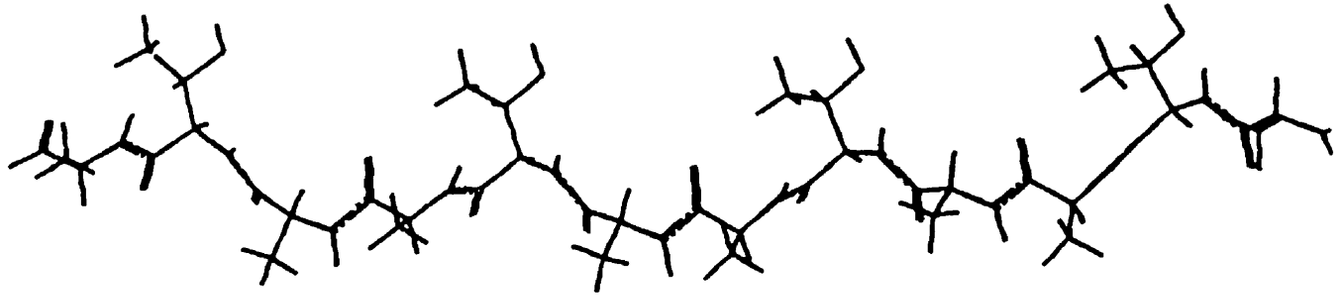
Type I AFPs are long α-helical, Ala-rich (> 60mol%) peptides found in righteye flounders (g. *Pleuronectes*) and some sculpins (g. *Myoxocephalus*) (1). Many different isoforms, varying in sequence and length have been isolated (Mw 3300-5000). At one extreme are the shorthorn sculpin serum AFP (29)(30) and the winter flounder (*Pleuronectes americanus*) skin isoform (31), which are both amphipathic, non

Fig. 3 **AFGP structures.** top: Four Ala-Thr-Ala repeats showing the γ-turn motif proposed as the most likely AFGP conformation. Thr hydroxyls are at the points of the turns, disaccharides are not shown. 2.92 refers to the hydrogen bond distance (Å) accross the turn. Van der Waals radii of four methyl hydrogens groups are shown with dots. bottom: Four Ala-Thr-Ala repeats showing the previously porposed polyproline II conformation. All Thr are pointing up. Figure taken from Drewes and Rowlen 1993 (26).

Gamma-Turn Motif



Polyproline Type II Helix



repetitious, and contain a relatively large proportion of Lys. In contrast, the winter flounder serum peptides are typified by the most abundant isoform, HPLC-6 (32)(33)(34). It is 37 amino acids long, contains three complete 11-amino-acid repeats of T-X_2-Asx-X_7, where X is generally alanine, and ends with the start of a fourth repeat (33)(34) (Figure 4). Other flounder serum isoforms with amino acid substitutions or different repeat lengths have been isolated (34)(35)(36). Winter flounder Type I AFP genes belong to a large multigene family with 30-40 members (37), of which half encode the abundant HPLC-6 component (38).

The HPLC-6 high resolution X-ray crystal structure (39)(40) confirmed the predicted helical structure (41) and regular spacing of the T and N/D residues along the same helix surface (Figure 5). The sequence LTAAN (repeated twice along the peptide) and DTASD at the N-terminus, were designated by Sicheri and Yang (40) as ice binding motifs (IBM). A fourth motif near the C-terminus is incomplete. The spacing of the hydrophilic residues making up the IBM matches the spacing of oxygen atoms on the {20-21} plane of ice in the <10-12> direction to which flounder Type I AFPs have been shown to bind by ice etching studies (42)(43).

The X-ray structure also confirmed the existence of intricately, internally hydrogen-bonded N- and C-terminal cap structures, which are known to increase helicity of α-helical peptides (44). These combined with the intra-helical hydrogen-bonds, the salt bridge (K18-E22) and the dipole interactions between the helix dipole and the N- and C-terminal charged amino acids, all contribute towards stabilizing the exceptionally long helical peptide (45). Circular dichroism studies indicated the helicity is sensitive to

Fig. 4 **HPLC-6 sequence.** The sequence of the most abundant winter flounder serum Type I AFP isoform, HPLC-6, is shown with its regular spaced threonines indicated by a dot. One of the three 11-amino-acid repeats is defined by a bar. A salt bridge between K18 and E22 is indicated by a hatched line. The IBMs designated by Yang and Sicheri (40) are underlined.

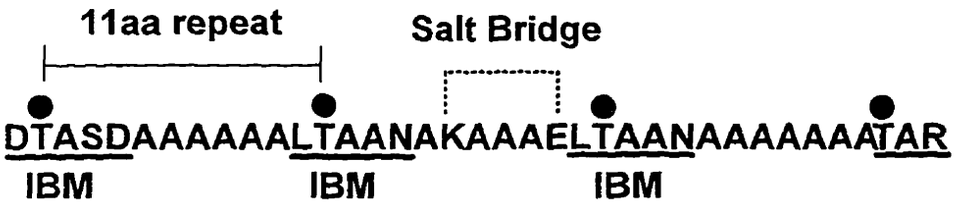
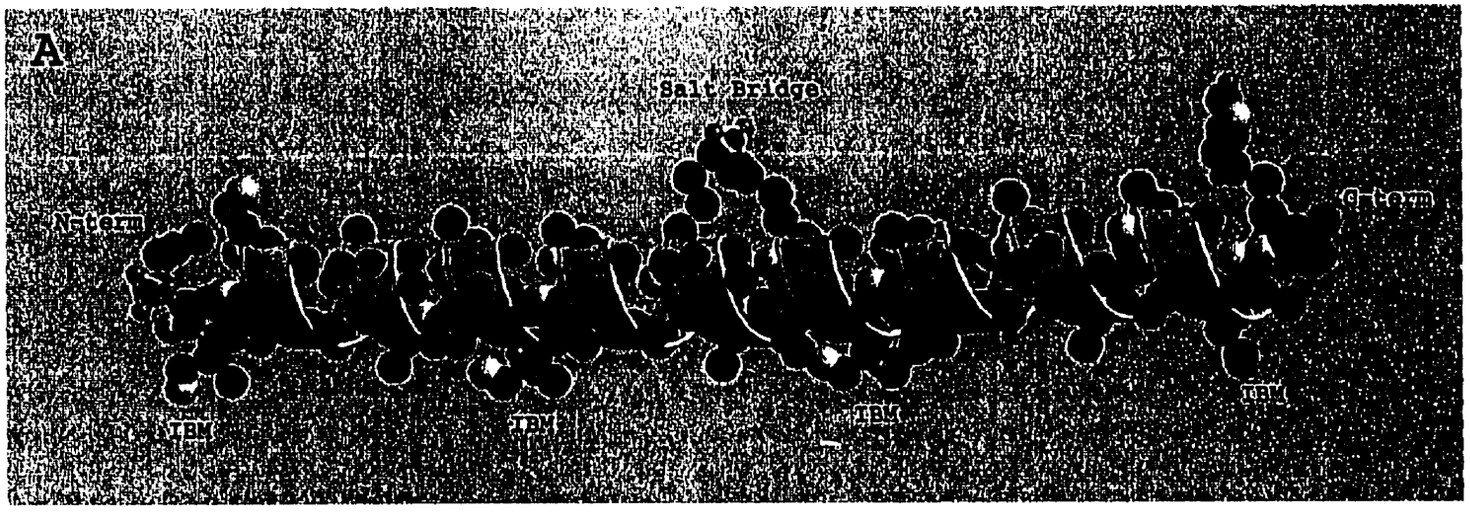
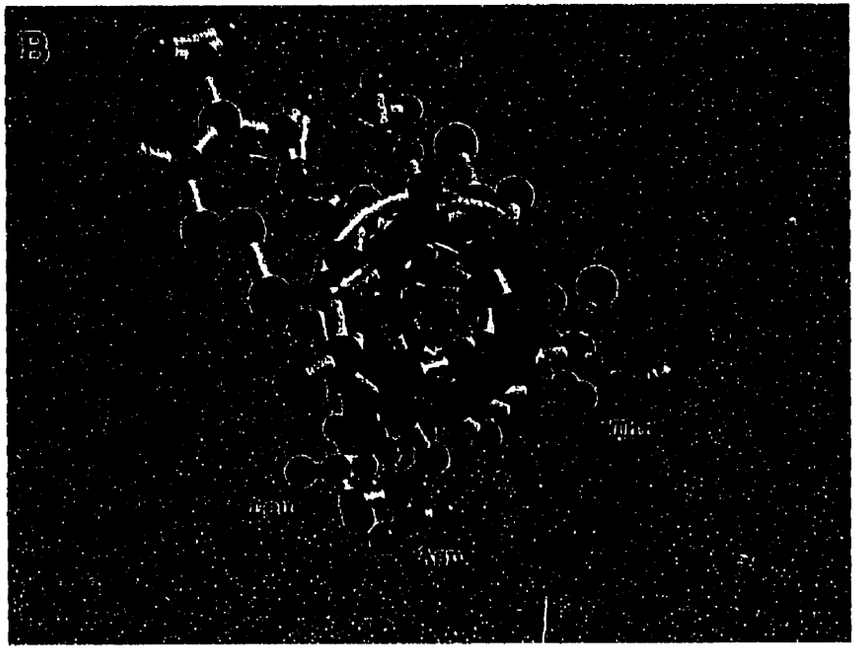

11aa repeat
Salt Bridge
DTASDAAAAAALTAANAKAAAELTAANAAAAAAAATAR
IBM
IBM
IBM

Fig. 5 **X-ray crystal structure of HPLC-6.** A: ball and stick side view of the helix with the salt bridge and four IBM regions identified. Colour scheme: carbon in black, nitrogen in blue, oxygen in red and hydrogen in white. B: ball and stick view down the helix from the C-terminal end. The colour scheme is the same as for A. The relative positions of the Thr, Asn and Leu residues of the IBMs on the side of the helix are indicated (40). Brookhaven protein data base identifier 1wfa. 2° structure was assigned using STRIDE (47). Diagrams were produced using Molscript (48) and Raster3D (49) on a Silicon Graphics terminal, printed on an OptraC LexMark colour laser printer.


A
Salt Bridge
N-term
C-term
TBM
TBM
TBM
TBM

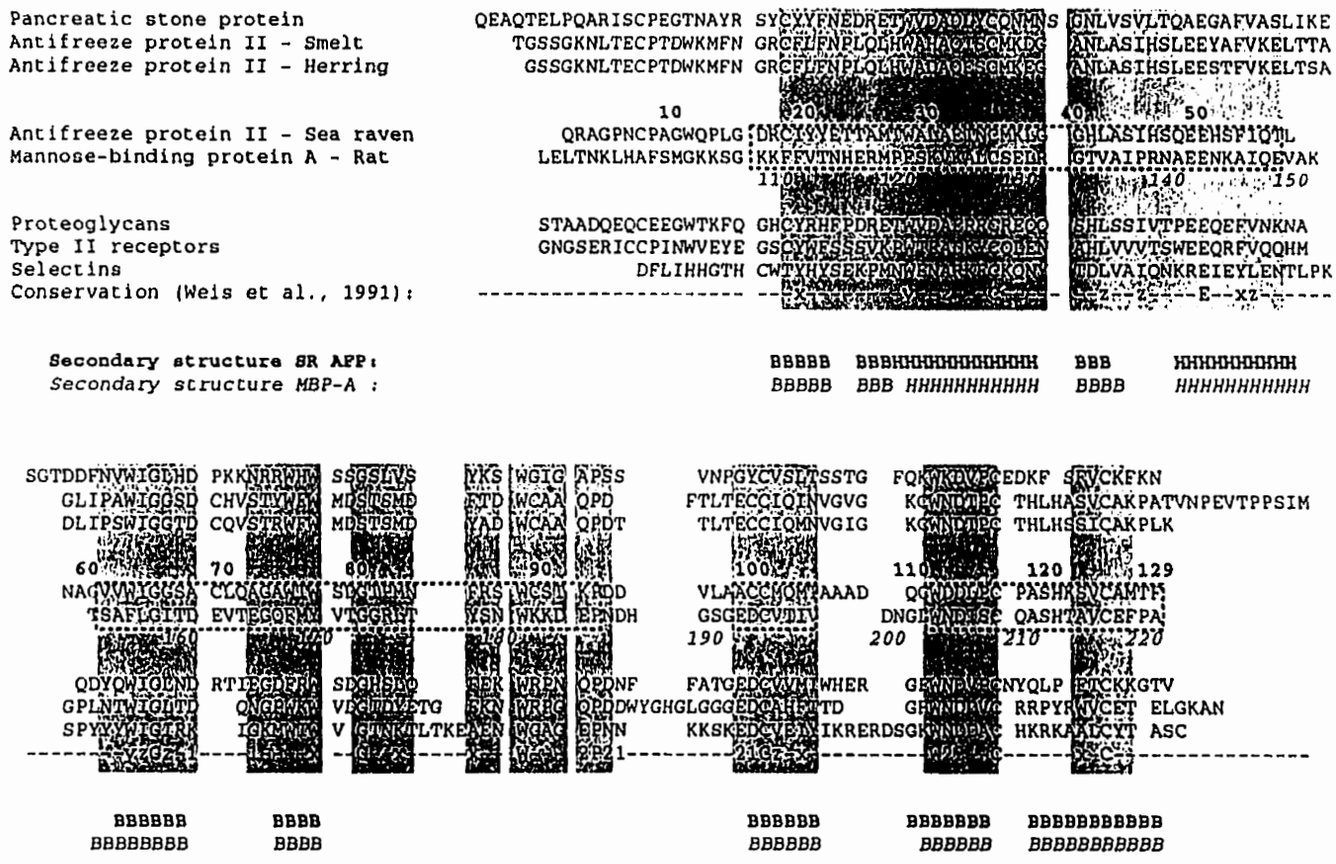

B

temperature. The peptide was close to 100% α-helix near 0°C, but was only 50% helical at 25°C, and resembled a random coil at 70°C (46). Helicity was also reduced by pH at values lower than 4.0 and greater than 9.0 (45). One attempt was made to increase the helicity of HPLC-6 by incorporation of two additional salt bridges (A7K-A11E and A29K-A33E) on the same helical surface as the original salt bridge (K18-E22). Although a more helical peptide was produced, the author found no increase in thermal hysteresis activity, but did see a reduction in ice crystal growth rates along the a- and c-axes compared to wild-type values (45).

Type II Antifreeze Protein

Type II antifreeze protein is a globular protein of Mw 14,000, that has been isolated from three very distantly related fish: sea raven (*Hemitripterus americanus*) (50), herring (*Clupea harengus harengus*) (51) and smelt (*Osmerus mordax*) (52) Type II AFPs show no homology to other AFPs, but are up to 30% identical to the carbohydrate-recognition domain (CRD) of Ca^{2+}-dependent (C-Type) lectins, as well as to the pancreatic stone protein (PSP) (53)(52) (Figure 6). CRDs are known to bind to carbohydrate residues in a Ca^{2+}-dependent (54) manner while PSPs bind to the surface of calcium-carbonate ($CaCO_3$) crystals and inhibit their growth (55). The amino acids that are conserved are principally those residues that define the tertiary structure of the CRD including: C35, C100, C117 and C125 involved in disulfide bridges; G65, G81 and P93 that make turns between loops; and W28, W63, W88 and W112, that help form the hydrophobic core (sea raven numbering). This suggests Type II may adopt a similar fold

Fig. 6 **Primary amino acid sequence alignment of Type II AFPs and homologues.** The sequences displayed are: three Type II AFPs, rat MBP, rat proteoglycan core protein (Proteoglycans), rat hepatic lectin (Type II receptors), lymphocyte homing receptor (Selectins) and PSP. Shaded regions represent sequences that are always consistently aligned. Gaps indicate deletions or insertions. Dashed boxes indicate regions defined as structurally conserved by the model.
Conservation x=aliphatic or aromatic, y=aromatic, z=aliphatic,
1=ligands for Ca^{2+} site 1, 2=ligand for Ca^{2+} binding site 2. 2^{o} structure from the SEQSEE program (58). B=β-sheet, H=α-helix. Figure taken from Sönnichsen et al., 1995 (53).


Pancreatic stone protein
Antifreeze protein II - Smelt
Antifreeze protein II - Herring
Antifreeze protein II - Sea raven
Mannose-binding protein A - Rat
Proteoglycans
Type II receptors
Selectins
Conservation (Weis et al., 1991):
Secondary structure SR APP:
Secondary structure MBP-A :

to the CRD (52)(56). This homology is also significant because the herring and smelt AFPs (83% identical to each other) conserve all of the five Ca^{2+}-binding residues of the CRDs (D120, E105, Q98, D100 (herring numbering))(51), and show metal ion- (Mn^{2+}, Ba^{2+}, Zn^{2+} and Ca^{2+}) dependent structure and activity (57). Sea raven AFP (SRAFP) is only 40% identical to the herring and smelt isoforms (51), conserves only two of the five CRD Ca^{2+}-binding residues (D94, D114 (sea raven numbering)) and does not require Ca^{2+} for activity (52).

The gene for sea raven AFP is present in 12-15 copies in the sea raven genome. A representative copy was sequenced and found to be split into six exons spanning 2.2kbp. DNA and protein sequence analysis indicated that the cDNA encodes a precursor protein of 163 amino acids (59). This included a 34-amino-acid-long N-terminal precursor sequence of which the first 17 residues comprise the signal peptide, and the second 17 the prosequence, which are both removed to produce a mature 129-amino-acid-long AFP (60). Both pro and mature forms of the protein have been isolated from sea raven serum (61). The pro form was also secreted by the natural fish signal peptide, when fall army worm cells were infected with a baculovirus expression vector containing the entire sea raven AFP cDNA (60).

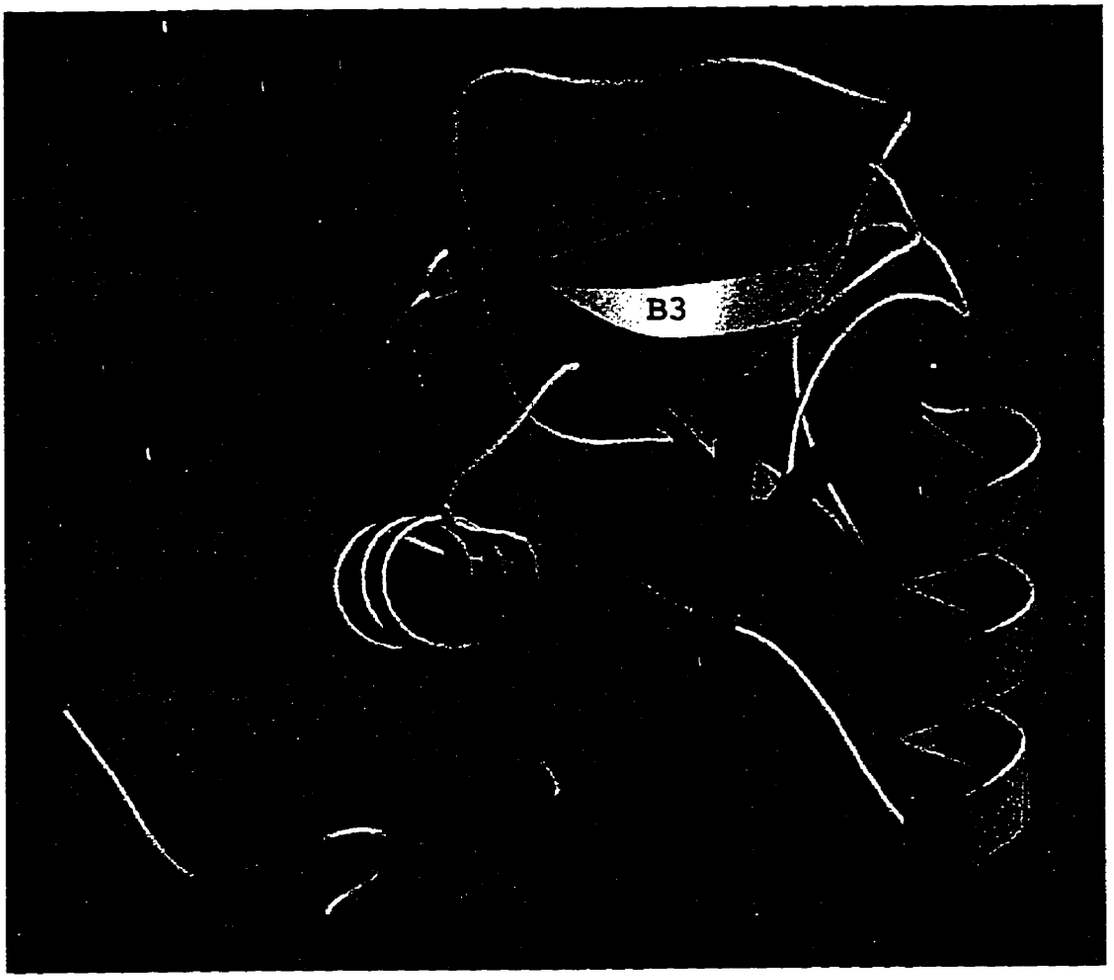
X-ray crystal structures of a number of CRDs, including mannose-binding protein (MBP) (56)(62) and E-selectin (63) have been solved. The lectin 3D structure is comprised of two antiparallel β-sheets, each made of β-strands 1 and 5, or 2, 3 and 4 respectively. β2 divides the entire molecule into two halves. The lower half contains the first β-sheet (1+5) and the only two helices in the structure, sandwiching the β-sheet.

The upper half contains β-strands 3 and 4 of the second β-sheet, as well as four large loops. The upper region including loops 3 and 4, and β-strands 3 and 4 contain all the residues required for binding of Ca^{2+} and carbohydrate to the lectin (56)(62). A similar fold is described for lithostathine, the human PSP (64).

The mature 129-amino-acid-long SRAFP includes ten cysteines which form five disulfide bridges (65). Four of these cysteines are conserved in the CRD and PSP, one pair linking the N- and C-termini (C35-C125), the other pair linking β-strands 3 and 4 (C101-C117)(62)(53). As well, the two cysteines, C7 and C18, are conserved in some CRD and PSP sequences to form a third bridge. Originally, peptide mapping was used to determine the disulfide bonding pattern in SRAFP (66). This map linked C117 to C89, and the adjacent cysteines C101 and C100 to either C69 or C111. This pattern maintains only the C35-C125 and C7-C18 bridges in common with the CRD X-ray structures (56)(62). Computational modelling on SRAFP (53) proposed an alternative map that maintains all three conserved disulfides in addition to C69-C100 and C89-C111. AFP activity can be largely inactivated by sulfhydryl reagents, indicating an important role for these disulfide bridges (50).

3D models of both sea raven (53) and herring (57) AFPs have been computed based on the X-ray crystal structure of the CRD of rat MBP (Figure 7). Both herring and sea raven models, despite sharing only 40% identity to each other, and no more than 30% identity with the CRDs, maintain a very similar fold to that described for the lectins. They maintain all five β-strands, the two α-helices, the four loops, and the two conserved disulfide bonds. The only significant variation from the lectin structure occurs

Fig. 7 **3D model of sea raven AFP.** α-helices (H1 and H2), β-strands (B1-5) and loops (L1-4) are identified. The putative ice binding surface is located at the top of the molecule, involving residues of L3, L4, B3 and B4. Structure file courtesy of F. Sönnichsen. 2° structure assigned using STRIDE (47). Diagrams produced using Molscript (48) and Raster3D (49) on a Silicon Graphics terminal, printed on an OptraC LexMark colour laser printer.


B3

in the conformation of loop 4, which makes up part of the Ca^{2+} binding surface (53)(56)(62). A putative ice binding site has been proposed for the AFPs on this same surface (53). The observed conservation in herring of all five residues involved in binding Ca^{2+} to lectin (52) prompted an interesting study on the effect of divalent metal cations on herring AFP activity (57). It indicated that binding of the cation causes a conformational change in the protein which is required for activity. Different divalent cations produced different levels of thermal hysteresis and had markedly different effects on ice crystal morphology. The wild-type spicular hexagonal bipyramid ice crystal formed in the presence of Ca^{2+} and Mn^{2+} changed to a rectangular form in the presence of Mg^{2+}, Ba^{2+} and Zn^{2+}. While the same surface was proposed for the ice-binding face of SRAFP, it conserves only two of the five residues for binding Ca^{2+} and shows no Ca^{2+}-dependent activity. The SRAFP wild-type ice crystal also differs significantly from that of the herring AFP. It produces a very rounded hexagonal bipyramidal ice crystal, with no distinct crystal planes in evidence (Figure 2). The putative ice-binding site of SRAFP has a large number of hydrophilic, surface-accessible, residues, oriented on a relatively planar surface. Cheng and DeVries (27) have reported that SRAFP binds to the {11-21} plane of ice. However, to date no spatial match has been made between SRAFP surface residues and water molecules on the {11-21} plane of ice.

Type III Antifreeze Protein

Type III AFPs, isolated from zoarcid fishes such as the ocean pout (*Macrozoarces americanus*) and wolffish (*Anarhichus lupus*) are 7kDa proteins, with no pronounced

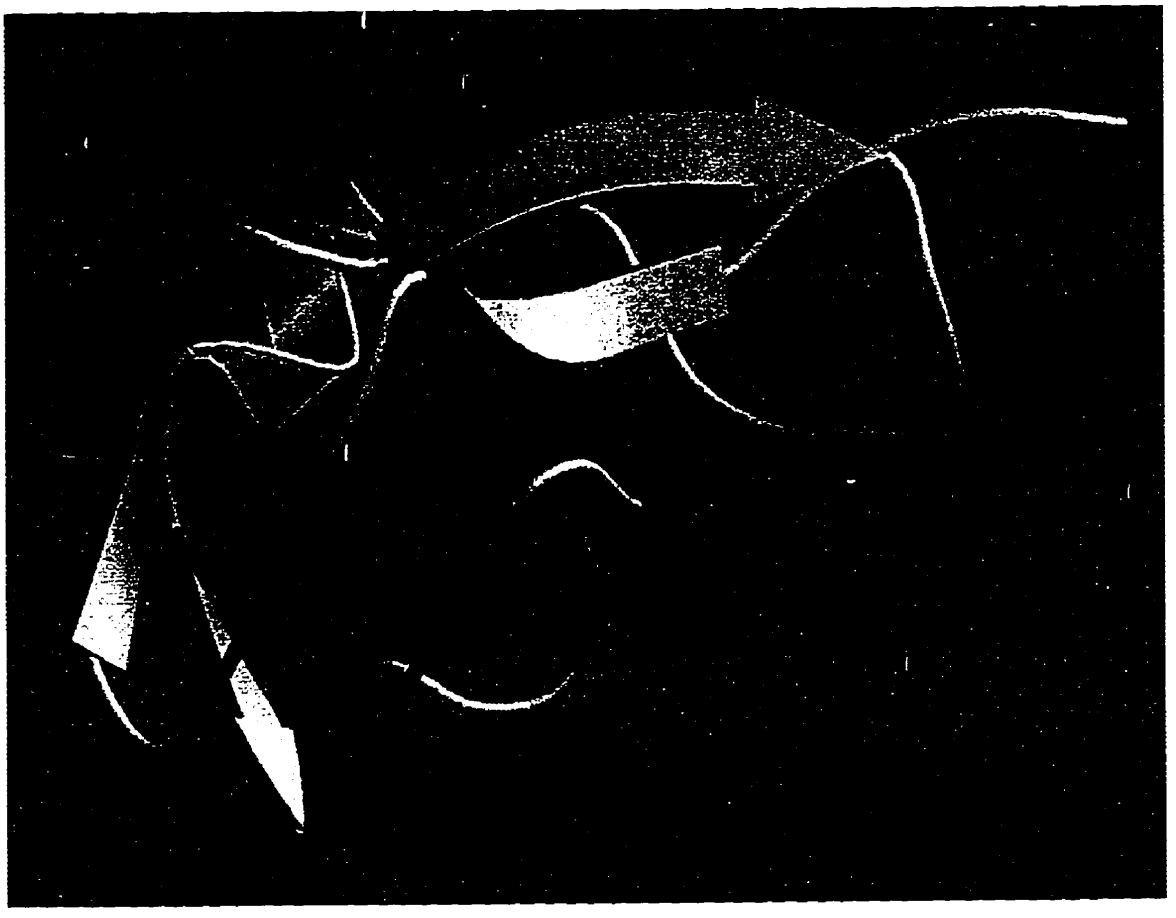
amino acid imbalances. Many isoforms of Type III have been isolated to date (1). They can be divided into two homology groups: QAE and SP (67)(68). The QAE components have pIs just below neutrality and show 75% sequence identity. The SP components have basic pIs, show 85% identity to each other, and are 50% identical to the QAE components. In the ocean pout, these different isoforms are the product of 150 linked, but irregularly spaced, genes (69).

A synthetic gene for Type III AFP was designed and over-expressed in *Escherichia coli* (70). AFP was purified from both the soluble fraction and the inclusion bodies , the latter after refolding into an active conformation. This system has provided ample protein for structural and functional analyses. Both the solution (NMR) and crystal (X-ray) structures of recombinant QAE Type III AFP have been determined (71)(72)(73) (Figure 8). This AFP is a compact angular protein, in which the overall fold comprises numerous short and imperfect β-strands, and one turn of α-helix (73). Extensive mutagenic studies have also been carried out using this over-expression system to study the ice-binding site of the globular protein (74)(16), which, together with the structural work, has provided a detailed model for Type III AFP binding to the prism plane of ice (73).

Type IV Antifreeze Protein

A new AFP (Type IV) has recently been purified from the serum of longhorn sculpin (*Myoxocephalus octodecimspinosis*) (75). Type IV AFP is 108 amino acids in length and contains 17% Gln. Circular dichroism studies and conformational analysis

Fig. 8 **X-ray crystal structure of Type III AFP.** The ice binding surface is comprised of the C-terminal β-sheet on the left side of the molecule (73). Structure file courtesy of C. DeLuca. 2° structure assigned using STRIDE (47). Diagrams produced using Molscript (48) and Raster3D (49) on a Silicon Graphics terminal, printed on an OptraC LexMark colour laser printer.



indicates a high content of helix. Sequence analysis indicates the protein has no homology to the αhelical Type I AFP, but that it is related to the low-density-lipoprotein receptor-binding domain of human apolipoprotein E, for which the crystal structure is known to be a four helix-bundle (75)(76).

INSECT, PLANT AND BACTERIAL ANTIFREEZE PROTEINS

Many insects have been reported to produce antifreeze proteins (4)(7)(77-82). In our laboratory the active components have been isolated and characterized from spruce budworm (*Choristoneura fumiferana*) and beetle (*Tenebrio molitor*) (Tyshenko *et al*,. unpubl., Graham *et al.*, unpubl.). Characterization of recombinant spruce budworm AFP (9kDa) indicates it has up to 10-30-fold greater thermal hysteresis activity than fish AFPs on a molar basis. The ice crystals shaped by spruce budworm are rounded-hexagonal bipyramids somewhat similar to those observed with SRAFP. *Tenebrio* AFP (8.4kDa) also produces a very rounded hexagonal bipyramid-like ice crystal, more closely approximating the shape of an American football and dilutional analysis indicates it too has exceptionally high activity (L. Graham and S. Gauthier personal communications). Both proteins are rich in Thr, Ser and Cys, are inactivated by reducing agents, and have an underlying repeat structure that is most obvious in the *Tenebrio* AFP. Nevertheless, there is no obvious homology between the two insect AFPs and the spacing of the disulfide-bonded Cys residues is completely different.

Both the bittersweet nightshade (*Solonum dulcanara*) and cold-acclimated winter

rye (*Secale cereale L.*) plants produce AFPs (83)(3)(84). In the latter, active proteins range from 19-36kDa, whereas the nightshade AFP is significantly larger at 67kDa. On a molar basis, the nightshade AFP is significantly less active than even the least active fish AFP on a molar basis (83). Thermal hysteresis activity of the winter rye AFPs has not been quantified, but they give rise to hexagonal bipyramidal-shaped ice crystals, similar to those produced in the presence of fish AFPs (85). The winter rye AFPs have also been identified as homologues of the plant pathogenesis-related proteins, endochitinases, endo-β-1,3-glucanases and thaumatin. The pathogenesis-related proteins themselves do not show thermal hysteresis activity (85).

The plant growth-promoting rhizobacterium, *Pseudomonas putida* GR12-2, was found to synthesize and secrete a protein with AFP activity. This protein has yet to be purified and characterized (2).

THE MECHANISM OF ACTION

From the very first few studies describing their isolation and characterization (18) (20) scientists have been proposing mechanisms and models of how AFPs function in lowering the freezing point. Original observations suggested that depression of the freezing point by AFGP was not in accord with the number of particles in solution *(ie.* it was not a colligative effect) (20). Instead the first mechanisms assumed that large volumes of water were affected. It was proposed that AFGPs might be sequestering and immobilizing water molecules, thereby depressing the apparent concentration of water

and subsequently depressing the freezing point (20)(86). This mechanism of sequestering water had previously been shown to lower freezing points with synthetic flexible polymers (87).

However, later studies, including measurement of AFP concentrations in ice and scanning electron micrographs of freeze-dried AFP solutions, indicated that the AFPs were incorporated into the ice (88). Further observations indicated AFPs had a significant effect on the crystal habit of ice growing in their presence (88). Based on this evidence, it was proposed that ice and not liquid water is the primary AFP ligand. Other studies, including surface-second-harmonic generation and ice etching (89)(27) provided direct evidence that ice is the AFP ligand.

Ice: The Antifreeze Protein Ligand

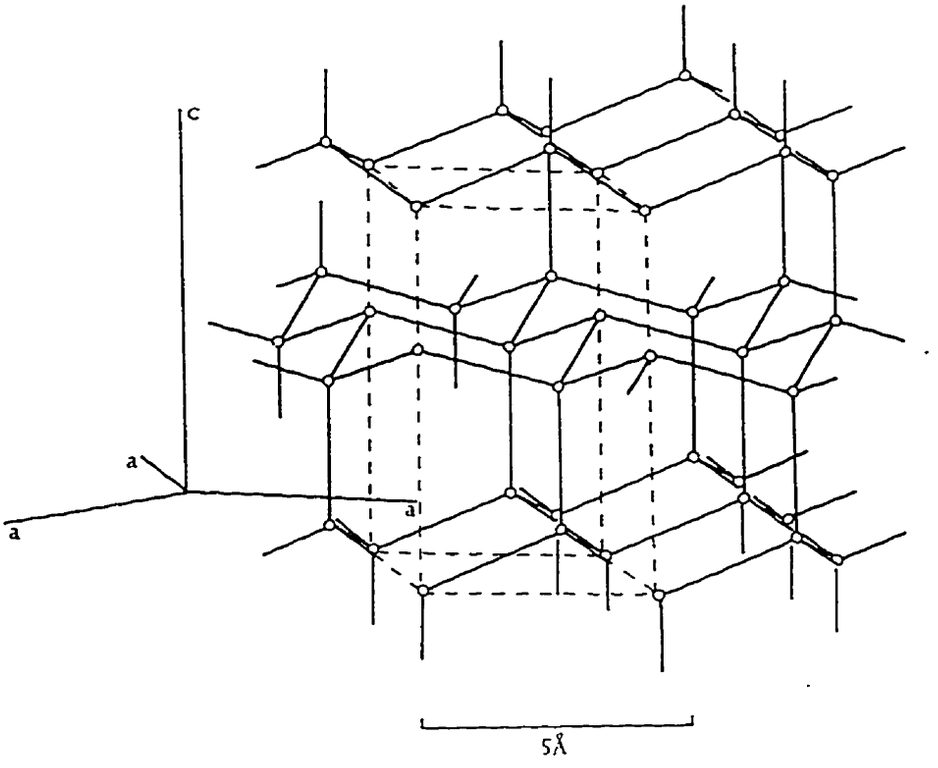
Ice is the solid state of water. Water molecules are comprised of one oxygen atom and two hydrogen atoms, arranged with the oxygen at the centre of a regular tetrahedron (90). The four vertices are occupied by the two lone pairs of electrons from the oxygen and the two hydrogen atoms, enabling the molecule to have both proton acceptor and proton donor sites. In liquid water the oxygen is coordinated to four other oxygens atoms through hydrogen-bonds (91). These interactions are weak and the overall network is unstable allowing diffusion of water molecules and distortion of the network (91). Hexagonal ice (I_h) has the same tetrahedral structure as liquid water, but with more stable H-bonds, as the rate of hydrogen bond exchange decreases with the depression of temperature and corresponding transition from liquid to solid (91).

The I_h ice lattice consists of puckered layers of water, perpendicular to the c-crystal axis (Figure 9). The oxygen atoms are ordered into hexagonal rings in the "chair" conformation, producing "shafts" running both perpendicular and parallel to the c-axis. These shafts account for the lower density of ice compared to water (92).

Supercooling occurs when water does not freeze below 0°C (91). Ice nucleators must be present in order for supercooled liquid to freeze (91). One source of ice nucleation is the spontaneous generation of ice embryos, or clusters of water molecules that randomly take on an ice-like configuration. If the seed crystal is large enough, crystallization will occur. The probability of producing these nuclei increases with the degree of supercooling (93). At minus 40°C the probability of a suitable nucleator occurring approaches unity and homogeneous nucleation occurs (91)(94). A second source of ice nucleators is the presence of foreign materials that can act as initiators for water organization (91)(95). This is known as heterogeneous nucleation.

The growth of ice crystals is complex. Different axes have different growth rates, which affects crystal morphology. The linear growth rate in the a-axial direction is greater than along the c-axial direction under supercooling (11-14) (Figure 10). The ratio of the two rates is not constant. This variation is the result of different growth mechanisms in the two directions. Growth along the a-axis, requires addition of water molecules to the prism faces. Prism faces have rough surfaces which allows easy addition of water molecules as the lattice already contains dislocations (95). Growth along the c-axis requires addition of water molecules to the relatively flat basal plane. A few nucleating molecules must attach to the flat surface, and then additional molecules

Fig. 9 **Structure of I_h.** top: Arrangement of oxygen atoms in I_h. Note the chair conformation of the hexagonal rings. bottom: Van der Waals radii of the oxygen atoms of I_h. View is down the c-axis to illustrate the shafts that occur in this structure. Figure taken from Eisenberg and Kauzman 1969 (92)).


c
a
a
a
5Å

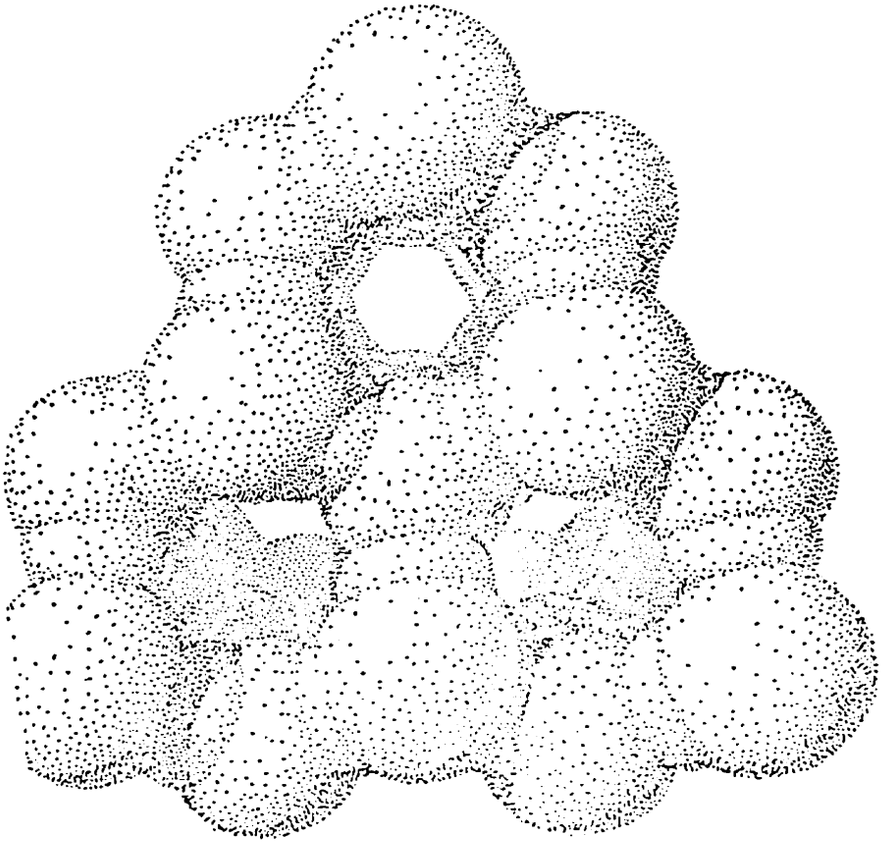
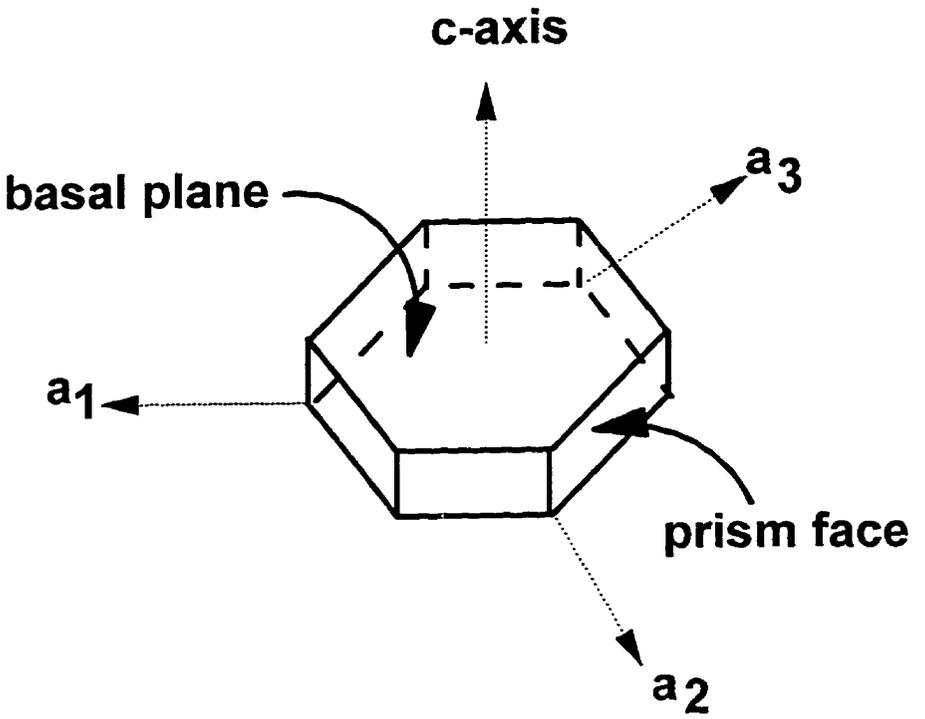


Fig. 10 **The *a* and *c* axes of I_h.** A hexagonal ice crystal with prism faces and basal planes identified. The c-axis extends up and down from the basal planes. The a-axes (a_1,a_2,a_3)extend out from the prism faces.


c-axis
basal plane
a3
a1
prism face
a2

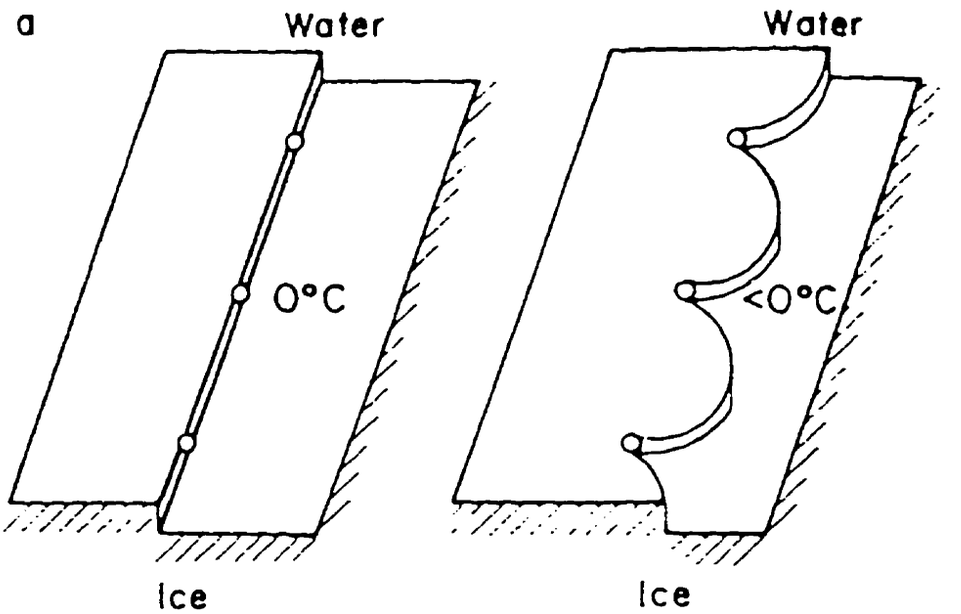
can join by networking with the original molecules, for lateral propagation to produce a new ice layer. This is known as two dimensional nucleation (95)(13). The screw dislocation process (spiral propagation) in the a-axis direction can generate new ice layers faster than the two dimensional nucleation process on the basal plane (96).

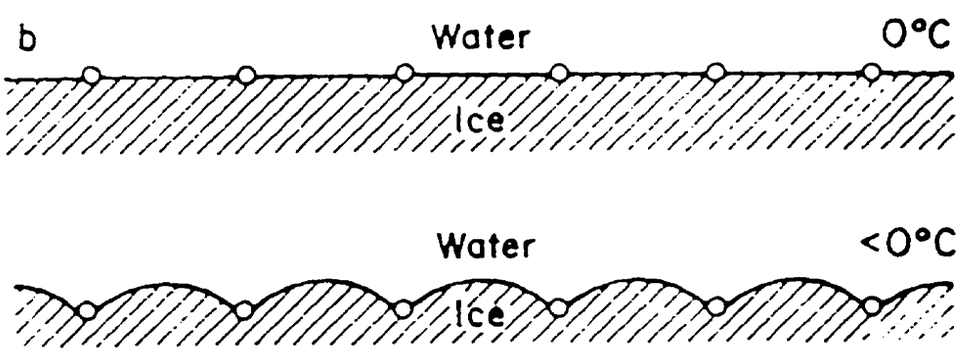
Adsorption/Inhibition

Observation of altered ice crystal growth habits and morphology in the presence of AFPs indicates predominant inhibition of the energetically favourable growth along the a-axis, allowing only the unfavourable c-axis growth (88)(17). Based on these observations, as well as the ice etching and surface-second-harmonic generation results, it was concluded AFPs bind to ice and an adsorption-inhibition mechanism for AFP action was proposed (88). According to this mechanism, AFPs adsorb to the surface of ice, but ice growth can still occur between bound AFPs. This allows a raised curvature on the surface of the ice between bound AFPs, which increases the surface free energy and therefore inhibits further water addition (Kelvin effect) and crystal growth, producing the freezing hysteresis (32) (Figure 11). In this model, adsorbed AFPs stop the growth of the ice crystal when the average spacing between adsorbed proteins is equal or smaller than twice the critical radius of curvature. The critical radius refers to the smallest curvature on the ice surface between bound AFPs that still leads to ice growth. Similar theories for inhibition of other types of crystal growth have been developed (97)(98).

An interpretation of the Gibbs-Thompson (Kelvin) effect (described above) allowed this "step" adsorption-inhibition model to be fitted to the AFP data reasonably

Fig. 11 **Adsorption/Inhibition.** a: An early model of "step growth" adsorption/inhibition (88). In this model AFPs bind to ice steps and inhibit growth by increasing the radius of curvature of the ice surface between bound AFPs. b: The original model was extended from "step growth" inhibition to inhibition "normal to the expressed surface", to reflect the fact that ice does not grow in discrete steps. Figure Taken from Knight et al., 1991 (42).


a
Water
Water
0°C
<0°C
Ice
Ice


b
Water
0°C
Ice
Water
<0°C
Ice

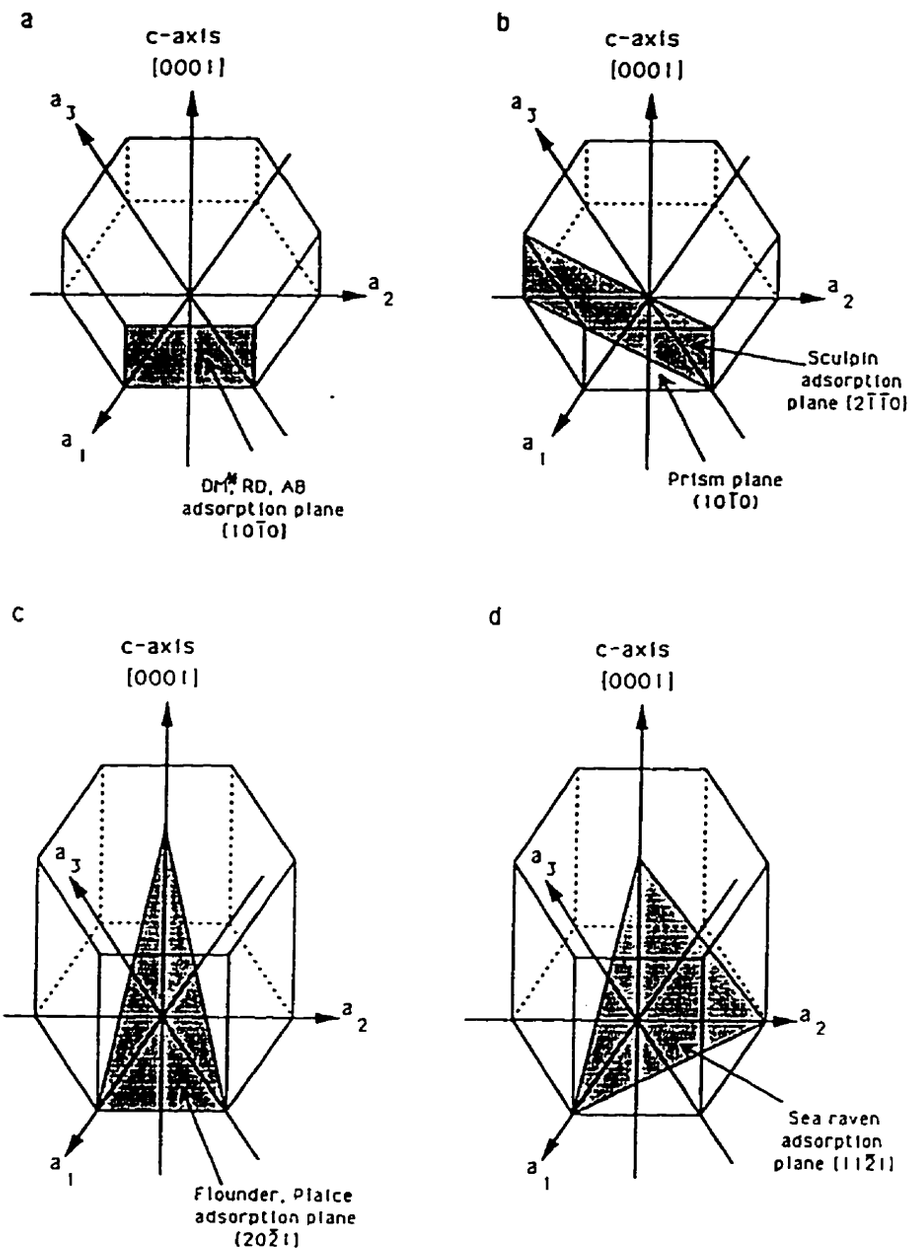
well (32). However, later observations, with the introduction of ice etching techniques, indicated that inhibition (binding) from any face, not just the step surface was required (42)(99). The proposal that bound Type I AFPs produced a "felted" or "mattress" type of ice surface allowed Wilson (100) to fit the "felted surface" adsorption-inhibition model to the AFP data with better success, based again on the Kelvin effect.

Ice Etching

Ice etching is a novel technique that revolutionized the study of AFPs. Developed by Charles Knight and Arthur DeVries (101) it allows determination of the ice crystal plane bound by AFPs in very dilute solutions. Initially an oriented crystal is allowed to grow as a hemisphere in a very dilute solution of AFP. The concentration of AFP must be dilute in order to allow ice growth. This crystal is then "etched" by evaporation or sublimation, until the ice surface is mirror smooth except where the AFP has been incorporated into the crystal (27)(42). From the etching pattern and the orientation of the ice crystal, the adsorption plane is derived.

This technique provides direct evidence of the AFP-ice interaction and also allows a unique method for characterization of AFPs. Analysis of AFGPs and Type I, II and III AFPs indicate these structurally diverse AFPs bind different planes of ice (Figure 12). A mixture of AFGPs 1-5 from *Dissostichus mawsoni* binds to a very high index plane (72-90), very close to the {10-10} and equivalent primary prism planes. Sculpin Type I AFP binds to the {2-1-10} plane, while flounder and plaice Type I AFPs bind to the {20-21} plane of ice. Sea raven Type II AFP binds the {11-21} plane and Type III AFP

Fig. 12 **AF(G)P planes of adsorption to ice.** a: Type III AFP from Antarctic eel pouts binding to the {10-10} prism plane of ice. A mixture of AFGPs 1-5 from notothenid fishes of the Antarctic Ocean bind {72-90}, a high index plane very close to the prism plane. b: sculpin Type I AFP binds the {2-1-10} plane. c: winter flounder Type I AFP binds to the {20-21} plane of ice. d: sea raven Type II AFP binds the {11-21} plane. All planes determined by ice etching techniques. Figure taken from Cheng and DeVries, 1991 (27).


a
c-axis
[0001]
a3
a2
a1
DM, RD, AB
adsorption plane
(10$\bar{1}$0)
b
c-axis
[0001]
a3
a2
a1
Sculpin
adsorption
plane (2$\bar{1}\bar{1}$0)
Prism plane
(10$\bar{1}$0)
c
c-axis
[0001]
a3
a2
a1
Flounder, Plaice
adsorption plane
(20$\bar{2}$1)
d
c-axis
[0001]
a3
a2
a1
Sea raven
adsorption
plane (11$\bar{2}$1)

from eel pout binds to the {10-10} prism planes (27). The binding of different planes of ice produces hexagonal bipyramidal ice crystals with different *c* to *a* axis ratios. The winter flounder Type I AFP produces ice crystals with a ratio of 3.3 to 1, referred to as simply "3.3".

Due to the asymmetric nature of the flounder Type I AFPs, it was possible to deduce a direction, <01-12>, to which the AFP binds on the {20-21} plane. Further ice etching experimentation with winter flounder AFPs indicated that the D-enantiomer binds the mirror image direction, <-1102>, as predicted, compared to the natural L-enantiomer which binds the <01-12> direction (102)(103).

MECHANISMS OF ADSORPTION TO ICE

Type I Antifreeze Protein Models

As the smallest and simplest AFP, Type I has been the most extensively studied AFP to date. The amino acid sequence of a Type I AFP isoform was originally identified in 1977 (32). Circular dichroism studies showed that most of the peptide is in an α helical form (46)(41). In a regular α-helix the Thr and Asx residues were predicted to be placed on the same side of the helix, 4.5Å apart (DeVries and Lin 1977). This spacing matches a repeat distance of 4.51Å between oxygens in the ice lattice on the ice prism planes (32)(95). This match was the first suggestion that binding of antifreeze peptides to ice occurs through hydrogen bonds between the hydroxyl and carboxyl groups on the hydrophilic side chains and the water molecules in the ice lattice

(32). A similar model for AFGP binding has been proposed whereby the hydroxyl groups of the dissacharide moieties, oriented by a poly-proline II helix structure, make a hydrogen bonding match to the lattice (32). These models both expose hydrophobic surfaces of the AF(G)Ps to the water, which was proposed to further inhibit water binding to ice and therefore help inhibit crystal growth.

Initial X-ray crystallographic studies resolved the 3D structure of Type I AFP to 2.5Å resolution (39). This confirmed the helical structure of the peptide. Unfortunately, at this resolution side chain atoms were not visible and the side chains were thought to be flexible in their location. A model for ice-binding was suggested by these authors in which the helix dipole moment was proposed to dictate the preferential alignment of the peptide to the (c-axis) prism plane of the ice nuclei by interaction with the ice dipole. Torsional freedom of the side chains was in this case proposed to facilitate hydrogen-bonding to the prism face in different directions dependent on AFP concentration (39).

Ice etching techniques later demonstrated that neither flounder, plaice nor sculpin Type I AFPs bind the prism faces of ice (27)(42). Focussing on the flounder AFPs, which bind the {20-21} pyramidal planes of ice in the <01-12> direction, a third model was proposed (42). This again was based on a hydrophilic residue match to the ice lattice. The 11-amino-acid repeat, of the flounder AFPs which places Thr residues equally spaced along the same surface of the helix every 16.5Å, matches the 16.7Å repeat spacing between ridges of water molecules along the <01-12> direction of the {20-21} plane of ice. Adsorption was proposed to occur via hydrogen bonding between the Thr residues and water molecules on the ice lattice (42). This model was further

refined by a structure/function study that confirmed the importance to ice binding, of not only the Thr residues, but also the regularly spaced Asx and Leu residues present in the winter flounder HPLC-6 isoform of Type I AFP (104). The topological ridge and valley repeat distance of 16.7Å matches not only the 16.5Å repeat distance between Thr (42), but also the same distance between repeated Asx and Leu residues along the helix (43). Computational modelling of the binding of HPLC-6 to ice supported Thr hydrogen-bonding to the 16.7Å spaced oxygens and also proposed a role for Asx as possibly making 2 hydrogen bonds to water molecules of a different rank than those bound by the Thr for a total of 13 hydrogen-bonds (43) (Figure 13). Leu was proposed to contribute binding energy through the formation of van der Waals contacts with the ice surface (43). A second similar model proposed a slightly different role for Asn as simply penetrating into the ice lattice (105), reducing the potential number of hydrogen-bonds to eight or nine. Two other computational docking models further corroborate the 16.5Å spacing match between Thr residues and the ice lattice. The model by Chou (106) presents a mechanism in which the Thr share a common rotamer and allow the peptide to "zipper" onto the ice surface. However, no attempt was made to ascribe roles to Asx or Leu residues in either of these latter two models (106)(107).

A common concern with these models is that there are insufficient hydrogen bonds (a maximum of 13) to irreversibly bind AFPs to ice. Through an extension to the adsorption-inhibition mechanism, a model was proposed in which co-operativity between neighbouring AFPs provides the energy required for binding to ice (43)(102)(108). According to this model, following initial AFP adsorption along <01-12> on the {20-

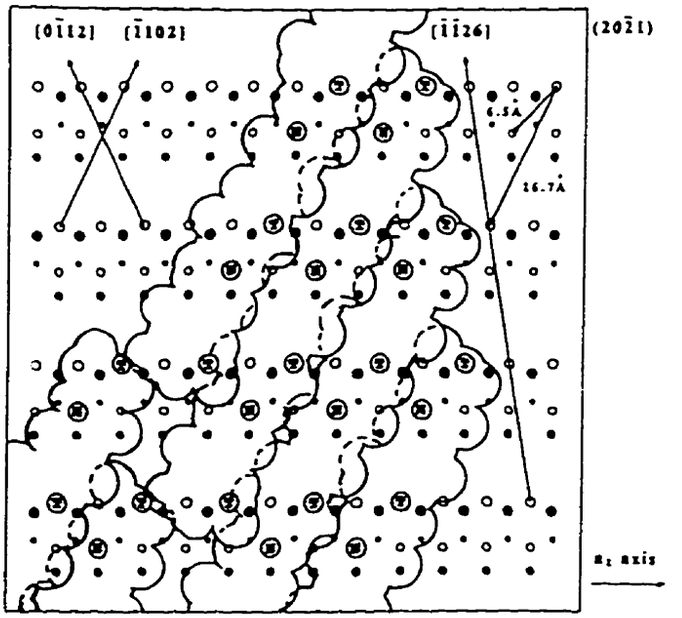
Fig. 13 **Computational model of HPLC-6 docked to ice.** HPLC-6 is depicted bound to the {20-21} plane of ice in the <-1102> direction. Asn and Thr are each predicted to make two hydrogen bonds to ice and Asp one bond. Leu side chains are suggested to make van der Waals contacts to ice in the valley of the "ridge and valley" topology of the ice plane. Figure taken from Wen and Laursen, 1992 (43).

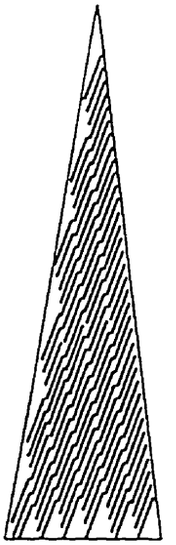
N
C
c-axis
a-axis

21} plane, more AFP molecules are directed to bind immediately adjacent to the first by hydrophobic interpeptide interactions and hydrogen bonding to ice (43) (Figure 14). The staggered alignment of the N-termini of the helices matches the <-1-126> direction of the edge of the bipyramidal crystal. According to this model permanent binding of AFPs to ice occurs when they form a patch, but not when they bind individually. This model was tested with only moderate success. Mutation of an Ala (A17 to Leu), on the side of the helix that would be making the hydrophobic contacts with other peptides, resulted in complete loss of activity (108) as predicted by the cooperativity model. However mutation of another Ala (A8), resulted in only a very minor loss of activity despite also being situated on a side of the helix proposed to make contacts with adjacent peptides (108). A second test of cooperativity, in which both D- and L-enantiomeric forms of HPLC-6 were synthesized and characterized, produced an equivocal result. It was predicted and later proven (102)(103) that the D-enantiomeric form bound to the mirror image direction of <01-12>. It was hypothesized that binding of a mixture of the D- and L- forms in the two different directions would interfere with cooperative binding in both directions and result in decreased activity. However no loss of activity was observed (102). This was not interpreted as evidence against cooperativity, but was explained by the formation of homogeneous patches of all L- or all D-enantiomers as a result of cooperative-binding (102).

A study, carried out in our laboratory, tested cooperativity by looking at the effect of mixing different types of AFPs together, on thermal hysteresis and ice crystal morphology (15). The mixtures of Type I, II and III AFPs produced ice crystals of

Fig. 14 **Proposed model of cooperativity.** top: Schematic view from above the {20-21} plane of ice. Filled circles represent surface oxygen atoms (largest are the highest). Open circles represent oxygens in the virtual ice layer above the ice surface. Those marked with T are Thr binding site, and those marked with N are Asx binding sites. The AFP van der Waals footprint is outlined. The potential cooperative packing arrangement of the AFPs is depicted producing the [-1-126] crystal edge alignment. bottom: complete coverage of the {20-21} plane of ice by cooperative AFP binding to ice. Figures taken from Wen and Laursen, 1992 (43).


$[0\bar{1}12]$
$[\bar{1}102]$
$[\bar{1}\bar{1}26]$
$(20\bar{2}1)$
6.5Å
16.7Å
a_1 axis

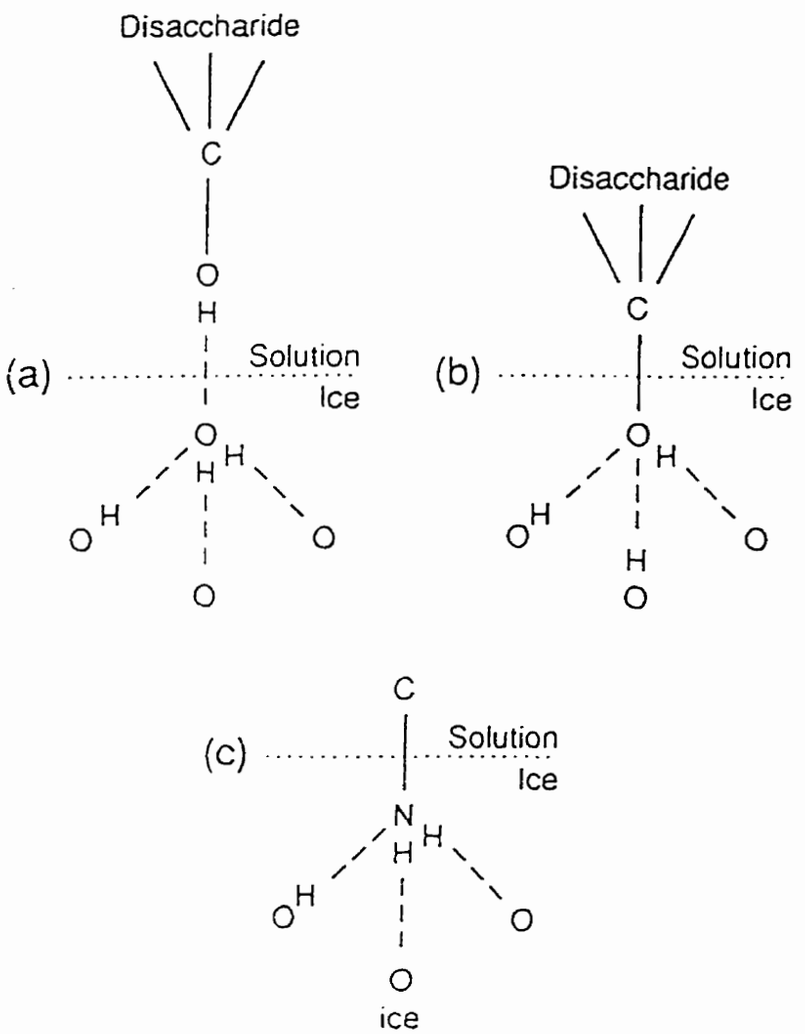


hybrid shapes and dimensions. The thermal hysteresis activity of the mixtures was found to be independent of the proportions of the different proteins present, each of equal specific activity. This indicated that the AFPs did not act to either activate or inhibit each others activity and suggested that protein-protein interactions are not required for ice binding.

Without invoking cooperativity, a second model attempted to explain "tight" AFP binding with few hydrogen-bonding groups, by modelling of the hydrophilic residues of the Type I AFP and the hydroxyl groups of AFGP disaccharides "within" the ice layer (Figure 15). This significantly increased the number of hydrogen-bonds between protein and the ice lattice by increasing the number of hydrogen bonds produced by each active group (eg. Thr would make three hydrogen-bonds instead of just two). This provided a second possible explanation for permanent binding of the AFP and subsequent inhibited of crystal growth (28).

However, in order for the functional groups of the side chains to occupy lattice oxygen positions as described in the last model (28), the side chains would have to be exceptionally well exposed. In their recent publication of the HPLC-6 high resolution X-ray structure where the side-chains were visable, Sicheri and Yang (40) showed that the surface containing Thr, Leu and Asx residues was very flat, and that the Thr hydroxyl groups did not protrude far enough to integrate into the ice lattice. In this structure all four Thr residues were observed to adopt a rigid side chain conformation with a $\chi 1$ rotation of -60°. This is the least accessible rotamer for external hydrogen-bonding. In this conformation the Thr residues are better positioned to form intra-helix

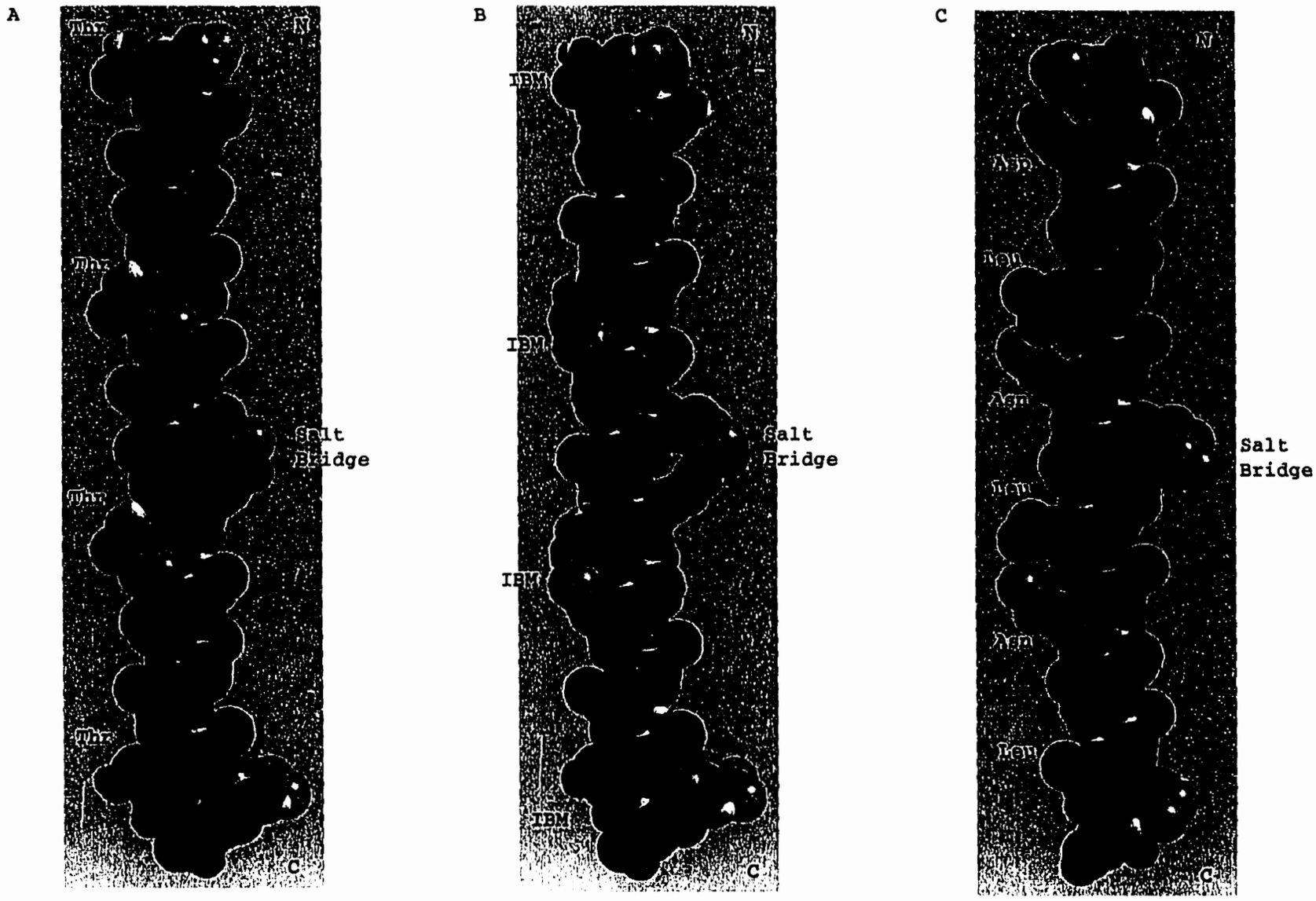
Fig. 15 **Side chain burial model.** a: A single hydrogen bond is made from the AFP hydroxyl group above the ice/water interface to the ice lattice below. b: The hydroxyl oxygen of the AFP is viewed as taking the position of one of the water oxygens within the ice lattice. Two addition hydrogen-bonds are made in this situation. c: A similar situation is described for the nitrogen of an amide group on an AFP, where the nitrogen takes the place of a water oxygen within the ice lattice. Figure taken from Knight et al, 1993 (28).


Disaccharide
C
O
H
(a)
Solution
Ice
O
H
H
O
H
O
O
Disaccharide
C
(b)
Solution
Ice
O
H
H
O
H
O
H
O
C
(c)
Solution
Ice
N
H
H
O
H
O
O
ice

bonds. The Asx residues were proposed to hydrogen-bond to the same rank of water molecules as the Thr residues, constrained through steric interactions with the $i+4$ side chain. In this model, Leu interacts with Asx to help stabilize the ice binding motif. Overall the rigidity and common orientation of these side chains are postulated to be critical features for ice binding (Figure 16). In contrast to this, solution NMR studies near the freezing point indicated that the ice-binding side chains of Type I AFP were free to adopt other conformations (109). The two central Thr (T13 and T24) were observed to have a 55% preference for the same -60° rotamer observed in the X-ray structure. However all four Thr sampled many rotameric states in this study. This solution study also provided evidence for a minor interaction between Leu and Asx, but not to the extent that they are fixed in a rigid conformation in solution (109) as proposed in the crystal structure (40).

Overall a number of different models for binding of Type I AFP to ice have been proposed. These models have gradually progressed over the decades, becoming more specific and refined with each new discovery about Type I AFP. The original Lin and DeVries model (32) was proposed prior to any information on ice-binding sites or modelling. Ice etching studies (27)(42) provided data critical to resolving the complete mechanism by providing a specific ligand for the AFPs. This allowed the computational modelling and docking of AFPs to ice (43)(105-107). However it was not until the most recent X-ray and NMR studies were completed (40)(109) that the AFP side chains were actually visualized. All these models are based on the assumption that the AFP is interacting with ice primarily through hydrogen-bonds between the hydrophilic side

Fig. 16 **HPLC-6 space filling structures.** Three views are shown, each rotated approximately 20° around the helix axis, with respect to each other. Motion of the rotation can be followed in the diagrams from the relative positioning of the salt bridge on the right hand side of the molecule. A: View of the maximal extension of the Thr residues on the ice binding surface. B: View of the ice-binding surface on the left hand side of the molecule. C: View of the maximal extension of the Leu and Asn residues projecting from the helix. Structural Brookhaven file identifier 1wfa. 2° structure was assigned using STRIDE (47). Diagrams were produced using Molscript (48) and Raster3D (49) on a Silicon Graphics terminal, printed on an OptraC LexMark colour laser printer.


A
N
Thr
Thr
Salt
Bridge
Thr
Thr
C
B
N
IBM
IBM
Salt
Bridge
IBM
IBM
C
C
N
Asp
Asn
Asn
Salt
Bridge
Asn
Asn
Asn
C

chains and the ice lattice water molecules. Very little has in fact been proven with respect to the precise role of the Thr, Asx and Leu residues in mediating the interaction.

Type III Antifreeze Protein Models

The preceding Type I models (and similar AFGP models) are all based on an ice lattice match to the regular repetitive structure of the antifreeze. Sequence repetition is not a characteristic of either Type II or III AFPs, and therefore is not available to suggest a way in which these globular AFPs bind to ice. The low precision solution structure of the QAE isoform of Type III AFP was originally solved by NMR analysis, which suggested a β sandwich-type fold in which six out of eight β strands formed two sheets of three antiparallel strands (71). Two hydrophilic surfaces were identified, one on each of the triple sheets, which exposed six to seven hydrophilic residues, that were potentially important for making hydrogen-bonds to the ice surface. The protein for the structural study was produced in a recombinant *E.coli* over-expression system using a synthetic gene with optimal *E.coli* codon usage (110) and added unique restriction sites (70). This expression system allowed easy mutagenic analysis for studies on the mechanism of protein action.

The cluster of hydrophilic, surface accessible residues near the C-terminal triple-stranded sheet were noted to be conserved in Type III isoforms (74). Several residues (N14, T18, Q44 and N46) were targeted for mutagenesis. Mutations that changed the position of the putative hydrogen-bonding group (eg. amide and hydroxyl exchanges) caused significant decreases in activity. Other changes (eg. amide to carboxylic acid)

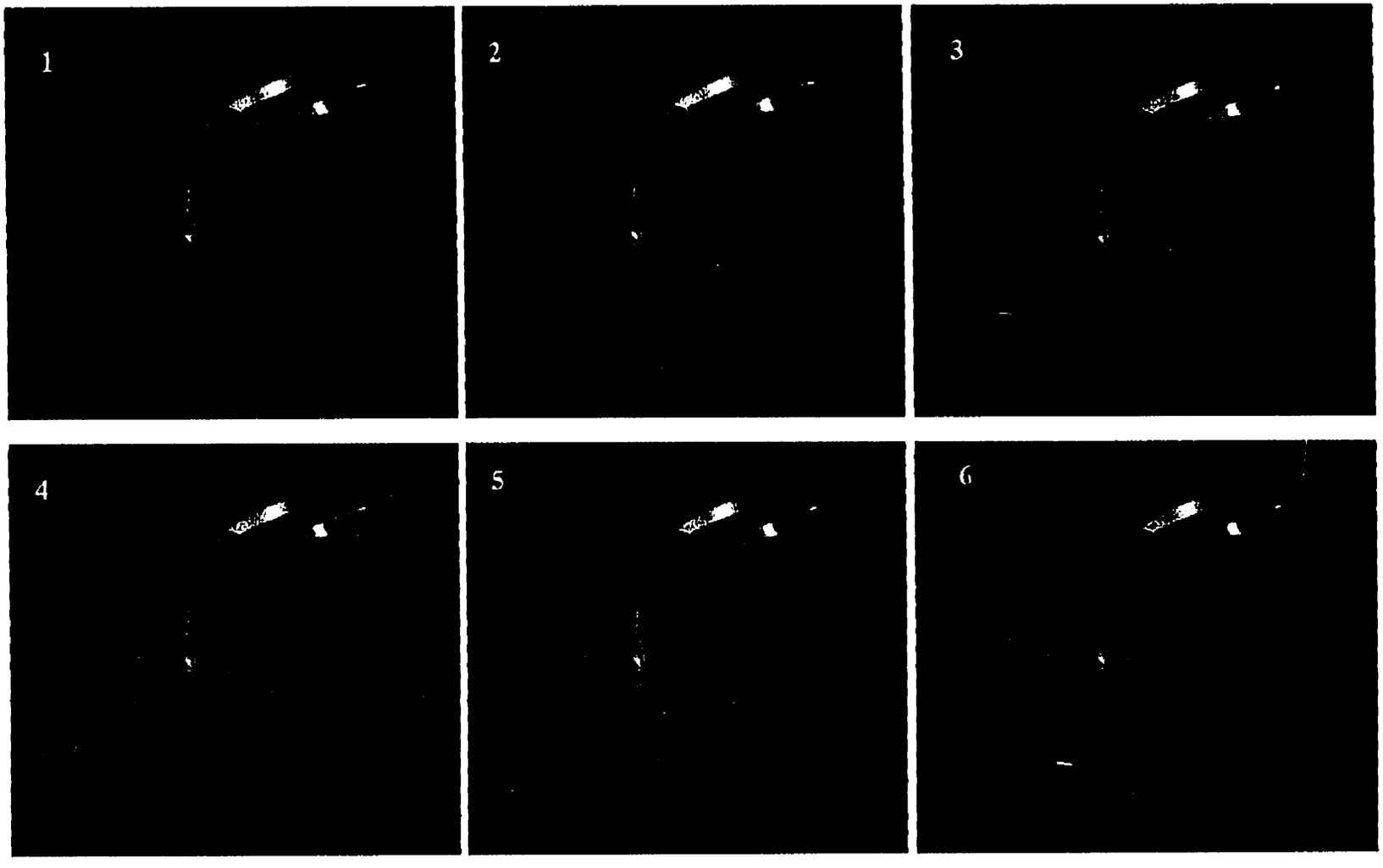
produced AFPs that were completely active, even at pH values on either side of the pKa value (74). Of the targeted residues, Q44, N14 and T18 were shown to be essential for activity. A triple mutant (Q44T, N14S, T18N) combining alterations at all three sites was completely inactive, but retained its 3D fold (74).

This identified surface of Type III AFP was further probed by a study in which A16, situated centrally to the three active residues, was systematically replaced by amino acids with larger side chains (16). The relationship between the size of the intruding side chain and AFP activity was complex. Branched and cyclic residues were more deleterious to the thermal hysteresis activity than straight chain residues (eg. Met). Observation of ice crystal morphology indicated that mutations with poorer thermal hysteresis activity formed ice crystals with longer *c* to *a* axis ratios, ranging from 2 in wild type to 10 in the least active mutant (A16H)(16). An increase in the ratio was also observed on dilution of the wild-type antifreeze. This differs from the behaviour of ice crystals in the presence of mutant (104) or dilute solutions of Type I AFP (16), which retained a *c* to *a* ratio of 3.3. This difference, in combination with the ice etching studies that indicate Type III binds principally to the prism plane and not a pyramidal plane (27), requires an explanation for the shaping of the ice into a bipyramid. This prompted development of the first model for binding of a globular AFP to ice (16). The model was based on a variation of the "step-growth-inhibition" mechanism. In this model the Type III AFP binds to the prism plane while covering the basal plane. Dissociation or overgrowth of the AFP allowed "filling in" of occupied niches at the prism surface, causing the ice crystal to form with a larger *c* to *a* axial ratio (16).

A third study on Type III was carried out to test the cooperativity model for antifreeze action (43)(102)(108). In this study, large domains were attached by cassette mutagenesis to the N-terminal end of Type III AFP. Two domains of different sizes were tested: maltose-binding protein at 42 kDa and thioredoxin at 12 kDa. The hypothesis was that large domains would sterically inhibit any cooperative effect between AFPs and should cause a decrease in AFP activity. In fact, the larger the domain, the more active was the AFP. This refutes the cooperativity model and also suggests that size is important for AFP activity (111). It was proposed that the larger the AFP diameter, the greater the potential increase in ice curvature between AFPs, at equivalent ice coverage, which according to the Kelvin effect will increase the surface free energy and depress the freezing point further (111)(100).

Finally, when the X-ray crystal structure of Type III AFP (QAE isoform) was resolved to 1.25Å (73), the structure revealed a remarkably flat, amphipathic ice binding surface (Figure 17). This degree of planarity had not been obvious from the low precision NMR solution structure (71). Five hydrogen bonding atoms were modeled to match two ranks of oxygens spaced 4.5Å apart on the prism face, consistent with high ice binding affinity and specificity. It was noted that N14 was not in the same ice binding plane; rather it was at the transition point between prism and basal planes. A proposal was made that initial binding through this one residue could slow the lateral spread of a single ice layer over the basal plane such that a second layer of ice could catch up to the first, making contact with the second ice-binding residue, Q44. At these early stages the AFP is predicted to have a significant off-rate. Growth of subsequent

Fig. 17 **Model of Type III AFP docking to ice.** 1: Type III AFP makes the first hydrogen bond to an expanding ice layer via N14. 2: This slows the growth of the ice layer such that a second layer catches up to the first and makes a second interaction with Q44 on the AFP. 3-6: This continues until all ice binding residues are bound to the prism face of the ice crystal and growth is inhibited. Figure courtesy of C. DeLuca.


1
2
3
4
5
6

ice layers continues to catch up to the AFP making more and more contacts with it, until all binding sites on the AFP are filled and the ice growth is completely inhibited (73) (Figure 17). In this model the AFP is to a certain extent inducing formation of its own binding site. Although the prism plane is a naturally expressed ice crystal plane, binding solely of this plane would not cause the tapering of the ice crystal that is always observed with this AFP.

RESEARCH GOALS

The overall goal of this research was to learn more about the general mechanisms involved in mediating AFP activity. In working towards that goal, mechanistic analyses on two structurally disparate fish AFPs have been carried out. For reasons of simplicity, the remaining chapters of this text have been divided into parts A and B. Part A deals with the structure/function analysis of sea raven Type II AFP, while part B reports on the mechanistic analysis of winter flounder Type I AFP, isoform HPLC-6.

PART A Objectives

Type II AFP is one of a number of globular proteins known to exhibit antifreeze activity. A recent homology modelled structure for Type II AFP indicates it shares no structural similarity to either the more extensively studied Type I and AFGP proteins, both of which have a very regular repetitive structure (1), or to the globular Type III AFP (53). Recent models for the mechanism of action of Type III and Type I AFPs indicate these proteins are likely to function though somewhat different ice binding

geometries (73)(43). Type II SRAFP does not bind to the same ice crystal surface as either of these AFPs (27), and it maintains a distinct, rounded hexagonal bipyramidal shaped ice crystal (15) not observed with either Type I or Type III or any of their mutants. This evidence strongly suggests yet another structural fit between AFP and the ice lattice.

SRAFP was specifically targeted for this study because it is the most active of the Type II AFP isoforms and the only one to show Ca^{2+}-independent activity (51)(52). As well, at the start of this particular project, it was the only Type II AFP for which genomic and cDNA sequences were known (65)(59). Goals for the Type II AFP structure/function study are outlined below.

1. Elucidation of the 3D structure of Type II AFP from sea raven. The solution structure and crystal structure of the protein are being attempted for comparison to the existing 3D models of SRAFP and herring AFP, as well as to the CRD and PSP structures. Studies on both the Type I and Type III AFPs indicate that significantly different, yet equally useful information can be derived about a protein from the application of both of these different techniques (73)(72)(40)(109). The solution structure is being solved in collaboration with Drs. Frank Sonnichsen[1], Brian Sykes[2], and Wolfram Gronwald[2], at [1]Case Western Reserve University and the [2]University of Alberta. The crystal structure is being attempted in collaboration with Dr. Zongchao Jia at Queen's University. Protein for both structural studies was provided though the development of a high level recombinant expression system, using the yeast *Pichia pastoris* (*P. pastoris*) as host. In collaboration with Dr. Andrew Daugulis in the Department of Chemical

Engineering at Queen's University, sophisticated fermentation techniques were employed to achieve levels of AFP production sufficient for crystallization (> 100mg) and for metabolic labelling of the protein for NMR studies.

2. Determination of the mode of binding of Type II AFP to ice. Mutagenic structure/function studies were carried out on Type II SRAFP to determine the importance of various hydrophilic residues to AFP activity. Single site changes, insertions and a domain swap were made to assess the putative ice binding site (IBS1) proposed in the SRAFP 3D model (53). Computer modelling was used to study the localized binding surfaces of SRAFP, CRD and PSP models to further analyze the specific spatial arrangements of residues, to better determine what features are important for binding. The computational analysis lead to the identification of a potential second binding site on the Type II molecule, which was further probed by more extensive mutagenesis. The *P. pastoris* recombinant expression system was used for production of mutant AFPs.

PART B Objectives

Type I AFP, specifically the HPLC-6 isoform from winter flounder, has been extensively studied. However at this time, specific roles for the ice binding residues have yet to be determined. It has simply been assumed for the last two decades that protein/ice interactions are predominantly, if not exclusively, mediated through hydrogen-bonds between hydrophilic amino acid side chains and water molecules of the ice lattice. One goal of this study was to determine more precisely the roles carried out by each

individual member of the IBM.

HPLC-6 was made by solid-phase peptide synthesis. Subsequently a number of variants were made in which the inner two identical IBMs were changed with the same substitution. The role of the Thr residue was probed by specific replacement with Ser and Val residues that maintain either the hydroxyl or the methyl group. Matching replacements at the Leu and Asn positions were also carried out. Other experiments were done to shorten the helix and determine the minimum size of the AFP for ice binding. And finally, a domain addition experiment was made to further test the cooperativity theory and to work towards making an effective synthetic AFP.

CHAPTER 2

MATERIALS AND METHODS

PART A

Bacterial and Yeast Strains

The methylotrophic strain of *P. pastoris*, GS115(*his4*) (Invitrogen, San Diego, CA)(112) was used for expression of recombinant Type II AFP. *E. coli* JM83 was used for all plasmid constructions and propagations. *E. coli* CJ236 was used for production of uracil-containing single-stranded DNA for mutagenic experiments.

Media Compositions

i) *E. coli*:

E. coli was grown in Luria Bertani (LB) (Difco) broth or on LB agar plates, containing 10g/L tryptone, 5g/L yeast extract and 10g/L sodium chloride. Ampicillin (75μg/mL) was added to media for selection of ampicillin-resistant plasmid-containing clones.

ii) Yeast:

Buffered Minimal Glycerol-complex medium (BMGY) contained the following (per litre): yeast extract (10g), meat peptone (20g)(Sigma), yeast nitrogen base without

amino acids (13.4g)(Difco), biotin (400μg)(Sigma), glycerol(10mL) in 100mM potassium phosphate buffer (pH 6.0). Buffered Minimal Methanol-complex medium (BMMY) contained the same components as BMGY except that glycerol was replaced by methanol (5mL/L). Minimal Methanol medium (MM) contained the following (per litre): yeast nitrogen base without amino acids (13.4g), biotin (400μg), methanol (5mL). Minimal Glycerol media (MG) contained the same as MM media except methanol was replaced by glycerol (10mL/L). Minimal Dextrose medium (MD) contained the same as MM medium except methanol was replaced with dextrose (10g/L). Agar plates of these media were made with the addition of 12-15g/L agar. Fermentation media contained the following (per litre): glycerol (50g), ammonium sulfate (20g), KH_2PO_4 (12g), $MgSO_4$-$7H_2O$ (4.7g), $CaCl_2$-$2H_2O$ (0.36g), plus trace elements as follows: $CaSO_4$-$5H_2O$ 0.2μM, KI 1.25μM, $MnSO_4$-$4H_2O$ 4.5μM, Na_2MoO_4-$2H_2O$ 2.0μM, H_3BO_3 0.75μM, $ZnSO_4$-$7H_2O$ 17.5μM, $FeCl_3$-$6H_2O$ 44.5μM. The pH was adjusted to 5.5 using 5M KOH.

Plasmid Isolation from *E. coli*

Plasmid DNA was prepared from 10mL LB ampicillin overnight cultures by alkaline lysis followed by two equal volume phenol/chloroform (1:1) extractions and two equal volume chloroform extractions and then precipitated with an equal volume of isopropanol (113). Plasmid DNA for sequencing and yeast transformations was prepared using the Wizard™ miniprep DNA purification system (Promega).

Gel Purification of DNA Fragments for Cloning

Restriction enzyme-digested DNA fragments were separated by gel electrophoresis. Fragments longer than 800 base pairs (bp) were isolated by the Prep-a-Gene™ (Biorad) DNA purification kit. Fragments shorter than 800bp were eluted from 1.0% agarose gels using DEAE paper Na 45 (Schleicher & Schuell) (114).

Double-Stranded DNA Sequencing

DNA sequencing was performed using the Sequenase® Version 2.0 DNA sequencing kit (USB)(115). Sequencing primers are listed in Table 1. Sequencing reactions were separated on a 47% w/v urea, 8% polyacrylamide gel (1500V, 25-35mA) and visualized by autoradiography (Kodac XAR) after fixing (10% methanol and 10% acetic acid) and drying.

Plasmid Construction

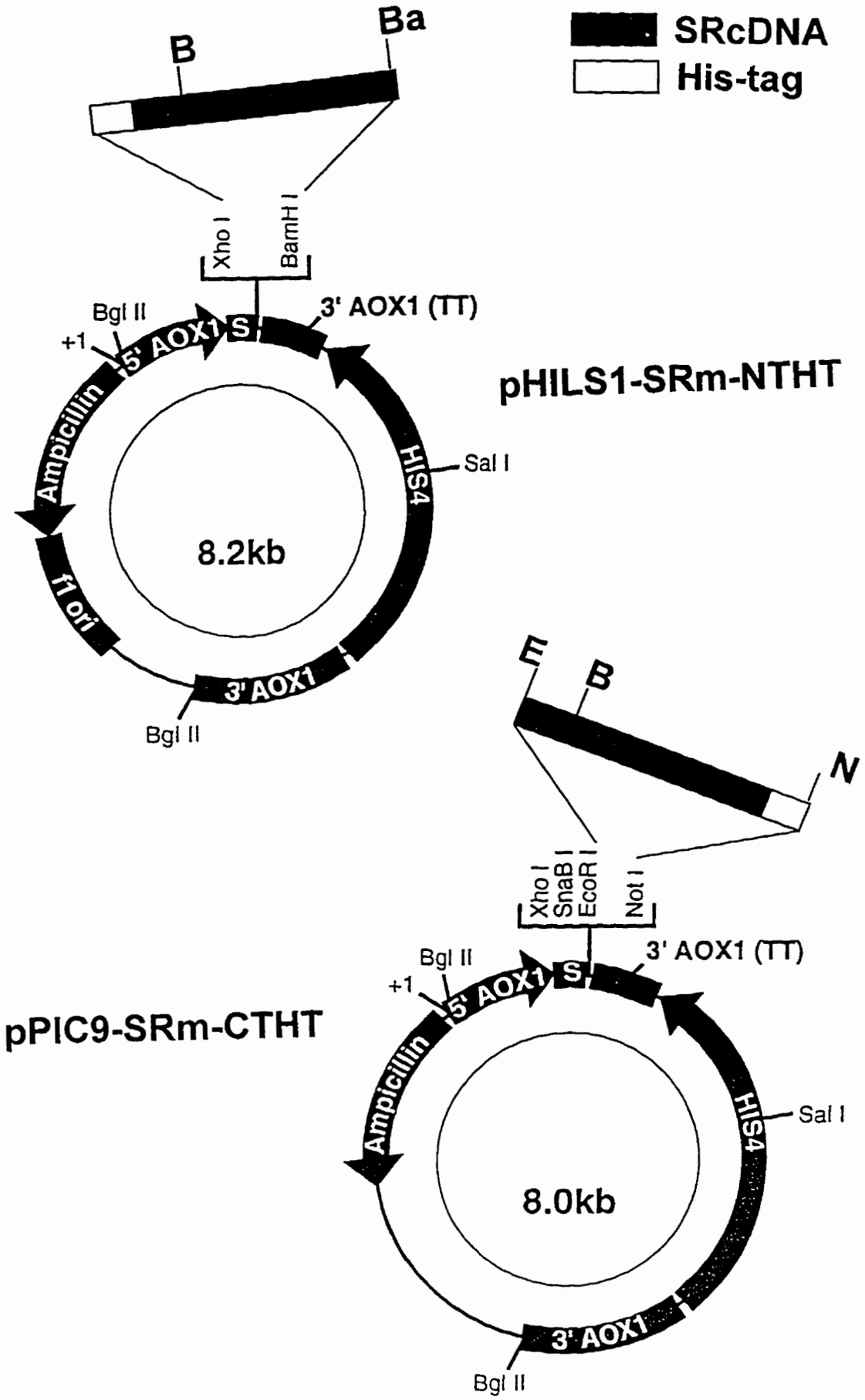
i) pPIC9-SRm-CTHT:

The vector pPIC9 (Figure 18), which contains the HIS4 gene for selection in the HIS4$^-$ strain, GS115(*his4*), and the ampicillin-resistance gene for selection in *E. coli* (112), was used to express the mature sea raven cDNA in *P. pastoris*. The 5'-end primer, SR1 (5'-GAATTCGAATTCCAGAGAGCCCCACCA-3'), which introduced tandem *Eco*RI sites (underlined) immediately upstream of the Gln codon at the N-terminus of mature SRAFP (59), and the 3' primer, SR2 (5'-ATACGTA

Table 1. Sequencing Primers

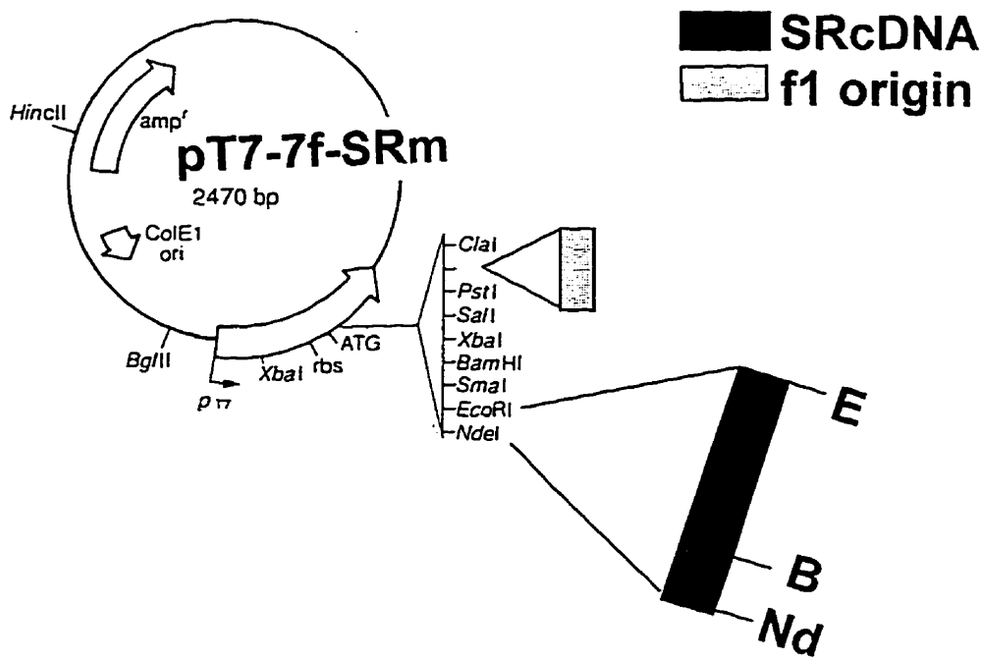
Name	(5'-3')	Vector
5'-antisense strand sequencing primers		
T7 promoter	TTTAATACGGACTCACTATA	pET/pT7-7f
α-mating	TATTGCCAGCATTGCTGC	pPIC9
Pichia	GACTGGTTCCAATTGACAAGC	pHIL-S1
3'-sense strand sequencing primers		
T7 terminator	GCTAGTTATTGCTCAGCGG	pET/pT7-7f
sea raven 3'	ACTAAAGTGTTTGTGTG	pET/pT7-7f/pPIC9/pHIL-S1
3'AOX	GCAAATGGCATTCTGACATCC	pPIC9/pHIL-S1

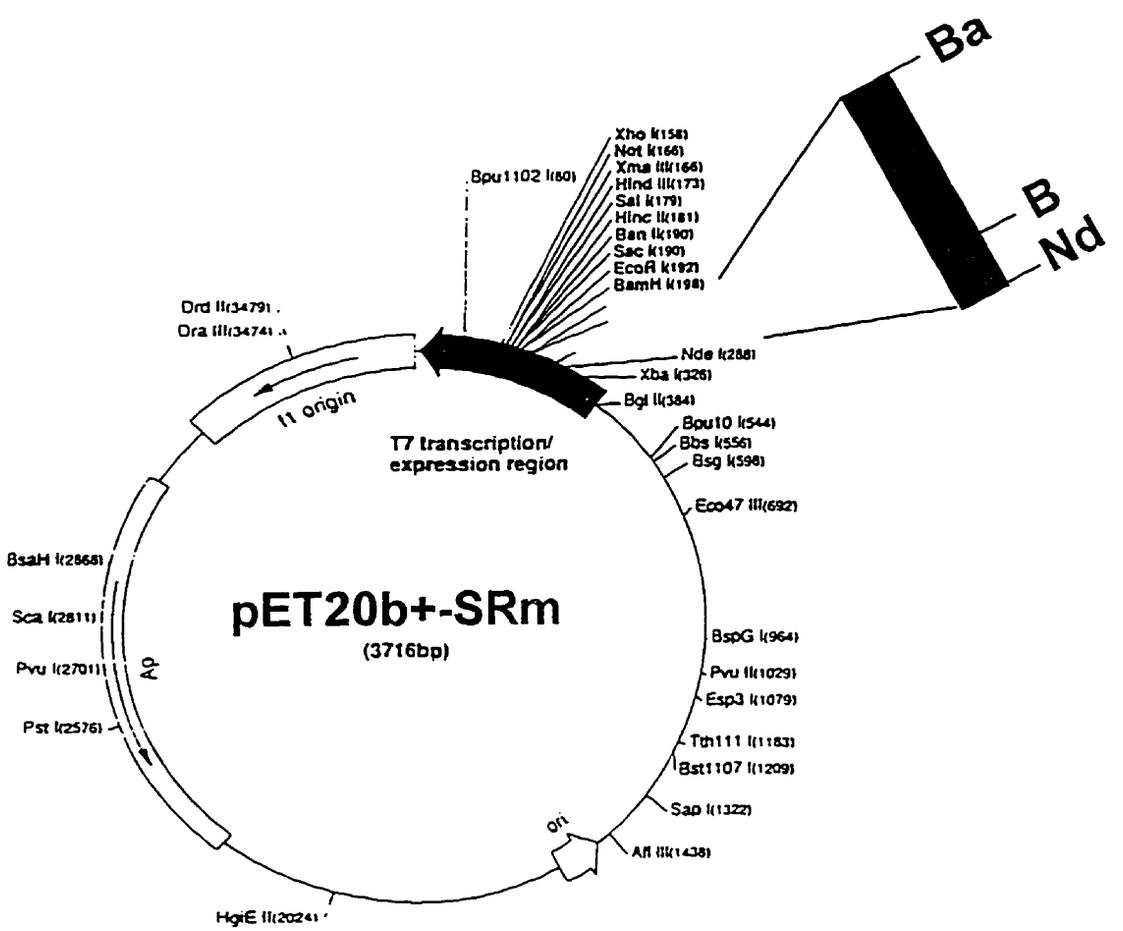
Fig. 18 **pPIC9 and pHIL-S1 vector diagrams.** top: pHIL-S1-SRm-NTHT final construct. The *Xho*I of the multiple cloning site is deleted in the final construct by incorporation of the N-terminal His-tag. bottom: pPIC9-SRm-CTHT final construct. B=*Bst*EII, Ba=*Bam*HI, E=*Eco*RI, N=*Not*I. Vector diagrams are from Invitrogen *P. pastoris* system manual (117).


SRcDNA
His-tag
B
Ba
Xho I
BamH I
Bgl II
+1
5' AOX1
S
3' AOX1 (TT)
Ampicillin
HIS4
Sal I
8.2kb
f1 ori
3' AOX1
Bgl II
pHILS1-SRm-NTHT
E
B
N
Xho I
SnaB I
EcoR I
Not I
Bgl II
+1
5' AOX1
S
3' AOX1 (TT)
pPIC9-SRm-CTHT
Ampicillin
HIS4
Sal I
8.0kb
3' AOX1
Bgl II

AGCGGCCGCCAAACACCCACTAAAGTG-3'), which introduced a *Not*I site (underlined) 47 bp downstream of the stop codon, were used to amplify the mature AFP cDNA by the polymerase chain reaction (PCR) from the clone pT7-7F-SRm (Figure 19). pT7-7F-SRm is a derivative of the *E. coli* expression vector pT7-7 (116), in which the mature cDNA of Type II AFP was cloned into the *Nde*I/*Eco*RI sites of pT7-7F. The pT7-7F vector contains an *ori*F site for production of single-stranded (ss) DNA as previously described (114). A 492 bp PCR product was obtained by DNA amplification and was cloned into the *Eco*RI and *Not*I sites of the pPIC9 vector, 24 bases (8 amino acids) downstream from the α-mating factor signal peptide cleavage site, to give pPIC9-SRm. Subsequently, the vector pT7-7F-SRm was mutated using a primer-directed mutagenic method (119) with primer CTHT1 (5'-pTGCGCCATGACATT CCACCACCACCACCACCAC**TAA***GCGGCCGC*TGAGCTAACACAGAGG-3'),which introduced a 6-histidine-long His-tag (underlined) at the C-terminal end of the cDNA, immediately prior to the stop codon (**bold**) to allow simplified protein purification by affinity chromatography on a nickel-agarose column (Novogen). The primer also adds a *Not*I site (*italics*) immediately following the stop codon. A fragment, that included the His-tag, was excised using the *Not*I site and a unique *BstE*II site farther upstream within the cDNA. This was cloned into the corresponding sites within the pPIC9-SRm vector to produce the final expression construct, pPIC9-SRm-CTHT, which differed from pPIC9-SRm by the presence of the C-terminally coded His-tag.

Fig. 19 **pT7-7f and pET20b+ vector diagrams.** top: pT7-7f-SRm final construct. Vector diagram is from Tabor et al., 1985 (116) bottom: pET20b+-SRm final construct (Vector diagram is from Novogen pET expression system manual (118). Nd=*Nde*I, B=*Bst*EII, E=*Eco*RI, Ba=*Bam*HI.


SRcDNA
f1 origin
pT7-7f-SRm
2470 bp
HincII
amp'
ColE1 ori
BglII
XbaI
rbs
ATG
pT7
ClaI
PstI
SalI
XbaI
BamHI
SmaI
EcoRI
NdeI
E
B
Nd


Ba
B
Nd
Xho I(158)
Not I(166)
Xma III(166)
Hind III(173)
Sal I(179)
Hinc II(181)
Ban I(190)
Sac I(190)
EcoR I(192)
BamH I(198)
Bpu1102 I(60)
Nde I(288)
Xba I(326)
Bgl II(384)
Bpu10 I(544)
Bbs I(556)
Bsg I(598)
Eco47 III(692)
BspG I(964)
Pvu II(1029)
Esp3 I(1079)
Tth111 I(1183)
Bst1107 I(1209)
Sap I(1322)
Afl III(1438)
ori
f1 origin
T7 transcription/
expression region
pET20b+-SRm
(3716bp)
Ap
BsaH I(2568)
Sca I(2811)
Pvu I(2701)
Pst I(2576)

ii) pHIL-S1-SRm-NTHT:

The vector pHIL-S1 (Figure 18) was also used to express mature sea raven cDNA in *P. pastoris*. The vector pT7-7F-SRm was mutagenized using primer directed mutagenesis with primer pHIL1 (5'-pGGAGATATACATGCTCGAGCCCCACCAA-3'), which incorporated a *Xho*I site (underlined) at the 5'end of the mature cDNA, while deleting the first glutamine at the N-terminal end of the protein. A *Xho*I/*Bam*HI fragment was cut from this mutagenized pT7-7F-SRm vector and subsequently cloned into the pHIL-S1 vector at the corresponding *Xho*I and *Bam*HI sites, producing the vector pHIL-S1-SRm. An N-terminal His-tag (underlined) was subsequently added to this construct by insertion of the NTHT-linker (5'-pTCGAATTCACCACCACCACCAC CACGG-3') complimentary to (5'-pTCGACCGTGGTGGTGGTGGTGGTGAAT-3') which contains *Eco*RI and *Xho*I overhangs for cloning immediately 5' to the AFP cDNA insert. This inserted the coding sequence for a His-tag at the N-terminal end of the protein, producing pHIL-S1-SRm-NTHT.

Transformation of DNA

i) *E. coli*:

Plasmids were introduced into *E. coli* by the calcium chloride method of transformation (120). Cells were grown to log phase and chilled on ice. These were centrifuged and resuspended in 1/5th volume of chilled 0.1M $CaCl_2$ and incubated on ice for 15min. Cells were centrifuged again and resuspened in 1/15th volume of 0.1M $CaCl_2$ and incubated on ice for 30min. An aliquot (50μL) of cell solution was added to

10-40ng DNA and incubated on ice for 30min. The DNA/cell solution was then incubated at 42°C for 90sec and LB broth was added to 200μL volume and incubated at 37°C for a further 60min. Cells were plated on LB medium with appropriate antibiotic for plasmid selection.

ii) Yeast

Vector DNA (10μg) produced in *E. coli* was linearized and transformed into GS115(*his4*). Vector linearized by *Bgl*II was integrated by homologous recombination into the AOX1 gene site producing clones with a His^+, "methanol utilization slow" (Mut^s) phenotype. Vector linearized at the *Sal*I site was integrated into the HIS4 gene site of the GS115(*his4*) genome, without disruption of the HIS4 plasmid gene, gave a His^+, Mut^+ phenotype. The linearized vector was introduced into *P. pastoris* using either the spheroplast transformation procedure essentially as described by Cregg *et al.*, 1985,(112) or by electroporation as described in Scorer *et al.*, 1994,(121) and then plated on MD plates for selection of HIS4-containing clones

Selection of Positive Yeast Transformants

Mut^s clones were selected first by plating of yeast transformation reactions on MD (no His) plates which allowed for selection of clones containing the HIS4 gene. Clones were picked off these plates and replica-plated onto MD and MM plates for selection of Mut^s phenotypes. Colonies that grew well on MD (dextrose as carbon source) but slowly on MM (methanol as carbon source) were selected and tested for AFP production. In

the case of Mut^+ selection, clones that grew well on both MM and MD were selected and tested for production of recombinant SRAFP (rSRAFP).

Induction of Yeast Alcohol Oxidase Promoter for AFP Expression

Transformant cultures of volume 10mL or 50mL were grown in 50mL conical tubes (with loosened tops) or shake flasks (250mL Erlenmeyers), respectively, at 30°C for 48h in BMGY which contained glycerol and therefore inhibited the alcohol oxidase promoter. Cells were harvested by centrifugation and resuspended in the same volume of BMMY, which contained methanol for induction of the alcohol oxidase promoter, and again incubated for 48h at 30°C with shaking. Cells were removed by centrifugation, leaving rSRAFP to be purified from the medium.

Fermentation

Inoculum cultures (600mL) were grown in MG medium for 48h at 30°C, in 2L Erlenmeyer flasks, at 250rpm. This 10% inoculum was used to inoculate a 10L Chemap bioreactor (Model FZ-3000), containing 6L of fermentation medium. Cell culture conditions were set at 30°C, 800rpm and 2.3 L-air/L-culture volume/min. The pH was maintained at 5.5 by the automatic addition of 5M KOH. The level of dissolved oxygen was measured using an Ingold galvanic electrode. Antifoam 204 organic (2mL)(Sigma) was added prior to inoculation, and as required throughout the fermentation. Methanol and glycerol levels were monitored by HPLC (Waters) on a Sugar-Pac (Waters) column using the Maxima 800 data acquisition system for peak integration. The different

fermentation strategies used are described in the Results section.

Production of ^{15}N-Labelled rSRAFP

The clone GS115 pPIC9-SRm-CTHT Mut^{s} was grown in 10mL BMMY in a 50mL conical tube at 30°C with shaking for 24h. This was used to inoculate a 100mL culture (in a 250mL Erlenmeyer) of MG media containing 3.6g/L yeast nitrogen base (Difco) without amino acids or ammonium sulfate, and 10g/L ^{15}N-labelled ammonium sulfate (CIL). This culture was grown at 30°C, with shaking at 250rpm for 48h and then used to inoculate a 1.5L New Brunswick Scientific Bioflow Fermentor (Model C-30), containing 1L of fermentation medium with ^{15}N ammonium sulfate as the sole nitrogen source. Cell culture conditions were set as described for the 10L Chemap Bioreactor and pH was again maintained at 5.5 by the addition of 5M KOH. The mixed-feed fermentation strategy described in the Results section was used to induce expression and secretion of the uniformly ^{15}N-labelled rSRAFP.

Analysis of Type II Antifreeze protein

i) SDS-PAGE and Western Blotting

Samples were analyzed by electrophoresis on 17% polyacrylamide-SDS gels, containing 0.1M sodium phosphate and 4M urea at pH6.8. These gels were either stained with Coomassie blue (0.25%) and amido black (0.1%) and then destained (10% methanol, 10% ethanol), or were Western blotted onto a nylon membrane (PVDF, NEN). The membrane was incubated with rabbit anti-SRAFP antiserum and then

horseradish peroxidase-linked goat anti-rabbit IgG (Promega). Detection was carried out using enhanced chemiluminescence (Amersham).

ii) Protein Mass Spectrometry and Quantitation

Mass spectrometric analysis was carried out on a Kompact MALDI III spectrometer by matrix assisted laser desorption/ionization time of flight mass spectrometry in 0.1% TFA. AFP was quantified by estimation from Western blot analysis, by visual inspection and comparison of band intensities generated by known quantities of SRAFP. Dry weight, A_{280nm} (using the predetermined standard for Type II AFP 4.5mg/mL=10_{OD}) and Bradford protein detection (122) were also used to quantify purified AFP.

iii) Antifreeze Activity Measurements and Photomicroscopy

Thermal hysteresis was used as a quantitative measure of AFP activity. It is defined as the temperature difference (°C) between the melting point and the non-equilibrium freezing point of a solution. This thermal hysteresis activity is primarily the result of inhibition of crystal growth by the AFP (88). Thermal hysteresis can be measured by placing nanolitre volumes (1-10μL) of an AFP solution on a cooling stage beneath a microscope for observation of ice crystals. The temperature of the stage is controlled by a nanolitre osmometer (Clifton Technical Physics, Hartford, New York) (1Osmol=1.86°C) (45). The growth of a single ice crystal in the presence of AFP can then be observed as the temperature is dropped from the melting point (where ice and

its melt are at equilibrium) to the freezing point (rapid ice growth). All measurements were made in 100mM NH_4HCO_3, except for the Type II pH and DTT studies. Ice crystal morphology was observed through a Leitz Dialux 22 microscope and recorded by a Panasonic CCTV camera linked to a JVC Super VHS video recorder. Still images were obtained from a Silicon Graphics INDY terminal using IRIS Capture Version 1.2 and printed on an HP Laserjet4. For ice growth rate analysis, samples were held at a fixed degree of undercooling and images were captured at 0min and 10min time points.

iv) NMR Characterization

1D and 2D-NOESY NMR spectra of purified recombinant, ^{15}N-labelled recombinant and some mutant SRAFPs were collected and analyzed by Drs. Frank Sönnichsen and Wolfram Gronwald at the University of Alberta, Edmonton, to compare their overall fold to that of wild-type SRAFP isolated from sea raven serum.

Purification of Type II AFP

The yeast media (BMMY) containing SRAFP was loaded directly onto a nickel-agarose column (10x50mm)(Novogen). The loaded column was washed extensively with Buffer N (20mM Tris-HCl (pH 7.6), 0.5M NaCl, 2% glycerol, 5mM imidazole) prior to elution of rSRAFP by an imidazole gradient (5mM - 125mM) in 75min at a flow rate of 1.5mL/min. Fractions containing rSRAFP were pooled and dialyzed (using 6,000-8,000 Mw cut off dialysis membrane (Fisherbrand)) against exchange buffer (25mM Tris-HCl (pH9.0)) and subsequently chromatographed on an FPLC (Pharmacia) anion

exchange column (Q-Sepharose 16/10 (16x100mm); Pharmacia) pre-equilibrated in exchange buffer. Bound AFP was eluted by a linear 0.2-0.4 NaCl gradient in exchange buffer at a flow rate of 2.5 mL/min. Fractions containing AFP were pooled and dialyzed into 100mM NH_4HCO_3 and lyophilized.

Samples of the FPLC-pure rSRAFP were applied onto a C18 reversed-phase analytical HPLC column (3.9X250mm)(Vydac). The column was equilibrated in 8% acetonitrile and 92% H_2O containing 0.1% TFA. The AFP was eluted by a gradient of the organic phase to 80% acetonitrile, at a flow rate of 1mL/min.

Large-scale purification of rSRAFP from the 10L bioreactor medium included a preliminary step in which the medium was concentrated 10-fold using an Amicon Spiral Cartridge (Model S1Y3) 2L concentrator (Model CH2PRS). The resulting solution was saturated with EDTA (20mL of 0.5M EDTA (pH9.0), in 100mL of media). The pH was increased to pH 8.0-8.5 with the addition of 1M Tris base. Residual EDTA was then saturated by the addition of excess $MgCl_2$ (20mL of 2M $MgCl_2$ in 100mL of media) and additional 1M Tris base to maintain the pH. The resulting solution was centrifuged at 10000 rpm for 20 minutes at 4°C and any residual precipitate removed. The final solution was treated as described above.

Recombinant SRAFP Crystallization

Recombinant SRAFP was eluted off the FPLC Q-Sepharose column and dialyzed into 10mM Tris-HCl (pH9.0). This was concentrated, using Millipore Ultrafree Biomax-5K (15mL) centrifugal concentrators, to 11mg/mL as determined by OD_{280}. A small

fraction of this stock solution was further concentrated to 22mg/mL. All solutions were filtered using a 0.2μm syringe filter (millipore). These protein solutions were used in hanging drop crystallization trials. An aliquot (2μL) of protein solution was added to an equal volume of crystallization solution on a siliconized cover slip and sealed upside down in a chamber equilibrated with crystallization solution. Hampton Research Crystal Screens I and II were used in tests, as well as three grid screens outlined in the Results section.

pH Dependence Study

The antifreeze activity of a 0.8 mg/mL solution of wild-type SRAFP isolated from sea raven serum (a gracious gift from Dr. Choy Hew, Hospital for Sick Children, Toronto) (50) was measured at pH values ranging from 1 to 13. An aliquot (20μL) of a pH-adjusted buffer mixture of 93.75mM sodium borate, 93.75mM sodium phosphate and 93.75mM sodium citrate was added to an aliquot (5μL) of SRAFP (dissolved in H_2O at 4.0mg/mL) for pHs 2-11. The final concentrations were 0.8mg/mL SRAFP and 75mM buffer. Activity at pH 1 was measured in a 100mM HCl/75mM NaCl solution at 0.8 mg/mL AFP. Similarly at pH 13, activity was measured in 100mM NaOH/75mMNaCl.

DTT Dependence Study

The antifreeze activity of a 1mg/mL solution of wild-type SRAFP isolated from sea raven was measured at a variety of temperatures in the presence of 10mM DTT. The AFP was dissolved in 100mM Tris-HCl (pH 8.0), to approximately 1.0 mg/mL. To this

was added DTT (from a 1M stock solution) to 10mM, or an equivalent volume of buffer. Samples (10μL) of this solution were incubated for timed intervals at 0°C, 22°C and 37°C.

Primer-Directed Mutagenesis

Mutants were all made by primer-directed mutagenesis (Kunkel et al., 1987). Single-stranded uracil-containing template of either pT7-7F-SRm or pET20b+-SRm vectors (Figure 19), was produced using the helper phage M13KO7 and was recovered by standard procedures (123)(124) and further purified on a hydroxylapatite column (124). All mutagenic primers were designed to have DNA lengths on either side of the miss-match, with Tms of at least 40°C. Mutagenic primer (10pmol) and single-stranded DNA (ssDNA) template (1μg) were combined in 20μL of annealing buffer: 20mM Tris-HCl (pH 7.5), 50mM NaCl and 2mM $MgCl_2$. The mixture was placed in a 95°C water bath and cooled to 22°C over 2h. Following the addition of 10mM dNTPs (2.5μL each), 10mM ATP (5μL), synthesis buffer: 10mM Tris-HCl (pH 8.0), 50mM $MgCl_2$, 20mM DTT (5μL) and water (4μL). The mixture was left on ice for 5min. At the end of this incubation gene 32 protein (3μg)(Pharmacia), T4 DNA polymerase (6.4U)(NEB) and T4 DNA ligase (4.0U)(NEB) were added sequentially. The mixture was incubated on ice for a further 5min and then at 23°C for 5min and then at 37°C for 2h. An aliquot (20μL) of this incubation mixture was used to transform *E. coli* JM83. Primers used in mutagenic reactions are described in Table 2. Positive clones, detected by dideoxy sequencing (115) were digested with *BstEII* and *NotI* and subcloned in the PHIL-S1-SRm

Table 2. Antisense Mutagenic Primers

Title	Sequence (5'-3')
Loop 4 Swap	pCGTTCCTGGTGTTCTGACCAACCTAACGATCACGG
	TTCCGGTGAATGCTGTATGCAGATG (Sense Primer)
C111A	pCAAGTCATCCCAGGCTTGGTCAGCTGC
C89A	pAGGTTTGGTAGAAGCCCAGGAACGAAA
S90A	pATCATCAGGTTTGGTAGCACACCAGGAACGAAA
D113A	pGGACAAGGCAAGTCAGCCCAGCATTGGTC
T105A/Q103A	pGGTCAGCTGCAGCAGCCATAGCCATACAGCACGCG
D79A/T81A	pACGAAATTCATAGGAGCACCAGCAGACCAGGTCCA
Q55A/T56A	pAACACCAGCATTCAAAGCAGCAATGAAACTATGCTC
S47A/E50A	pCTGAATGAAACTATGAGCCTCCTGAGCGTGGATGGATGCAAG
T23N/T24N/T27A	pGCCAGAGCCCAAGCCATCGCATTATTCTCATAATAGATACAGC
S120H	pGCAGACTGATTTGTGATGTGCCGGACAAGGCAA
T23H	pCAAGTCATCGCTGTATGCTCATAATAGATACAG
T24H	pCCCAAGTCATCGCATGTGTCTCATAATAGATA
INSERTS	
119/120	pGACTGATTTGTGGGAACCACCTTGGTACCCACCTGCCGGACAAGGC
107/108	pCAGCATTGGTCAGCACCACCTTGGTACCCACCTGCAGCAGTCATCT
94/95	pCGGCCAGTACCATCACCACCTTGGTACCCACCATCAGGTTTGGTAGA
81/82	pACGAAAATTCATAGGACCACCTTGGTACCCACCTGTACCATCAGACC
INSERT LINKER	
INSERT A	pTTGCCAGATGCCTAGGCCTGCTAAACCACTGTAC
INSERT B	pAGTGGTTTAGCAGGCCTAGCATCTGGCAAGTAC

NTHT or pPIC9-SRm-CTHT vectors for expression.

Double Mutagenesis

Mutants that included two single site changes at different positions along the cDNA were mutated by the addition of two mutagenic primers to the same mutagenic reaction. Other double mutants, made up of single site changes plus the Loop 4 mutant, were produced using ssDNA of the pETt20b+-SRm-CTHT-Loop4 mutant. All mutants were subcloned and screened as described above.

Linker Insertions

The insertion-linkers primers A and B (Table 2) were resuspended to 10pmoles/μL and 10μL of each of these primers was added to annealing buffer (5μL) (500mM Tris-HCL (pH 8.0), 0.1M $MgCl_2$) plus 25μL H_2O to 50μL total volume. This solution was heated to 95°C for 3min and cooled slowly over 2.5h to 22°C. An aliquot (5μL) of this annealed primer solution was added to pPIC9-SRm-CTHT insertion mutants (either 94/95 or 107/108) vectors digested with *Kpn*I and previously dephosphorylated using alkaline phosphatase. This mixture was ligated overnight with T4 DNA ligase and buffer (NEB) at 14°C. The resulting mixture was transformed into JM83 *E. coli* and plated on LBamp plates

PART B

Type I Peptide Synthesis, Purification, Quantitation, Circular Dichroism and Sedimentation

Peptide synthesis and purification, CD studies and sedimentation experiments were performed at the University of Alberta, Edmonton, by Dr. Heman Chao. Dr. Michael Houston (University of Alberta, Edmonton) was responsible for production of lactam bridged samples as well as their conjugation to bovine serum albumin (BSA).

Antifreeze Activity Measurements and Photomicroscopy

Thermal hysteresis measurements and photomicroscopy were carried out as described in Part A. Ice crystal *c* to *a* axis ratios were determined by measurement and averaging of between 11 and 16 different ice crystals for each variant.

NMR Spectroscopy

2D-NOESY NMR experiments were performed at the University of Alberta, Edmonton, by Dr. Wolfram Gronwald and Dr. Frank Sonnichsen.

CHAPTER 3

RESULTS

PART A

Expression of Type II rSRAFP

In order to test the model of the 129-amino-acid-long SRAFP, a structure-function study was carried out. An efficient expression system was required to produce the protein for both the structural and the mutagenic studies. Initial attempts at expressing Type II AFP using both baculovirus-infected insect cells and *E. coli* produced very low yields of soluble recombinant protein; only 0.8mg rSRAFP/L of media accumulated in the baculovirus system (60) and less than 2.0mg/L was obtained in the *E. coli* system (not shown). Significant difficulties were encountered during purification of the soluble *E. coli* rSRAFP product, and the insoluble fraction could not be refolded into an active conformation, rendering the system virtually useless (unpublished data). However, by using the yeast α-mating factor signal peptide for secretion, soluble, active rSRAFP was successfully obtained from *P. pastoris* at levels up to 3.0mg/L in shake flask cultures.

Yeast Expression Vectors and Transformed Strains

Two *P. pastoris* expression vectors containing different signal peptides, were tested for secretion of rSRAFP. These included the yeast α-mating factor signal peptide

in the pPIC9 vector and the yeast alkaline phosphatase signal peptide sequence in the pHIL-S1 vector. Both vectors contained the DNA sequence coding for the 129-amino-acid mature SRAFP, situated downstream of the *P. pastoris* AOX1 (alcohol oxidase) promoter (112). As well, these constructs included a sequence encoding a histidine tract, six histidines in length. In pPIC9 the His-tag was inserted by mutagenesis at the 3'end of the cDNA immediately upstream of the stop codon; in pHIL-S1, cassette mutagenesis was used to introduce the His-tag at the 5'end of the cDNA. The latter manipulation resulted in a cloning artifact in which the N-terminal glutamine codon was replaced by a glycine codon. The N- and C-terminal insertions of the His-tag were carried out to allow comparison of the effects of the different insertions on expression levels. These vectors also contained the *P. pastoris* HIS4 gene for positive selection in the $HIS4^-$ strain, GS115(*his4*), and the ampicillin resistance gene for positive selection in *E. coli* (112). Following production of the vectors in *E. coli*, they were linearized and transformed into GS115(*his4*). Vector linearized by *BglII* and integrated by homologous recombination into the AOX1 gene site produced clones with a His^+, Mut^s phenotype. The vector linearized at the *SalI* site integrated into the HIS4 gene site of the GS115 (*his4*) genome, without disruption of the vector HIS4 gene, and gave a His^+, Mut^+ phenotype. Positive transformants were selected and tested for rSRAFP production by measurement of thermal hysteresis activity and Western blot analysis of the spent medium.

Shake Flask Expression of rSRAFP.

Recombinant SRAFP produced and secreted by GS115-pPIC9-SRm-CTHT in

shake flasks, was detected in the medium by Western analysis at a level of only 3.0 mg/L. The recombinant protein, which included the eight N-terminal additional amino acids and six C-terminal histidines, had a calculated Mw of 15777 (Figure 20). It gave rise to a single band on a Western blot that ran with approximately the same mobility as the native SRAFP mature standard (calculated Mw 13999) (Figure 21A, lanes 2 and 3). The bell-shaped band in lane two is the result of overloading. A thermal hysteresis reading of approximately 50 mOs (0.1°C) was detected in the medium from the pPIC9-SRm-CTHT culture (grown to a final OD_{600} of 10-15 in glycerol-containing media and subsequently induced with methanol for 48 hours) indicating that the rSRAFP was active. Ice crystals produced in the presence of rSRAFP had a slightly rounded hexagonal bipyramidal morphology characteristic of SRAFP (15) (Figure 22). Expression levels from the Mut^+ clone containing pPIC9-SRm-CTHT (Figure 21A, lane 5), were similar to those from the corresponding Mut^s clone (Figure 21A, lane 3). When Mut^s pPIC9-SRm-CTHT was switched to growth on a minimal medium in order to label rSRAFP with ^{15}N for NMR structural studies, the yield of AFP decreased by at least half as detected by Western blotting (Figure 21A, lane 6).

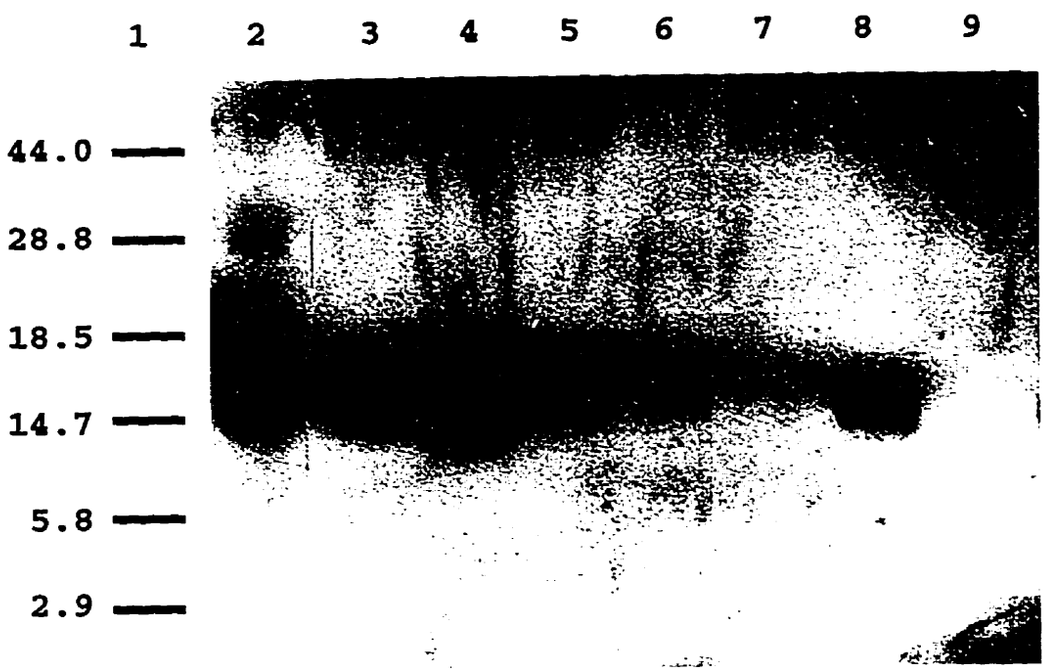
rSRAFP produced by GS115-pHIL-S1-SRm-NTHT was detected near the expected 15kDa size (calculated Mw 15020) (Figure 21A, very faint band lane 7), at a level of less than 1.0mg/L. A thermal hysteresis value of approximately 10mOs (0.02°C) was detected in the medium from the pHIL-S1-SRm-NTHT culture, confirming the lower yield. These results suggest that the α-mating factor signal sequence in the pPIC9 vector is more effective in directing the expression and secretion of rSRAFP from yeast. These

Fig. 20 **Primary amino acid sequence of rSRAFP.** The predicted sequence of the rSRAFP product secreted from *P. pastoris* by the α-mating factor signal peptide from the pPIC9-SRm-CTHT construct is shown. The eight addition amino acids are underlined (solid) at the N-terminal end. The C-terminal His-tag is marked with a dotted line. The residues are numbered according to the mature N-terminus =1.


1
10
20
EAEAYVEFQRAPPNCPAGWQPLGDRCIY
30
40
50
YETTAMTWALAETNCMKLGGHLASIHSQEE
60
70
80
HSFIQTLNAGVVWIGGSACLQAGAWTWSDG
90
100
110
TPMNFRSWCSTKPDDVLAACCMQMTAAADQ
120
CWDDLPCPASHKSVCAMTFHHHHHH

Fig. 21 **Western Blot and SDS-PAGE of rSRAFP.** A) Western blot, lane 1: low molecular weight prestained markers (kDa)(GIBCO/BRL), lane 2: approximately 150ng SRAFP from sea raven serum, lane 3: pPIC9-SRm-CTHT (Mut^s), lane 4: pPIC9-SRm (Mut^s), lane 5: pPIC9-SRm-CTHT (Mut^+), lane 6: pPIC9-SRm-CTHT (Mut^s) from defined medium (flask), lane 7:pHIL-S1-SRm-NTHT (Mut^s), lane 8: pHIL-S1-SRm (Mut^s), lane 9:GS115 media (BMMY) after growth. Equivalent volumes of media (20μl) were loaded in each lane (3-9) following growth,induction and removal of cells by centrifugation. Culture volumes were identical therefore band intensities represent relative [AFP] secreted in each case. This Western blot represents only one of a number of blots that were analyzed for estimation of rSRAFP quantities in each case. B) SDS-PAGE, lane 1: low molecular weight prestained markers kDa) (GIBCO/BRL), lane 2: purified rSRAFP (from pPIC9-SRm-CTHT (Mut^s)) after nickel-agarose column and ion-exchange purification steps.

A



B

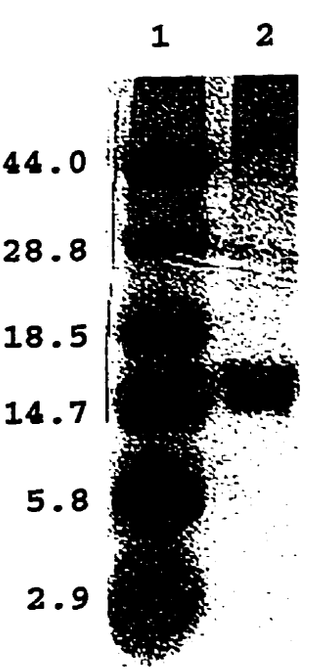


Fig. 22 **rSRAFP ice crystal morphology.** The rounded hexagonal bipyramidal ice crystal produced by purified rSRAFP (Recombinant) is compared to that produced by SRAFP (Wild-Type) purified from sea raven serum (concentrations unknown).

Wild-Type

Recombinant

results could also reflect the C- versus N-terminal orientation of the His-tag in the two different constructs. However, yields from GS115 pHIL-S1-SRm (Mut^s) and GS115 pPIC9-SRm (Mut^s), which both lacked the His-tag, reflected the difference seen with the His-tagged proteins (Figure 21A, lanes 4 and 8). This indicates that the N- or C-orientation of the His-tag has little effect on expression levels compared to the nature of the signal sequence.

Purification of rSRAFP

Recombinant SRAFP produced from pPIC9-SRm-CTHT (Mut$^+$ or Mut^s), was not significantly larger (143 amino acids) than many of the components in the complex BMMY medium. This, combined with the high salt and complex nature of the yeast medium, made most of the traditional chromatographic purification techniques ineffective. Using a nickel-agarose column to affinity purify the His-tagged rSRAFP from the medium provided a simple alternative step for the isolation of this protein, as the medium could be loaded directly through the column with no prior dilution or dialysis. The His-tagged rSRAFP was retained on the column during extensive washing to eliminate all residual medium, but was easily eluted with 60 mM imidazole (pH 9.0) (Figure 23). In the subsequent ion-exchange chromatography, SRAFP was the main protein peak in the elution profile at 0.27M NaCl, although a number of minor impurities were removed in this step (Figure 24). The analytical HPLC profile (Figure 25), as well as a Coomassie blue-stained SDS-gel, of the FPLC product, indicated that the protein was essentially pure after the ion-exchange step (Figure 21B, lane 2). Recovery from

Fig. 23 **Nickel-agarose affinity elution profile for purification of rSRAFP.** Up to 500ml of BMMY media containing rSRAFP was loaded onto a nickel-agarose column (10x50mm). The column was washed with a gradient of imidazole (0 to 125mM). 5ml fractions were collected at a flow rate of 1mL/min. Thermal hysteresis activity was measured directly from the fractions.

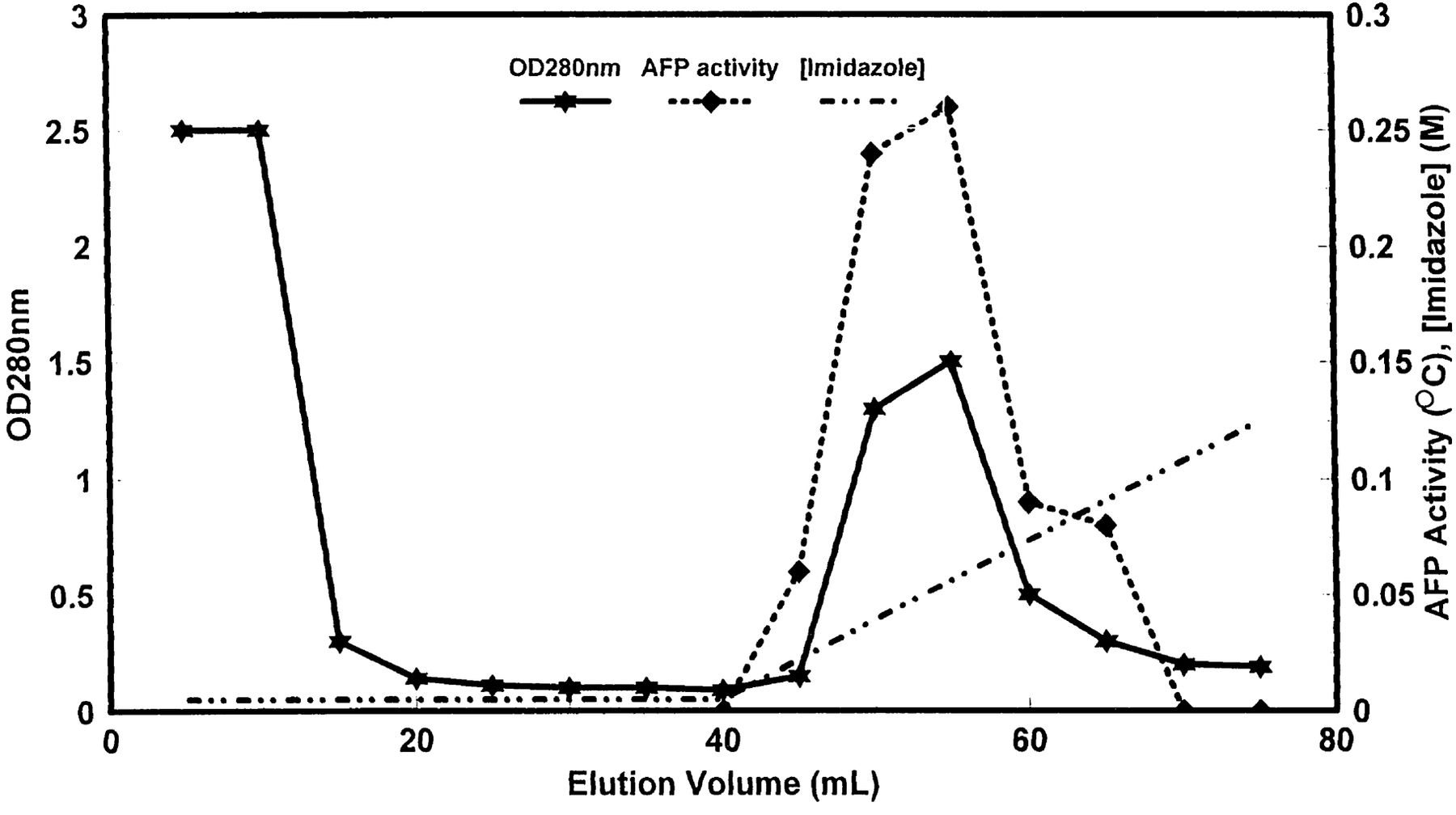

OD280nm
AFP activity
[Imidazole]
OD280nm
AFP Activity (°C), [Imidazole] (M)
Elution Volume (mL)
0
0.5
1
1.5
2
2.5
3
0
0.05
0.1
0.15
0.2
0.25
0.3
0
20
40
60
80

Fig. 24 **FPLC Q-Sepharose ion-exchange elution profile for purification of rSRAFP.** 2.5mg of rSRAFP eluted in 150ml from the nickel column, was loaded onto a Q-Sepharose ion-exchange column.The bed volumne of the column is 16x100mm. 2.5mL fractions were collected at a flow rate of 2.5 mL/min.

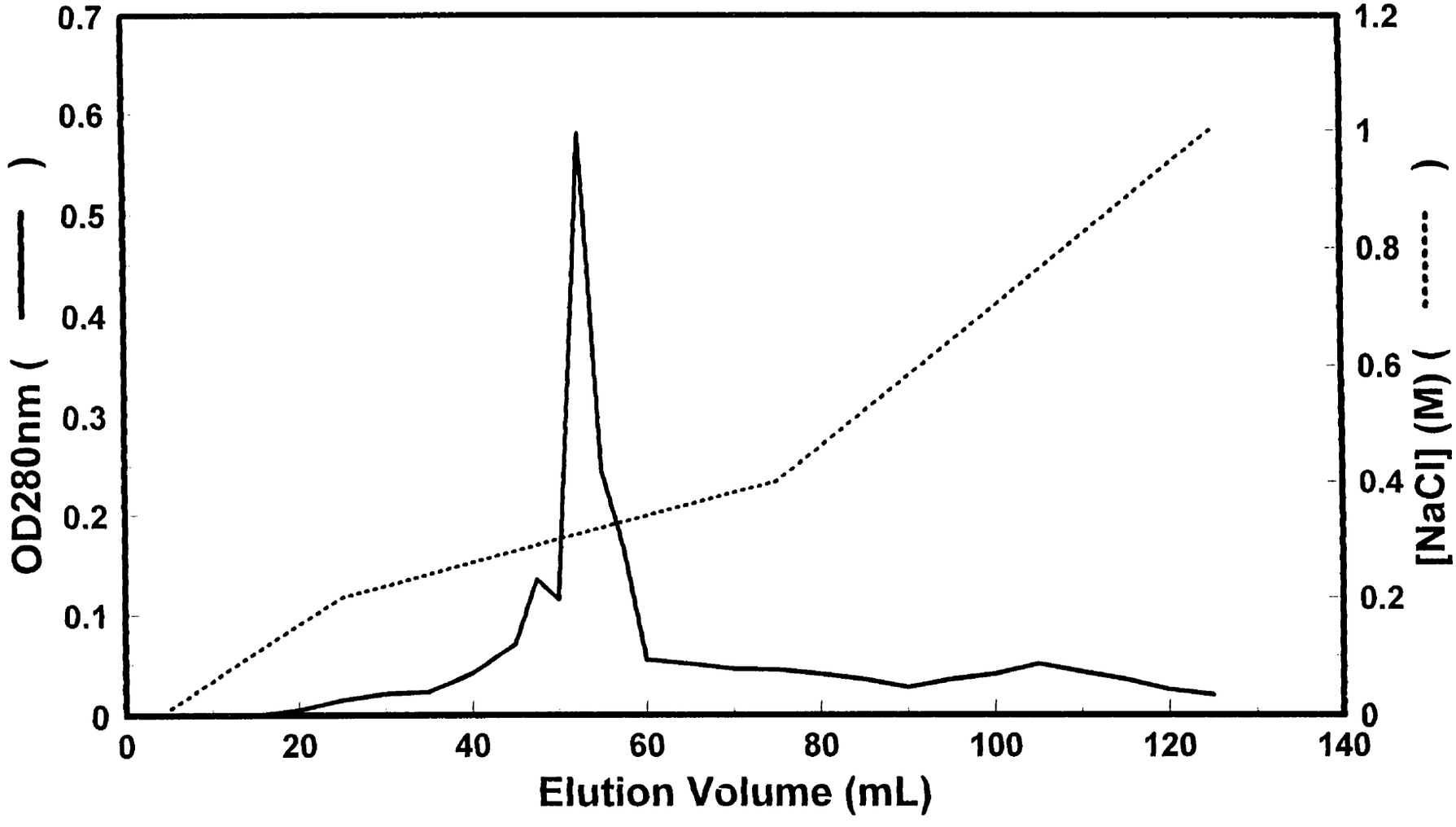
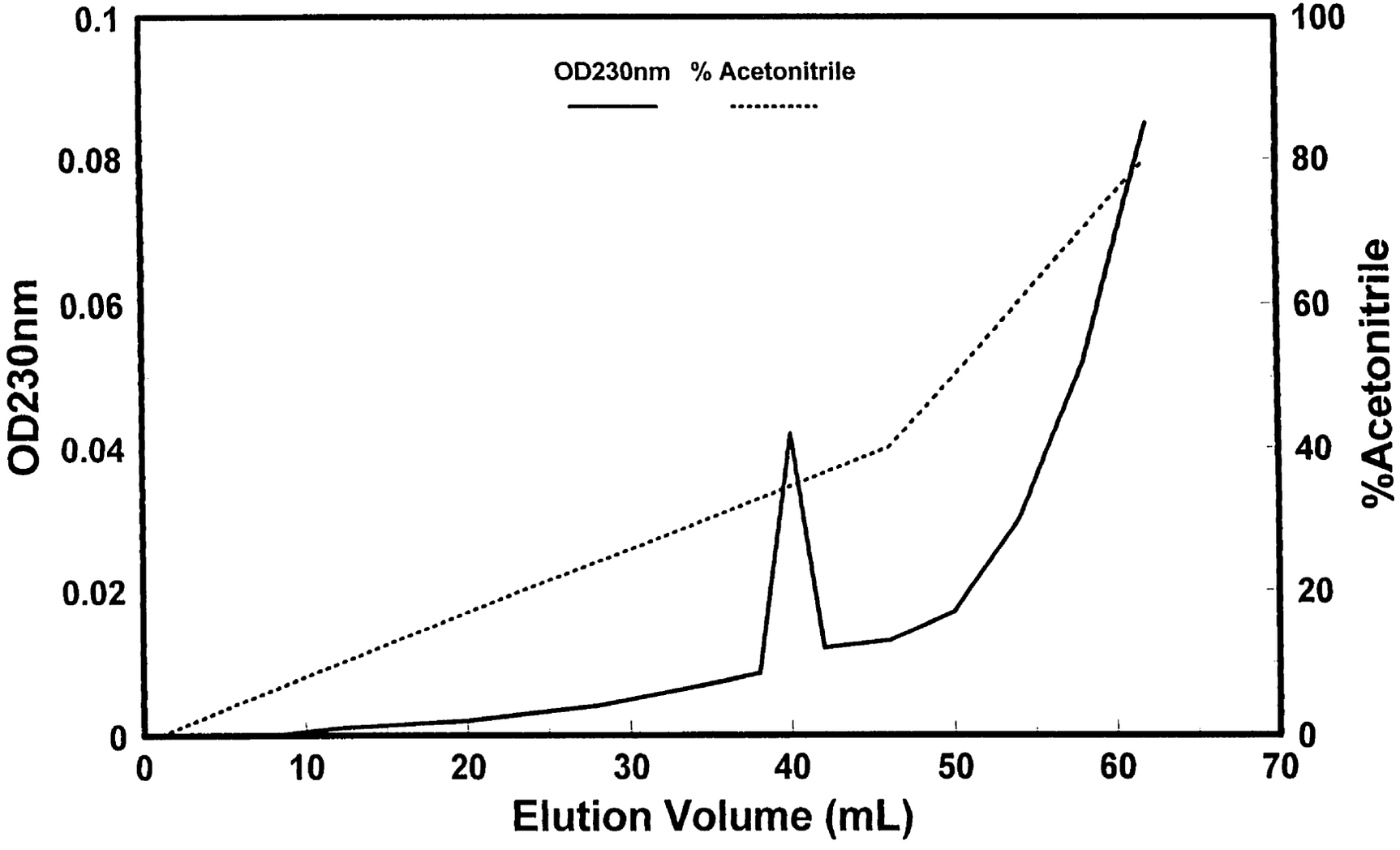

0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
OD280nm (——)
1.2
1
0.8
0.6
0.4
0.2
0
[NaCl] (M) (------)
0
20
40
60
80
100
120
140
Elution Volume (mL)

Fig. 25 **HPLC C18-reversed-phase elution profile for purification of rSRAFP.** A small sample (100μg) of rSRAFP eluted from the ion-exchange column was loaded onto an HPLC C18-reversed-phase column. A single peak was eluted at approximately 36% acetonitrile from an elution gradient of 0.8% acetonitrile/min.


OD230nm % Acetonitrile
OD230nm
%Acetonitrile
Elution Volume (mL)
0.1
0.08
0.06
0.04
0.02
0
100
80
60
40
20
0
0
10
20
30
40
50
60
70

the nickel-agarose column was approximately 2.5 mg/L of medium (83% recovery of total rSRAFP loaded) as determined by thermal hysteresis and Western blotting analyses (Table 3). Recovery of rSRAFP off the subsequent ion exchange column provided a final yield of 2 mg purified rSRAFP per litre of complex flask produced media for a final total recovery yield of 66% (Table 3).

Characterization of Purified rSRAFP

The antifreeze activity of purified rSRAFP was compared to that of the natural SRAFP, from sea raven serum, by assessing its thermal hysteresis values over a range of concentrations. The activity profile for the rSRAFP was indistinguishable from that of the SRAFP (125) (Figure 26). This indicates that neither the eight-amino-acid-long N- nor the six-amino-acid long-C-terminal addition to this rSRAFP interfered with its activity. As well, purified rSRAFP altered ice crystal morphology to the rounded hexagonal bipyramidal form characteristic of SRAFP (15) (Figure 22). The fold of rSRAFP, as indicated by 1D and 2D-NOESY NMR spectra, matched that of SRAFP purified from sea raven serum (see Appendix A for NMR data).

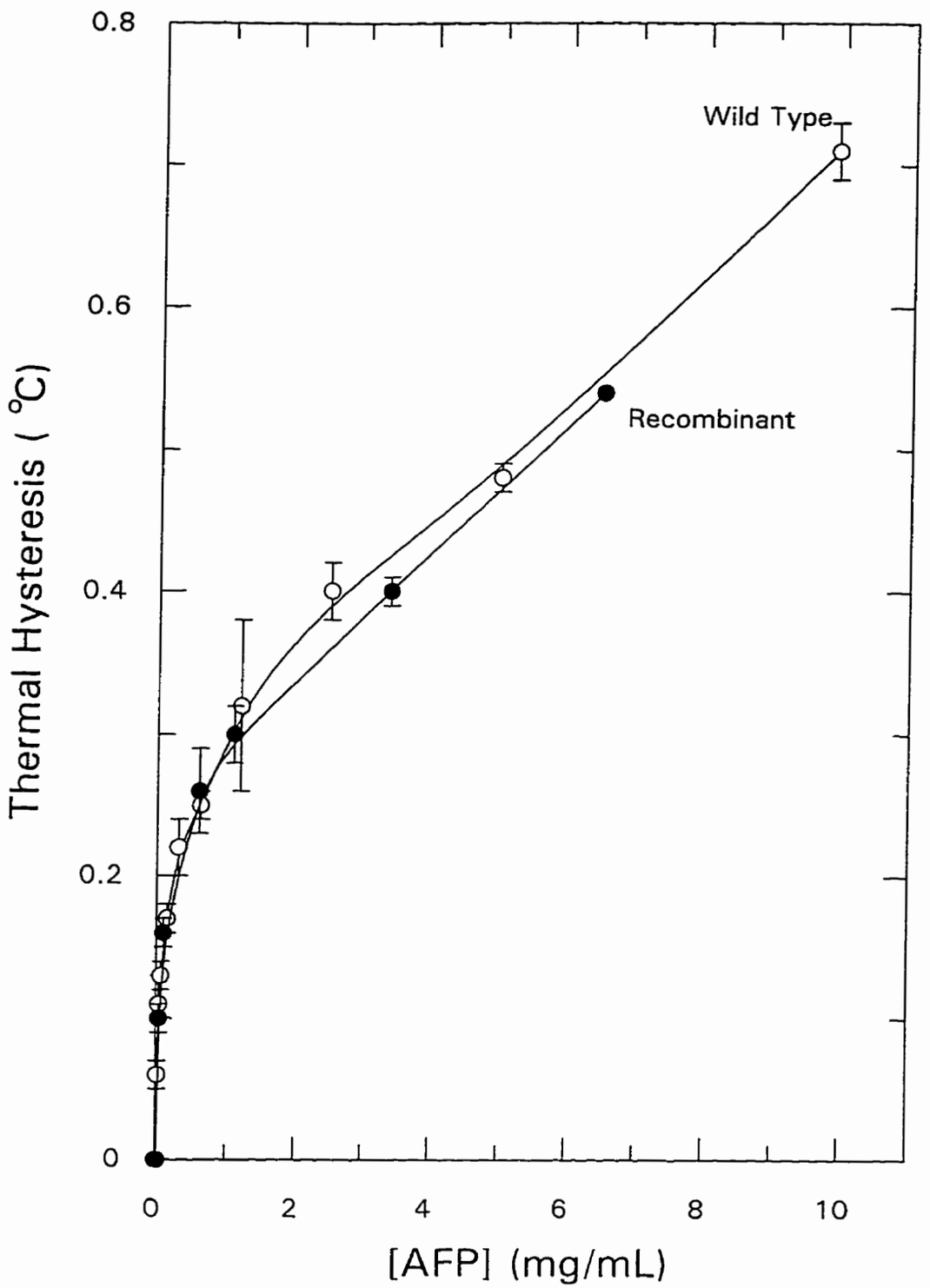
Fermentation

In order to improve the yield of rSRAFP, a shift to fermentation techniques was made to allow for better control of environmental conditions such as pH and dissolved oxygen. Four different strategies were tested at the 10L scale. The first was to improve the yield of rSRAFP in the defined fermentation medium, simply by control of pH and

Table 3. Purification Yields from 1 Litre Yeast Media

Sample	[AFP] (mg AFP)	% Recovery	Overall Recovery
Media	3.0	-	100
Nickel Column Eluate	2.5	83	83
FPLC Column Eluate	2.0	80	66

Fig. 26 **Thermal hysteresis activity of rSRAFP.** Thermal hysteresis values for SRAFP (Wild-Type) purified from sea raven serum was compared to rSRAFP values (Recombinant) purified from *P. pastoris*, at various concentrations. Each point represents the mean of at least three determinations. Standard deviations are shown as vertical bars.

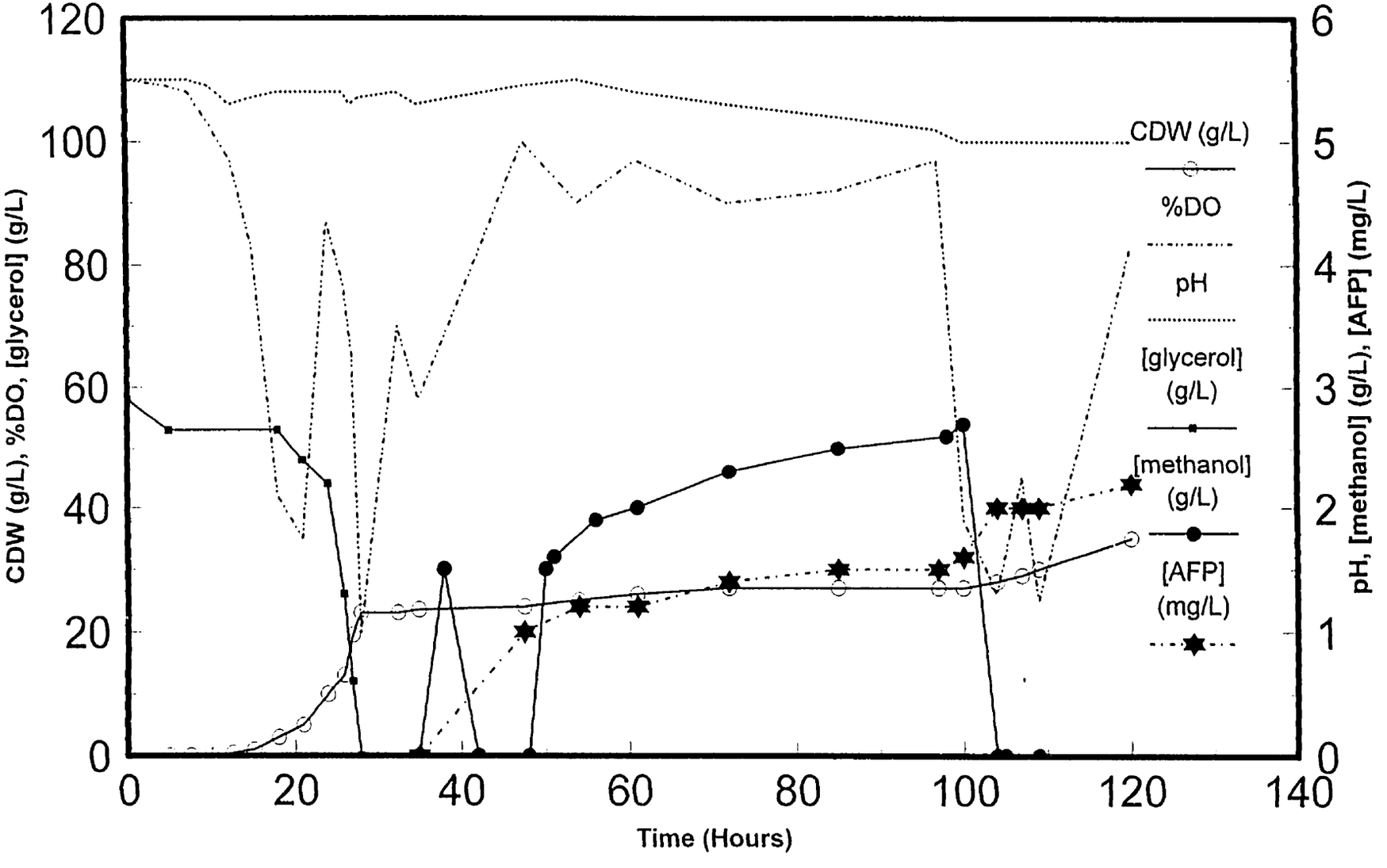

0.8
0.6
0.4
0.2
0
Thermal Hysteresis (°C)
Wild Type
Recombinant
0
2
4
6
8
10
[AFP] (mg/mL)

dissolved oxygen (DO) levels (Figure 27). The pH was maintained at 5.5 and the DO at non-limiting (> 10% saturation) values. The fermentation began with a batch growth phase, with 50g/L glycerol as a carbon source and excess (20g/L) ammonium sulfate in minimal medium. Twenty five hours after inoculation the cells had consumed all the glycerol and had reached 22 g/L cell dry weight (CDW). After depletion of the glycerol, induction with methanol was initiated. Methanol was fed at a rate of 1.0g/L per hour for a subsequent period of 95h. Due to the Mut^s phenotype, cells grow poorly on methanol, and the CDW increased only slightly up to 35g/L by the end of the fed batch phase. rSRAFP accumulated in the medium gradually over time, starting at 24h after the start of the methanol feed, to a final level of 2.0mg/L at 120h, as estimated by Western blotting (described in Methods and Materials). This was a slight improvement over shake flask minimal medium yields, but was still too low for cost effective labelling with ^{15}N and for production of protein for X-ray crystallographic studies.

Optimization of Fermentation

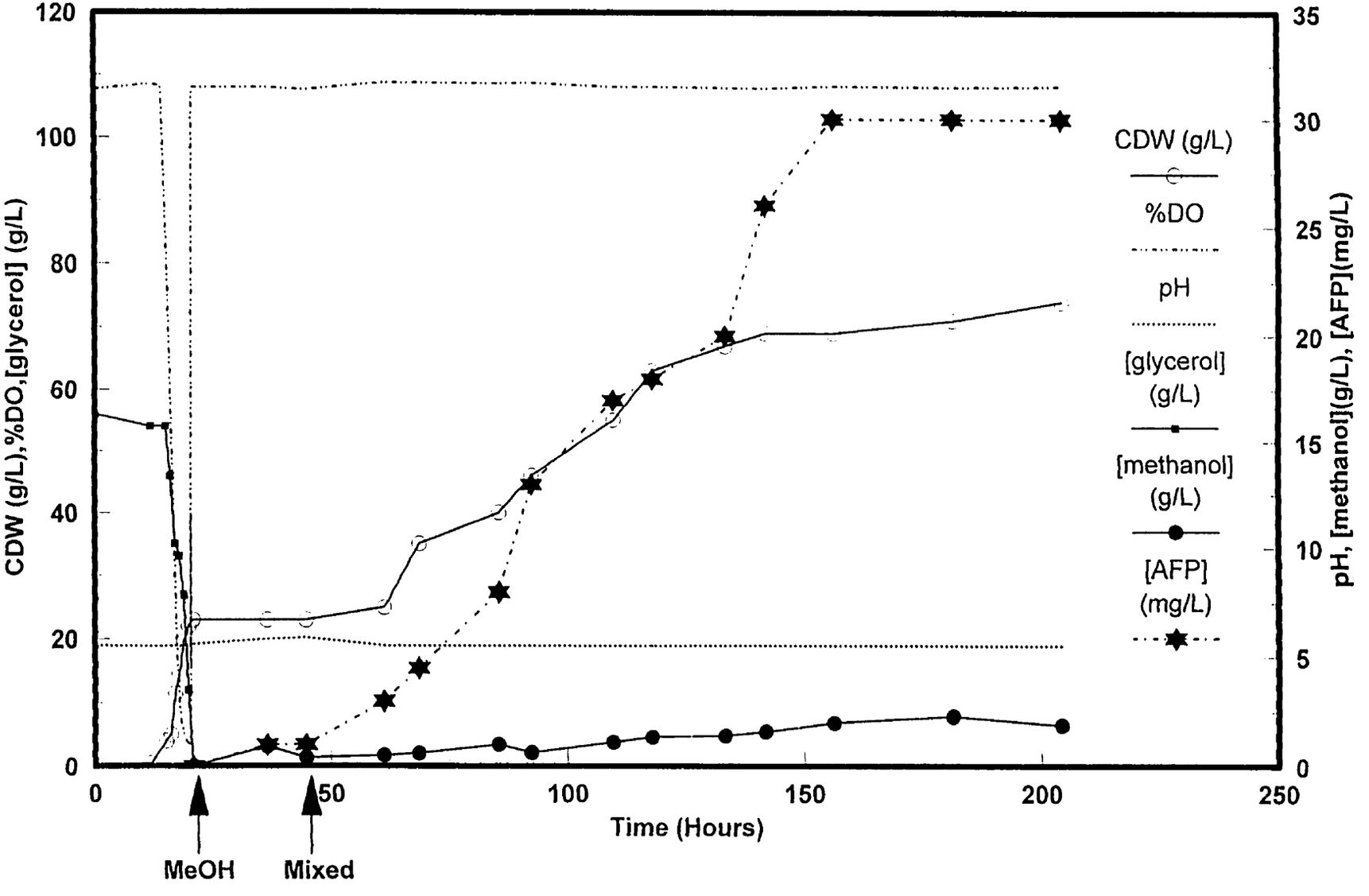
Following the marginal success of the first fermentation profile, attempts were made to increase yields of rSRAFP through optimization of the fermentation conditions. The second experiment included a mixed-feed protocol. This strategy of feeding both glycerol and methanol during the induction phase involved providing a readily usable carbon source (glycerol) for the Mut^s cells for biosynthesis, while at the same time providing an inducer (MeOH) for the AOX1 promoter. Care had to be taken in determining the proportion and amounts of each substrate to be fed, since any glycerol

Fig. 27 **Fermentation profile of Mut^S/MeOH feed strategy.** In this strategy, a batch growth using glycerol as the carbon source was followed by a MeOH-feed to 120 hours from the time of innoculation.


120
100
80
60
40
20
0
0
20
40
60
80
100
120
140
6
5
4
3
2
1
0
CDW (g/L), %DO, [glycerol] (g/L)
pH, [methanol] (g/L), [AFP] (mg/L)
Time (Hours)
CDW (g/L)
%DO
pH
[glycerol] (g/L)
[methanol] (g/L)
[AFP] (mg/L)

accumulation would repress the AOX1 promoter. As well, although methanol had to be present, levels could not reach inhibitory concentrations (>1-2%), because deletion of the AOX gene during the homologous recombination event allows only very slow metabolism of MeOH by these Mut^s cells. Overall, this mixed-feed approach was designed to allow the cells to propagate, thereby increasing the cell mass and maintaining an active culture which might encourage protein expression. The fermentation began with batch growth on 50g/L glycerol which gave 22g/L CDW (Figure 28). The pH was again maintained at 5.5 and DO at non-limiting values. At this time a methanol feed of 1.0g/L per h was started for 24h, to ensure the alcohol oxidase promoter was induced. After this methanol feeding period, the feed was switched to the mixed substrate of glycerol/methanol (5:2 volume basis). The feeding rate was started as 1.0g/L per h methanol and 4.0g/L per h glycerol. The rate of feeding was gradually increased to the maximum value permitted without residual glycerol accumulation, and without methanol rising above the 0.5% used in the shake flask cultures. The latter precaution was to avoid methanol toxicity, as noted above. This maximum rate was approximately 2.6g/L per h for methanol and 10.4g/L per h for glycerol. Fermentation proceeded for 156h from the time of inoculation. At this point the CDW was approximately 70g/L and the rSRAFP, which accumulated gradually over the induction period starting within 20h of the methanol feed, reached 30mg/L by 156h. This represents a 15-fold increase of rSRAFP over the original fermentation. By the end of 156h, a total of 185g/L glycerol had been consumed. Based on a yield coefficient of 0.42 (estimated from the batch growth phase) this would indicate an expected CDW of 76.8g/L very close to our

Fig. 28 **Fermentation profile of the Mut^S/mixed-feed strategy.** In this strategy, a batch growth using glycerol as the carbon source, was followed by a mixed-feed of MeOH and glycerol in a ratio of 5:2 (volume basis) to 210 hours.


120
100
80
60
40
20
0
CDW (g/L),%DO,[glycerol] (g/L)
35
30
25
20
15
10
5
0
pH, [methanol](g/L), [AFP](mg/L)
0
50
100
150
200
250
Time (Hours)
MeOH
Mixed
CDW (g/L)
%DO
pH
[glycerol] (g/L)
[methanol] (g/L)
[AFP] (mg/L)

observed value of approximately 70g/L.

This mixed-feed protocol was repeated to include the addition of peptone as a complex amino acid source (Figure 29). This technique was previously observed to dramatically increase yields of an extracellular protein from *Bacillus brevis* 47 (126). As well, the expression of rSRAFP from *P. pastoris* in complex medium in flasks was shown to produce much higher yields than in minimal medium. The fermentation was supplemented with peptone (Gibco BRL) at 48h (5g/L), 70 h (5g/L), 90 h (10g/L), and a final addition at 120 h (5g/L), for a total of 25g/L peptone. The pH was again maintained at 5.5 with DO at non-limiting values. At the end of 160h the cell mass had reached 74 g/L CDW and rSRAFP had again accumulated uniformly over the induction period, to 35mg/L, only marginally better than in the previous fermentation.

The final optimization strategy shifted back to the original methanol feeding profile, using the Mut^+ clone, which is capable of metabolizing methanol as a carbon source (Figure 30). This final strategy eliminated the possibility of any inhibition of the AOX1 promoter by residual glycerol in the mixed-feed profiles, even though no glycerol accumulation had been detected by HPLC. The initial conditions were identical to those outlined for the original Mut^s fermentation: a batch growth to 22g/L CDW on glycerol, followed by a methanol feed starting at a rate of 1.0g/L per h. This rate was increased to 2.6g/L per h without the accumulation of methanol above 0.5%. Higher rates of methanol feeding (up to 3.3g/L per h) led to accumulation of 1.6% methanol in the medium (as shown by the peak at about 100h (Figure 30)) which temporarily depressed cell growth and rSRAFP production. Upon decreasing the feeding rate, the cells

Fig. 29 **Fermentation profile of the Mut^S/mixed-feed/peptone strategy.** In this strategy, the mixed-feeding protocal (Figure 28) was repeated with the addition of peptone, as a complex amino acid source, throughout the mixed-feed phase.

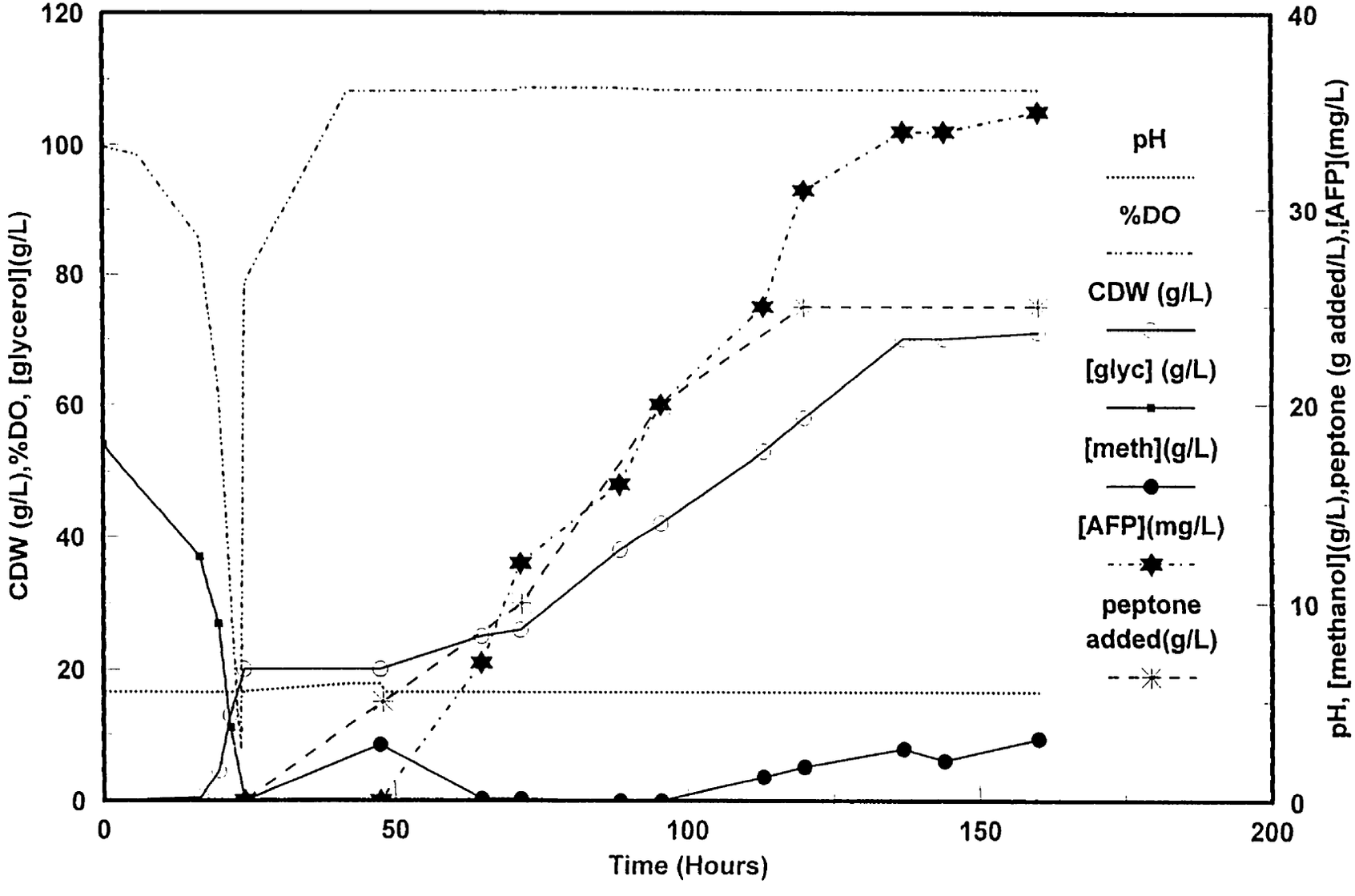
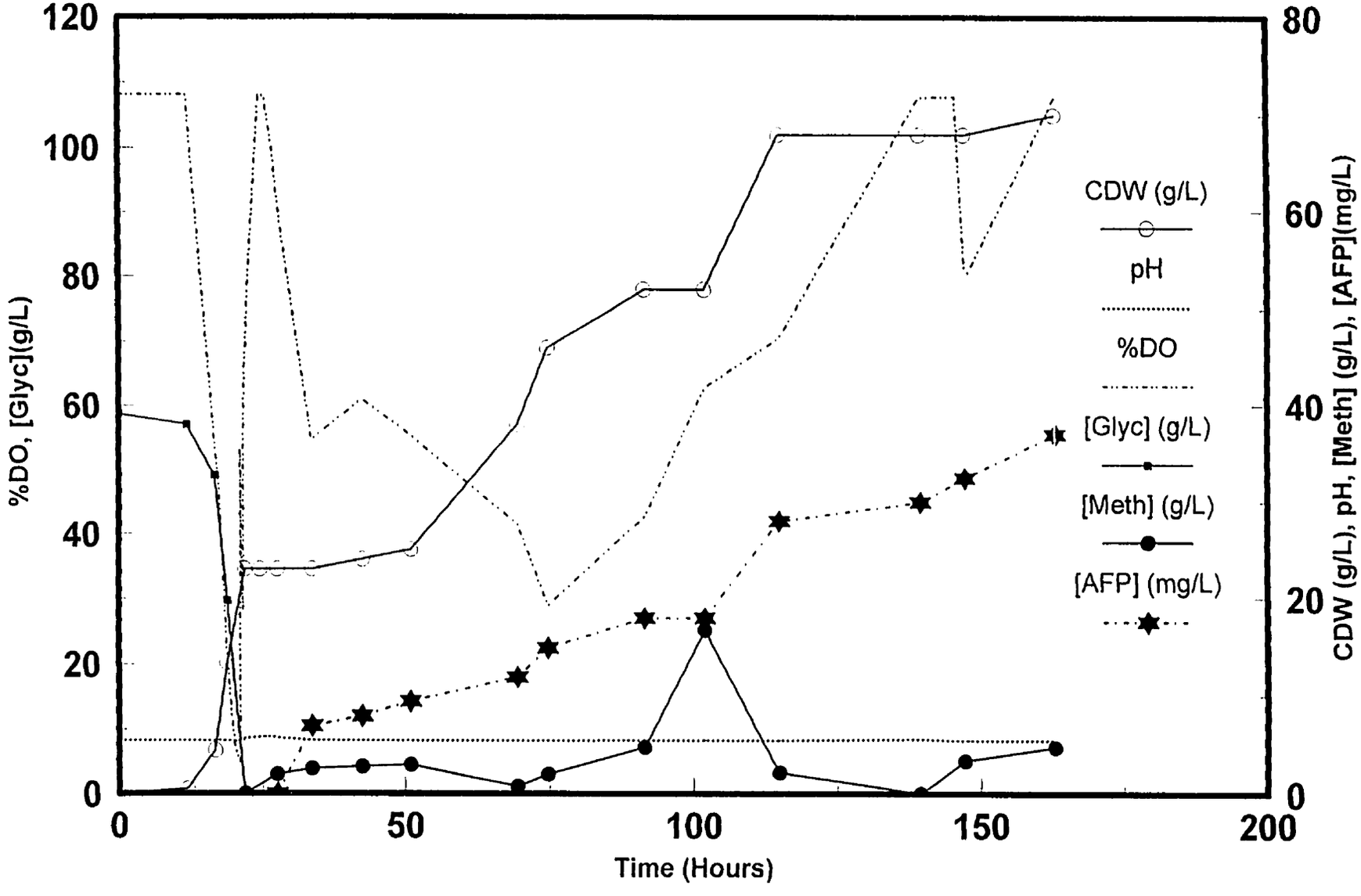

120
100
80
60
40
20
0
0
50
100
150
200
Time (Hours)
CDW (g/L),%DO, [glycerol](g/L)
40
30
20
10
0
pH, [methanol](g/L),peptone (g added/L),[AFP](mg/L)
pH
%DO
CDW (g/L)
[glyc] (g/L)
[meth](g/L)
[AFP](mg/L)
peptone added(g/L)

Fig. 30 **Fermentation profile of the Mut$^+$/MeOH-feed strategy.** In this strategy, a Mut$^+$ GS115-pPIC9-SRm-CTHT clone was used in the original MeOH-feed strategy (Figure 27). The Mut$^+$ phenotype grows well using MeOH as a carbon source.


120
100
80
60
40
20
0
%DO, [Glyc](g/L)
0
50
100
150
200
Time (Hours)
80
60
40
20
0
CDW (g/L), pH, [Meth] (g/L), [AFP](mg/L)
CDW (g/L)
pH
%DO
[Glyc] (g/L)
[Meth] (g/L)
[AFP] (mg/L)

resumed growth with no apparent ill effects. Again pH and DO were maintained as described previously. At 160h the cells reached 71g/L CDW and the rSRAFP had gradually accumulated to 37mg/L. Therefore, no significant increase in AFP yield was obtained using Mut^+ over Mut^s. By the end of the fermentation 58.7g/L of glycerol were consumed to theoretically produce 24.6g/L CDW at a yield coefficient of 0.42. 309.6g/L methanol were consumed to theoretically produce 46.3 g/L CDW (yield coefficient 0.15 for methanol (127)). This total theoretical cell mass of 70.9g/L CDW closely matches the observed value of 71g/L. The use of oxygen was much greater in the Mut^+ profile, and was reflected in lowered DO levels. This is consistent with the more reduced state of methanol compared to glycerol.

Purification of rSRAFP from Bioreactor Preparations

The fermentation culture volume increased from 6L to over 8L due to the continuous feeding over the 150h growth and induction phases. When this culture was centrifuged, up to 1.5L of cell pellet was collected, reducing the final media volume to 6+ litres. To make this large volume manageable for biochemical procedures it was concentrated approximately 10-fold to give a final rSRAFP concentration of approximately 300mg/L and a final volume of 600-800mL. The pH of this solution was raised from 5.5 to 8.5 to facilitate nickel column affinity chromatography, which functions optimally at pHs between 8 and 9. Unfortunately raising the pH by simple addition of base or even concentrated buffer led to significant precipitation, which caused a substantial loss of rSRAFP. The precipitate was very fine rather than flocculant and

occurred between pHs 7 and 8. It stuck to the sides of glassware and was only removed by the use of a calcium-lime remover (CLR Jamie Industries). This suggested the precipitate was possibly Ca^{2+} based. Another possibilty was that other metal ions (eg. Fe^{3+}) formed hydroxides at the alkaline pH and trapped some of the rSRAFP in the precipitate. To counter these possibilties excess EDTA was added to chelate all divalent ions prior to the pH change. Subsequent addition of Tris base to pH 8.5 resulted in no precipitation, indicating ions were the primary source of the problem. Since excess EDTA might strip the Ni^{2+} from the nickel-agarose column it was saturated by the addition of excess $MgCl_2$ prior to chromatography. Any residual precipitate was removed by centrifugation before loading onto the nickel-agarose column. The rSRAFP that eluted from the nickel-agarose column was loaded onto an FPLC ion exchange Q-Sepharose column. Elution with a salt gradient gave rise to a single peak at approximately 0.25M NaCl. An aliquot (200μg rSRAFP) of this was loaded onto an HPLC C18-reversed-phase column and eluted with an exceedingly shallow gradient (increase 0.1 %acetonitrile/min), from which a number of closely spaced, overlapping peaks were resolved (Figure 31). MALDI mass spectrometric analysis of the fractions comprising these peaks showed that peak 1 contained only low molecular weight material (614 and 799). Peaks 2,3,4 and 5 contained at least three species of Mw 14,785, 14,753 and 14,817 (Table 4). Peak 1 overlaps significantly with peak 2 such that it is impossible to separate it. As well significant losses have been noted because recoveries of rSRAFP off the reversed-phase column are only 50% or less. For these reasons, the HPLC purification step was omitted. Of the four large-scale fermentations described

Fig. 31 **HPLC C18-reversed-phase elution profile of purified rSRAFP.** rSRAFP from the Mut^+ fermentation profile previously, eluted from nickel and Q-Sepharose columns was loaded onto a C18-reversed-phase column. The acetonitrile elution gradient was run 8 times slower (0.1% acetonitrile/min) than in the original profile (Figure 25).

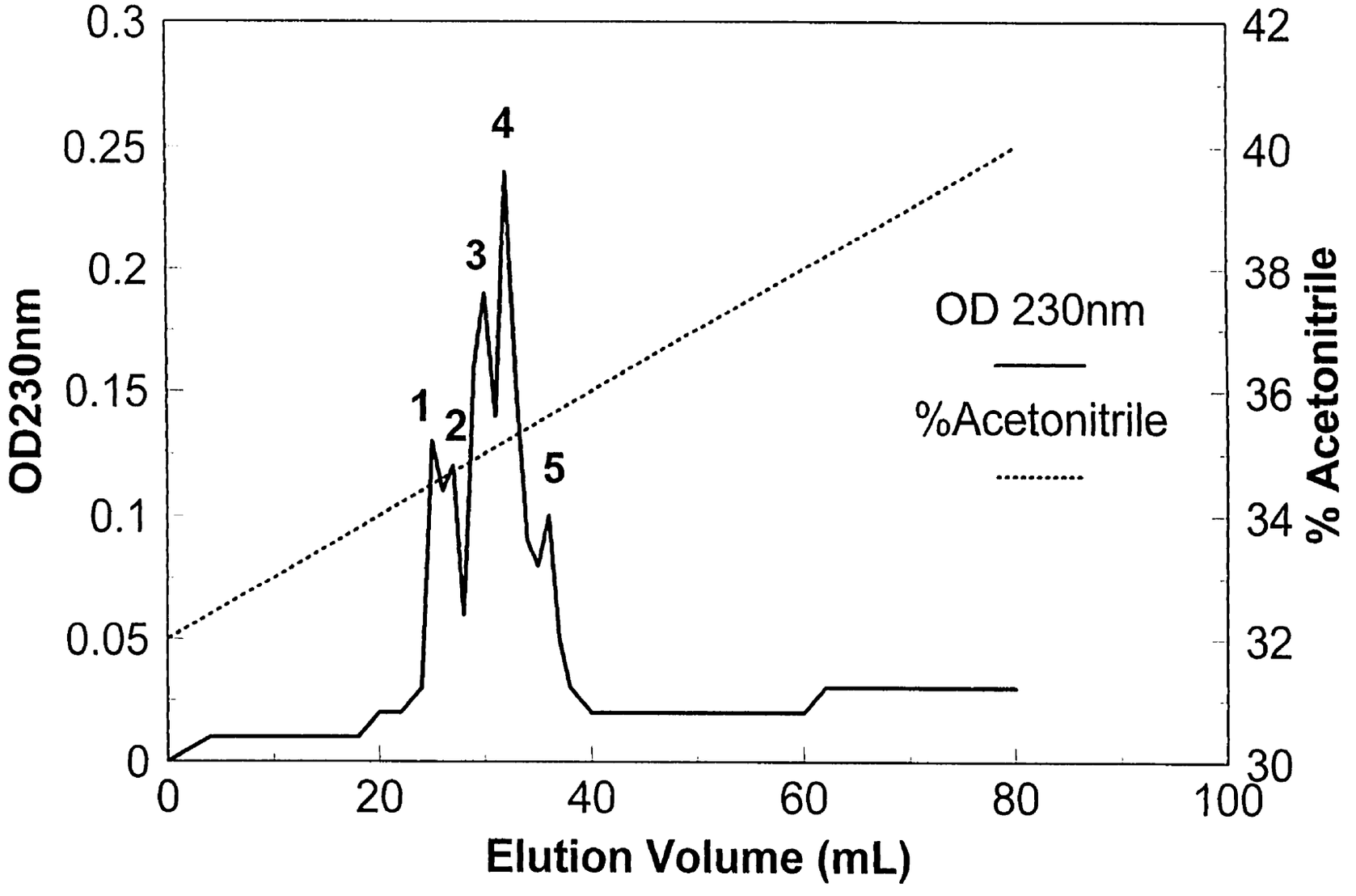

0.3
0.25
0.2
0.15
0.1
0.05
0
OD230nm
42
40
38
36
34
32
30
% Acetonitrile
0
20
40
60
80
100
Elution Volume (mL)
1
2
3
4
5
OD 230nm
%Acetonitrile

Table 4. Molecular Weights of HPLC Reverse Phase Peaks

Fraction Number	Peak Number	Molecular Weight
34	1	613.7
35	1	798.9
36	2	14,785.0
40	3	14,818.3
41	3	14,753.0
43	4	14,753.0
46	5	14,753.0
47	5	14,816.4

above, the Mut$^+$ preparation was used entirely for purification of rSRAFP for crystallization trials and other studies. The yield after ion-exchange chromatography was 106mg rSRAFP. ^{15}N-labelled rSRAFP was also produced in a 1.5L bioreactor and was prepared as described above except for the initial concentration step which was not required.

X-ray Crystallographic and NMR Structural Studies

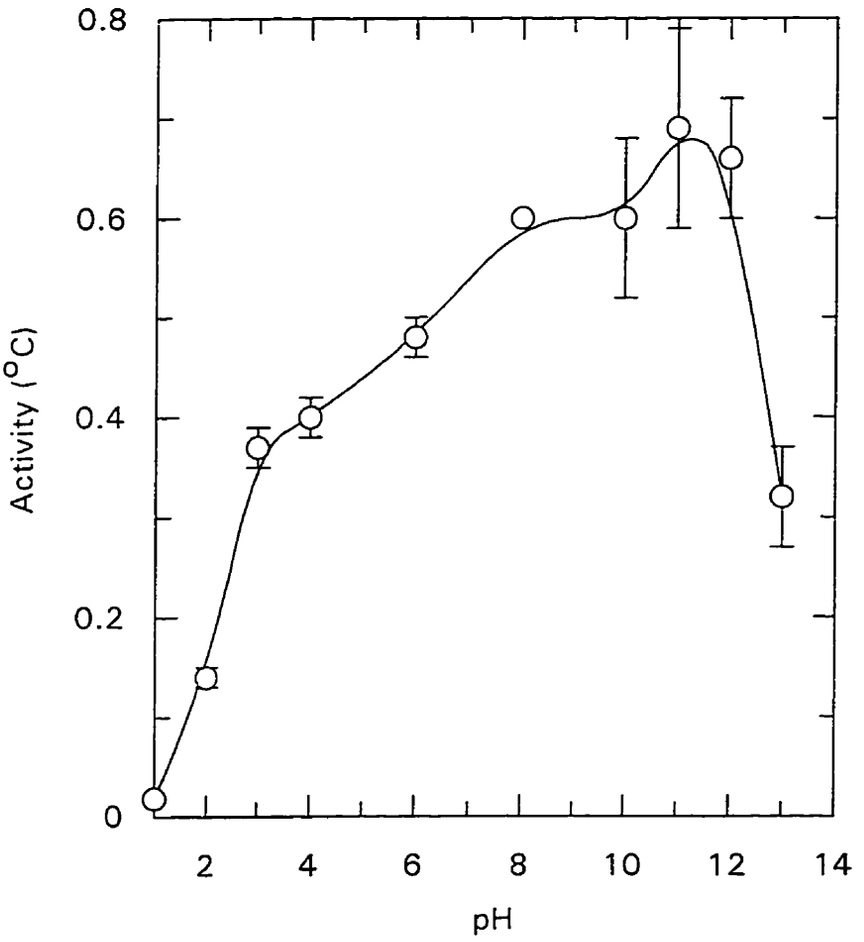
rSRAFP purified from the Mut$^+$ fermentation was used in extensive crystallographic trials. A stock solution of rSRAFP was made by dialysis of the ion-exchange purified product into 10mM Tris-HCl (pH 9.0). This was concentrated to 11mg/mL. A small aliquot was further concentrated to 22mg/mL as determined by OD_{280nm}. Crystal Screens I and II (Hampton Research) were tested at room temperature and at 4°C. As well, three grid screens including: [ammonium sulfate] (0.8 to 3.2M) versus pH (4-9), [polyethylene glycol 6000] (5-30%) versus pH (4-9) and [2-methyl-2,4-pentanediol] (10-65%) versus pH (4-9), were tested. All trials were set up using the "hanging drop" vapour diffusion crystallization technique with the 11mg/mL stock rSRAFP solution. No crystals were produced for at least three months after the trials were set up. Precipitation was variable, with no significant trends evident. Very little heavy precipitation was observed and therfore the effect of increasing AFP concentration was tested. The use of a 22mg/mL stock rSRAFP solution produced approximately twice the amount of precipitation. Future efforts will lie in using higher concentrations of rSRAFP and precipitants, lower or higher pH, and different temperatures.

Almost 12mg of ^{15}N-labelled rSRAFP has been produced from two 1.5L bioreactor preparations. The cost of production was over 470$/mg of labelled rSRAFP for label alone. NMR analysis of the protein from the first batch indicated it was virtually identical in fold to the natural SRAFP (see Appendix A for NMR data) and allowed up to 70-80% assignment of residues by 2D and NOESY NMR. Comparison of the second batch to the first indicated that it also maintained an identical fold, and could safely be pooled with the first batch to double the working concentration. This significantly improved the resolution of the spectral data. A complete 3D structure is expected to be available in the near future (personal communications with W. Gronwald).

Effect of pH on SRAFP activity

Characterization of recombinant Type III AFP indicated very little pH dependence to its activity (74). This AFP was fully active from pH 2 to pH 12, suggesting that charged residues were not important for the activity and stability of the protein. In order to better assess whether this is a general phenomenon, the pH activity profile (Figure 32) of SRAFP was determined. The profile showed a gradual increase in activity between pHs 3.0 and 11.0, with 100% activity at pH 11.0 and 50% at pH 3.0. Activity was 0% at pH 1.0 and began dropping at pH 12.0, declining to 43% of the maximum at pH 13.0. Deviations in measurements were larger at the alkaline pH values. Ice crystal morphology showed no change from the stubby rounded hexagonal bipyramid at pHs 8.0 through 13.0. At lower pHs (pH 1.0-2.0) the crystal lost its rounded shaped and was significantly faceted and more elongated.

Fig. 32 **The effect of pH on SRAFP activity.** Thermal hysteresis activity of (°C) SRAFP purified from sea raven serum (0.8 mg/mL) was measured at pH values ranging from 1 to 13 at intervals of 1 or 2 pH units. Each data point represents the mean of at least three determinations. Standard deviations are shown as vertical bars.

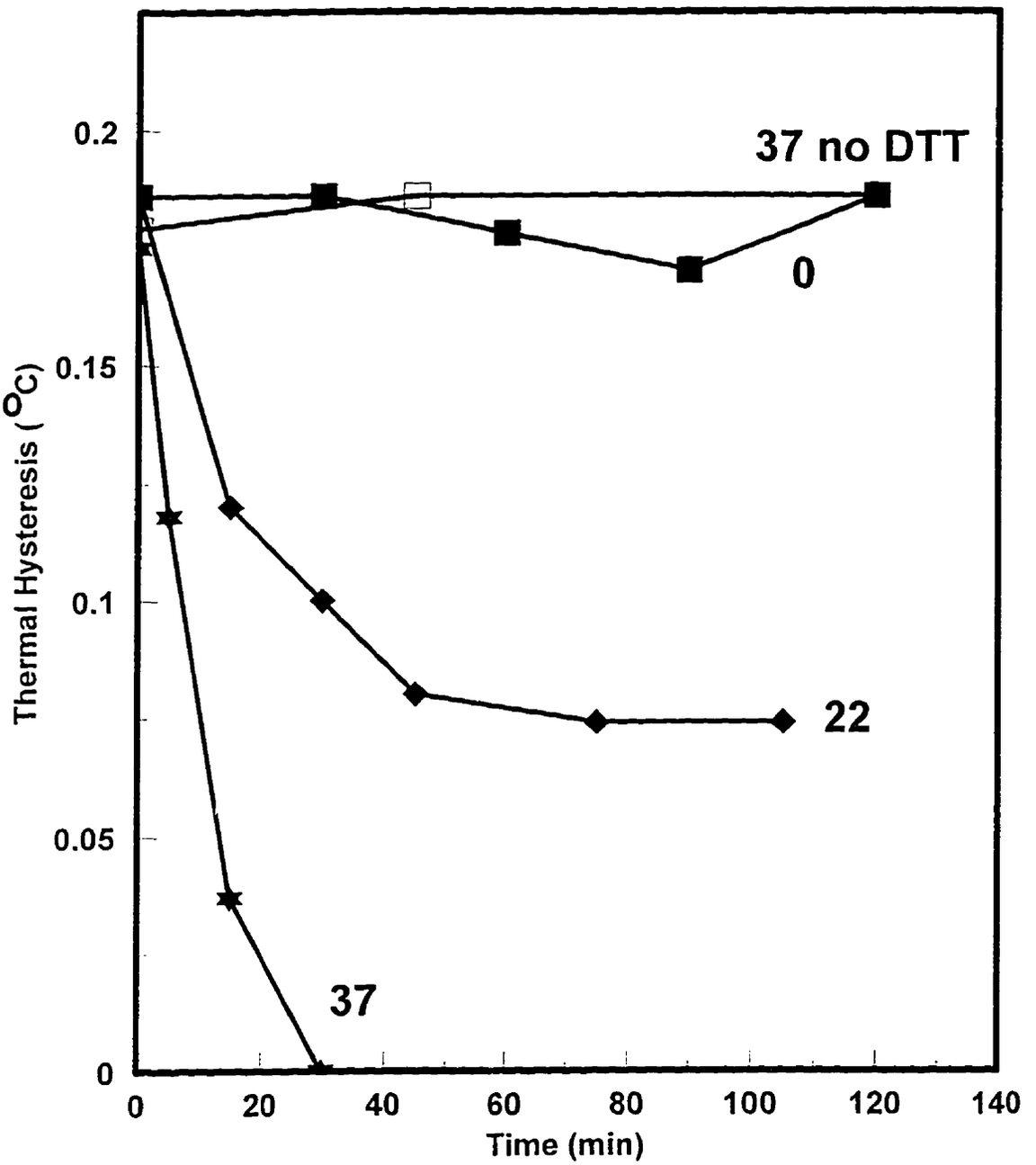

0.8
0.6
0.4
0.2
0
Activity (°C)
2
4
6
8
10
12
14
pH

Studies on the Role of Disulfide Bridges in SRAFP

Two different maps have been proposed for the disulfide bonding patterns of the five bridges in SRAFP; one based on a protein chemical analysis by Ng and Hew (66), and the second based on modelling from the CRD structures (53). To check this discrepancy, and better assess the role of disulfides in structure and function, plans were made to delete the bridges one at a time by site-directed mutagenesis. Neither the lectin nor the PSP conserve C69-C100 and C89-C111 suggesting their removal from SRAFP might not completely disrupt the stability of the fold. Initially C89 and C111 were mutated to Ala. In the case where the chemical map was correct, the double mutation would disrupt two disulfide bridges, C117-C89 and C111 to either C100 or C101, and potentially cause more significant destabilization of the protein. This is in contrast to the modelled map where only one bridge would be deleted (C89-C111). Analysis of the overall stability of the native SRAFP compared to these and other disulfide mutants would provide insight into the correct disulfide pattern. Unfortunately, no disulfide mutant protein was ever recovered using the *P. pastoris* host.

In a second study of the disulfide bridges, the reducing agent dithiothreitol (DTT) was used to assess the contribution of the disulfides to protein stability. Previous studies showed that DTT only partially inactivated Type II AFP at 37°C (50). We expanded on this observation to study the time course of inactivation (Figure 33). A 1mg/mL solution of SRAFP lost 100% activity in under 30 minutes when treated with 10mM DTT at 37°C. No loss of activity was observed at this temperature in the absence of DTT (Figure 33, open box). This loss of activity was temperature dependent. At 22°C, only

Fig. 33 **The effect of DTT on SRAFP activity.** The activity of SRAFP, purified from sea raven serum (1.0mg/mL), was assessed over time, in the presence of 10mM DTT following incubation at 0°C, 22°C 37°C and in the absence of DTT at 37°C. Thermal hysteresis readings were carried out near freezing temperatures as described in the methods and materials. Each data point represents only one determination.


37 no DTT
0
22
37
Thermal Hysteresis (°C)
Time (min)
0.2
0.15
0.1
0.05
0
0
20
40
60
80
100
120
140

60% of the activity was lost after 110 minutes and no activity was lost at 0°C (Figure 33). These results indicate a significant structural role for the five disulfide bridges and the possible presence of a very resistant core that is not reduced unless heat is added.

Ice Binding Mutations

At 129 amino acids, SRAFP is the largest of the four fish AFPs. The model of SRAFP proposes a putative ice binding site (IBS1) on the same surface as the Ca^{2+} binding site of the MBP (Figure 34). The surface is comprised of loops 3 and 4 and β-strands 3 and 4. It is hydrophilic compared to the rest of the protein surface (which is overall 65% hydrophobic) and contains a large number of hydrogen-bonding amino acids (53).

i) Specific Mutations at Putative IBS1

To test this putative IBS1, two double site mutants were made, T105A/Q103A and S90A/D113A (Figure 34A). These residues were chosen based on their surface accessibility, hydrogen-bonding potential and their proximity to the putative ice binding surface outlined in the model. Alanine was chosen as a non radical substitute that would eliminate hydrogen-bonding potential. Alanine replacement had previously been used to identify ice-binding residues in Type III AFP (T18A and T15A) (73)(111). Surprisingly, thermal hysteresis values in the double mutants were virtually identical to those of rSRAFP (Figure 35). Only T105A/Q103A showed a slightly altered crystal morphology,

Fig. 34 **Sites of mutation on SRAFP.** A) Residues of the putative IBS1 surface are shown in one view of the SRAFP model. B) A second view identifies residues subsequently tested in the search for a second ice binding surface relative to the original surface shown in A. In both views red=oxygen, blue=nitrogen, black=carbon. Structure file courtesy of F. Sönnichsen. Diagram was produced as descibed in Figure 7.

Fig. 35 **Thermal hysteresis activity of IBS1 double site mutations.** Thermal hysteresis values for rSRAFP and mutants at various concentrations were compared. Each data point represents the mean of three determinations. Standard deviations are shown as vertical bars. rSRAFP (○), Q103/T105 (●), S90A/D113A (□).

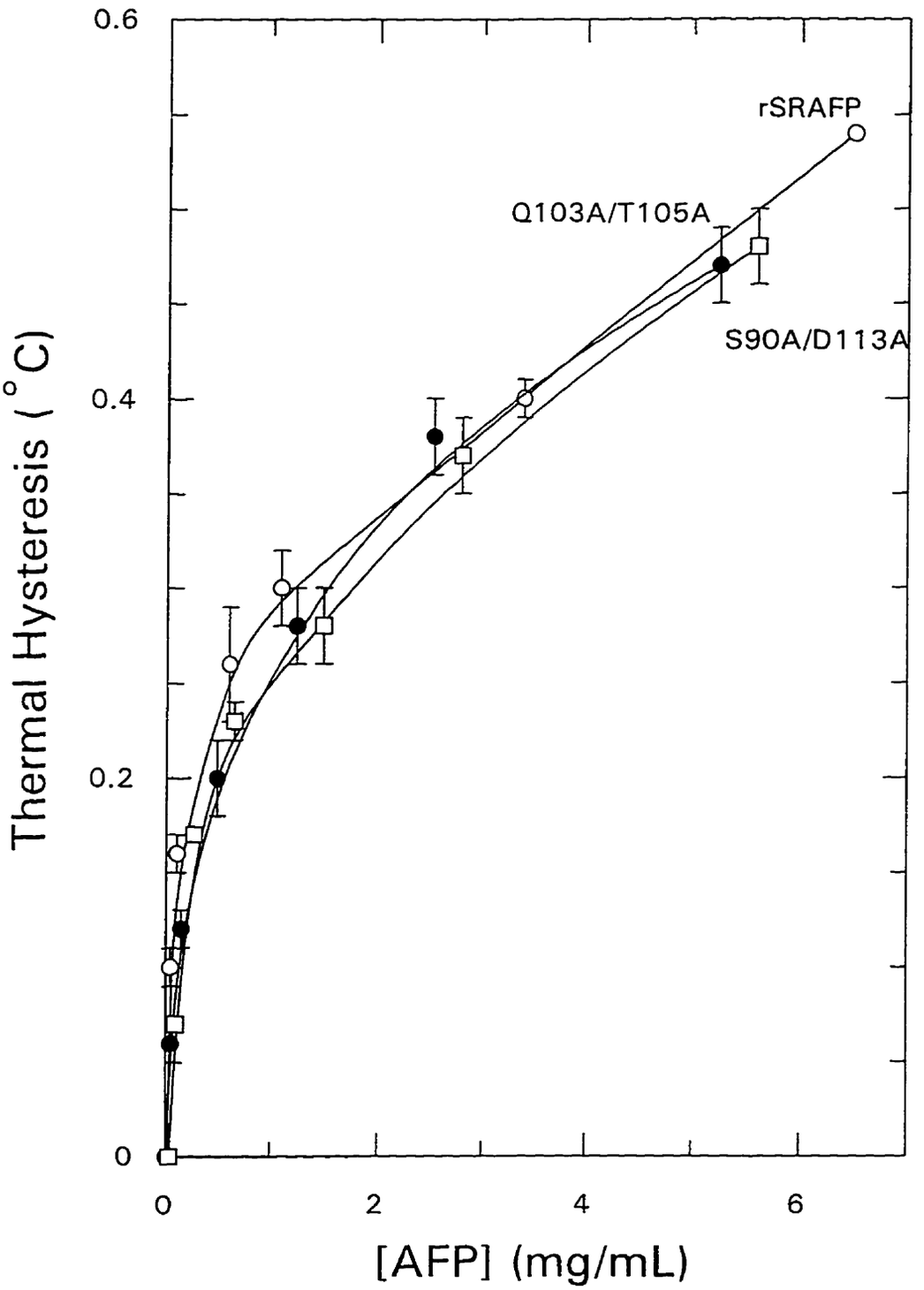
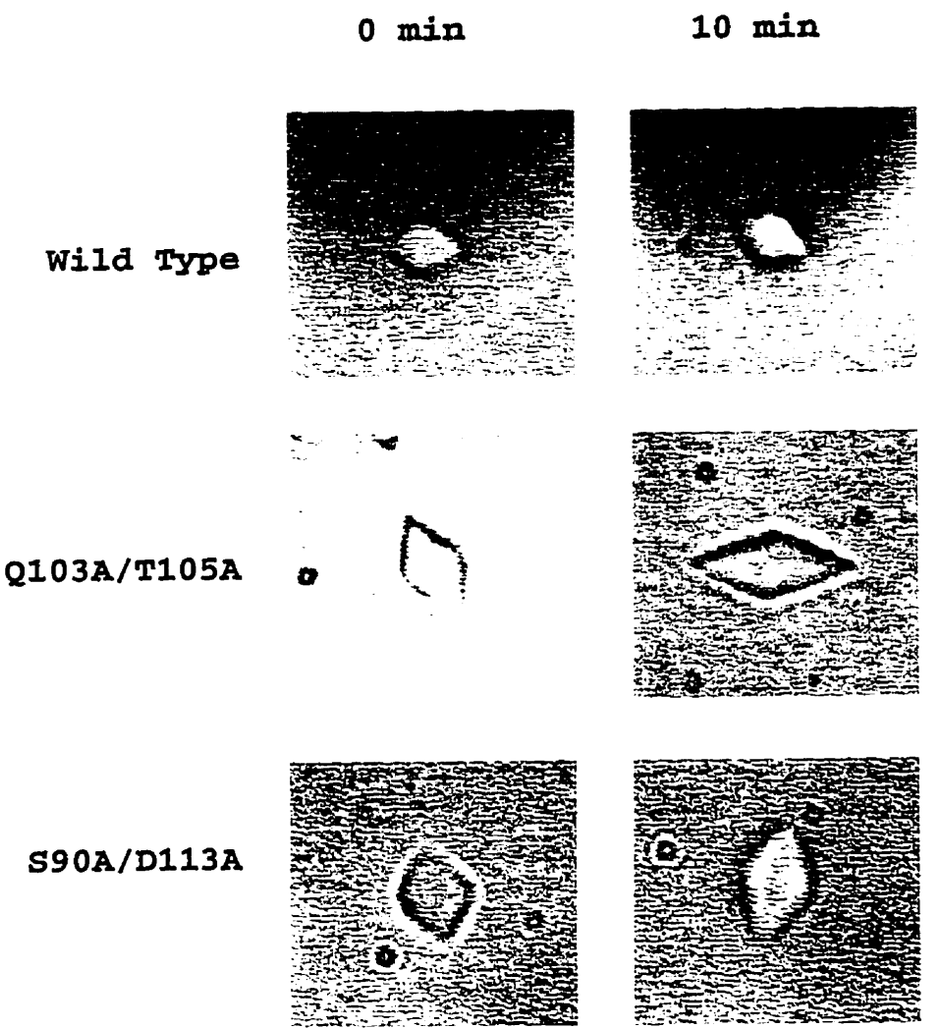

0.6
0.4
0.2
0
0
2
4
6
rSRAFP
Q103A/T105A
S90A/D113A
Thermal Hysteresis (°C)
[AFP] (mg/mL)

Fig. 36 **Ice crystal morphology of IBS1 double site rSRAFP mutants.** Ice crystals were formed in the presence of approximately 1mg/mL protein at 0.1°C cooling. Images were made at 0 and 10min time points to evaluate shapes and growth of ice crystals over time.


0 min
10 min
Wild Type
Q103A/T105A
S90A/D113A

with the crystal being more faceted, with somewhat defined, angular edges, compared to rSRAFP or S90A/D113A (Figure 36). Neither mutant allowed any ice crystal growth over time. These results indicated that while T105 and Q103 might play some role in mediating AFP-ice interactions, the overall binding surface could be somewhat more extensive than that described in the SRAFP model.

ii) Insertion Mutations at IBS1

Due to the lack of effect the double mutations had on activity, more intrusive mutations were designed that involved peptide insertions. Two mutants, an insertion at position 94/95 (94/95) and a second at 107/108 (107/108) (Table 5)(Figure 34A), were produced based on the hypothesis that a peptide insertion, near the ice-binding site might sterically inhibit interactions with ice, without interfering with AFP structure. The insertion sequence was GGYQGG. The Gly were chosen to allow a flexible link between the protein and the insert. Tyr and Gln were chosen for their bulkiness and hydrophilic nature, the latter to minimize solubility problems. As well, codons for Tyr and Gln produced a useful restriction site (*Kpn*I) for future insertion of longer peptides. The exact sites of insertion on the AFP molecule were chosen based on their proximity to the putative ice binding site and their likelihood of structurally accepting an insertion (Figure 34). Initially four sites were tested: 81/82 in loop 2, 94/95 in loop 4, 107/108 between β-strands 3 and 4 and 119/120 between beta-strands 4 and 5. Mutations at positions 81/82 and 119/120 did not yield protein in the yeast system. Insertions 94/95 and 107/108, which most closely test the putative ice binding site produced mutant proteins

at reasonable levels. The 94/95 mutant lost 25% of wild-type rSRAFP thermal hysteresis

Table 5. Amino Acid Insertion Sequences

Insertion Site	Insertion Length (AA)	Insertion Sequence (AA)
94/95	6	GGYQGG
94/95	17	GGYLPDARPAKPLYQGG
94/95	17	GGYSGLAGLASGKYQGG
94/95	28	GGYLPDARPAKPLYLPDARPAKPLYQGG
94/95	28	GGYSGLAGLASGKYSGLAGLASGKYQGG
94/95	61	GGYSGLAGLASGKYLPDARPAKPLYLPDARPA KPLYLPDARPAKPLYSGLAGLASGKYQGG
107/108	6	GGYQGG
107/108	17	GGYLPDARPAKPLYQGG
107/108	28	GGYLPDARPAKPLYLPDARPAKPLYQGG
107/108	28	GGYSGLAGLASGKYSGLAGLASGKYQGG

activity (Figure 37) and its ice crystal morphology showed more sharply defined facets on a crystal which grew slowly over time (Figure 38). The 107/108 insertion, placed the extra domain on the same surface as the 94/95 insert, but at the opposite end of the putative ice-binding surface. This mutant had slightly better thermal hysteresis activity than 94/95 (*ie.* only a 15% loss) (Figure 37). The ice crystals formed in the presence of 107/108 were somewhat larger than wild-type, but maintained the more rounded shape and grew less over time than 94/95 (Figure 38). These results suggested that while 107/108 is close to the ice binding site, 94/95 is likely to be closer.

Attempts were made to insert a linker, coding for 11 additional amino acids, into the *Kpn*I site of the previous insertion mutants. This strategy provided a variety of different sequences and insert lengths, summarized in Table 5. The variety was achieved due to the incorporation of the linker into the single *Kpn*I site. This provided no directional control and gave rise to two possible amino acid sequence inserts depending on the insertion orientation (Table 5), as well as multiple insertion events. Unfortunately these larger insertions appeared to destabilize the overall fold of the AFP to such an extent that these mutant proteins did not accumulate in the yeast medium.

iii) Loop-4 Swap Mutation at IBS1

The SRAFP model deviates from the X-ray structure of the MBP CRD on its putative ice-binding surface. Loop 4 of the MBP structure is flipped up such that it produces a pocket in which the Ca^{2+} binds. Loop 4 of the AFP is thought to be flipped down such that the surface is relatively planar, and potentially ideal for ice interactions

Fig. 37 **Thermal hysteresis activity of rSRAFP with IBS1 insertion and swap mutations.** Thermal hysteresis values for rSRAFP and mutants at various concentrations were compared. Each data point represents the mean of three determinations. Standard deviations are shown as vertical bars. rSRAFP (○), 107/108 (■), Loop4 (●), 94/95 (□).

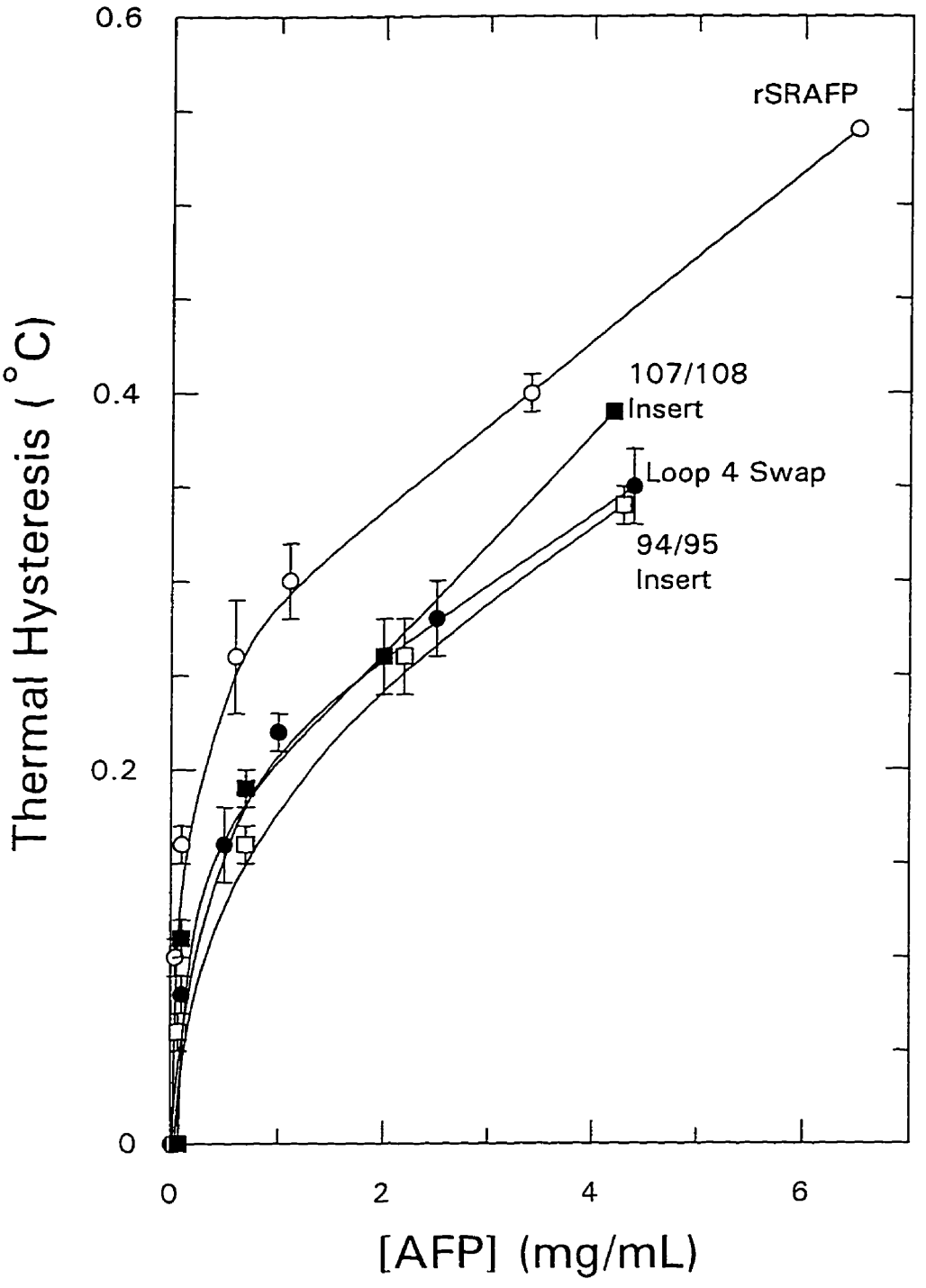
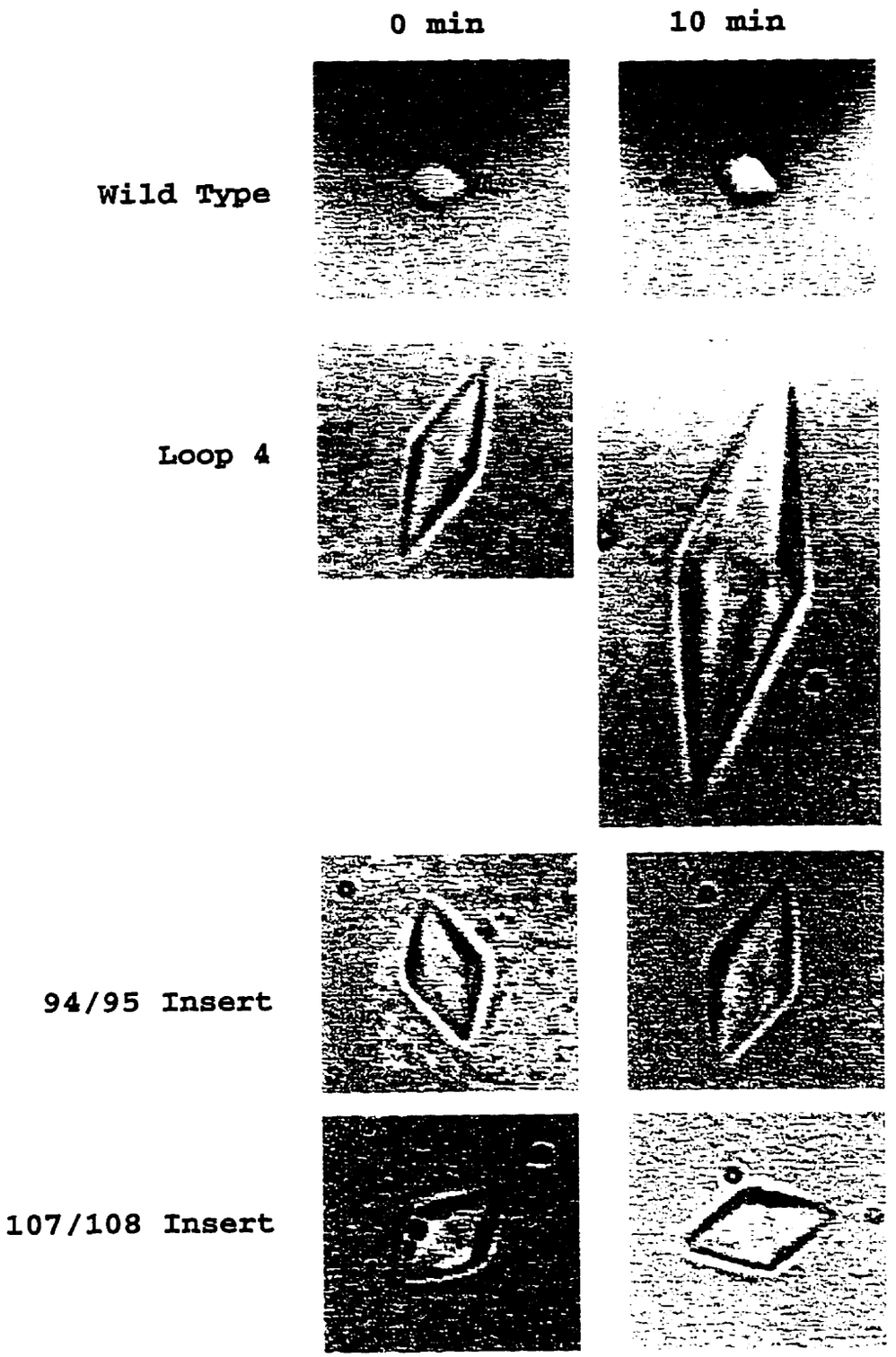

0.6
0.4
0.2
0
Thermal Hysteresis (°C)
rSRAFP
107/108
Insert
Loop 4 Swap
94/95
Insert
0
2
4
6
[AFP] (mg/mL)

Fig. 38 **Ice crystal morphologies of IBS1 insertion and swap mutants.** Ice crystals were formed in the presence of approximately 1mg/mL protein at 0.1°C cooling. Images were made at 0 and 10 min time points to evaluate shapes and growth of ice crystals over time.


0 min
10 min
Wild Type
Loop 4
94/95 Insert
107/108 Insert

(53). This conformation is similar to that described for the E-selectin CRD (128)(63) as well as the PSP (64)(129). The loop 4 swap mutation (Loop4) was designed to take advantage of this proposed structural deviation. It is a domain swap of the nine amino acids (91-TKPDDVLAA-99) making up loop 4 of the SRAFP model (Figure 34A), with the corresponding ten residues in the MBP (DEPNDHGSGE). This swap potentially eliminates the ice-binding site, by deleting ice binding residues and causing a change of conformation (loop 4 flip up) that might block access to the ice-binding surface. Unlike the peptide insertion mutations where the structural changes were unpredictable, the Loop 4 mutation involves the replacement of a structurally homologous region. Yield from shake flask expression improved to 5mg/L of yeast medium, compared to 3mg/L rSRAFP. 1D NMR of this mutant protein indicated no extensive alteration in protein fold compared to rSRAFP (see Appendix A for NMR data). However, a 25% decrease in thermal hysteresis activity was observed (Figure 37). Ice crystal morphology shows an even more dramatic change over that of 94/95, with complete loss of the characteristic small, rounded hexagonal bipyramid, to a much more elongated, faceted crystal, that grew more rapidly over time (Figure 38). Despite the re-introduction of four of the five Ca^{2+} MBP binding residues there was no change in activity, or 1D NMR spectra in the presence or absence of Ca^{2+} (see Appendix A for NMR data).

Taken together these mutants indicated that the putative-IBS1 does in fact play some role in mediating protein-ice interactions. However, it is remarkable that despite these rather drastic mutations at IBS1, the rSRAFP still maintained at least 75% of wild-type activity.

iv) Second Site Mutations

The inability to completely knock out activity by mutation at IBS1 is one indicator that a second ice-binding surface might exist on the SRAFP molecule. Other evidence includes: a) the rounded hexagonal bipyramidal ice crystals characteristically produced in the presence of SRAFP (15) (Figure 22), which could be interpreted as a competition for two or more different ice binding planes on the ice crystal; b) the recently proposed calcium-carbonate ($CaCO_3$) binding surface on the PSP (lithostathine) molecule (64) which differs in position from that of the CRD's. Based on these considerations, a search was initiated for a second ice-binding surface on SRAFP.

Initially, surface accessible, hydrophilic residues conserved between the sea raven, herring and smelt AFPs were targeted for mutagenesis (Figure 6). These mutants included three double mutants: D79A/T81A situated in loop 2 of the SRAFP model, Q55A/T56A in helix 2 and S47A/E50A at the very N-terminus of helix 2 (Figure 34B). Of these mutants, only Q55A/T56A produced protein in the *P. pastoris* system. When it was purified and characterized, Q55A/T56A maintained identical activity compared to wild-type, both in terms of thermal hysteresis and ice crystal morphology (Figure 39A and 40).

A second set of mutants was then tested based on the genomic alignment, which indicated that SRAFP is more closely related to PSP than CRDs (see Discussion and Figure 50). 1° and 3° structure comparisons of SRAFP to lithostathine suggested a number of interesting and previously overlooked sites to be tested based on the putative $CaCO_3$ binding site of the PSP. As well, recent modelling of a proteoglycan tandem

Fig. 39 **Thermal hysteresis activity of rSRAFP with mutations probing for IBS2.** Thermal hysteresis values for rSRAFP and mutants at various concentrations were compared. Each data point represents the mean of three determinations. Standard deviations are shown as vertical bars. A) rSRAFP(○), Q55A/T56A(■), T23H and S120H(●). B) rSRAFP(●), Loop4(○), T23H+Loop4(■), S120H+Loop4(□).

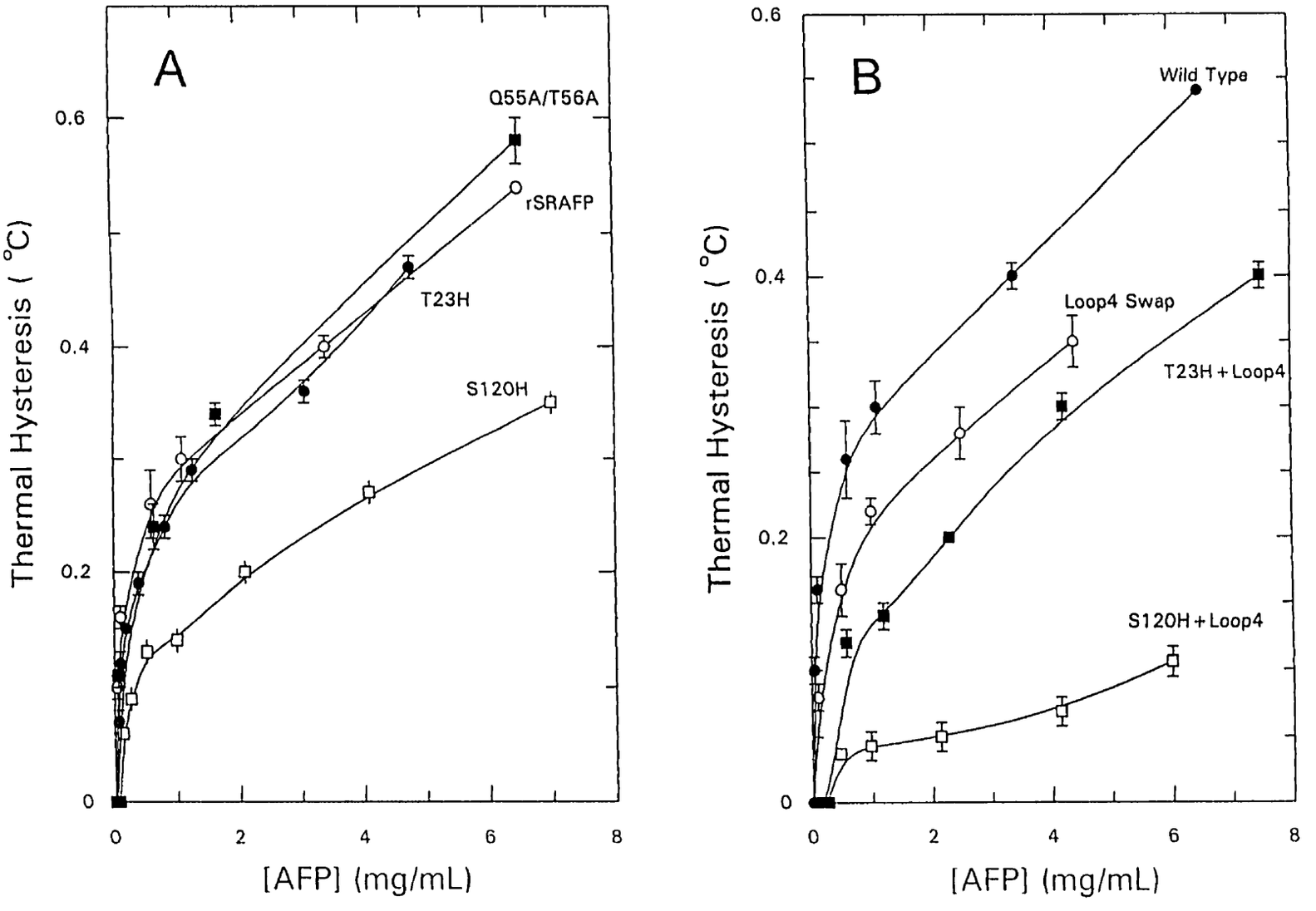
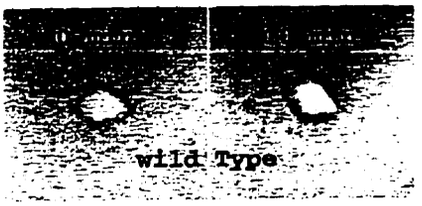
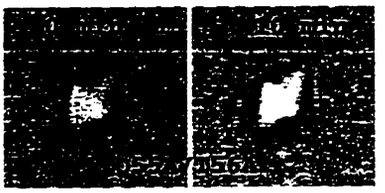
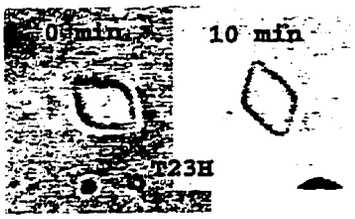
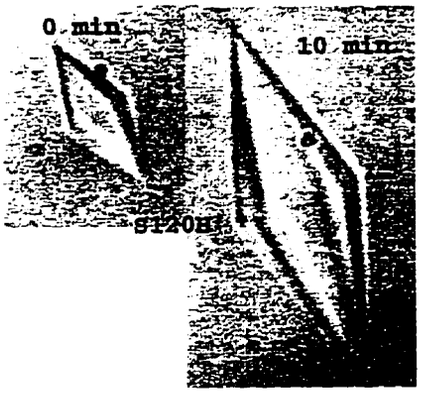

A
Q55A/T56A
rSRAFP
T23H
S120H
Thermal Hysteresis (°C)
[AFP] (mg/mL)
0
0.2
0.4
0.6
0
2
4
6
8
B
Wild Type
Loop4 Swap
T23H + Loop4
S120H + Loop4
Thermal Hysteresis (°C)
[AFP] (mg/mL)
0
0.2
0.4
0.6
0
2
4
6
8

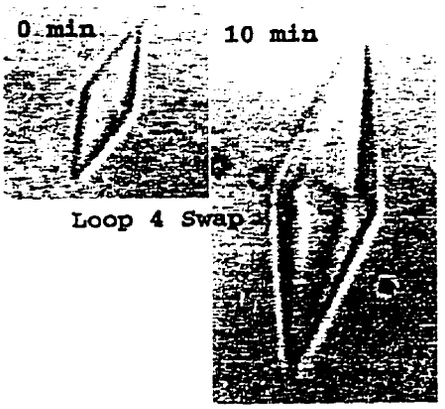
Fig. 40 **Ice crystal morphologies of rSRAFP with second site mutations.** Ice crystals were formed in the presence of approximately 1mg/mL protein at 0.1°C cooling. Images were made at 0 and 10 min time points to evaluate shapes and growth of ice crystals over time. The activity of S120H+Loop4 was so low that even at 6mg/mL, 0.1°C cooling was not achieved, and 0 and 10 min time points were not obtained. The presented S120H+Loo4 crystals show the shape of the ice crystal before and after the burst in a 6mg/mL solution during a regular thermal hysteresis assay. The burst occured at 0.05°C undercooling.

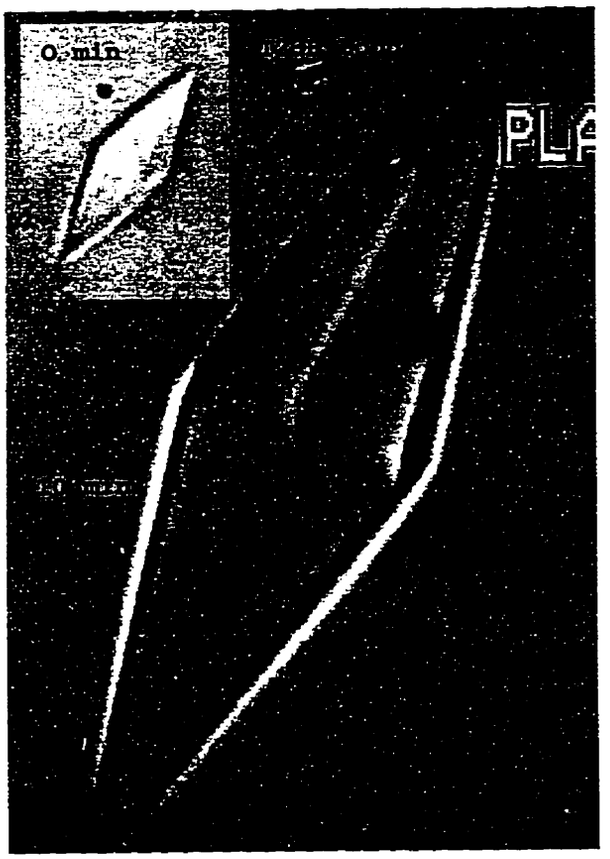

wild Type



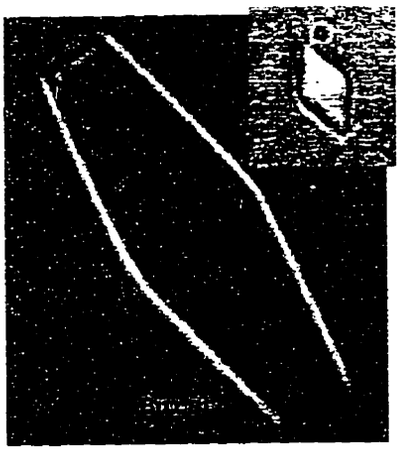

0 min
10 min
T23H


0 min
10 min
S120H


0 min
10 min
Loop 4 Swap


0 min

S120H/Loop4



repeat (PTR) domain, known to bind to hyaluronate in Link protein (130), demonstrated a fold similar to that of the CRD of C-Type lectins. In this PTR domain model a putative binding surface was proposed that included residues from both the CRD binding region, as well as the PSP putative binding region (130). Residues postulated to be part of the PSP $CaCO_3$ binding surface included E30, D31, E33, D37, D72, D73, E133 and D134 (PSP numbering)(64). E30 and D31 correspond to residues T23 and T24 in the SRAFP model (Figure 6), situated immediately prior to helix 1 and in close proximity to T27 within helix 1. Lithostathine E133 and D134 are in a loop region connecting the two final β strands to one another. S120 in SRAFP is also localized to this linker region. Together these residues (T23, T24, T27, S120) could potentially form part of a second ice-binding surface (IBS2), or at least in the case of S120, significantly extend the original binding surface (Figure 34B). Three mutants were initially made at these sites within the loop 4 swap (Loop4) mutant background: S120H+Loop4, T23N/T24N/T27A+Loop4, and S120H/T23N/T24N/T27N+Loop4. Histidine was chosen for substitution at the S120 site because it is a large residue that might cause steric hinderance between the AFP and ice. Asn was chosen for replacement of T23 and T24 for the same reasons and because it maintains the hydrophilic nature of the residues. Ala was used to replace T27 because it is in an α-helix and Ala is the best residue for promoting the formation of α-helical structure (131)(132). Of these mutants, only S120H+Loop4 was successfully expressed by *P. pastoris*. Characterization of purified S120H+Loop4 indicated it maintained only 10% of wild-type activity by the criterion of thermal hysteresis (Figure 39B). Ice crystal monitoring indicated significant growth

over time and loss of the rounded character for a more angular crystal similar to that formed by the Loop4 mutant (Figure 40).

Subsequent mutations including S120H in the wild-type background, as well as T23H, T23H+Loop4 and T24H+Loop4 were produced and characterized. T24H+Loop4 did not express in the *P. pastoris* system. S120H maintained only 50% of wild-type activity (Figure 39A) and demonstrated an ice crystal morphology that grew significantly over time with a *c* to *a* axis ratio somewhat lower (stubbier) than that of the Loop4 or the S120H+Loop4 mutants (Figure 40). T23H had virtually no effect on either thermal hysteresis or ice crystal morphology (Figure 39A an 40), but T23H+Loop4 did exhibit a depressed level of thermal hysteresis activity compared to the Loop4 (Figure 39B) as well as a significantly greater rate of ice crystal growth (Figure 40).

PART B

THE ROLES OF THR, ASN, AND LEU IN HPLC-6

The complexity due to the large size and globular nature of Type II AFP made studying the AFP-ice interaction very problematic. This was further compounded by the many mutants that did not produce protein in the yeast expression system. During the course of this work a second system of AFP-ice interaction was studied, this one using Type I AFP, a significantly smaller protein (37 amino acids), with a regular repetitive α-helical structure (Figures 4 and 5). This structure places putative ice-binding residues

Thr, Asx and Leu in ice binding motifs, at regular intervals every 16.5Å along the same surface of the helix (Figure 16)(40). In contrast to the Type II study, the production of Type I AFP variants was possible by solid-phase synthesis, which avoided the vagaries of biosynthesis, and the amino acid replacements were designed on the basis of a well characterized and highly predictable protein structure.

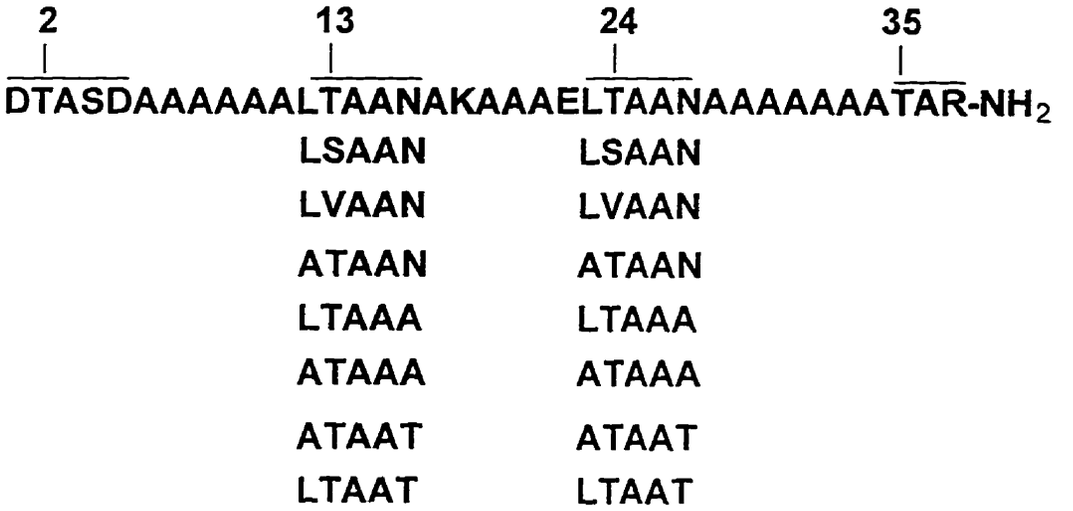
Design of Type I Antifreeze Protein Variants

A systematic replacement of Thr, Asn and Leu residues within the two central ice-binding motifs (LTAAN) allowed the same substitution to be made twice, thereby magnifying the effect of the change, without disrupting the internally hydrogen-bonded N- and C- terminal helix-cap structures (Figure 41).

Specifically, Thr was mutated to Ser (LSAAN). This deleted the methyl group on the side chain, to assess its contribution to ice binding. Similarly Thr was mutated to Val (LVAAN). This maintained the methyl group and changed the hydroxyl group to another methyl group. This change eliminates any possible hydrogen bond formation, while still maintaining the same basic size and shape of the side chain.

Three mutations to test the role of Asn and Leu were made: LTAAA, ATAAN and ATAAA (Figure 41). Interestingly this third variant, in which both residues were replaced by Ala, has been naturally selected for in yellowtail flounder (*Limanda ferruginea*) Type I AFP (133). Another naturally selected ice binding motif (ATAAT) in the less abundant AFP-9 isoform in winter flounder serum (35) was tested. AFP-9 has higher activity than HPLC-6 on a molar basis. Whether this is due to the presence of

Fig. 41 **Type I AFP mutant primary sequences.** The two central IBMs (all IBM are overlined) received the same alteration.


2
13
24
35
DTASDAAAAAALTAANAKAAAELTAANAAAAAAATAR-NH2
LSAAN
LVAAN
ATAAN
LTAAA
ATAAA
ATAAT
LTAAT
LSAAN
LVAAN
ATAAN
LTAAA
ATAAA
ATAAT
LTAAT

an additional 11-amino-acid-repeat, or the alternative ice-binding motif is not known. By inserting this motif into the HPLC-6 peptides the contribution of this motif can be compared in a three repeat Type I AFP (Figure 41). One final variant was designed based on this ATAAT motif, in which the Leu was reintroduced (Figure 41), based on the observation by Wen and Laursen (104) that replacement of L12 and L23 by Ala caused a 33% loss of activity.

Antifreeze Activity

The thermal hysteresis activity of synthetic wild-type HPLC-6 was indistinguishable from that of the natural protein from flounder serum (109). However, LSAAN exhibited a surprising loss of AFP activity, being completely unable to halt ice growth below a concentration of 2mg/mL. Above 2mg/mL LSAAN showed weak thermal hysteresis (up to 0.1°C) within which the ice crystal was held as a hexagonal bipyramid with an average *c* to *a* axis ratio of 3.2 (Figure 42 and 43). This crystal was indistinguishable from that obtained with wild-type synthetic AFP. In contrast, the variant LVAAN was extremely active. It maintained up to 80-90% of the activity of wild-type AFP (Figure 42) over the entire concentration range tested. The ice crystal formed in the presence of 1mg/mL LVAAN also had an average *c* to *a* axis ratio of 3.2 (Figure 43). This crystal grew imperceptibly both during the thermal hysteresis readings as the temperature was lowered in stages, and when held at 0.1°C below the melting point for 10 min. In both cases growth was only observed by time lapse video microscopy. The crystal obtained in 1mg/mL LTAAN did not grow under either

Fig. 42 **Thermal hysteresis activity of Type I AFP Ser and Val mutants.** T13 and T24 were mutated to either Ser (LSAAN) or Val (LVAAN) and their thermal hysteresis values compared at various concentrations to wild-type HPLC-6 (LTAAN). Each data point represents the mean of at least three determinations and standard deviations are shown as vertical bars.

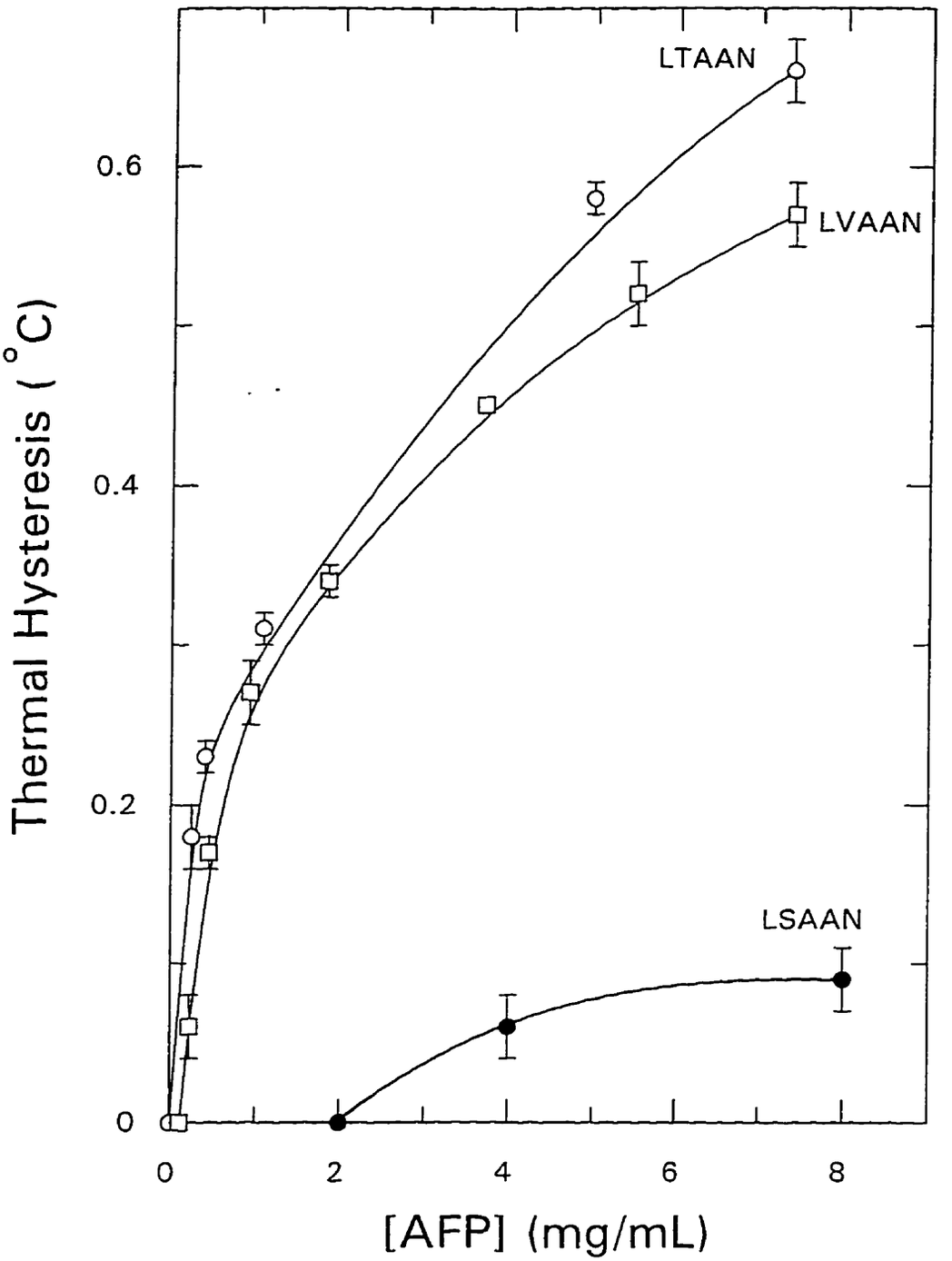
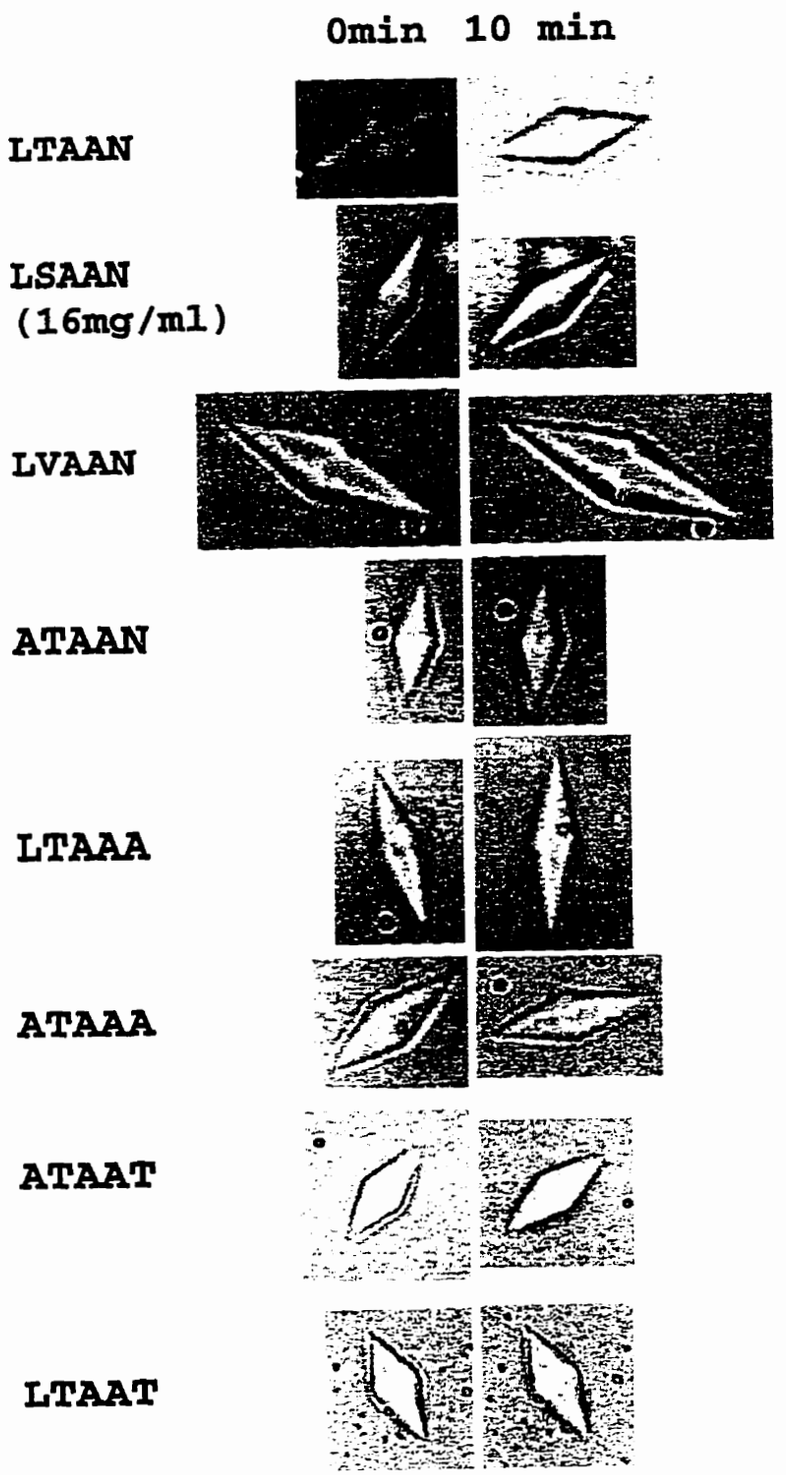

LTAAN
LVAAN
LSAAN
0.6
0.4
0.2
0
0
2
4
6
8
Thermal Hysteresis (°C)
[AFP] (mg/mL)

Fig. 43 **Ice crystal morphologies of Type I AFP mutants.** Ice crystals were formed in the presence of approximately 1mg/mL protein at 0.1°C cooling. Images were made at 0 and 10 min time points to evaluate shapes and growth of ice crystals over time. LSAAN (the least active mutant) was assayed using a 16mg/mL protein solution to achieve the 0.1°C cooling.


0min 10 min
LTAAN
LSAAN
(16mg/ml)
LVAAN
ATAAN
LTAAA
ATAAA
ATAAT
LTAAT

circumstance.

At a concentration of 1mg/mL, ATAAN maintained 70% of the activity of the wild-type AFP, compared to 61% observed for LTAAA and 80% for ATAAA (Figure 44). All ice crystals formed in the presence of these variants over a range of concentrations, maintained average *c* to *a* ratios close to the 3.2 value observed for wild-type synthetic AFP. Only the LTAAA crystal grew imperceptibly to the naked eye when held at 0.1°C below the melting point for 10 minutes (Figure 43). All three variant ice crystals grew imperceptibly during thermal hysteresis readings as the temperature was lowered in stages.

Over the concentration range tested, ATAAT maintained virtually identical thermal hysteresis activity to that of the wild-type AFP, at levels below 2mg/mL (Figure 45). At higher concentrations (to 4mg/mL) an increase in activity was observed to 20% over LTAAN. The incorporation of Leu into this motif caused a decrease in activity compared to ATAAT (Figure 45). Activity for LTAAT was virtually identical to wild-type at levels below 3mg/mL, but demonstrated a 20% increase compared to LTAAN at 7mg/mL. The ice crystals of these two variants had average *c* to *a* ratios of 2.7 and 2.9 respectively, that were significantly shorter than the 3.2 ratio obtained with wild-type HPLC-6 (Figure 43). Both variants produced ice crystals that grew imperceptibly during thermal hysteresis readings as the temperature was dropped, but only the more active LTAAT variant grew when held at 0.1°C below the melting point for 10 minutes. In all cases, growth was only observed by time-lapse video-microscopy, and *c* to *a* ratios were determined by statistical analysis of between 10 and 16 crystals for each variant.

Fig. 44 **Thermal hysteresis activity of Type I AFP with Leu, Asn and Leu/Asn mutations.** Thermal hysteresis of wild-type HPLC-6 (LTAAN)(○) was compared at various concentrations to Leu (ATAAN)(●), Asn (LTAAA)(□) and the double Leu/Asn (ATAAA)(■) mutants at various concentrations. Each data point represents the mean of at least three determinations. Standard deviations are shown as vertical bars.

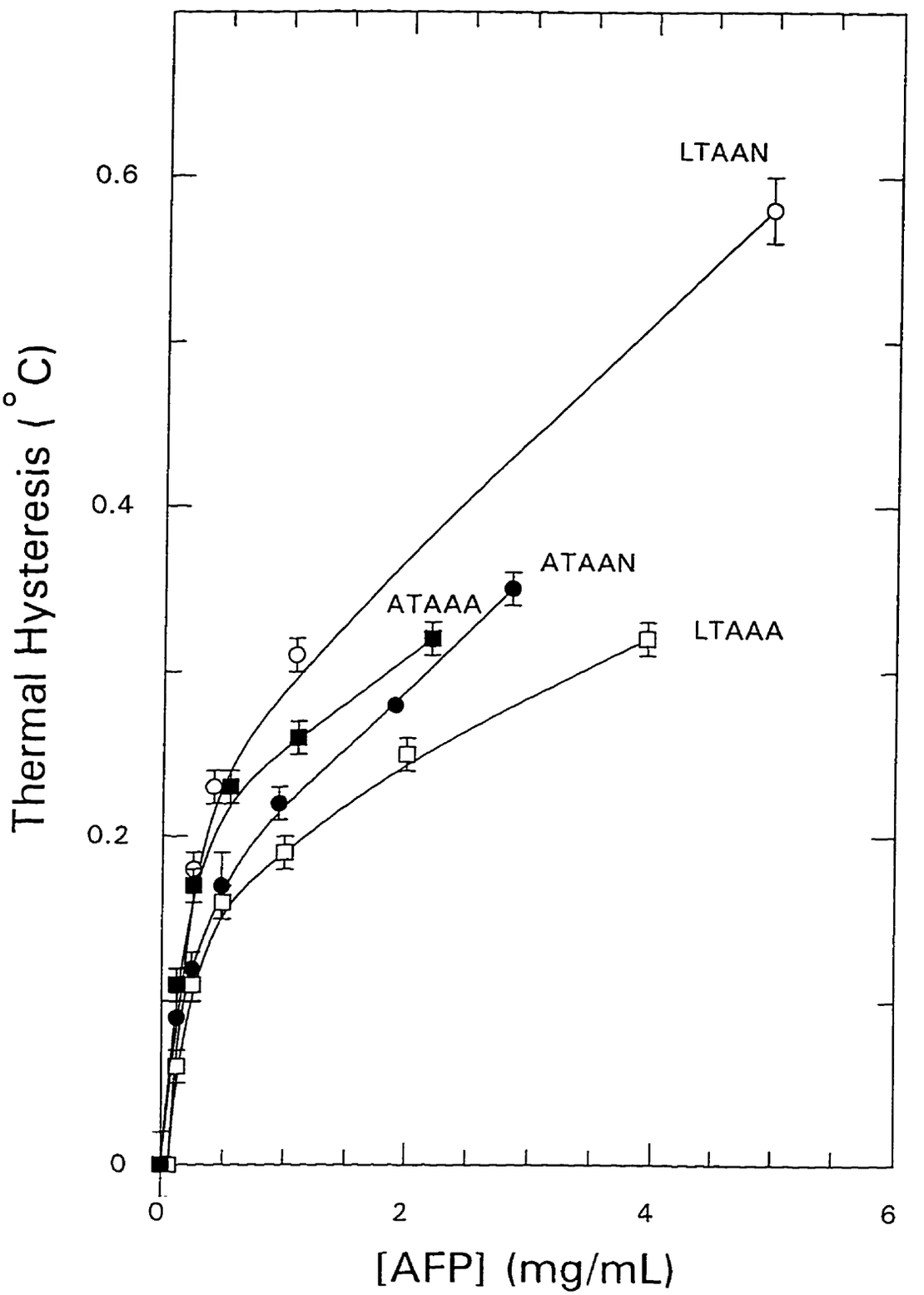
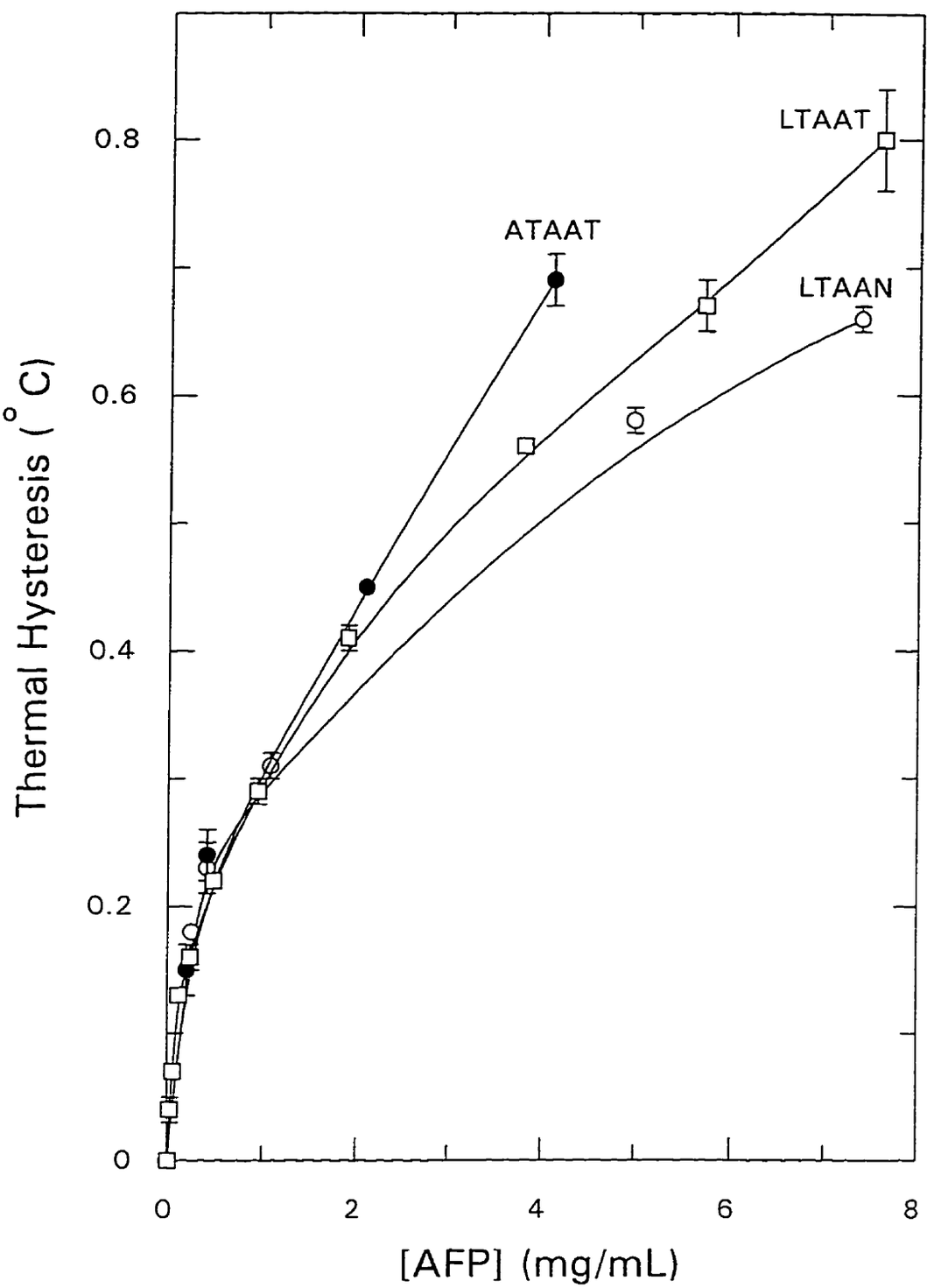

LTAAN
ATAAN
ATAAA
LTAAA
0.6
0.4
0.2
0
0
2
4
6
Thermal Hysteresis (°C)
[AFP] (mg/mL)

Fig. 45 **Thermal hysteresis activity of Type I AFP variants LTAAT and ATAAT.** Thermal hysteresis values of wild-type HPLC-6 (LTAAN) (○) was compared to the double Thr (ATAAT) (●) and the Leu/double Thr (LTAAT)(□) mutants at various concentrations. Each data point represents the mean of at least three determinations. Standard deviations are shown as vertical bars.


LTAAT
ATAAT
LTAAN
0.8
0.6
0.4
0.2
0
0
2
4
6
8
Thermal Hysteresis (°C)
[AFP] (mg/mL)

CD, Sedimentation and NMR Analysis

Circular dichroism studies on LTAAN (41)(45)(131), LSAAN and LVAAN indicated that the two variants were slightly more helical than wild type, which is in accordance with the helix propensities of Ser and Val compared to Thr (131)(132). Sedimentation analyses showed that both variants remained monomeric in solution up to concentrations of at least 6.5mg/mL. NMR showed that there was no significant change in the overall structure of the helical variants and also demonstrated that the Val in the LVAAN variant had a 70% preference for the χ^1-60° rotamer. This is the same rotamer preferred up to 55% by Thr at the same position. These biophysical analyses were carried out by Drs. Heman Chao, Wolfram Gronwald and Frank Sönnichsen. A more detailed summary of these results is described in Appendix B.

As predicted, none of the mutations at Leu and Asn had any significant impact on helicity compared to wild-type AFP. Again detailed CD results for these variants are available in Appendix B. Significant solubility/aggregation problems were encountered when certain variants, primarily those containing Leu to Ala changes, were redissolved at high concentration for CD and activity analysis. Self association or aggregation of the peptide in solution could account for the observed depressions in activity levels. To investigate this, attempts at determining molecular weights of all the remaining variants were carried out by sedimentation equilibrium analysis. Mw were not obtained for any variant containing a Leu to Ala Change. All these variants underwent a slow and irreversible aggregation during dialysis prior to the sedimentation. A Mw of 3,886 was obtained for the LTAAA variant although higher order species were present in the

solution as well. LTAAT was completely soluble. This data suggested that loss of activity at higher concentrations is likely due to aggregation in solution and suggests a role for Leu, and to a lesser extent Asn, in increasing solubility of these Ala-rich helices.

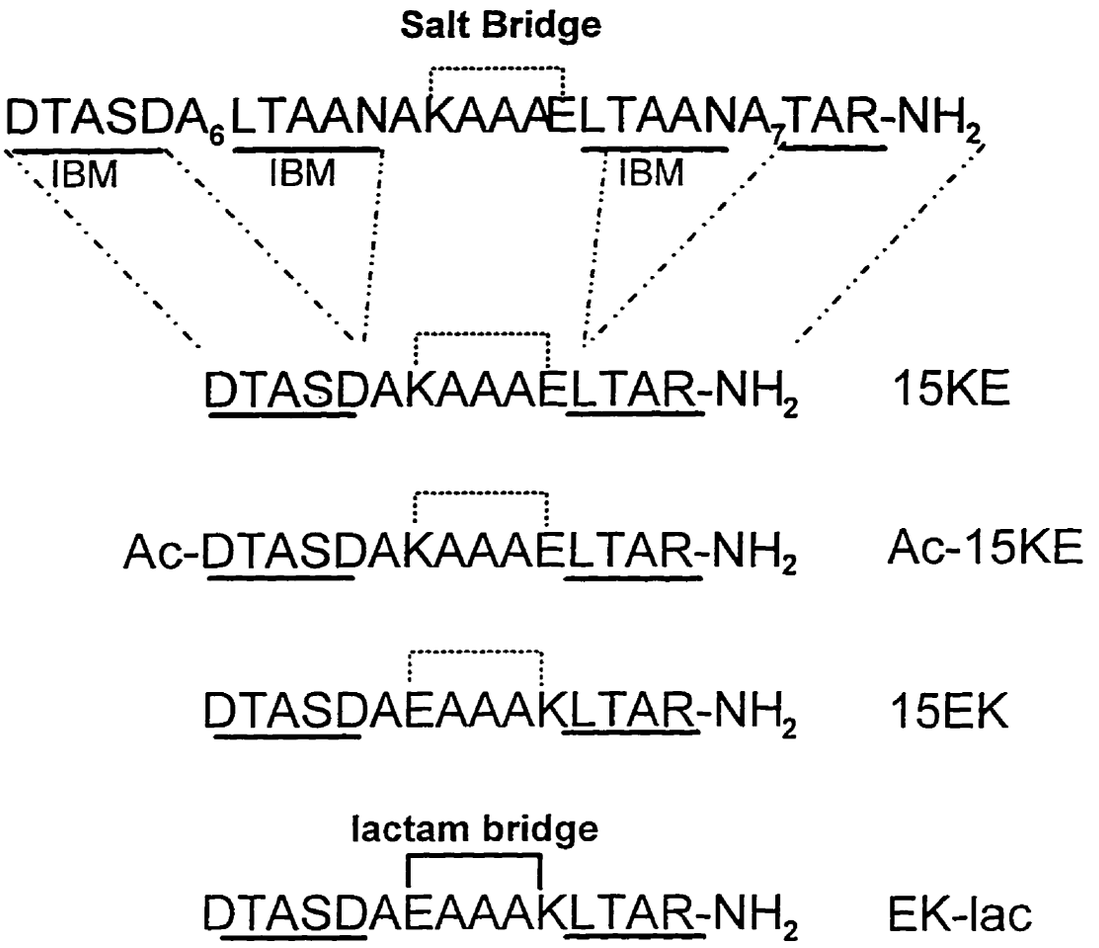
MINIMIZATION OF HPLC-6

Design of the Minimized Antifreeze Peptides

Type I AFP is an ideal protein for minimization because there are few structural restraints to making it smaller by simply shortening the helix. This is not possible with the globular AFPs, where the entire molecule is required to make the overall 3D fold. The initial minimized AFP structure (15KE) was formed by simply removing the two central 11 amino acid repeats (residues 13 to 34 inclusively) while maintaining the internal salt bridge on the non-ice binding face of the helix (Figure 46). The remodelling maintained the critical 16.5Å spacing between the remaining two Thr assuming the peptide retained its helicity. Helicity was the primary concern in designing the minimization. By shortening the protein from the center, the helix stabilizing N- and C-terminal cap structures were not disturbed.

Three separate adjustments were made to this basic design in an effort to improve helicity. In one peptide the amino terminus was acetylated (Ac-15KE). Previous work reported that the unfavourable effect of protonation at the N-terminus is approximately 0.5kcal/mol and that acetylation of the free amino group can improve helix stability (44). In a second synthesis, the order of the ion pair residues K7 and E11 was reversed to E7

Fig. 46 **Design of minimized Type I AFP peptides.** The primary sequences of four minimized peptides are shown, each shortened from 37 amino acids to 15. A lactam bridge was introduced in the EK-lac peptide, replacing the salt link with a covalent bond. 15EK had the orientation of the salt bridge switched with respect to the wild type. Ac-15KE included an acetyl group at the N-terminus and 15KE is the basic unmodified minimized peptide.


Salt Bridge
DTASDA6LTAANAKAAAELTAANA7TAR-NH2
IBM
IBM
IBM
DTASDAKAAAELTAR-NH2
15KE
Ac-DTASDAKAAAELTAR-NH2
Ac-15KE
DTASDAEAAAKLTAR-NH2
15EK
lactam bridge
DTASDAEAAAKLTAR-NH2
EK-lac

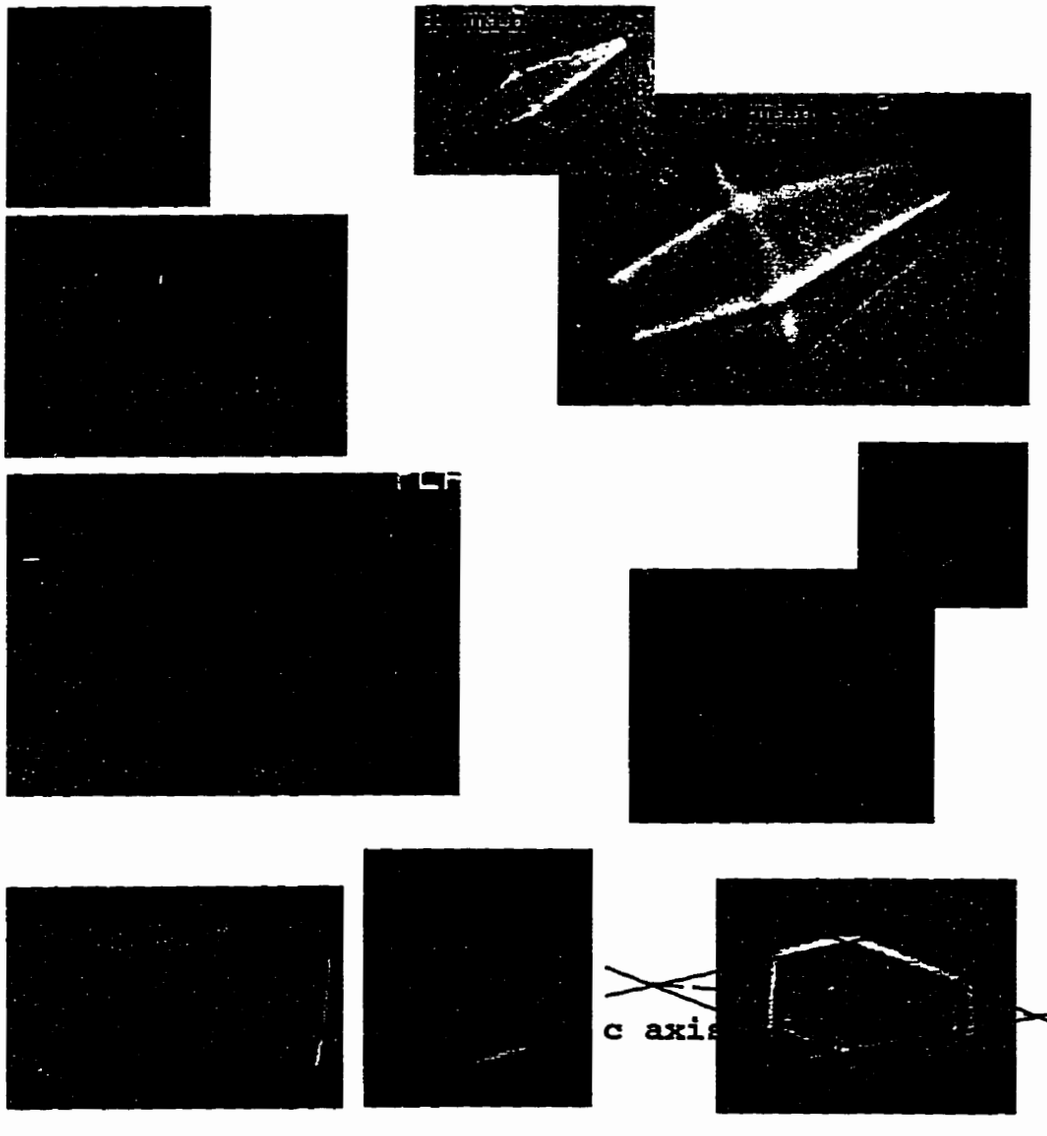
and K11 (15EK), in order to orient them more favourably to the helix macrodipole (134)(135). A third peptide was designed using the 15EK as a basis in which the salt bridge was made into a covalently bound lactam bridge (Lac-15EK). The EK peptide was chosen, as the E7/K11 order simplifies production of the lactam bond. Lactam bridges have been previously shown to significantly increase the helicity of α-helical peptides (136) (Figure 46).

Antifreeze Activities

None of the 15mer peptides were abie to depress the freezing point below the melting point at concentrations of 50-100mg/mL (approx. 50mM). Even the slightest undercooling caused seed ice crystals to grow steadily (Figure 47). 15KE and 15EK produced ice crystals with the typical rounded disk-like shape observed in pure water. Ac-15KE produced crystals with very slight shaping into a hexagonal disk. The Lac-15EK peptide produced ice crystals with significantly altered morphologies (Figure 47). They were shaped into incomplete hexagonal bipyramids which generally lacked both apices. By extrapolation, the *c* to *a* axial ratio of these imperfect hexagonal bipyramids averaged about 3.3, which is identical to that observed in previous studies with the native full-length HPLC-6 peptide (104)(16). This value was maintained even as the ice crystal grew over a time course of 1 to 10min, indicating a persistent influence of the peptide over the ice surfaces expressed.

Fig. 47 **Ice crystal morphologies of the minimized peptides.** Ice crystals were produced in the presence of between 50 and 100mg/mL peptide. A variety of examples of the EK-lac crystals are given in the bottom three quarters of the figure, including three time courses identified with the 1 min, 5 min and 10 min time point labels to illustrate the growth over time. The *c* and *a* axes depicted in the bottom right hand crystal provides an example of how *c* and *a* axis ratios were determined.

EK-1ac



CD of the Minimized Peptides

CD studies indicated that the 15KE and the 15EK peptides had significant instability in their helical structures. The Ac-15KE peptide had even lower helical values while the Lac-15EK peptide maintained the highest helicity (90% compared to the 100% theoretical value for a 15 residue α-helix (137). CD studies were carried out by Dr. Heman Chao at the University of Alberta. Detailed CD results are described in Appendix C.

CONJUGATION OF A MINIMIZED PEPTIDE TO BSA

Design of the Peptide and Conjugation to BSA

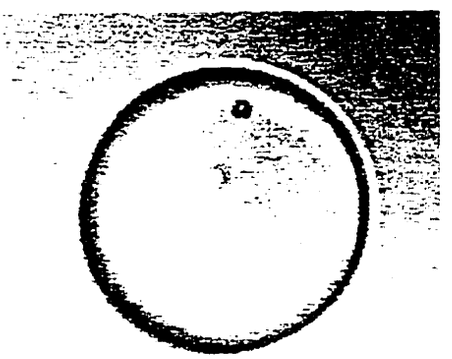
Results of Type III AFP domain addition experiments indicated that large adducts can increase AFP activity (111). This observation was applied to the minimized lactam-bridged sample. A slightly modified sequence was used, that replaced A8 with a Ser residue (SLac-15EK). This was done because polar residues in the vicinity of the bridge residues can simplify formation of the lactam link (personal communication, M. Houston). The overall sequence was DTASTAESAAKLTAR-amide, with a lactam bridge between E7 and K11. This peptide was conjugated to BSA with a final average ratio of five peptides bound to every BSA molecule. BSA was chosen as the additional domain due to its large size (67kDa), its availability at relatively low cost, and its familiarity as a serum protein.

Antifreeze Activity

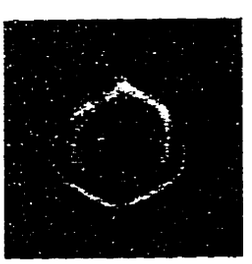
At a concentration of 66.7mM the SLac-15EK peptide showed no thermal hysteresis and no alteration of ice crystal morphology. This is in contrast to the original Lac-15EK (without the A8S change) that did significantly alter ice crystal morphology at 50mM concentration (Figure 47). At 0.2mM, BSA showed no thermal hysteresis activity nor any alteration of ice crystal morphology. The 5:1 conjugated product dissolved at 0.5mM BSA and an apparent peptide concentration of 2.7mM, had up to 0.06°C thermal hysteresis activity and produced hexagonal bipyramidal ice crystals that grew slowly over time (Figure 48).

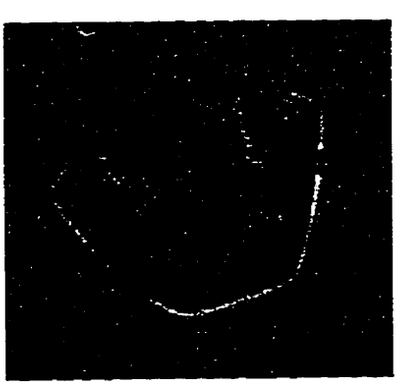
Fig. 48 **Ice crystal morphologies of the conjugated BSA-peptide construct.** Ice crystals were produced in the presence of 100mg/mL BSA-peptide (1:5) conjugate.

Water



BSA-Peptide Conjugate





CHAPTER 4

DISCUSSION

PART A

Optimization of Antifreeze Protein Expression in *P. pastoris*

A number of different methods were tested to optimize the expression of rSRAFP from the yeast *P.pastoris*. While both the alkaline phosphatase signal peptide and the α-mating factor signal peptide are intrinsic to yeast, the latter was found to be significantly more effective at increasing rSRAFP yields irrespective of whether or not the AFP had an N-terminal His-tag. This observation was belatedly confirmed by the manufacturer (Invitrogen) based on experience with expression of a variety of other proteins. The addition of the His-tag did not affect the activity of the rSRAFP in any manner. This was not unexpected because the potential ice binding site (IBS1) is located at the other end of the molecule from either the N- or C-termini according to the model of Sönnichsen *et al.*, 1995 (53). The His-tag provided a very effective first purification step from the peptide-rich, rSRAFP-dilute yeast media. A similar comment could have been made for the N-terminal extension, except that mass spectrometry results strongly suggest that this peptide sequence was removed by post-translational processing.

A number of different fermentation strategies were tested to optimize the expression of rSRAFP from *P. pastoris*. The goal was to reach levels of rSRAFP that

would provide quantities (>100mg) of purified protein sufficient for research and development applications and X-ray crystallography. NMR structural studies, while not requiring large quantities of protein, are facilitated by labelling the protein with ^{15}N. In order to produce labelled proteins, the host system must be grown in a minimal media with ^{15}N as the sole nitrogen source. It was demonstrated that at the shake flask level, yields dropped significantly (at least two-fold) in minimal media. However shifting to fermentation with a defined minimal medium and a mixed-feed strategy, improved yields 10-fold, up to 35mg/L. This appears to be due largely to an increase in cell mass and not to any increased productivity of the yeast. The levelling off of cell growth towards the end of the cultivation period is likely due to nitrogen depletion. It was anticipated that the addition of peptone during the induction phase of the fermentation might increase yields dramatically, as was observed in *B. brevis* 47 (126). In the bacterial situation the expression of extracellular protein was increased five-fold from 3g/L to 15g/L. Unfortunately, peptone did not have the same effect on foreign protein expression from *P. pastoris*. Similarly, the Mut^+ fermentation, with the continuous methanol feed protocol showed no significant increase in rSRAFP yields over the Mut^s mixed feed. Due to the large amount of methanol consumed by Mut^+ to grow to a CDW of 70g/L (because of its low yield coefficient of 0.15 compared to 0.42 for glycerol) it is preferable to use the Mut^s with the mixed-feed fermentation profile.

These results suggest that varying medium formulations, feeding strategies and cultivation conditions are largely ineffective in improving individual cell performance in expressing the rSRAFP protein; however these techniques are able to generate much

higher final protein levels due to the substantially increased cell concentrations obtained. It is very likely that the nature of the heterologous protein itself is an important factor in the levels of gene product expressed by *P. pastoris*. Expression levels can be one or more orders of magnitude different depending on which protein is produced, not withstanding the use of similar expression vectors, strains and cultivation conditions (138).

Future optimization should focus on the development of more productive genetic variants. This will include the isolation of clones with multi-copy insertions of the rSRAFP expression cassette into the yeast genome. A second possibility is the production of fusion proteins, where the cDNA of the SRAFP is fused to the gene of a protein known to be expressed at high levels in the *P. pastoris* system (138).

Comparison of Expression Systems

A comparison of cystine-rich SRAFP expression in three different biosynthetic systems (Table 6) has led to the development of an effective expression/purification scheme. Originally a baculovirus-infected insect cell system was tested based on its ability to duplicate many post-translational modifications, including signal peptide cleavage and disulfide bond formation (60). This system met with only modest success, producing up to 0.8mg/L (S. Gauthier unpublished results) of glycosylated, yet active, pro-Type II AFP secreted into the medium (60). Subsequent attempts at expression in *E. coli*, using both the pT7/pGP1-2 heat-inducible system (116) (M. Loewen unpublished results) and the pET-IPTG inducible system (Novogen), produced

Table 6. Summary of Expression Yields

Host System	Medium	Yield
P. pastoris		
Flask: Mut^s,pPIC9-SRm CTHT	BMMY	3mg/L
	Defined	Undetectable
Fermentation:		
Mut^s MeOH feed	Defined	2mg/L
Mut^s Mixed feed	Defined	30mg/L
Mut^s Peptone	Defined	35mg/L
Mut$^+$ MeOH feed	Defined	37mg/L
Baculovirus*	Complex	<1.0mg/L
*E. coli***	Complex	2mg/L

* ProType II AFP was secreted from fall army worm Sf21 cells by transfected pBlueBac-srAFP vector as previously described (60).

** Mature Type II AFP was produced intracellularly in *E. coli* JM83 cells by isopropyl ß-D-thiogalactopyranoside induction of the T7 promoter on the pet20b+ (Novogen, Madison WI) expression vector, containing the mature SRAFP cDNA inserted into the *NdeI* and *BamHI* sites by PCR cloning (unpublished results).

yields of up to 4.0mg/L of rSRAFP unpurified (M. Loewen unpublished results). At least 50% of this was present in inclusion bodies, and was not inclined to refold into active AFP following solubilization. In contrast, expression of the herring Type II AFP (40% identical to SRAFP) in a similar *E. coli* system, yielded greater than 20mg/L in the insoluble fraction which was successfully refolded and purified (V. Ewart, personal communication). Comparison of herring and sea raven AFP expression levels demonstrates the variability in expression levels that can exist between even closely related proteins. This observation suggests that there exists some intrinsic characteristics to each individual protein which play a role in determining expression levels. Expression strategies for proteins should therefore be developed on a case-by-case basis. Expression of rSRAFP in the yeast, *P. pastoris*, as described here, met with the best success, achieving levels of up to 5mg/L in shake flask, which were improved 7-fold through the use of fermentation, even with switching to defined media.

A Putative Ice Binding Site (IBS1)

The 3D model for SRAFP led to a proposal for an ice-binding surface (IBS1) for ice/protein interactions (53). The surface is made up of loops 3 and 4 and β-strands 3 and 4 of the SRAFP model (Figures 7 and 34). The surface is highly hydrophilic, contains a large number of hydrogen-bonding amino acids, both charged and polar, and is relatively planar. These traits reflect the traditional view of ice-binding surfaces as judged from AFP Types I (40) and III (74). IBS1 also corresponds to the Ca^{2+} binding site on the CRD structure (53).

Loop4 and the two insertion mutants were produced to probe IBS1 in a very general manner. Both the Loop4 and 94/95 mutants conclusively demonstrate that the loop 4 region of the putative ice binding site is important for mediating ice-binding interactions. For Loop4 at least, NMR analyses suggest that the swap caused only local changes to the structure. The 107/108 mutant suggests that the opposite end of that surface is also important, although not to the same degree as the loop 4 region. We can deduce little about the exact residues involved in hydrogen bonding based on these mutations. It is possible that T91, K92 or D94 might play a role, but this will have to be further tested by site-specific mutagenesis. One of the double site-specific mutants implicated T105 and Q103 in mediating ice binding interactions. S90 and D113 were shown to play a minimal role, if any. Overall this work indicates that IBS1 is involved in the ice-AFP interaction.

The recently published X-ray crystal structure for a PSP, lithostathine, has an overall fold similar to that of the CRD and the AFP model (64). However, the PSP has a proposed $CaCO_3$-binding site on a surface other than that corresponding to the Ca^{2+}-binding surface of the CRD. Based on the organization of genomic sequences SRAFP is most closely related to PSP. This, together with the observed lack of Ca^{2+}- dependent activity, suggests that potential SRAFP ice-binding sites are not necessarily limited to the original IBS1.

Sea Raven Versus Herring and Smelt Antifreeze Proteins

It is interesting that of all the implicated IBS1 residues, only Q103 is conserved

between herring, smelt and sea raven AFPs. S90 and T91 are Ala in herring and smelt, K92 is Gln, D94 is Thr or Phe and T105 is Asn. Other differences include the recently published observation that herring AFP has metal ion-dependent conformation and activity, implicating the Ca^{2+} binding surface in its activity (51). SRAFP does not have any requirement for metal ions. Finally, comparison of ice crystal morphologies produced by herring and sea raven AFPs, indicate that the herring AFP ice crystal is very elongated and strongly faceted (57) while the SRAFP crystal is very stubby and round. As well, the herring ice crystals took on significantly different shapes in the presence of alternative divalent metal ions, ranging from the hexagonal bipyramid to basic rectangular forms. This observation is further evidence that the Ca^{2+} site is involved in the ice-AFP interaction (57). The significance of these observations is that SRAFP may possibly have a different binding site(s) and ice-binding residues than the herring and smelt AFPs even on the same face of the protein. This concept of different binding sites to carry out the same function in related proteins would not be unique to this group of AFPs. Sculpin and flounder have related AFPs of the Ala-rich α-helical type, but they contain different hydrophilic residues and have been shown to bind different ice planes (27). As well, a study on three hydrophobic-molecule-binding proteins, that show up to 90% identity to each other, demonstrated a virtually identical fold for each (139). Despite this, the three proteins have very different ligand specificities and their binding sites are in very different regions of the molecules.

A Second Ice Binding Site? IBS2

Several attempts to disrupt IBS1 by mutagenesis involving insertions, a sequence swap and amino acid replacement caused at the most a 25% loss of antifreeze activity. A total of seven hydrophilic residues in the IBS1 region including all the Thr and Ser residues and one of the two Gln and two of the four Asp were mutated in these experiments. This is in contrast to Type III AFP, where mutagenesis of only one binding residue (T18N) resulted in a loss of 90% of activity and other single mutations saw losses of 75% (N14S) and 50% (Q44T) (73)(74). It is possible that the few unchanged hydrophilic residues of the IBS1 are important to ice binding, but the model indicates that they are not positioned in a surface accessible manner (53). Another possibility is that elimination of all of the implicated residues (T105, Q103,T91,K92 and D94) at once or in a particular combination would result in greater loss of activity. However this seems inherently unlikely in view of the drastic activity losses seen with AFP Types I and III when only one or two residues are altered (74)(104)(111). Thus the minimal loss of activity observed to date suggests that even as a group, the hydrophilic residues of the IBS1 do not represent all the binding residues. Therefore, a more plausible explanation, in light of these mutagenic results and comparisons of SRAFP to PSP and to herring and smelt AFPs, is that there are at least two binding sites which interact with ice independently of each other, of which IBS1 is only one.

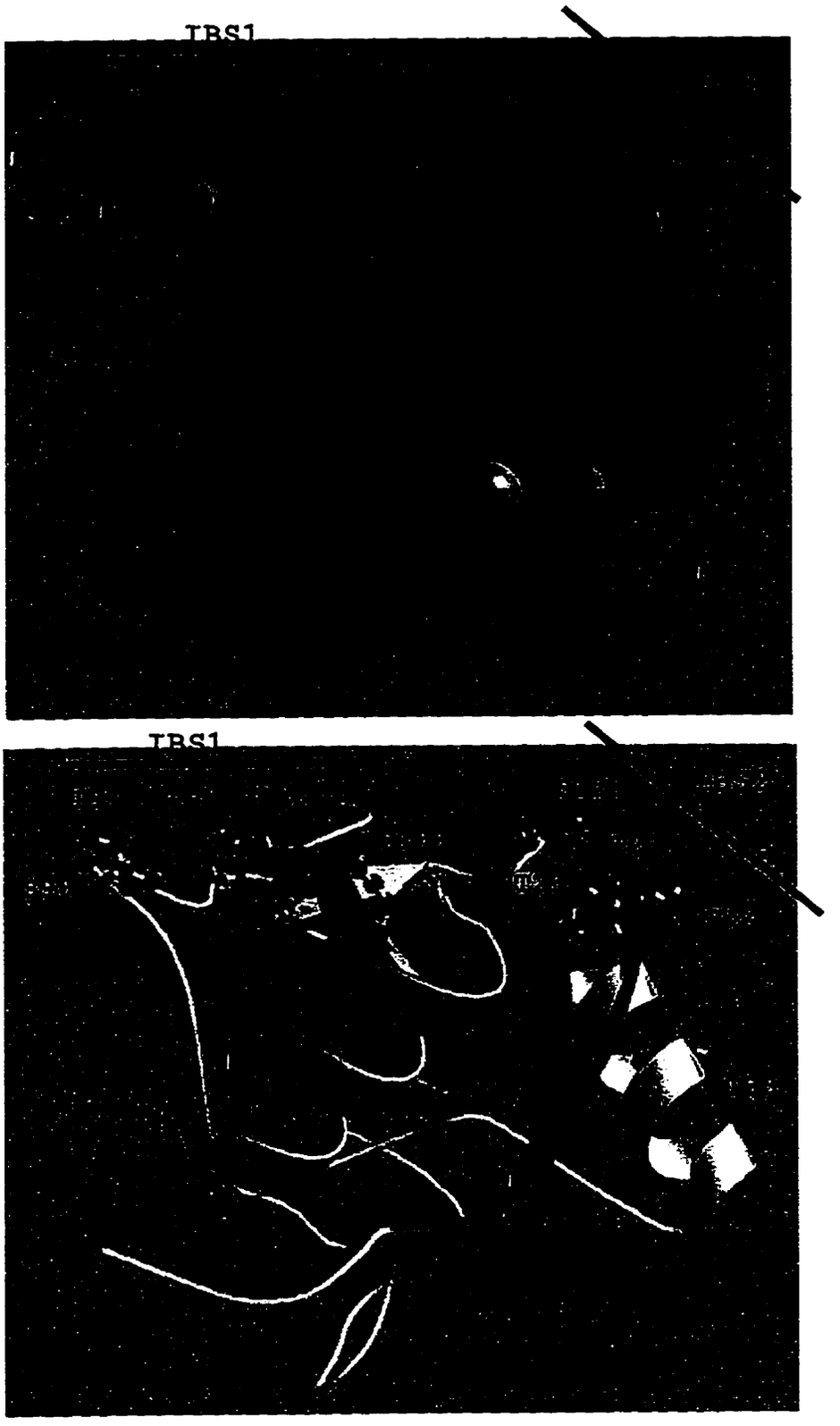
This notion that there exists at least two distinct IBSs on the SRAFP is further supported by the AFP's unique ice crystal morphology. Type I, III and herring AFPs have very crisp, clearly faceted crystals with well defined edges. The wild-type SRAFP

ice crystal is very round in shape with no single facet in evidence. This distinct characteristic might be explained if there were two separate IBSs, with comparable affinities for ice, competing for two different ice planes. The result could conceivably be a rounding of the crystal. The loss of the rounded characteristic during mutagenesis might be the result of eliminating one of the two competing IBSs, leaving only the second intact. The ice etching studies which identified only a single binding plane for SRAFP (27), might be the result of the low concentration of AFP required to carry out the experiment. Only one of the ice-binding surfaces might bind preferentially at low SRAFP concentration.

Searching for IBS2

Two strategies were used to search for a second ice-binding site. In the first strategy, surface accessible residues conserved between sea raven, herring and smelt were targeted. Unfortunately, out of several mutants only the double mutation Q55A/T56A in helix 2 resulted in rSRAFP production and it showed no alteration of antifreeze activity. In the second strategy, residues identified in the PSP putative $CaCO_3$-binding surface that maintain surface accessibility and hydrophilicity in the SRAFP model, were targeted (Figure 49). These mutants were made in the Loop4 background, so that IBS1 would be at least partially inactivated. Again, a low success rate was encountered in attempts to produce this set of mutants. Only S120H+Loop4 was expressed and purified to a level that permitted characterization. The low level of activity observed for this mutant was the first positive indication of a second, or at least

Fig. 49 **Structural diagrams of IBS1 and IBS2.** Two views of the SRAFP model are shown. top: Space filling van der Waals structure showing the relative surface topology and orientation of IBS1 and IBS2. bottom: View from the side to illustrate that residues comprising IBS2 are located on a different face than those of IBS1. In both views red=oxygen, blue=nitrogen, black=carbon yellow=sulfur. Structure file courtesy of F. Sönnichsen. Other details are described in the legend to Figure 7.


IBS1
IBS1

a significantly extended IBS. Unfortunately, S120 is situated in a hinge position (similar to N14 in Type III AFP), that might allow it to bind to ice in conjunction with either IBS1 or the putative PSP-like ice-binding site (IBS2). Subsequent mutants made in the IBS2 binding region, including S120H in wild type background, T23H in Loop4, and T23H in wild-type, indicate that S120 is a pivotal side chain, as it is the only residue where a single site mutation resulted in a significant loss of activity. While T23H by itself gives no indication of being involved in ice binding, the T23H/Loop4 mutant, which allows visualization of the effect of T23H over the background activity of IBS1, indicates that T23 is required for activity. It is possible that if the original double site mutations in IBS1 were made in a T23H background, their effect would be more easily identified. T23 is positioned around the corner on a different plane from IBS1, and could not physically interact with the same plane of ice as IBS1. This result provides strong evidence for the presence of a second ice-binding plane. The significant role of S120 is likely to reflect its position with respect to both IBS1 and IBS2 and its ability to contribute to the binding energy of either IBS. Mutation of S120 therefore results in a loss of activity on both planes at the same time. Additional mutants knocking out one or the other IBS, in combination with ice etching studies, could help determine if these two sites do in fact bind ice independently or if they bind simultaneously in some sort of corner between two ice planes, to produce the observed {11-21} plane of ice (27).

Overall, we conclude that there exist two ice-binding sites on the SRAFP molecule. Work to determine the full extent of the binding surfaces and the roles the different residues play in mediating the ice/AFP interaction must now be carried out.

The difficulties experienced in expressing various mutants to levels sufficient for purification and characterization will be a significant challenge in working towards this future goal. The application of fermentation techniques might be of some value in overcoming the low expression levels. Determination of the complete, high resolution 3D-structure by NMR or X-ray crystallography would be a valuable adjunct to this study.

The Role of Charged and Disulfide-Bonded Residues in AFP Activity

The effect of pH on SRAFP activity and ice crystal morphology suggests some role for charged residues in activity. A large drop in activity was observed at low pH (2.0 through 4.0). This region of the pH profile is just below the isoelectric point of the protein (at approximately 5.5). Although the observed changes might be the result of deprotonation and/or aggregation, no precipitate was detected during the experiment. This range also corresponds closely to the pKa (3.86) of the side chain of Asp. There are four Asp (94,95,113,114) on the putative IBS1 and others spread out over the rest of the surface. Their protonation from the ionized state may cause the loss of observed activity and the change in ice crystal morphology. Unfortunately, the role negatively charged residues might play in the ice-protein interaction remains unknown. The change from ionized state would not be expected to drastically alter hydrogen-bonding, for while the ionized form of Asp can act only as a hydrogen acceptor, the protonated form can act as both hydrogen donor and acceptor. Moreover, alteration of N14 to Asp in Type III AFP was not accompanied by any change in activity even at pH values of 2-5. A second possibility might include salt bridges being critical to the 3D structure of the

protein. A salt bridge would be lost upon protonation, but could be partially compensated for by hydrogen-bond interactions. However, the modelled structure does not indicate any likely salt bridge interactions, although confirmation of this awaits better resolution of the 3D structure. One final possibility is that charge repulsion might cause some perturbation of the structure at the low pH causing decreased activity, although aggregation might be expected in this case.

The five disulfide bridges of SRAFP are conserved in both herring and smelt AFPs. Only two of these are conserved in all the PSPs and CRDs (53). In the case of SRAFP, some of these disulfide links might serve to replace Ca^{2+}-binding in the stabilization of loop regions, although no proof of this exists. As these disulfides have been selectively conserved in all three Type II AFP isoforms, a more plausible role for the disulfides might be stabilization of the protein against cold denaturation. An initial attempt to delete a pair of cysteines by mutagenesis were unsuccessful, because these mutants would not express in the *P. pastoris* system. This may have been due to improper folding, leading to degradation of the recombinant product. Reducing agent was previously shown to lower SRAFP activity (10mM DTT at 50°C for 4h caused a 67% decrease in activity of a 1mg/mL AFP solution) (50). In the experiments reported here, the SRAFP appeared to be more sensitive to 10mM DTT because all activity was lost at 37°C after 30min. The DTT effect was also found to be temperature dependent, with no loss of activity at 0°C. This temperature dependence suggests a more compact structure at lower temperatures, which corresponds to the working environment of the protein in the fish. This might indicate some value in checking the X-ray and NMR

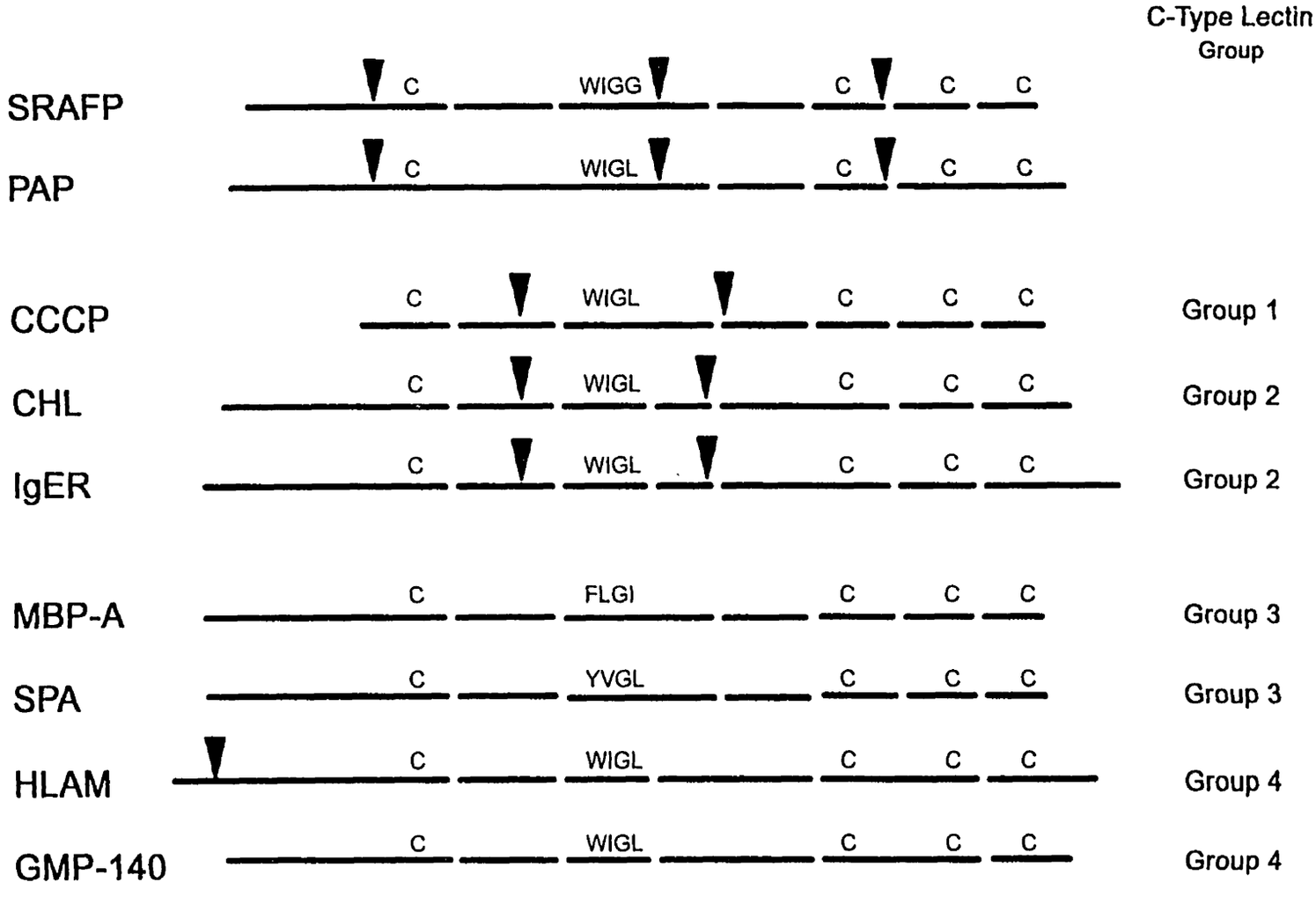
structures at temperatures close to 0°C. However, no significant change in structure has been observed for the globular Type III AFP when crystallized at 4°C versus 22°C (E. Leinalar, personal communication).

The Evolution of Type II AFPs

The intron-exon boundaries in the SRAFP gene were previously determined from a comparison of cDNA and genomic sequences (59). From an alignment of this structure with the genome structures of related proteins and domains (Figure 50), it becomes evident that this family of homologous proteins can be separated into three evolutionary groups: Group A) comprised of PSP and SRAFP with 4 exons each, Group B) comprised of Groups 1 and 2 C-type lectins with 3 exons each and Group C) comprised of Groups 3 and 4 C-type lectins containing 1 exon each. Based on a similar alignment carried out in the absence of SRAFP, PSP was identified as belonging to a new group of lectins which had evolved from the same CRD ancestral precursor, along a separate branch of the family tree (140). Our alignment of genomic structures and the similar loss of Ca^{2+}-binding residues, suggests that the SRAFP is in fact more closely related to the PSP than the CRD.

Unfortunately, the genome structures of the herring and smelt AFPs are unknown at this time. It would be interesting to know where they might fit into this ancestral tree. There is a definite subdivision between SRAFP and the herring and smelt forms, based on the Ca^{2+}-dependence. This might suggest the SRAFP has evolved independently with the PSP, while herring and smelt AFPs evolved from the CRD. However the Type II

Fig. 50 **Comparison of genome structures for SRAFP and homologues.** The amino acid sequences of SRAFP, PSP and CRDs were aligned using the tetrapeptide WIGL as reference. SRAFP (59), PAP: rat pancreatitis-associated protein (140), CCCP: core protein of chicken cartilage proteoglycan (141), CHL: chicken hepatic lectin (142), IgER: lymphocyte IgE Fc receptor (143), MBP: human mannose binding protein (144), SPA: pulmonary surfactant apoprotein (145), HLAM: human lymphocyte homing receptor (146), GMP-140: platlet granule protein (147). The triangles indicate intron positions. Gaps in the gene line represent gaps in the sequence alignment.


C-Type Lectin
Group
SRAFP
C
WIGG
C
C
C
PAP
C
WIGL
C
C
C
CCCP
C
WIGL
C
C
C
Group 1
CHL
C
WIGL
C
C
C
Group 2
IgER
C
WIGL
C
C
C
Group 2
MBP-A
C
FLGI
C
C
C
Group 3
SPA
C
YVGL
C
C
C
Group 3
HLAM
C
WIGL
C
C
C
Group 4
GMP-140
C
WIGL
C
C
C
Group 4

fish AFPs all conserve five disulfide bridges, while the PSP and CRDs have only two or three. Because loss of disulfides is more likely than gain, the three AFPs probably shared some ancestral precursor in common with PSP and lectins, that contained five (or more) disulfides. The split in Ca^{2+}-dependency likely occurred prior to the loss of the disulfides. Following this split, the PSPs and CRDs lost the less structurally important disulfides, while for some reason, possibly to avoid cold denaturation, the AFPs were independently selected to maintain all the disulfides. This hypothetical tree maintains the independent evolution of the Type II AFPs, which would allow for the different properties of each.

AFGPs in both northern cod and notothenioid fishes are believed to have evolved independently of each other as well (23)(148). The notothenioid AFGP gene evolved from a pancreatic trypsinogen precursor. They maintain up to 96% homology in the signal peptide and 3'untranslated regions (21). The cod AFGP gene shows no identity to the trypsinogen gene indicating that it evolved from a different precursor. Therefore in contrast to the Type II AFPs, the AFGPs have evolved independently by convergent evolution, rather than by divergent evolution (141). Finally, a similar division among homologous AFPs is also observed between the very distantly related flounder and sculpin Type I AFPs (1). Although no specific precursor has been identified, one explanation in this case might be that an ancestral sculpin-AFP-like α-helix was present intracellularly in all fish (31). A liver specific isoform arose in flounders to become the major serum form by picking up a signal sequence.

PART B

Roles for Thr in Binding Type I AFP to Ice.

Since the introduction of the adsorption-inhibition mechanism in 1977, it has been assumed that the four regularly spaced Thr on winter flounder Type I AFP are key residues for ice binding. The spacing of Thr along the helix suggested some form of lattice matching mechanism (32)(149). Although, the precise mechanism of attachment to ice has not been defined, it was assumed for the last two decades that the Thr hydroxyl makes hydrogen bonds with water molecules on the ice surface, providing the necessary energy of binding (42). However, as described in the Introduction, there are a number of inconsistencies with the theory. The AFP makes an insufficient number of hydrogen-bond contacts to account for the irreversible binding to ice (102)(108)(28). As well the preferred rotameric conformation of Thr orients the hydroxyl group such that it preferentially makes contacts with the helix and not the ice (40)(109). The Thr replacement study described in the Results section was carried out to gain further insight into the roles of both the hydroxyl and methyl groups of Thr.

Thr Hydroxyl vs Methyl Groups in Ice Binding

Substitution of Thr by Ser, which has similar hydrogen-bonding potential through its hydroxyl, is not detected in any natural helical antifreeze (40)(35). This suggests there may be more to the role of Thr than just hydrogen-bonding. The replacement of Thr by Ser in the HPLC-6 variant LSAAN resulted in a catastrophic loss of activity.

Prior to this study only the introduction of proline (a helix breaker) into the helix led to such a complete loss of activity by a single or double residue change (104). Any trivial explanation for the loss of LSAAN activity was eliminated when the peptide appeared to be very soluble, monomeric and completely helical as judged by CD, sedimentation analysis and NMR.

Over the concentration range tested, LSAAN was observed to bind to ice and produce hexagonal bipyramidal crystals identical to those produced by much lower concentrations of the wild-type peptide. This indicates that the mutant antifreeze binds to the same pyramidal plane and likely in the same direction. The activity loss must then be considered the result of a decrease in affinity for the ice site, since increases in AFP concentration compensated for the decreased affinity. The chief difference between Thr and Ser lies in the loss of the γ-methyl groups, suggesting the methyl groups play an important role in mediating the affinity. The NMR data clearly indicates that while Ser loses all conformational constraint, the Thr was only partially constrained to begin with, suggesting the loss of conformational constraint on the hydroxyl group by deletion of the methyl group is not likely to be the primary source of decreased activity.

To further probe the relative importance of the methyl and hydroxyl groups, Thr was changed to Val, which restored 80-90% of wild-type activity compared to Ser. Val is a very close match to Thr in shape and size of the side chain, except Val replaces the hydroxyl with a second methyl group. The NMR data indicated that Val had an even greater rotameric preference for the same conformation preferred by Thr. Overall these results demonstrate that restoration of the methyl group restores activity, indicating a

very significant role for the Thr methyl group in binding to ice. These results also demonstrate that elimination of the hydroxyl (replacement by methyl) does not cause significant loss of activity, in a molecule already proposed to form too few hydrogen-bonds to bind irreversibly to ice. This study of Thr clearly indicates that the overall hydrogen-bonding hypothesis is not based on solid experimental demonstration of hydrogen-bonding.

A New Role for Thr

These observations shift the assumed binding mechanism from hydrogen-bonding to one of van der Waals contacts between surfaces that conformationally match one another. Val is a very good space filling match for Thr, and NMR indicates it has a similar rotameric preference, which would provide a similar overall surface presentation for comparable van der Waals interactions with the ice. This role is in line with a recent suggestion that places an increased emphasis on the importance of van der Waals interactions in mediating AFP binding to ice (72), as well as the earlier proposal that Leu might contribute binding energy through van der Waals contacts in the {20-21} valleys (43).

Another contributing factor might be the entropic contributions of burying protein surfaces on the ice surface. Entropically, the burial of hydrophobic residues and surfaces is a favourable event. The Thr to Val change reduces the hydrophilicity of the helix surface, increasing its hydrophobicity, and would therefore increase the entropy gain in the event of the LVAAN surface being buried between helix and ice. In the case of the

Thr to Ser, and even the Leu to Ala changes, the hydrophobicity of the surface is decreased such that there would be less entropy gain upon binding of AFP to ice. As well, the smaller side chains in these latter cases might lead to a loss of complementarity or shielding of the interface, producing an overall decrease in the total buried surface area and a subsequent comparative decrease in entropy gain on binding.

Overall, the residual ice crystal growth with LVAAN and the slightly higher activity of LTAAN indicate that the hydrogen bonds likely contribute to the binding energy. However these results suggest that other energetic contributions through non-polar interactions of at least equal magnitude must exist in the protein-ice-water system. Recent studies on a number of sugar-binding proteins have lent support to the concept of greater roles for van der Waals contacts in mediating protein interactions. In the CRD of rat MBP (related to Type II AFP), the protein-sugar interaction is comprised of hydrogen-bonds to the hydroxyls of the sugar as well as a number of critical van der Waals contacts. Some of these include, interactions between β-carbons of the side chains and carbons of the sugar ring, as well as between imidazole rings of His and hydroxyl groups (150). A third type of non-polar interaction in this case involves an interaction between the B face of galactose and a tryptophan ring, discovered in an experiment to convert the specificity of MBP from mannose binding to galactose (151). In the case of cellulases binding to cellulose, extensive studies indicated that in all cases, hydrogen-bonding, together with van der Waals interactions provide the major driving force for the binding event (152). The structure of these cellulose-binding domains is significantly different from the helical type I AFPs, being comprised of jelly-roll β-sandwiches, that

contain a strip of hydrophobic side chains flanked by polar residues (153). The non-polar residues (Tyr and Trp) were believed to make van der Waals stacking contacts with the sugar rings (152)(153)(154). Finally in the case of antibodies binding to sugar epitopes (eg. Ab SE155-4 binding to a *Salmonella* trisaccharide (155)) "complementarity" or the overall shape and fit of the surface is considered critical for molecular recognition. This final example is particularly reminiscent of the recognition of the ice surface by AFPs. Non-polar interactions, and not cooperativity or hydroxyl group burial in ice, likely account for AFP irreversibly binding in the absence of sufficient hydrogen-bonding groups.

Roles for Asn, Asp and Leu in Binding AFP to Ice

It was originally assumed (149) that the Asn and Asp residues would hydrogen-bond with ice in a similar fashion to Thr, potentially making even more hydrogen-bonding contacts than Thr. However, despite their organization in a similar arrangement and spacing along the AFP as the Thr residues, later studies often left Asn and Asp out of the model building (42)(106)(107). The result is that the roles of Asn and Asp are much less well appreciated at this time. In the case of Leu, a mutagenic study (104) indicated that replacement of the two central Leu with Ala resulted in a loss of 33% antifreeze activity, with no perturbation of the helical structure. Two suggested roles for Leu included providing additional binding energy through van der Waals contacts with ice (43)(105), or for maintaining Asn in an optimal binding configuration (40).

Analysis of ATAAA and ATAAN indicate 20% and 30% loses of activity

respectively. LTAAA lost 39% of its activity. The ATAAN result closely matches the 33% value previously observed for the same variant (104). Overall these results might imply roles for Leu and Asn residues in mediation of the AFP-ice interaction, in the form of van der Waals contacts or hydrogen bonds to ice. As well the observed higher activity for ATAAA compared to ATAAN or LTAAA lends support to the proposed interation between Leu and Asn. It is proposed that in the single change variants, the remaining side chain is less constrained and therefore occupies a less ideal conformation for binding. Removal of the second side chain then results in improved activity. This hypothesis is further supported by the fact that the ATAAA ice-binding-motif is present in the natural AFP of the yellowtail flounder (133). This natural AFP, which contains an additional 11-amino-acid repeat and IBM, is just as active as wild-type HPLC-6 on a molar basis (35).

A Role For Asn and Leu in Solubility

Observation, by equilibrium sedimentation analysis, indicated that all variants with the Leu to Ala change, and to a lesser extent those with a Asn to Ala change, had a tendency to aggregate slowly over time. Thus the observed decreases in activity for these variants is likely due to aggregation. No specific ice-AFP interaction can be proposed for Leu or Asn at this time; rather this implies a role for these residues, in prevention of aggregation. The high alanine content of HPLC-6 and the other Type I AFPs, is in itself conducive to aggregation by coil-coil interactions. Studies have shown the existence of helical bundles, known as "Alacoils", comprised of antiparallel coiled-

coils of α-helices, with Ala in every seventh position (156). The structure allows a very close spacing between helices (7.5-8.5Å between helix axes) over four or five helical turns. It is reminiscent of leucine zippers and fibrous-protein coiled-coils, except that the Alacoils are antiparallel as opposed to the parallel leucine zippers and the Alacoils are more closely packed (156). Analysis of the parameters for Alacoil formation suggest even wild-type HPLC-6 should have a propensity to form these bundles. The Leu and Asn replacement results indicate that with the removal of these side chains the propensity becomes a probability. This raises an interesting issue about wether multimers retain thermal hysteresis activity, or maybe even show higher activity due to increased size, which off sets the loss of AFP concentration. It is interesting, in this context, to note that Type IV AFP appears to be a four-helix bundle (75).

A Better Ice Binding Motif

AFP-9 is a minor isoform in winter flounder serum, which has an ATAAT IBM. This isoform has decisively higher activity on a molar basis that HPLC-6 (35), but again it includes an additional 11-amino-acid repeat and IBM. To determine the extent to which the modified IBM contributes to the higher activity, we incorporated it into the HPLC-6 molecule as the ATAAT variant. This variant, lacking Leu, actually had 20% higher activity compared to wild-type at 4mg/mL. In light of the solubility data, this is remarkable. The higher activity was likely observed prior to any aggregation because this particular peptide was assayed for thermal hysteresis activity immediately following resuspension. The aggregation occurs slowly over a number of days and the assay was

performed in a matter of three hours. This is the only peptide that was treated in this manner; all other variants were assayed three to four days after resuspension. Overall this ATAAT variant indicates that replacement of Asn with Thr provides additional binding energy, that compensates for the loss of Leu. In the native AFP-9, while one of the IBMs is ATAAT, the peptide in fact maintains other IBMs that contain Leu and Asn residues that are likely sufficient to confer decent solubility to the peptide (35).

The re-introduction of Leu into this motif (LTAAT) resulted in a decrease in activity compared to ATAAT, but completely eliminated any tendency to aggregate. This further supports the role Leu plays in solubility.

Ice-Binding Specificity

The ice crystal morphology of LTAAT and ATAAT showed interesting features. The ATAAT variant had a decidedly stubbier crystal with a *c* to *a* ratio of 2.7 (+/- 0.1) compared to wild-type at 3.2 (+/- 0.1). This is suggestive of AFP binding to an alternative or additional surface. When the Leu residue was present, the ratio shifted back to 2.9 (+/- 0.5), closer to the wild-type value. The Leu residue also caused the ratio to become highly variable with a standard deviation of +/- 0.5. This variability might be indicative of a competition for two different ice binding planes: a second plane peculiar to ATAAT binding and the {20-21} plane to which LTAAN binds. The depressed activity of LTAAT compared to ATAAT suggested that this competition does not confer maximal antifreeze action.

Overall these observations indicate a significant role for Leu and Asn in the

specificity of ice binding. Very little is actually known about how ice plane specificity is achieved. One suggestion has been that Thr and Asn residues bind by hydrogen bonding to distinct sites on the {20-21} plane, providing a means for selection (104)(157). It was further suggested that a contour fit to the ice surface might be important for protein/ice interactions (105). From X-ray studies (40), an alternative view was that the Type I AFP is flat, with only partially accessible Thr hydroxyl groups fixed in space. The fixed geometry specifies the binding direction, although the authors conclude that the geometric match is so poor, that there should be no specificity for one or the other direction along <01-12>, which contradicts ice etching studies (42).

Based on these observations, as well as the LSAAN and LVAAN results, we propose that ice-binding specificity is directed primarily by the contour fit of the molecule with the ice plane. Leu and Asn might contribute to specificity by the specific shape of their side chains, which when eliminated (Leu to Ala) and/or modified (Asn to Thr) cause a significant change in ice crystal morphology, indicative of different ice binding specificity. Thr and Asn may contribute further to the specificity through hydrogen bond matches to water molecules on the given plane.

Minimization of HPLC-6

Wild-type HPLC-6 has an uncommon structure in that it is one long helix. The natural development of an Ala-rich peptide (Ala is the best helix promoting residue (131)), the intricate N- and C- terminal cap structures, and the internal salt bridge are all geared toward maintenance of helicity, and suggest that helicity is critical for

mediating AFP activity. Studies have indicated that the helical structure is required to maintain the regular spacing between the repeated Thr, Asx and Leu residues, for matching to the ice lattice (43)(45)(104). When HPLC-6 was redesigned into a 15mer peptide it had a significantly lower mol% of Ala (potentially decreasing helicity) but maintained the cap structures, two appropriately spaced Thr residues and the salt bridge. Overall this design resulted in a significant loss of helical structure, primarily due to the ends of the helix "breathing". This is consistent with an NMR study carried out on a 17mer α-helical peptide, that was even more helical than the 15mers designed here (158). This study indicated that in a population of small helical peptides, a significant population of non-helical conformers exist (158). In the case of the lactam bridged 15mer, the helical conformer is locked in place, This is reflected in its very high helical content. This higher helical content results in significant binding to ice. Likely this represents a lowering of the conformational entropy between peptides bound to ice and those in solution. In the salt-bridged samples, the free energy of binding to ice does not compensate for the loss of conformational entropy upon immobilization of the peptide into a helix.

Induced Fit

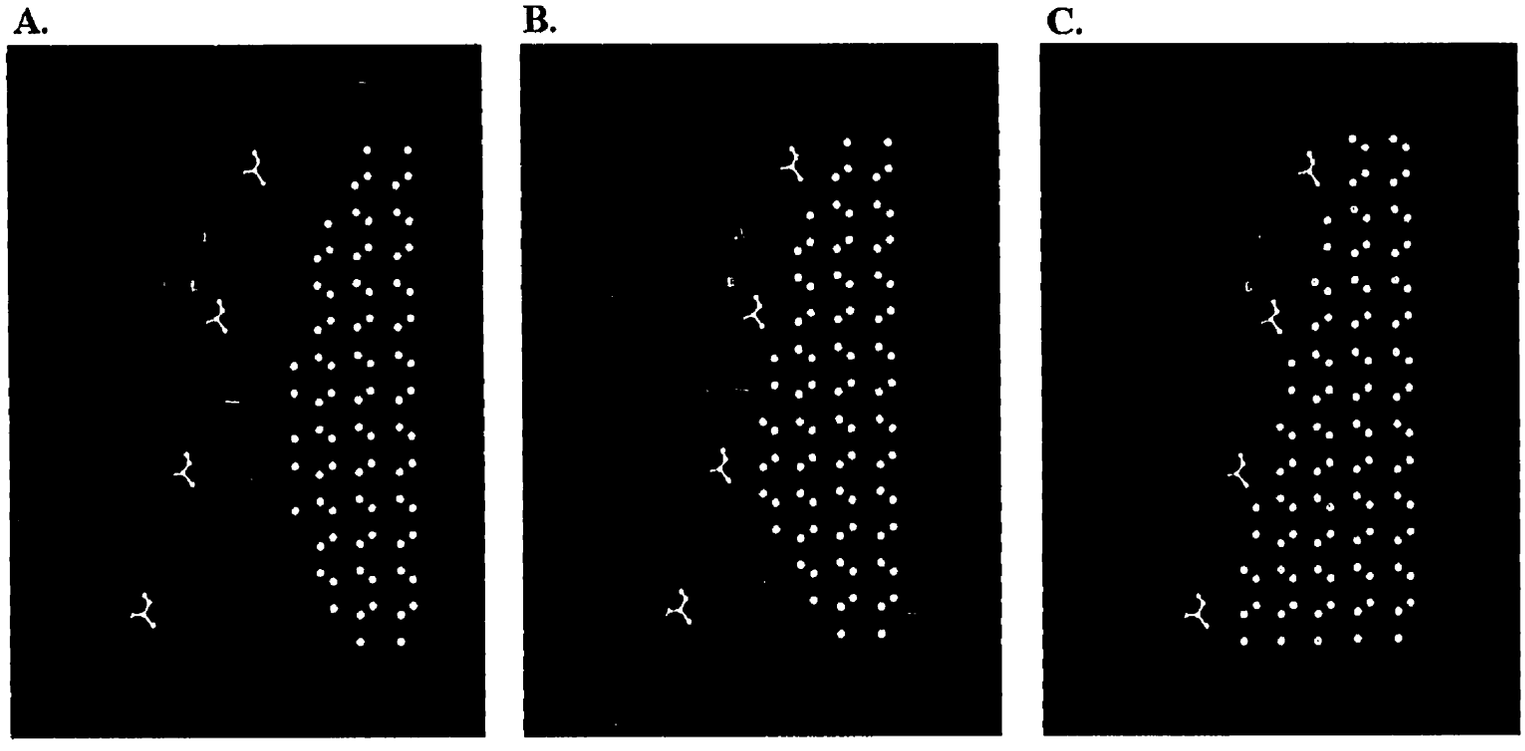
The primary paradox in the mechanism of HPLC-6 binding to ice is that the facets of ice that are bound by the AFP are not macroscopically expressed on the growing crystal; nor is the Type I AFP always optimally configured (Thr χ^1=-60°) to bind to ice. This minimization study suggested that in a helical peptide, only two of the four Thr are

actually required for initial binding to ice, although this binding may well be reversible, as demonstrated by the inability of the AFP to stop growth. This means that initially only two of the four Thr need be in the correct conformation for binding, and that a significantly smaller surface of the growing ice crystal is required to randomly produce the {20-21} plane. This random event of exposing the {20-21} plane is potentially going to occur as water molecules are added in an irregular manner to the non basal surfaces of ice. If two Thr (or one 11-amino-acid repeat) find and bind this small section it might potentially remain bound long enough for the ice front to grow in and make additional contacts. Alternatively the AFP might detach and reattach quickly enough to modify the growth of ice such that the {20-21} plane is extended. At this time a third and possibly even the fourth Thr (second and third 11-amino-acid repeats) might fall into the correct rotameric conformation and bind to the plane. Ultimately the outcome is irreversible binding of wild-type AFP to ice. Irreversibility can be envisioned because, unlike binding to a dynamic ice surface, detachment of a fully bound AFP would require the simultaneous loss of all binding forces, which is unlikely to occur.

This describes an "induced fit" mechanism for HPLC-6 binding to ice (Figure 51). The AFP is inducing the formation of its {20-21} plane ligand and the ice is selecting the rotameric conformation of the four Thr into positions that optimize binding to ice. This is a dynamic and synergistic process which improves binding sites and leads to a more efficient binding, which further improves the binding site.

This model is supported by the LTAAN NMR data that indicates all four Thr are free to occupy different rotameric states in solution, although the two middle Thr do have

Fig. 51 **Induced fit binding of Type I AFP to ice.** A) Type I AFP represented in green, with four $i,i+11$ threonyl sidechains is shown situated near the prism surface of a seed ice crystal. B) Water is added to the ice lattice creating a {20-21} plane (red oxygen atoms) large enough to bind one repeat of helix (T13 to T24). C) Reversible binding of two Thr (shown in B) leads to further shaping of the {20-21} plane as demonstrated by EK-lac. Wild-type AFP will gain more contacts as the ice continues to grow in the a-direction leading ultimately to irreversible binding.


A.
B.
C.

a 55% preference for the -60° rotamer (109). The preferential orientation of the two central Thr might "nucleate" binding of AFP to ice for subsequent induction of a better fit, as they are more likely to represent the smallest possible unit that could define the specific plane and direction of binding. In the minimized peptide it is the two terminal Thr (which presumably have no rotameric preference) and not the two internal Thr that are conserved. But as such, these Thr are still sufficient for initiating binding to ice, further supporting the idea of rotameric selection by the ice.

Conjugation to BSA

The conjugation of a lactam-bridged peptide to BSA significantly increased activity of the peptides on a molar basis. This result further refutes the cooperativity model for Type I action at the ice surface. The relatively large size of the globular BSA (67kDa) attached directly to the small peptide sterically prevents any possible helix-helix interaction that might promote cooperative packing on the ice surface. In fact the short lactam-bridged peptide used in this particular study contained a single site mutation (A8S), which facilitated the bridge formation, but eliminated the ice shaping activity observed for the wild-type EK-lac sample. The reason for this loss of activity is unknown as S8 is on a side of the helix that theoretically does not contact ice. Overall the increased size of the conjugated product compensated for the decreased affinity. According to the Kelvin effect (100)(42), the critical radius of curvature of the ice surface between bound AFPs (above which ice growth can occur), is proportional to the radius of the bound particle. Thus if the particle is very small, the critical radius of

curvature will be small and easily overcome upon undercooling of the solution. If the size of the particle is large, the critical radius of curvature will be significantly greater, which causes depression of the freezing point. Obviously the 67kDa BSA is a very large molecule that not only covers more of the surface of ice, but also rises high above the surface, further inhibiting the ice growth by preventing the fusion of curved ice surfaces. This theory is supported by the Type III AFP additional domain study (111) which indicates that the larger the added domain the greater the increase in AFP activity.

OVERALL CONCLUSIONS

Overall, this work on AFP Types I and II introduced a number of new and very significant findings, that have altered our views on how AFPs function; 1) the potential existence of more than one binding plane on a single antifreeze molecule, 2) the shift from hydrogen-bonding to non-polar interactions as being the primary mechanism of ice binding, 3) demonstration that better AFPs, both in terms of improved IBMs and overall more active AFPs can be designed, 4) the "induced fit" model, solving the paradox of how an AFP that does not maintain an optimal binding site in solution can bind to a crystal that does not naturally present the required ligand plane. However, as with all good scientific research, answers often lead to even more questions. If van der Waals forces contribute significantly to the binding of Type I AFP to ice, what role do these forces play in binding of the globular AFPs? If the primary role of Thr is not hydrogen-bonding, what is the primary role of Asn and Asp? Can Leu and Asn have roles in

binding to ice, solubility and specificity? If so how will these aspects be separated and demonstrated? What is so special about the conformation of Thr with respect to the general surface of ice that makes it a primary choice for AFP binding sites, considering many other residues can make hydrogen-bonds and have unique conformations? Is a general conformational fit between Type II AFP and the {11-21} plane of ice the primary factor contributing to binding energy? How can this possibility be efficiently tested? How will the complete Type II AFP binding surface be properly mapped in this scenario? If larger AFPs are more active, why has nature selectively chosen small peptides as antifreezes? These and other questions can only be resolved by continued long term study of the AFPs and their diverse mechanisms of action. Future work combining mutagenesis with ice etching analysis will be critical.

REFERENCES

1. Davies PL, Hew CL .1990. Biochemistry of fish antifreeze proteins. *FASEB J.* 4:2460-2468.

2. Sun X, Griffith M, Pasternak JJ, Glick BR. 1995. Low growth, freezing survival and production of antifreeze protein by the plant growth promoting rhizobacterium *Pseudomonas putida* GR12-2. *Can. J. Micro.* 24 41:776-784.

3. Griffith M, Ala P, Yang DSC, Hon WC, Moffat BA .1992. Antifreeze protein produced endogenously in winter rye leaves. *Plant. Physiol.* 100:593-596.

4. Tomchaney AP, Morris JP, Kung SH, Dunman JP .1982. Purification, composition and physical properties of a thermal hysteresis "antifreeze" protein from larvae of the beetle *Tenebrio molitor. Biochem.* 21:716-721.

5. Carpenter JF, Hansen TN .1992. Antifreeze protein modulates cell survival during cryopreservation: Mediation through influence on ice crystal growth. *Proc. Natl. Acad. Sci. USA* 89:8953-8957.

6. DeVries AL. 1982. Biological antifreeze agents in cold water fishes. *Comp. Biochm. Physiol.* 73A:627-640.

7. Duman JG, Xu L, Neven LG, Tursman D, Wu DW. 1991. Hemolymph proteins involved in insect subzero temperature tolerance: ice nucleators and antifreeze proteins. In RE. Lee Jr. DL. Denlinger eds. *Insects at low Temperature.* Chapman and Hall, New York, pp. 94-127.

8. DeVries AL. 1971. Glycoproteins as biological antifreeze agents in antarctic fish. *Science* 172:1152-1155.

9. Feeney RE, Hofman R. 1973. Depression of freezing point by glycoprotein from an antarctic fish. *Nature* 243:357-358.

10. Duman JG, DeVries AL. 1974. Freezing resistance in winter flounder *Pseudopleuronectes americanus. Nature (Lond.)* 247:237-238.

11. Hallet J. 1964. Experimental studies of the crystallization of supercooled water. *J. Atom. Sci.* 21:971-682.

12. Hobbs PV. 1974. *Ice Physics.* Oxford University Press.

13. Hillig WB and Turnbull D. 1956. Theory of crystal growth in an undercooled pure liquid. *J. Chem. Phys.* 24:912.

14. Kallungal JP. 1975. The growth of single ice crystals parrallel to the a-axis in subcooled quiescent and flowing water. *PhD dissertation*, Syracuse University.

15. Chao H, DeLuca CI, Davies PL .1995. Mixing antifreeze protein type changes ice crystal morphology without affecting antifreeze activity. *FEBS lett.* 357:183-186.

16. DeLuca CI, Chao H, Sonnichsen FD, Sykes BD, Davies PL. 1996. Effect of Type III antifreeze protein dilution and mutation on the growth inhibition of ice. *Biophys. J.* 71:2346-2355.

17. Knight CA, DeVries AL, Oolman LD. 1984. Fish antifreeze protein and the freezing and recrystallization of ice. *Nature* 308:295-296.

18. DeVries AL, and Wohlschlag DE. 1969. Freezing resistance in some antarctic fishes. *Science* 163:1073-1075.

19. Feeney RE and Yeh Y. 1978. Antifreeze proteins from fish bloods. *Adv. Prot. Chem.* 32:191-282.

20. DeVries AL, Komatsu SK, Feeney RE. 1970. Chemical and physical properties of freezing point depressing glycoproteins from Antarctic fish. *J. Biol. Chem.* 245:2901-2908.

21. Morris HR, Thompson MR, Osuga DT, Ahmed AI, Chan SM, Vandenheede JR, Feeney RE. 1978. Antifreeze glycoproteins from the blood of an antarctic fish. *J. Biol. Chem.* 253:5155-5162.

22. Hsiao K-C, Chen CHC, Fernandez IE, Detrich W, DeVries AL. 1990. An antifreeze protein gene from the antarctic cod *notothenia coriiceps neglecta*, encodes a polyprotein of high copy number. *Pro. Natl. Acad. Sci. USA.* 87:9265-9269.

23. Chen L, DeVries AL, Cheng CH. 1997. Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proc. Natl. Acad. Sci. USA* 94:3811-3816.

24. Bush CA, Feeney. 1986. Conformation of the glycopeptide repeating unit of antifreeze glycoprotein of polar fish as determined from the fully assigned proton NMR spectrum. *Int. J. Peptid. Prot. Res.* 28:386-397.

25. Roa BNN, Bush CA. 1987. Comparison by 1H-NMR spectroscopy of the conformation of the 2600 dalton antifreeze glycoprotein of polar cod with that of the high molecular weight antifreeze glycoprotein. *Biopolymers. J.* 26:1227-1244.

26. Drewes JA and Rowlen KL. 1993. Evidence for a gamma-turn motif in antifreeze glycoproteins. *Biophys. J.* 65:985-991.

27. Cheng CC, DeVries AL. 1991. The role of Antifreeze glycoproteins and peptides in the freezing avoidance of cold-water fish. In *Life Under Extreme Conditions.* G. diPrisco, editor Springer-Verlag, Berlin 1-14.

28. Knight CA, Driggers E, DeVries AL. 1993. Adsorption to ice of fish antifreeze glycopeptides 7 ad 8. *Biophys J.* 64:252-259.

29. Hew CL, Fletcher GL, Ananthanarayanan VS. 1980. Antifreeze proteins from the shorthorn sculpin, *Myoxocephalus scorpius*: isolation and characterization. *Can. J. Biochem.* 58:377-383.

30. Chakrabartty A, Hew CL, Shears M, Fletcher G. 1988. Primary structure of the alanine rich antifreeze polypeptide from gruby sculpin, *Myoxocephalus aenaeus. Can J. Zool.* 66:403-408.

31. Gong Z, Ewart KV, Hu Z, Fletcher GL, Hew CL. 1996. Skin antifreeze protein genes of the winter flounder, *Pleuronectes americanus*, encode distinct active polypeptides without the secretory signal and prosequences. *J. Biol. Chem.* 271:4106-4112.

32. DeVries AL and Lin Y. 1977. Structure of a peptide antifreeze and mechanism of adsorption to ice. *Biochem. Biophys. Acta* 495:388-392.

33. Davies PL, Roach AH, Hew CL. 1982. DNA sequence coding for a Antifreeze protein precursor from winter flounder. *Proc. Natl. Acad. Sci. USA* 79:335-339.

34. Pickett M, Scott, Davies PL, Wong N, Joshi S, Hew CL. 1984. Sequence of an antifreeze protein precursor. *Eur. J. Biochem.* 143:35-38.

35. Chao H, Hodges RS, Kay CM, Gauthier SY, Davies PL. 1996. A natural variant of Type I antifreeze protein with four ice-binding repeats is a particularly potent antifreeze. *Prot. Sci.* 5:1150-1156.

36. Scott GK, Davies PL, Shears MA, Fletcher GL. 1987. Structural variations in the alanine rich antifreeze proteins of the *Pleuronectinae. Eur. J. Biochem.* 168:629-633.

37. Davies PL, Hough C, Scott GK, Ng N, White BN, Hew CL. 1984. Antifreeze protein genes of the winter flounder. *J. Biol. Chem.* 259:9241-9247.

38. Davies PL. 1992. Conservation of antifreez protein encoding genes in tandem repeats. *Gene.* 112:163-170.

39. Yang DSC, Sax M, Chakrabartty A, Hew CL. 1988. Crystal structure of an antifreeze polypeptide and its mechanistic implications. *Nature* 333:232-237.

40. Sicheri F, Yang DSC. 1995. Ice binding structure and mechanism of an antifreeze protein from winter flounder. *Nature (Lond.)* 375:427-431.

41. Annanthanarayanan VS, Hew CL. 1977. Structural studies on the freezing-point-depressing protein of the winter flounder *Pseudopleuronectes Americanus*. *Biochem. Biophys. Res. Commun.* 74:685-689.

42. Knight CA, Cheng CC, Devries AL. 1991. Adsorption of alpha-helical antifreeze peptides on specific ice crystal surface planes. *Biophys. J.* 59:409-418.

43. Wen D, Laursen RA. 1992b. A model for binding of antifreeze polypeptide to ice. *Biophys. J.* 63:1659-1662.

44. Doig AJ, Chakrabartty A, Klingler TM, Baldwin RL. 1994. Determination of free energies of N-capping in alpha-helices by modification of the Lifson-Roig helix-coil therapy to include N- and C-capping. *Biochem.* 33:3396-3403.

45. Chakrabartty A, Hew CL. 1991. The effect of enhanced alpha-helicity on the activity of a winter flounder antifreeze polypeptide. *Eur. J. Biochem.* 202:1057-1063.

46. Raymond JA, Radding W, DeVries AL. 1977. Circular dichroism of protein and glycoprotein fish antifreezes. *Biopolymers* 16:2575-2578.

47. Frishman D, Argos P. 1995. Knowledge based secondary structure assignment. *Prot: Struct. Funct. Genet.* 23: 566-579.

48. Kraulis PJ. 1991. MOLSCRIPT: a program to produce both detailed schematic plots of protein structures. *J. Appl. Crystallography.* 24: 946-950.

49. Merrit EA, Murphy MEP. 1994. Raster 3D Version 2.0: A program for photorealistic molecular graphics. *Acta Crystallography* D50: 869-873.

50. Slaughter D, FLetcher GL, Ananthanarayanan VS, Hew CL. 1981. Antifreeze proteins from the sea raven *Hemitripterus Americanus*. *J. Biol. Chem.* 256:2022-2026.

51. Ewart KV, Fletcher GL. 1993. Herring antifreeze protein: primary structure and evidence for a C-type lectin evolutionary origin. *Mol. Marin. Biol. Biotech.* 2:20-27.

52. Ewart KV, Rubinsky B, Fletcher, GL. 1992. Structural and functional similarity between fish antifreeze proteins and calcium-dependent lectins. *Biochem. Biophys. Res. Comm.* 185:335-340.

53. Sonnichsen FD, Sykes BD, Davies PL. 1995. Comparitive modeling of the three-dimensional structure of Type II antifreeze protein. *Prot. Sci.* 4:460-471.

54. Drickamer K. 1988. Two distinct classes of carbohydrate recognition domains in animal lectins. *J. Biol. Chem.* 263:9557-9560.

55. Bernard JP, Adrich Z, Montalto D, DeCaro a, DeReggi M, Sarles H, Dagorn JC. 1992. Inhibition of nucleation and crystal growth of calcium carbonate by human lithostathine. *Gastroenterology*. 103:1277-1284.

56. Weis WI, Kahn R, Fourne R, Drickamer K, Hendrickson WA. 1991. Structure of the calcium-dependent lectin domain from a rat mannose binding protein determined by MAD phasing. *Science* 254:1608-1615.

57. Ewart KV, Yang DSC, Ananthanarayanan VS, Fletcher GL, Hew CL . 1996. Ca2+ dependent antifreeze proteins: modulation of conformation and activity by divalent metal ions. *J. Biol. Chem.* 271:16627-16632.

58. Wishart DS, Boyko RF, Willard L, Richards FM, Sykes BD. 1994. SEQSEE: A comprehensive program suite for protein sequence analysis. *Comput. Appl. Biosci.* 10:121-132.

59. Hayes PH, Scott GK, Ng NFL, Hew CL, Davies PL. 1989. Cystine-rich Type II antifreeze protein precursor is initiated from the third AUG codon of its mRNA. *J. Biol. Chem.* 264:18761-18767.

60. Duncker BP, Gauthier SY, Davies PL. 1994. Cystine rich fish antifreeze is produced as an active proprotein precursor in fall army worm cells. *Biochm. Biophys. Res. Comm.* 203:1851-1857.

61. Duncker BP, Gauthier SY, Davies PL. 1995. Evidence for a proprotein intermediate during maturation of Type II antifreeze protein in sea rave, *Hemitripterus americanus*. *Biochi. Biophys, Acta*. 1292:312-316.

62. Weis WI, Crichlow GV, Murtly HM, Hendricksen WA, Drickamer K. 1991. Physical characterization and crystallization of the carbohydrate-recognition domain of a mannose binding protein from rat. *J. Biol. Chem*. 266:20678-86.

63. Bradford JG, Crowther RL, Chandran C, Rumberger JM, Li S, Huang K-S, Presky DH, Familletti PC, Wolitsky BA, Burns DK. 1994. Insight into E-selectin ligand interaction from the crystal structure and mutagenesis of the lec/EGF domain. *Nature* 367:532-538.

64. Bertrand JA, Pignol D, Bernard J-P, Verdier J-M, Dagorn J-C, Fontecilla-Camps JC. 1996. Crystal structure of human lithostathine, the pancreatic inhibitor of stone formation. *EMBO J*. 15:2678-2684.

65. Ng HFL, Trinh K-Y, Hew CL. 1986. Structure of an antifreeze polypeptide precursor from the sea raven *Hemitripterus americanus*. *J. Biol. Chem*. 261:15690-15695.

66. Ng NFL, Hew CL. 1992. Structure of an antifreeze polypeptide from the sea raven. *J. Biol. Chem*. 267:16069-16075.

67. Li x, Trinh KY, Hew CL, Buettner B, Baeziger J, Davies PL. 1985. Structure of an antifreeze peptide and its precursor from the ocean pout, *Macrozoarces americanus*. *J. Biol. Chem*. 260:12902-12909.

68. Scott GK, Hayes PH, Fletcher GL, Davies PL. 1988. Wolffish antifreeze protein genes are primarily arranged as tandem repeats that each contain two genes in inverted orientation. *Molec. Cell. Biol*. 8:3670-3675.

69. Hew CL, Wang NC, Joshi S, Fletcher GL, Scott GK, Hayes PH, Buettner B, Davies PL. 1988. Multiple genes provide the basis for antifreeze protein diversity and dosage in the ocean pout, *Macrozoarces americanus*. *J. Biol. Chem*. 263:12049-12055.

70. Chao H, Davies PL, Sykes BD, Sonnichsen FD. 1993. Use of proline mutants to help solve the NMR solution structure of the Type III antifreeze protein. *Prot. Sci*. 2:1411-1428.

71. Sonnichsen FD, Sykes BD, Chao H, Davies PL. 1993. The non-helical structure of antifreeze protein Type III. *Science* 259:1154-1157.

72. Sonnichsen FD, DeLuca CI, Davies PL, Sykes BD. 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.

73. Jia Z, DeLuca CI, Chao H, Davies PL. 1996. Structural basis for the binding of a globular antifreeze to ice. *Nature* 384:285-288.

74. Chao H, Sonnichsen FD, DeLuca CI, Sykes BD, Davies PL. 1994. Structure-function relationship in the globular Type III antifreeze protein: identification of a cluster of surface residues required for binding to ice. *Prot. Sci.* 3:1760-1769.

75. Deng G, Andrews DW, Laursen RA. 1997. Amino acid sequence of a new type of antifreeze protein from the long horn sculpin myoxocephalus octodecimspinosis. *FEBS lett.* 402:17-20.

76. Wilson C, Wardell MR, Weisgraber KH, Mahley RW, Agard DA. 1991. Three-dimensional structure of the LDL-receptor binding domain of the human apolipoprotein E. *Science* 252:1817-1822.

77. Duman JG, Olsen MT. 1993. Thermal hysteresis protein activity in bacteria, fungi and phylogenetically diverse plants. *Cryobiology* 30:322-328.

78. Patterson JL, Duman JG. 1982. Purification and composition of protein antifreezes with high cysteine contents from larvae of the beetle, *Tenebrio molitor*. *J. Exp. Zool.* 219:381-384.

79. Scheppenheim R, Theede H, 1980. Isolation and characterization of freezing point depressing peptides from the larvae of Tenebrio molitor. *Comp. Biochem, Physiol.* 67B:561-568.

80. Wu DW, Duman JG, Cheng CC, Castellino FJ. 1991. Purification and characterization of antifreeze proteins from larvae of the beetle *Dendroides canadensis*. *J. Comp. Physiol.* B161:271-278.

81. Hew CL, Kao MH, So YP. 1983. Presence of cysteine-containing antifreeze proteins in the spruce budworm, *Choristoneura fumiferana*. *Can. J. Zool.* 61:2324-2328.

82. Patterson JL, Kelly TJ, Duman JG. 1981. Purification and composition of a thermal hysteresis producing protein from the milkweed bug, *Oncopeltus fasciatus*. *J. Comp. Physiol.* 142:539-542.

83. Duman JG. 1994. Purification and characterization of the thermal hysteresis protein from a plant, the bittersweet nightshade, *Solanum dulcamara*. *Biochi. Biophys. Acta*. 1206:129-135.

84. Hon W-C, Griffiths M, Chong P, Yang DSC. 1994. Extraction and isolation of antifreeze proteins from winter rye leaves. *Plant Physiol.* 104:971-980.

85. Hon W-C, Griffith M, Mlynarz A, Kwok YC, Yang DSC. 1995. Antifreeze proteins in winter rye are similar to pathogenesis-related proteins. *Plant. Physiol.* 109:879-889.

86. Tamford C. 1967. *Physical chemistry of macromolecules*. John Wiley and Sons, New York. pp 343-347.

87. Bloch R, Walters DH, Kuhn W. 1963. Structurally caused freezing point depression of biological tissues. *J. Gen. Physiol.* 46:605-615.

88. Raymond JA, DeVries AL. 1977. Adsorption-inhibition as a mechanism of freezing resistance in polar fish. *Proc. Nat. Acad. Sci. USA.* 74:2589-2593.

89. Brown RA, Yeh Y, Burcham TS, Feeney RE. 1985. Direct evidence for antifreeze glycoprotein adsorption onto an ice surface. *Biopolymers* 24: 1265-1270.

90. Kern CW, Karplus M. 1972. The water molecule. In *Water- A Comprehensive Treatise*, vol1,ed F.Franks, pp. 21-91. New York: Plenum Press.

91. Franks F. 1985. *Biophysics and Biochemistry at Low Temperature*. Cambridge Press Syndicate of the University of Cambridge.

92. Eisenberg D, Kauzmann W. 1969. *The Structure and Properties of Water*. Oxford:Clarendon Press, pp.71-77.

93. Dufour L. and Defay R. 1963. *Thermodynmics of clouds*. New York: Academic Press. pp. 174

94. Turnbull D, Fisher JC. 1949. Rate of nucleation in condensed systems. *J. Chem. Phys.* 17:71-73.

95. Fletcher NH. 1970. *The Chemical Physics of Ice*. pp. 26-27. Cambridge University Press.

96. Keizer J, Mazur P, Morita T. 1985. Theory for the anomalous light scattering in growing ice crystals *Phy. Rev. A* 32:2944-2962.

97. Cabrera N, Vermilyea DA. 1958. In *Growth and perfection of crystals*. eds Darem RH, Roberts BW and Turnbull D. (John Wiley and Sons, New York) pp. 393-408.

98. Berner RA, Morse JW.1974 Disolution kinetics of calcium carbonate in sea water iv. theory of calcite dissociation. *Amer. J. Sci.* 274:108-134.

99. Raymond JA, Wilson PW, DeVries AL. 1989. Inhibition of growth of non-basal planes in ice by fish antifreezes. *Proc. Natl. Acad. Sci.* USA 86:881-885.

100. Wilson PW. 1993. Explaining thermal hysteresis by the Kelvin effect. *Cryo-Lett.* 14:31-36.

101. Knight CA, DeVries AL. 1988. The prevention of ice crystal growth from water by antifreeze proteins. In Wagner PE, Valli G (eds) *Atmospheric aerosol and nucleation*. Springer, Berlin, Heidelberg, New York, Tokyo pp. 717.

102. Wen D, Laursen RA. 1993a. A D-antifreeze polypeptide displays the same activity as its natural L-enantiomer. *FEBS lett.* 317:31-34.

103. Laursen RA, Wen D, Knight CA. 1994. Enantioselective adsorption of the D- and L-forms of an α-helical antifreeze polypeptide to the {20-21} planes of ice. *J. Amer. Chem. Soc.* 116:12057-12058.

104. Wen D, Laursen RA. 1992a. Structure-function relationships in an antifreeze polypeptide. *J. Biol. Chem.* 267:14102-14108.

105. Madura JD, Wierzbicki A, Harrington JP, Maughon RH, Raymond JA, Sikes CS. 1994. Interaction of the D and L forms of winter flounder antifreeze peptide with the {201} plane of ice. *J. Am. Chem. Soc.* 116:417-418.

106. Chou K-C. 1992. Energy-optimized structure of AFP and its binding mechanism. *J. Mol. Biol.* 223:509-517.

107. Jorgensen H, Mori M, Matsui H, Kanooka M, Yanagi H, Yabusaki Y, Kikuzono Y. 1993. Molecular dynamics simulation of winter flounder antifreeze protein variants in solution: correlation between side chain spacing and ice lattice. *Prot. Eng.* 6:19-27.

108. Wen D, Laursen RA. 1993b. Structure-function relationships in an antifreeze polypeptide. The effect of added bulky groups on activity. *J. Biol. Chem.* 268:16401-16405.

109. Gronwald W, Chao H, Reddy DV, Davies PL, Sykes BD, Sönnichsen. 1996. NMR characterization of side chain flexibility and backbone structure in the Type I antifreeze protein near freezing temperatures. *Biochemistry* 35:16698-16704.

110. deBoer HA, Kastelein RA. 1986. Biased codon usage: An exploration of its role in optimization of translation. In- *Maximizing gene expression*. (Reznikoff W, Gold L, eds.) Bulworth Publishers, USA. pp. 225-285.

111. DeLuca CI (1997) The high resolution structure of Type III AFP and its implications for ice binding. *PhD dissertation*. Queen's University, Kingston, ON, Canada.

112. Cregg JM, Barringer KJ, Hessler AY. 1985. Pichia pastoris as a host system for transformations. *Mol. Cell Biol.* 5:3376-3385.

113. Birnboim HC, Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids. Res.* 7:1513-1523.

114. DeLuca CI, Davies PL, Samis JA, Elce JS. 1993. Molecular cloning and bacterial expression of cDNA for rat calpain II 80kDa subunit. *Biochim. Biophys. Acta.* 1216:81-93.

115. Sanger F, Nicklen S, Carlson AR. 1977. DNA sequencing with chain termination inhibitors. *Proc. Nat. Acad. Sci. USA* 74:5463-5467.

116. Tabor S, Richardson CC (1985) A bactriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* 82:1074-1078

117. Invitrogen's *Pichia* Expression Kit Version 1.9: A manual of methods for expression of recombinant protein in *Pichia pastoris*. Invitrogen, San Diego, CA.

118. Novogen. 1993. Technical Bulletin #37 (2/3/92). Novogen, Maidson, WI.

119. Kunkel TA, Roberts JD, Zarkour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154:367-382

120. Mandel M, Hig A. 1970. Calcium dependent bacteriophage DNA infection *J. Mol. Biol.* 53:154.

121. Scorer CA, Clare JJ, McCrambie WK, Ramanos MA, Sreekrishna K. 1994. Rapid selection using G418 of high copy number transformants of *Pichia pastoris* for high-level foreign gene expression. *Bio/Technology* 12:181-184.

122. Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* 72:248-254.

123. Mead DA, Szczesna-Skarupae E, Kemper B. 1986. ssDNA "blue" T7 promoter plasmids: a versitile tandemn promoter system for cloning and protein engineering. *Prot. Eng.* 1:67-74.

124. Richards EJ, 1989. Chromatography of nucleic acids. In *Current Protocols in molecular biology*. eds: FA Ausubel, R Brent, RE Kingston, DD Moore, JG Seidman, JA Smith, Struhl K. New York: Greene Publishing and Wiley-Interscience pp. 2.10.1-2.10.3.

125. Kenward KD. 1995. Expression of fish antifreeze protein genes in transgenic tobacco for increased plant freeze resistance and as a model for molecular farming. *PhD dissertation*. Queen's University, Kingston, Ontario, Canada.

126. Wight CP, Whitney GK, Daugulis AJ, White BW (1992) Enhancement and regulation of extracellular protein production by *Bacillus brevis 47* through manipulation of cell culture conditions. *Biotech. Bioeng.* 40:46-522

127. Duff SJB, Murray WD. 1988. Production and application of methylotrophic yeast *P. Pastoris*. *Biotech. Bioeng.* 31:44-49.

128. Mill A. 1993. Modelling the carbohydrate recognition domain of human E-selectin. *FEBS Lett.* 319:5-11.

129. Patard L, Stoven V, Gharib B, Bontems F, Lallemand J-Y, DeReggi M, 1996. What function for human lithostathine?: structural investigations by three dimensional structure modelling and high resolution NMR spectroscopy. *Prot. Eng.* 9:949-957.

130. Brisset NC, Perking SJ. 1996. The protein fold of the hyaluronate binding proteoglycan tandem repeat domain of link protein, aggrecan, and CD44 is similar to that of the C-type kectin super family. *FEBS lett.* 388:211-216.

131. Chakrabartty A, Kortemme T, Baldwin RL. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions *Prot. Sci.* 3:843-852.

132. Zhou NE, Monera OD, Kay CM, Hodges RS. 1994. α-helical propensity of amino acids in the hydrophobic face of an amphipathic α-helix. *Protein and Peptide Letters* 1:114-119.

133. Scott GK, Fletcher GL, Davies PL. 1986. Fish antifreeze proteins: recent gene evolution. Can. *J. Fish Aquat. Sci.* 43:1028-1034.

134. Marqusee S, Baldwin RL. 1987. Helix stabilization by $Glu^{-}...Lys^{+}$ salt bridges in short peptides of *de novo* design. *Proc. Natl. Acad. Sci. USA.* 84:8898-8902.

135. Huyghues-Despointes BM, Scholtz JM, Baldwin BL. 1993. Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Prot. Sci.* 2:80-85.

136. Houston ME, Campbell AP, Lix B, Kay CM, Sykes BD, Hodges RS. 1996. Lactam bridge stabilization of alpha-helices: the role of hydrophobicity in controlling dimeric versus monomeric alpha-helices. *Biochem.* 35:10041-10050.

137. Chen Y-H, Yan JT, Chau KH. 1974. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. *Biochemistry* 13:3350-3359.

138. Cregg JM, Vedvick TS, Raschke WC (1993) Recent advances in the expression of foreign genes in *Pichia pastoris*. *Bio/Technology* 11:905-910.

139. Monaco HL, Zanotti G. 1992. Three dimensional structure and active site of three hydrophobic molecule-binding proteins with significant amino acid sequence similarity. *Biopolymers* 32:457-465.

140. Dussetti NJ, Frigerio JM, Keim V, Dagorn JC, Iovanna JL. 1993. Structural organization of the gene encoding the rat pancreatitis-associated protein. *J. Biol. Chem.* 268:14470-14475.

141. Tanaka T, Har-El R, Tanzer ML. 1988. Partial structure of the gene for chicken cartilage proteoglycan core protein. *J. Biol. Chem.* 260: 12523-12527.

142. Bezouska K, Crichlow GV, Rose JM, Taylor ME, Drickamer K. 1991. Evolutionary conservation of intron positions in a subfamily of genes encoding carbohydrate-recognition domains. *J. Biol. Chem.* 266: 11604-11609.

143. Suter U, Bastos R, Hofstetter H. 1987. Molecular structure of the gene and the 5' flanking region of the human lymphocyte immunoglobulin E receptor. *Nuclei. Acid Res.* 15: 7295-7308.

144. Drickamer K, McCreary V. 1987. Exon structure of a monnose-binding protein gene reflects its evolutionary relationship to asialoglycoprotein receptor and nonfibrillar collagens. *J. Biol. Chem.* 263: 2582-2589.

145. White RT, Dam D, Miller J, Spratt K, Shilling J, Hawgood S, Bensen B, Cordell B. 1985. Isolation and charaterization of the human pulmonary surfactant apoprotein gene. *Nature* 317: 361-363.

146. Ord DC, Ernst TJ, Zhou LJ, Ranbaldi A, Spertini O, Griffin J, Tedder TF. 1990. Structure of the gene encoding the human adhesion molecule-1 (TQ1,Leu8) of lymphocytes and neutrophils. *J. Biol. Chem.* 265:7760-7767.

147. Johnston GI, Bliss GA, Newman PJ, McEver RP. 1990. Structure of the human gene encoding granule membrane protein-140, a member of the selctin family of adhesion receptors for leukocytes. *J. Biol. Chem.* 265:21381-21385.

148. Chen L. DeVries AL, Cheng CH. 1997. Convergent evolution of antifreeze glycoproteins in Antarctic notothnioid fish and Arctic Cod. *Proc. Natl. Acad. Sci USA* 94:3817-3822.

149. DeVries AL. 1984. Role of glycopeptides and peptides in inhibition of crystalization of water in polar fishes. *Philos. Trans. R. Soc. B Biol. Sci.* B304:575-588.

150. Iobst ST, Wormald MR, Weis WI, Dwek RA, Drickamer K. 1994. Binding of sugar ligands to Ca^{2+}-dependent animal lectins. *J. Biol. Chem.* 269:15505-15511.

151. Iobst ST, Drickamer K. 1994. Binding of Sugar Ligands to Ca^{2+}-dependent animal lectins. II. *J. Biol. Chem.* 269:15512-15519.

152. Tomme P, Creagh AL, Kilburn DG, Haynes CA. 1996. Interaction of polysaccharide with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 1. Binding specificity and calorimetric analysis. *Biochemistry* 35:13885-13894.

153. Johnson PE, Tomme P, Joshi MD, McIntosh LP. 1996. Interaction of soluble cellooligosaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 2. NMR and ultraviolet absorption spectroscopy. *Biochemistry* 35:13895-13906.

154. Din N, Forsythe IJ, Burtnick LD, Gikes NR, Miller RC, Warren RAJ, Kilburn DG. 1994. The cellulose-binding domain of endoglucanase A (CenA) from *Cellulomonas fimi*: evidence for the involment of tryptophan residues in binding. *Molec. Microbiol.* 11:747-755.

155. Bundle DR, Eichler E, Gidney MAJ, Meldal M, Ragauskas A, Sigurskjold BW, Sinnott B, Watson DC, Yaguchi M, Young NM. 1994. Molecular recognition of a *Salmonella* trisaccharide epitope by monoclonal antibody Se155-4. *Biochemistry* 33:5177-5182.

156. Gernet KM, Surles MC, Labean TH, Richardson JS, Richardson DC. 1995. The alacoil: a very tight antiparallel coiled coil of helices. *Prot. Sci.* 4:2252-2260.

157. Wierzbicki A, Taylor MS, Knight CA, Madura JD, Harrington JB, Sikes CS. 1996. Analysis of shorthorn sculpin antifreeze protein stereospecific binding to (201) faces of ice. *Biophys. J.* 71:8-18.

158. Merutzka G, Morikis D, Bruschweiler R, Wright PE. 1993. NMR evidence for multiple conformations in a highly helical model peptide. *Biochemistry* 32:13089-94.

159. Schultz JM, Qian H, York EJ, Stewart JM, Baldwin RL. 1991. Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* 31:1463-1470.

APPENDIX A

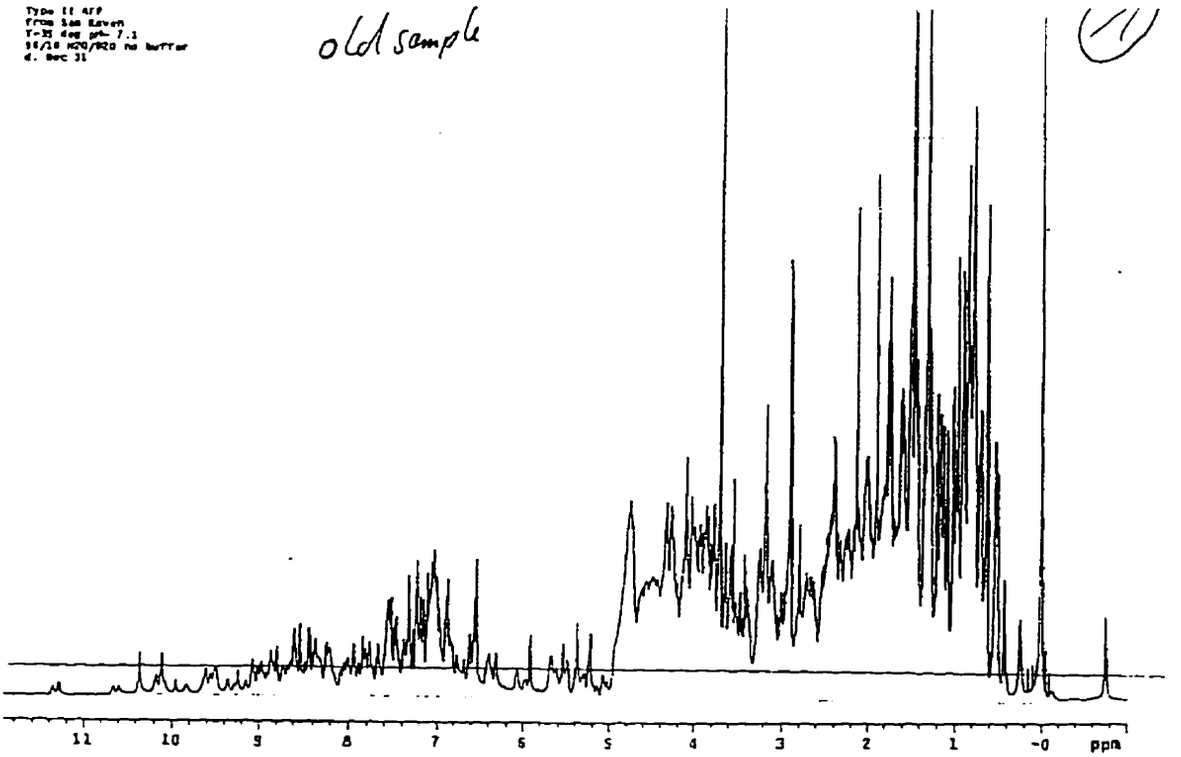
RESULTS

PART A

Type II AFP NMR Data

Samples of SRAFP purified from sea raven serum (Figure A1 and A5), rSRAFP purified from yeast media (Figure A1 and A5), Loop4 rSRAFP mutant (Figure A2), two individual samples of ^{15}N-labelled samples of rSRAFP (Figure A3 and A6), the pooled sample of ^{15}N-labelled rSRAFP (Figure A4), and S120H+Loop4 rSRAFP mutant were analyzed by 1D ^{1}H-NMR and 2D-NOESY NMR to compare folds and ensure that the mutants were correctly folded. The S120H+Loop 4 mutant was the most highly mutated and least active mutant produced. NMR indicates its fold is still virtually identical to rSRAFP as were all other samples tested. A Ca^{2+} titration experiment was carried out on the Loop4 mutant (Figure A2), which has four of the five lectin Ca^{2+} binding residues, to test for Ca^{2+} binding and any conformational change upon binding. A few minor peak shifts were observed in the amide region, but no significant conformational changes were seen that would suggest Ca^{2+}-binding occured.

Fig. A1 **1D-^{1}H-NMR spectra of SRAFP and rSRAFP.** top: nSRAFP purified from sea raven serum. The presence of the A68 methyl group peak at -0.77ppm is a unique shift resulting from orientation of the side chain of A68, 3.5Å over the centre of the aromatic rings of W75. This causes an unusual up-field shift of 2.1ppm for the A68 methyl, due to the strong anisotropic effect of the W rings. This and the amide region from 10-6ppm are useful for comparing the overall folds. bottom: rSRAFP purified from yeast media. This spectrum maintains the -0.77ppm A68 methyl peak and the amide region is virtually identical to that of nSRAFP. A few additional and shifted peaks in the α-proton region (from 6-3ppm) are likely due to the additional N- and C terminal regions. Some of the addition peaks were attributed to a very minor impurity.


old sample
11
10
9
8
7
6
5
4
3
2
1
-0
ppm

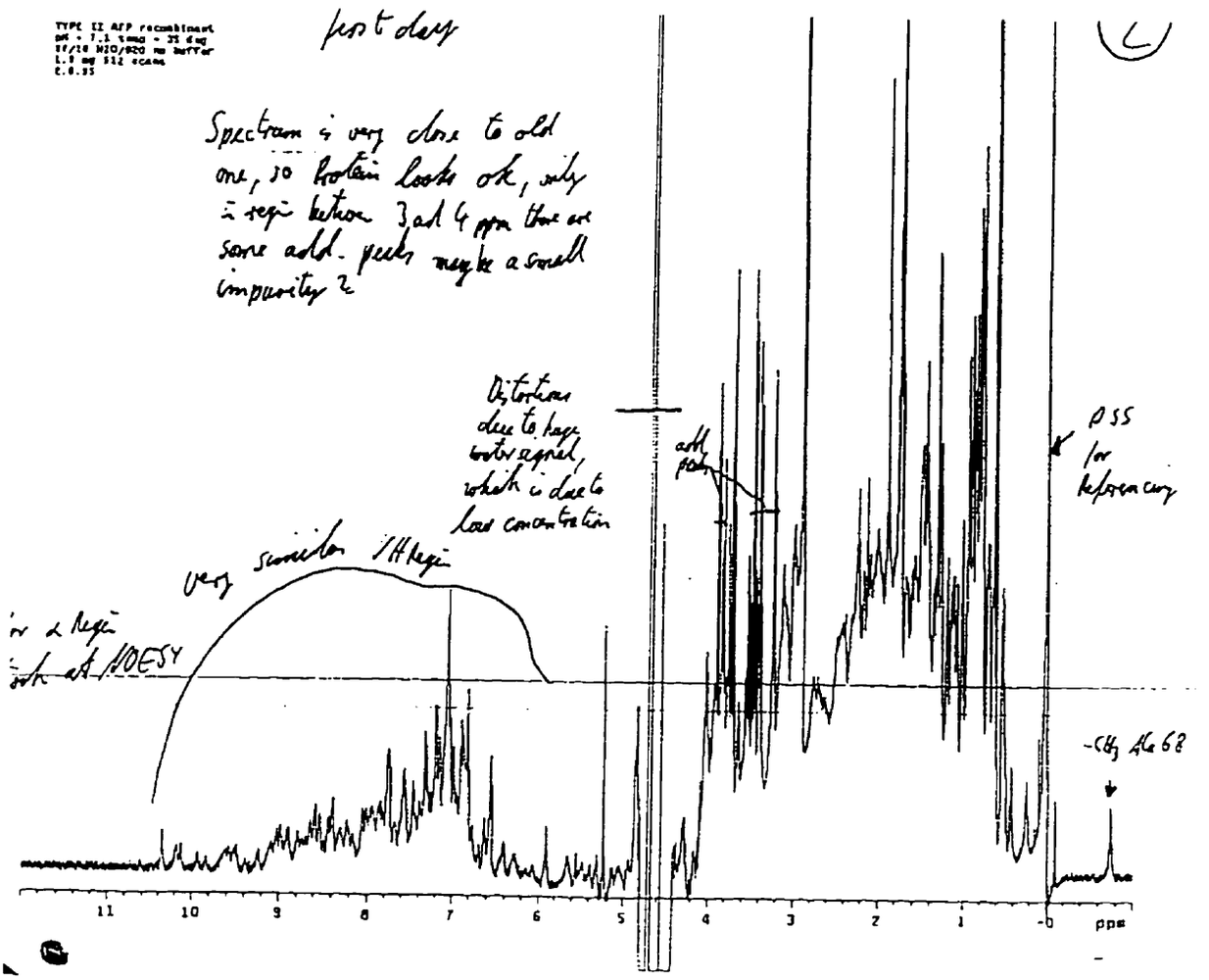
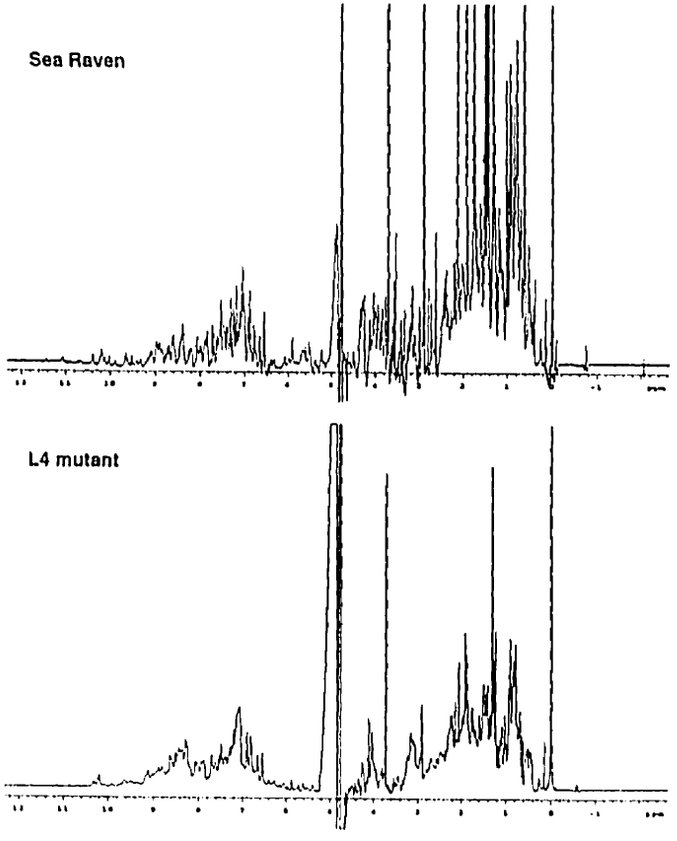

fresh sample
Spectrum is very close to old one, so protein looks ok, only a region between 3 and 4 ppm there are some add. peaks, maybe a small impurity ?
Distortions due to huge water signal, which is due to low concentration
add peaks
DSS for reference
very similar NH region
-CH3 Ala 68
11
10
9
8
7
6
5
4
3
2
1
-0
ppm

Fig. A2 **1D ^{1}H-NMR spectra of the Loop4 rSRAFP mutant and its Ca^{2+} titration.** left: Comparison of the Loop4 mutant to rSRAFP. The presence of the peak at -0.77ppm in Loop4 indicated the A68-W75 interaction is maintained. Amide region comparison indicated that the mutant maintains the same basic fold as rSRAFP.


Sea Raven
L4 mutant

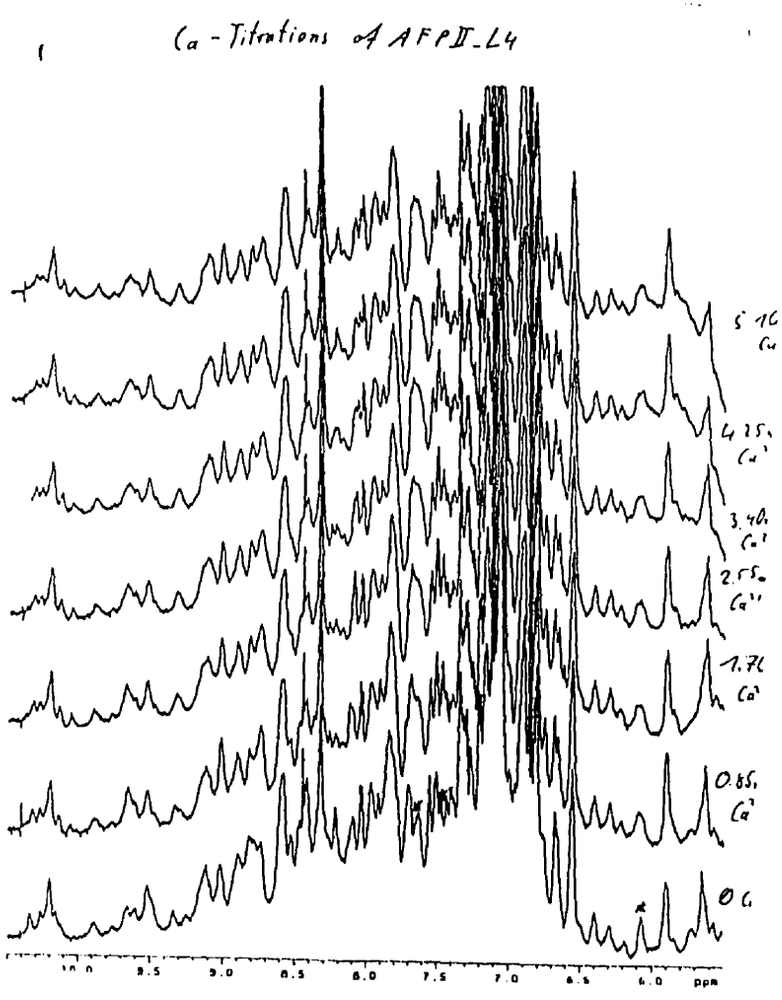
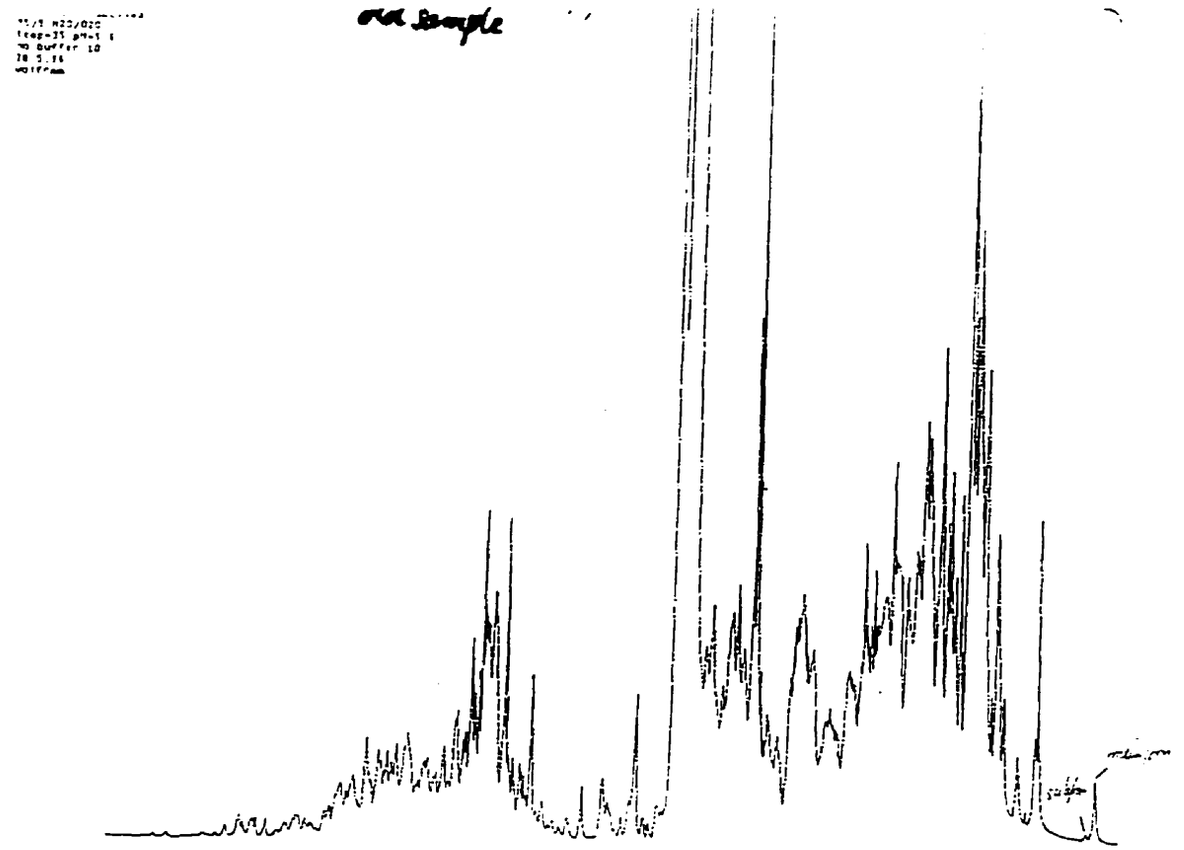

Ca - Titrations of AFPII_L4
ppm

Fig. A3 **1D ^{1}H-NMR spectra of two individual ^{15}N-labelled rSRAFP samples.** top: The first ^{15}N-labelled rSRAFP sample. bottom: The second ^{15}N-labelled rSRAFP sample. Both spectra maintain the shifted A68 methyl peak at -0.77ppm and comparison of the amide region to spectra in Figure A1 indicate an identical fold in both cases. These two samples were pooled.


Sample

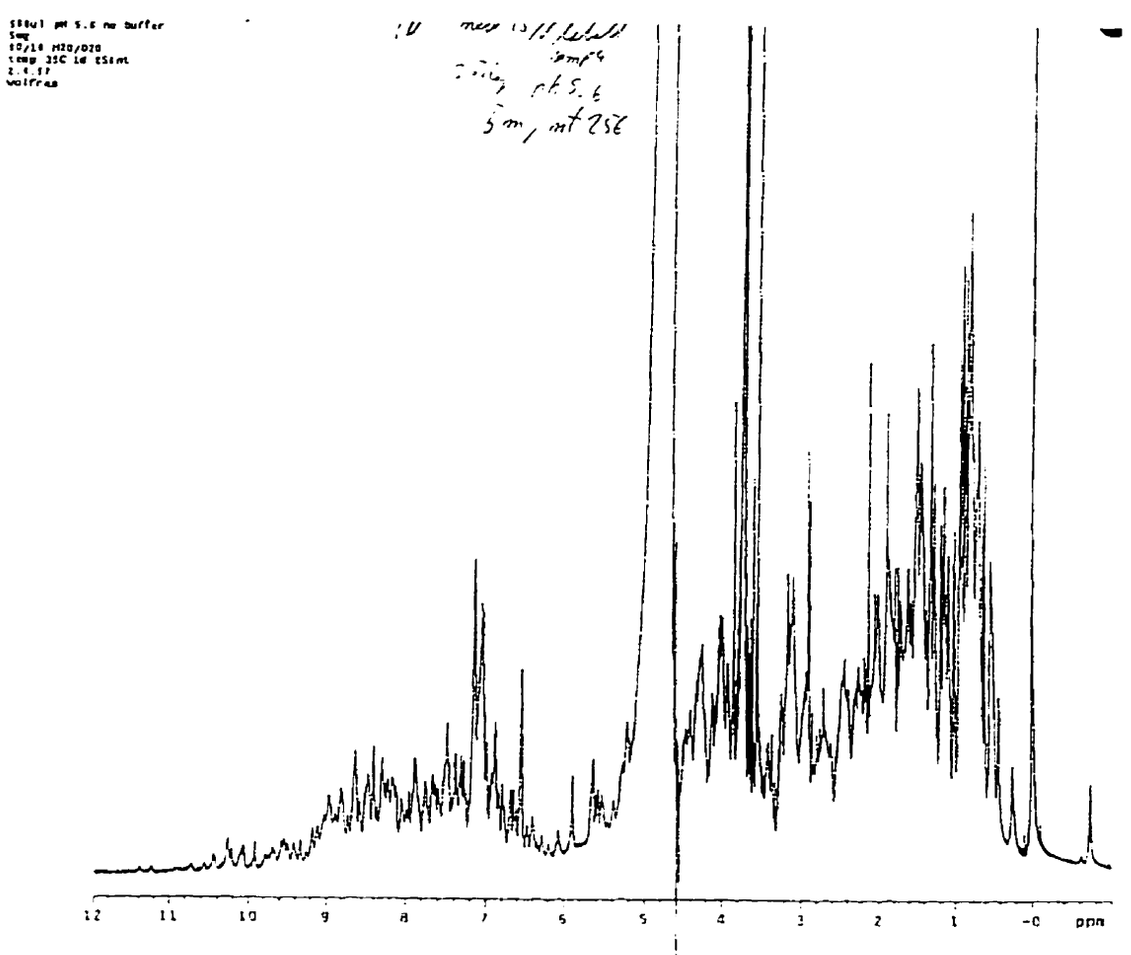
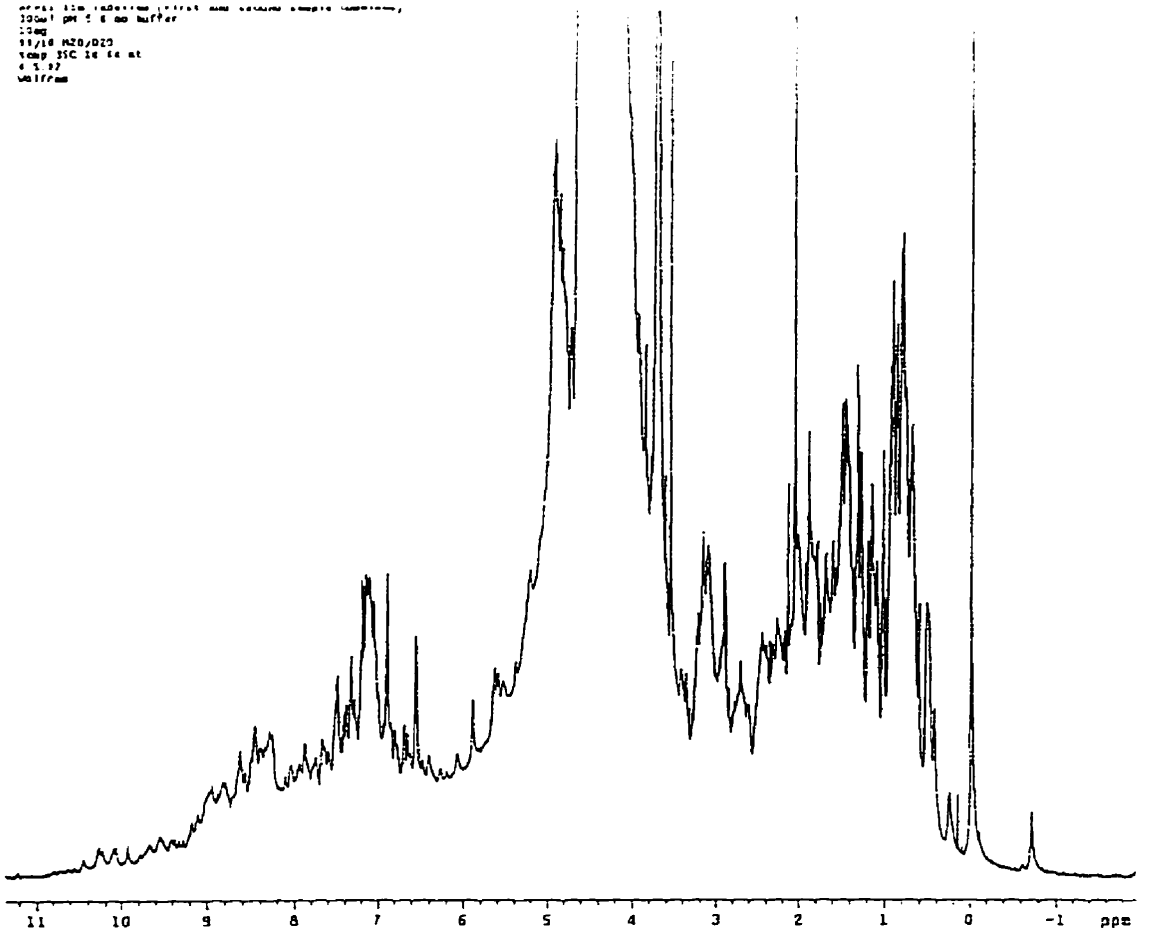

12
11
10
9
8
7
6
5
4
3
2
1
-0
ppm

Fig. A4 **1D-^{1}H-NMR spectra of the pooled ^{15}N-labelled rSRAFP and the S120H+Loop4 rSRAFP mutant samples.** top: The pooled ^{15}N-labelled rSRAFP sample. bottom: The S120H+Loop4 rSRAFP mutant. Both samples maintained the shifted A68 methyl peak at -0.77ppm and comparison of the amide regions indicates and identical fold to unlabelled wild-type samples (Figure A1).


11
10
9
8
7
6
5
4
3
2
1
0
-1
ppm

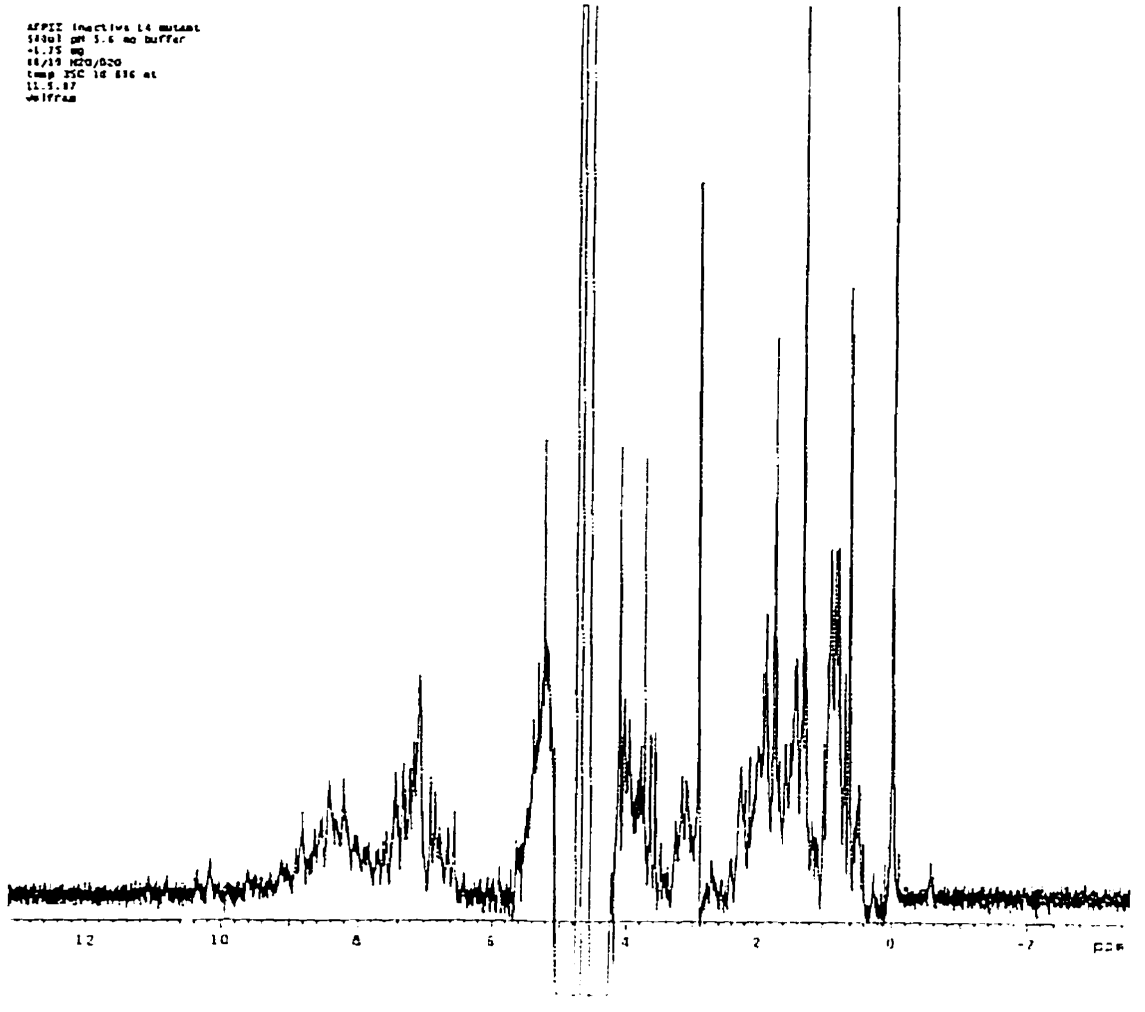
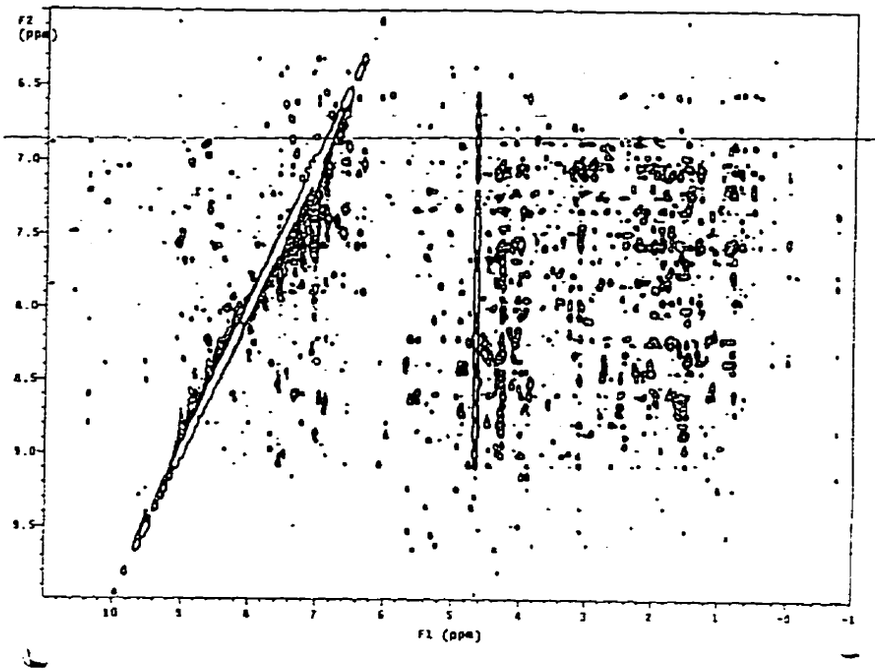

12
10
8
6
4
2
0
-2
ppm

Fig. A5 **2D-NOESY NMR spectra of SRAFP and rSRAFP.** top: SRAFP purified from sea raven serum. bottom: rSRAFP purified from yeast media. The shifted A68 methyl peaks (with NOEs to the protons on W75) were present. Peaks missing from the amide region of the rSRAFP are due to a significantly lower concentration of AFP in the recombinant experiment.

(4)

OLD NOESY



(5)

NEW NOESY

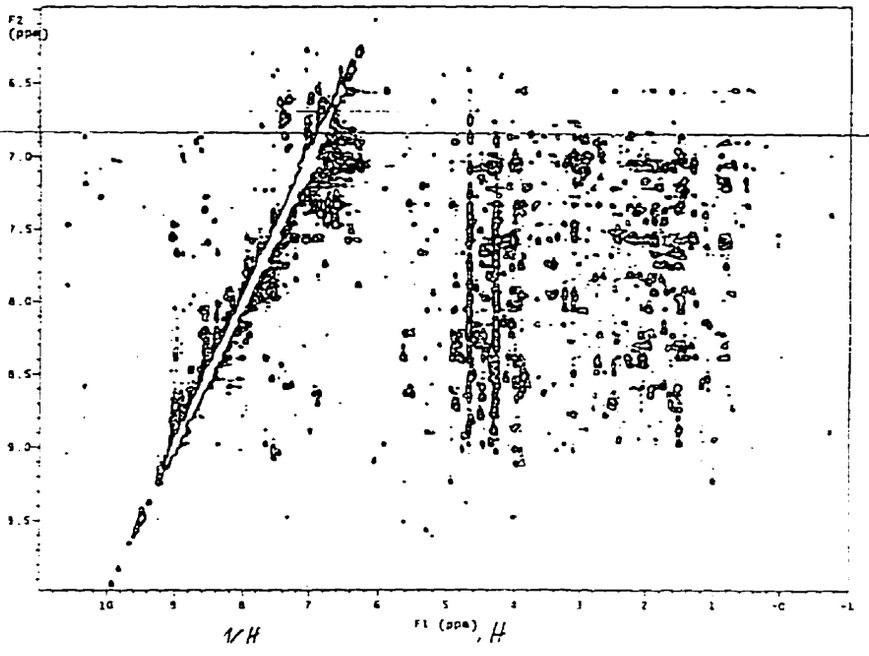
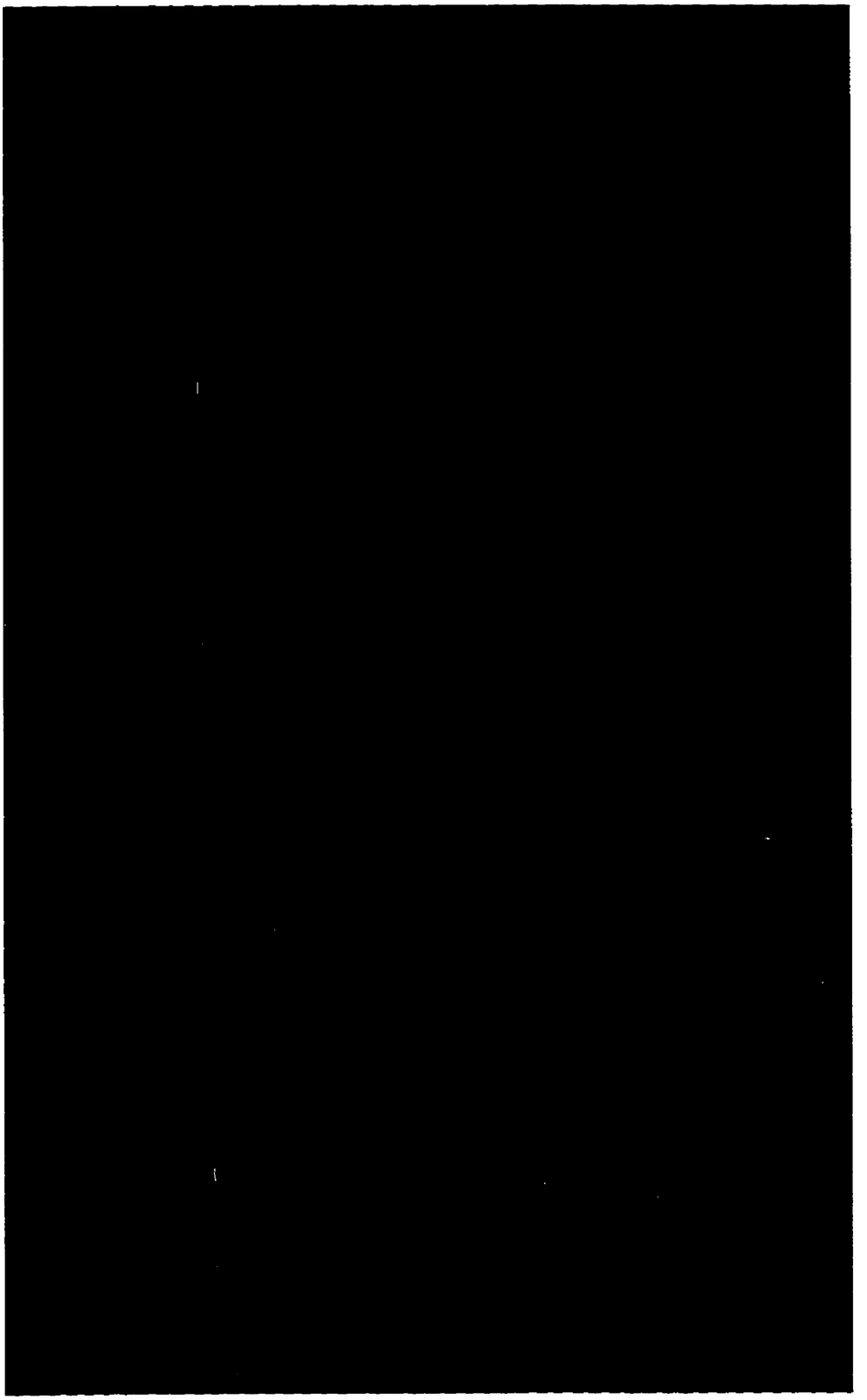


Fig. A6 **NMR solution structure of ^{15}N-labelled rSRAFP at low resolution.** An overlay of four independent solution structures of rSRAFP is shown. Structures were aligned using helix 2 as a starting point. The protein retains all the basic features of the model, including the 2 helices and the 5 β-sheets. A high resolution structure will be available in the near future. Structure files courtesy of W. Gronwald.



APPENDIX B

RESULTS

PART B

Table B1. Helicity of Type I AFP Variants

Sample	$[\Theta]_{222nm}$[1] (deg.cm^2.dmol^{-1})	Tm[2] (°C)
LTAAN	-39,370	21
LSAAN	-40,533	24
LVAAN	-40,800	25
ATAAN	-36,060	21
LTAAA	-36,250	28
ATAAA	-38,500	N/A
ATAAT	N/A	N/A
LTAAT	-37,900	19

1. The observed molar elipticity at 1°C.
   [Θ] values for ATAAN, LTAAA and ATAAA are slightly less than 100% helical Shultz et al 1991) due to aggregation in solution.

2. Apparent Tm was calculated assuming the observed elipticity at 1°C was essentially 100% (expected theoretical value is near 38,000 deg.cm^2.dmol^{-1}) Shultz et al 1991) and the observed elipticity value at 60°C was taken as the approximation for random coil.
   The unusually high Tm for LTAAA migh reflect the presence of stable multimeric complexes rather than loose aggrgate.

Table B2. Molecular Weight Determination by Sedimentation Equilibrium Ultrcentrifugation

Sample	Molecular Weight (Expt)	Molecular Weight (Theor)
LTAAN	N/A*	3,243
LSAAN	3,292	3,215
LVAAN	3,226	3,239
ATAAN	N/A**	-
LTAAA	3,887***	3,157
ATAAA	N/A**	-
ATAAT	N/A**	-
LTAAT	3,261	3,217

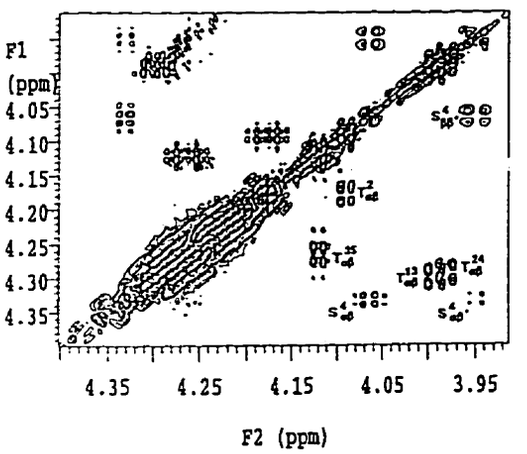
* NMR and X-ray analyses indicate LTAAN exists as a monomer in solution

** Peptide precipitated slowly and irreversibly during a three day dyalysis at 4°C prior to the centrifugation.

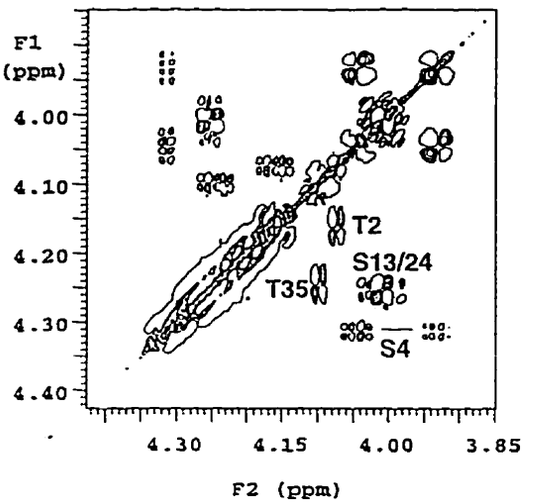
*** Sample contained higher order species but an average molecular weight for 0.7mg/mL to 3.1 mg/mL was obtained.

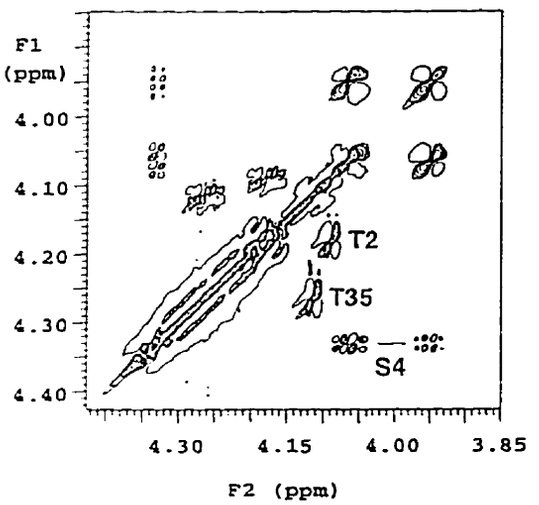
Fig. B1 **Type I AFP variant NMR data.** 2D-NOESY-NMR spectra of wild-type (LTAAN) (*109) and variants LSAAN, LVAAN. In LSAAN T13 and T24 peaks are replaced with S13 and S24 peaks. In LVAAN T13 and T24 peaks disappear and V13 and V24 aliphatic peaks appear at a different position on the spectrum. Most other points in the spectrum reamined constant indicating the overall structure of the variants is identical to wild-type.

* LTAAN

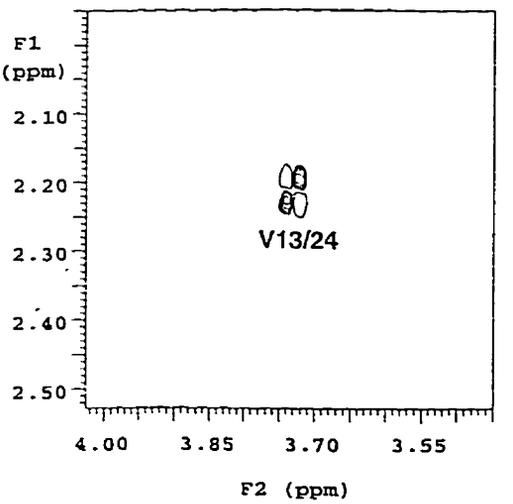


LSAAN





LVAAN



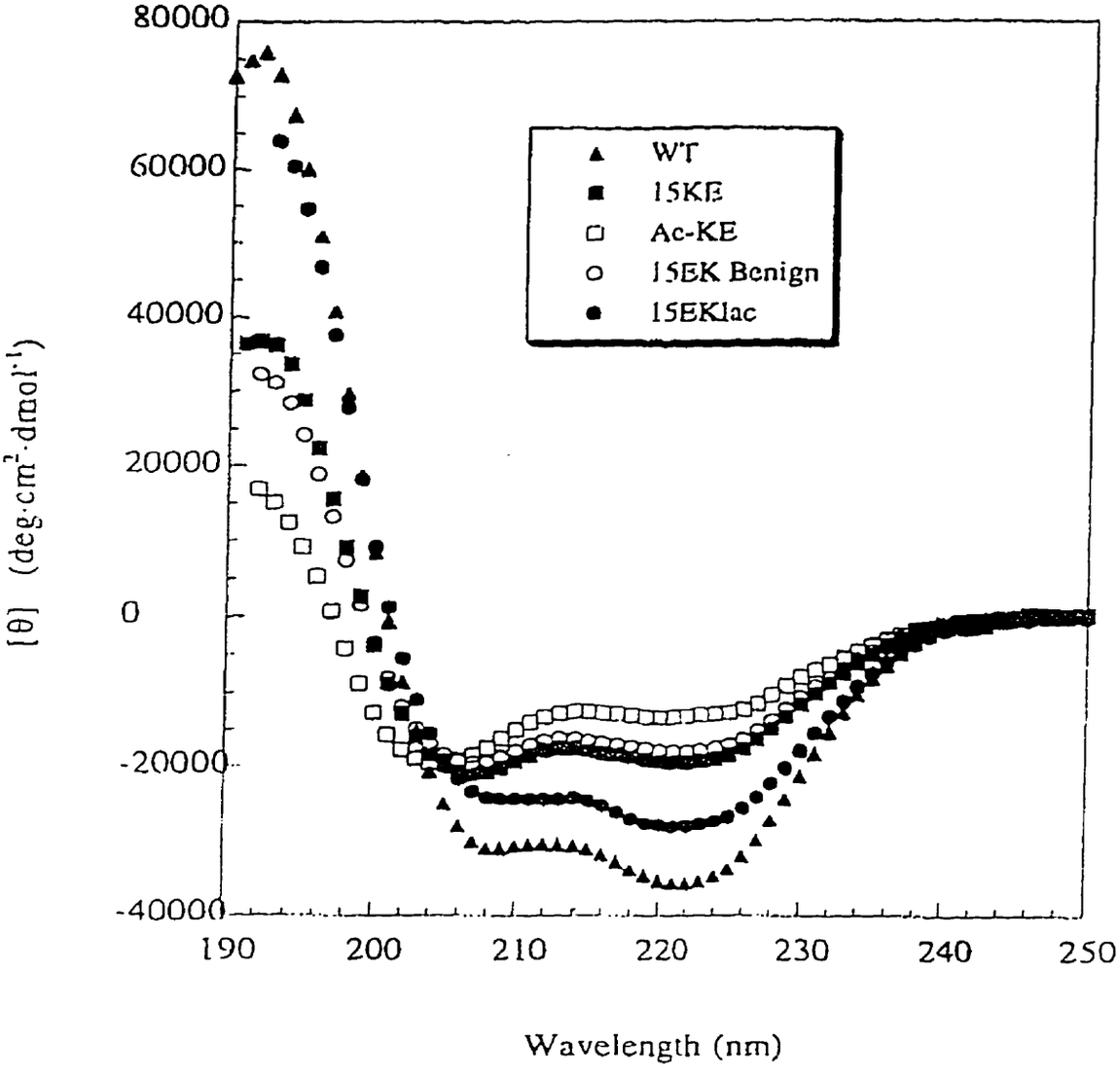
LVAAN

APPENDIX C

RESULTS

PART B

Fig. C1 **Circular dichroism spectra of the minimized peptides.** WT refers to the 37-amino-acid AFP (See Table B1 (LTAAN)). The spectrum from 15EK-lac indicates a high proportion of helical structure. The three salt-bridged minimized peptides all demonstrated significantly less helicity than the 15EK-lac sample.


80000
60000
40000
20000
0
-20000
-40000
190
200
210
220
230
240
250
[θ] (deg·cm²·dmol⁻¹)
Wavelength (nm)
WT
15KE
Ac-KE
15EK Benign
15EKlac